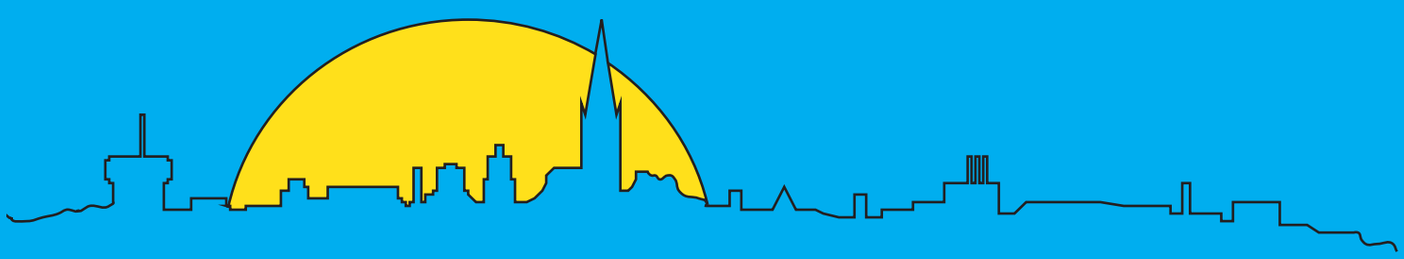
The 19th Scandinavian International Conference on Fluid Power, SICFP'25

Linköping, Sweden
June 2 - 4, 2025

Conference Proceeding



LINKÖPING UNIVERSITY

River Publishers

Description:

The proceedings of the 19th Scandinavian International Conference on Fluid Power (SICFP'25) present a collection of papers—both peer-reviewed and non-peer-reviewed—delivered during the conference. Held from June 2–4, 2025, at Linköping University in Sweden, this event continued a longstanding tradition of dialogue and collaboration among diverse perspectives within the fluid power community. SICFP'25 was organized by the Division of Fluid and Mechatronic Systems at Linköping University. SICFP is a biannual conference that alternates between Linköping University and Tampere University.

SICFP'25 took place under the motto *"Shaping the Future of Fluid Power,"* reflecting the conference's focus on the development of fluid power towards a sustainable future. As a forum for sharing the latest advancements, SICFP'25 brought together colleagues from both academia and industry to exchange information, ideas, and experiences.

The papers in these proceedings cover a broad range of topics, including innovative components and intelligent systems, diagnostics, noise and vibration, and AI-driven development. They highlight the ongoing transformation of fluid power technology to meet new demands for performance and environmental responsibility.

We thank all authors and reviewers for their contributions and dedication. We hope that these proceedings will serve as a valuable resource and inspire further research and collaboration within the fluid power field.

Keywords:

Hydraulics, fluid power, noise, vibrations, mobile applications, industrial applications, hydraulic components, digitalization, AI development, simulation and modelling, efficiency, sustainability, control systems

Table of content

Article 1

A Multi-Criteria Approach to Electro-Mechanical Actuator Selection in Non-Road Mobile Machines

By Matias Alenius, Vili Vironmäki, Vinay Partap Singh and Tatiana Minav

Article 2

SLM 3D printed proportional directional water hydraulic valve, tested with optimized mounting plate

By Jan Bartolj, Ana Trajkovski and Franc Majdič

Article 3

A Foundation for the Development of Deaeration-Efficient Hydraulic Tanks

By Charlotte Becker, Marius Hofmeister, Yamen Alhaddad and Katharina Schmitz

Article 4

Design and efficiency modelling of a fixed-ratio hydraulic transformer

By Pierre Bernard and Vincent Langlois

Article 5

Preemptive Scheduling of Multi-rod Hydraulic Infinite Linear Actuator

By Robert Braun, Magnus Landberg, Magnus Sethson and Petter

Article 6

Development of a hydraulic supply system with pressurized reservoir for a digital hydraulic actuator

By Henrique Frassetto Cardoso, Túlio de Abreu Burgos Gonçalves, Marcos Paulo Nostrani and Victor Juliano De Negri

Article 7

A Pump Decoupled Architecture to Allow Increasing Energy Efficiency of Hydrostatic Transmissions Solution based on Fixed Displacement Secondary Units

By Prithvi N. Chandiramani, Andrea Vacca and Ryan Jenkins

Article 8

Energy management of hydraulic flight control actuation systems using multi-mode control"

By Alessandro Dell'Amico

Article 9

Impact of Impeller Structural Parameters on Cavitation of High-speed Aviation Piston Pump

By Hua Fang, Hanyu Gao, Bo Yang, Yijie Lu, Hao Liu and Xiaoping Ouyang

Article 10

Simulative controller evaluation for an electro-hydrostatic excavator actuator with load-holding capability using active hysteresis and throttling control

By Felix Figge and Katharina Schmitz "

Article 11

Physics-Informed Neural Network for Solving Hydrodynamic Lubrication Characteristics of Piston Pump Slipper Pair

By Hanyu Gao, Hua Fang, Hao Liu, Yan Tong, Dayong Yang and Xiaoping Ouyang

Article 12

Simulation of Hydrostatic Pockets Between the Cylinder Block and Valve Plate of a Piston-type Pump

By Haotian Han, Thomas Heeger, Lizhi Shang and Liselott Ericson

Article 13

Simulation Study on Harmonic Torque Injection forSuppressing Flow Pulsations in Electrified Axial Piston

By Thomas Heeger and Liselott Ericson

Article 14

Measurement of a rotary mechanical pulsation compensator under hydraulic resonance

By David Holzer and Gudrun Mikota

Article 15

Towards Full Electrification: Virtual Thermal Prototyping of a Dual-Mode Electro-Hydrostatic Actuator for a Hybrid Wheel Loader

By Baoyan Hu, Chao Wang and Jian Fu

Article 16

An electrohydraulic actuator for bipedal robots with active-passive combination

By Junru Hui, Bo Yang, Haiyang Chu, Suyang Xiao, Yingwei Bao, Junhao Chen, Yang Chen, Xiaoping Ouyang and Mingjie Li

Article 17

Health Aware Control Design for Extending Remaining Useful Life of Electro-Hydrostatic Actuator Based Steering

By Simo Käki, Vinay Partap Singh, Mikko Huova and Tatiana Minav

Article 18

Experimental Investigation of Control Strategies for an Asymmetric Cylinder with Two Individually Controlled Pumps

By Samuel Kärnell, Axel Jemsson, Simon Johansson and LiselottEricson"

Article 19

Automatic Sequential Actuation of Multiple Cylinders in Mobile Cranes Using a Single Electro-Hydraulic Converter

By Timofei Komarov, Victor Zhidchenko and Heikki Handroos

Article 20

Potential of Hydraulic Waste Heat Recovery for Cabin Comfort

By Fabian Lagerstedt, Samuel Kärnell, Marcus Rösth and Liselott Ericson"

Article 21

An Elastohydrodynamic Lubrication (EHL) Model for Radial Piston Motors

By Jinhwan Lee, Lizhi Shang and Pierre Bernard

Article 22

Particle Dampers for Pump Noise Mitigation: Experimental Investigation of the Effects on Airborne and Structure-Borne Noise and Sound Perception on an Axial Piston Pump

By Michael Lenz, Peter Zuppke and Jürgen Weber

Article 23

An Inline Hydro-Mechanical Transmission (i-HMT) with a Rotatable Valve Cam to Improve Efficiency

By Perry Li, Evan Sand, Zihao Chen and Dani Putra

Article 24

Extreme Variable-Speed Lubrication Dynamics of Cylinder Block/Valve Plate Interface Based on Neural Network Method

By Dingchong Lyu, Jian Fu, Chao Zhao, Baoyan Hu, Siyuan Chen and Shoujun Zhao"

Article 25

New rotary-actuated proportional valves for humanoid robot

By Franc Majdič, Dominik Božič, Jan Bartolj and Ana Trajkovski

Article 26

Experimental and analytical investigation of a highly integrated electro-hydraulic pump motor unit

By Lukas Matias, Andrew Plummer and Nigel Johnston

Article 27

A Multi-Domain Simulation Model for Gerotor Pump Performance Prediction including Thermal Modeling

By Kai Ping Qwah, Ajinkya Pawar and Andrea Vacca

Article 28

Energy Efficient Semi-Automatic Control of Loader Cranes using Reinforcement Learning

By Amy Rankka and Alessandro Dell'Amico

Article 29

Accelerating Product Development: The Effort of Asset Administration Shells for aggregated Systems

By Denis Ritz, Malte Becker, Jürgen Weber and Katharina Schmitz

Article 30

Data-Driven Serial Testing of Axial Piston Units: Integrating Signal Processing and Machine Learning for Roller Bearing Fault Detection

By Maximilian Romeser and Reinhold von Schwerin

Article 31

First Steps Towards Model-Based Valve Design: Measuring Pressure Drop and Flow Forces of Industrial 2/2 Directional Poppet Valves

By Timo Rothhardt and Katharina Schmitz

Article 32

Model Predictive Control with Discrete-valued Input for Digital Valve Unit - Comparison of Differential Evolution and Genetic Algorithm

By Yuhao Song and Wataru Kobayashi

Article 33

The Multi-Pump System Combinatorial Problem: A Filtering Approach Using Genetic Algorithms

By Artur Tozzi de Cantuaria Gama, Samuel Kärnell and Liselott Ericson

Article 34

Sliding contacts of 3D-printed direction control valves - tribological performance of suitable material pairs

By Ana Trajkovski, Jan Bartolj, Žan Spačal and Franc Majdič

Article 35

An Approach for Evaluating Structure-Borne Noise Emission

of Axial Piston Units Using Blocked Forces

By Matthias Vogt and Marcus Geimer

Article 36

A Class-incremental Learning Method based on Knowledge Distillation for Multi-fault Diagnosis of Axial Piston Pump

By Dandan Wang, Shihao Liu, Junhui Zhang, Fei Lyu, Weidi Huang and Bing Xu"

Article 37

Sensorless Position Estimation of Electro-Hydraulic Proportional Valve Using Dual Solenoids

By Jiasheng Wang, Qi Su, Junhui Zhang, Xukun Zhang, Jinheng Liu and Zhi Qiu"

Article 38

A Vision-Based Vibration Measurement Method for Large Hydraulic Manipulator

By Shuwei Yang, Weidi Huang, Junhui Zhang, Bing Xu, Yaoxing Wang and Qi Wang"

Article 39

CFD Analysis of Hydraulic Oscillator for Flow Separation Control

By Ryuto Yoshida, Wataru Kobayashi and Seiichiro Tanibuchi

A Multi-Criteria Approach to Electro-Mechanical Actuator Selection in Non-Road Mobile Machines

Matias Alenius*, Vili Vironmäki, Vinay Partap Singh, Tatiana Minav

Faculty of Engineering and Natural Sciences, Tampere University, Tampere, Finland
* Corresponding Author: matias.alenius@tuni.fi

Abstract

The electrification of non-road mobile machines has gained momentum due to tightening emissions regulations. However, linear actuating systems in these machines often remain hydraulic, despite advancements in electro-mechanical actuators (EMAs), which offer benefits such as improved energy efficiency, controllability, and noise reduction. Selecting EMA for diverse applications involves complex trade-offs, requiring detailed simulations that can increase time-to-market. To address this, a novel EMA selection tool has been developed, incorporating a scaling algorithm to adjust operating points for lower force and velocity constraints. The tool evaluates EMA alternatives based on productivity, size and cost using a multi attribute decision-making approach, supported by a modified analytic hierarchy process to weight design criteria. The tool is tested on two non-road mobile machines—an excavator, and a forest forwarder—using duty cycle data generated using the co-simulation models based on Mevea and Matlab/Simulink simulation environments. This structured approach aims to streamline EMA selection, accelerate electrification adoption, and promote sustainable solutions in non-road mobile machines.

Keywords: Electro-mechanical actuators, Multi-attribute decision-making, Non-road mobile machines, Actuator selection, Design objectives

1 Introduction

Non-Road Mobile Machines (NRMMs) contain a large variety of machines used for different applications such as construction, forestry, agriculture, and mining, to name a few. These machines usually have application-specific linear actuation systems, which have requirements that are dependent on the specific application or purpose of the system. The differing requirements imposed on the linear actuators, including energy efficiency, cost, accuracy, size, productivity, and thermal behavior, vary greatly between applications, and the requirements are often conflicting. Thus, there does not exist a single actuation system that fits all NRMMs, and the designing of linear actuating systems involves trade-offs between the requirements.

Environmental concerns and economic factors have heightened the demand for energy efficiency, as valve-controlled hydraulics, which are frequently used, often operate with low efficiency [1]. In the literature, a large variety of technologies have been proposed to be used in the linear actuators of NRMMs [2]. Individually powered actuator concepts such as Electro-mechanical Actuator– (EMA) [3, 4] and Electro-hydrostatic Actuators (EHA) [5] have shown significant improvements in energy efficiency compared to their conventional hydraulic counterparts. However, the integrated nature of such actuators means that the power needed for the actuator's actions needs to be generated by the connected electric motor, unlike in centrally powered actuating systems, where power generated by the prime mover can be distributed flexibly to any actuator. Local power generation leads to higher total installed power, if similar peak productivity is required from every actuator in multi-actuator system. Additionally, the use of fixed transmission ratios in individually powered actuators restricts the velocity range, where the actuator can reach its maximum power, thus limiting further the plausible region of operation [6].

The total installed power needs to be kept at levels that do not make the systems economically inaccessible; hence, a compromise must be made with the peak power capacity of individual actuators. This compromise shall be made in a way that the productivity and cost of the machine are acceptable for the end user. The dimensioning of the actuator and its components therefore has multiple conflicting objectives. For example, actuator costs, size, and productivity, all of which cannot be minimized simultaneously, making the

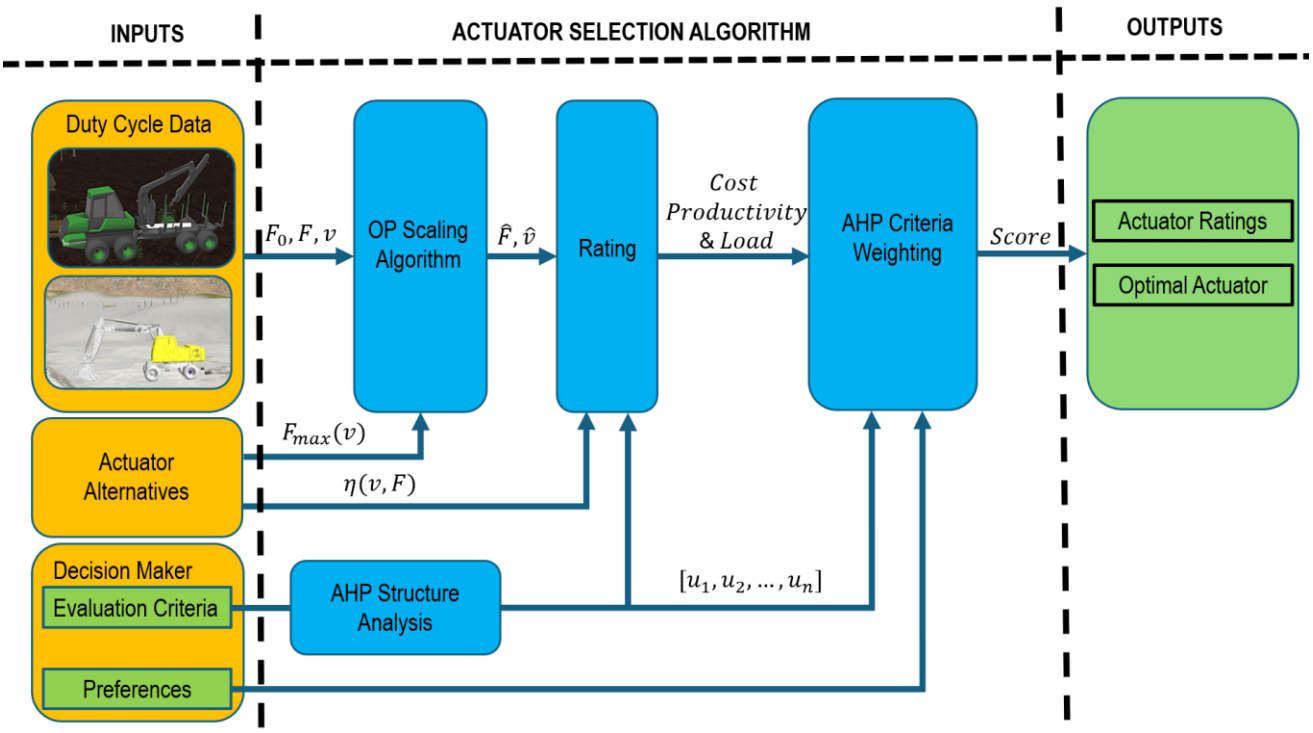


Figure 1: Overview of the actuator selection process

dimensioning problem a Multi-Objective Optimization (MOO) problem with conflicting objectives.

Multi-objective optimization methods have been widely used for integrated hybrid vehicle drivetrain topology, and component and control optimization [7, 8, 9] in both the automotive industry and for NRMMs. However, the integrated system and control design using dynamic simulations and MOO require expertise in modelling, model integration, and simulations in multiple fields, which is time-consuming [10]. The problems of the integrated MOO approach are amplified by the increased search space of possible solutions and the multi-DOF nature of manipulators driven by linear actuators. Additionally, the lack of standard duty cycles used for the simulations may lead to optimizing the linear actuators for conditions different from real world operating conditions [11]. Xue et al. [12] have used simulation-based parameter search for Pareto optimal solutions using multi-objective particle swarm optimization, though no selection method is used there. The actuator selection problem is formulated so that predefined selection of actuators is compared, thus the use of more general MOO methods is not necessary as all possible combinations can be evaluated, thus eliminating the need to find the Pareto optimal set. The scope of the said study is to develop methods to structure the decision-making process. Multi-Attribute Decision Making (MADM) methods have been applied to dimensioning and selection of components in many fields [13, 14]; however, these methods rely on low-resolution metrics to rate the alternatives. Fassbender et. al. [15] presents a broad evaluation algorithm for actuator technology selection in heavy-duty mobile machines, which again relies on subjective low resolution rating scales and the expertise of the designer rating the alternatives. Yu et al. [16] used MOO methods to generate Pareto frontier and Analytic Hierarchy Process (AHP), a method to systemize and solve MADM problems [17], to choose the optimal design for EHA to be used in aeronautical environment. However, the approach cannot be used to evaluate the productivity of the actuator as the load torque and velocity of the actuator during the duty cycle are considered as known design parameters. An exhaustive search for optimal hydraulic system topology is used by Jiao et al. [18] for aircraft control actuation system selection using MOO and AHP. Nonetheless, the said study focuses on the safety of the system and the metrics used for selection are not capable of evaluating productivity of actuator accurately, as the duty cycle is a fixed design parameter.

Considerable research has been conducted for the traction system optimization of hybrid-electric vehicles. Whereas similar optimization methods have been studied to a lesser extent for design parameter optimization of linear actuators. Moreover, the system-level optimization and decision-making frameworks for selecting linear actuators in NRMM have not been presented.

The present paper bridges this research gap for the selection of distributed linear actuating systems in NRMM applications, where trade-offs between systems productivity and other performance metrics are assessed without using high-order models. This paper aims to present a measurement data-driven multi-criteria approach selection framework that aids in the selection of EMAs to match the power demand in a predetermined duty cycle to actuators' power output capacity.

Figure 1 shows the overview of the proposed actuator selection process described in this study. An MADM algorithm based on AHP, and a novel operation point scaling algorithm is proposed to find the optimal trade-offs between productivity and other selection criteria. The developed OP scaling algorithm based on reduced order model of the system to predict EMA alternatives' OPs for recorded duty cycle is presented. The framework is then used for selecting EMAs for two different NRMM a forest forwarder and an excavator

using duty cycle data recorded from detailed co-simulation models of the machines. The use of the selection algorithm is described in detail with example use case of the forest forwarder.

The next sections present the details about the use case machines and their modeling and simulation. The third section presents the algorithms used in this study and an example use case for the EMA selection algorithm, selecting forest forwarders' lift and tilt cylinders. In the fourth chapter, the results for the presented example use case of the forest forwarder and excavator are presented, with concluding remarks and future prospects at the end.

2 Simulation models of the non-road mobile machinery

The data for the operation points used in the study is generated using co-simulation models of the two use-case NRMMs: a forest forwarder and an excavator. The multibody dynamics models of the said machines are prepared in the Mevea modeling environment, while the models of EMA are made in MATLAB/Simulink and validated through experimental data. Mevea is a digital twin technology and simulation platform featuring its own physics engine, which models the mechanics, hydraulics, power transmission, and operating environment of machines [19].

The communication between MATLAB/Simulink and Mevea software is established using a socket interface provided by Mevea [20]. This interface leverages the TCP/IP protocol to facilitate synchronous data transfer between Mevea and Simulink models at predetermined time intervals. The following sub-sections elaborate more on individual machine case models and the EMA model, respectively.

2.1 Use case NRMM

2.1.1 Forest Forwarder

A forest forwarder is an articulated NRMM used in logging operations to transport logs from the harvesting site to a roadside landing. It is equipped with a hydraulic crane and grapple for loading and unloading logs, along with lift and tilt mechanisms for precise handling. The conventional forwarders are equipped with Load-Sensing (LS) hydraulics, and the control of the linear hydraulic actuators is achieved using the throttling principles of hydraulic valves. The lift and tilt actuators are the most power-intensive ones among all others, consuming the majority of the power delivered by centralized LS hydraulics [21]. Hence, this study focuses on these two actuators, replacing them with EMA. The system shown in fig. 2 employs EMAs as lift and tilt actuators, while the rest of the system operates using LS hydraulics.

A log lifting and loading cycle is considered for further investigation, as this represents the most common operation of a forest forwarder. Figure 3 shows the lifting cycle used in the study where actuator position and forces are plotted. Initially, the boom is in a standby position, and a log of 250 kg is positioned 7 meters away from the boom's pillar, and the forwarder's load space is empty. During the cycle, the log is lifted from the ground and loaded into the load space of the forwarder machine, and the boom goes back to its initial position.

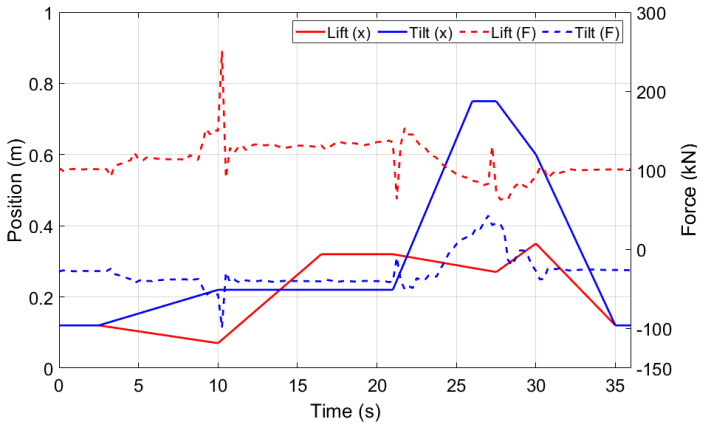
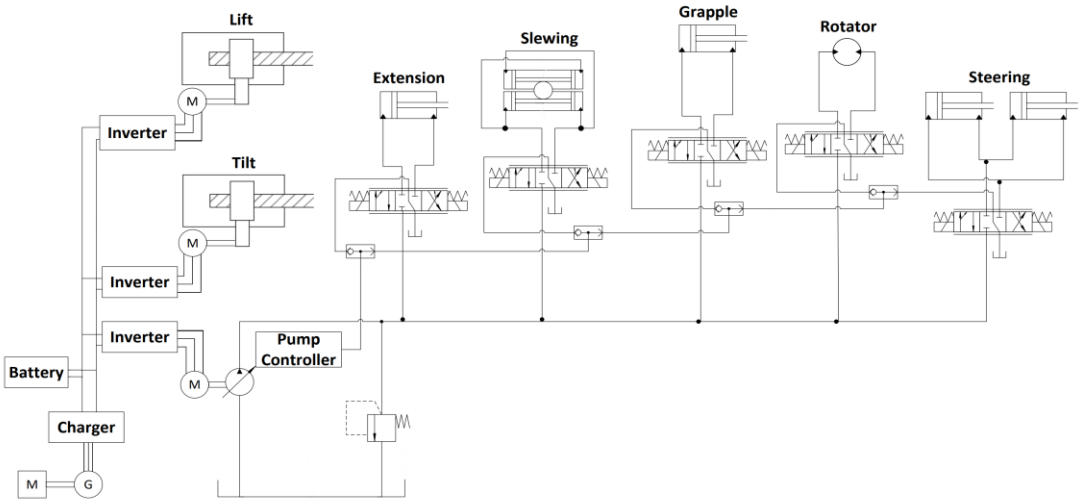


Figure 3: Duty cycle for the forest forwarder



2.1.2 Excavator

An excavator is an NRMM used for a variety of operations which involve digging, lifting, and material handling in various earthmoving and industrial applications. In conventional excavators, these operations are primarily achieved using hydraulic actuators powered by LS hydraulics and controlled using throttling valves. Since the operation of an excavator is more versatile and hence involves all actuators intensively, for this study, all the major linear actuators are replaced with EMA. It shall be noted that the model is based on a 20-ton class backhoe excavator using two parallel-connected boom actuators. The system topology is shown in fig. 4, where two boom lift actuators, an arm actuator, and a bucket actuator are replaced with EMAs. The only major hydraulic component unchanged is a hydraulic motor used for the swing motion, as the focus of the current study is linear actuators.

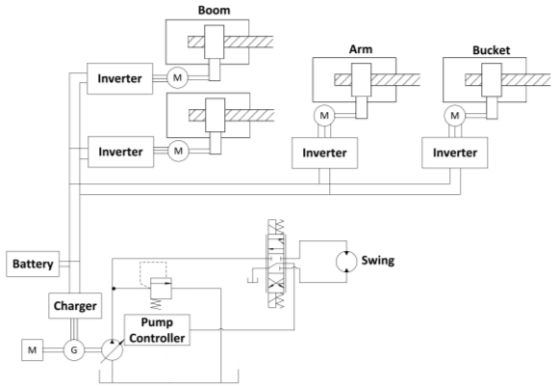


Figure 4: Schematic of EMA-based topology of the excavator

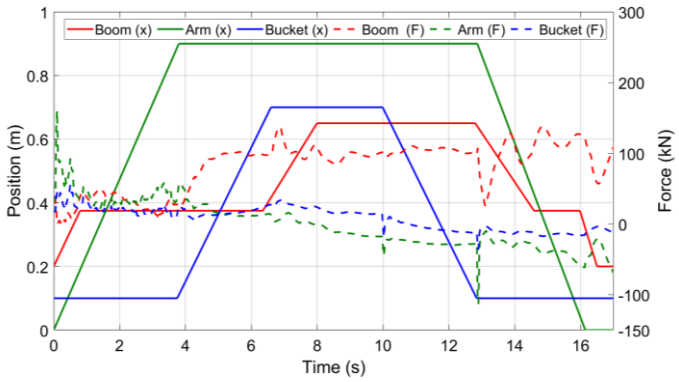


Figure 5: JCMAS duty cycle of the excavator

An excavator performs a wide variety of operations, and unlike the previous case of the forwarder, it is not possible to generalize its duty cycle. Hence, a standard digging cycle widely recommended for measurements of fuel consumption is used [22, 23]. The duty cycle, commonly referred to as JCMAS digging cycle is modified by creating load force in the bucket by using Mevea native earth model. The cycle position trajectory and the resulting load force for lift, tilt, and bucket actuators are illustrated in fig. 5.

2.1.3 EMA model

Electro-mechanical actuator (EMA) converts electrical energy directly to mechanical work, using gears and some mechanism such as ball screw or roller pitch [24]. This effectively eliminates the throttling and transmission losses which are inherently present with hydraulic actuators, giving them superior efficiency as compared to its hydraulic counterparts [25]. The EMA considered in this study uses a screw mechanism to convert the rotary motion of EM to linear motion. A simulation model of the permanent-magnet synchronous motor (PMSM) and the linear transmission mechanism are developed within the MATLAB/Simulink environment. The model is represented by the following equations.

The dynamics of the electric motor are derived from the fundamental principles governing its electrical and mechanical behavior. The electrical part of the model is described by voltage equations (1) and (2):

$$u_d = R_s i_d + L_d \frac{di_d}{dt} - \omega_s L_q i_q, \quad (1)$$

$$u_q = R_s i_q + L_q \frac{di_q}{dt} + \omega_s (L_d i_d + \Psi_r), \quad (2)$$

where R_s is the stator resistance, which may be considered identical for both the d- and q-axes. The parameters L_d and L_q are the d- and q-axis inductances, ω_s is the electrical rotor speed, and i_q and i_d are the stator currents. The rotor flux linkage along the d-axis, denoted as Ψ_r, is a constant. The flux linkage along the q-axis is assumed to be zero. [26]

As inverter tries to keep current i_d at zero, the output torque of the motor is modelled according to equation (3)

$$\tau_{EM} = K_t i_q - b_{EMA}\omega, \quad (3)$$

where b_{EMA} is the viscous friction coefficient and K_t is the motors torque constant. The force is calculated from the torque produced by the EM according to equation (4)

$$F_{EMA} = \frac{p}{2\pi} N \tau_{EM}, \quad (4)$$

where p is the pitch of the linear transmission and N is the gear ratio of the reduction gear between EM and linear transmission. The EMA output force F_{EMA} is then used as input in the mechanical model realized in Mevea simulation environment.

The main parameters used in the aforementioned equations are sourced from publicly available datasheets of commercial products [27, 28]. The primary objective of the current study is to propose the methodology for the selection of the EMA, as described earlier. Hence, the parameters used are solely for the example calculations for the present study. It is worth noting that the said parameters need to be changed according to the available products.

3 Actuator selection algorithm

The actuator selection algorithm is a high-level MADM algorithm that includes multiple subtasks, outlined in this chapter. It takes as input the data of the intended system's duty cycle, actuator characteristics, evaluation criteria including their relative importance, and design constraints. Based on these inputs and constraints, the algorithm suggests the optimal actuator from the set of alternatives. It uses AHP to model the MADM problem and rank the alternative EMAs. After that, the OP scaling algorithm is used to evaluate the selected aggregates. A flowchart of the selection process is shown in fig. 6, whereas the individual sub-tasks are described in the following sub-sections.

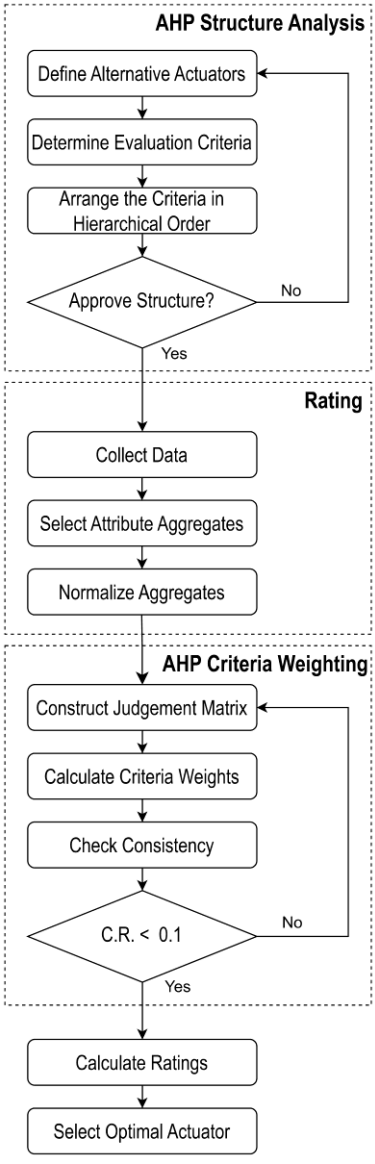


Figure 6: Flowchart of the selection algorithm

The first sub-task of the actuator selection algorithm is the structure analysis of the decision-making process. This involves defining the actuators to be compared, determining the evaluation criteria, and organizing them into a hierarchical structure, which defines the utility function form for the decision-making model. The hierarchical model follows the structure used in AHP.

The rating sub-task involves defining the actuator's duty cycle and collecting force and velocity data during this cycle for an example actuation system. The duty cycle data is then used to evaluate and normalize the aggregates in the decision-making model, with normalization based on reference values provided by the decision maker (DM). To assess each actuator alternative under its specific force and velocity constraints, the OP scaling algorithm is used. This algorithm scales down unreachable operating points due to lower force or velocity limits, estimating OPs for each actuator during the duty cycle.

The objective of criteria weighting sub-tasks is to rate the importance of utilities of sub-criteria $u_j(x_j)$ such that the total utility function U approximates the preference function P, an imaginary function describing DMs preferences ideally, as the following equation describes.

$$\nabla_x U(X) = \nabla_x P[F(X)]. \tag{5}$$

The AHP framework is used to approximate the preference functions gradient by systematically rating the importance of the criteria. Ultimately, the Actuator Selection Algorithm rates the utility of each alternative actuator, with the total utility indicating the actuator's suitability for the given application in the compared machine.

3.1 Analytic Hierarchy Process

The analytic hierarchy process (AHP) is a systemic decision-making framework that facilitates complex MADM by breaking down a problem into smaller, more manageable components [29]. AHP employs pairwise comparisons to assess the relative importance of criteria, allowing for an intuitive and structured approach to decision making.

The process is built on a hierarchical structure that typically consists of three levels: the overall goal at the top, the criteria for evaluation in the middle, and the alternatives at the bottom, as shown in fig. 7 [29].

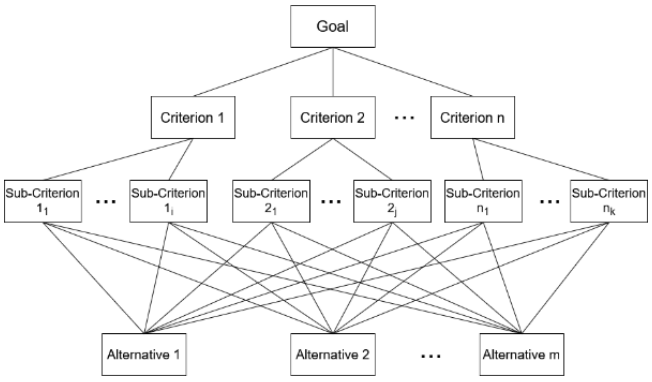


Figure 7: General structure of the hierarchy

The main steps of the AHP are as follows:

1. Define the problem and structure the hierarchy.
2. Conduct pairwise comparisons using a judgment matrix.

3. Calculate the consistency ratio (C.R.) to ensure judgment consistency.
4. Derive and normalize the priority vector.

Once the hierarchy is structured, the utility function modelling the decision makers preferences can be derived as follows,

$$U(X) = \sum_{j=1}^{n} \bar{w}_j u_j(x_j), \tag{6}$$

where $u_j(x_j)$ is the normalized utility aggregate of an attribute j and $\bar{w}_j$ is the weight of that attribute [30, 31].

Once the problem hierarchy has been defined, a comparative judgment of the criteria is performed. The relative importance of two criteria is compared on a scale from 1 to 9 in the judgment matrix, where 1 is least important and 9 is most important criteria.

If there is a set of n criteria (C), the judgment matrix A is formed as an $n\ x\ n$ matrix, $A = [a_{i,j}]$, where $i,j \in \{1, 2, \ldots, n\}$. In the judgment matrix, the relative importance of criterion C_i compared to criterion C_j is represented by the element $a_{i,j}$ [17]. When $i = j$, the value of $a_{i,j}$ is 1, as a criterion is equally important when compared to itself. Additionally, the matrix is reciprocal, meaning that $a_{i,j} = 1/a_{j,i}$.

Matrix A is considered consistent if $a_{i,k} = a_{i,j} \cdot a_{j,k}\ \forall i, j, k \in \{1,2, \ldots, n\}$ [29]. This condition typically holds only in cases where the relative importance values are derived from exact measurements.

In the case of exact measurements, where the precise relative importance of different criteria is known, the elements of the judgment matrix A can be expressed as $a_{i,j} = w_i/w_j$, where w_i represents the priority of the criterion C_i. The judgment matrix A can then be written as [17]

$$\mathrm{A} = \begin{bmatrix} \frac{w_1}{w_1} & \frac{w_1}{w_2} & \cdots & \frac{w_1}{w_n} \\ \frac{w_2}{w_1} & \frac{w_2}{w_2} & \cdots & \frac{w_2}{w_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{w_n}{w_1} & \frac{w_n}{w_2} & \cdots & \frac{w_n}{w_n} \end{bmatrix} \tag{7}$$

For a consistent matrix, the priority vector w is an eigenvector as well as priority vector of A, and the eigenvalue decomposition satisfies the equation

$$Aw = nw. \tag{8}$$

In practical applications, however, it is often impossible to determine the exact relative importance of criteria. Instead, the relative importance values are estimated, introducing some inconsistency into the judgment matrix. In such cases, the following approximation is used

$$A'w' = \lambda_{max} w', \tag{9}$$

where A' is perturbation of A and λ_{max} is the largest eigenvalue of the matrix A'. The closer λ_{max} is to n, the more consistent the matrix is. The consistency of A' may be examined by calculating the consistency index ($C.I.$) for A'

$$C.I. = \frac{\lambda_{max} - n}{n - 1}. \tag{10}$$

The $C.I.$ is then compared to random index $R.I.$ to obtain the consistency ratio $C.R.$, a relative measure of consistency

$$C.R. = \frac{C.I.}{R.I.}. \tag{11}$$

The random consistency index R.I. is the average C.I. calculated for randomly generated same-dimension judgment matrices. Average random consistency indices are presented in table 1 according to different dimensions of A [28].

Table 1: Average random consistency index

n	1	2	3	4	5	6	7	8
R.I.	0	0	0.52	0.89	1.11	1.25	1.35	1.40

The normalized criteria weights $\bar{w}$ are then derived from the elements of λ_{max}

$$\bar{w} = \frac{w}{\sum_{i=1}^{n} w_i}. \tag{12}$$

As a result of this process, the priority vector $\bar{w}$ is obtained, representing the weights of each criterion. These weights can then be used to rank the alternatives and identify the most suitable option.

3.2 Operation Point (OP) Scaling Algorithm

The OP scaling algorithm is used for predicting the operation points of alternative EMAs with differing maximum force and velocity constraints for a set task from single measurement of the intended duty cycle. The measured cycle may be seen as the power demand for the actuator and the algorithm tries to estimate where the operating points are located when the device's force or velocity capacity is exceeded in the original dataset.

As inputs from the measured duty cycle, the algorithm uses force F, velocity v, and gravitational load force component F_0, together with the maximum force constraint of studied actuator F_{max}. An example force constraint curve is depicted in red in fig. 8.

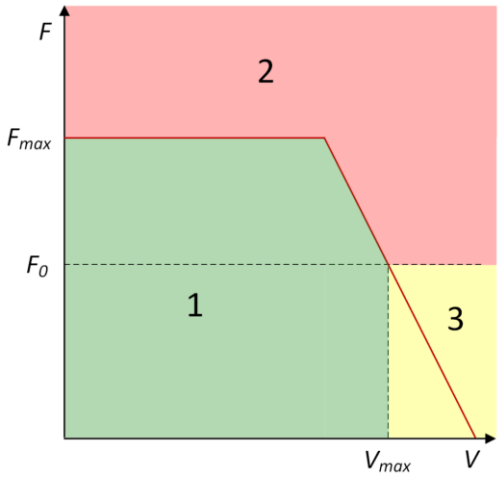


Figure 8: Example of force limit of actuator and operation point categorization

The measured operation points are reflected on the first quadrant and then divided into three categories by their location on the force-velocity graph:

1. Reachable operating points (Green).
2. Unreachable operating points constrained by maximum force (Red).
3. Unreachable operating points constrained by maximum velocity (Yellow).

The reachable operating points are not constrained by the force and velocity constraints; thus operating points inside area 1 are transferred into the set of scaled operating points unchanged.

The operating points constrained by maximum force are mapped on the force constraint curve such that the work done in one timestep into the system is equal for both; the measured and the scaled down operation points. For mapping these, similar force dynamics are assumed. The resulting scaled down operation points will have variable timesteps in order to compensate for the different force constraint and therefore, lower instantaneous power level. The steps of the algorithm for OPs in unreachable operating points constrained by maximum force (area 2) are graphically presented in fig. 9.

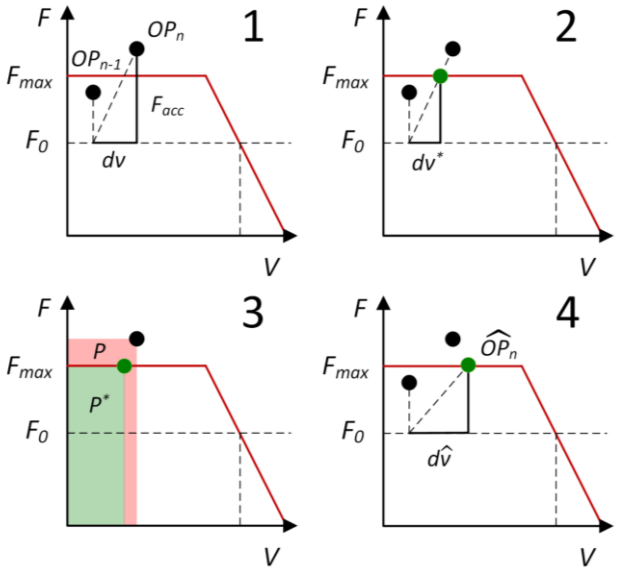


Figure 9: Steps of OP scaling algorithm for OPs in unreachable operating points constrained by maximum force

The algorithm assumes that the measured system follows the following finite difference approximation of the kinematic equation,

$$\frac{v(n) - v(n-1)}{\Delta t} = m_{eq}(n) \cdot F_{acc}(n), \tag{14}$$

where v is the velocity value of the operating point from measured dataset, Δt is the sampling time of measurement, and F_{acc} is the accelerating component of the force at actuator rod. The algorithm estimates the equivalent mass $m_{eq}(n)$ during one timestep, by solving gradient $k(n) = \Delta t \cdot m_{eq}(n)$ from the equation (14) depicted in step 1 of fig. 9 with formulae

$$k(n) = \frac{v(n) - v(n-1)}{F_{acc}(n)}. \tag{15}$$

Assuming that the difference in velocity is directly proportional to the accelerating component of force, the difference in velocity dv^* after the sampling time with the force constraint F_{max} is given by (fig. 9 step 2) the following formulae

$$dv^*(n) = k(n)\left(F_{max}\left(\hat{v}(n-1)\right) - F_0(n)\right). \tag{16}$$

After the original sample time Δt, the alternative actuator has done a different amount of work into the system compared to the actuator used for the measurement, thus, to end up at equal state energetically the new force must be applied for longer time. The following relationship between power and time can be established if the actuators output equal amount of energy.

$$\frac{\Delta\hat{t}(n)}{\Delta t} = \frac{P(n)}{P^*(n)} = \frac{v(n)F(n)}{\left(\hat{v}(n-1) + dv^*(n)\right)\hat{F}(n-1)}, \tag{17}$$

where $\Delta\hat{t}$ denotes the new timestep between scaled down operating points n and $n-1$, $\hat{v}$ and $\hat{F}$ are the velocities and forces of scaled down operating points, respectively. Whereas P and P^* are output powers of measured and scaled down actuators, respectively at a measurement sampling time of Δt after $n-1$ sample (fig. 9 step 3). The new velocity after the new time step $\Delta\hat{t}$ can be calculated as (fig 9 step 4)

$$d\hat{v}(n) = \frac{P(n)}{P^*(n)}\, k(n) \cdot \left(F_{max}\left(\hat{v}(n-1)\right) - F_0(n)\right) \tag{18}$$

and, therefore, the new scaled operating point is

$$\left[\hat{v}(n), \hat{F}(n)\right] = \left[\hat{v}(n-1) + d\hat{v}(n),\ F_{max}\left(\hat{v}(n-1)\right)\right]. \tag{19}$$

The operating points constrained by maximum velocity seen in fig. 8 are mapped on the maximum velocity border denoted by v_{max}, since for operation points with positive powers cannot cause additional increases in velocity as actuator in area 3 is not able to produce positive net force. For operation points with negative power values, the area 3 shall not be entered even though it is theoretically reachable due to risk of overrunning loads. Therefore, the OPs in area 3 are mapped on the velocity limit seen in fig. 8. Thus, the scaled OP is

$$\left[\hat{v}(n), \hat{F}(n)\right] = \left[v_{max}\left(F_0(n)\right), F(n)\right] \tag{20}$$

and the new time step between samples n and $n-1$ is

$$\Delta\hat{t}(n) = \frac{v(n)}{\hat{v}(n)}\Delta t. \tag{21}$$

The algorithm's estimated operation points and cycle time estimation were verified against simulation data of a 1 DOF crane realized in Simulink environment and the experimentally verified EMA model.

4 Algorithm Example Use Case: Forest Forwarder

This section demonstrates the use of the selection framework through an example. The goal is to optimize the combination of lift and tilt actuators by minimizing lifetime costs and load while maximizing productivity in a forest forwarder described in section 2.1.1. The PMSM servo motor rated power ranging from 3.2 to 11.6 kW, gear ratios of 3 and 7, and linear

transmission screw pitch of 10, 15, and 20 mm are selected for the forest forwarder's lift and tilt cylinders using the actuator selection algorithm [25, 27]. This makes a total of 48 alternative EMA configurations for the machine, and in turn a total combination of 2304 (48x48) to select from. The main parameters of servo motors used in the selection are presented in table 2.

Table 2: The specifications of the servo motors used

	A	B	C	D	E	F	G	H
Rated Power (kW)	3.2	6.0	8.3	10.4	10.7	11.0	11.4	11.6
Rated Torque (Nm)	10	19	27	33	34	35	36	37
Stall Torque (Nm)	11	23	34	45	54	63	71	77
Peak Torque (Nm)	34	68	101	134	162	189	213	231

To evaluate the alternative actuators' productivity, the total load on the boom structure at the lift and tilt booms joint and the total investment cost were used as top-level criteria. These criteria were organized in the hierarchical structure shown in fig. 10.

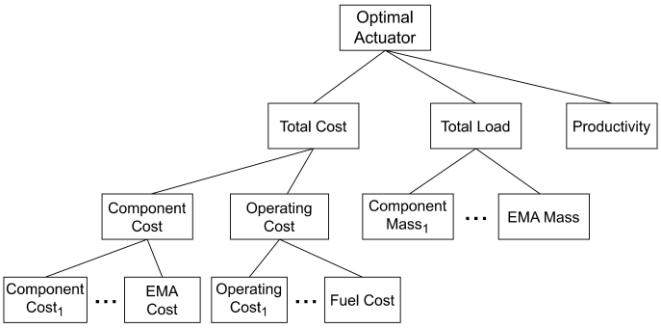


Figure 10: Hierarchy used in forest forwarder example

The fundamental idea of hierarchical modeling is to break the problem down into its elementary parts. The total load can be considered as the sum of all component masses connected to the boom. Similarly, the total cost can be divided into elementary cost components, some of which are influenced by the actuator selection. However, the forwarder's boom operation involves the simultaneous operation of multiple actuators. If a single actuator cannot complete its task in time, the cycle time increases directly. Therefore, this hierarchical approach does not apply to the productivity of the forwarder.

To evaluate the performance of each actuator alternative, the net present value of the investment cost, torque at the lift and tilt booms joint, and forwarder productivity are used. The duty cycle data for evaluation is generated using the simulation model of the forest forwarder. The collected data is input into the OP scaling algorithm, and the resulting operating points are used to calculate the fuel cost estimate during the duty cycle using the eq. (22). It shall be noted that all the values and numbers used in this example and results throughout are taken from the existing references [32, 33, 34]. Moreover, the actual values with respect to each machine case are presented in next section, focusing on results. The said values and numbers must be carefully considered for individual case. The cost calculations are as follows,

$$Cycle\ Cost = Cost_D \sum_n \left(\frac{\hat{v}(n) \cdot \hat{F}(n) \cdot \Delta\hat{t}(n)}{\eta_{EMA}\left(\hat{v}(n), \hat{F}(n)\right) \cdot \eta_I \cdot \eta_G} \right), \quad (22)$$

where $Cost_D$ is the cost of diesel per unit energy, the η_{EMA} represents the actuator's efficiency obtained from the EMAs efficiency map, η_I is the efficiency of the inverter, and η_G is the average efficiency of electric power generation from calorific value of diesel fuel. The total net present value of diesel consumption caused by the actuator over observation period used for evaluating the operation cost is,

$$NPV_{FC} = \sum_{y=0}^{y_{end}} Cycle\ Cost\ (Cycles/Year)\ (1 - k_D)^y\ , \quad (23)$$

where the k_D is the discount factor of future investments and y denotes the year when the cost occurs. For the component cost, public sources are used to estimate the servo motors cost, and the cost of the gear reduction system is set as constants for all alternatives.

If the productivity of the NRMM is defined as the time to complete a set duty cycle, cycle time can be used to evaluate it. To measure the productivity of an EMA alternative, the total estimated time to complete the duty cycle, as estimated by the OP scaling algorithm, is used. Since actuators work concurrently, the maximum increase in cycle time is considered at each time step.

$$t_{tot} = \sum_n \max\left(\Delta\hat{t}_{Lift}(n), \Delta\hat{t}_{Tilt}(n)\right). \quad (24)$$

The actuators' impact on the total load is estimated by EMA alternatives torque at the lift and tilt booms joint,

$$T_{EMA} = mean(d) \cdot m_{ema}, \quad (25)$$

where $mean(d)$ is the torque-inducing component of distance between the joint and actuator center of mass during the duty cycle and m_{ema} is the mass of the actuator alternative.

The thermal behavior of the servo motor is used as a constraint. The fulfilment of thermal constraint is assessed by the RMS value of the OP, which must be inside the nominal operating area of the motor to satisfy the constraint.

The aggregates are then normalized by reference values which are defined by the DM. The reference values are set for the duty cycle time t_{cycle}, the investment cost $Cost_{ref}$, and the mean torque at the lift and tilt booms joint during the duty cycle $mean(T_{Load})$.

Next, the criteria weights are determined using AHP, which evaluates the attribute weights w_j through pairwise comparisons between criteria. The pairwise comparison w_1/w_2 represents the ratio of partial differentials of the preference function relative to the aggregate values.

$$\frac{\frac{\partial P[F(X)]}{\partial x_1}}{\frac{\partial P[F(X)]}{\partial x_2}} \approx \frac{W_1}{W_2}. \quad (27)$$

To aid in assigning the values accurately in the AHPs pairwise comparisons, the normalization ranges with effect to utility between shared metric may be used if a known shared metric M_{12} exists of which gradient is,

$$\nabla_u M_{12}(u_1(x_1), u_2(x_2)) \approx \frac{\partial P[F(X)]}{\partial x_1} + \frac{\partial P[F(X)]}{\partial x_2}, \quad (28)$$

By evaluating the shared metric at the normalization points, the pairwise comparison can be calculated using formulae,

$$\frac{M_{12}(1,0) - M_{12}(0,0)}{M_{12}(0,1) - M_{12}(0,0)} = \frac{W_1}{W_2}. \quad (29)$$

The pairwise weight between productivity W_P and investment cost W_C is determined using a reference value for cost and an estimate for expected economic output based on typical forest forwarder values from the literature, as explained beforehand.

$$\frac{W_P}{W_C} = M_{PC} = \frac{NPV_{output}}{Cost_{ref}} \quad (30)$$

The pairwise comparison between total load and productivity is estimated based on the load's effect on machine productivity. While the load from the boom structure reduces productivity in a 1:1 relationship, the additional actuator mass also increases power consumption. Therefore, only a lower limit for the weight parameter can be estimated.

$$\frac{W_L}{W_C} \geq M_{LP} = 1. \quad (31)$$

For the pairwise importance between the remaining criteria, total cost, and total load, no suitable metric is found. With additional information, it may be possible to derive; however, for the example a consistent judgement matrix is assumed, and the weight parameter is solved from the already set weight factors. This assumption leads to a complete reliance on the previous weighting factors, and consistency of the judgement cannot be analyzed. Subsequently, weights of 1 are used between the different cost components.

5 Results

Based on the use case example presented in the preceding section, the EMA alternatives on both the machine cases are assessed. The following sub-sections present the results of forest forwarder and excavator cases. The results are represented as non-dominated solutions for the selection problem. A non-dominated solution is one where no objective can be improved without worsening another, thus representing an optimal trade-off. Solutions that are worse in all criteria compared to a non-dominated solution are excluded.

5.1 Forest Forwarder

For the actuator selection of the forwarder case, the following initial values were used, as presented in Table 3.

Table 3: Initial values used in forest forwarder actuator selection

Designation	Denotation	Value	Unit
$Cost_D$	Cost of diesel	$8.88 \cdot 10^{-8}$	€/J
t_{cycle}	Cycle time reference	41	s
-	Yerly operating time	2500	h
$Cost_{ref}$	Cost reference value	$1.08 \cdot 10^6$	€
NPV_{output}	Economic output	$1.76 \cdot 10^6$	€
y_{end}	Considered time	7	y
η_G	Engine-generator set efficiency	0.35	-
η_I	Inverter efficiency	0.95	-
k_d	Discount factor	0.1	-

Using the initial values, the following criteria weights are calculated $\bar{w}_C = 0.1701$, $\bar{w}_L = 0.5532$, and $\bar{w}_P = 0.2766$, where the factors are weights of total cost, total load, and total productivity, respectively. In the tilt function all PMSM alternatives fulfilled the RMS OP constraint unlike in the lift function, where out of the 8 PMSM alternatives only 5 (D-H) fulfilled the constraint using any effective transmission ratio.

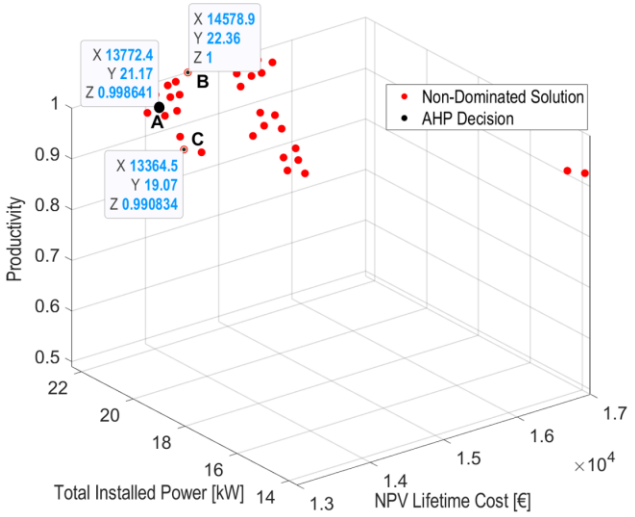


Figure 11: Non-dominated solutions for the forest forwarder

In fig. 11 productivity, total installed power and total investment cost values for all non-dominated lift and tilt actuator combinations are presented. Point A represents the optimal tradeoff given by AHP, based on weights defined here.

The A solution corresponds to a motor option with 10.7 kW nominal power using a gear ratio of 7 and a linear transmission pitch of 15 mm in the lift actuator, and in for the tilt actuator nominal power of 10.4 kW gear ratio of 3 and linear transmission pitch of 15 mm. Point B is the productivity maximizing solution at lowest lifetime cost. The solution B corresponds to 11.6 kW, 7 and 20 mm, and 10.7 kW, 3 and 15 mm for the lift and tilt cylinders nominal power, gear ratio and linear transmission pitch, respectively. Point C is a solution that minimizes the lifetime costs. It corresponds to 10.7 kW, 7 and 15 mm, and 8.3 kW, 3 and 15 for the lift and tilt cylinders nominal power, gear ratio and linear transmission pitch, respectively.

The results show that the lifetime costs of the system are higher than the lifetime costs for the productivity maximizing solution B for a large proportion of compared EMA combinations even with systems with lower component costs. This indicates that under-sizing the electric motor increases the lifetime costs of the system, which is explained by the lower efficiency of the EMA operating at higher velocities.

5.2 Excavator

In the excavator example use case the selection algorithm is used to select boom, arm and bucket cylinders resulting in 48^3 alternatives for the system, however only 6 (C-H) of the total 8 PMSM alternatives fulfilled the thermal constraint using any effective transmission ratio on the boom function. The selection hierarchy is applied as in the forest forwarder case, although the initial values—such as economic output and cost reference values—are updated using data from public sources [35, 36]. The adjusted initial values are presented in Table 4.

Table 4: Initial values used in excavator actuator selection

Designation	Denotation	Value	Unit
t_{cycle}	Cycle time reference	18	s
-	Yerly operating time	1800	h
DR	Duty Ratio	0.5	-
$Cost_{ref}$	Cost reference value	$3.71 \cdot 10^5$	€
NPV_{output}	Economic output	$5.64 \cdot 10^5$	€

Using the initial values, the following criteria weights are calculated: $\bar{w}_C = 0.1799$, $\bar{w}_L = 0.5468$, and $\bar{w}_P = 0.2734$, where the factors are weights of total cost, total load and total productivity, respectively. In fig. 12 the resulting non-dominated combinations for boom, arm, and bucket cylinders are presented. Point A represents the optimal trade-off determined by AHP based on the present weights. Solution A corresponds to 11.6 kW, 3 and 10 mm, 10.4 kW, 3 and 10 mm and 6.0 kW, 3 and 10 mm for the 2 boom, arm, and bucket cylinders' nominal power, gear ratio and linear transmission pitch, respectively. Point B maximizes productivity at the lowest lifetime cost. Solution B corresponds to 11.6 kW, 3 and 10 mm, 11.0 kW, 3 and 10 mm and 6.0 kW, 3 and 10 mm for the 2 boom, arm, and bucket cylinders nominal power, gear ratio and linear transmission pitch respectively. Point C minimizes lifetime costs. Solution C corresponds to 10.7 kW, 7 and 20 mm, 8.3 kW, 3 and 15 mm and 3.2 kW, 7 and 20 mm for the 2 boom, arm, and bucket cylinders' nominal power, gear ratio and linear transmission pitch respectively.

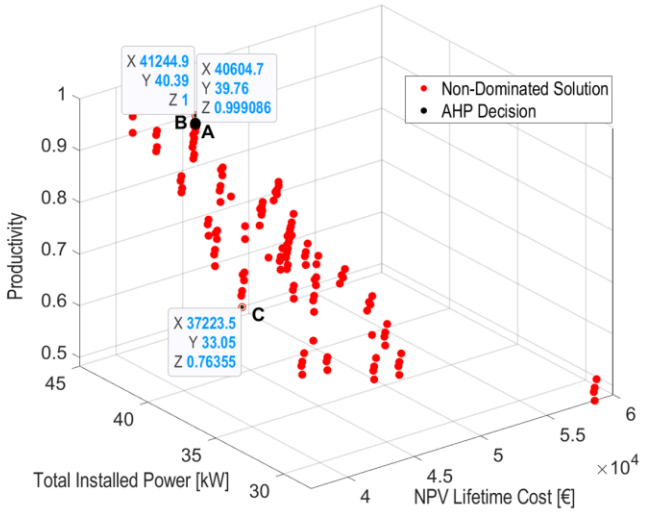


Figure 12: Non-dominated solutions for the excavator

In the excavator's case, a similar trend is seen where productivity tends to decrease as lifetime costs increase after point B. The selected actuation system is close to the productivity maximizing solution in terms of values selected for representation, however in the selection the mean position of the increased weight of the actuator is considered, thus on the mean load axis the difference between the points is greater.

6 Conclusion

In this paper, a multi criteria selection framework for selection of electro-mechanical actuators (EMA) for non-road mobile machinery is presented. In the framework operation points of actuator alternatives are estimated according to their force constraints using novel operation point (OP) scaling algorithm and an analytic hierarchy process (AHP) based Multi-Attribute Decision Making (MADM) algorithm to select the optimal actuator and the combination in machine.

As the case study, two NRMMs are used, a forest forwarder and an excavator. For the forwarder's lift and tilt functions, most of the alternative solutions evaluated exhibited a higher lifetime cost compared to the productivity-maximizing alternative, indicating how oversizing or undersizing can significantly affect costs. The best trade-off solution from the initial 2304 alternatives had nominal powers of 10.7 kW for the lift and 10.4 kW for the tilt, with respective gear ratios of 7 and 3, and linear transmission pitches of 15 mm each. By comparison, the excavator scenario produced 110,592 alternative configurations, where the best trade-off solution had nominal powers of 11.6 kW for the two boom actuators, 10.4 kW for the arm actuator and 6.0 kW for the bucket actuator, with gear ratios of 3, and linear transmission pitches of 10 mm for all actuators. The resulting non-dominated solutions for the excavator indicate that productivity tends to drop sharply when undersizing the actuators, showcasing the effect of higher utilization rate of each actuator in the measurement data on the productivity score.

These results indicate that optimizing one criterion often compromises another. Productivity emerges as a vital metric, as lifetime energy costs can outweigh initial investments. The chosen weights favor reduced lifetime costs and weight over minimal productivity losses, underscoring the importance of balancing multiple factors for EMA selection in NRMM.

In the future, more features can be added to the evaluation criteria. Particularly, thermal behavior should be included using the estimated OPs. Additionally, larger data samples should be utilized to ensure fully representative input data for decision-making, as there is a risk of over-optimization to match the performance needed for the input duty cycle. The applicability of input measurement data gathered from conventional hydraulically powered NRMMs should be tested. Unfortunately, such data was not available for this study. The concept is also extendable for electro-hydrostatic actuators (EHAs) with minor modifications, and with larger modifications, it can be studied for selecting between actuator technologies.

Acknowledgments

This research is a part of Business Finland project EMMA-2 (2471/31/2021). The authors gratefully acknowledge this funding and support by Business Finland.

References

[1] Casoli, P., Riccò, L., Campanini, F., & Bedotti, A. (2016). Hydraulic hybrid excavator - Mathematical model validation and energy analysis. Energies (Basel), 9(12), 1002-. https://doi.org/10.3390/en9121002

[2] XU, B., & CHENG, M. (2018). Motion control of multi-actuator hydraulic systems for mobile machineries: Recent advancements and future trends. Frontiers of Mechanical Engineering, 13(2), 151–166. https://doi.org/10.1007/s11465-018-0470-5

[3] Volvo CE unveils 100% electric compact excavator prototype. (2017, May 16). Volvo Construction Equipment. Retrieved February 13, 2025, from https://www.volvoce.com/global/en/news-and-events/news-and-stories/2017/volvo-ce-unveils-100-percent-electric-compact-excavator-prototype/

[4] Bahari, M., Paz, A., & Mattila, J. (2024). System-Level Performance Metrics Sensitivity of an Electrified Heavy-Duty Mobile Manipulator. IEEE Vehicular Technology Conference, 1–5. https://doi.org/10.1109/VTC2024-Fall63153.2024.10757845

[5] Qu, S., Fassbender, D., Vacca, A., & Busquets, E. (2021). A high-efficient solution for electro-hydraulic actuators with energy regeneration capability. Energy (Oxford), 216, 119291-. https://doi.org/10.1016/j.energy.2020.119291

[6] Heybroek, K. (2017). On Energy Efficient Mobile Hydraulic Systems: with Focus on Linear Actuation. https://doi.org/10.3384/diss.diva-142326

[7] Goswami, G., Tupitsina, A., Jaiswal, S., Nutakor, C., Lindh, T., & Sopanen, J. (2022). Comparison of Various Hybrid Electric Powertrains for Non-Road Mobile Machinery Using Real-Time Multibody Simulation. IEEE Access, 10, 107631–107648. https://doi.org/10.1109/ACCESS.2022.3213034

[8] Hu, X., Murgovski, N., Johannesson, L. M., & Egardt, B. (2014). Comparison of Three Electrochemical Energy Buffers Applied to a Hybrid Bus Powertrain with Simultaneous Optimal Sizing and Energy Management. IEEE Transactions on Intelligent Transportation Systems, 15(3), 1193–1205. https://doi.org/10.1109/TITS.2013.2294675

[9] Murgovski, N., Johannesson, L., Sjöberg, J., & Egardt, B. (2012). Component sizing of a plug-in hybrid electric powertrain via convex optimization. Mechatronics (Oxford), 22(1), 106–120. https://doi.org/10.1016/j.mechatronics.2011.12.001

[10] Silvas, E., Hofman, T., Murgovski, N., Etman, L. F. P., & Steinbuch, M. (2017). Review of Optimization Strategies for System-Level Design in Hybrid Electric Vehicles. IEEE Transactions on Vehicular Technology, 66(1), 57–70. https://doi.org/10.1109/TVT.2016.2547897

[11] Padovani, D., Dimitriou, P., & Minav, T. (2024). Challenges and solutions for designing Energy-Efficient and Low-Pollutant Machines in Off-Road hydraulics. Energy Conversion and Management. X, 21, 100526-. https://doi.org/10.1016/j.ecmx.2024.100526

[12] Xue, L., Wu, S., Xu, Y., & Ma, D. (2019). A simulation-based multi-objective optimization design method for pump-driven electro-hydrostatic actuators. Processes, 7(5), 274-. https://doi.org/10.3390/pr7050274

[13] Poole, A. D., & Booker, J. D. (2011). Design methodology and case studies in actuator selection. Mechanism and Machine Theory, 46(5), 647–661. https://doi.org/10.1016/j.mechmachtheory.2010.12.009

[14] Zhang, Y., & Cai, M. (2010). Mechanical Properties Evaluation for Pneumatic Actuator and Electric Actuator Based on AHP-FCE. 2010 WASE International Conference on Information Engineering, 3, 75–79. https://doi.org/10.1109/ICIE.2010.196

[15] Fassbender, D., & Minav, T. (2021). An algorithm for the broad evaluation of potential matches between actuator concepts and heavy-duty mobile applications. Actuators, 10(6), 111-. https://doi.org/10.3390/act10060111

[16] Yu, B., Wu, S., Jiao, Z., & Shang, Y. (2018). Multi-objective optimization design of an electrohydrostatic actuator based on a particle swarm optimization algorithm and an analytic hierarchy process. Energies (Basel), 11(9), 2426-. https://doi.org/10.3390/en11092426

[17] Saaty, T. L. (1980). The analytic hierarchy process : planning, priority setting, resource allocation. McGraw-Hill.

[18] Jiao, Z., Yu, B., Wu, S., Shang, Y., Huang, H., Tang, Z., Wei, R., & Li, C. (2019). An intelligent design method for actuation system architecture optimization for more electrical aircraft. Aerospace Science and Technology, 93, 105079-. https://doi.org/10.1016/j.ast.2019.03.048

[19] Mevea. (n.d.). Mevea. Retrieved February 12, 2025, from https://mevea.com

[20] Abid Abdul Azeez, Zakharov, V., Orifjonov, A., Niemelä, J., & Minav, T. (2022). Multi-physics Co-simulation of an Electric Reach Truck. In K. Schmitz (Ed.), The 13th International Fluid Power Conference (pp. 438-447). International Fluid Power Conference. https://doi.org/10.18154/RWTH-2023-04575

[21] Geiger, C., & Geimer, M. (2017). Efficiency optimisation of a forestry crane by implement hydraulics with energy recovery. VDI Berichte, 2017(2300), 175–184.

[22] Zhang, S., Minav, T., Pietola, M., Kauranne, H., & Kajaste, J. (2019). The effects of control methods on energy efficiency and position tracking of an electro-hydraulic excavator equipped with zonal hydraulics. Automation in Construction, 100, 129–144. https://doi.org/10.1016/j.autcon.2019.01.003

[23] Japan Construction Mechanization Association. (2007). Earth-moving machinery—Fuel consumption on hydraulic excavator—Test procedures. JCMAS H 020.

[24] Qiao, G., Liu, G., Shi, Z., Wang, Y., Ma, S., & Lim, T. C. (2018). A review of electromechanical actuators for More/All Electric aircraft systems. Proceedings of the Institution of Mechanical Engineers. Part C, Journal of Mechanical Engineering Science, 232(22), 4128–4151. https://doi.org/10.1177/0954406217749869

[25] Pustavrh, J., Hočevar, M., Podržaj, P., Trajkovski, A., & Majdič, F. (2023). Comparison of hydraulic, pneumatic and electric linear actuation systems. Scientific Reports, 13(1), 20938–20938. https://doi.org/10.1038/s41598-023-47602-x

[26] Krishnan, R. (2001). Electric Motor Drives: Modeling, Analysis, and Control. Prentice-Hall.

[27] UNIMOTOR FM - AC Brushless Servo motor. (n.d.). Nidec Corporation. Retrieved February 13, 2025, from https://cdn-static.nidec-netherlands.nl/media/3693-servomotoren-unimotor-fm-brochure-uk-iss13-0702-0032-13.pdf

[28] High performance actuator catalogue. (2025, January). EWELLIX. Retrieved February 13, 2025, from https://medialibrary.ewellix.com/asset/16172

[29] Saaty, T. L., & Vargas, L. G. (2012). Models, Methods, Concepts & Applications of the Analytic Hierarchy Process (2nd ed. 2012, Vol. 175). Springer Nature. https://doi.org/10.1007/978-1-4614-3597-6

[30] Marler, R. T., & Arora, J. S. (2010). The weighted sum method for multi-objective optimization: new insights. Structural and Multidisciplinary Optimization, 41(6), 853–862. https://doi.org/10.1007/s00158-009-0460-7

[31] Herrmann, J. W. (2015). Engineering decision making and risk management (1st ed.). Wiley. ISBN: 9781118919330

[32] Malinen, J., Laitila, J., Väätäinen, K., & Viitamäki, K. (2016). Variation in age, annual usage and resale price of cut-to-length machinery in different regions of Europe. International Journal of Forest Engineering, 27(2), 95–102. https://doi.org/10.1080/14942119.2016.1171964

[33] Eriksson, M., & Lindroos, O. (2014). Productivity of harvesters and forwarders in CTL operations in northern Sweden based on large follow-up datasets. International Journal of Forest Engineering, 25(3), 179–200. https://doi.org/10.1080/14942119.2014.974309

[34] Dvořák, J., Jankovský, M., Chytrý, M., Nuhlíček, O., Natov, P., Kormanek, M., & Löwe, R. (2021). Operational costs of mid-performance forwarders in czech forest bioeconomy. Forests, 12(4), 435-. https://doi.org/10.3390/f12040435

[35] Excavator Total Cost of Ownership Calculator. (n.d.). A Volvo Construction Equipment Blog. Retrieved February 13, 2025, from https://volvoceblog.com/tco-calculator/

[36] McGrillen, C. (2017). Case Study: Diesel Excavators Compared to Hybrid Excavators.

SLM 3D printed proportional directional water hydraulic valve, tested with optimized mounting plate

Jan Bartolj[1], Ana Trajkovski[1], Marko Hočevar[1], Katharina Schmitz[2] and Franc Majdič[1]*

1. Laboratory for fluid power and controls, Faculty of Mechanical Engineering, University of Ljubljana, Aškerčeva ulica 6, 1000 Ljubljana, Slovenija
2. Institute for Fluid Power Drives and Systems (IFAS), RWTH Aachen University, 52074 Aachen, Germany

E-mail: jan.bartolj@fs.uni-lj.si

Abstract

In this study, we investigate the influence of pressure transducer installation and mounting plate geometry on the accuracy of flow characterisation in a 3D printed proportional directional control valve system. The valve itself, which was developed in our previous research work, has been confirmed numerically and experimentally. To allow measurements to be as accurate as possible, we designed and manufactured highly optimised mounting plate with integrated pressure transducers using SLM additive manufacturing process.

The main objective was to assess whether the position of pressure transducer and the internal geometry of the mounting plate affect the accuracy of the measurement and cause interference that could affect the determination of the flow characteristics of the valve. Conventional and 3D printed mounting plate configurations were tested under controlled conditions, including versions with external and integrated pressure transducers. Experiments were carried out using two fluids: ISO VG 46 mineral hydraulic oil and standard tap water, which allowed us to estimate performance under flow conditions with high and low viscosity.

The results of our measurements have shown that 3D printed mounting plate with integrated pressure transducers provided the lowest pressure drop in the system and allowed the most accurate characterization of the valve performance. Specifically, the integrated transducer configuration reduced pressure drop by 47 % compared to settings with external transducers. Reduction of pressure drop significantly reduced the effect that disturbing factors have on measurement, resulting in more reliable results. In addition we determined that location where transducer is integrated has great influence and it becomes increasingly important at higher flow rates, where differences in measurement location resulted in more significant deviations. The study highlights the benefits of additive manufacturing in creating application-specific hydraulic components that improve transducer integration and by that measurement accuracy.

Keywords: water hydraulics, valve, mounting plate, optimisation, additive manufacturing

1 Introduction

Recently, it has become clear that systems, machines and parts should be as efficient as possible in order to conserve material and energy resources. This is being repeated constantly in recent years, but fulfils a lot of global climate and continual development goals. Especially in the field of hydraulics, the use of new manufacturing technologies has proven to be extremely useful and in direct correlation with the development of components. Additive manufacturing (AM) allows users to make a variety of improvements based on more or less extensive numerical optimization [1].

2 Theoretical background

Nowadays, it should no longer be a question of whether to use optimisation software or not, but which one to use and to what extent. As AM is a way of manufacturing that imposes few to no constraints on designers, many extremely optimised components and systems have already been introduced in the field of hydraulics [2-4] and beyond.

2.1 Additive manufacturing

When we talk about AM in hydraulics, we emphasize 5 main advantages. AM allows us to:

1. Create complex geometries:
    - Internal and external geometry optimisation,
    - Customization.
2. Create lightweight, optimised designs:
    - Reduce the volume of parts compared to conventional manufacturing (CM), so that less material is used and the same functionality is achieved while less energy is consumed in manufacturing [5], all in the light of global climate policies.
3. Rapidly produce prototypes:
    - The production of prototypes and small series could be carried out quickly and cost-effectively.
4. Efficiently use base materials:
    - Since AM uses only the material that is critical to the structural integrity of the component, AM is much more efficient compared to CM. In addition, material that was not used in one process can be used in the next process.
5. Consolidate multiple parts into a single structure:
    - One of the strongest, but most overlooked points is definitely the consolidation of components. This allows designers to merge multiple components into a single body unit that has the same functionality, but in most cases is significantly more efficient.

All in all, AM is transforming the way that hydraulic components are designed and produced by enabling more efficient use of materials and enhancing performance. Main disadvantage could be found in the relatively high cost of machines for AM and so the process itself, but with the development and introduction of new and more optimised machines, combined with improved performance of components, even slightly higher cost of manufacturing should be justified [6].

2.2 Water hydraulics

Water hydraulics is a sub-discipline of hydraulics in which mineral hydraulic oil or other types of hydraulic fluids are replaced by pure tap water. Water has a higher density than the commonly used hydraulic fluids, but a lower viscosity. The use of water as a hydraulic fluid has many advantages and some disadvantages. The main advantage is environmental friendliness, because in the event of a leakage or replacement of the working fluid, the water can simply be disposed of as it has no negative impact on nature. The main advantages are therefore: environmental friendliness, availability and fire resistance. On the other hand, water poses some challenges that need to be overcome in order for users to replace mineral or synthetic oils with water, such as the use of suitable non-corrosive materials that have at least some self-lubricating properties. Since water also has a lower viscosity than oil, it is more susceptible to cavitation. The disadvantages drive up costs, so most industries still use mineral hydraulic oils, but there are exceptions where cleanliness or fire safety are of the most importance. It has already been shown in some studies that there are a number of materials that are compatible with water hydraulics and can also be interchanged between oil and water hydraulics [7]. In this study, authors have tried to compare the results of the same component used in oil and water hydraulics to ensure that it can be easily interchanged and used in both.

2.3 Pressure loss as an identificator of performance

The pressure drop could be determined mathematically with the help of equations. However, this approach only provides us with an approximation or a range in which we can expect the result of the measurement.

To determine the pressure drop as a function of the flow of the hydraulic fluid, the density and the cross-sectional area of the flow, we need to use the basic Bernoulli equation (1). In the Bernoulli equation, we have pressure, velocity and potential component, all of which must remain in correlation, regardless of how the geometry changes. This ensures basic continuity.

$$p + \frac{1}{2}\rho v^2 + \rho g h = const. \quad (1)$$

Derived from the Bernoulli equation of energy conservation (1) and using a shape coefficient of the orifice respectively, we get the equation from which we can determine the flow rate as a function of the pressure drop for a known orifice cross-sectional flow area (2).

$$q_v = C_d \cdot A_{orif} \sqrt{\frac{2 * \Delta p}{\rho}} \quad (2)$$

Also surface roughness of internal structure of 3D printed components is usually considered, but since our experience have shown that most of hydraulic pressure drop is due to local and not linear losses, we have neglected it in this study.

3 Experimental work

In 2022, 2 iterations of 3D printed highly optimised valve body geometries were developed in our Laboratory of fluid power and controls (*Slo: Laboratorij za fluidno tehniko, LFT*) [8,9]. The main focus of the recent work was on the design and CFD tests, while real tests were not performed on a large scale. In this paper, we will present the tests we conducted with an AM produced valve body and compare the different effects of using mineral hydraulic oil and pure tap water. We have also designed and 3D printed a hydraulic mounting plate with integrated pressure transducers to measure the pressure difference between the inlet and outlet channels of the valve. Until this new and improved base plate, we were not able to measure the pure pressure difference of the valve, but the measurement was affected by many couplings, hoses, T-junctions and the base plate itself, which were located between the transducer and the measured valve. This adds a lot of uncertainty to the measurement, which we tried to eliminate by placing the transducer as close as possible to the valve body. This meant that the measuring point had to be integrated

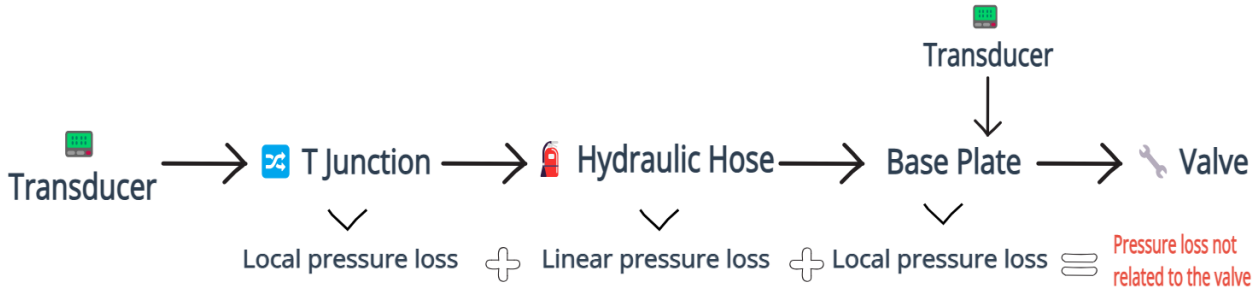


Figure 1: Measurement chain

into the base plate, close to the valve itself. We also tried to design the base plate to be as easy to handle as possible without the hydraulic couplings being too close to make screwing easier. The diagram in figure (1) shows the way we have installed the pressure sensors in the inlet branch of the valve. A pressure transducer is mounted at the beginning and has 3 elements (T-junction, hydraulic hose and base plate) that add to the pressure drop before measuring the pressure drop of the actual valve. We have therefore integrated pressure transducers directly under the valve itself (as shown by the second transducer connected directly to the base plate). We expect a slightly lower pressure drop compared to the measurement with the conventionally placed pressure transducer. Additional hydraulic losses are indicated below the measuring line of the pressure transducer on the far left which contribute to the measurement uncertainty and are avoided if a transducer is used below the valve body.

Testing was carried out at two testing facilities. Oil test facility has 20 kW of hydraulic power and was limited to 350 bar of pressure. For water hydraulic tests we used self-developed testing facility that has 11 kW of hydraulic power. Again, highest pressure with water hydraulic tests was limited to 350 bar that gave us approximately 30 l/min flow rate.

3.1 Valve and plate assembly

Figures 2 and 3 show the comparison between the conventionally manufactured (CM) and the 3D-printed base plate. The CM base plate is bulkier, and since all the connections for hydraulic fittings are from the underside, it is quite difficult to mount the pipes to close the hydraulic circuit. And since the pipes are mounted from below, the base plate cannot stand alone but must be attached to some kind of bracket. The base plate produced using additive manufacturing (AM) (Figure 3) is optimised for easier assembly and is much easier to handle as each connection points in its own direction. It can be operated as a stand-alone element, even when installed in a hydraulic circuit. The main advantage of the AM base plate is the possibility of integrating transducers close to the directional control valve that is mounted on it.

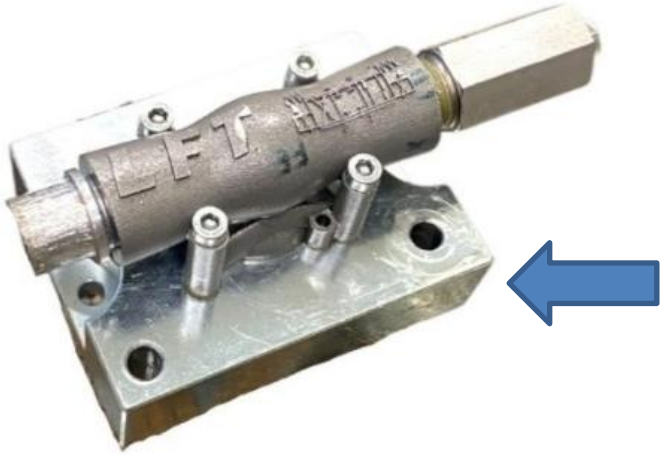


Figure 2: Conventionally manufactured (CM) base plate with spool valve attached.

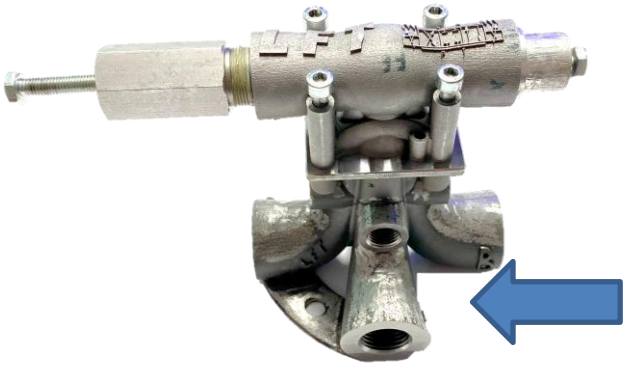


Figure 3: Additively manufactured (AM) base plate with spool valve attached.

As already mentioned, we used a hydraulic valve body that was optimised and 3D-printed in 2021 [8]. We used a spool that has 3 control notches with a notch width of 2 mm. To position the spool in the valve, we used two screws, one from each side, to ensure that the spool is not displaced by the flow forces. The structure and assembly of the valve is shown in the Figure 5 and consists of the following elements: 1. valve body, 2. core of the positioning screw, 3. positioning screw, 4. controlling spool, 5. washer that pushes the slider into the centre position, 6. stainless steel spring, 7. two 1 mm washers to ensure spring preload, 8. sealing washer, 9. core of the

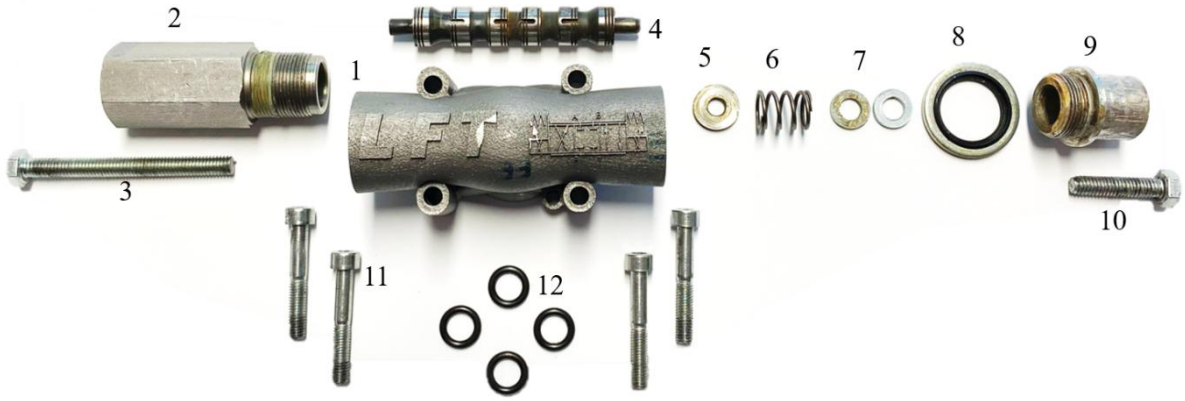


Figure 5: Directional spool valve assembly.

counterpositioning screw, 10. counter-positioning screw, 11. fixing screws, 12. static O-ring seals.

Figure 4 shows where the pressure transducers were installed in the mounting plate. The lower holes with the corresponding letters on the side, indicating the individual channels P, A, B and T, were used to connect the hydraulic G 1/2 fittings, and the upper holes were used to insert the standard Minimess fittings that connect the mounting plate to the pressure transducer. The transducers and all the measuring equipment were manufactured by Parker. We used Parker Service Master Plus (K-SCM series) to collect the data from the transducers and export the measurements.

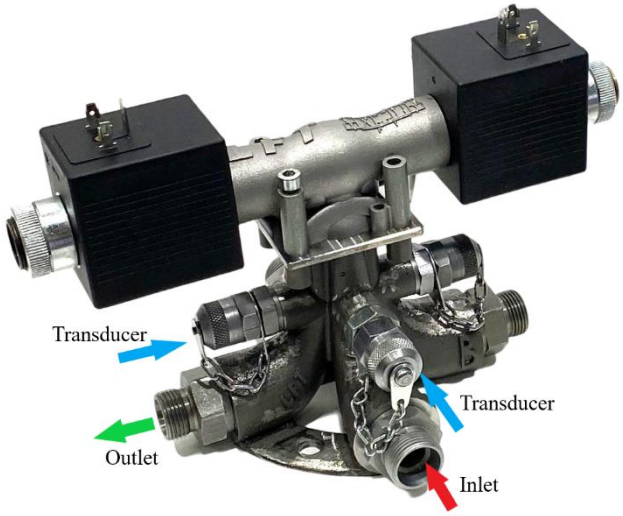


Figure 4: Installation of pressure transducers integrated into additively manufactured mounting plate.

Figure 6 shows the actual measurement setup with all connectors and measuring transducers fitted. The picture also shows the flow direction and the way in which the 4-point measurement of the mounting plate manufactured by AM was carried out.

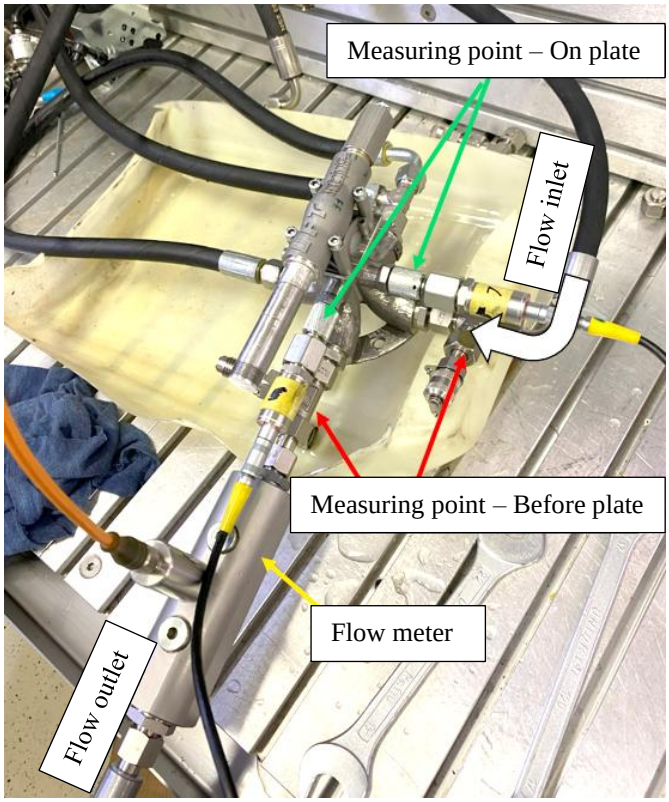


Figure 6: Valve assembly on the water testing rig, with pressure transducers and flow meter attached.

4 Results

During the measurement, we realized that the valve felt quite hot. We therefore used a thermal imaging camera (FLIR E6 Pro) to determine how much the valve and the AM plate were heated by the flow of ISO VG46 (Figure 7). The highest temperature measured during the tests was around 65°C at the valve body. The oil in the reservoir had a secondary cooling circuit that was set to maintain the temperature of the reservoir

at 40 °C, which was in perfect condition. The fluid flow heats up valve body due to the narrow paths and pressure drop since this is the main heat sink for approximately 22 kW of hydraulic power. We can assume that the spool inside the valve body heats up even more, as the notches are the places where the flow is most restricted due to the small cross-sectional area.

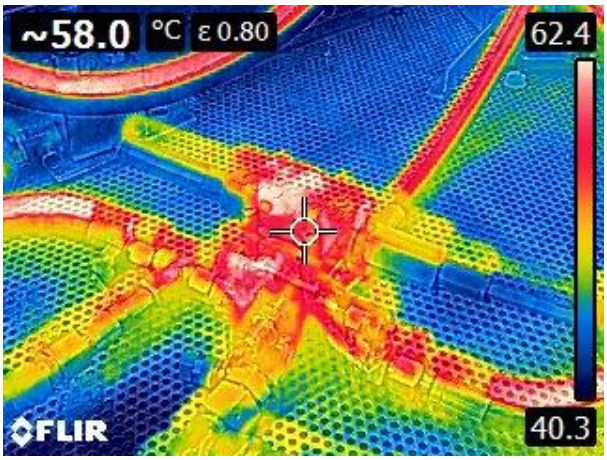


Figure 7: Image from thermal imaging camera, showing heat buildup in the valve and AM base plate assembly during ISO VG46 mineral hydraulic oil testing.

4.1 ISO VG46 measurement

Firstly, we present the measurement carried out with ISO VG46 mineral hydraulic oil using the CM and AM mounting plate. Simple calculation using equation 2, gives us estimated pressure drop of 67,8 bar at the control notches of the spool at 50 % opening and 30 l/min flow rate. AM mounting plate had 4 sensors measuring simultaneously, two on plate and two before plate, as it can be seen in the Figure 6. The pressure drop depends strongly on the position of the spool in the valve body and the flow rate of the liquid. The further away the spool is from the center position, the larger the cross-sectional area of the fluid flow and the lower the pressure drop through the spool control notches and the valve itself. Conversely, the higher the flow rate at constant spool displacement, the higher the pressure drop. In all cases with a higher flow rate (30 l/min), we measured a higher pressure drop (Figure 8) than with a lower flow rate (20 l/min). In all cases, the CM mounting plate performed the worst when comparing the same spool displacement. When comparing the measurement of the pressure drop before and on plate, we can see the similarity, as the difference is only about 1 to 2 bar pressure drop. At a flow rate of 30 l/min, we measured a constant or slightly increased pressure drop after 60 % opening. This was due to the effect of pressure compensator on the variable displacement pump. It tries to increase the flow rate and consequently increases measured pressure drop at the valve, but in the case of this measurement it was the disturbing factor. In future we should use a constant displacement pump.

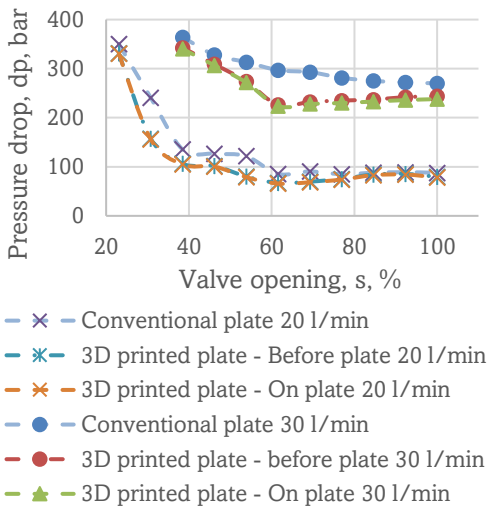


Figure 8: ISO VG46 mineral hydraulic oil measurement of pressure drop at 20 and 30 l/min at variable valve opening.

In the Figure 9, we have shown the pressure drop at two different opening states of the valve at 30 % and 60 %. Logically, a larger valve opening also reduces the pressure drop since flow area is larger. This figure also shows that the pressure drop at constant valve opening increases with increasing volume flow. A clear difference between the CM and AM mounting plate can also be seen here. The AM mounting plate significantly reduces the pressure drop, but the difference between the measurement before and on the plate is quite small.

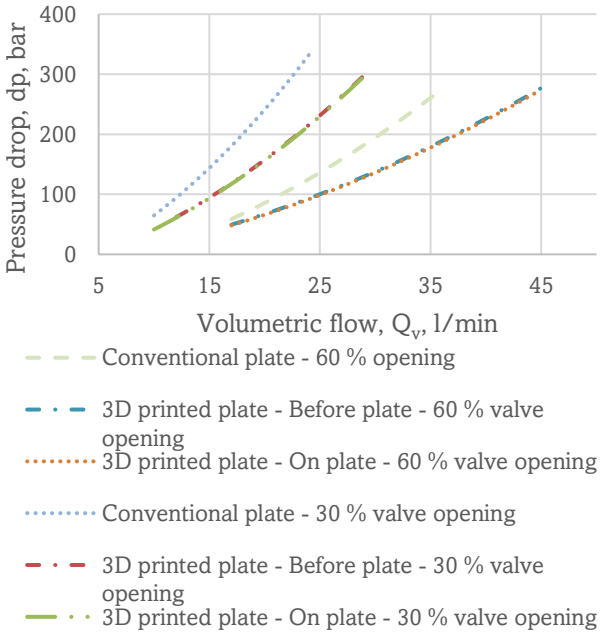


Figure 9: ISO VG46 mineral hydraulic oil measurement of pressure drop at 30 % and 60 % valve opening at variable flow rate.

To compare the differences between the measurement before and on the plate, we have plotted the deduction of the pressure

drop before plate and on plate (Figure 10). We can see that the difference between the two values increases almost linearly for both a 30 % and a 60 % opening of the valve. This means that the real added value of the optimised AM mounting plate becomes significant at the higher volumetric flows, which also leads to a more accurate measurement of the valve itself. However, when comparing the opening at the same flow rate, it can be seen that the difference between pressure drops upstream and downstream of the plate is quite similar, differing by only 8 % on average.

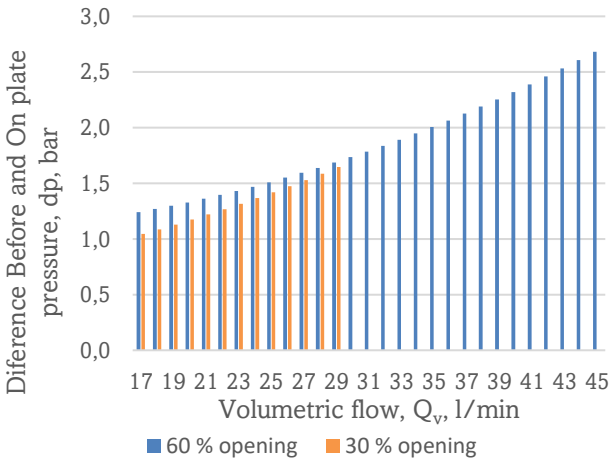


Figure 10: ISO VG46 mineral hydraulic oil. Calculated difference between measurements obtained during simultaneous 4-point measurements on AM mounting plate at 30 and 60 % valve opening and variable volumetric flow.

4.2 Tap water measurement

In the second part, we repeated the measurement, but using tap water as the flow medium and a specially adapted 11 kW water test system. The maximum capacity of this system is 30 l/min at pressure 350 bar. It would therefore not be possible to achieve the same properties as in the first tests with an oil system that has a capacity of 20 kW. Nevertheless, we expected a certain degree of comparability with the same flow rates that were achieved. Using equation 2, again gives us estimated pressure drop at the control notches of the spool. Using water properties at 50 °C and flow rate of 20 l/min gives us the 34,7 bar of pressure loss at 50 % opening. As our CM mounting plate was not intended for use with water, we were unable to measure the properties of both plates in water, but we did manage to do so with the AM plate.

The results show that the shape of the characteristic curve (Figure 11) is similar to the one we obtained when measuring with oil. The pressure drop is lower at the same flow rate when we open the valve and thus increase the flow area through the control notches. The measurements were carried out at a flow rate of 5 l/min and 20 l/min, with the data recorded both before and on plate at the same time.

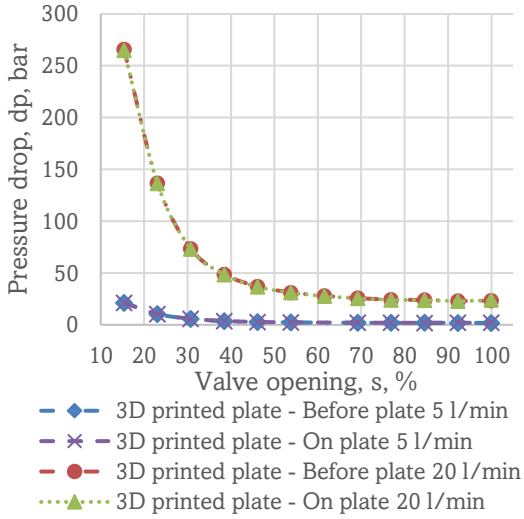


Figure 11: Tap water measurement of pressure drop at 5 l/min and 20 l/min at variable valve opening.

As with the first measurements, we set a valve opening of 30 % and 60 % and measured the pressure drop with increasing volumetric flow. Again, we measured an almost perfect second order curve of the pressure drop through the valve (Figure 12).

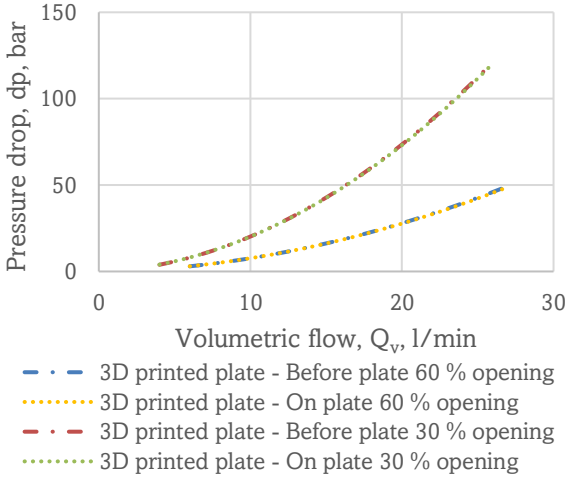


Figure 12: Tap water measurement of pressure drop at 30 % and 60 % valve opening at variable volumetric flow.

The difference between the measurements before and on the plate is shown in Figure 13. Here too, we can see that the difference increases at higher flow rates, similar to the oil measurement. It therefore makes sense to use this plate for more accurate measurements at higher flow rates.

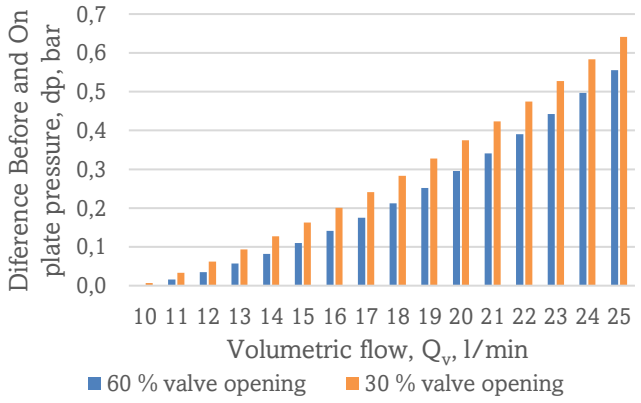


Figure 13: Tap water calculated difference between measurements obtained during simultaneous 4-point measurements on AM mounting plate at 30 % and 60 % valve opening and variable flow rates.

5 Comparing results of ISO VG46 and water

In this chapter we will discuss the differences between oil and water measurements. In Figure 14 we have shown the pressure drop for different flow rates at a constant valve opening of 30 %. It can be clearly seen that water has a lower pressure drop, which is directly related to the lower viscosity. Figure 15 shows a similar trend when comparing oil and water, but at a valve opening of 60 %, and the difference is almost identical. In all cases, it is again evident that the CM plate performs worst in absolute terms, while the 4-point simultaneous measurements of the AM plate always show very similar properties, which is to be expected.

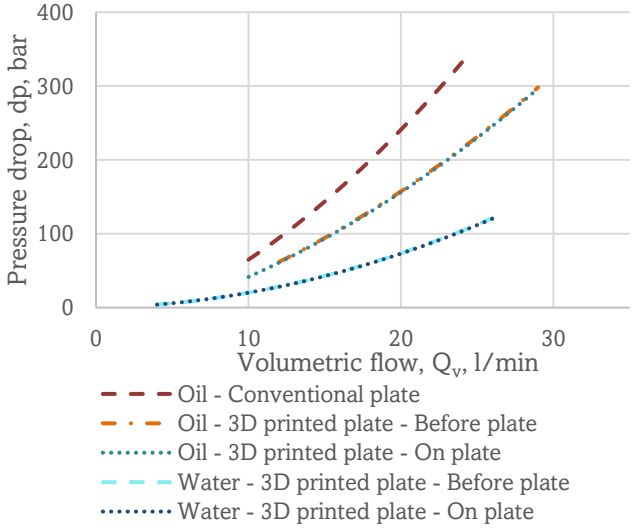


Figure 14: Comparison between water and ISO VG46 mineral hydraulic oil and CM to AM mounting plate measured characteristics at constant 30 % valve opening and variable flow rates.

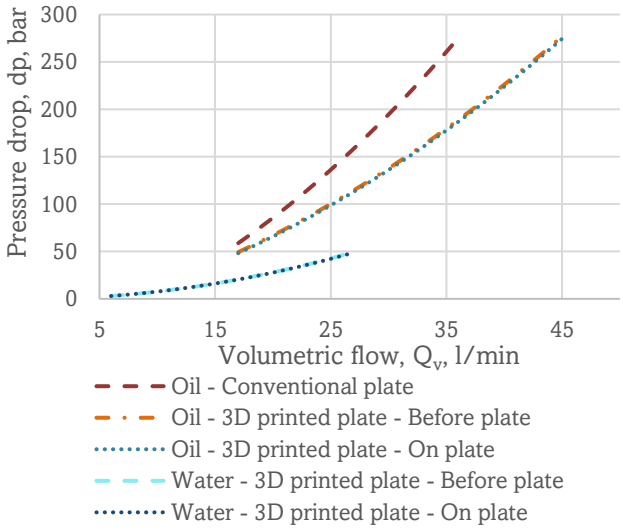


Figure 15: Comparison between water and ISO VG46 mineral hydraulic oil and CM to AM mounting plate measured characteristics at constant 60 % valve opening and variable flow rates.

6 Conclusions

From the results we can conclude that it simply makes sense to use an AM mounting plate over a CM plate. As already mentioned, there are many advantages, from easier mounting to better performance and more accurate measurement. The AM plate is only partly held back by the cost, but as this work clearly shows the level of measurement accuracy achieved with the CM plate, the use of the AM plate with integrated sensors makes more sense, especially at higher flow rates.

Future work

The next steps in this study would be to optimise the shape of the internal geometries to further improve the performance of the AM valves. We believe that the development of self-supporting structures and their flow characteristics would lead to a next generation of high-performance hydraulic components developed using AM. Another step forward, could also be, to make AM mounting plate even more user-friendly and more optimised.

Acknowledgements

We would like to thank our colleagues at Solidworld d.o.o. 3D advanced solutions for their co-operation and for providing AM services for this work. Their work was an essential part of this research, and we hope for further co-operation.

Nomenclature

Variables used in equations are presented in the table below.

Designation	Denotation	Unit
A_{orif}	orifice cross-section area	m^2
C_d	orifice contraction coefficient	/
g	gravity	m/s^2

h	hydraulic height	m
p	pressure	Pa
Δp	pressure difference	Pa
q_v	volumetric flow	l/min
v	velocity	m/s
ρ	fluid density	kg/m^3

References

[1] O Diegel, A Nordin, D Motte. A Practical Guide to Design for Additive Manufacturing. Springer Singapore, Singapore, 2019. ISBN 978-981-13-8281-9.

[2] U Hofmann, M Fankhauser, S Willen, D Inniger, C Klahn, K Löffel, M Meboldt. Design of an additively manufactured hydraulic directional spool valve: an industrial case study. *Virtual and Physical Prototyping*, 18(1):1-18, 2023.

[3] M Arena, P Ambrogiani, V Raiola, F Bocchetto, T Tirelli, M Castaldo. Design and Qualification of an Additively Manufactured Manifold for Aircraft Landing Gears Applications. *Aerospace*, 10(69):1-18, 2023.

[4] G Xie, Y Dong, J Zhou, Z Sheng. Topology optimization design of hydraulic valve blocks for additive manufacturing. *Proc. IMechE Part C: Journal of Mechanical Engineering Science*, 234(19):1899-1912, 2020.

[5] Y Wang, T Peng, Y Zhu, Y Yang, R Tang. A comparative life cycle assessment of a selective-laser-melting-produced hydraulic valve body using design for Property. Procedia CIRP, 90(2020):220-225, 2020.

[6] G Matthiesen, D Merget, M Rückert, K Schmitz, J H Schleifenbaum. Additive manufacturing processes in fluid power – properties and opportunities demonstrated at a flow-optimized fitting. Proceedings of 2018 International Conference on Hydraulics and Pneumatics – HERVEX, ISSN 1454 – 8003, 2018.

[7] F Majdič, N Novak, J Pustavrh, B Laznik, A Trajkovski. Tribological properties of hydraulic cylinder piston sealings in water and oil hydraulics. *14th International Fluid Power Conference, (14 IFK)*, 115-130. 10.13052/rp-9788770042222C09, 2024.

[8] J Bartolj, A Čelik, F Majdič. Development of metallic 3d-printed water hydraulic proportional directional control valve. *International conference fluid power 2021 conference proceedings*, 10.18690/978-961-286-513-9.4, 2021.

[9] J Bartolj, A Trajkovski, F Majdič, A Čelik. Designing, manufacturing and testing of proportional 3D-printed hydraulic directional control valve with topology optimization. *Ventil*, 28(3):190-196,2022.

A Foundation for the Development of Deaeration-Efficient Hydraulic Tanks

Charlotte Becker*, Marius Hofmeister , Yamen Alhaddad and Katharina Schmitz

RWTH Aachen University, Institute for Fluid Power Drives and Systems (ifas), Campus-Boulevard 30, 52074 Aachen, Germany
*E-mail: charlotte.becker@ifas.rwth-aachen.de

Abstract

As the sole component capable of releasing entrained air, a well-designed hydraulic reservoir is crucial for the health of any hydraulic system. This is especially important in mobile machinery, where even more air is entrained via the free surface. Entrained air in hydraulic fluids poses an unavoidable obstacle within mobile machinery, forming a complex multiphase liquid-gas system. This entrained air influences the behaviour of the hydraulic fluid, causing fluctuations in various parameters such as density, viscosity and bulk modulus. These fluctuations often result in a loss of efficiency, as well as reducing the dynamic performance of the system. As electrification of mobile machinery increases, the installation space for hydraulic components shrinks, demanding spatially efficient redesign. The hydraulic tank has long since been designed with a spectrum of rules of thumb - overestimating the required space - to complex CFD simulations. As multiphase CFD already requires ample computing power, additionally considering the relative motion a hydraulic tank experiences, is currently outside the scope of feasibility, coming either at the cost of three dimensionality, or exceeding reasonable computing times - with no guarantees for accuracy. A test-rig, however, would allow for the investigation into the effect of relative motion on deaeration. This paper provides a foundation for a test-rig to re-examine the design of deaeration efficient mobile hydraulic tanks. For this purpose, the main aeration and deaeration mechanisms occurring within the tank have been outlined, and current design elements used to promote deaeration, gathered.

Keywords: Deaeration, Entrained air, Hydraulic reservoir, Hydraulic tank, Tank design, Test-rig development

1 Introduction

Whilst air can enter a hydraulic system in many places, deaeration only takes place within the hydraulic reservoir. Hydraulic systems are constantly pulling in air via cylinders or across seals which must then be released within the tank to prevent damages to the hydraulic system. Within stagnant oil deaeration occurs naturally but slowly, in flowing systems however, turbulence can prevent the rise of the bubbles, creating a race against the clock for the bubbles to reach the free surface lest they are once again sucked into the hydraulic circuit. Even though the size of tanks in mobile machinery is vastly smaller than those employed within stationary hydraulics, this time-pressure is still being exacerbated by shrinking installing space for tanks, especially in the face of steadily increasing electrification of mobile machinery. In addition to the influx of air via the remaining hydraulic system, tanks in mobile machinery also face additional air entrainment via the free surface due to relative movement. These many challenges - especially those of relative movement - have yet to be addressed holistically.

The current design methodologies mostly sit at two extremes of complexity - ranging from rules of thumb to CFD simulation - with the middle ground mostly empty. This middle remains insufficient, even as the number of publications on CFD tank simulations steadily rise, owing to their focus narrowing to machine-specific design. The lack of holistic design guidelines is reflected in the methodologies currently used in the industry. Often times, the selection of a tank is limited to simply finding an appropriately sized tank using a simple flow-demand calculation combined with a large safety factor. Following this, the tank may then also undergo trial-and-error experiments to place inserts such as baffles and sieves. This process may be lengthy and thus costly, particularly if no prior iterations of the mobile machinery or personnel experienced with tank designs are available. The other end of the spectrum, CFD simulations, require extensive computing power and experience, and are thus often outsourced. Additionally, CFD is currently not capable of depicting the system fully. As multiphase CFD already requires ample computing power, considering the relative motion a hydraulic tank experiences in

addition to the miscibility between the air and hydraulic oil, remains currently outside the scope of feasibility. Including relative motion into the simulation would come at either the cost of three-dimensionality or exceed reasonable computing times - with no guarantees for accuracy.

In light of these limitations, further experimental investigation is required to develop design guidelines for hydraulic tanks in mobile machinery. First, the mechanisms of changes in air entrainment that can occur within the tank is outlined in Section 2. Following this, the current reservoir designs will be discussed in Section 3 with special attention on the tank geometry and inserts commonly used to support deaeration and prevent additional entrainment. Finally, the concept for a test rig will be presented; where the challenges associated with the measurement of the change in air content and possible solutions will also be discussed.

2 Air Entrainment

The relevant air content within the design of hydraulic tanks, is the undissolved, dispersed air, also called the "free air". Dissolved air can never be removed entirely, and if the air remains in its dissolved state, only negligible damage will be caused by its presence as it does not change the physical properties of the oil [1]. The dissolved air may be released due to pressure drops as the saturation limit decreases, this can only be influenced via the tank if it were pressurised or there was no air to dissolve in the tank. So-called closed tanks are in use to prevent pump cavitation and increase efficiency, however since the tank continuously sits at this elevated pressure level, the oil would simply saturate at this new limit. Since the tank must remain under the least pressure within the system, to ensure its functionality, this dissolved air would not be released. Free air however can cause significant problems within hydraulic systems and must therefore be minimised. Accordingly, any further mentions of air entrained by the hydraulic oil will refer to the undissolved, free air. In order to reduce the amount of free air within a system, the entrainment of air must be minimised and the oil must be allowed to de-aerate. The problems caused by entrained air arise mostly from the changes in the behaviour of the hydraulic medium, as the air content causes fluctuations in parameters such as density, viscosity and bulk modulus. These fluctuations often result in a loss of efficiency, as well as reducing the dynamic performance of the system. The presence of entrained air can lead to increased wear and damage to the components and increased ageing of the oil, through processes such as cavitation and oxidation. Aside from the ageing of the oil, the damages occur when the entrained air enters the remaining hydraulic circuit, consequently it must be removed prior to leaving the tank.

In most hydraulic tanks, the oil is exposed to a significant air layer. The interfacial boundary between the gas and liquid phase is referred to as the free surface. Due to their lower density, air bubbles are buoyant and naturally tend towards the free surface, where they can escape. The precise trajectory of the bubbles is governed by the shape and size of the bubble, as well as the flow conditions within the tank, with bubbles rising faster in lower viscosity oils. Multiple functions approximate the bubble rise velocity, depending on the bubble size, and fluid conditions such as temperature and pressure, these have been expansively reviewed in [2]. Currently available additives, including defoaming agents cannot improve the air-release of hydraulic oils [3]. Local turbulences often cause the bubbles to be swept along and downwards, resulting in continued entrainment within the hydraulic system. Air entering the system in the form of bubbles via the free surface must only traverse a short distance to be released. This means that the penetration depth of the bubbles and the proximity to the tank outlets, influences deaeration.

As previously mentioned, mobile hydraulic systems are not only more vulnerable to increased air contents due to the low residential times, but also due to the relative motion of the tank causing aeration within the tank. A positive aeration rate is especially critical, as the tank itself is the only place within the hydraulic system where air is able to leave the oil. Thus, mobile hydraulic tanks must not only be optimised to support deaeration but also to prevent additional aeration. Six aeration mechanisms will be described in the following section. For some of these mechanisms countermeasures are widely available, their efficient application is however still disputed. Specific countermeasures will be discussed in the Section 3 . The aeration rate AR in this work is defined as the difference between the void fraction β_{out} of the flow leaving the tank and the void fraction β_{in} of the flow returning to the tank (see Equation 1).

$$AR = \beta_{out} - \beta_{in} \tag{1}$$

The six aeration mechanisms or events shown in 1 will be outlined. In the tank, free air enters the oil via local turbulences, where the surface tension of the oil is smaller than the turbulent shear stress [4] or via the trapping of large bubbles. The former case may take place due to the flow of a jet (Nr. 3 in Figure 1) or splashes of liquid impacting the free surface whilst the latter occurs when pockets of air are enveloped by the folding of the free surface. This can occur due to waves breaking when they reach critical heights or due to their impact with the tank walls (Nr. 2 and 1 in Figure 1). The latter phenomenon is called wall- or roof-slamming [5] and like most of the other aeration events in this section, it requires "sloshing".

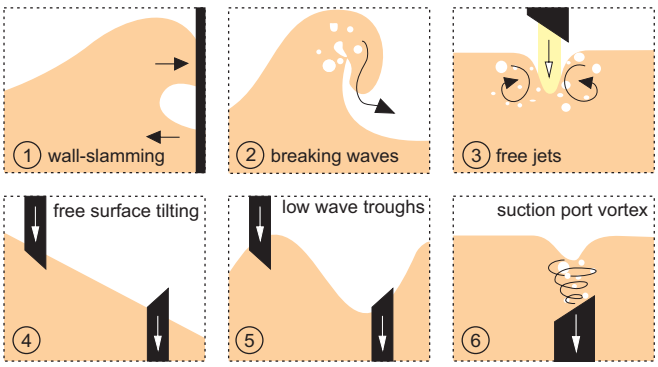


Figure 1: Overview of selected air entrainment mechanisms within the tank

2.1 Sloshing

Sloshing is broadly defined as any movement of a liquid within a container it does not fully fill [6]. Aeration mechanisms influenced by sloshing are susceptible to resonance phenomena. Sloshing at resonance frequencies results in high strain on the containers due to concentrated energy transfers when the liquid impacts the walls [7]. Resonance also introduces nonlinear behaviours causing more air entrainment as the amplitude rises exponentially and should thus be avoided. Estimations of the eigenfunction of a sloshing liquid can be made by simplifying the shape of the container without significant accuracy losses [6], with analytical solutions existing for rectangular, conical, and cylindrical tanks. The estimation of the resonant frequencies may be particularly beneficial for the design of tanks in mobile machinery with oscillation working modes. The resonant frequency of a tank can be increased using longitudinal partitions or ring baffles [6] which provide a dampening effect on sloshing [5]. The severity of the air entrainment outside resonance is dependent on the frequency and amplitude of the excitation. Even without the trapping of large pockets, the shaking of the tank can disrupt the surface tension of the free surface, trapping small bubbles within the liquid [8]. In extreme cases the liquid flow disperses and the gas is suspended as small bubbles within the liquid, forming a foam. Although the effects of the above mentioned aeration events might be contained to the tank itself, the following mechanisms will result in free air entering the hydraulic system. This is due to their immediate involvement with the suction outlets of the tank.

2.2 Suction port

If the outlet is exposed to air or insufficiently covered, pumps may directly suction air into the system. The simplest way this may happen, is due to the free surface tilting at low oil levels. Generally this is accounted for in the design of a new tank, should however be re-examined when additional outputs, especially cylinders are added to machines without a full tank-redesign. The free surface may tilt due to the entire vehicle tilting - the degree of which is generally limited within mobile machinery - or due to inertial forces acting upon the liquid [9]. In the latter case, the acceleration of the vehicle a in relation to the acceleration due to gravity g determines the tilt angle Φ (Equation 2).

$$\Phi = \arctan(\frac{a}{g}) \tag{2}$$

A less predictable event may however arise due to the previously mentioned resonance phenomena. If the trough of a wave is below the suction outlet, air will enter the system. Whilst the level within the tank may not drop below the opening of the tank outlet, air may still be entrained via a fluid vortex.

The final aeration event is the formation of a vortex, where the oil level does not drop below the opening of the tank outlet but where the free surface is locally deformed, splitting the oil layer. The formation of a vortex on a free surface is shown in Figure 2 where the vortex types (VT) are classified using the "Hecker vortex classification scale" and are qualitatively

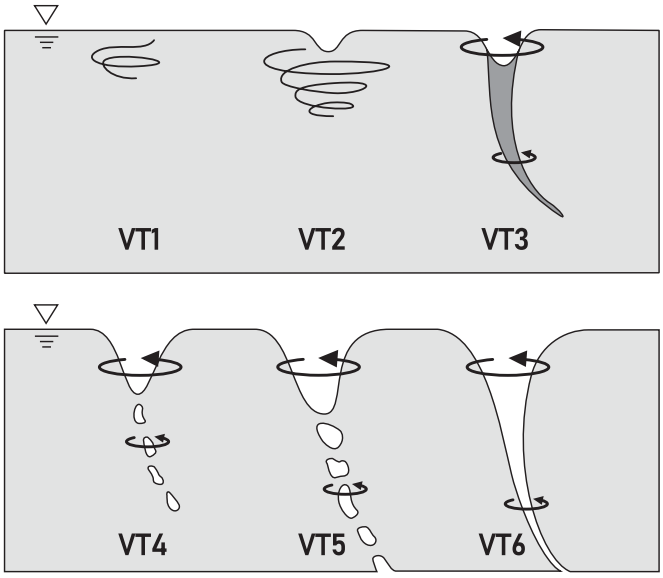


Figure 2: Hecker Vortex Classification Scale [10]

categorised from VT1 to 6. A free surface vortex generally develops from a "coherent surface swirl" (VT1) which proceeds to dimple the surface (VT2) then connects the axis of rotation to the suction intake (VT3) . At first, only debris is entrained (VT4), as the vortex develops air bubbles (VT5) follow and finally an air core is formed (VT6). Air entrainment thus only occurs in VT5 and VT6, with significantly more air entering the system due to VT6 [11]. Once the tip of the air core reaches the (pump) inlet depth, air enters the inlet [12]. The vortex may not stabilise fully, instead cyclically forming and collapsing due to the pump flow dropping as air is sucked in, and reforming once the pump flow increases again [11]. To prevent the formation of an air core - unstable or not- the critical Submergence is calculated in water hydraulics. This is the distance between the inlet and the free surface at which the vortex makes contact. Vortices are not exclusive to tank outlets in the floor of the tank and are highly dependent on the geometry of the outlet and the flow conditions, as well as the viscosity of the liquid.

In water hydraulics close attention has been paid to the critical submergence, with several models and equations existing to control vortex formation in reservoir designs. Generally the critical submergence is considered to be a function of the Froude Number Fr; which is the ratio between the inertia of the flow to gravity; and the geometric configuration of the tank. Multiple analytical models, mainly based on work by Rankine, and the Navier-Stokes equation have been proposed to model the vortex velocity [13] or gas-core length [14] and estimate the critical submergence . These models have to be separated into the positioning of the suction ports as only a vortex directly above the port can have an axially symmetrical shape. When considering multiple different liquids however, the Weber We and Reynolds Re Numbers also have been used to estimate the critical submergence [15], incorporating effects from surface tension and viscosity. Unlike the models available for inviscid liquids however they require knowledge of

difficult to measure parameters, such as the circulation number, or use empirical factors. The simpler equations used in water-reservoir design cannot be transferred to oil hydraulics, in part due to the vastly different geometries, but also due to the models neglecting the effects of viscosity [10]. The energy decay along the vortex increases with viscosity, making oil-hydraulic systems less susceptible to the formation of stable air-cores. However the formation of vortices within hydraulic machinery has not been disproven but merely neglected. In spite of the lower risk, this aeration event may still take place within mobile hydraulic tanks, especially at low tank levels.

3 Hydraulic reservoir design

The main task of hydraulic reservoirs is to store hydraulic fluid and to compensate for uneven oil requirements (e.g. due to the movement of differential cylinders). Besides this primary function, hydraulic reservoirs cool the hydraulic fluid and remove contaminants like particles, water, and air. In this regard, the hydraulic reservoir is the only place in a hydraulic system where air can be removed in significant quantities. Still, it can also be entrained due to effects like sloshing.

The basic mechanisms highlighted in Section 2, which lead to the entrainment or separation of air, can be influenced by different design elements in a targeted manner. A distinction can be made between passive and active elements, whereby active elements, like cyclones, are currently not planned to be implemented in the test bench and are therefore not considered in this section.

Passive elements include the general reservoir geometry, the use of different inserts and the design of the reservoir inlet and outlet. The test rig to be designed should be able to map the influence of the above-mentioned elements and allow a combination of different configurations. For this reason, the general function of the named passive elements is described shortly in the following and the corresponding requirements for the test bench are derived.

3.1 Hydraulic reservoir geometry

Sufficient residential time of the hydraulic fluid in the reservoir is crucial for effective air separation. The longer the fluid stays in the reservoir, the more time bubbles are given to rise to the surface and be separated from the fluid. The mean residential time $t_{residential}$ is calculated by dividing the reservoir volume $V_{reservoir}$ by the maximum volume flow Q_{pump}.

$$t_{residential} = \frac{V_{reservoir}}{Q_{pump}} \tag{3}$$

Depending on the application the suggested residential times varies as shown in Table 1.

In addition to the volume, the reservoir height in relation to the base area is particularly relevant for the separation of air. If the fill level is low, air bubbles can rise to the surface in a shorter time and be separated more quickly.

Furthermore, the basic shape has a crucial influence on the field of flow. In trapezoidal containers per example, a significantly decreased slosh response is observed when compared to a cylindrical or rectangular tanks with non-inclined walls [16]. The flow dynamics within the tank also impact air-release directly, as shown in 3 the rectangular shape of the left reservoir leads to dead zones (red) in each corner. In these areas, air bubbles have a lot of time to rise to the surface, but fluid flowing from the return line can only reach these areas with difficulty. For this reason only small amounts of air can be separated effectively in these areas.

Table 1: Suggested residential times for different hydraulic applications

Application	Residential time $t_{residential}$
Stationary hydraulic systems	3 ... 5 min
Stationary hydraulic systems (HFC and HFD)	8 min
Mobile hydraulic systems open loop systems	2.5 min
Mobile hydraulic systems closed loop systems	1.5 ... 2 min
Aircraft applications	0.5 ... 1 min

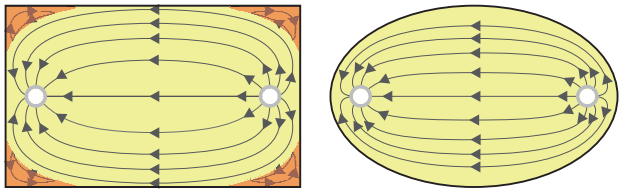


Figure 3: Influence of basic reservoir geometry on the flow field

To enable the test bench to map the described effects on air release, it should be possible to adjust both the reservoir volume and the filling level. This way also varying residential times for different hydraulic applications can be considered. The entire reservoir should be interchangeable to investigate the influence of the different base geometries.

3.2 Inlet and outlet design

In addition to the general reservoir geometry, the design of inlet and outlet geometry has a significant influence on the entrainment and the separation of air. To avoid a vortex, as described in Section 2, it should be ensured that the suction line is always sufficiently immersed during operation. Otherwise, air can be dragged along by the flow and enter the system through the suction line (Figure 4 l.).

The stream distance between the return line and suction line should be as large as possible to give air bubbles more time to rise to the surface, and ideally creating an even velocity field [17]. Belov et al. recommend that the flow length should be balanced with the costs, as the unevenness of the flow field decreases hyperbolically with an increase in length.

The return line should end below the fluid level so that no air can be entrained via the free jet. To increase the cross sec-

tion of inlet and outlets, the pipe opening is often cut at a 45° angle [18]. The geometry of the tank inlet can also be adjusted from a simple pipe to instead include perforations on the side. [19] found that in a CFD simulation of their specific tank geometry, a perforated inlet resulted in significantly better air release than a bottom opened inlet. This improvement was attributed to the perforations resulting in better flow distribution and a quicker momentum exchange between the fluid entering and the fluid already within the tank. This effect can also be achieved using diffuser-attachments for the pipe, which often come integrated into the return filter.

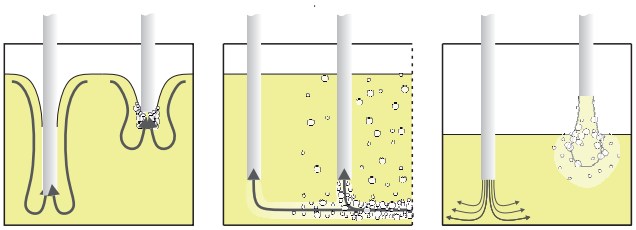


Figure 4: inlet and outlet design

To investigate the effects of the design of the tank ports on deaeration, the height and position of the suction and return pipes in the test-reservoir should be variable.

3.3 Inserts

To improve air separation and to prevent the entrainment of air, different reservoir inserts can be used. The use of baffles, partitions and swash plates in sloshing has been examined closely for liquid storage containers, especially in marine applications, where large external excitations occur (for further reading see [7]. Although the liquid within the tank does not experience a global flow rate, this research is still a vital boost to understand insert interaction within mobile machinery, as many publications consider the effects of both air entrainment and viscosity. Baffles and partitions are used to divide the reservoir into a sections and a return flow section or to guide the flow of hydraulic oil inside the reservoir. This way short-circuiting between suction and return line is avoided. Simultaneously, the fluid can be directed to the surface by bottom mounted baffles, thus shortening the path of air bubbles to the surface. Vertical baffles also reduce the sloshing amplitude by blocking the fluid motion and causing local turbulences at the sharp baffle edges, resulting in a dissipation of energy. Bottom mounted baffles are less effective in reducing impact pressures and sloshing amplitudes, than those touching the free surface, as the highest horizontal velocity manifests at the free surface in the middle of the tank and reduces towards the bottom of the tank [7]. Whilst vertical baffles can combat heavy sloshing modes, care must be taken when sizing them, as they may severely impact the flow, causing cascading flows down the sides of a baffle [20], which also causes air entrainment. The total number of partitions used should also be limited to avoid creating more dead-zones [21]. Investigations into elastic baffles has also found that when the baffle is not immersed, those with smaller stiffnesses result in higher damping, whilst for the immersed baffles the inverse was observed [22].

Another way to improve air separation is to install screens on which air bubbles can accumulate, travel along the screen and finally rise to the surface. The performance of screens depends largely on the mesh size and the angle. The grain of the screen is recommended to not excessively hinder the flow and may be coated to promote bubble collection [17]. For this reason, the test rig should be facilitate changing between screens with different mesh sizes and vary their installation angle.

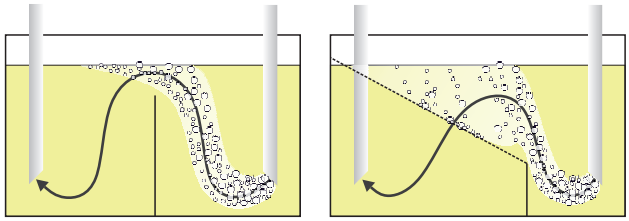


Figure 5: influence of baffles and screens on air separation

3.4 Requirements for the test bench design

The requirements for the test bench derived from the design elements described above are summarized in Table 2 below.

Table 2: Requirements for test bench design derived from tank geometry

Design Element	Test Bench Variables
Base container	Exchangeable reservoir
	Adjustable fluid level
Inlet and outlet design	Adjustable length of inlet/outlet line
	Variable positioning of inlet and outlet
	Variable opening geometry
Inserts	
Baffles and partitions	Changeable position and inclination
Screens	Changeable position and inclination
	Variable mesh size
Slit sheets and perforated baffles	Variable opening sizes

4 Test-rig Design

To account for the incompatibility of some design choices, the test rig concept is modular, enabling the swapping of the tank. In addition to the possibility of easily exchanging the test-tank,

the design will also minimise the effort required to change the hydraulic fluid, by separating the hydraulic circuit carrying out the movement and generating the hydraulic work, from the fluid flowing through the test-tank. This will not only allow the investigation into the influence of different fluid characteristics, but also prevent damage to the components caused by the excess air within the test fluid . The free air content of the oil will be measured before and after the tank, allowing the calculation of the deaeration rate. These measurement methods will also be briefly discussed and evaluated below. The air content measurement will additionally be used in the control of the air content in the inflowing fluid. To control the air content, additional air will be injected. The injection of additional air is to observe the effect of the tank geometry and movement on the deaeration rate. In order to reduce the air within the tank, a second larger tank will be used as a settling tank, from which fresh oil can be sourced. A bypass will then be used if the air content is to be increased from the current air content within the test-tank. The test-circuit of the test-rig will furthermore be kept at low pressures below 10 bars to allow for the active injection of compressed air. The test rig will also investigate typical operating conditions and volume flows but will be restricted to 100 l tanks.

4.1 Movement

The intended movement capabilities of the test-rig have been adapted to simulate typical working motions of mobile hydraulic machinery. To assess the most critical external excitation of the tanks, the movements will initially be carried out in isolation. Another motive for the initial focus on isolated movement is that the overlaying of excitation directions has already been shown to increase sloshing dramatically. The most vital movement types were considered to be:

- Translational in Z (up/down)
- Rotational around Z (yaw)
- Rotational around X or Y (roll/pitch)

In order to reduce the degrees of freedom required by the test rig, motions in the X and Y direction will be achieved by switching the mounting direction of the tank. The yawing movement is especially vital due its high use in construction machinery, per example in dig-and-dump-cycles. Acceleration in the direction of travel (X) will be depicted by making use of the gravitational force and tilting the tank. Movement in the direction of travel will not be depicted within the test-rig as most hydraulic machinery can reach limited driving velocities which do not justify the additional space, costs and risks associated with implementing horizontal movement. The movement types specified above, will be examined in isolated impulses as well as oscillations to examine the effect sudden impacts and sloshing.

4.2 Measurement Methods

To accurately determine the change in aeration of the tank, the difference in the amount of free air entrained in the in- and outflow of the tank must be measured. The "amount" can be either specified by the number of bubbles of significant diameters or by measuring the void fraction. The advantage of the former is that the bubble size discerns the type of damage that may be caused and if a broad enough range of bubble diameters are captured, it may then be used to estimate to void fraction.

Whilst there has been significant research into the measurement of entrained air not all is transferable to the desired operating conditions. Much of the available technology on measuring multi-phasic flows originates within the medical field, where dynamic imaging is vital. Unfortunately the volume flows within the body are much lower than those within hydraulic systems and available sensors are for high volume flows are much rarer, nevertheless some of the measurement methodologies can be adapted for use within oil hydraulics.

Another significant restriction is high void fractions which may occur during measurement as many methods cannot accurately capture larger bubbles or slug flow. In some cases the exposure of probes to large air pockets result in sensor failures, whilst in others the signals become nonsensical. In the latter case, the measurement method could be combined with a secondary measuring system to allow for its use. In addition to the demands arising from the high flow speed and void fraction requirements, the sensors must also be suitable for use in a closed-loop control to regulate the free air within flow entering the tank. This means that the measurement must be inline, at an adequate sample rate. Other requirements for the sensors originate in the use of oil-hydraulics; the sensor-materials must be compatible with various hydraulic oils and must provide robust measurements, despite small temperature fluctuations or vibrations carried through the machine. In addition to the requirements listed above, economic restrictions, as well as safety concerns - primarily in regards to radiation - were utilized to screen current measurement systems. A selection of the evaluated measurement methodologies is listed in Table 3. The measurement methods for gas-liquid two-phase flow either emit energy into the flow or observe differences in flow behaviour, mostly in the pressure differentials, when it passes resistance components. The first group can be separated into categories by the type of energy used and will be discussed below. The sensors generally emit either sound or electromagnetic waves which are then altered by passing through the flow. The type of disturbance depends on the material characteristics and the interfaces between the components. In liquid-gas flows the waves are usually impacted the most by traversing the interfacial barrier and depending on the type of wave are either absorbed or reflected to some degree [30]. The sensor then either receives the reflected or remaining wave and compares it to a reference. The degree of change, depends on the characteristics of the gaseous phase. Scattering vs reflecting per example is determined by the size of the object relative to the wavelength, with significantly smaller isolated objects acting as point scatters [31]. Sensors working with reflected or scattered waves usually have a reflective material on the other side of the pipe or container to reflect any remaining waves back as an echo. Methods which utilise the remaining

Table 3: Overview of measurement methodologies to determine air content

Method	Pro	Contra
Compressibility [23]	Accurate	Only offline
High-speed photography [23]	Non-Invasive, High Resolution	Struggles with opaque liquids, limited frame-rate
Optical Probe [24], [25]	Versatile	Expensive tracers, Ineffective for larger bubbles
Electrocapacitance Tomography [26]	Inexpensive, high sample rate	Low spatial resolution
Electrical Impedance [23]	Inexpensive, high sample rate	Incompatible with oil
Thermal Anemometry [27]	Simple, Inexpensive	Inaccurate, invasive
Pulse Echo Technique [28]	Can measure bubble size and number	Limited void fraction measurable
Pulse Wave Ultrasonic Doppler [29]	Can measure in stratified flow (high void fractions)	May not achieve sufficient depth penetration
Continuous Wave Ultrasonic Doppler [29]	High resolution and sample rate	Precise placement of transducers required

waves are called shadowgraph methods, the most popular being shadowgraphy, an optical method. These methods often use a pulsing, rather than a continuous emission, to minimise interference between outgoing and incoming waves. The received waves are analysed on their amplitude, frequency and phase, depending on the method. From this information, the void fraction or bubble size is then reconstructed using models such as the ECAH model, which is used for ultrasonic methods [31]. In addition to optical waves, the other commonly used electromagnetic methods use electricity, heat or forms of radiation, namely X- and gamma-rays. The techniques using radiation generally require a strong insulation layer to absorb remaining rays safely, this results in larger setups and higher costs.

Since the change in the waves are observed and referenced, the depth penetration of the wave is of utmost import. The wave must be noticeably changed by the air bubbles, but not fully absorbed or reflected in order to accurately measure the air content. This is especially relevant in high volume flow rates as the pipe sizes increases with the flow. Increasing the depth penetration by increasing the energy of the wave is limited as adverse effects, such as heat damage in the components or fluid, may occur [32]. One example of this is in sound waves which, when the amplitude is too high, can result in cavitation due to pressure drops in the rarefaction of the wave. Increased power can also lead to changes in the fluid movement. In larger pipes - especially at higher void fractions- more bubbles may also "hide behind" another if the wave used is completely dispersed by a single phase transition. This is usually counteracted by the use of several probes creating a net in a 2 dimensional plane. The most literal example of this are wire-mesh sensors, which use a mesh of wires and electricity to measure the passing bubbles. Probes generally exist in sender-receiver pairs or single units, with many sensors also using bundles of probes.

Ultrasonic methods can be used to measure both void fraction and particle size distribution, due to their non-invasive nature and low costs their application is widespread in the medical and industrial field [33] . Within the ultrasonic methods, the Doppler Techniques and the Pulse Echo Technique can be used to determine the void fraction within a liquid, a comprehensive review can be found in [29]. The Continuous Wave Ultrasonic Doppler (CWUD) method appears to be the method with the fewest drawbacks amongst all evaluated methods. Unfortunately the market availability does not reflect the breadth of published articles on void fraction measurements. Instead it appears that many facilities have developed their own sensors specific to their application, providing descriptions of the used components and signal processing methods, examples of this can be seen in [34], [35]. In light of this, the development of a project specific air content sensor is a viable option if no commercial sensors can be acquired.

5 Conclusion

Due to movement of mobile machinery, additional air can be entrained in the hydraulic oil within the tank. This air may be captured in the form of small bubbles via local turbulences or jet streams, in the form of large bubbles captured during violent sloshing events where waves break, or directly suctioned into the pump inlets. Preventing these events, as well as encouraging deaeration, is crucial to maintaining the health of the hydraulic system. Several approaches to minimise the amount of entrained air by optimising the tank geometry and using inserts were discussed. The shortcomings of the current design guidelines were called to attention, especially the lack of research into the relationship between relative motion and deaeration. This work outlined the current challenges in the design of deaeration efficient mobile hydraulic tanks to highlight the need for more experimental investigation. Following this, a concept for a test-rig was introduced. The required variables and possible air-content measurements were discussed. With the research gathered from this test-rig, the design of hydraulic tanks for mobile machinery is to be illuminated, creating design guidelines.

6 Acknowledgments

The IGF research project 01/F23342N of the research association Forschungskuratorium Maschinenbau e. V. - FKM, Lyoner Straße 18, 60528 Frankfurt am Main was supported from the budget of the Federal Ministry of Economic Affairs and Climate Action through the DLR within the scope of a program to support industrial community research and development (IGF) based on a decision of the German Bundestag.

References

[1] Dietmar Findeisen and Siegfried Helduser. *Ölhydraulik: Handbuch der hydraulischen Antriebe und Steuerungen*. VDI-Buch. Springer Vieweg, Berlin, 6., neu bearbeitete auflage edition, 2015.

[2] Amol A. Kulkarni and Jyeshtharaj B. Joshi. Bubble formation and bubble rise velocity in gas−liquid systems: A review. *Industrial & Engineering Chemistry Research*, 44(16):5873–5931, 2005.

[3] Theo Mang and Wilfried Dresel. *Lubricants and lubrication*. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, third completely revised and enlarged edition edition, 2017]- // 2017.

[4] Hubert Chanson. Hydraulics of aerated flows: qui pro quo ? *Journal of Hydraulic Research*, 51(3):223–243, 2013.

[5] L. Constantin, J. J. de Courcy, B. Titurus, T.C.S. Rendall, and J. E. Cooper. Sloshing induced damping across froude numbers in a harmonically vertically excited system. *Journal of Sound and Vibration*, 510:116302, 2021.

[6] R. A. Ibrahim. *Liquid sloshing dynamics: Theory and applications*. Cambridge University Press, New York, 2005.

[7] Mi-An Xue, Jinhai Zheng, Pengzhi Lin, and Zhong Xiao. Violent transient sloshing-wave interaction with a baffle in a three-dimensional numerical tank. *Journal of Ocean University of China*, 16(4):661–673, 2017.

[8] Emma Frosina, Adolfo Senatore, Assunta Andreozzi, Gianluca Marinaro, Dario Buono, Gennaro Bianco, Domenico Auriemma, Francesco Fortunato, Fulvio Damiano, and Pino Giliberti. Study of the sloshing in a fuel tank using cfd and efd approaches. In *Proceedings of the ASMEBATH Symposium on Fluid Power and Motion Control - 2017*, New York, N.Y., 2017. The American Society of Mechanical Engineers.

[9] Ronja Hoch. *Zweiphasenströmung in Kühlmitteltanks von Sportwagen*. Dissertation, Universität Rostock, 2021.

[10] Jost Knauss. *Swirling Flow Problems at Intakes*. Routledge, 2017.

[11] Robert Ettema, David Z. Zhu, Susan J. Gaskin, Troy C. Lyons, Robert Boes, Shu Ning Chan, and Joseph H.W. Lee. *Vortex-Flow Intakes*. ETH Zurich, 2023.

[12] Rahul Kumar Mondal and Parmod Kumar. Experimental study of entrained air-core structures induced by a pump intake vortex. *Physics of Fluids*, 34(5), 2022.

[13] Yun-liang Chen, Chao Wu, Mao Ye, and Xiao-ming Ju. Hydraulic characteristics of vertical vortex at hydraulic intakes. *Journal of Hydrodynamics*, 19(2):143–149, 2007.

[14] Kei Ito, Takaaki Sakai, Yuzuru Eguchi, Hideaki Monji, Hiroyuki Ohshima, Akihiro Uchibori, and Yongze Xu. Improvement of gas entrainment prediction method—introduction of surface tension effect—. *Journal of Nuclear Science and Technology*, 47(9):771–778, 2010.

[15] Carlos Julián Gavilán Moreno. Determining critical submergence in tanks by means of reynolds & weber numbers. *World Journal of Engineering and Technology*, 02(03):222–233, 2014.

[16] S. Mitra, P. P. Upadhyay, and K. P. Sinhamahapatra. Slosh dynamics of inviscid fluids in two–dimensional tanks of various geometry using finite element method. *International Journal for Numerical Methods in Fluids*, 56(9):1625–1651, 2007.

[17] N. Belov and N. Sosnovsky. Design of an optimal hydraulic tank configuration. *IOP Conference Series: Materials Science and Engineering*, 779:012041, 2020.

[18] Vito Tič and Darko Lovrec. Air-release and solid particles sedimentation process in a hydraulic reservoir. *Tehnički vjesnik*, 20(3):407–412, 2013.

[19] Gorle Jagan. Coupled fluid-thermal-structural analysis of two-phase flow in a hydraulic tank. In *The Sixteenth Scandinavian International Conference on Fluid Power*, 2019.

[20] Ronja Hoch and Frank-Hendrik Wurm. Numerical and experimental investigation of sloshing under large amplitude roll excitation. *Journal of Hydrodynamics*, 33(4):787–803, 2021.

[21] V. Tic and D. Lovrec. Design of modern hydraulic tank using fluid flow simulation. *International Journal of Simulation Modelling*, 11(2):77–88, 2012.

[22] Yaru Ren, Mi-An Xue, and Pengzhi Lin. Experimental study of sloshing characteristics in a rectangular tank with elastic baffles. *Journal of Fluids and Structures*, 122:103984, 2023.

[23] Katharina Schrank, Hubertus Murrenhoff, and Christian Stammen. Investigation of different methods to measure the entrained air content in hydraulic oils. *ASME/BATH 2014 Symposium on Fluid Power and Motion Control*, 2014.

[24] G. Rojas and M. R. Loewen. Fiber-optic probe measurements of void fraction and bubble size distributions beneath breaking waves. *Experiments in Fluids*, 43(6):895–906, 2007.

[25] Christian Poelma. Measurement in opaque flows: a review of measurement techniques for dispersed multiphase flows. *Acta mechanica*, 231(6):2089–2111, 2020.

[26] Jing Lei and Shi Liu. An image reconstruction algorithm based on the regularized minimax estimation for electrical capacitance tomography. *Journal of Mathematical Imaging and Vision*, 39(3):269–291, 2011.

[27] Volfango Bertola. Two-phase flow measurement techniques. In Manuel Garcia Velarde, Mahir Sayir, Wilhelm Schneider, Bernhard Schrefler, Giovanni Bianchi, Carlo Tasso, and Volfango Bertola, editors, *Modelling and Experimentation in Two-Phase Flow*, volume 450 of *CISM International Centre for Mechanical Sciences*, pages 281–323. Springer Vienna, Vienna, 2003.

[28] Antonin Povolny, Hiroshige Kikura, and Tomonori Ihara. Ultrasound pulse-echo coupled with a tracking technique for simultaneous measurement of multiple bubbles. *Sensors (Basel, Switzerland)*, 18(5), 2018.

[29] Chao Tan, Yuichi Murai, Weiling Liu, Yuji Tasaka, Feng Dong, and Yasushi Takeda. Ultrasonic doppler technique for application to multiphase flows: A review. *International Journal of Multiphase Flow*, 144:103811, 2021.

[30] Chunwei Gu, Qian Li, Rui Ma, Yingsong Lin, Xiangfang Li, Yiming Li, Aixia Zhang, Yingjie Li, and Bangtang Yin. Propagation characteristics of doppler ultrasonic wave in gas–liquid two-phase flow in an offshore deepwater riser. *Natural Gas Industry B*, 8(6):615–621, 2021.

[31] Jinhui Fan and Fei Wang. Review of ultrasonic measurement methods for two-phase flow. *The Review of scientific instruments*, 92(9):091502, 2021.

[32] Christian Poelma. Ultrasound imaging velocimetry: a review. *Experiments in Fluids*, 58(1), 2017.

[33] Fria Hossein, Massimiliano Materazzi, Paola Lettieri, and Panagiota Angeli. Application of acoustic techniques to fluid-particle systems – a review. *Chemical Engineering Research and Design*, 176:180–193, 2021.

[34] Zihui Wei, Xinyue Zhao, Lintao Han, Ning Zhao, and Lide Fang. Ultrasonic probes and echo time algorithms in ultrasonic gas flow measurement systems: Progress and perspectives. *Physics of Fluids*, 37(1), 2025.

[35] Eduardo Nunes dos Santos, Natan S. Reginaldo, Jean N. Longo, Roberto da Fonseca Jr, Marco G. Conte, Rigoberto E. M. Morales, and Marco Jose da Silva. Advancing oil and gas pipeline monitoring with fast phase fraction sensor. *Measurement Science and Technology*, 35(12):125302, 2024.

Design and efficiency modelling of a fixed-ratio hydraulic transformer

Pierre Bernard and Vincent Langlois

Poclain Hydraulics Industrie, Verberie, France
E-mail: pierre.bernard@poclain.com, vincent.langlois@poclain.com

Abstract

A hydraulic transformer, sometimes also described as a pressure booster or pressure amplifier, is a component used to transform a hydraulic power, from a flow Q1 and pressure P1 to another state of flow Q2 and pressure P2. The pressure booster's aim is usually to multiply the pressure level by a given ratio, thus dividing the amount of flow by the same ratio. Common challenges in such components are transformation efficiency, outlet pressure and outlet flow dynamics, and space claim.
The present work investigates the design and modelling of a – compact – fixed ratio pressure amplifier based on a radial pistons cam-lobe architecture. This type of architecture has been used for years in direct drive mobile and industrial applications, and its technology is here adapted to a hydraulic transformer concept.
The component's concept and design are described along with its basic kinematic and dynamic behaviour. Then, models of the hydromechanical and volumetric losses are presented, allowing the complete simulation of the power transformation's efficiency. A confrontation of the efficiency modelling with experimental results constitutes the last part of the paper, showing good agreement of simulated results with measurements on a prototype version of the pressure intensifier. The model developed in this work can be used to vary design parameters, and optimise the design of the pressure amplifier to maximize the efficiency of the power transformation.

Keywords: Hydraulic transformer, pressure booster, radial pistons, cam-lobe, efficiency

1 Introduction

In a context of global drive towards energy economy, there is a need for hydrostatic transmissions and hydraulic systems to evolve towards more efficient architectures and components [1, 2]. The general trend of electrification of the prime mover, replacing diesel engines, puts a focus on uptime of the machines, also pushing towards greater system efficiency. One way for hydrostatic transmissions to get more efficient is to move to higher pressure levels. Indeed, increasing pressure will lower the required displacement for a given torque, which will allow a reduction of flow in the system, thus lowering pressure drops and increasing the transmission's overall efficiency. However, as shown by Achten in [3], the maximum rated pressure in hydraulic systems has not changed for the past 40 years, and despite a few components available for higher pressure levels [4, 5], no ecosystem of components is yet available to move to the next level of pressure.
To make a step in the direction of increased pressure, Poclain proposed a concept of boosted radial pistons cam-lobe motor [6, 7], which incorporates a pressure amplifier. This boosted motor is able to internally increase its working pressure beyond that of the hydraulic system when the need arises, thus allowing a reduction in the installed displacement. The present work investigates the concept of the pressure amplifier used in the proposed boosted motor. Indeed, although several different concepts of pressure amplifiers have been proposed along the years, see for example [8–11], including it inside a hydraulic motor to amplify the pressure without impacting the rest of the system is - to the best of the authors' knowledge - a novelty.
Amplifying the pressure in the hydraulic circuit of a mobile application, or directly inside a hydraulic motor poses some challenges, which existing pressure boosters do not meet. Indeed, in a hydraulic motor directly driving a machine's wheel, pressure and flow ripples can easily be felt by the machine's operator, as they are directly creating a jerky behaviour on the machine's movement, and can also lead to vibrations and noise. The amplification ratio and inlet pressure level are also quite specific, and no commercially available amplifier was found that met the criteria. Finally, the space claim of existing pressure boosters is quite important, and integrating them inside a motor that must fit in a wheel is not easy. This is to try and solve these shortcomings that the radial pistons cam-lobe pressure booster has been designed. This paper will first describe the basic design and kinematics of the concept. The second part of the paper is dedicated to the modelling of losses

and efficiency of the proposed concept, while the final part discusses the results of the efficiency modelling, and compares experimental results with simulations. In a further work, the models presented in this paper will be used to improve and optimise the design of the component, to maximize the efficiency of the power transformation.

2 A radial pistons cam-lobe pressure amplifier

In this first section, the concept and patented [12] design of the radial pistons cam-lobe pressure booster (RPCLB) are presented.

2.1 Concept

The RPCLB uses radial pistons with two separate pushing stages in order to create a motor and a pump in the same rotating group. Figure 1 shows a schematic view of the operating principle of the RPCLB, in which the top stage of the piston (or head) is used as a motor, while the lower stage (or foot) is used as a pump. In fig. 1, the orange colour indicates high system pressure HP_{sys} which is the inlet pressure to the booster, the blue colour is low system pressure LP_{sys} which is also the pressure on the secondary outlet port of the device, and the red colour is the amplified pressure on the RPCLB's outlet HP_{boost}. Due to the difference in pressurised areas between the motoring and the pumping stages of the pistons, the pressure is amplified from HP_{sys} to HP_{boost}.

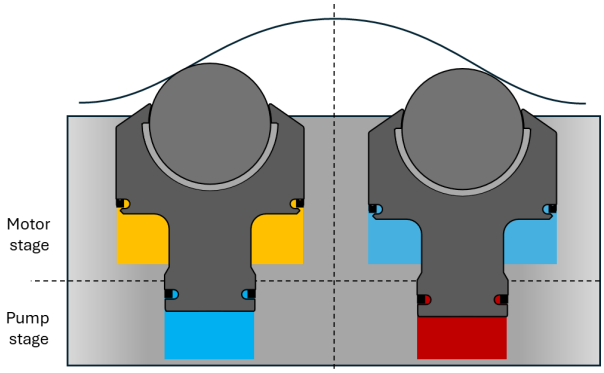


Figure 1: RPCLB *operating principle*

2.2 Design and parameters

The specific design of the RPCLB in this study is composed of 3 pistons and 4 cam-lobes, as illustrated by fig. 2. The ratio of displacements of the two parts (motor and pump) is 3.27, with a motor displacement of 34.9 cc/rev, and a pump displacement of 10.7 cc/rev. Figure 2 introduces the different elements composing the torque module of the RPCLB. The cam (1) is fixed, and as the cylinder-block (2) rotates, the pistons (3), pushed by pressurised fluid, move in and out of their bore in a direction radial to the cylinder-block. As this movement occurs, the cylindrical rollers (4) roll along the cam profile, their rotation movement with respect to the pistons being facilitated by the bushings (5). Sealing rings (6, 7) define two fluid volumes under each piston, both volumes alternately suck in fluid during the outwards stroke of the piston, and discharge fluid during the inwards stroke. Although the rotation of the cylinder-block is described as counter-clockwise on fig. 2, the RPCLB is perfectly capable of rotating in clockwise direction, thus reducing pressure to increase flow output. In the design studied in this

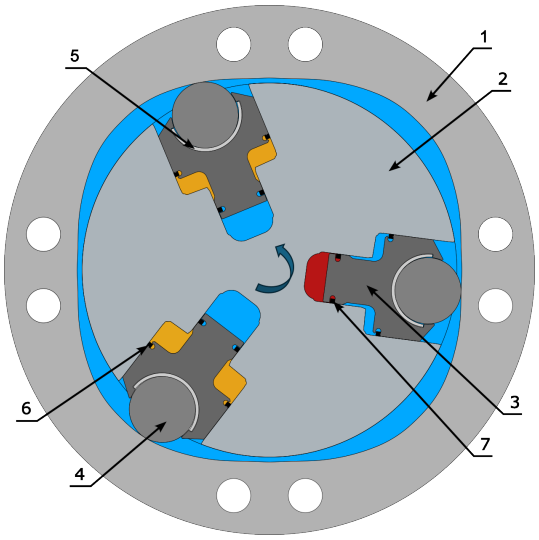


Figure 2: RPCLB *operating principle*

work, the case of the RPCLB is linked to the low pressure port P_3 of the amplifier (see fig. 3), which means that the case is subjected to the low system pressure (symbolised by the blue colour between cylinder-block and cam on fig. 2). Because of this case pressurisation, the net hydraulic force applied on the piston is the force due to the pressures in the two piston chambers, to which is subtracted the force due to the pressure in the case.

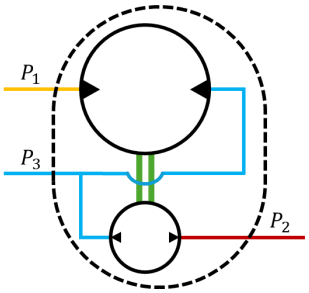


Figure 3: RPCLB *schematic representation*

3 Modelling

In this section, the kinematics of the RPCLB is described, and models of hydrodynamic and volumetric losses are presented in order to obtain a model of the energetic transformation occurring in the component.

3.1 Kinematics

Figure 4 represents one piston with the parameters describing the cam profile and roller coordinates, as well as the different forces at play. As the roller rolls along the cam profile, contacting it in a point described by the (α, R_{CRC}) coordinates, its centre follows a trajectory described by the (θ, ρ) coordinates.

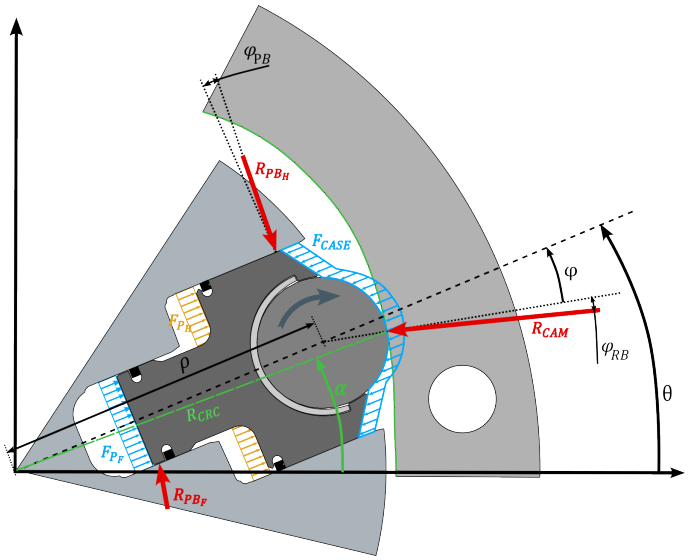


Figure 4: RPCLB *kinematics, and piston force balance*

The direction normal to the cam-roller contact forms an angle φ with the piston axis.
The laws of motion (speed V_P, and acceleration Γ_P) of the piston in its bore can directly be derived from the trajectory of the roller centre.

$$V_P(\theta) = \frac{d\rho}{dt} = \frac{d\rho}{d\theta} \cdot \omega \tag{1}$$

$$\Gamma_P(\theta) = \frac{d^2\rho}{dt^2} = \frac{d^2\rho}{d\theta^2} \cdot \omega^2 \tag{2}$$

The fluid flow displaced by each chamber of a piston (Q_{Pist_M};Q_{Pist_P}) are derived from eq. (1, 2), and the sum of volumes displaced by all pistons then gives the motor and pump stages volume displacements (D_M;D_P).

$$Q_{Pist_M} = V_P \cdot \frac{\pi({D_{P_h}}^2 - {D_{P_f}}^2)}{4}; D_M = \sum_{N_P} Q_{Pist_M} \tag{3}$$

$$Q_{Pist_P} = V_P \cdot \frac{\pi {D_{P_f}}^2}{4}; D_P = \sum_{N_P} Q_{Pist_P} \tag{4}$$

In equations (3, 4), the quantities D_{P_h} and D_{P_f} represent the diameters of the head and tail of the piston. The theoretical pressure amplification ratio $r_{amp_{th}}$ is defined as the ratio of the motor and pump stages displacement, as given by eq (5).

$$r_{amp_{th}} = \frac{D_M}{D_P} \tag{5}$$

3.2 Internal forces

The equipped piston force balance is illustrated on fig. 4, with the different reaction forces R_{PB_H}, R_{PB_F}, and R_{CAM}, resulting from the pressure forces applied to the piston F_{P_H}, F_{P_F}, and F_{CASE}. The force balance equation of the equipped piston is represented by eq. (6), applied at the centre of the roller.

$$\begin{pmatrix} 0 \\ F_{P_H} + F_{P_F} - F_{CASE} \\ 0 \end{pmatrix} = [M] \begin{Bmatrix} R_{CAM} \\ R_{PB_H} \\ R_{PB_F} \end{Bmatrix} \tag{6}$$

The matrix $[M]$ accounts for distances from force application points to roller centre, and for the angles φ_{PB} and φ_{RB} which are related to friction coefficients in a way described in §3.3.2. A more detailed description of the piston force balance, albeit applied to a standard piston design instead of the stepped piston of the RPCLB can be found in [13]. In the RPCLB, a rotary planar distributor is used to distribute the pressurised fluid to the piston chambers at the appropriate time, in the same way as on radial pistons cam lobe motors. The principle and simulation of such a valve is detailed in [14]. The friction losses in such a distributor valve are low, thanks to the total separation of surfaces by the fluid film. However, to balance the forces in the film, an axial counter-balance force is needed. This force F_{dist} directed from the valve to the cylinder-block, is created by the differential sections of valve grooves, and is transferred through the cylinder-block to the tapered roller bearing, as illustrated on fig. 5.

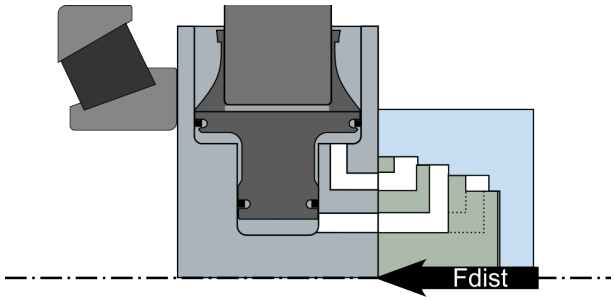


Figure 5: Distribution force impact on the bearing

3.3 Hydromechanical losses model

In order to correctly evaluate the transformation ratio of the RPCLB, both the hydromechanical and volumetric losses must be taken into account. Hydromechanical losses relate to losses having an impact on the amount of torque generated by a motor, or the torque required by a pump to generate a given flow at a certain pressure level. These losses can further be broken down into pressure drops and friction related losses.

3.3.1 Pressure drops

Due to the tortuous geometry of the channelling in the distribution part of the RPCLB, the minor head losses (pressure losses due to perturbations in the channels: bends, convergents, divergence,...) dominate the major head losses (pressure losses that occur along the length of pipes or channels) in the component. These minor head losses can be calculated using eq. (7), in which ρ_{oil} is the oil density, ξ is a coefficient related to the geometry, and V_{fluid} is the oil speed in the reference area corresponding to the pressure drop coefficient ξ.

$$\Delta p = \frac{1}{2} \rho_{oil} \xi V_{fluid}^2 \tag{7}$$

The pressure drop in the different parts of the oil distribution channels is taken into account as a succession of orifices. Figure 6 illustrates the different orifices used to represent the pressure drop along the path of the fluid. Four valve channels are

represented, linked to the RPCLB's inlet port P_1 by a circular channel usually referred to as a groove. The modelling is

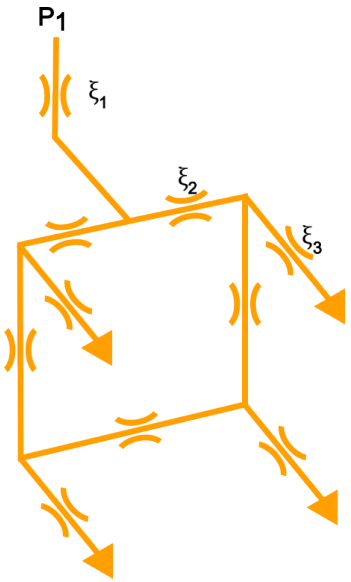


Figure 6: Inlet groove representation for pressure drops calculation

identical for the two outlet ports $P_2$and P_3. On fig. 6, three coefficients are noted, as the pressure drop coefficient for all channels ξ_3 are identical, and the pressure drop coefficient ξ_2 is the same for all orifices along the groove. The oil flow rate entering at port P_1 is equal to the sum of the oil flow rates exiting at the four valve channels (as symbolised by the arrows on fig. 6).From these, and the pressure drop coefficients ξ_i and corresponding flow areas of the groove, the conservation of mass flow allows the calculation of the flow in each of the areas, and thus the global pressure drop between port P_1 and each of the valving ports.

The modelling of pressure losses in the cylinder-block's channels is done in a similar way. However, at the interface of the cylinder-block and valve channels, the planar distributor requires a special modelling. Indeed, the flow passage area at the interface between each valve channel and cylinder-block drilling, defined as the intersection of the cylinder-block and valve orifices, varies with the angular position of the cylinder-block, as illustrated in fig. 7. For this reason, eq. (7) is adapted for this calculation and becomes eq. (8), where $A(\theta)$ is the flow passage area at the valving interface for a given valve and cylinder-block orifice couple.

$$\Delta p(\theta) = \frac{1}{2}\rho_{oil}\xi(\theta)\left(\frac{Q_{Pist}(\theta)}{A(\theta)}\right)^2 \tag{8}$$

It should also be noted, that due to the evolving flow passage area, the pressure drop coefficient $\xi(\theta)$ at this interface is also variable.

3.3.2 Friction losses

There are three main sources of frictional loss in the RPCLB.

Roller/Bushing friction The interface between the roller and

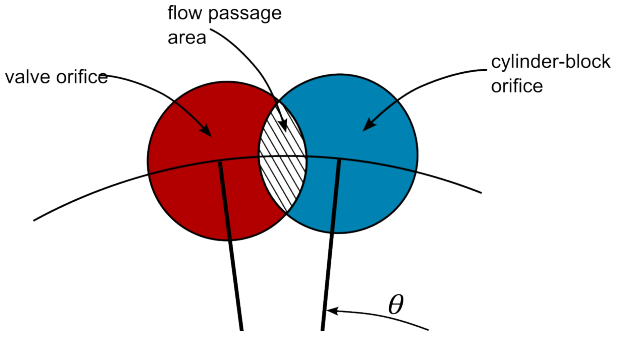


Figure 7: Valving interface flow passage area

bushing can be described as a partial arc journal bearing. The sliding motion is facilitated by the bushing's plastic material in low speed conditions, and a hydrodynamic fluid film separates the surfaces in higher-speed operation conditions. A semi-empirical model is used, building upon elementary test results realised on a radial pistons cam lobe motor. This model relates the friction coefficient to the Sommerfeld number defined by eq. (9), and leads to friction coefficient values shown in fig. 8. A particularity that is not captured by the standard Stribeck curve is a higher dependence on the load applied to the contact. In fact, the load parameter P^* is experimentally found to have an impact on friction coefficient, beyond the Sommerfeld parameter. This is due to the pressure-induced deformation of the bushing, which tends to increase the friction coefficient at lower loads, and which would require an elasto-hydrodynamic modelling approach, coupling parts deformation and fluid film pressure to be simulated accurately. Due to simulation running time concerns, and scope of the work, this is not attempted here, and the Stribeck curve model is adapted to reflect the pressure dependence induced by deformations. Hence, several curves appear on fig. 8, to highlight this stronger pressure dependence.

$$S = \frac{\mu \cdot V}{p^*} \tag{9}$$

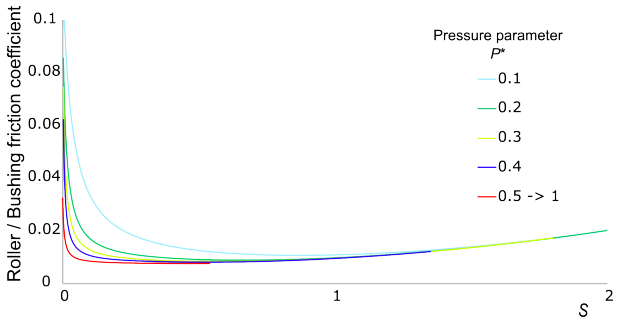


Figure 8: Pressure dependent Stribeck curves for the Roller/Bushing friction coefficient cof_{RB}

Using the friction coefficient given in fig. 8, the friction force in the roller/bushing contact can easily be obtained as the product of the friction coefficient cof_{RB},and the theoretical roller/bushing reaction. However, the model does not cal-

culate friction individually, but rather uses the friction coefficient in the global piston force balance of eq. 6. To that effect, the angle φ_{RB} is calculated through eq. 10.

$$\varphi_{RB} = atan(cof_{RB}) \tag{10}$$

Piston/Cylinder friction In the same way, the piston/bore contact friction is modelled using a semi-empirical model based on the Sommerfeld number of eq. (9), where V now stands for the piston sliding speed $V_P(\theta)$. It must be noted that since the piston has a reciprocating motion in its bore, it constantly travels up and down the Stribeck curve of fig. 9. An additional particularity is that since the maximum piston sliding speed remains low compared to the roller/bushing sliding speed, the Sommerfeld number remains low, and the curve does not show a re-increase of the coefficient of friction at higher Sommerfeld values, indicating that the hydrodynamic regime is never reached. In a similar manner to the roller/bushing friction described previously, the angle φ_{PB} is obtained by way of eq. 11, and subsequently used in eq. 6.

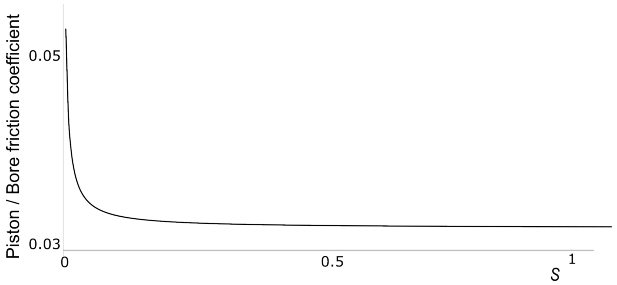


Figure 9: Stribeck curve for the Piston/Bore friction coefficient cof_{PB}

$$\varphi_{PB} = atan(cof_{PB}) \tag{11}$$

Bearing internal friction In addition to the two previous friction sources, the RPCLB includes a third main source of frictional loss in the bearings, which is much more significant in the booster design than it is in a radial pistons cam-lobe motor. The reason for the significance of the losses in the bearing is due to the higher speeds that can be reached by the RPCLB, and also to the fact that the net torque on the rotor of the RPCLB is zero, so any torque loss in the bearing results in a lower pressure amplification ratio. Bearings modelling and losses have been a subject of study for many years, and good engineering models are proposed by suppliers (see for example [15], and [16]). For the purposes of this study, because of its easy implementation, and the availability of the required parameters and equations, the SKF model [16] is used. It is an analytical model, which calculates frictional moment with the inputs of geometry, rotation speed and axial and radial forces on the bearing.

3.4 Volumetric losses model

The volumetric losses in the RPCLB are mostly internal leakages, as the leakage flow can only follow a limited number of paths:

- From high pressure (port P_1) to low pressure (port P_3).
- From boosted pressure (port P_2) to low pressure (port P_3).
- From boosted pressure (port P_2) to high pressure (port P_1).

These losses happen at different locations of the RPCLB, but the vast majority of the leakage flow is concentrated at the piston sealing rings (6, 7) on fig.2). An empirical model is used to relate leakage flow to oil sealing ring differential pressure and fluid viscosity. Indeed the leakage flow through an individual sealing ring is found to be linearly correlated to the pressure level, and the effect of piston sliding speed is very limited when looking at leakage values averaged over one RPCLB revolution.

$$Q_{leak} = \alpha_r \cdot \Delta p \cdot \nu^{0.4} \tag{12}$$

Equation (12) is used in the model to calculate the leakage flow through an individual sealing ring. In this equation, α_r is a constant determined by the sealing ring type, and ν is the kinematic viscosity of the hydraulic fluid. The individual leakage flows through the different sealing rings are summed in order to obtain the leakage flow through each of the three paths described previously.
In addition to these leakage flows, the compressibility of the fluid also leads to a volumetric inefficiency, which can be assimilated to an internal leakage flow. Compressibility losses impact fluid volumes that see a switch between pressure levels. In the RPCLB, this is the case of the oil volume contained in the cylinder-block, between the valving interface and the bottom surface of the piston, as illustrated on fig. 10 for the piston's foot.

$$\Delta V = \frac{V_0 \cdot \Delta p}{\beta} \tag{13}$$

This volume switches from boosted to low pressure N_L times per revolution. Using eq.(13), the volume that is lost due to compressibility effects can be calculated. It is then transformed in a volumetric flow rate by multiplying the value of ΔV by $N_L \times L_P = 12$ times (due to the $N_L = 4$ commutation events for each of the $N_P = 3$ pistons) the rotation speed. In a similar manner, there is a volumetric flow rate that can be calculated, corresponding to the head part of the piston, where the pressure switches from low to high at bottom dead centre (where the oil volume under the piston is $V_{0,H}$, or dead volume under the piston's head), and from high to low at top dead centre (where the oil volume under the piston is $V_{D,H}$, which

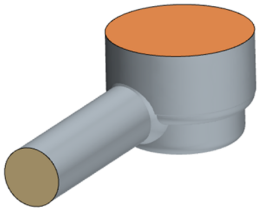


Figure 10: Dead volume under the piston's foot

is the volume displaced by the head part of the piston). Taking into account the compressibility effects leads to the well known *p-V* diagrams depicted in fig.11 for the motoring cycle (orange), and the pumping cycle (red).

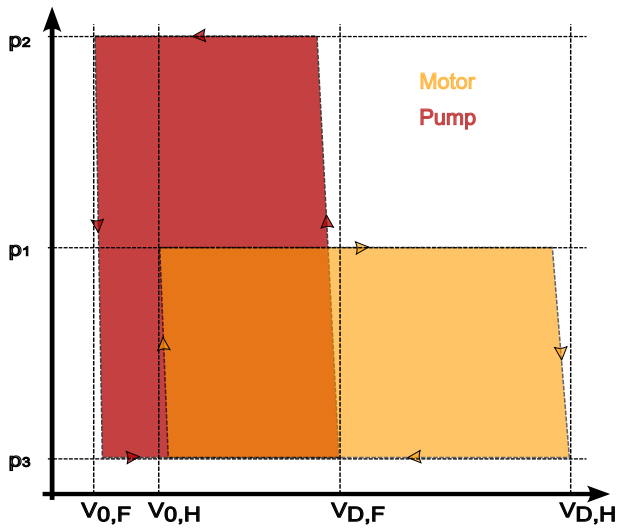


Figure 11: p-V diagram for motor and pump cycles

3.5 Complete simulation

Putting together all the individual bricks presented in the previous sections allows the simulation of the behaviour of the RPCLB for a given set of inlet conditions (flow, pressure and temperature), and low pressure setting. The complete simulation process is described in fig. 12. After convergence of the

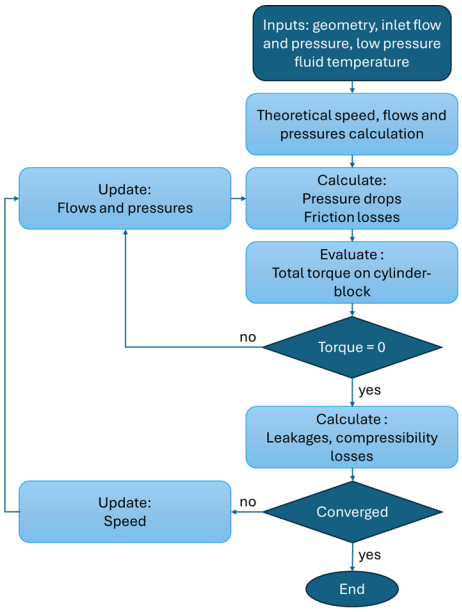


Figure 12: Complete simulation process

pressure and flow values on the RPCLB's outlet ports P_2, and P_3, the pressure amplification ratio, the ratio of outlet (P_2) flow over inlet (P_1) flow , and the global transformation efficiency can be calculated. In addition to these average values, some angle dependant values can also be interesting. Indeed, the simulation model is able to output the flow and pressure values in function of the cylinder-block's angular position, therefore allowing a calculation of flow and pressure stability metrics.

4 Results and discussion

4.1 Transformation ratio

The transformation efficiency links the increase of pressure to the reduction of flow rate, as seen from P_1 to P_2. This efficiency is calculated with the help of eq. (14).

$$\eta_{trans} = \frac{p_{P_2} \cdot Q_{P_2}}{p_{P_1} \cdot Q_{P_1}} \tag{14}$$

Figure 13 shows the transformation efficiency map for the RPCLB, as defined by eq. 14, in function of the pressure and flow at the inlet port P_1, under a hydraulic fluid inlet temperature of $50^{\circ}C$, and a low pressure on the outlet port P_3 set to 20 bar. Figures 14,15 respectively show the corresponding pressure increase ratio, equal to p_{P_2} over p_{P_1}, and the flow rate at the outlet port P_2.

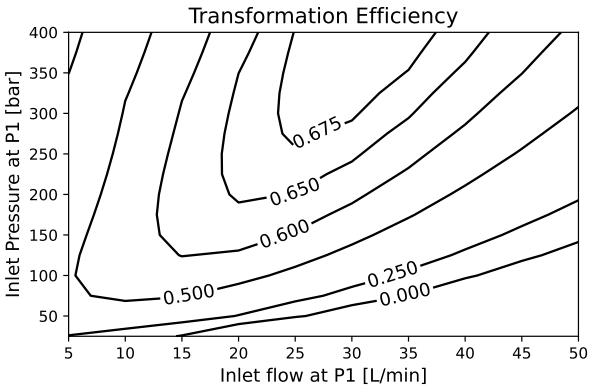


Figure 13: Overall transformation efficiency η_{trans}

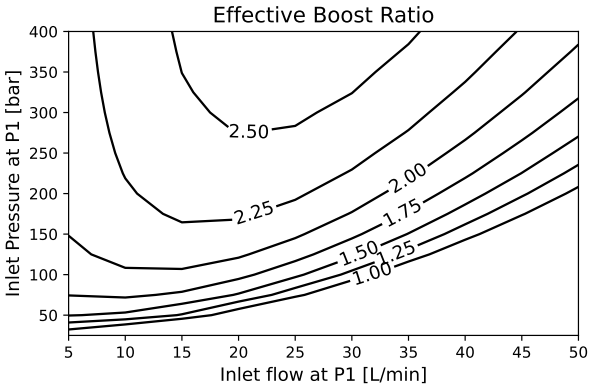


Figure 14: Effective pressure increase ratio ($p_{P_2} \div p_{P_1}$)

4.2 Valid operating conditions

The transformation efficiency map of fig.13 shows that the efficiency can be equal to zero in some conditions of high inlet flow and low inlet pressure. In addition, fig.14 also displays ratios lower than one in a slightly more extended zone.

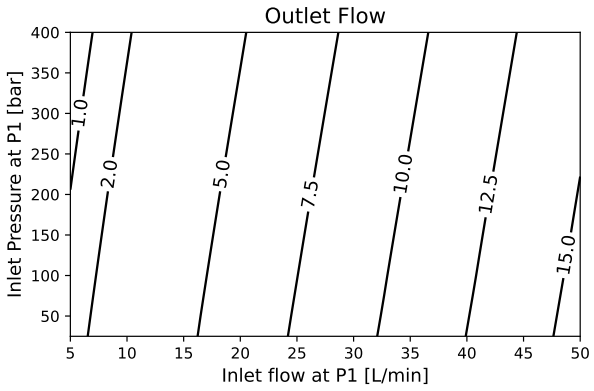


Figure 15: Outlet flow Q_{P_2} (L\min)

This clearly defines operating conditions where the RPCLB is – at best – counter-productive, and perhaps even inoperable. Figure 16 shows in a red zone the operating conditions where

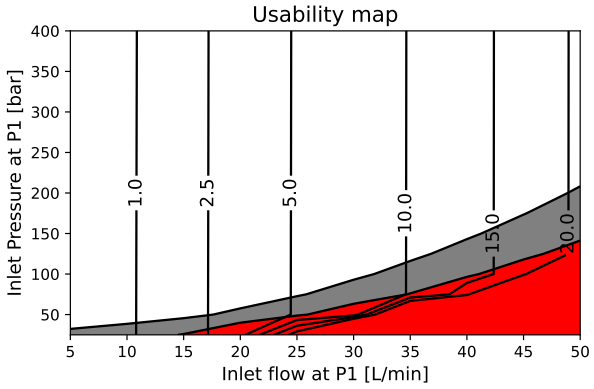


Figure 16: Possible operating conditions of RPCLB

the RPCLB is inoperable, and in grey the operating conditions where it is counter-productive, and the outlet pressure at port P_2 is actually lower than the inlet pressure at port P_1. Superimposed on these coloured areas, are contour lines indicating the minimum pressure requirement on the outlet port P_3 to ensure that no cavitation occurs inside the component.

4.3 Flow stability

Using the simulation model, it is also possible to investigate the variations of outlet flow during one rotation of the RPCLB's cylinder-block. Indeed, as explained in §1, the flow variation on the outlet line of the pressure amplifier is an important criterion for the pressure booster. The outlet flow variation can be calculated by the developed simulation model, and shows low values of flow ripple. The variation is in the range of ±3% of the nominal flow value, which is a good improvement on the pulsating outlet flow that is created by standard – commercially available – pressure amplifier. Figure 17 compares the outlet flow, plotted in function of time, for a commercially available pressure amplifier and the RPCLB in identical conditions: 200bar inlet pressure and 20L/min inlet flow.

As shown by fig. 17, the RPCLB shows a much lower flow

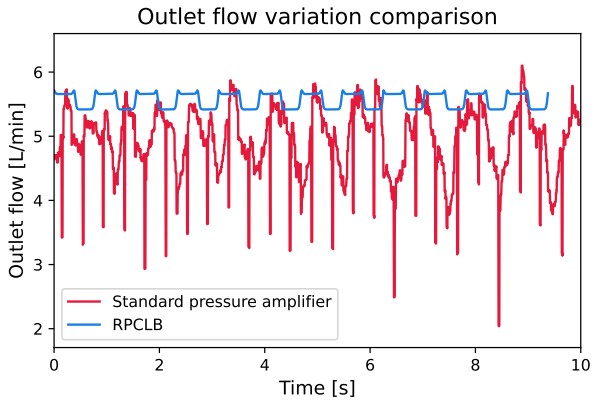


Figure 17: Comparison of simulated RPCLB *and measured commercially available pressure booster outlet flow. 200 bar and 20L/min inlet*

ripple than the standard pressure amplifier, making it a better fit for integration in a the hydrostatic transmission of a mobile application.

4.4 Experimental results comparison

The design of RPCLB presented in this paper, and simulated by the model has been fabricated and tested on an elementary test bench (fig. 18). The inlet flow at port P_1 is controlled,

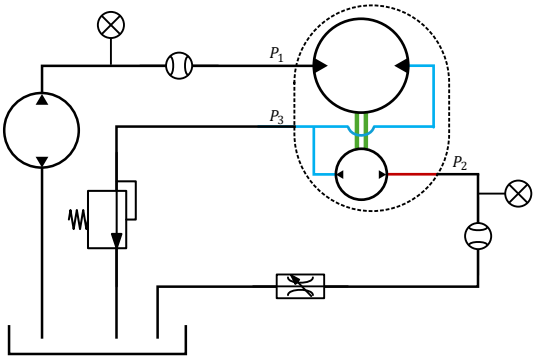


Figure 18: Elementary RPCLB *test bench setup*

and the pressure on the outlet port P_2 is set using a variable orifice. Due to test bench limitations, the boosted pressure level is limited to 450 bar. Figures 19, and 20 respectively show the outlet pressure and flow at port P_2, comparing simulated and measured values. The value of inlet pressure (at port P_1) is indicated by way of a colour scale, and the dots are simulated points, whereas the '×' markers represent the measured points. For both the outlet pressure and flow, the simulation shows reasonably good agreement with the measured data. The difference between simulated and measured values, in percent, as calculated by eq. (15), is listed in tab. 1.

$$error_{\%} = 100 \times \frac{\bar{x}_{sim} - \bar{x}_{exp}}{\bar{x}_{exp}} \tag{15}$$

The simulation model correctly captures the overall trends of the RPCLB's volumetric and mechanical efficiencies. It can be

Table 1: Simulated vs measured value errors

error in %	pressure	flow
min	-8.64	-3.13
max	6.62	1.32
average	-1.28	-0.43

noted that the model seems to underestimate the outlet pressure for low flow values, which could indicate that the friction losses in the boundary lubrication regime are in fact lower than expected. As indicated in §3.3.2 the friction models for roller-bushing and piston-cylinder are based on measurements made on a radial pistons cam lobe motor of similar design to the RPCLB. However the pistons and rollers of the RPCLB are slightly smaller than their existing counterparts of the motor range, and some size effects could lead to this overestimation of boundary lubrication friction levels. Alternately, the real friction of the tapered roller bearing could be less than the SKF model predicts for low rotation speeds, as this kind of engineering model is more suited to higher speed operations commonly encountered by bearings.

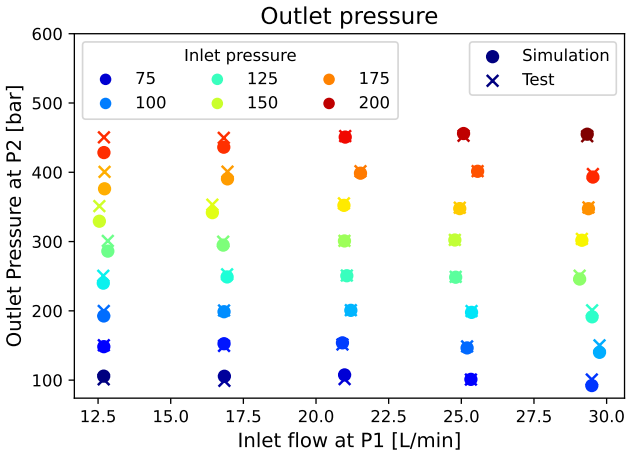


Figure 19: Comparison of simulated and measured outlet pressure

5 Conclusions and next steps

In this work, a concept of radial pistons cam lobe pressure intensifier is introduced, and a simulation model is described, which allows the calculation of outlet pressure and flow for given inlet conditions. The different parts of the model are described, accounting for hydromechanical and volumetric losses. Different outputs of the model are detailed, showing the amplification ratio, the transformation efficiency, and highlighting the limits in operating conditions. The simulation results are shown to be in good agreement with experimental results, with some possible modelling improvements to be made in the low flow regions.

The first tests realised on the RPCLB, and the simulation results, have shown a potential for improvements in the design of the component. In fact this was to be expected, as the

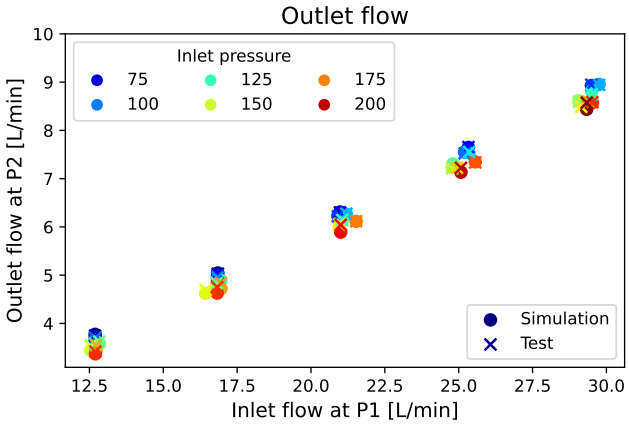


Figure 20: Comparison of simulated and measured outlet flow

tested version of the component is a preliminary design, with un-optimised parts and fabrication processes. The target was to have a functional component, and to evaluate it in light of the model. The simulation model will now be utilised to vary design parameters, in order to optimise the pressure intensifier's efficiency. The size and number of pistons, the size of the channels inside the RPCLB, as well as the internal distribution of fluid can be modified in the model to investigate the effect on outlet flow and pressure, and ultimately, to optimise its overall efficiency.

Nomenclature

Symbol	**Description**
N_L	Number of lobes
N_P	Number of pistons
HP_{sys}	High system pressure
LP_{sys}	Low system pressure
HP_{boost}	Amplified pressure
D_{P_h}	Piston diameter (head)
D_{P_f}	Piston diameter (foot)
R_r	Roller radius
L_P	Piston total length
θ	Angular position
ρ	Roller centre position
α	Cam-roller contact point angular position
R_{CRC}	Cam-roller contact radial position
φ	Cam-roller contact incidence angle vs piston axis
ω	Rotation speed
V_P	Piston radial speed
Γ_P	Piston radial acceleration
R_{CAM}	Roller/Cam reaction force
R_{PB_H}	Piston/Bore contact force - piston head
R_{PB_F}	Piston/Bore contact force - piston foot
φ_{PB}	Piston/Bore friction coefficient

φ_{RB}	Roller/Bushing friction coefficient
β	Bulk modulus of the hydraulic fluid
P_1	1^{st} port of the RPCLB
P_2	2^{nd} port of the RPCLB
P_3	3^{rd} port of the RPCLB
Q_{Pist_M}	Flow for each piston in the motor stage
Q_{Pist_P}	Flow for each piston in the pump stage
D_M	Motor displacement
D_P	Pump displacement
$r_{amp_{th}}$	Theoretical amplification ratio
F_{P_H}	Pressure in piston's head chamber
F_{P_F}	Pressure in piston's foot chamber
F_{CASE}	Case pressure
F_{dist}	Distribution valve axial pushing force
Δ_P	Pressure drop
ρ_{oil}	Hydraulic fluid density
ξ	Pressure drop coefficient
V_{fluid}	Fluid speed in reference flow area
ξ_1	Inlet/Outlet pressure drop coefficient
ξ_2	Pressure drop coefficient in grooves flow area
ξ_3	Pressure drop coefficient in valve channel flow area
$Q_{Pist}(\theta)$	Piston flow requirement at a given angular position
$A(\theta)$	Flow passage area at interface
S	Sommerfeld number
μ	Oil dynamic viscosity
V	Oil speed in reference area
p^*	Pressure parameter for friction coefficient model
Q_{leak}	Leakage flow
V_0	Dead volume under piston
$V_{0,H}$	Dead volume under piston's head
$V_{D,H}$	Dead and displaced volume under piston's head
$V_{0,F}$	Dead volume under piston's foot
$V_{D,F}$	Dead and displaced volume under piston's foot
η_{trans}	Transformation efficiency
p_{P_1}	Pressure at port P_1
p_{P_2}	Pressure at port P_2
Q_{P_1}	Flowrate at port P_1
Q_{P_2}	Flowrate at port P_2
$error_{\%}$	Simulation result error compared to measurement
$\overline{x}_{sim}$	Simulated value
$\overline{x}_{exp}$	Measured value

References

[1] Thomas Heeger, Samuel Kärnell, and Liselott Ericson. Challenges for multi-quadrant hydraulic piston machines. *Energy Conversion and Management: X*, 22:100578, April 2024.

[2] Damiano Padovani, Pavlos Dimitriou, and Tatiana Minav. Challenges and solutions for designing Energy-Efficient and Low-Pollutant Machines in Off-Road hydraulics. *Energy Conversion and Management: X*, 21:100526, January 2024.

[3] Peter Achten. Innovation in the fluid power industry. *IFK Aachen*, 2014.

[4] HAWE. Radial piston pump type R, RG product documentation, 2024. [Accessed 2025-02-14].

[5] Poclain. MHP20/MHP27 Hydraulic Motors technical catalog, 2024. [Accessed 2025-02-14].

[6] Pierre Bernard, Vincent Langlois, and Maxime Maurer. Improving transmission efficiency via a novel radial piston cam lobe motor architecture. *Proc. of the 2024 International Maha Fluid Power Conference*, pages 20–38, 2024.

[7] Poclain. Advanced drive technology. *IVT Off-Highway Annual 2025*, pages 54–56, 2025.

[8] G M Barrett. Fluid pressure intensifier, 1973. Patent GB1317742A.

[9] Hammelmann Paul. Hydraulic pressure transformer, 1983. US Patent US4401415A.

[10] Feng Wang, Linyi Gu, and Ying Chen. A hydraulic pressure-boost system based on high-speed on-off valves. *IEEE-ASME Transactions on Mechatronics*, 18(2):733–743, 2013.

[11] Peter Achten. The design of the FCT80 hydraulic transformer. In *SICFP'23*, Tampere,Finland, June 2023.

[12] Gilles Lemaire, Yoann Rousseau, and Paul Fournet. Machine hydraulique multicylindrees amelioree, 2024. Patent WO2024231635A1.

[13] Chao Zhang, Hao Tan, Yu Fang, Xiaolong Zhang, Yu Yang, Yiman Duan, Min Han, Shaojian Cui, Bing Xu, and Junhui Zhang. Deformation pre-compensated optimization design of cam ring for low pulsation hydraulic motors. *Journal of Zhejiang University-SCIENCE A*, 24(2):130–145, 2023.

[14] Leon Dahlén and Peter Carlsson. Numerical optimization of a distributor valve. *International Journal of Fluid Power*, 4(3):17–25, November 2003.

[15] Schaeffler Technologies AG & Co. KG. Rolling bearings technical - principles and product data for the design of rolling bearing arrangements. pages 56–61, 2020. [Accessed 2025-02-03].

[16] SKF. The SKF model for calculating the frictional moment. 2009. [Accessed 2025-01-06].

Preemptive Scheduling of Multi-rod Hydraulic Infinite Linear Actuator

Robert Braun , Magnus Sethsson , Magnus Landberg*, and Petter Krus

Department of Management and Engineering, Linköping University, Sweden
E-mail: robert.braun@liu.se, magnus.landberg@saabgroup.com, magse13@gmail.com
*Saab Aeronautics, Linköping, Sweden

Abstract

The Hydraulic Infinite Linear Actuator (HILA) is an actuator with a through-rod design, where the piston is able to grip or release the rod during operation. If multiple pistons, connected in series, take turn in gripping, moving, and releasing the rod, a continuous linear motion without the stroke limitations of conventional hydraulic actuators is achieved. This allows the actuator to extend or retract indefinitely while maintaining force continuity. The HILA system can be extended so that each piston is able to grip or release multiple rods, called HILA Multi-Rod (HILA MR). In this way, a few hydraulic pistons are able to control the motion of several rods. In order to determine which rod should be moved by which piston and when, a scheduling algorithm is required. This paper investigates two scheduling algorithms: time-domain multiplexing and round-robin scheduling. The round-robin algorithm is extended with an early-return feature and greedy job selection. All algorithms are evaluated on a simulation model of a HILA with three rods and one or two pistons running a reference cycle. The simulation results show that greedy job selection outperforms the other algorithms, both with one and two servos. As a final test, the scheduler was extended with the ability to give priority to one of the rods. The results show that the rod with priority performs very well and that the priority feature is fairly easy to implement.

Keywords: Actuator, piston, control, scheduling

1 Introduction

The Hydraulic Infinite Linear Actuator (HILA) is a specialized mechanism that enables continuous linear motion without the stroke limitations of conventional hydraulic actuators [1]. This is achieved through a system of interlocking hydraulic pistons that alternates between gripping and releasing the piston rod in a synchronized fashion. This allows the actuator to extend or retract indefinitely while maintaining force continuity. HILA combines the power density of hydraulic actuators with the benefits of a long stroke, enabling for compact and efficient solutions. Experiments have been conducted on applications in various fields, including agricultural and forest machinery [2], elevators in mid-rise buildings [3], and not least aerospace applications [4] where weight and volume reduction is of great importance. In this paper, four different scheduling algorithms for controlling three rods with one or two hydraulic pistons are evaluated against a simulation model.

In its simplest form, the HILA system, with a single piston rod, utilizes a hydraulic membrane clamping mechanism to connect the piston and rod when pressurized. The clamping mechanism relies on friction to transmit force between the piston and the rod. This concept has been thoroughly investigated and confirmed in previous experiments. This configuration integrates two short-stroke cylinders with engaging and disengaging clamping mechanisms to function as a long-stroke actuator. Through the use of simple directional valves, the piston and rod can be securely and reliably engaged or disengaged. The sequential movement of each short-stroke piston, linked through the clamping mechanisms, produces the overall motion of the piston rod. This design enables a stepwise motion along the rod, similar to the mechanics of rope climbing actuators [5].

1.1 HILA Multi Rod

A variation known as HILA Multi-Rod (HILA MR) [6] employs multiple piston rods operating within a shared piston inside a common cylinder, see fig. 1. Each piston rod is independently controlled through separate clamping elements or static clamping members, utilizing individual logic valves. A servo valve regulates the movement of the shared piston, while independent pressurization of the clamping elements allows precise control over each linear rod. By alternating between continuous and discrete incremental positioning steps, HILA MR achieves a form of motion control that parallels the time-multiplexing schemes found in computer operating systems.

HILA MR technology replicates the functionality of hydraulic

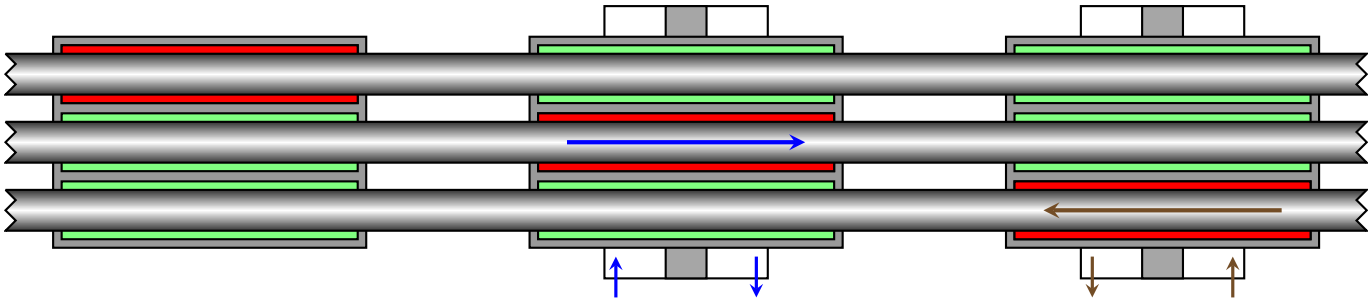


Figure 1: Multi-rod HILA with two servos controlling three rods. Red indicates that a clamp is engaged, and green disengaged. The top rod is parked, while the first servo is moving the second rod towards the right and the second servo is moving the third rod towards the left.

servo cylinders but operates by sequentially actuating multiple rods through clamping elements. Each servo piston can grip one or more rods at the same time and then move them in the desired direction. The technology enables a controlled transition between continuous motion and discrete incremental positioning, governed by precise timing. This paper investigates the special case where each piston only grips one rod at a time. This time-multiplexing approach resembles scheduling mechanisms in computer science. HILA MR technology provides a novel method for generating and distributing linear mechanical motion in various applications that require precise and scalable actuation. It bridges the gap between traditional high-fidelity hydraulic servo systems and applications with less stringent time constraints. A key advantage is its ability to enable stroke independence, reducing the force exerted on the rod while extending the stroke length without altering torque and stiffness. This flexibility enhances both design and performance, making HILA technology a promising solution for modular and efficient actuation.

In aviation, this technology offers significant advantages, including a reduced number of servo valves, the integration of well-established clamping techniques for both stepping and locking functions, and fewer position feedback sensors. A single HILA MR actuator has the potential to replace an entire actuator and sensor network for aircraft surface control. This enables substantial rationalization by reducing mass and volume per actuator function, leading to more efficient and compact aircraft control systems.

The possibilities of using HILA MR in aerospace applications were investigated by Landberg et al. [4]. The long stroke technology, combined with a high gear ratio lever arm, offers very high forces for surface control in slender wings. In this configuration, the left and middle piston rods are connected to the control cylinder. Movement and force are transmitted to the control surface through a cable or push rod and a lever arm. When applied to flight control, HILA MR shares similarities with fly-by-wire systems but differs in key aspects. The actuator is positioned within the fuselage or inside the wing box and is integrated with an electro-hydraulic power pack or linked to a redundant central hydraulic system. Rudder movements are transmitted through rigid cables or push rods that run along the wings, reminiscent of the traditional mechanical cable systems used in early aviation [7].

1.2 Clamping Mechanism

The system includes six or nine clamps, if using a single or a dual servo, respectively. Of these, three are *parking clamps*, which are attached to a fixed frame and prevents motion of the rod when engaged. The others are *transfer clamps*, which locks the rod onto a movable piston to enable transfer of motion. Figure 2 shows a simulation model of the system, where (c_{01}, c_{02}, c_{03}) are the parking clamps and $(c_{11}, c_{12}, c_{13}, c_{22}, c_{22}, c_{23})$ are the transfer clamps. The scheduling component sends control signals to each one of the clamps, as well as reference position signals to the hydraulic position servos. For simplicity, each servo is controlled by a proportional controller through closed-loop feedback.

2 Scheduling Algorithms

Four different scheduling algorithms have been implemented and analyzed in the hydraulic model; Time Division Multiplexing, Round Robin, Early-Return Round Robin, and Greedy Round Robin. They differ mainly in how they select which rod to grip and how long they wait until they grip the next rod. All algorithms are implemented in C++ and can operate on either one or two hydraulic servos. When only one servo is used, the servo must return to neutral position between each motion. If two servos are used, one servo moves the selected rod while the other performs the return motion.

2.1 Time Division Multiplexing

Time Division Multiplexing (TDM) is a technique used in telecommunications to transmit multiple signals over a single communication channel. It works by dividing the channel's available time into separate time slots and assigning each signal a unique time slot for transmission. This allows multiple data streams to share the same physical medium efficiently, avoiding interference between signals. However, TDM can conceptually be adapted as a general task scheduler, which have been studied in several fields such as controlling multiple solar panels [8], air traffic signals [9] and surgical robots [10]. The key idea of TDM - dividing available time into discrete time slices and allocating them to specific tasks - makes it a suitable framework for controlling a multi-rod HILA, where multiple servos take turns in moving the rods. The rods are selected in sequential order, and each rod is gripped for a fixed

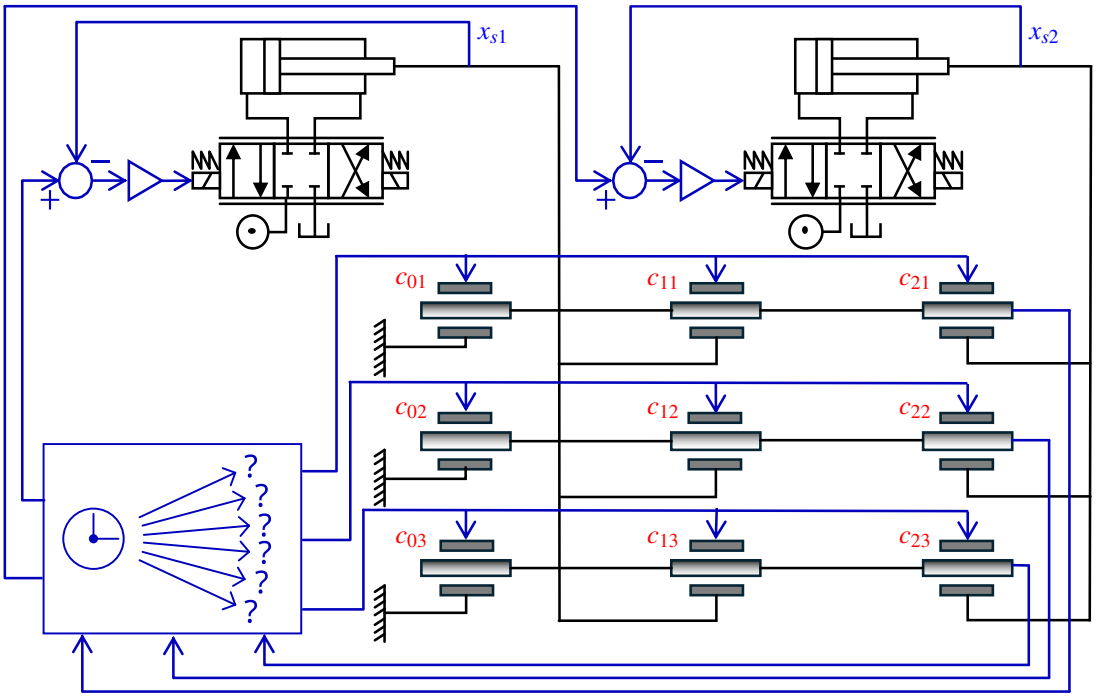


Figure 2: Schematic view of the simulation model.

period of time before it is released, a shown in figs. 3 and 4. A benefit with this approach is *predictable waiting time* - the time required for a specific motion can be calculated in advance. A critical parameter is the size of the time slots, which balances between movement resolution and overhead due to switching between rods and servos.

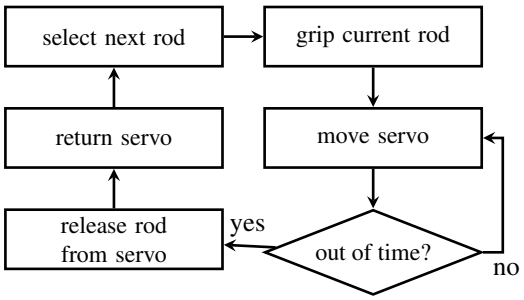


Figure 3: Flowchart of TDM algorithm for a single servo.

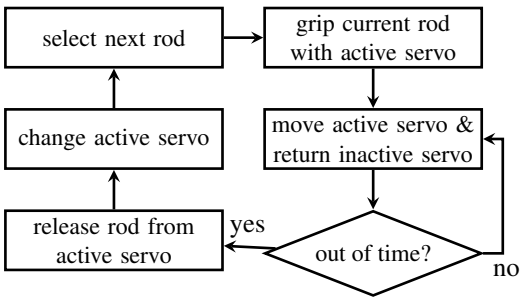


Figure 4: Flowchart of TDM algorithm for a dual servo.

2.2 Time Equidistant Round Robin

A more traditional scheduling technique is to balance the utilization by means of need and availability. Round Robin (RR) scheduling is a widely used CPU scheduling algorithm that allocates a fixed time quantum to each process in a cyclic order [11]. This approach ensures fairness by allowing all processes to execute for an equal time slice before moving onto the next. RR scheduling is particularly effective in time-sharing systems, as it minimizes starvation and provides predictable response times.

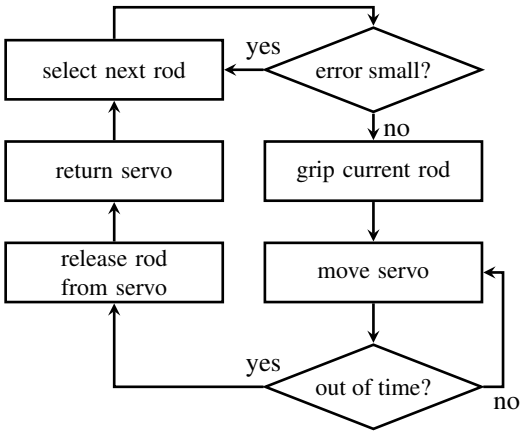


Figure 5: Flowchart of RR algorithm for a single servo.

The main difference between TDM and RR scheduling is that rods where the position error is below a given tolerance will be skipped over, so that only those rods with a significant error will be moved. Figures 5 and 6 shows flowcharts for the algorithm with one and two servos, respectively.

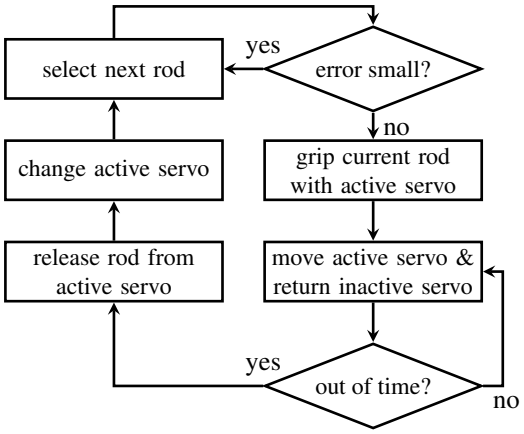


Figure 6: Flowchart of RR algorithm for a dual servo.

2.3 Early-Return Round Robin

Assigning a fixed time slice for each rod will be wasteful in case the rod reaches its reference position before the time slice has ended. The servo would then have to wait until the end of the current time slice before it is allowed to move any of the other rods. This can be avoided by aborting the motion early and changing to the next rod as soon as the position error is within a given tolerance. In this way the next rod can start moving earlier, improving the resolution of the motion, albeit at the cost of sacrificing the predictable waiting time. However, the worst-case time to complete a task can still be calculated. We denote this algorithm the *Early-Return Round Robin (ERR)* algorithm. When using two servos, it must be ensured that the second servo has returned to neutral position before the next rod can start to move. Otherwise the positions of the hydraulic servos might start to drift, eventually reaching their end-of-stroke limitations. Figures 7 and 8 shows flowcharts for the algorithm with one and two servos, respectively.

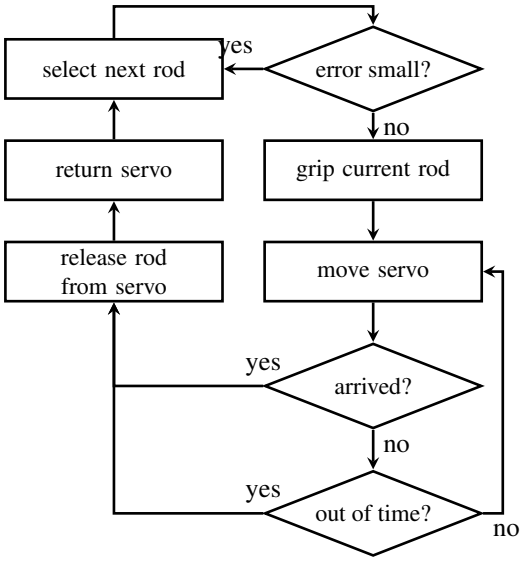


Figure 7: Flowchart of ERR algorithm for a single servo.

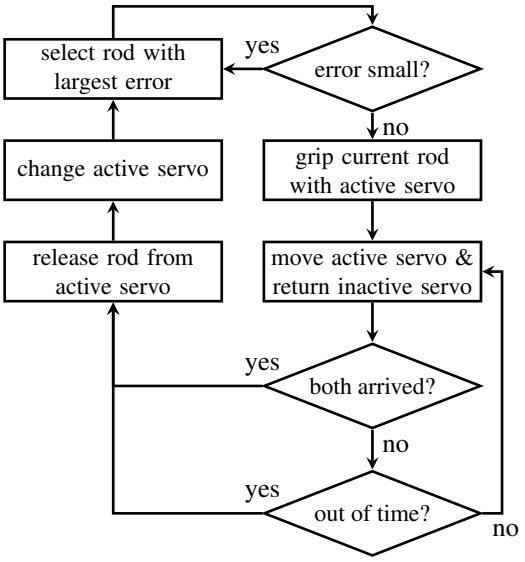


Figure 8: Flowchart of ERR algorithm for a dual servo.

2.4 Greedy Round Robin

The previous algorithms all select one rod at a time in a specific order, without priority. A possible improvement is to give priority to the rods with the largest position error. The fourth algorithm is an attempt to improve the fairness of the algorithm by adding a "greedy" rod selection. The servo always moves the rod with the largest error, calculated as the absolute difference between reference and feedback signals. Similarly, to the ERR algorithm, it also aborts the motion and selects a new rod once the current rod is close enough to its reference position. Note that this algorithm is not starvation free, and that waiting times are unpredictable. In the worst case, a rod with a small position error may need to wait very long until it is moved. Figures 9 and 10 shows flowcharts for the algorithm with one and two servos, respectively.

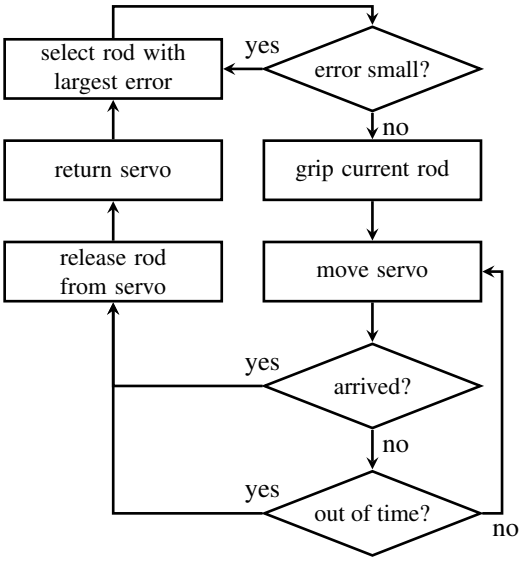


Figure 9: Flowchart of GRR algorithm for a single servo.

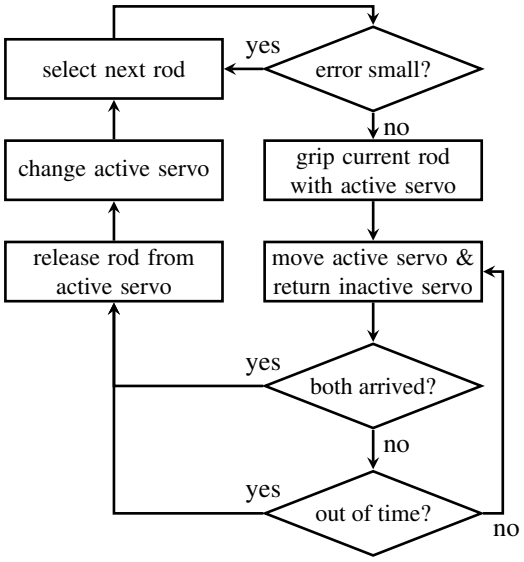


Figure 10: Flowchart of GRR algorithm for a dual servo.

3 Simulation Experiments

All algorithms are tested in the Hopsan simulation tool [12] with the model illustrated in fig. 2. Submodels for inertia and friction are left out of the figure for simplicity. It is important to evaluate the algorithms when moving one, two, or three rods simultaneously. For this reason, predefined input reference signals are used where all rods are moving, sometimes at the same time, and sometimes not. These references are used to study the deviation of the actual position of each piston rod against the corresponding input reference. The same reference input signals are used for all simulations and evaluations, as shown in fig. 11. One rod is moving all the time, one starts moving after 1 second, and the third only moves between 4 and 8 seconds. This makes it possible to compare how the different algorithms behave when moving multiple rods simultaneously. The rods also move with different frequencies. The fastest frequency is intentionally chosen high compared the time slices in the scheduling algorithm, to demonstrate the difference in dynamic response between one and two servos.

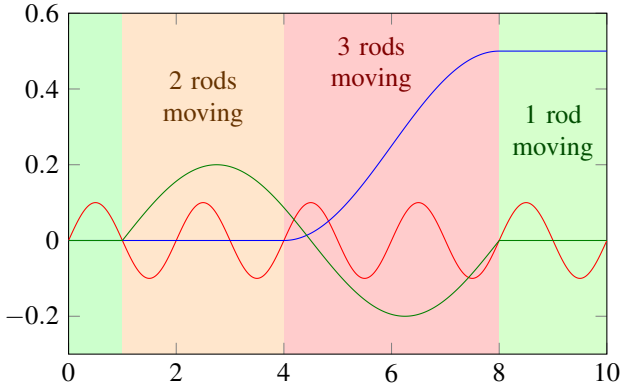


Figure 11: Reference signals for simulation experiments.

3.1 Results with Single Servo

The results for each algorithm with a single servo are shown in figs. 12 to 15. It is clear that TDM is not able to capture the fast dynamics of the first rod. The time resolution limits the maximum frequency response of the system. The other three algorithms work better, but still have problems with fast dynamics while moving all three rods at the same time. It should be noted that the time slices used are intentionally large compared to the fast dynamics of the model. In reality, one would choose time slices that are sufficiently small to capture the dynamics of the reference signals. However, the minimum size of the time slices is also limited by the response time of the hydraulic position servos.

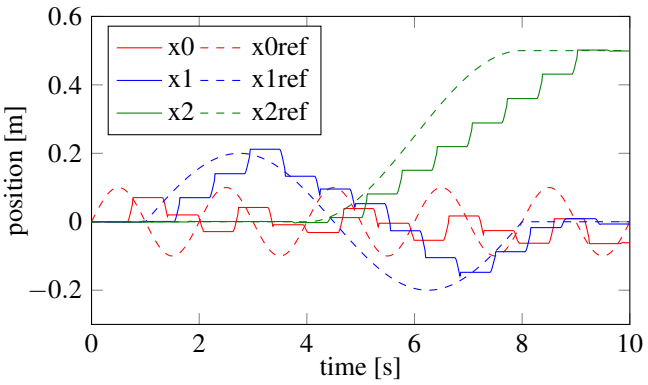


Figure 12: Simulation results for a single servo with TDM scheduling.

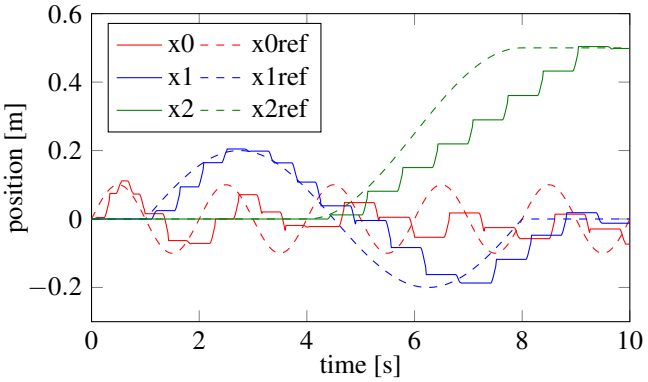


Figure 13: Simulation results for a single servo with RR scheduling.

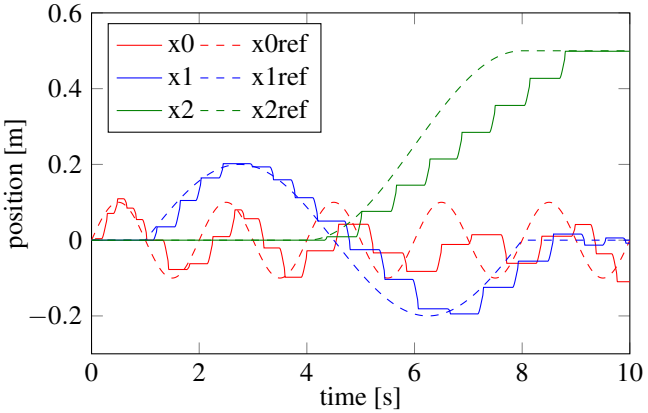


Figure 14: Simulation results for a single servo with ERR scheduling.

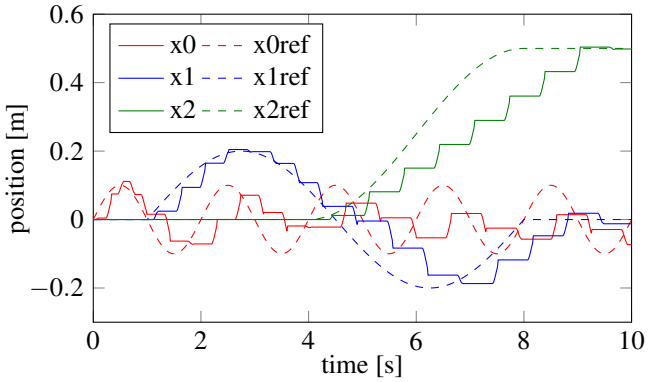


Figure 15: Simulation results for a single servo with GRR scheduling.

3.2 Results with Dual Servo

Results for a dual servo is shown in figures figs. 12 to 15 for each algorithm, respectively. All algorithms are now able to capture all frequencies in he references signals, but with different degree degrees of smoothness ranging from worst (TDM) to best (GRR). With TDM the high frequency motion suffers from a phase shift, as it is never really able to catch up with the reference signal.

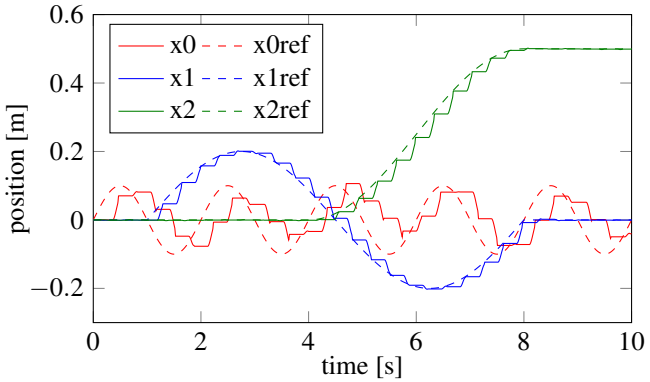


Figure 16: Simulation results for a dual servo with TDM scheduling.

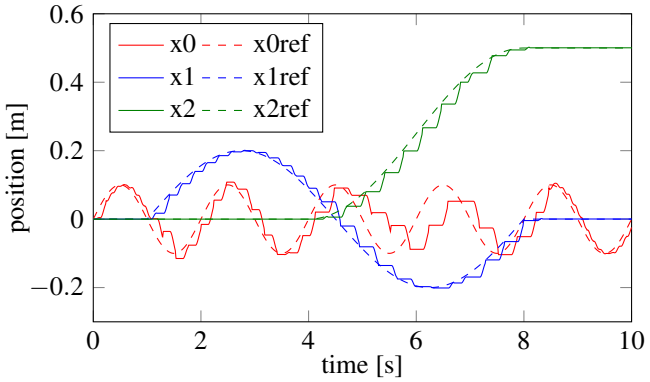


Figure 17: Simulation results for a dual servo with RR scheduling.

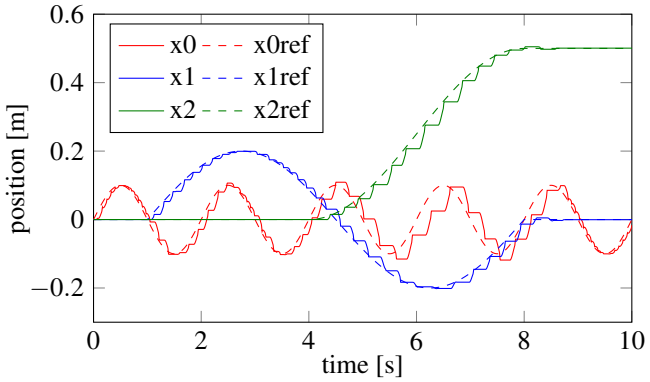


Figure 18: Simulation results for a dual servo with ERR scheduling.

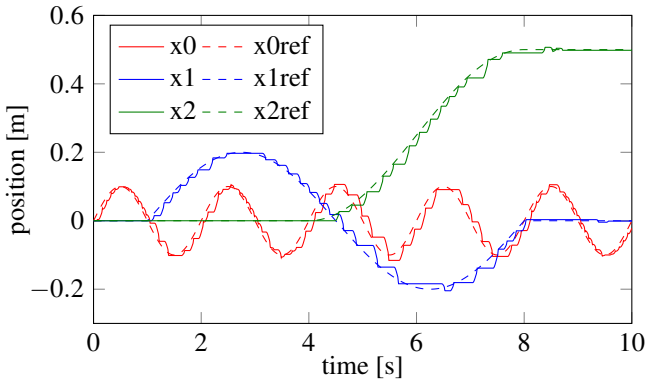


Figure 19: Simulation results for a dual servo with GRR scheduling.

3.3 Prioritization of Rods

In safety critical systems, not least aerospace applications, some motions are more important than others. For example, the control surfaces on an airplane is more important than the flaps and the landing gear. Hence, the scheduler must be able to give priority to these motions. This feature was added to the GRR algorithm. If a prioritized rod needs to move, no other rod is allowed to move until the prioritized rod has reached its reference position. Results with a single and a dual servo is shown in figs. 20 and 21, respectively.

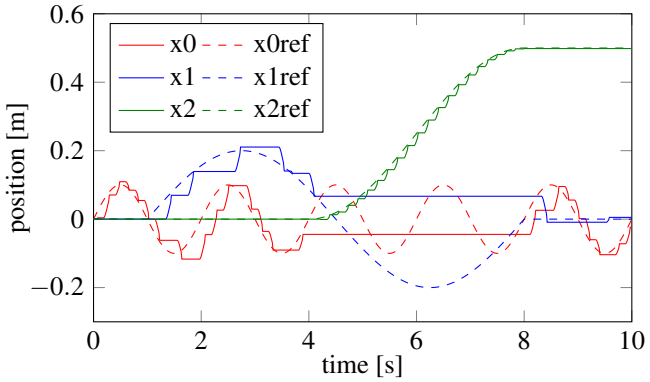


Figure 20: Simulation results for a single servo with GRR scheduling and priority for x2.

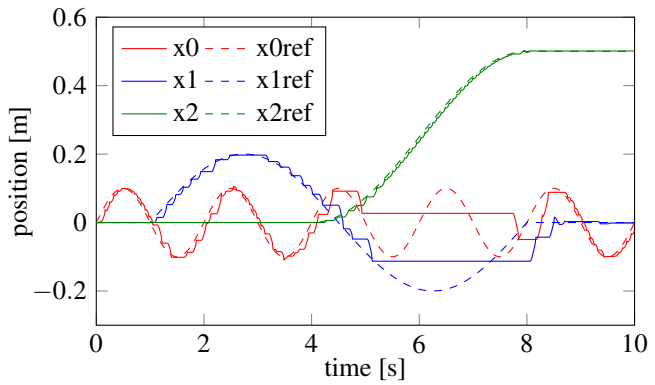


Figure 21: Simulation results for a dual servo with GRR scheduling and priority for x2.

4 Evaluation

The simulation results indicate that time-sharing scheduling algorithms, such as TDM and RR, can be effectively applied to HILA MR systems, both with one and two servos. From a scheduling perspective, it would be straightforward to extend the system to incorporate a third servo or additional rods. However, it is important to acknowledge that the simulation model contains several unknown model parameters. These include both design-related variables, depending on component selection, and physical effects, such as friction and leakage, which could influence the accuracy and predictive capability of the model.

The addition of a second servo significantly improves the dynamic response of the system. However, this enhancement comes at the cost of increased complexity, weight, size, and cost. Furthermore, the dynamic behavior of the system is highly dependent on the time slices allocated by the scheduling algorithm. If the time slices are sufficiently short relative to the fastest dynamics in the system, even a single-servo configuration could adequately capture the required motion characteristics.

Among the evaluated algorithms, GRR demonstrated the best overall performance. However, it is not starvation-free, meaning that all rods cannot be guaranteed to receive actuation within finite time. Although this issue is unlikely to occur in practice, it poses a potential risk in safety-critical applications. To mitigate this limitation, we have shown that GRR can be enhanced with a priority mechanism, ensuring that safety-critical motions are always executed when necessary.

To quantitatively compare the algorithms, maximum and average errors were extracted from the simulation data. Table 1 presents the maximum positional error observed for any individual rod across different algorithm and servo configurations, while Table 2 reports the average total error accumulated across all rods during a full cycle. Both tables further confirm that GRR consistently outperforms the other algorithms in terms of accuracy and control precision.

algorithm	single servo	dual servo
TDM	203 mm	136 mm
RR	205 mm	130 mm
ERR	188 mm	97 mm
GRR	154 mm	48 mm

Table 1: Maximum error from any one of the rods during the reference cycle.

algorithm	single servo	dual servo
TDM	177 mm	77 mm
RR	155 mm	55 mm
ERR	132 mm	39 mm
GRR	103 mm	32 mm

Table 2: Average total error during the reference cycle.

5 Conclusions

The HILA MR approach enables the design and control of more compact actuation systems, as a single piston or a small number of pistons can control multiple piston rods. In addition, the stroke of each cylinder is not constrained by the stroke of the rods, allowing short cylinders to perform long-range motions. This characteristic is particularly valuable in applications where weight and volume are critical constraints, such as aerospace engineering.

This study demonstrates the feasibility of controlling HILA MR systems using various time-sharing scheduling algorithms. Moreover, it highlights the possibilities of assigning priority to specific rods, an essential feature for safety-critical applications. Among the tested algorithms, Early-Return Round Robin and Greedy Round Robin exhibited the most promising performance.

Future research directions include experimental validation of these algorithms on a physical hardware system, as well as the integration of more advanced and realistic models into the simulation framework. In addition, exploring more sophisticated scheduling strategies, such as algorithms that allow a single piston to grip multiple rods simultaneously, could further enhance system efficiency and performance.

References

[1] Magnus Landberg, Martin Hochwallner, and Petter Krus. Novel linear hydraulic actuator. In *Fluid Power Systems Technology*, volume 57236, page V001T01A054. American Society of Mechanical Engineers, 2015.

[2] Magnus Landberg, Magnus Sethson, and Petter Krus. The hydraulic infinite linear actuator for efficient and flexible timber and agricultural logistics. In *Scandinavian International Conference on Fluid Power*, pages 309–321, 2021.

[3] Magnus Landberg, Robert Braun, Magnus Sethson, and Petter Krus. The hydraulic infinite linear actuator applied to elevators in mid-rise buildings. In *Fluid Power Systems Technology*, volume 51968, page V001T01A030. American Society of Mechanical Engineers, 2018.

[4] Magnus Landberg, Magnus Sethson, Robert Braun, and Petter Krus. An actuator system to control multiple surfaces in a morphing wing. In *:*, pages 8–, 2023.

[5] Martin Hochwallner, Magnus Landberg, and Petter Krus. The hydraulic infinite linear actuator: Properties relevant for control. In *10th International Fluid Power Conference (10. IFK, Dresden, Germany*, volume 3, pages 411–424, 2016.

[6] Magnus Landberg, Magnus Sethson, and Petter Krus. The hydraulic infinite linear actuator with multiple rods. In *The 15th Scandinavian International Conference on Fluid Power, SICFP'17, June 7-9, 2017, Linköping, Sweden*, volume 144, pages 279–293. Linköping University Electronic Press, 2017.

[7] Martin Wehr, Andreas Pott, and Karl-Heinz Wehking. Bending fatigue strength and lifetime of fiber ropes. In Clément Gosselin, Philippe Cardou, Tobias Bruckmann, and Andreas Pott, editors, *Cable-Driven Parallel Robots*, pages 73–84, Cham, 2018. Springer International Publishing.

[8] Oscar López-Lapeña. Time-division multiplexing control of multi-input converters for low-power solar energy harvesters. *IEEE Transactions on Industrial Electronics*, 65(12):9668–9676, 2018.

[9] Ben Ahmed Mohamed, Boudhir Anouar Abdelhakim, Bouhorma Mohammed, and Ben Ahmed Kaoutar. A statistical multiplexing method for traffic signal timing optimization in smart cities. *TELKOMNIKA Indonesian Journal of Electrical Engineering*, 15(1):107–113, 2015.

[10] Yashuai Zuo, Glen Merritt, and Xuefeng Wang. Design of a novel surgical robot with rigidity-adjustable joints based on time-division multiplexing actuation. In *2020 8th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob)*, pages 885–890. IEEE, 2020.

[11] Taghreed Balharith and Fahd Alhaidari. Round robin scheduling algorithm in cpu and cloud computing: a review. In *2019 2nd international conference on computer applications & information security (ICCAIS)*, pages 1–7. IEEE, 2019.

[12] Robert Braun, Peter Nordin, Liselott Ericson, L. Viktor Larsson, Petter Krus, and Maria Pettersson. Hopsan: An open-source tool for rapid modelling and simulation of fluid and mechatronic systems. In *Proceedings of the BATH/ASME 2020 Symposium on Fluid Power and Motion Control (FPMC2020) :*. American Society of Mechanical Engineers (ASME), 2020.

Development of a hydraulic supply system with pressurized reservoir for a digital hydraulic actuator

Henrique Frassetto Cardoso, Túlio de Abreu Burgos Gonçalves, Marcos Paulo Nostrani and Victor Juliano De Negri

Laboratory of Hydraulic and Pneumatic Systems (LASHIP), Federal University of Santa Catarina, Florianópolis, Santa Catarina/Brazil
E-mail: henriquefrassetto2001@gmail.com, tulio.burgos@laship.ufsc.br, marcos.nostrani@ufsc.br, victor.de.negri@ufsc.br

Abstract

The aviation sector actively pursues enhanced energy efficiency and performance, as these improvements trigger a cascading effect that iteratively reduces system weight and size, ultimately lowering fuel consumption, a priority for minimizing operational costs and addressing environmental concerns. In flight control surface actuation systems, where hydraulics is widely employed for its maturity and reliability, such advancements may translate to reductions in hydraulic pumps and conditioning components, fewer and smaller hydraulic lines (and the fluid mass they contain), smaller heat exchangers, reduced fuel reserves required for thermal management, and minimized leakage risks at connection points. In this context, digital hydraulics has been in development in recent years as a promising approach to significantly improve the energy efficiency of flight control systems while maintaining the reliability critical for the application. This paper presents a hydraulic supply system capable of efficiently delivering three different constant-pressure lines, which can support digital hydraulic actuators under development for aviation and other high-performance systems. Since the technology is intended for aviation, the presented architecture includes a pressurized reservoir obtained through a combination of actuators connected to the reservoir and pump outlet. The proposed supply matches pressure reducing and relief valves together with a pressure-compensated pump and three accumulators to generate and maintain the demanded pressure lines. Via simulations on the Hopsan software, the solution is tested alongside with the digital hydraulic actuator for step and sine position inputs. The pressure lines show little fluctuation and the actuator performs as expected, which indicates that the proposed supply is capable of providing the necessary conditions of operation. The simulated system as a whole (supply and actuator) achieves efficiencies of up to 59%, far greater than the obtained by traditional servo hydraulics, and cavitation is avoided in all times.

Keywords: digital hydraulics, supply system, energy efficiency, flight control systems, pressurized reservoir.

1 Introduction

In the recent decades, the growing concern over the environmental situation has affected how engineering decisions are taken. This shift pushes attention to solutions that present distinguished energy efficiency, not only across entire machines, but also within each of the systems that comprise them.

The transport sector relies on oil for more than 95% of its energy needs, which leads to a consumption of about 60% of the worldwide petroleum supply as fuel [1]. Aviation has grown faster than any other transport mode. However, due to the time taken for new technologies to penetrate the sector, efficiency may lag up to 20 years. [1]

According to ICAO [2], international aviation by itself contributes around 2% of the global CO_2 emissions and 12% of the total transport emissions. Nevertheless, the emissions from aviation are projected to increase by two- to fourfold by 2050. In this scenario, the sector could be responsible for 4.1 to 11.3% of the total carbon budget set by the Paris Agreement

to limit global warming to 1.5°C [2]. Therefore, the industry must not only switch to sustainable aviation fuels, however also increase the efficiency of all onboard systems. [2]

Hydraulics is one of the systems heavily applied in aviation. This technology is most often responsible for braking systems, cargo doors, wheel steering, landing gear movement and flight control surfaces operation [3]. According to [4], hydraulic systems are constantly engaged in critical tasks, especially in the context of flight control surfaces, due to their proven reliability, high stiffness and precise control.

Hydraulic systems applied in aircraft are conventionally based on servo-hydraulic actuators (SHA). Such technology is mature and trusted, however it can be improved in energy efficiency. This energy efficiency is defined mostly by throttle losses and internal leakage in the components [4]. The authors in [5] simulated the hydraulics in flight control under commonly used maneuvers in commercial flights and found that the efficiency of a conventional hydraulic positioning system in such cases can be lower than 3%, depending on the mission.

An alternative to improve the energy efficiency is the use of a Digital Hydraulic Actuator (DHA) presented in [6] [7]. In this conception, the servo valves are replaced by a set of on/off valves, which can select between different pressure lines for each chamber of a hydraulic cylinder. In this configuration, the system responds similarly to servo-hydraulics while reducing the energy dissipation on the valves, minimizing the throttling and internal leakage. The combination of the different pressures with the actuator areas allows a spectrum of forces with p^A discrete force values, where p is the number of pressures and A is the number of actuator chambers. In the specify configuration shown in [6], the four actuator chambers and three pressure lines results on 81 discrete force possibilities. Furthermore, the DHA offers enhanced reliability and potentially lower production costs, as it replaces a single complex component (servovalve) with multiple simpler ones operating in parallel. Although this increases the total number of valves, it improves fault tolerance and each unit is significantly simpler and cheaper to manufacture. For a comprehensive discussion of additional aspects and advantages of the DHA, readers are referred to [5], [6] and [7].

One of the challenges for the implementation of digital hydraulic actuators, such as the one presented in Figure 1, is the hydraulic supply system. It must provide multiple constant pressure values while maintaining energy efficiency. In [6], the author proposes a configuration where the high-pressure line is connected to one of the aircraft's centralized hydraulic supplies, the medium-pressure line to another, and uses the return to reservoir as the low-pressure line. However, this approach as an impact on reliability, as the supply lines are no longer segregated and, therefore, a failure in one of the systems could propagate to all. Furthermore, the return line could introduce issues regarding cavitation and fluid availability at the pumps if used as a pressure line.

In [7], the authors suggest that the supply system for DHA could be composed with digital pumps, as they improve efficiency. Such technology is presented in [8] and [9] with different approaches. For both cases, the control systems involved are velocity controlled and, therefore, have a strong reliance in flow rate regulation. Since the DHA is force controlled, as approached in [6], the solution presented would not be ideal.

This paper presents the development of a hydraulic supply power unit for the DHA presented in [6], while ensuring segregation and without compromising energy efficiency. Maintaining the pressures constant is critical, as large variations affect the force spectrum. In addition, the system must be capable of pressurizing the reservoir line in order to avoid cavitation in the pump, as the proposed application is aircraft. In this case, the actuation system requires three different pressures and, therefore, the supply was developed considering lines for high-pressure, medium-pressure and low-pressure; and a reservoir line.

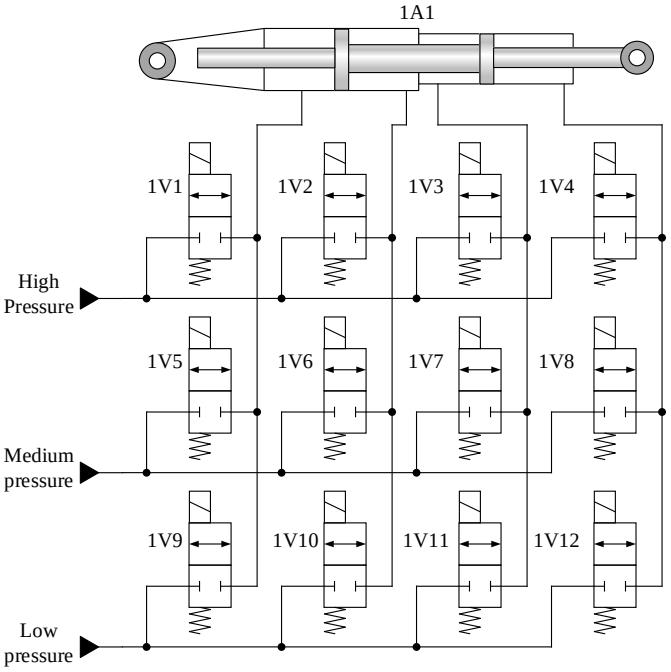


Figure 1: DHA in details with its pressure lines (Based in [6]).

This paper is divided into five sections. Section 2 presents the supply system layout and the function of each selected component. Section 3 describes the efficiency approach and equations. Section 4 shows the simulation model, the system parameter definition and the results obtained. Finally, Section 5 brings conclusions and, briefly, further advances using such technology.

2 System layout

2.1 Pressure supply system

In order to achieve a hydraulic source with multiple pressure supply lines the following principles were considered. The high-pressure line is directly connected to the pump outlet. A pressure-compensated variable displacement pump is used, in which the outlet pressure is applied to regulate the pump volumetric displacement. When perfectly adjusted, the flow rate provided will match the system needs. The power consumption of the pump is, therefore, minimized whenever possible, which increases the systems energy efficiency [10] [11].

For the medium and low-pressure lines an arrangement of pressure valves and accumulator is assumed: A pressure reducing valve regulates the minimum pressure by diverting fluid from a higher-pressure line, while a relief valve controls the maximum pressure by returning fluid to the reservoir line. This concept results on a pressure hysteresis, allowing small pressure fluctuations, without compromising the force spectrum. The hydraulic diagram of the complete supply system can be visualized in Figure 2.

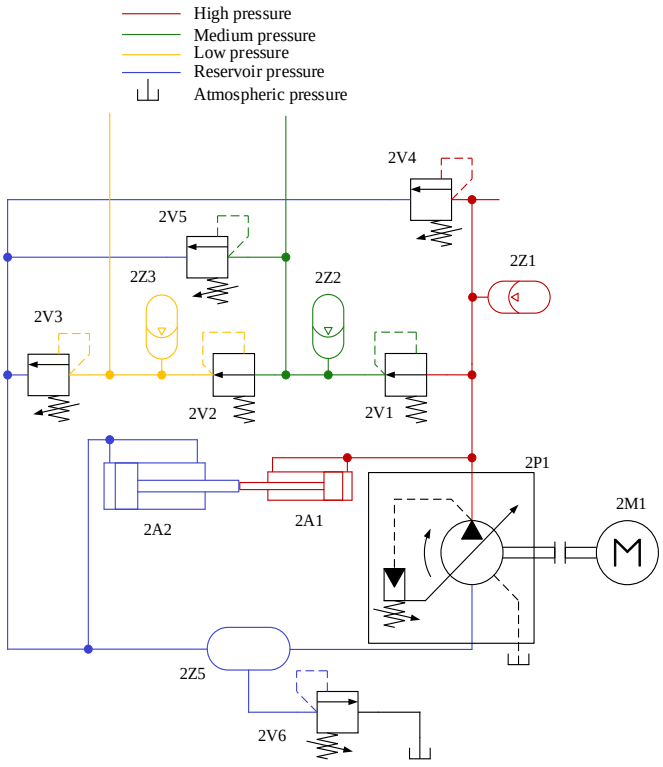


Figure 2: Supply system with pressurized reservoir.

In aviation, the return line is part of a more complex hydraulic system, may exceed hundreds of meters and handles the influence of multiple actuators. Given that, it may be poorly suited to operate as a pressured supply line, since it is susceptible to pressure fluctuations and cavitation. Further research is needed to analyze this use possibility and, if possible, provide the required adaptations.

Accumulators are included in all three pressure lines to attenuate pressure oscillations and store hydraulic energy within each line specified pressure range. In conventional systems, fluid from the actuator retracting chamber is directed to the reservoir making its energy unavailable for further use. This configuration prioritizes energy dissipation rather than storage. Conversely, the proposed approach maximizes energy use by directing any incoming flow (whether from the actuator retracting chamber or the pump) to the highest-pressure accumulator capable of accommodating it within its pressure range before discharges to the reservoir.

Also, if needed, additional pressure lines can be incorporated by applying the same valve scheme present in the medium and low-pressure lines and defining a new range.

2.2 Pressurized reservoir

In aircraft, since the vertical orientation of the systems change, the reservoir line must always be pressurized in order to avoid cavitation in the pump inlet and to ensure fluid supply disregarding position.

It is known that the reservoir must absorb fluid volume fluctuations within the system. For instance, when actuator movement reduces the amount of fluid in the pressure lines, the excess must be stored in the reservoir until the system requires it again. A single accumulator will not meet this function effectively, as its pressure varies greatly with volume changes [12].

The solution designed for this requirement is analogous to a bootstrap reservoir, using two actuators with different resultant areas mechanically coupled by their rods. The high-pressure line connects to the chambers of the smaller area actuator. The chambers of the second actuator are connected to, and thus pressurize, the reservoir line. The pressure ratio between the hydraulic lines is inversely proportional to the resultant area ratio of the two actuators. The arrangement can be observed in Figure 3.

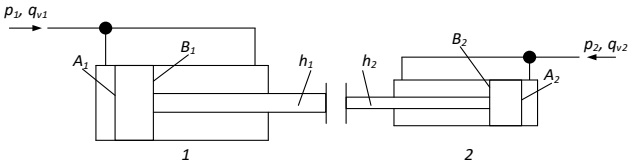


Figure 3: Detailed model of the actuator assembly in the pressurized reservoir.

In the implemented system, an area ten times smaller on the high-pressure side resulted in a pressure ten times smaller on the reservoir side.

Since the actuators are mechanically connected, there is a pressure relation between lines that can be represented as eq. (1), where p_1 is the reservoir pressure, p_2 is the high pressure and k is the area ratio between the larger area and the smaller area:

$$p_1 = \frac{p_2}{k}\,, \tag{1}$$

where

$$k = \frac{A_1 - B_1}{A_2 - B_2} \ . \quad (2)$$

As previously mentioned, one of the key functions of the system is to compensate volume variations, which is achieved through the movement of the piston rods. When fluid enters the return line, the rods consequently shift to the right, accommodating the variation. Conversely, when more fluid is in the pressure lines, the rods move to the left. This dynamic adjustment ensures proper reservoir volume while using the pump's outlet pressure to maintain reservoir pressure.

Additionally, the atmospheric pressure reservoirs shown in Figure 2 are intended for the pump drain and an emergency over-pressurization of the pressurized reservoir. The amount of fluid expected through these directions is, therefore, minimal. Such reservoir could be comprised of a small canister, to be emptied only in scheduled maintenance or following an emergency event.

3 Efficiency

3.1 Energy supplied by the pump

To calculate the system efficiency, the energy delivered to the system by the pump is first defined. The equation for the hydraulic power supplied by the pump is presented in eq. (3):

$$P_h = q_v . \Delta p \ , \quad (3)$$

where q_v [m³/s] is the pump flow rate and Δp [Pa] the pressure differential between the outlet and the inlet. The relationship between energy E [J] and power P [W] is given as shown in eq. (4):

$$P = \frac{dE}{dt} . \quad (4)$$

By substituting eq. (3) into eq. (4) the hydraulic energy E_h [J] supplied by the pump is obtained:

$$E_h = \int_{t_i}^{t_f} q_v . \Delta p \, dt \ . \quad (5)$$

3.2 Energy delivered by the actuator

The hydraulic power delivered by the actuator $P_{h,A}$ [W] is given by eq. (6):

$$P_{h,A} = F_h . v, \quad (6)$$

where F_h [N] is the hydraulic force applied by the actuator and v [m/s] is the velocity. By substituting eq (6) into eq. (4), eq, (7) is obtained, which expresses the hydraulic energy consumed by the actuator $E_{h,A}$ [J]:

$$E_{h,A} = \int_{t_i}^{t_f} F_h . v \, dt \ . \quad (7)$$

3.3 System efficiency

The efficiency η of a hydraulic system is given by the ratio of the energy supplied by the pump to the energy delivered by the actuator, as shown in eq. (8):

$$\eta = \frac{E_{h,A}}{E_h} . \quad (8)$$

4 Simulation

To ensure the proposed system works as expected and evaluate its dynamics and energy aspects, a simulation model of the scheme, shown in Figure 4, was implemented using the Hopsan® software, an open-source multi-domain system simulation tool developed at the division of Fluid and Mechatronic Systems at Linköping University. Also, this practice allowed for a finer tunning and testing of component parameter prior to a future test bench assembly.

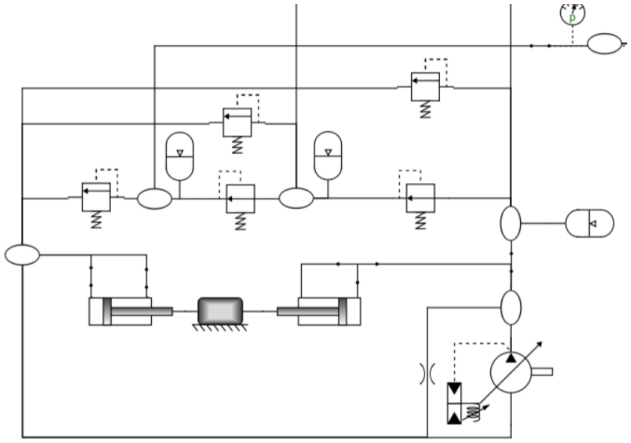


Figure 4: Hopsan simulation model.

4.1 System parameters

As mentioned before, the required pressure ratio (k), in eq. (2), is around 10. Which means that the pressure p_1 is 10 times smaller than p_2. To achieve that, the actuators were selected based on components available in the market. The result was the cylinder 2A1 with 63 mm in diameter and piston rod h_1 with 45 mm in diameter, and the cylinder 2A2 with 32 mm in diameter, with 14 mm in h_2 diameter. Beyond that, the cylinders are implemented as C-type pistons, with a 0.4 m stroke and 7.27 × 10^{-5} m³ of dead volume in all chambers. Default values are assigned for other parameters.

The DHA is simulated based in [6], with the aim of evaluating the proposed supply system performance operating with the digital hydraulic actuator. It presents a multi-chamber actuator with areas A_A 13.44 × 10^{-4} m², A_B 7.07 ×10^{-4} m², A_C 11.02 × 10^{-4} m², and A_D 15.07 × 10^{-4} m². The stroke is 0.2 m. The actuator is operated by the set of 12 on/off valves presented in [6], providing 81 possible force outputs. Following the

approach detailed in the prior publication, the controller employs a cost function to select the valve combination providing the nearest feasible force output relative to the system demand.

Additionally, the system proposed by [6] is supplied by a high-pressure line of 75 bar, a medium-pressure line of 45 bar and 7.5 bar low-pressure line that is also the reservoir line. In the present model, as aforementioned, the reservoir line is separated from the low-pressure line. Thus, a fourth line is added and the low-pressure line is set to 10 bar. The 7.5 bar line simulated is the reservoir one.

To achieve these pressures, as mentioned in Section 2.1, the pressure relief and pressure reducing valves are adjusted to cause pressure hysteresis. The values of 43 to 45 bar are set in the medium line and 8 to 12 bar in the low pressure. The upper limits are set to the reference pressures of the hydraulic pressure relief valves in the simulation. Such components are parametrized with 10mm of spool diameter, 18 kN/m of spring constant, and a flow coefficient of 0.7. The lower pressure limits are set to the reference pressure of the hydraulic pressure reducing valves in Hopsan. A nominal flow rate of 0.001 m^3/s at the reference pressure are assigned to the valves. Other parameters are left as default.

Accumulators are simulated as gas accumulators, all with 0.001 m^3 as volume and a polytropic index of 1.4. Preload pressures of 72 bar, 40 bar, and 8 bar are assigned to the high, medium and low pressure accumulators, respectively.

The high pressure is maintained by the pressure-compensated variable displacement pump, simulated as a Q-type pump with pressure control in Hopsan. A reference pressure difference of 75 bar, a rotational frequency of 1800 rpm, and a time response of 0.15 s from min to full displacement and of 0.12 s in the opposite direction are parametrized to the pump. The minimal flow rate is set to 0 and the nominal flow rate is set to give 23 Lpm (0.00038 m^3/s) at maximum displacement, 13 cm^3/rev, running at 1800 rpm, which is able of maintaining the maximum speed required by the actuator of 0.15 m/s.

Although other pump control strategies can be used, in this paper will only focus on providing and maintaining the constant pressure levels as originally required by the hydraulic system in [6]. In the future, different strategies could be implemented and optimized to further develop the total system's efficiency, performance, or functionality.

4.2 Step input responses

Considering the hydraulic supply system (Figure 2) coupled to the digital hydraulic actuator (Figure 1), results for a sequence of step inputs are discussed. As shown in Figure 5, a reference step input is applied at every five seconds. The reference signal is shown in blue, the actuator position in red and in yellow the response of an equivalent SHA system (dimensioned for the same use case) for comparison.

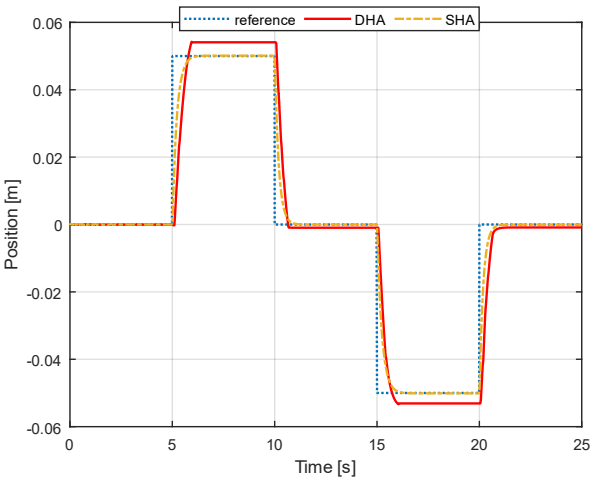


Figure 5: Actuator position for step input.

Figure 6 shows the pressure behavior in each of the supply lines for the same simulation.

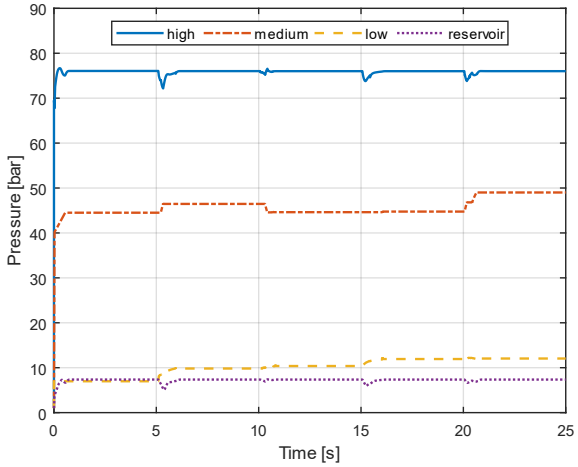


Figure 6: Pressure behavior for step input.

As it can be seen, at five-second intervals, the actuator movement causes a slight pressure drop in the high-pressure line as the actuator demands fluid supply. The fluctuation, however, is less than 5 %, demonstrating the system ability to maintain the defined pressure limits.

Additionally, at 5 and 20 seconds, the medium-pressure line exhibits an increase in pressure value. This fluctuation indicates that the line receive fluid from the actuator, indicating that the flow (and, consequently, the energy) is recovered at the intermediary pressure. A similar behavior occurs in the low-pressure line. The reservoir pressure line follows the high-pressure behavior due to the pressurizing device shown in Figure 3. Figure 7 shows the piston rod displacement in the bootstrap reservoir.

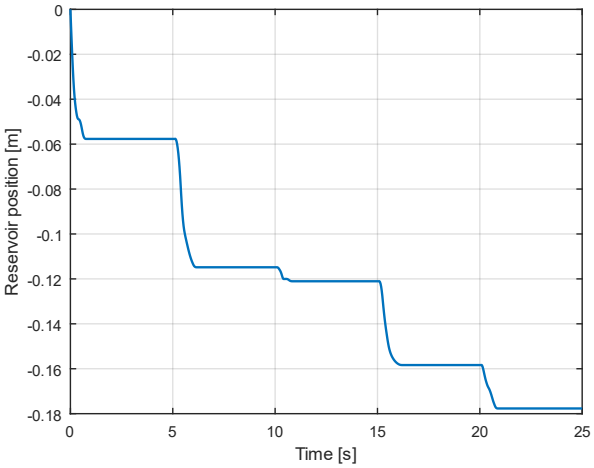


Figure 7: Reservoir position for step input.

It is noteworthy that, as the DHA moves, the reservoir cylinder rod displaces to the left (negative displacement). Therefore, the 2A2 actuator rod occupy part of the internal volume of the actuator, meaning that the reservoir is reducing the fluid volume. Consequently, the pressure lines are increasing their fluid volume. As shown in Figure 6, the medium- and low-pressure lines experience a pressure increasing during the DHA movement and, consequently, store fluid in the supply system. Figure 8 provides a detailed view of the fluid volume stored in the accumulators.

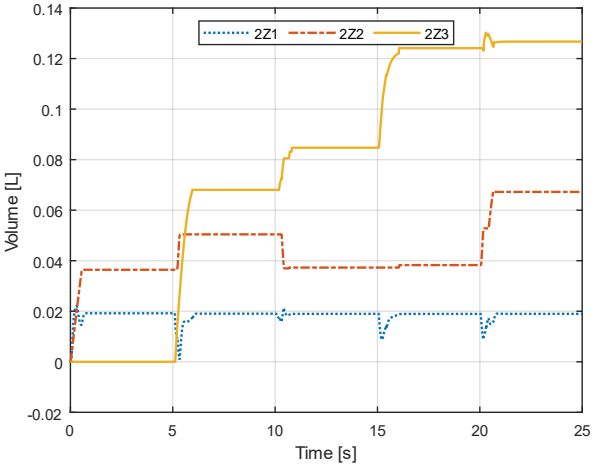


Figure 8: Fluid volume stored in the accumulators.

Figure 9 show the displacement for the sequence of step inputs (parametrization shown in section 4.1). As shown in Figure 9, the pump changes its displacement to provide fluid to the system when the pressure in the high-pressure line reduces, that is, when the DHA moves.

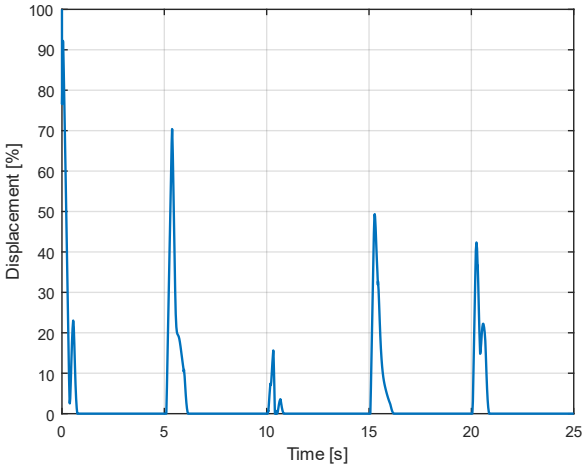


Figure 9: Pump displacement for the step inputs.

The Figure 10 shows the flow rates through the pressure valves. The valve labels are identified in in Figure 2. The valves stay closed practically the whole time except for the relief valve of the low-pressure line, as this line receives enough fluid for its pressure to reach the set pressure and it opens to reduce the pressure.

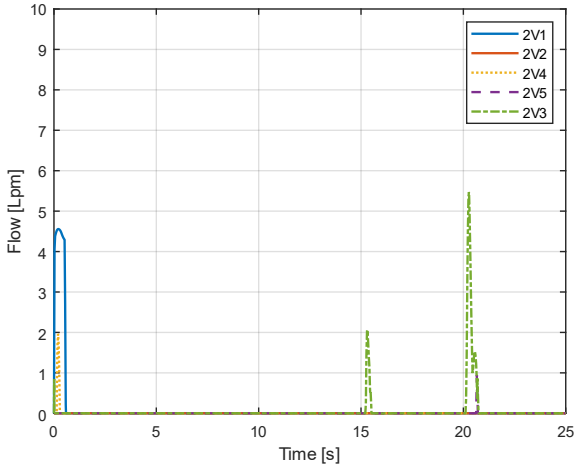


Figure 10: Valve flow for the step input

The variables used for efficiency calculations as shown in Section 3 are extracted from the simulation. In the simulated scenery, the pump provided 2295 J of energy to the system. The DHA consumed around 628 J, which results in around 27% of efficiency. Figure 11 shows the energy supplied by the pump and consumed by the actuator.

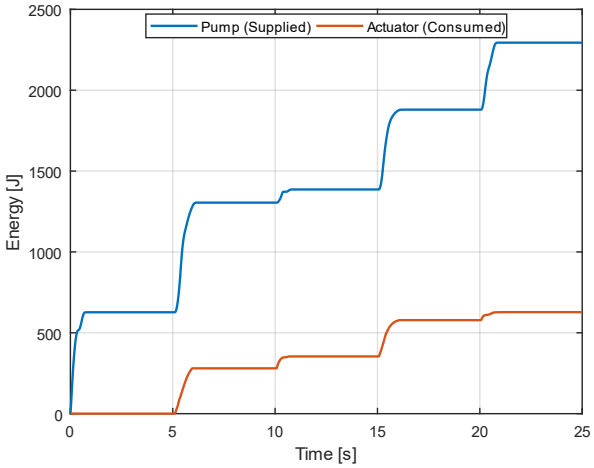


Figure 11: Energy Supplied by the Pump and Consumed by the Actuator for step input.

4.3 Sinusoidal input responses

Results for a sinusoidal input for the actuator positioning system are discussed in this section. The sine input has a frequency of 0.2 Hz and amplitude of 0.09 m. Figure 12 shows, in blue, the reference signal, in red, the actuator position and in yellow the response of an equivalent SHA system for comparison:

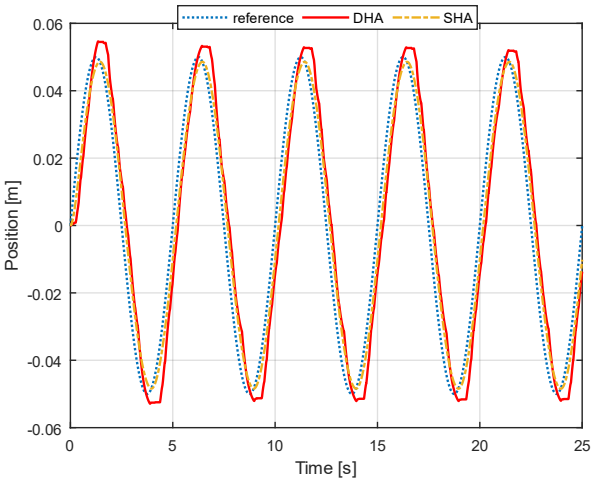


Figure 12: Actuator position for sinusoidal input.

The proposed supply system does not affect the system ability to follow the reference, as the simulation result is similar as the ones shown in [7].

Figure 13 shows the pressure behavior for the sinusoidal input. As can be seen, the pressure varies in all four pressure lines throughout the simulation. This behavior is expected, as the actuator continuous movement at different velocities enables the system to optimize the fluid utilization. To enhance efficiency, the system dynamically switches the pressure lines connected to each actuator chamber and, consequently, the lines constantly receive and supply fluid.

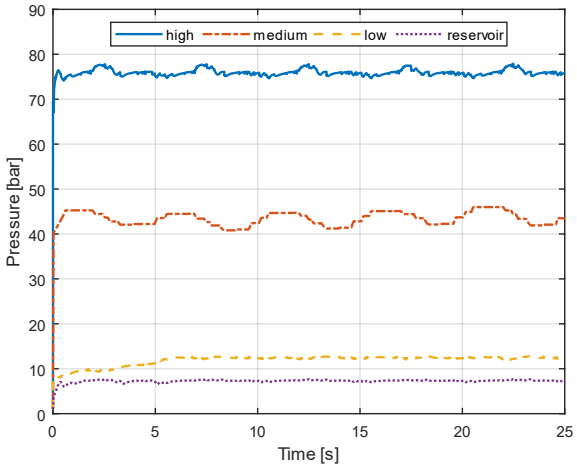


Figure 13: Pressure behavior for sinusoidal input.

Figure 14 shows the position of the bootstrap reservoir cylinder. Similar to what was observed in Figure 7, Figure 14 shows that, as the system pressurizes the lines, the reservoir is displaced to the left, reducing its fluid volume and supplying it to the system. However, it can be noted that after 5 seconds, the system begins to reach a steady-state condition. As seen in Figure 13, the pressures oscillate within a defined range. Consequently, the fluid volume stored in the accumulators does not change considerably, instead of that, it fluctuates along with the pressure oscillations. Figure 15 provides a detailed view of this behavior.

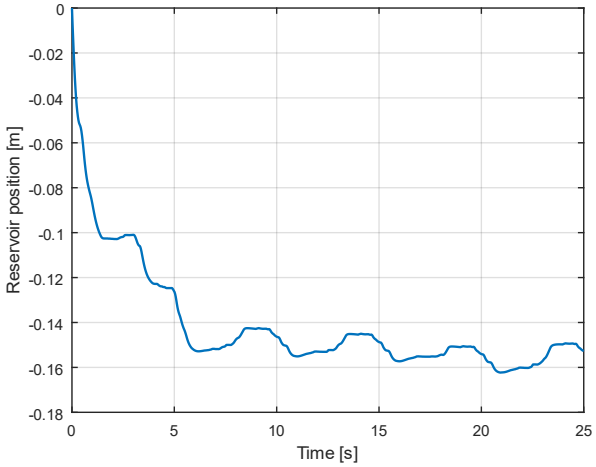


Figure 14: Reservoir position for sinusoidal input.

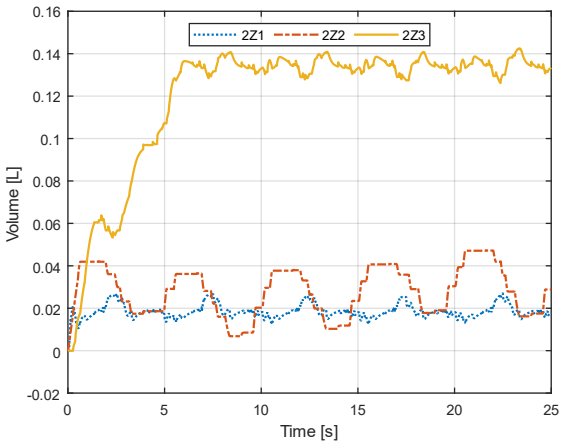


Figure 15: Fluid volume stored in the accumulators for sinusoidal input.

Figure 16 shows the pump displacement for the sinusoidal input.

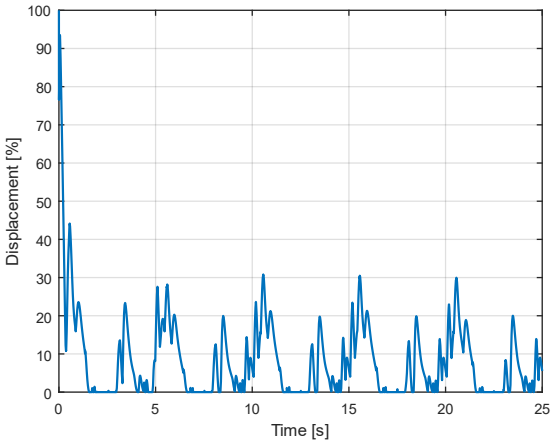


Figure 16: Pump displacement for the sinusoidal input.

As can be seen in Figure 16, the pump is constantly changing its displacement in response to the DHA movement. Since the pressure in the high-pressure line is constantly changing, the pump is constantly compensating the pressure drops.

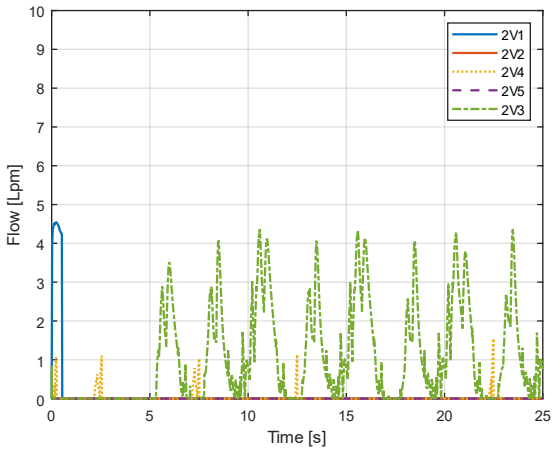


Figure 17: Valve flow for the sinusoidal input.

As it is shown in Figure 17, in the sinusoidal input the valves also stay closed the whole time, except for the relief valve of the low-pressure line. As the DHA is constantly discharging fluid to this line, its pressure increases due to the increase volume in the accumulator, so the relief valve is constantly opening to maintain the pressure in this line.

As previously discussed, the system was designed with pressure hysteresis in each line to facilitate fluid to be moving between the pressure lines rather than discharging to the reservoir. For the sine input the pump provided 5280 J and the DHA consumed around 3100 J. In this case, the system achieved an energy efficiency of approximately 59%. The energy supplied by the pump and consumed by the actuator is shown in Figure 18:

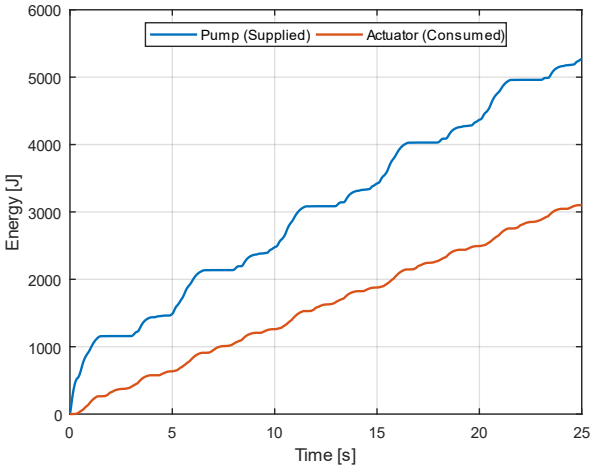


Figure 18: Energy supplied by the pump and consumed by the actuator.

The difference in energy efficiency from the previous case can be explained by the fact that in the sine input, the actuator is constantly moving, thus, the system is constantly using the energy that was provided.

5 Conclusion

Based on the simulation results, the proposed multiple pressure supply system successfully achieves its intended purpose of supplying hydraulic power to the DHA without compromising its performance and maintaining smooth operation. This facilitates the application of digital hydraulics with reliable power supplies, especially with mobile and aeronautical applications as the system is simple and lean.

Additionally, the system kept the reservoir line pressurized and avoided cavitation at all times. This prevents potential damage to hydraulic components, disregarding the orientation of the vehicle, and ensures fluid supply whenever the pump's displacement needed is higher.

Beyond that, the system also maintained the pressure within the designated range for all pressure lines. This validates its ability to handle pressure fluctuations and ensure fluid distribution when needed.

The overall efficiency of the system was significantly higher than conventional hydraulics applied in aviation (as seen in [5], that can be as low as 2%) in all cases simulated, achieving up to 59% as seen in section 4.3. This confirms the efficiency benefits shown by DHA when the supply scheme is also considered, and infatuates the promising alternative that digital hydraulic is for flight control actuation.

The DHA concept and its Hopsan model have been experimentally validated in prior work [6]. The development of this novel supply system proposed in this study is scheduled for experimental validation, with a dedicated test bench currently under assembly to evaluate its functionality and efficiency under real-world operating conditions.

Acknowledgement

This work was carried out with the support of CAPES - Coordination of Improvement of Higher-Level Personnel, CNPq - National Council for Scientific and Technological, FAPESC - Foundation for the Support of the Scientific and Technological Research of Santa Catarina State, and FINEP - Project Multimode Actuating Systems (MMAS) n.01.22.0478.00.

References

[1] OECD. Transport Energy and CO2: Moving Towards Sustainability. OECD/IEA Publishing, 2009

[2] International Civil Aviation Organization (ICAO). *2022 Environmental Report: Innovation for a Green Transition*. ICAO, 2022.

[3] J C Maré. Aerospace Actuators 1: Needs, Reliability and Hydraulic Power Solutions. John Wiley & Sons, Inc., London, UK, 2016.

[4] J C Maré. Aerospace Actuators 2: Signal-by-Wire and Power-by-Wire. John Wiley & Sons, Inc., London, UK, 2017.

[5] R S Lopes Jr, D O e Silva, M P Nostrani, A Dell'Amico, P Krus, and V J De Negri. A Comparative Analysis of Innovative Digital Hydraulic Actuators for Primary Flight Control. Proc. of the 33rd Congress of the International Council of the Aeronautical Sciences, Stockholm, pp. 1-18, 2022.

[6] H C Belan. Development of digital hydraulic actuators for controlling aircraft control surfaces [Sistemas de atuação hidráulicos digitais para aviões com foco em eficiência energética]. Thesis, Federal University of Santa Catarina, 2018. Available at: https://laship.ufsc.br/site/wp-content/uploads/2018/07/Tese_Belan_2018.pdf.

[7] H C Belan, B Lantto, P Krus, and V J De Negri. Digital Hydraulic Actuator (DHA) Concept for Aircraft Actuation Systems. Proc. of Recent Advances in Aerospace Actuation Systems and Components, Toulouse, 2016.

[8] M P Nostrani, H Raduenz, A Dell'Amico, P Krus, and V J De Negri. Multi-Chamber Actuator Using Digital Pump for Position and Velocity Control Applied in Aircraft. International Journal of Fluid Power, pp. 1-28, 2023.

[9] D O e Silva, M P Nostrani, R S Lopes Jr, G Waltrich, P Krus, and V J De Negri. Digital Electro Hydraulic Actuator with Variable Speed Digital Hydraulic Pump: A Design Overview. Proc. of Global Fluid Power Society Symposium, Napoli, 2022.

[10] D Kolář, A Bureček, L Hružík, M Ledvoň, T Polášek, J Jablonská, and R Lenhard. Static Characteristics and Energy Consumption of the Pressure-Compensated Pump. Processes, 12:1081, 2024. https://doi.org/10.3390/pr12061081.

[11] M Axin, B Eriksson, and P Krus. Flow versus pressure control of pumps in mobile hydraulic systems. Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering, 228(4):245-256, 2014. doi:10.1177/0959651813512820

[12] J D Van de Ven. Constant pressure hydraulic energy storage through a variable area piston hydraulic accumulator. Applied Energy, 105:262-270, 2013. https://doi.org/10.1016/j.apenergy.2012.12.059

[13] R S Lopes Jr. Evaluation of the dynamic performance of a digital hydraulic actuator for aeronautical applications under fault conditions [Avaliação do desempenho dinâmico de um atuador hidráulico digital para aplicações aeronáuticas em condição de falha]. Dissertation, Federal University of Santa Catarina, 2021. Available at: https://laship.ufsc.br/site/wp-content/uploads/2023/03/Dissertacao_Lopes_2021.pdf

[14] L Tao. Study on Bootstrap Reservoir Type Pressurized System for Civil Aircraft Hydraulic Supply System. Proc. of the 2016 IEEE/CSAA International Conference on Aircraft Utility Systems, Beijing, China, 2016. IEEE.

A Pump Decoupled Architecture to Allow Increasing Energy Efficiency of Hydrostatic Transmission Solution based on Fixed Displacement Secondary Units

Prithvi N. Chandiramani, Andrea Vacca, Ryan Jenkins*

School of Mechanical Engineering, Purdue University, West Lafayette, IN, USA
E-mail: pchandir@purdue.edu, avacca@purdue.edu, ryan.jenkins@cnh.com
*R&D Hydraulics Competence Center, CNH Industrial, Oak Brook, IL, USA

Abstract

Despite being inherently throttle less system, Hydrostatic Transmission (HT) for off-highway applications suffers from low efficiency when the primary and secondary units do not operate at their optimal efficiency point. This paper proposes a novel HT architecture that decouples the primary unit's operating point from the load (i.e., external torque and rotational speed) allowing it to always operate at a more favorable point, while maintaining roughly the same energy efficiency for the secondary unit, resulting in an overall increase of the total HT efficiency. Essentially, the decoupling is achieved by integrating a hydraulic accumulator into the supply line and controlling the primary unit to regulate supply pressure within a narrow range rather than delivering a reference flow. The supply line is connected to a fixed displacement secondary unit, with its angular velocity controlled by another hydraulic rotary actuator connected to the same shaft. The paper details the architecture design, the power flow management and the architecture's simple and scalable implementation for a range of applications.

To assess the energy efficiency advantage of the proposed solution, the reference case of the basecutter system in a sugarcane harvester is considered. A lumped-parameter model for both the baseline commercial solution, validated based on available experimental data, and for the proposed system were developed. Results for a realistic drive-cycle show a 15% reduction in energy consumption compared to the baseline flow-supply series HT used in the reference machine, thus proving the effectiveness of the proposed system.

Keywords: Hydrostatic Transmission, Off-Road Machines, Pump-Load Decoupling, Power Management

1 Introduction

A hydrostatic transmission (HT) is a throttle-less hydraulic control system for driving rotary actuators. Unlike conventional mechanical transmissions, it provides a continuously variable transmission (CVT) ratio, making it ideal for high-load applications. HTs are widely used to power the drive and work functions of off-highway machinery in the earthmoving and agricultural industries. However, HTs are also characterized by relatively low efficiency compared to mechanical or electric transmissions predominantly used in on-road vehicles, leading to higher fuel consumption and emissions in machines that utilize them. Rising fuel costs and growing environmental concerns in recent years highlight the need for more efficient HT architectures for off-highway machines.

Conventionally, HTs are classified into two configurations: Flow Supply Primary Control (FSPC) and Pressure Supply Secondary Control (PSSC), as illustrated with simplified schematics in Figure 1. In the FSPC configuration, the fractional displacement of the pump (ϵ_P) determines the flow going to the fixed displacement motor, essentially controlling the motor' shaft speed (n_M). Also, the supply line pressure (p_H) is governed by the external torque (T_{Ext}) applied on the motor's shaft. The design simplicity of this configuration enables ease of implementation and high reliability, making it a common choice in off-road machinery. However, it exhibits slow dynamic response due to oscillations caused by the high compliance of the connecting hose and its expansion under pressure variations [1]. Additionally, the direct flow coupling between the pump and motor restricts its application to single-drive function.

In the PSSC configuration, the supply line pressure can is kept constant. Thus, fractional displacement of the motor (ϵ_M) determines the torque generated by motor (T_M) and is controlled to get a desired motor's shaft speed using a feedback loop. The fractional displacement of the pump (ϵ_P) is controlled to maintain the supply line at a fixed pressure using hydro-mechanical feedback. To stabilize the supply line pressure during transients, when pump response may be insufficient, an accumulator is often integrated into the supply line to increase its capacitance. This configuration enables fast dynamic response with minimal oscillations. Additionally, the absence of direct flow coupling between the pump and motor allows a single pump to drive multiple actuators connected to a common supply line. However, the increased system complexity due to additional components and advanced control requirements limits its application to niche domains demanding precise control, such as large winches, marine cranes, and offshore heave compensation systems [2].

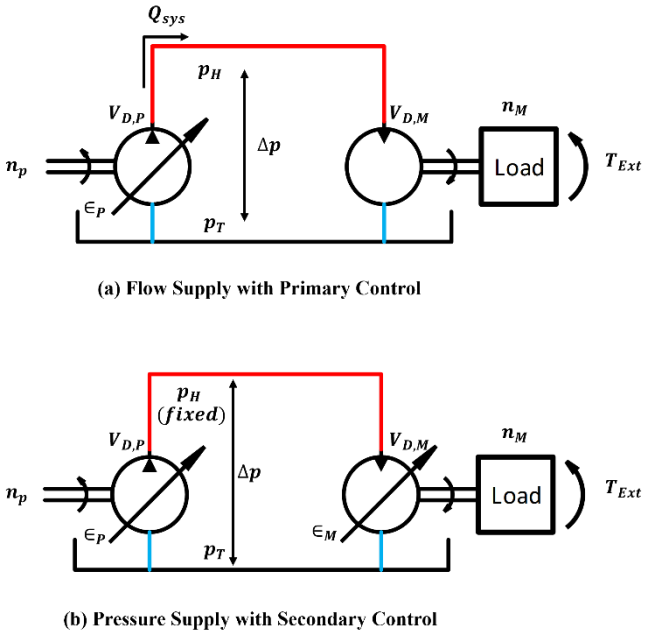


Figure 1: Conventional HT configurations.

The steady-state transmission efficiency of an HT is largely affected by the energy losses in the primary and secondary unit used [3]. These units attain peak efficiency within a narrow operating range, typically near full displacement of the hydraulic unit and rated pressure. However, in both configurations, the operating conditions of the pump and motor are dictated by the HT's operating condition, i.e., motor's shaft speed and external torque. As a result, the wide operating range typical of off-highway machines often forces these units to operate under suboptimal conditions, such as partial displacements and lower pressure differentials, which can significantly reduce HT transmission efficiency, in some cases to below 40% [4]. Therefore, HT efficiency is highly dependent on machine operating conditions. Additional factors impacting efficiency include pressure drop across directional control valve (in open-loop HTs), flushing and low-pressure charge systems (in closed-loop HTs) and the inability to regenerate and reuse overrunning energy.

To improve the transmission efficiency over the traditional HT, it is necessary to reduce the energy losses of the positive displacement units. Literature mentions three main strategies to achieve this, which are outlined in the following subsections: -

1.1 Digital Displacement Hydrostatic Transmission

A potential approach in this context is to implement HT with pumps and motors based on the digital displacement principle. The digital displacement pump (DDP) and digital displacement pump/motor (DDPM) are radial piston machines that regulate flow by individually connecting or disconnecting each piston to the high- and low-pressure sides using actively controlled high-speed solenoid valves [5]. This mechanism significantly enhances overall efficiency at partial displacements, compared to conventional variable displacement units. Swapping conventional pumps with digital displacement pumps (DDPs) in the 16-tonne DEXTER excavator resulted in 19% fuel savings during a typical loading cycle, with 8% attributed to reduced pump losses, without compromising productivity [6]. Implementing a similar approach for HTs, incorporating both DDP and/or DDPM, is expected to have similar fuel savings. Additionally, each piston in a DDP, known as a pumplet, can be accessed and controlled like an individual variable displacement unit. This capability enables the DDP to function as an "elastic pump" system [7], facilitating multi-actuator HT implementations with reduced total pump capacity. Despite the fact that DDP machines are not yet commercially available, there is an opportunity for solutions able to increase HT efficiency by utilizing cost-effective and consolidated hydraulic components.

1.2 Hydro-Mechanical Power-Split Transmission

A hydro-mechanical power-split approach for increasing HT's efficiency is to combine the HT with a direct mechanical transmission within a single transmission system, utilizing a planetary gear train (PGT) [8,9]. This allows routing majority of power through the highly efficient mechanical path while keeping the CVT feature of transmission intact. Macor and Rosetti achieved 84% mean overall efficiency [10] by optimizing a hydro-mechanical power-split transmission used for agricultural tractor propulsion application.

Nevertheless, the use of bulky PGT necessitates for a sperate PGT for every actuator, which can be impractical for implementation with multiple actuators. This is why this architecture is mostly limited to propulsion powertrains.

1.3 Design Modification and Optimal Sizing of HT

Another approach to enhance efficiency of HTs is to allow the hydraulic units to operate in more efficient regimes, i.e., higher displacements and pressure differential, thus reducing energy losses associated with them. Conventional HT configurations, which have a single control degree of freedom, lack the flexibility to optimize operating conditions of these components. Guo and Vacca proposed a split-pump design with a dual displacement motor as an alternative flow-supply HT configuration, along with genetic algorithm based optimal

sizing methods [4]. This design introduces additional control degrees of freedom to enhance operating conditions of the components at the most frequent operating condition, thereby improving transmission efficiency. Simulation results for a heavy-duty harvester demonstrated 14% improvement in efficiency and 20% reduction in power consumption, compared to a reference commercial system.

This approach does not require expensive components for implementation, making it a cost-effective solution suitable for off-highway machines. However, the operating conditions of the hydraulic units remain coupled to the HT's overall operation, limiting efficiency gains in drive cycles with multiple frequently occurring operating conditions. Moreover, adding control degrees of freedom often requires replacing fixed displacement secondary units with variable or dual displacement units, which is impractical for applications utilizing series hydrostatic drives.

While the above-mentioned approaches improve transmission efficiency over conventional HT configurations their adoption is constrained by inherent limitations and implementation challenges. This paper proposes an alternative HT implementation aimed at improving transmission efficiency while ensuring broad applicability in off-highway applications. The architecture decouples the operating conditions of the hydraulic units from that of the HT, allowing them to operate at optimal efficiency at all times and enhancing transmission efficiency across various drive cycles. Additionally, it utilizes fixed displacement secondary units and commercially available hydraulic components, facilitating simpler and more adaptable implementation across a wide range of applications.

The paper is structured as follows: Section 2 outlines the proposed Pump Decoupled Hydrostatic Transmission (PDHT) architecture and its underlying motivation. Section 3 presents a case study comparing its performance against a conventional flow-supply series HT system. Finally, Section 4 provides concluding remarks and scope for future work.

2 Pump-Decoupled Hydrostatic Transmission (PDHT) Architecture

2.1 Description of Proposed Architecture

The objective is to develop an HT architecture using a variable displacement primary unit and a fixed displacement secondary unit, where the primary unit consistently operates at maximum displacement ($\epsilon_{P,Max}$) during power delivery, and a constant pressure differential (Δp_c) across both units is maintained regardless of the HT's operating conditions (i.e., motor's shaft speed and external torque). This would enable appropriately selected units (i.e., those with Δp_c within optimal pressure range) to operate at peak efficiency, therefore minimizing energy losses and maximizing overall transmission efficiency.

Such a specific decoupling of the positive displacement units' operating conditions from those of the HT would classify the architecture as a pressure-supply HT with two key requirements:

1. The pump must regulate the supply line pressure within a specific (narrow) range while operating at maximum displacement ($\epsilon_{P,Max}$) or zero displacement.
2. The fixed displacement motor's shaft speed must be controlled in the pressure-supply HT.

To accomplish these requirements, the open-circuit PDHT architecture illustrated in Figure 2 is proposed. A variable displacement pump regulates the supply line pressure, which is governed by equations 1-3. To maintain the line pressure within a narrow optimal range [p_{Min}, p_{Max}], the pump fractional displacement (ϵ_P) alternates between zero and maximum, following the pressure control logic shown in Figure 3. The accumulator integrated into the supply line to reduce its hydraulic stiffness is sized adequately to prevent rapid stroking and de-stroking of the pump, which could otherwise render the proposed architecture infeasible. A fixed displacement motor is connected to the supply line, with its shaft speed controlled by an independently controlled hydraulic rotary actuator connected on the same shaft. This rotary actuator is referred to as speed-control unit (SCU). The motor's shaft speed is governed by eq. 4-5.

$$\frac{dp_H}{dt} = \left(\frac{p_0^{\frac{1}{\gamma}} \cdot V_o}{\gamma \cdot p_H^{\left(\frac{\gamma+1}{\gamma}\right)}} + \frac{V_{line}}{k_{oil}} \right)^{-1} \cdot (Q_P - Q_M) \quad (1)$$

$$Q_P = n_P \cdot V_{D,P} \cdot \epsilon_P \cdot \eta_{v,P} \quad (2)$$

$$Q_M = \frac{n_M \cdot V_{D,M}}{\eta_{v,M}} \quad (3)$$

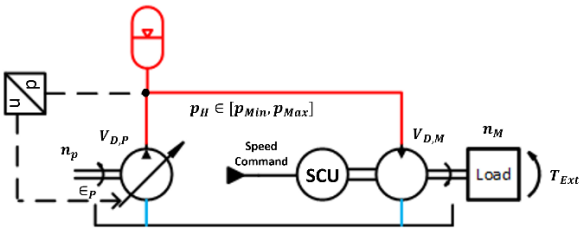


Figure 2. Generic layout of a PDHT architecture

$$T_{SCU} + T_M - T_{Ext} = J \frac{dn_M}{dt} \quad (4)$$

$$T_M = (p_H - p_T) \cdot V_{D,M} \cdot \eta_{HM,M} \quad (5)$$

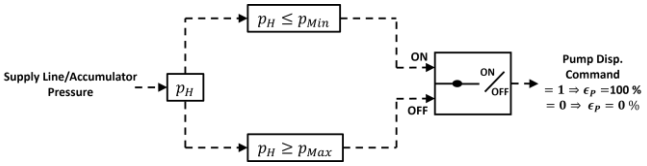


Figure 3. PDHT supply pressure control logic.

The strategy of alternating the pump displacement between zero and maximum, to regulate the supply line (or accumulator) pressure within an optimal range, as illustrated in Figure 4, is similar to that used in constant pressure systems [11-13]. This approach ensures that the pump operates at

maximum efficiency whenever supplying hydraulic power to the system, and only incurs parasitic losses when not delivering any hydraulic power. Alternatively, a fixed displacement unit with an accumulator unloading valve can be used. However, it presents certain limitations, such as throttling losses through the unloading valve during unloading and reduced control flexibility for operating in a flow-supply mode if required, which is demonstrated as beneficial in the case study presented later in the paper.

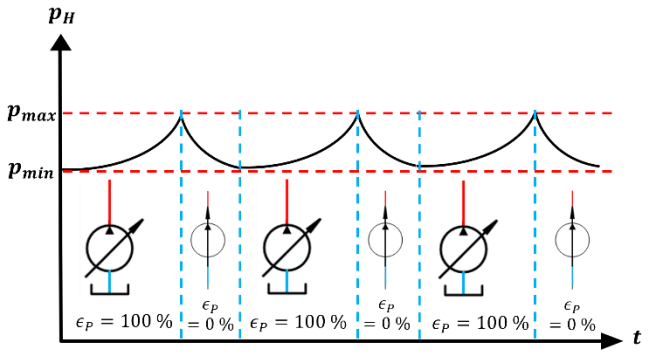


Figure 4. Pump supply pressure regulation by alternating between max and zero displacement.

The fixed displacement motor operates at its peak efficiency, converting hydraulic power from the supply line into mechanical power (P_{DHT}) by delivering torque (T_M) to the rotary function, as defined by eq. (6). As the pressure differential across the motor varies within a narrow range, the generated torque also remains within a narrow band, as defined by eq. (5). Additionally, the SCU provides the remaining torque (T_{SCU}) required for shaft speed control, contributing mechanical power (P_{SCU}), as given by eq. (7). Therefore, the PDHT architecture features two distinct power transmission paths from the input power source to the rotary function: a highly efficient decoupled HT path and a speed-control path, whose efficiency depends on the operating conditions, as shown in Figure 5. Consequently, the transmission efficiency of the PDHT architecture (η_{PDHT}) is described by eq. (8).

$$P_{DHT} = T_M \cdot n_M \tag{6}$$

$$P_{SCU} = T_{SCU} \cdot n_M \tag{7}$$

$$\eta_{PDHT} = \frac{Output\ Power}{Net\ Input\ Power} = \frac{P_{DHT} + P_{SCU}}{\left(\frac{P_{DHT}}{\eta_{DHT}}\right) + \left(\frac{P_{SCU}}{\eta_{SCU}}\right)} \tag{8}$$

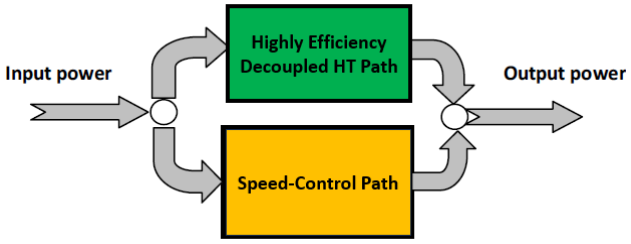


Figure 5.Two paths for power transmission in PDHT.

Eq. 6-8 indicate that high transmission efficiency of the PDHT architecture can be achieved if either of the following conditions is met:

1. Majority of power delivered to the rotary function should be transmitted through the decoupled HT path, i.e., $P_{DHT} \gg P_{SCU}$ (or $T_M \gg T_{SCU}$).

or

2. Efficiency of speed-control path (η_{SCU}) is high when power transmitted through speed-control path (P_{SCU}) is substantial, or T_{SCU} is substantial.

Therefore, a higher supply line pressure range $[p_{Min}, p_{Max}]$ is desirable for improved transmission efficiency, as it would increase the proportion of the total power transmitted through the highly efficient decoupled HT path.

However, the selection of the SCU in this architecture may impose a limit on the maximum allowable supply line pressure if the SCU is unable to regulate shaft speed under negative torque conditions. The steady-state relationship between the torque generated by the fixed displacement motor (T_M), the speed-control unit (T_{SCU}), and the external torque on the motor shaft (T_{Ext}) is defined by eq. (9):

$$T_{SCU} = T_{Ext} - T_M \tag{9}$$

If the chosen SCU can't handle negative torque, the maximum torque ($T_{M,Max}$) that can be supplied by the fixed displacement motor is constraint by eq. (10), also graphically illustrated in Figure 6. The corresponding maximum allowable supply line pressure ($p_{Max,Allowed}$) in the supply line is related to the maximum torque, as described by eq. (11).

$$T_{SCU} = T_{Ext,Min} - T_{M,Max} = 0 \tag{10}$$

$$T_{M,Max} = \left(p_{Max,Allowed} - p_T\right) \cdot V_{D,M} \tag{11}$$

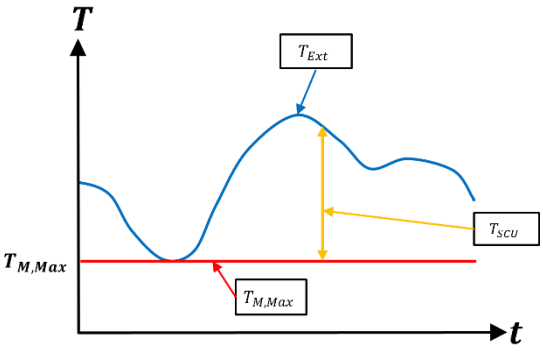


Figure 6. Max. allowable fixed disp. motor torque for cases where SCU can't handle negative torque.

Hence, selecting an appropriate SCU is crucial for maximizing the transmission efficiency of the PDHT architecture, as it may constrain the power transmitted through the efficient decoupled HT path. Additionally, an SCU that results in low speed-control path efficiency when delivering substantial power would introduce significant losses, thereby adversely affecting PDHT architecture's efficiency. Therefore, to achieve maximum transmission efficiency across all operating conditions, the desired characteristics/technology for the SCU are:

1. It should be capable of regulating the motor's shaft speed under negative torque conditions.

2. It should maintain high efficiency of the speed-control path (η_{SCU}) while delivering substantial power (P_{SCU}) or torque (T_{SCU}).

2.2 PDHT Architecture using a Secondary-Controlled Motor as SCU

To maximize PDHT transmission efficiency across various operating conditions, different SCU alternatives can be considered based on the requirements outlined in the previous section. This paper proposes a PDHT architecture variant incorporating a secondary-controlled motor with over-center capability as the SCU, as shown in Figure 7. The inlet of the secondary controlled motor is connected to the supply line with the accumulator, where pressure is regulated within a narrow optimal range [p_{Min}, p_{Max}] by the decoupled pump, ensuring a constant pressure differential across the secondary controlled motor. Therefore, the torque generated by the SCU (or the secondary controlled motor) is described by eq. (12). The motor's shaft speed is controlled by varying its fractional displacement (ϵ_{SM}) using a feedback loop.

$$T_{SCU} = (p_H - p_T) \cdot \epsilon_{SM} \cdot V_{D,SM} \cdot \eta_{HM,SM} \quad (12)$$

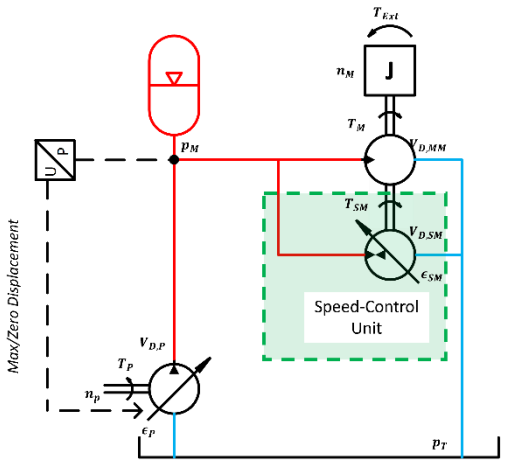


Figure 7. PDHT using a Secondary-Controlled Motor as an SCU

In this variant, power transmission through the speed-control path occurs in two stages: mechanical-to-hydraulic and hydraulic-to-mechanical conversion. The decoupled pump, operating at high efficiency, performs the mechanical-to-hydraulic conversion, while the secondary controlled motor, operating at varying frequency, handles the hydraulic-to-mechanical conversion. At high torque output, the secondary-controlled motor operates at high fractional displacement (ϵ_{SM}), ensuring high efficiency. Consequently, the speed-control path maintains high efficiency when the motor generates high torque.

A secondary-controlled motor with over-center capability can regulate shaft speed under negative torque conditions, allowing for a higher supply line pressure. This increases the proportion of power transmitted through the decoupled HT path. Additionally, a higher supply line pressure enables the use of a smaller displacement speed-control motor and, consequently, a smaller displacement decoupled pump while achieving the same combined maximum torque of the fixed displacement and secondary-controlled motors. During negative torque conditions, the secondary-controlled motor pumps flow into the accumulator, facilitating efficient regeneration of excess energy transmitted through the decoupled HT path, as illustrated in Figure 8.

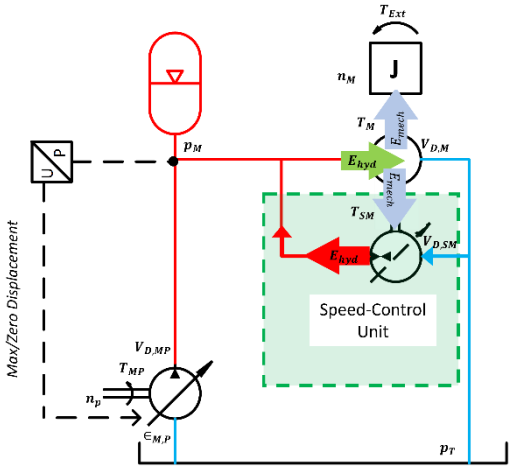


Figure 8. Regeneration of excess mechanical energy delivered by decoupled HT line.

2.3 Multi-Actuator Implementation with PDHT

The proposed variant of PDHT architecture incorporates an additional variable displacement motor per actuator compared to conventional HT. Furthermore, both the fixed displacement motor and the secondary-controlled motor receive power from a common supply line. This design enables straightforward implementation in systems with multiple actuators, as shown in Figure 9. Multiple actuators can be powered through a common pressure rail or a set of pressure rails, allowing them to share a set of pumps and accumulators. This centralized power supply reduces the number of pumps required, thereby lowering overall system costs.

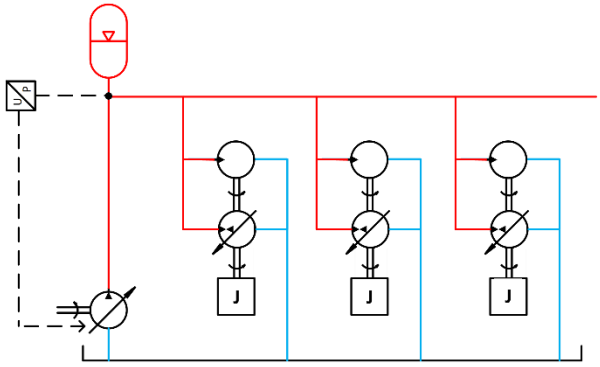


Figure 9. Multi-actuator implementation with PDHT.

Furthermore, the PDHT architecture's use of a fixed displacement motor makes it compatible with series hydrostatic drives. A set of motors connected in series can be directly connected to the supply line, whose pressure is regulated by the decoupled pump, with the SCU regulating the speed of the first motor in the series. Since the first motor's speed is controlled by the SCU, the subsequent motors receive the required flow, operating as they would in a conventional series configuration. Additionally, this implementation

requires only an additional variable displacement motor and accumulator for the entire set of motors, as described in Figure 10.

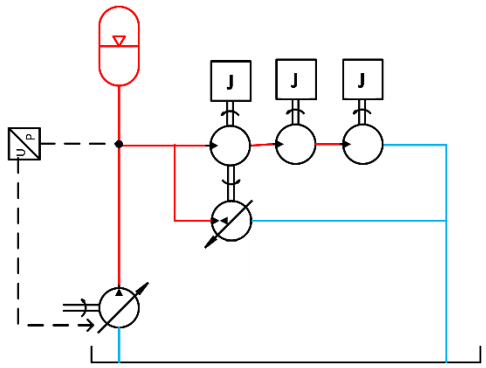


Figure 10. PDHT implementation for a series hydrostatic drive.

3 Case Study

To demonstrate the PDHT architecture's potential for improved efficiency with simpler implementation, it is compared to a conventional flow-supply primary control series HT system for a sugarcane harvester application. The comparison is based on simulations for a reference cycle, with measurements provided by an OEM.

3.1 Reference Machine/System/Cycle

The reference machine used in this study is a Case IH Austoft 9900 sugarcane harvester, with a focus on its basecutter system, a primary energy-consuming subsystem for the whole vehicle. The baseline version of this subsystem is an open-circuit flow-supply with primary control HT featuring multiple series of motors, as illustrated in simplified hydraulic schematic in Figure 11 . In this open-circuit layout, flow direction is controlled by the basecutter valve, which functions similarly to a directional control valve while maintaining roughly a constant pressure drop of 3 bar across a range of flow rates.

The system primarily operates in the forward mode, i.e., where the pump supplies flow to the basecutter motor. This flow is then directed through the motor, divided into subsequent branches, and delivered to the following motors. The speed of the motors is exclusively controlled in this mode. The system switches to reverse mode, i.e., reverse the direction of rotation of motors, by connecting the basecutter motor to tank using the DCV.

Measurement data for a 140-seconds drive cycle at a particular pump shaft speed (n_{Pump}) was provided, including pressures in each line, pump suction pressure, and basecutter motor shaft speed. To compare the baseline system with the PDHT architecture implementation, the external torque on the motor shaft was estimated from the measurement data and used as inputs for simulation model for both systems. The drive cycle defines the operation of the motors at a desired speed against the external torque on the motor shaft. As the system is a series hydrostatic drive with speed-synchronized motors, specifying the desired speed of one motor is sufficient. For the reference drive cycle, the desired basecutter motor shaft speed ($n_{Basecutter,desired}$) was provided.

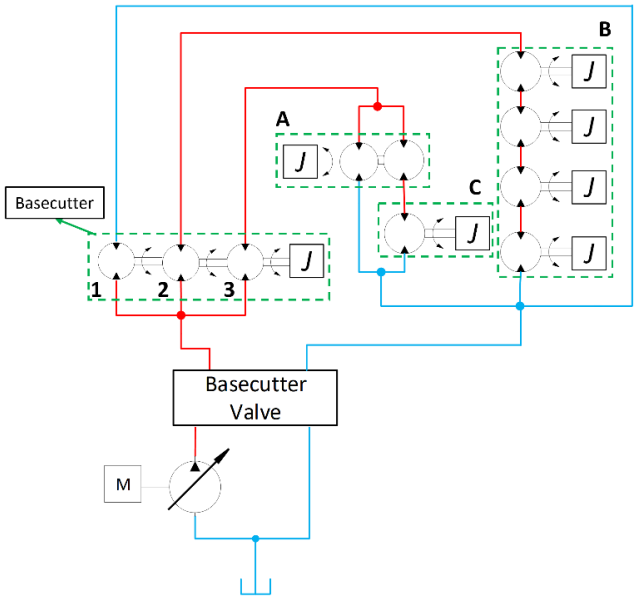


Figure 11. Simplified schematic of basecutter system.

3.2 PDHT Architecture Implementation for Basecutter System

The PDHT architecture is implemented on the basecutter system, as shown in Figure 12. To minimize modifications to the baseline system, the original layout is retained along with the components. The supply line connects to the basecutter motor inlet, while the SCU (i.e., secondary controlled motor) is coupled to the motor's shaft.

In the baseline system, the largest gear section (section-1) in the basecutter motor accounts for half of its total displacement and is incorporated solely to generate additional torque without increasing the motor's inlet pressure or pressure differential. In the PDHT architecture, however, the motor inlet pressure is regulated, and the SCU supplies the additional torque required to move against any load. Therefore, in the PDHT implementation, this gear section is replaced with an SCU of the same displacement. This modification preserves the total motor and pump displacement from the baseline system with minimal changes.

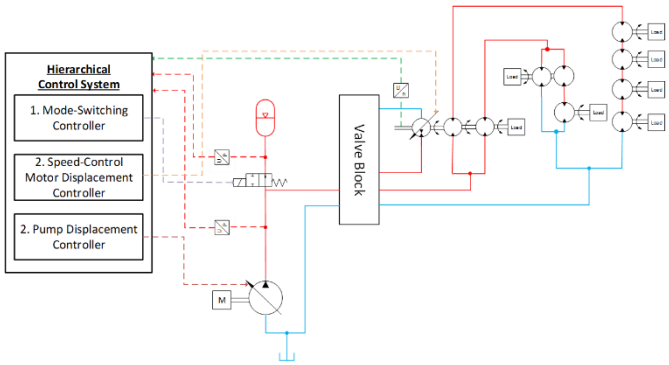


Figure 12. PDHT architecture implemented on basecutter system.

The upper limit of supply line pressure (p_{Max}) is set to the rated pressure of the basecutter motor ($p_{rated,Basecutter}$) to ensure that the maximum torque generated by the basecutter motor in the PDHT system matches that of the baseline system, as both systems have the same total motor displacement. Additionally, to keep the supply line pressure within a narrow range, the lower limit (p_{Min}) is set to $0.8 * p_{Max}$. Also, to keep the frequency of pump stroking/de-stroking low, the accumulator size (V_{Acc}) is set to be 20 liters.

The basecutter motor shaft may experience external torque spikes due to factors such as stuck cane, debris, or the basecutter blade being lowered excessively, causing it to hit the ground. In such an event, the torque demand on the motor shaft may momentarily exceed the motor's maximum capacity, potentially causing the motor to stall or rapidly decrease its shaft speed.

In a flow-supply HT system, such torque spikes are managed effectively without stalling the motors or causing a sharp speed drop. As the motor speed decreases, the inlet pressure rapidly increases above the rated motor pressure, resulting in the motor torque overcoming the torque spike with minimal speed reduction before restoring speed control. However, in a pressure-supply HT system, this response is not possible because the supply line pressure is regulated. Even if the system temporarily overrides the pressure control logic and switches to a flow-supply mode, the pressure increase remains limited due to the high capacitance of the supply line, primarily caused by the integrated accumulator.

To manage external torque spikes, the PDHT system for the basecutter implements a speed-control strategy. A solenoid-actuated, two-position ON/OFF valve, referred to as the switching valve, connects or disconnects the accumulator from the supply line. Under normal conditions, when the external torque on the motor shaft remains within the motor's maximum capacity, the system operates in pressure-supply mode with the accumulator connected and the system following PDHT logic as outlined in Section-2. However, during a torque spike, when the torque demand exceeds the motor's capacity, the system switches to flow-supply mode, disconnecting the accumulator to enable a rapid pressure rise. Additionally, the motor's shaft speed is regulated through primary control at the pump, while the secondary controlled motor (i.e., SCU) shifts to maximum displacement, functioning as a fixed displacement motor.

This strategy is implemented by the hierarchical control system, also shown in Figure 12, comprising of:

- Level 1: Mode-Switch Controller – for deciding which mode to operate on based on motor's shaft speed and pump outlet pressure feedback, as illustrated in Figure 13.
- Level 2: Pump/Motor Displacement Controller – to adjust the displacement of the pump and the speed-control motor to regulate the motor's shaft speed based on system mode and shaft speed feedback, as illustrated in Figure 14.

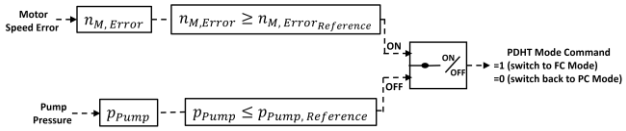


Figure 13. PDHT implementation for basecutter system: Mode-Switch Controller

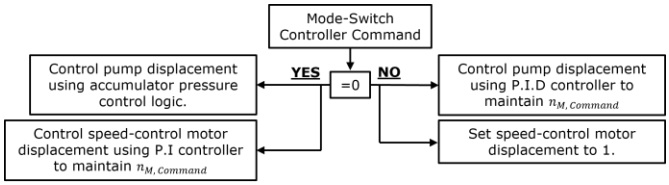


Figure 14. PDHT implementation for basecutter system: Pump/Motor Displacement Controller.

3.3 Simulation Model

The lumped-parameter simulation model for both the baseline and PDHT systems is developed in Simcenter Amesim. Key components include the basecutter pump, motors, rotating components, and hydraulic lines, with the accumulator and switching valve incorporated for the PDHT implementation. The accumulator is assumed to follow the ideal gas laws and operates under adiabatic conditions. Each component is modeled individually and integrated into the complete system model.

Both models use identical efficiency maps for common components to attribute efficiency improvements to architectural changes, not component differences. In the PDHT system model, the speed-control unit, a variable displacement motor with over-center capability, employs the efficiency map of a commercially available motor with similar over-center functionality. Moreover, it is assumed that the pump transient timescale is significantly shorter than that of stroking and de-stroking, justifying the use of steady-state efficiency maps.

However, the pump shaft input energy, derived from efficiency maps, does not account for parasitic losses at zero displacement. To incorporate these losses in the PDHT system's energy calculations, the input energy at zero displacement is assumed to be equal to that at 1% fractional displacement for the same pump outlet pressure and shaft speed. This approach provides a conservative estimate of parasitic losses, as energy consumption at zero displacement is generally lower than, or at most equal to, that at 1% fractional displacement under the same outlet pressure.

The relevant parameters assumed for PDHT simulation are: -

i. $p_{Max} = p_{rated,Basecutter}$
ii. $p_{Min} = 0.8 * p_{Max}$
iii. $\gamma = 1.4$
iv. pump stroking/de-stroking response time = 200 ms
v. $\Delta p_{switching\ valve}$= 10 bars at 325 L/min
vi. switching valve spool dynamics: 1^{st} order

vii. $\tau_{switching\ valve\ spool}$ = 100 ms

3.4 Simulation Results

The model takes the estimated external torque on each motor shaft as input and tracks the specified basecutter motor speed ($n_{Basecutter,desired}$) while maintaining the pump at the given shaft speed (n_{Pump}). The model outputs the pump's shaft torque (T_{Pump}), which is then used to compute the pump's shaft input power ($P_{shaft\ input,P}$) and input energy ($E_{shaft\ input,P}$), as defined in eq. 13-14 respectively. Additionally, for each motor, the model provides the total shaft torque (T_{M-i}) and shaft speed (n_{M-i}) for each motor, which are used to determine the motor's shaft output power ($P_{shaft\ output,M-i}$) and output energy ($E_{shaft\ output,M-i}$), as given in eq. 15-16 respectively.

$$P_{shaft\ input,P} = T_{Pump} \cdot n_{Pump} \tag{13}$$

$$E_{shaft\ input,P} = \int_0^t P_{shaft\ input} \cdot dt \tag{14}$$

$$P_{shaft\ output,M-i} = T_{M-i} \cdot n_{M-i} \tag{15}$$

$$E_{shaft\ output,M-i} = \int_0^t P_{shaft\ output,M-i} \cdot dt \tag{16}$$

The useful work for either of the two systems is defined by eq. (17). Both systems generate approximately 2400 kJ of total work over the 140-seconds drive cycle

$$\text{Usefull work} = \sum E_{shaft\ output,M-i} \tag{17}$$

The average system efficiency ($\eta_{sys-i_{Average}}$) for both the systems is defined by eq. (18). Also, the reduction in energy consumption of PDHT system over the baseline is computed using eq. (19).

$$\eta_{sys-i_{Average}} = \frac{Useful\ Work}{E_{shaft\ input,P}} \tag{18}$$

$$\Delta E_{PDHT} = \left(\frac{E_{shaft\ input,P_{Baseline}} - E_{shaft\ input,P_{PDHT}}}{E_{shaft\ input,P_{Baseline}}} \right) \tag{19}$$

Table 1 provides the summary of the comparison between the baseline and PDHT systems.

Table 1. Energy Comparison Summary Results

System #	$E_{shaft\ input,P}$ (Norm.) [-]	$\eta_{sys,Average}$[%]	ΔE_{PDHT}[%]
Baseline	100	53 %	--
PDHT	85	63 %	15 %

Figure 15-16 illustrate the energy flow and losses in the baseline and PDHT systems. The diagrams show that the PDHT system significantly reduces pump losses without substantially increasing motor or other losses, resulting in 15% lower overall energy consumption compared to baseline system. Energy analysis of the PDHT system further indicates that approximately 2% of pump's shaft input energy is lost due to excess energy recirculation during the SCU's negative torque condition, highlighting the potential for additional efficiency gains by mitigating these instances. Moreover, selecting pumps and motors with optimal pressure ranges aligned with the supply line pressure range could further improve system efficiency and reduce overall energy consumption.

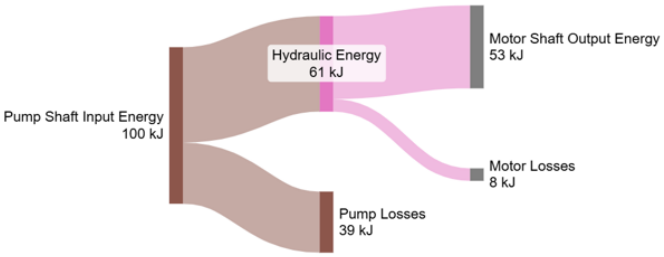


Figure 15. Normalized energy flow diagram of baseline system.

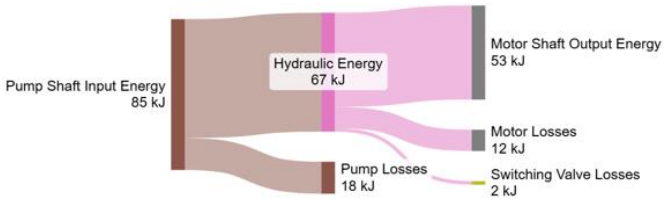


Figure 16. Normalized energy flow diagram of PDHT system.

The stroking/de-stroking frequency of the PDHT system for the given drive cycle is approximately 0.3 Hz, as indicated by Figure 17.

Due to the lower hydraulic stiffness of the accumulator, the PDHT system exhibits significantly fewer speed oscillations compared to the conventional flow-supply primary control HT. This is evident in Figure 18, which compares the basecutter motor's shaft speed tracking of PDHT system against baseline system.

4 Conclusion and Future Work

A novel hydrostatic transmission architecture for off-highway machines is presented, which decouples the operating conditions of positive displacement units from the HT, thereby enhancing transmission efficiency and improving performance across a range of drive cycles. The architecture enables straightforward and economical implementation for diverse applications. A comparative case study demonstrates that the proposed architecture reduces overall energy consumption by 15% compared to flow-supply primary control series HT, while minimizing modifications and additional components, for a sugarcane harvester application.

The study of the PDHT architecture is in its early stages. Current research focuses on: (1) assessing the accuracy of using steady-state efficiency maps in energy calculations for the PDHT architecture, (2) reducing the frequency of pump stroking/de-stroking or increasing its duty ratio without

widening the supply line/accumulator pressure range, (3) minimizing negative torque conditions in the SCU to reduce conversion losses during energy regeneration, and (4) optimizing the architecture to lower fuel consumption by enabling the engine to operate in a more efficient range.

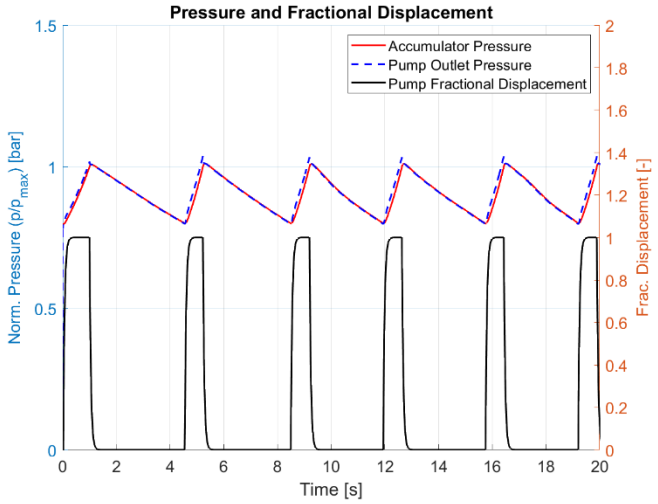


Figure 17. Pump stroking/de-stroking along with supply line/accumulator pressure variation of PDHT system.

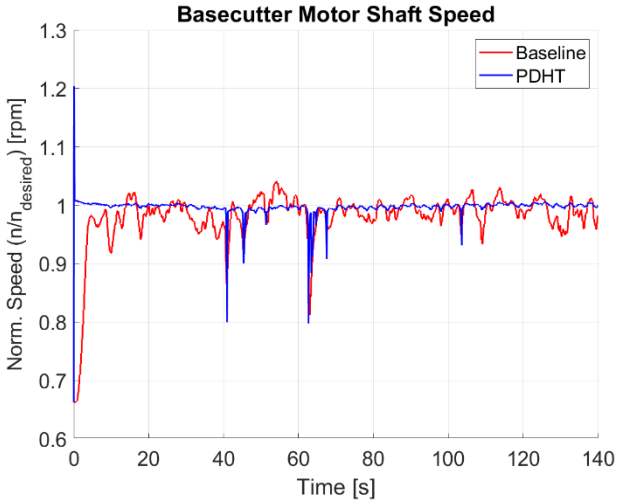


Figure 18. Motor's shaft speed oscillations comparison of PDHT against baseline system.

Acknowledgements

The authors would like to acknowledge Siemens Digital Industries Software for the use of Simcenter Amesim in the development of the simulation models for the study.

Nomenclature

Designation	Denotation	Unit
k_{oil}	Equivalent line bulk modulus	Pa
p_H	Supply line pressure	Pa
p_T	Return line pressure	Pa
p_{Pump}	Pump outlet pressure	Pa
p_0	Accumulator pre-charge pressure	Pa
$p_{Max,Allowed}$	Max allowable supply line pressure in PDHT	Pa
p_{Max}	Max pressure of supply line in PDHT	Pa
p_{Min}	Min pressure of supply line in PDHT	Pa
$p_{rated,Basecutter}$	Basecutter motor rated pressure	Pa
$\Delta p_{switching\ valve}$	Pressure drop across switching valve	Pa
V_0	Total volume of accumulator	Liter
V_{Acc}	Accumulator size	Liter
ϵ_P	Fractional displacement of pump	--
ϵ_M	Fractional displacement of motor	--
ϵ_{SM}	Fractional displacement of secondary-controlled motor	--
$V_{D,P}$	Max volumetric displacement of pump	cc/rev
$V_{D,M}$	Max volumetric displacement of motor	cc/rev
$V_{D,SM}$	Max volumetric displacement of secondary-controlled motor	cc/rev
T_P	Pump's shaft torque	Nm
T_M	Motor's shaft torque	
T_{SCU}	Torque generated by SCU	Nm
$T_{M,Max}$	Max. torque generated by fixed displacement motor in SCU	Nm
T_{Ext}	External torque on motor's shaft	Nm
n_P	Pump's shaft speed	rad/s
n_M	Motor's shaft speed	rad/s
$n_{M,Error}$	Motor's shaft speed error in PDHT	rad/s
$E_{shaft\ input,P}$	Pump's shaft input energy	J
$E_{shaft\ output,M}$	Motor's shaft input energy	J

ΔE	Reduction in energy consumption	J
$P_{shaft\ input,P}$	Pump's shaft input power	W
$P_{shaft\ output,M}$	Motor's shaft input power	W
P_{DHT}	Power transmitted through decoupled HT path	W
P_{SCU}	Power transmitted through speed-control path	W
η_{DHT}	Efficiency of decoupled HT path	--
η_{SCU}	Efficiency of speed-control path	--
η_{PDHT}	Transmission efficiency of PDHT	--
$\eta_{sys,Average}$	Average system efficiency	--
γ	Polytropic coefficient	--
$\tau_{switching\ valve\ spool}$	Time constant for switching valve dynamics	s

References

[1] Burrows, C. R., 1972, Fluid Power Servomechanisms, Van Nostrand Reinhold, New York.

[2] Vacca, A., Franzoni, G. Hydraulic Fluid Power: Fundamentals, Applications, and Circuit Design; Wiley: Hoboken, NJ, USA, 2021.

[3] Manring, N. D. (January 12, 2016). "Mapping the Efficiency for a Hydrostatic Transmission." ASME. *J. Dyn. Sys., Meas., Control*. March 2016; 138(3): 031004.

[4] Guo, X., & Vacca, A. (2021). "Advanced Design and Optimal Sizing of Hydrostatic Transmission Systems." *Actuators*, *10*(9), 243.

[5] Caldwell, N. "Review of Early Work on Digital Displacement® Hydrostatic Transmission Systems." *Proceedings of the BATH/ASME 2018 Symposium on Fluid Power and Motion Control*. *BATH/ASME 2018 Symposium on Fluid Power and Motion Control*. Bath, UK. September 12–14, 2018. V001T01A061.

[6] Green, M, Macpherson, J, Caldwell, N, & Rampen, WHS. "DEXTER: The Application of a Digital Displacement® Pump to a 16 Tonne Excavator." *Proceedings of the BATH/ASME 2018 Symposium on Fluid Power and Motion Control*. *BATH/ASME 2018 Symposium on Fluid Power and Motion Control*. Bath, UK. September 12–14, 2018. V001T01A047.

[7] Matteo Pellegri, Matthew Green, Jill Macpherson, Callan McKay, and Niall Caldwell. "Applying a Multi-Service Digital Displacement ® Pump to an Excavator to Reduce Valve Losses." In *12th International Fluid Power Conference (IFK'20).* Dresden, Germany, 2020.

[8] Kress, J. H. (1968). "Hydrostatic Power-Splitting Transmissions for Wheeled Vehicles — Classification and Theory of Operation." *SAE Transactions*, *77*, 2282–2306.

[9] Mistry, SI, & Sparks, GE. "Infinitely Variable Transmission (IVT) of John Deere 7000 TEN Series Tractors." *Proceedings of the ASME 2002 International Mechanical Engineering Congress and Exposition*. *Fluid Power Systems and Technology*. New Orleans, Louisiana, USA. November 17–22, 2002. pp. 123-131. ASME.

[10] Macor, A., & Rossetti, A. "Optimization of Hydro-mechanical Power Split Transmissions." *Mech. Mach. Theory* 2011, *46*, 1901–1919.

[11] Xu, Z., Bertolin, M., & Vacca, A., and Nilsson, J., "A Hydraulic Architecture Based on Multi-Common Pressure Rail Principle Using Multi-Chamber Cylinders for Excavators", Proceedings of the Global Fluid Power Society PhD Symposium, Hudiksvall, Sweden, Jun. 2024.

[12] Li, PY, Siefert, J, & Bigelow, D. "A Hybrid Hydraulic-Electric Architecture (HHEA) for High Power Off-Road Mobile Machines." *Proceedings of the ASME/BATH 2019 Symposium on Fluid Power and Motion Control*. *ASME/BATH 2019 Symposium on Fluid Power and Motion Control*. Longboat Key, Florida, USA. October 7–9, 2019. V001T01A011.

[13] Vukovic, M, Sgro, S, & Murrenhoff, H. "STEAM: A Mobile Hydraulic System With Engine Integration." *Proceedings of the ASME/BATH 2013 Symposium on Fluid Power and Motion Control*. *ASME/BATH 2013 Symposium on Fluid Power and Motion Control*. Sarasota, Florida, USA. October 6–9, 2013. V001T01A005.

Energy Management of Hydraulic Flight Control Actuation Systems using Multi-Mode Control

Alessandro Dell'Amico

Saab AB, Linköping, Sweden
Linköping University, Sweden

Abstract

This paper presents an analysis of a multi-mode control hydraulic flight control actuation system. The multi-mode control enables the actuator to be active in different modes and thereby reduces the required pump flow. The hydraulic supply system can be down-sized saving weight, space, and power consumption. This directly translates to aircraft fuel savings. The multi-mode control actuator is applied to an electro-hydraulic local supply system where the pumps are driven from electric motors instead of the engine mounted gearbox. An additional mode setting is the pump pressure setting. The work compares the multi-mode system to a traditional hydraulic flight control system. A simulation framework is developed that allows to analyze the required pump flow for various flight conditions and maneuvers. The information is used to estimate the total weight of both systems. A strategy is proposed how to integrate the system simulation information into an aircraft design framework that includes the systems architecture effects on aircraft sizing. Although the multi-mode system requires more components, the results show that the weight of both systems are similar. This is thanks to a 60% pump flow savings using the mode shifting strategy.

Keywords: Hydraulic actuator, energy management, flight control

1 Introduction

This paper investigates a multi-mode hydraulic flight control actuation system compared to a traditional system based on a proposed methodology. A traditional hydraulic flight control system is based on a centralized redundant architecture with gearbox driven pumps supplying oil at constant pressure. Here the flow requirements are derived from the maximum actuator load and speed. However, since the maximum load and speed do not occur at the same time, introducing mode-control can downsize the pump flow requirements and save system weight and size. The maximum load varies with the flight conditions, varying the system pressure can save on the power delivered by the pumps in certain flight conditions.

The multimode system is combined with an electrified local supply system. Electrifying the hydraulic supply system gives benefits such as increased efficiency and a higher degree of flexibility in terms of control, packaging, and integration with other electrical systems. Electrical systems are in general heavier than their hydraulic counterparts and more components are needed. The introduction of mode-control can overcome these drawbacks to some extent.

Mode shifting actuators is uncommon in aircraft applications. A mode shifting actuator was developed and successfully tested in the demonstrator aircraft YF-23 developed by Northrop/McDonnell Douglas, with hydraulic flow savings up to 60%, [1]. This actuator uses a tandem configuration cylinder with different area sizes. Switching modes is done for a low aerodynamic pressure range and high pressure range.

This type of cylinder with different area sizes is also known as a multi-chamber cylinder. There are many studies within the field of digital hydraulic implementing such a cylinder in mobile machines in order to overcome the typical throttling losses of traditional hydraulic systems. This was proven in [2] where a sophisticated system was implemented based on the multi-mode cylinder. The concept of digital hydraulics using the multi-chamber cylinder has been investigated in several studies for flight control actuation. A summary of several concepts is found in [3]. This approach has been proven to drastically increase the energy efficiency of the hydraulic actuator but comes with an increased complexity and requires advanced control strategies.

Local electro-hydraulic systems have been developed in [4] and [5]. These units are defined as being important steps towards more electrification of the aircraft systems. For large commercial aircraft the local hydraulic systems can be placed near the consumers avoiding long pipes. The electro-hydrostatic unit in [5] handles all the nose gear functions. Local electro-hydraulic systems are already flying on the Airbus A380 as a back-up system, thus eliminating one of the hydraulic systems [6].

Other means to save hydraulic power is presented in [7], with variable system pressure as one of the biggest contributors. However, varying the system pressure is complex and requires knowledge on the required actuator force, according to [8].

A framework is prepared for a conceptual comparison of the two architectures both at a subsystem level and an aircraft level. This framework is based on work presented in [9]. At a subsystem level the power extraction, cooling needs, and mass are estimated. This information is integrated into an aircraft model where the impact of the system characteristics is accounted for. The main focus in this work is the subsystem level, but a discussion is made of the aircraft level integration studies. The intended case-study is a fighter aircraft, for which the systems weight are estimated and compared.

2 Delimitations

This work only considers the primary flight control actuation system with the intention to highlight possible benefits with a mode-controlled hydraulic actuator. Although synergies exist several other subsystems and airframe, these are not included in order to narrow the scope for the sake of this study. The aircraft and mission used for integration are solely fictive but serve the purpose of creating a realistic integration environment for the actuation system.

The intention with the study is to remain at a higher abstraction level, meaning that higher order dynamics, details, and advanced control are excluded to facilitate the comparison at an architectural level. The integration at aircraft level is primarily illustrative for this work but highlights the process.

3 Method

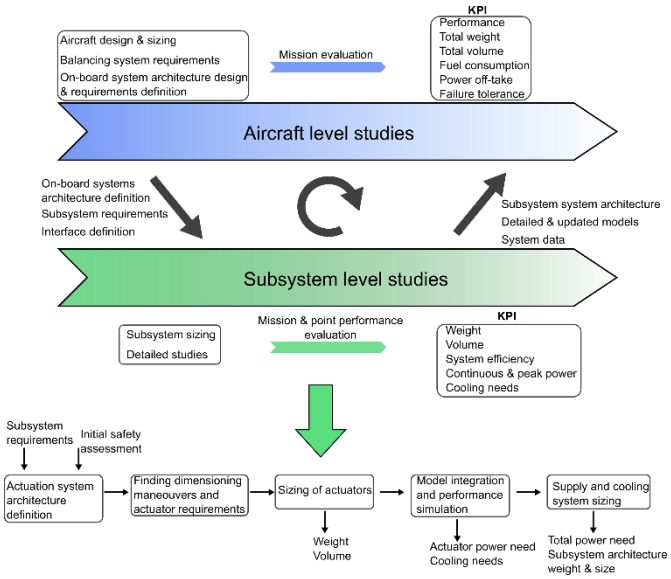


Figure 1: Aircraft system design and on-board systems interaction, [9]

A framework and methodology were developed in [9] for comparing different flight control actuation system architectures during the early aircraft conceptual design phase, where also a comprehensive description of the framework is presented. As shown in figure 1 studies should take place both at subsystem level and aircraft level providing a complete analysis of the system's performance, characteristics, and impact on the aircraft design. Following the steps of the defined methodology, an integrated environment allows to size and analyze the complete architecture, including actuators and supply system.

Information on the actuation system's characteristics and power requirements are propagated to the aircraft model. The design process is iterative during the entire phase. At an early stage, subsystem studies can benefit from a generic aircraft model of a certain configuration to analyze power and cooling needs, as well as initial sizing. This means that while the aircraft design is shaping up, subsystem studies can continuously provide the necessary information.

A generic fighter aircraft model of delta-canard configuration is used as a baseline for the actuation system study. The model is defined and described in [10]. By integrating the actuation system configurations an understanding of how the power and cooling needs vary for different flight maneuvers can be analyzed. Although a different design might eventually take place at aircraft level, this information is very useful to include early on. A proposal on how this is made is provided. The subsystem modelling and analysis have been done using the Matlab/Simulink tool version 2022b, while the aircraft integration has been done using the Desmo 9.0 tool from Pace.

4 System description and mode analysis

4.1 Mode control hydraulic actuator description

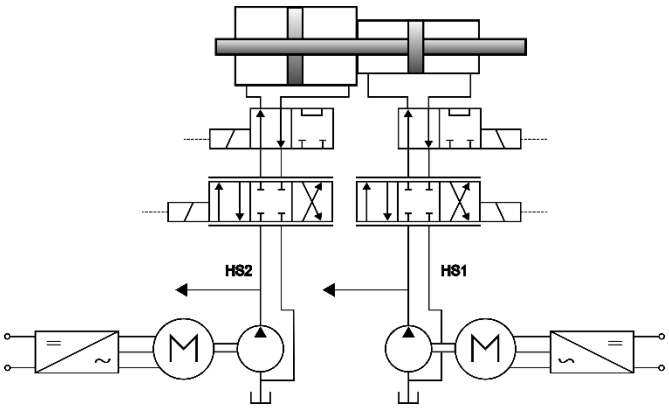


Figure 2: Multi-mode hydraulic actuator.

The reference hydraulic actuator consists of a symmetrical tandem cylinder supplied by two separated systems with gearbox driven pumps (the gearbox is attached to the engine). This is driven by safety requirements and the tandem cylinder ensures operation, although with limited power, in case of single failure. Two servo valves control the position of each half of the cylinder. For each side there is a by-pass valve connected both cylinder chambers in case of loss of hydraulic power.

The proposed system is shown in figure 2. It is similar to the reference with the main difference that each half of cylinder have different area sizes. Since the by-pass valves enable to deactivate one side of the cylinder, they can be used to choose

how to operate the actuator: either the small side, the big side, or both sides. The actuator can be set in three different modes depending on the operating conditions. The total actuator force is defined by the required stall load, as for the reference actuator. The respective size of each cylinder side can be defined to minimize the total flow requirement.

Although the mode-controlled actuator works in a traditional centralized supply system, local electro-hydraulic drives are studied here. It is assumed that the motors run from a 270 VDC electric supply system.

4.2 Requirement analysis

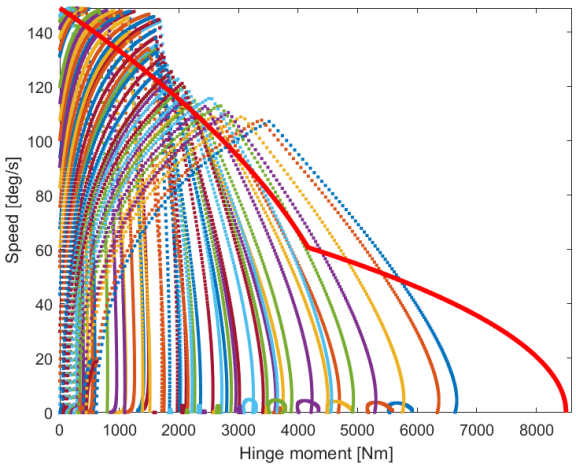


Figure 3: Force and speed requirements.

The required actuator force and speed can be found by simulating the aircraft model for various flight conditions and recording the control surface hinge moment and rate. An example is shown in figure 3, where a distinct separation is notable for higher aerodynamic pressure. As the aerodynamic pressure increases, the maximum control surface deflection is reduced and thereby the maximum speed is also reduced. It should be noted that the required speed and force looks different for other aircraft configurations, this is just one example. An approximative function is defined bounding the requirements region, shown as the bold red curve. The function is used for defining the actuator flow requirements.

Different control modes of the actuator are possible, a few are presented here in figure 4. The first approach, similar to the case with the YF-23 [1], is to choose control mode with respect to the maximum aerodynamic pressure, (q). The maximum q is given by the current flight condition: flight altitude and speed. In this way the mode shifting is not very frequent and depends on relatively slow changing variables compared to the actuator performance. Two strategies are evaluated based on q with two modes and three modes respectively. In both cases the maximum force is given by activating both cylinder sides. For the two-mode case, either the smaller piston area is activated or both, depending on the q. The piston and valve sizes are selected to fulfill the required actuator speed and the resulting flow and speed are marked as “2M, q-cntrl” in the figure. The three-mode case is similar with three different q zones. The required speed has been traded to keep the required flow low. This case is marked as “3M, q-cntrl”.

The other two strategies are based on a continuous mode shift depending on the actual actuator load. The load depends on the flight conditions and actuator position. This means that the actuator changes mode during its own operation. As the actuator move and the load increases, a mode shift occurs when the load reaches a certain limit. Both two modes and three modes are evaluated where the piston area and valve sizes are selected to minimize the required flow to maintain the required speed and fulfill the required stall load.

The most interesting approaches are the two-mode q-control and the three-mode continuous control strategies. Both fulfill the required speed while reducing the required flow by 50% and 62% respectively. The advantage with the q-control strategy is that mode-shifting only depends on the flight conditions. It will take place as the aircraft transition from one region of the flight envelope to another. Mode shifting takes place less frequently and not during actuator movement, thus, it could be easier to handle and less noticeable in actuator performance. The continuous control strategy shifts mode during actuator movement and will be more sensitive to handle. Nevertheless, the load is still dependent on the flight conditions and up to three mode shifts will only occur in the highest q-region. The main advantage is the additional reduction in required flow, that will affect the entire supply chain and installation.

The investigation continues with the three-mode strategy, focusing on integration aspects. Steady-state models are used for the analysis, while dynamic aspects and mode-shifting are left for future work.

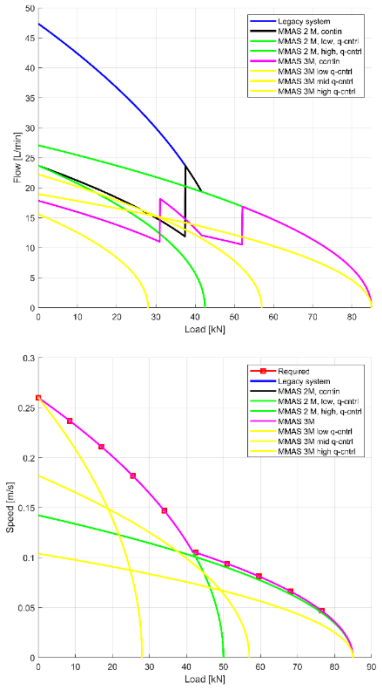


Figure 4: Various mode selection strategies.

4.3 Architecture definition

Two actuation system architectures are considered in this work as illustrated in figures 5, the reference system, and 6, the multi-mode system. Since the example aircraft has one engine, only one gearbox is available. Although no detailed safety analysis has been made, the two architectures have been designed with resilience to single fault failure in mind, that is, no single fault shall lead to a catastrophic event. There are in total seven control surfaces, each connects to a dual redundant actuator.

The reference system is supplied by a centralized hydraulic system. There a two separated system supplying each actuator. The pumps are attached to the same gearbox driven by the engine. Cooling is considered both for the hydraulic oil and the gearbox. The pump pressure is set to a constant 28 MPa and has a variable displacement, providing only the required amount of oil to move the control surface, as well as to cover for internal leakage. The two separate hydraulic systems ensure functionality in case of any fault in the hydraulic chain. An electrohydraulic unit supplied from a 270 VDC battery provides hydraulic power to one system if gearbox drive is lost. Additional safety measures could be necessary in practice.

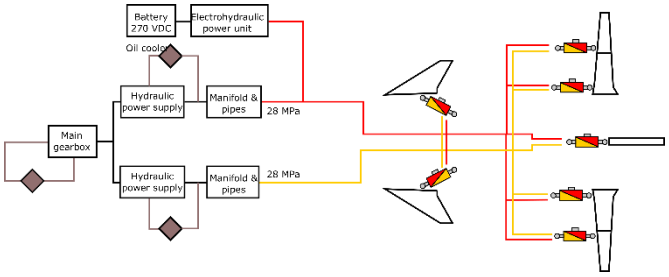


Figure 5: Reference hydraulic architecture.

The multi-mode system is designed in a similar way. Two redundant electro-hydraulic units supply oil to the rear actuators (elevators, ailerons, and rudder) and a forward unit supplies the canard actuators. Each unit consists of two fixed displacement electric motor driven pumps that share the same reservoir. A 270 VDC electric system is assumed. The electric system consists of a wild frequency generator and rectifying control unit, both systems driven by the same gearbox. Oil cooling is assumed for the electric power generation and gearbox. Two separate electric power systems ensure actuator drive in case for any single fault. If the gearbox drive is lost a battery pack can supply one of the electric systems.

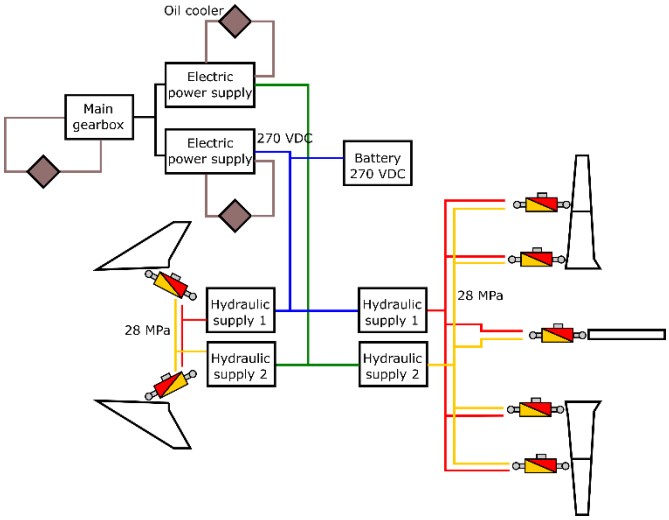


Figure 6: Multi-mode actuation system architecture.

5 System integration

5.1 Actuation system modelling

The system modelling has taken place in the Matlab/Simulink package using the Simscape library, shown in figure 7. Both the reference actuator and the multi-mode actuator are modelled in a similar way. The difference is in the piston sizing and control strategy.

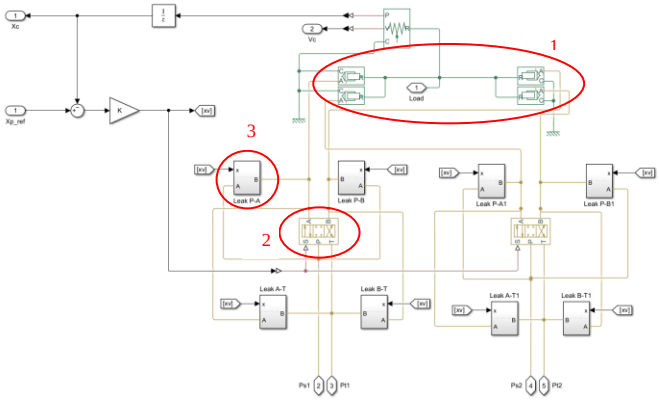


Figure 7: Actuator simulation model. 1) shows the multi-chamber cylinder, 2) the servo valve model, and 3) the leakage model.

The servo valves are represented by a 4-way directional valve considering the transition from laminar to turbulent flow. The discharge coefficient is set to 0.7 and a maximum valve opening to 1 mm. The cylinder is represented by four hydro-mechanical converters. No oil compression is considered. The piston size is defined by the stall load requirements according to eq. 1, assuming zero tank pressure.

$$A_p = \frac{F_{max}}{2p_s} \tag{1}$$

A leakage function is inserted for each of the valve spool lands, as shown in figure 8. This function gives maximum leakage in the centered and opening position, but decays as the spool

closes on the corresponding side. The leakage will thus add to the total output flow, although other approaches are possible.

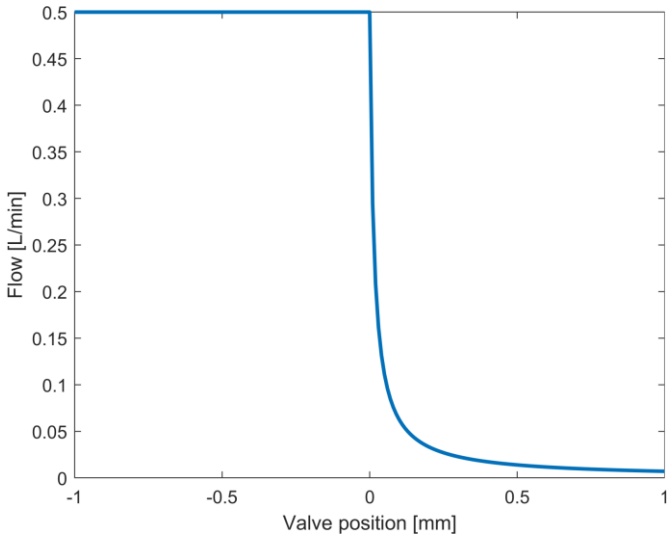


Figure 8: Servo valve leakage curve.

The actuator control loop is based on a generic approach. The actuator with its controller can be represented by the block diagram in figure 9, where Kp is the controller gain, Kq the plant gain, Ap the piston size, and s the Laplace operator.

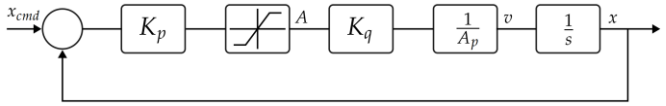


Figure 9: Hydraulic actuator block diagram.

The block diagram can be transformed into a first order transfer function on the form in eq. 2. The details of this operation are covered in reference [9].

$$G = \frac{1}{\tau \cdot s + 1} \tag{2}$$

By identification the controller gain is defined as in eq. 3.

$$K_p = \frac{A_p}{\tau K_q} \tag{3}$$

The maximum opening area is defined by eq. 4.

$$A_{max} = \frac{q_{max}}{C_q \sqrt{\frac{1}{\rho}(p_s - p_{nom})}} \tag{4}$$

The maximum flow, qmax, is defined by the speed requirements and piston size, the pump pressure is set to 28 MPa (tank pressure is assumed zero but is in practice a fraction of the pump pressure), and the nominal pressure, pnom, could be set to any desired value but is here selected to zero.

Although the control strategy is similar to the multi-mode actuator, the mode shifting needs to be considered. Besides from the maximum load, two more load levels are introduced, Flow and Fmid, with the relation according to eq. 5.

$$F_{max} = F_{low} + F_{mid} \tag{5}$$

These two force levels depict when to change to a new mode, that is, to either activate one of the cylinder sides or both. The control strategy needs the actual actuator force to determine which mode to select. If the force is low than Flow than mode 1 is selected, if it is between Flow and Fmid mode 2 is selected, if it is higher than Fmid than mode 3, which activates both cylinder sides, is selected. The actuator is only allowed to stall when the maximum force for mode 3 is reached. For mode 1 and 2 the shift has to occur before Flow and Fmid are reached respectively. Therefore, the shift occurs at 1 kN before the predefined levels.

The definition of the maximum valve opening differs also to the reference actuator. The reference actuator is defined to give maximum flow at maximum pressure drop and is allowed to drop off as the load pressure reaches the stall load. The multi-mode actuator has to consider the performance for each mode and has to deliver sufficient flow for the entire mode 1 and 2 regions. The maximum flow is therefore defined to occur at a pressure drop corresponding to 1 kN of load force. To prevent the actuator to deliver more flow corresponding to the maximum speed requirements, a dynamic saturation is introduced. The saturation calculates the maximum allowed valve position based on the current force and mode. Equation 4 can again be used here considering that the pressure drop is calculated from the current measured force. To better shape the response of the actuator a pre-filter is introduced to the reference position. The pre-filter is a first order filter where the time response is tuned so that the actuator closed loop response matches the required response.

A validation test of the actuators' performance confirms the behavior of the reference actuator and mode-controlled actuator compared to the requirements. A step of 10 mm is applied as reference position and a generic spring load with a rate of $7 \cdot 10^6$ N/m is working against the actuators. The performance is shown in figure 10. The required response is valid for the no-load case and therefore the reference actuator's performance drops off as the load increases. The mode shifting of the multi-mode actuator is clearly seen in both position and speed, although, no oil compression is included that could cause increased disturbance in position and speed control. At this point the multi-mode actuator's behavior is similar to the reference actuator.

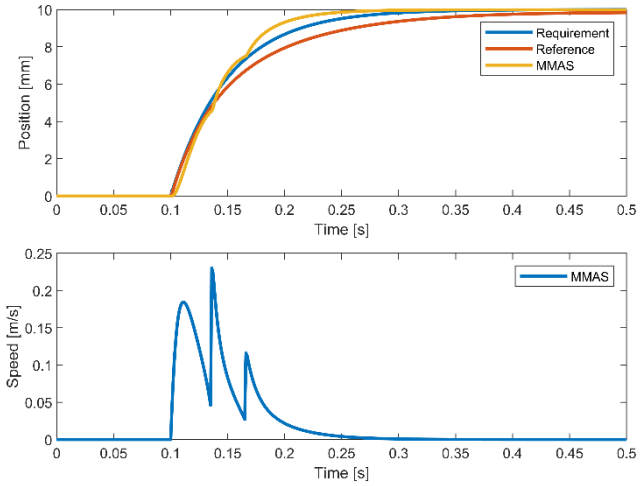


Figure 10: Actuator response and speed.

5.1.1 Varying the supply pressure

A brief analysis of introducing variable pressure with a mode-controlled actuator is provided here.

Varying the supply pressure is a basic feature in any load sensing hydraulic system used to reduce the energy consumption, typically seen on mobile hydraulic machines. Some studies can be found on this for servo hydraulic systems in aircraft, as mentioned by [8]. The challenge is to anticipate the actuator power need without implementing a very complex system. Varying the system supply pressure has a very positive effect on life cycle cost savings, as reported in [7].

Anticipating the actuator power needs for controlling the supply pressure could be done in the same way as for the mode shifting based on the aerodynamic pressure. By controlling the supply pressure based on the flight conditions the load dependency between actuators is eliminated. Assuming only two pressure levels give an indication on possible benefits. The advantage is limited for a mode-controlled actuator since the full system pressure is used for a larger portion of the actuator's working envelope. The analysis is shown in figure 11. The required actuator load pressure is calculated from the load force and piston size, top plot. The two pressure levels are 15 MPa and 28 MPa. The required actuator power (ideally, since no losses are considered here) is shown in the bottom plot. Since the actuator remains in the low force region for the larger portion of the mission, great power savings can be expected. This would not affect the supply system sizing, but is interesting from an energy management perspective, increasing the overall system efficiency. Additionally, saving power could also reduce the heat load and thereby save on the heat exchanger weight. This possibility is part of future work, and a constant supply pressure is assumed for the remainder of the work, including more advanced pressure control strategies.

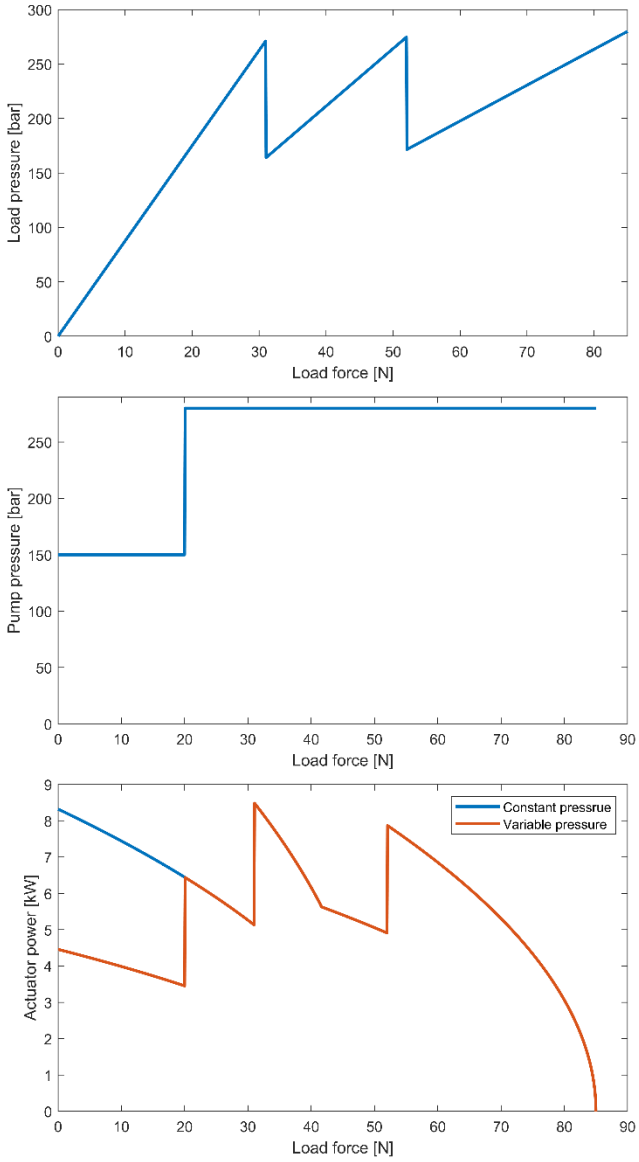


Figure 11: Varying supply pressure analysis.

5.2 System simulation integration

The system integration model was developed in reference [9], shown in figure 12. The framework is used both for supporting the system sizing and performance analysis. A 6 DOF fighter model is included, as well as aerodynamic data for calculating the control surface hinge moments. The model includes the seven control surfaces and the supply system. It is possible to simulate single maneuvers or an entire mission. The advantage of this approach is that a more realistic scenario for which the system can be analyzed is given. It is also advantageous from a sizing perspective since all actuators will not work equally all the time. A more accurate flow demand is estimated.

The actuator models are the previously defined. The hydraulic and electrical supply systems are represented by taking the sum of the actuator power, that is pressure times flow and applying a constant efficiency term. The efficiencies are

assumed to be 80% for the pumps, and 90% for the electric motor and generators, and 90% for the gearbox.

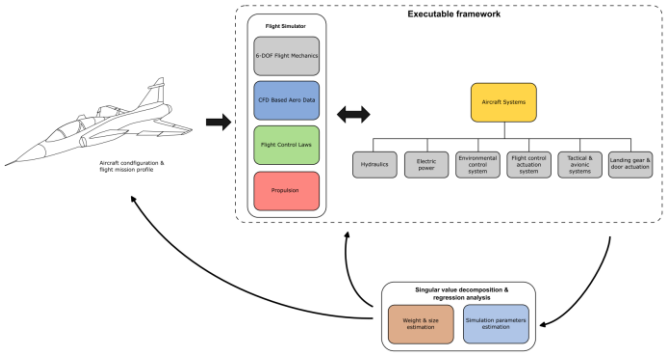


Figure 12: System simulation framework, [9].

5.3 System sizing and weight estimation

Among the most important characteristics when considering aircraft integration are the subsystems' weight, size, power consumption, and cooling needs. The focus here is all the mentioned besides the size.

The force and speed requirements are previously defined in chapter 4 by simulating a full stick input (pitch and roll) flight maneuver for various flight conditions. A safety factor of 1.2 is applied. This results in the following force and speed values used for sizing:

Control surface	Force (kN)	Speed (m/s)
Elevator	85	0.26
Ailerons	115	0.26
Canards	85	0.1
Rudder	50	0.26

The approach in this work is to use open data to develop weight estimation models. The advantage is that several samples are available to develop a more accurate model. The downside is that installation effects are not included. The availability of open data, especially for aircraft components, is also limited which will skew the results. Since this is a relative comparison between a traditional system and the multi-mode system, the lack of information applies to both systems. The weight estimation of each component follows the respective architectures. All estimation models are derived from linear regression analysis.

The pump model is based on data from [11]. Both the weight and displacement are estimated with the pump flow as input. The displacement is needed to calculate the electric motor torque for the electro-hydraulic system. The relations are defined in eq. 6 and 7.

$$W_{pump} = 0.066 \cdot q + 0.2 \text{ kg} \quad (6)$$

$$D_{pump} = 0.0006q^2 + 0.15q + 2.16 \text{ cm}^3/\text{rev} \quad (7)$$

The reservoir is assumed to be a bootstrap configuration. The weight estimation is based on data from reference [12] with total oil volume, V, as input. The relation is given in eq. 8.

$$Weight_{reserv} = 0.4 \cdot V + 5.7 \text{ kg} \quad (8)$$

The total oil volume is calculated from each actuator's piston size and stroke, and the oil in the hydraulic pipes, in addition 30%.

The gearbox weight is based on only one sample found in [13] and scaled linearly with a weight to power ratio of 0.5 kg/kW.

The electric power generation is also based on one sample from reference [14] and scaled linearly with a weight to power ratio of 0.23 kg/kW. This includes both generator and rectifier unit.

The electric motor weight estimation is based on several samples of permanent magnet motor for industrial use, without any forced cooling from reference [15]. The weight scales linearly with motor torque according to eq. 9.

$$Weight_{motor} = 0.33 \cdot T + 0.7 \text{ kg} \quad (9)$$

The electric motor torque is given from eq. 10 with the maximum pump pressure and a mechanical efficiency of 90%.

$$T = \frac{D_p p_s}{2\pi \cdot \eta} \quad (9)$$

The oil cooler is based on one sample from reference [16] and scaled linearly with the required cooling power using a weight to power ratio of 0.13 kg/kW. The required cooling power is given from the internal leakage in the system. Since the power required to drive the actuators is only intermittent during maneuvering, it is assumed that the main loss is from the leakage.

The weight for the hydraulic piping and electrical cables are given from the aircraft level integration model in chapter 5.4. This also gives the total weight of the oil.

5.4 Aircraft level integration

The Desmo 9.0 tool from Pace, [17], is used for aircraft level integration. The tool allows initial aircraft sizing based on a defined mission, as well as on-board systems architecture design and analysis. The architecture representation includes the weight, size, power consumption and bleed-air extraction from the engine. The physical representation of the components also gives a weight distribution along the aircraft. There is a built-in function that routes and sizes the power distribution elements, like hydraulic pipes and electrical cables.

The information on the subsystems' characteristics needs to be incorporated into the aircraft model. A full aircraft and system sizing is not provided in this work but a discussion on how the integration can take place. A fictive aircraft is designed based

on one of the example models provided by Desmo, shown in figure 13. The figure shows the reference architecture, where all the green components represent the actuators and the hydraulic supply system. The red lines are the hydraulic distribution elements used for weight estimation of pipes and cables.

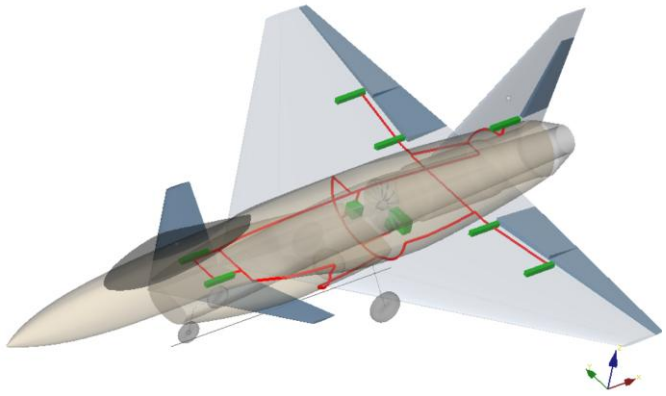


Figure 13: Example aircraft model implemented in Desmo 9.0.

Each component representation stores its own weight and power consumption. The flight mission is divided into several segments, like taxi, take-off, climb, cruise, etc. It is up to the user to define the granularity of the mission segments. The power consumption of each component is then defined for each segment. Since Desmo is not a time-based simulation tool, one approach to include the flight control actuators' power usage is to analyze the behavior from the time-based simulation for various maneuvers for a similar aircraft, and taking average power usage as input to that specific flight segment. An example of how this can be done is shown in figure 14. It shows the pump flow for the reference system during a sustained turn at Mach 0.9 at 5000 meters. A sustained turn segment can be implemented in Desmo where the average flow, or power, from the system simulation is used. In this way the different mission segments defined in Desmo can be populated with information from the system simulation model.

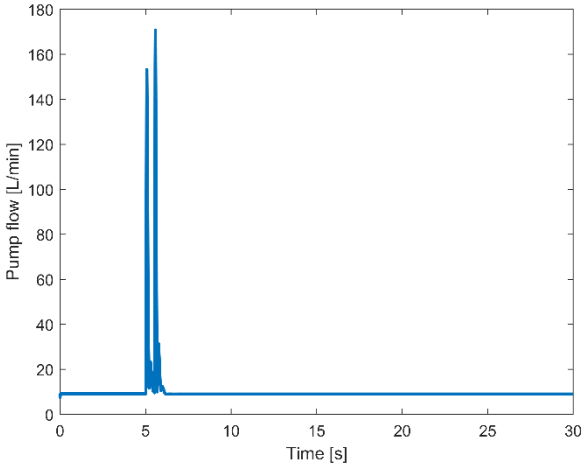


Figure 14: Required pump flow during a sustained turn.

6 Analysis

This analysis is focused on the weight comparison between the two systems shown in figures 15 and 16. The data from the systems simulation and aircraft integration has been used to estimate the weight of each component of the architectures. The total weight is similar, 115 kg compared to 125 kg, for the reference and multi-mode systems respectively. These numbers should not be seen as absolute. The quality of the results depends on the available data. Open reliable data is hard to find and for some components only one sample was used. The electrical motors are based on industrial data and are not optimized for aircraft applications. Installation effects are also not included since it would require detailed knowledge about the actual aircraft. Nevertheless, the results show that by using a multi-mode control approach, a local electro-hydraulic system can be significantly down-sized. This would, of course, also apply to a traditional architecture, though an electro-hydraulic system benefits from an increased flexibility and could easily share power between systems.

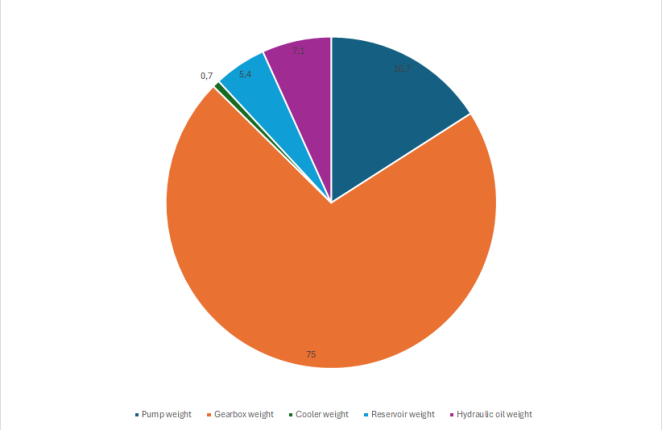


Figure 15: Weight distribution reference system.

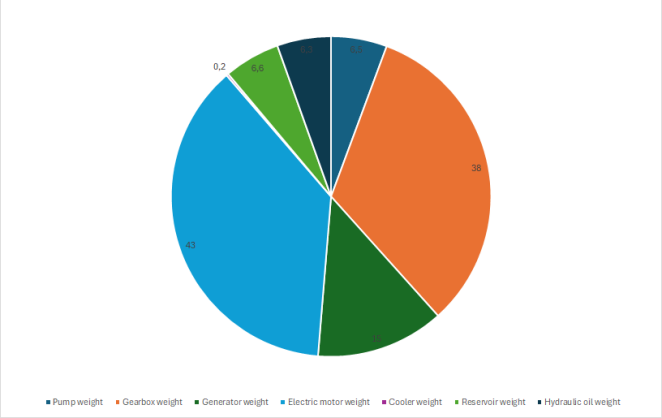


Figure 16: Weight distribution multi-mode system.

The presented work applies to top level studies in the design process. Later stages should include more detailed studies that might alter the outcome. Examples of such effects to include are hysteresis in the switching strategy, influence from natural dynamics and dynamic stiffness, excitation of natural modes coming from transients generated by the commutation, etc.

7 Conclusions

This work has shown an approach to size and analyze an hydraulic flight control system. The weight of traditional system was compared to a local electro-hydraulic multi-mode system. The multi-mode approach helps to significantly down-size the power requirements, up to 60% is possible to save. The work also highlights the interaction between subsystem level and aircraft level studies, with the aim to analyze the complete on-board systems architecture and its impact on aircraft performance. A method is proposed where system simulation is used to populate an aircraft sizing model with information on the flight control actuation power usage.

References

[1] K. W. Vieten, High performance fighter fly-by-wire flight control actuation system, AIAA 92-1123, Aerospace Design Conference, 1992

[2] K. Heybroek and M. Sahlman, A hydraulic hybrid excavator based on multi-chamber cylinders and secondary control – design and experimental validation, *International Journal of Fluid Power*, *19*(2), 91–105. https://doi.org/10.1080/14399776.2018.1447065

[3] Simoes Lopes Junior et al., A COMPARATIVE ANALYSIS OF INNOVATIVE DIGITAL HYDRAULIC ACTUATORS FOR PRIMARY FLIGHT CONTROL, 33rd Congress of the International Council of the Aeronautical Sciences, 2022

[4] L.Z. Lima et al., DEVELOPMENT OF A HIGH EFFICIENT HYDRAULIC POWER PACK, 33rd Congress of the International Council of the Aeronautical Sciences, 2022

[5] M. N. Duval and T. Lammering, DEVELOPMENT OF AN ELECTRO-HYDROSTATIC NOSE LANDING GEAR ACTUATION AND STEERING SYSTEM, 34th Congress of the International Council of the Aeronautical Sciences, 2024

[6] S. Dellac and D. Ternisien D., Airbus 380 Electro-Hydraulic back-up architecture for Braking and Steering Systems, Proceedings of the 2nd International Conference on Recent Advances in Aerospace Actuation Systems and Components, Toulouse, France, pp. 103–108, 24–26 November 2004

[7] J.B. Greene et al., Low energy consumption hydraulic techniques, Final Technical Report AD-A206 285, 30 August 1988

[8] J-C. Maré, Aerospace Actuators 2, ISTE Ltd and John Wiley & Sons, Inc, 2017

[9] A. Dell'Amico and C. Reichenwallner, A Conceptual Comparison of Hydraulic and Electric Actuation Systems for a Generic Fighter Aircraft, Aerospace 2025, 12, 1. https://doi.org/10.3390/ aerospace12010001

[10] Forsell, L.; Nilsson, U. ADMIRE The Aero-Data Model in a Research Environment Version 4.0, Odel Description; Technical Report FOI-R–1624–SE; FOI—Swedish Defence Research Agency, Systems Technology: Kista, Stockholm, 2005

[11] Aircraft Engine-Driven Pumps, Parker Aerospace, https://www.parker.com/content/dam/Parker-com/Literature/Hydraulic-Systems-Division/HSD-literature-files/HSD-OTS-Aircraft-Engine-Driven-Pumps.pdf, accessed 2025-03-02

[12] Bootstrap reservoirs, Parker Aerospace, https://www.parker.com/content/dam/Parker-com/Literature/Hydraulic-Systems-Division/HSD-literature-files/HSD_Bootstrap-Reservoir.pdf, accessed 2025-03-02

[13] Aircraft gearboxes, https://www.pbs.cz/en/Aerospace/Aviation-devices-andgine-components/Aircraft-gearboxes, accessed 2025-03-02

[14] Honeywell Aerospace Technologies, https://aerospace.honeywell.com/us/en/products-and-services/product/hardware-and-systems/electric-power/power-generation-and-conversion, accessed 2025-03-02

[15] Bosch Rexroth Synchronous Servomotors MSK, Project planning manual R911296289, Edition 12, 2018-06, Document type DOK-MOTOR*-MSK********-PR12-EN-P

[16] Hughes-Treitler Aerospace & Defense, https://www.hughes-treitler.com/tech-corn/blog/dynamics-of-fuel-cooled-oil-coolers, accessed 2025-02-12

[17] Desmo, https://pace.txtgroup.com/products/preliminary-design/pacelab-apd-for-engineers-by-engineers, accessed 2025-05-14

Impact of Impeller Structural Parameters on Cavitation of High-speed Aviation Piston Pump

Hua Fang, Hanyu Gao, Bo Yang, Yijie Lu, Hao Liu*, and Xiaoping Ouyang*

State Key Laboratory of Fluid Power & Mechatronic Systems, Zhejiang University, Hangzhou, China
E-mail: Hua Fang. fhua@zju.edu.cn, Hanyu Gao. henrygao@zju.edu.cn, Bo Yang. yang_bo@zju.edu.cn, Yijie Lu. luyijie@zju.edu.cn, Hao Liu*. haoliu7850052@zju.edu.cn, Xiaoping Ouyang*, ouyangxp@zju.edu.cn
* State Key Laboratory of Fluid Power & Mechatronic Systems, Zhejiang University, Hangzhou, China

Abstract

Impellers can effectively suppress cavitation in high-speed aviation piston pumps. However, the effects of different impeller structural parameters on cavitation are still unclear. In this paper, a CFD model of a high-speed aviation piston pump with an integrated impeller is established, and the effects of various structural parameters on cavitation are analyzed in detail based on the orthogonal test method. Simulation results show that the blade outlet width has the greatest effect on cavitation, followed by the blade outlet angle, while the impeller inlet diameter has the least effect. Within a certain range, increasing the impeller inlet diameter and blade outlet angle can effectively suppress cavitation, but the cavitation suppression effect will decrease after exceeding critical thresholds. Increasing the blade outlet width can suppress pump cavitation. The effects of different impeller structural parameters on cavitation under different outlet pressures and speeds are further analyzed. Simulation results show that the speed has a greater impact on cavitation than the outlet pressure. The results of this paper can be used to guide the design and optimization of impellers for high-speed aviation piston pumps.

Keywords: Aviation piston pump, impeller, cavitation, suction port, CFD model

1 Introduction

With many advantages such as high-power density, high output pressure, high efficiency, and variable control, axial piston pumps are a type of volumetric hydraulic pump widely used in aircraft hydraulic systems (also known as aviation piston pumps (APPs)). The development trend of efficient and lightweight aircraft has prompted the development of aviation hydraulic systems and their key components, APPs, towards higher power density. For APPs in electrical motor-driven pumps, increasing the speed is one of the effective ways to increase power density [1]. However, as the speed of the APP increases, the flow rate in the oil suction passage increases, causing the pressure in these regions to drop sharply below the air separation pressure or saturated vapor pressure, resulting in the precipitation of gas dissolved in the oil, and even the gasification of the oil itself, i.e., cavitation. When bubbles enter the high-pressure regions with the oil, they will shrink or collapse rapidly, causing cavitation damage such as erosion or detachment on the surface of components, thereby reducing the reliability and service life of APPs [2].

At present, there are two main solutions to suppress cavitation in high-speed APPs. The first is to optimize the suction passage structure. The main principle is to reduce the pressure loss in the suction passage by optimizing the suction passage structure between the suction port and the piston chamber[3-6]. The second is to increase the suction pressure. By increasing the pressure at the suction port, the pressure loss in the suction channel can be compensated. Studies have shown that the effect of optimizing the suction passage structure on suppressing cavitation is very limited [7], and increasing the suction pressure is a more effective solution. There are three main methods to increase the suction pressure, namely, pressurizing the tank [8], adding a booster pump to the suction pipeline, and adding an impeller in the suction port of APP[9]. Among them, the constraints on the size and weight of the aircraft hydraulic system limit the diameter of the suction pipeline, and it is difficult to compensate for the pressure loss between the tank and the suction port by increasing the tank pressure. Meanwhile, adding a booster pump to the suction pipeline will reduce the reliability of the aircraft hydraulic system. In contrast, integrating an impeller in the suction port allows the APP to operate normally, even under very low suction pressure conditions[10]. Therefore, compared with the other two methods, adding an impeller in the suction port is a better method.

At present, the research on impellers mainly focuses on the relationship between structural parameters (such as blade

outlet width [11], impeller inlet diameter [12-13], blade outlet angle [14-16]) and impeller performance, and various methods are used to optimize the impeller structural parameters to obtain the optimal solution of boost pressure, efficiency, and energy loss [17-23]. However, there are few studies on impellers in APPs. Chen et al. [24] studied the energy loss and distribution of the impeller based on the entropy generation theory. Dong et al. [25] verified the cavitation suppression effect of the impeller in the speed range of 1000 rpm-5000 rpm through simulation and experiments. However, the research on the influence of impeller structural parameters on cavitation in high-speed APPs is still limited.

In this paper, a CFD model of a high-speed APP with an integrated impeller is established, and the effect of impeller structural parameters, such as impeller inlet diameter, blade outlet width, and blade outlet angle, on cavitation, is analyzed based on the orthogonal test method, and reveals the effects of impeller structural parameters on cavitation under different pressure and speed conditions. The results of this paper can be used for the design and optimization of impellers in high-speed APPs.

The rest of the paper is organized as follows. In Section 2, a CFD model of a high-speed APP with an integrated impeller is established, and the optimal impeller structural parameter combination is obtained based on the orthogonal test method. In Section 3, the effects of different impeller structural parameters on cavitation are analyzed in detail, and the effect of structural parameter changes on cavitation under different operating conditions is analyzed. The conclusions are provided in Section 4.

2 Modeling

2.1 APP and impeller

The structure of the APP studied in this paper is shown in Figure 1. The piston is located in the cylinder chamber of the cylinder block and is connected to the swash plate through a slider hinged on the spherical heads. Driven by the coordinated motion of the shaft and the swash plate, the piston can both rotate around the axis and reciprocate in a straight line. The reciprocating motion of the piston causes volume in the cylinder chamber to change, thereby completing the suction and discharge stroke of the piston pump.

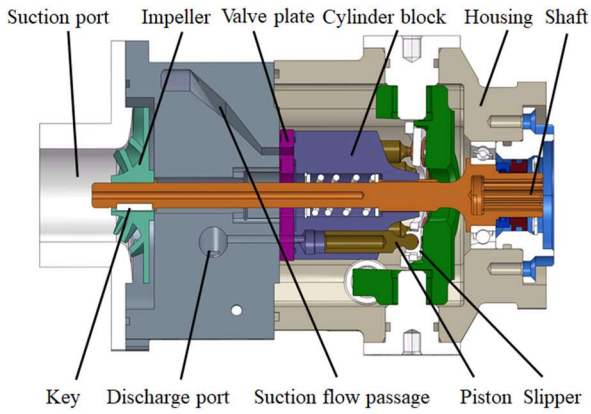


Figure 1: Structure of APP integrated with impeller.

Compared to traditional APPs, the high-speed APP studied in this paper integrates an impeller in the suction port. The impeller structure is shown in Figure 2. The impeller is rigidly connected to the APP through a flat key and rotates synchronously with the APP shaft. The main structural parameters of the impeller include the hub diameter d_h, the impeller inlet diameter D_j, the blade inlet diameter D_1, the impeller outlet diameter D_2, the blade inlet width b_1, the blade outlet width b_2, the blade inlet angle β_1, the blade outlet angle β_2, and the blade wrap angle φ. Under the action of centrifugal force, the oil in the impeller is transported to the suction flow passage of the APP. Simultaneously, the oil in the suction port is continuously replenished into the impeller under the action of the pressure difference.

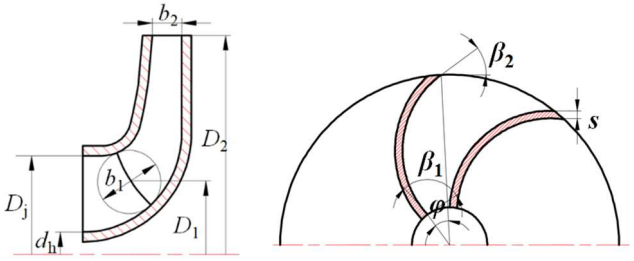


Figure 2: Impeller structure.

Table 1: Parameters of APP and impeller

Parameters	Values	Units
Displacement	2	mL/r
Number of pistons	9	null
Inclination of the swash plate	17	°
Impeller outlet diameter	49	mm
Hub diameter	11	mm
Blade inlet angle	15	°
Blade wrap angle	150	°

2.2 CFD modeling

The numerical model used in this paper integrates a turbulence model[26]and the full cavitation model proposed by Singhal et al.[27], and considers three liquid properties, namely cavitation, aeration and liquid compressibility, which have been successfully applied to CFD models of APPs [28]. The transport equation for cavitation vapor can be described as

$$\frac{\partial}{\partial t}\int_{\Omega(t)}\rho f_V \mathrm{d}\Omega + \int_\sigma \rho\left((\vec{v}-\overrightarrow{v_\sigma})\cdot\vec{n}\right)f_V \mathrm{d}\sigma = \int_\sigma \rho\left(D_V + \frac{\mu_t}{\sigma_f}\right)\cdot(\nabla f_V\cdot\vec{n})\mathrm{d}\sigma + \int_\Omega (R_e - R_c)\mathrm{d}\Omega \quad (1)$$

$$R_e = C_e\cdot\rho_l\rho_V\left[\frac{2}{3}\frac{p-p_V}{\rho_l}\right]^{1/2}(1-f_V-f_g) \quad (2)$$

$$R_c = C_c\cdot\rho_l\rho_V\left[\frac{2}{3}\frac{p-p_V}{\rho_l}\right]^{1/2} f_V \quad (3)$$

where $\Omega(t)$ is the control volume as a function of time, σ is the surface area of the control volume, $\vec{n}$ is the surface normal of

σ, ρ is the density of the hydraulic oil, $\vec{v}$ is the fluid velocity, $\overrightarrow{v_\sigma}$ is the surface motion velocity, μ_t is the turbulent dynamic viscosity, p is the static pressure, f_v is the mass fraction of vapor, D_v is the diffusivity of vapor, σ_f is the turbulent Schmidt number, R_e is the vapor generation rate, R_c is the vapor condensation rate, p_V is the liquid vapor pressure, ρ_l is the liquid density and ρ_V is the vapor density. Among them, c_e=0.02 and c_c=0.01 are the coefficients of R_e and R_c, respectively.

The full cavitation model takes into account the mass fraction of non-condensable gas in the liquid f_g, which can be calculated as

$$\frac{\partial}{\partial_t}\int_{\Omega(t)}\rho f_d d\Omega + \int_\sigma \rho\left((\vec{v}-\overrightarrow{v_\sigma})\cdot\vec{n}\right)f_d d\sigma = \int_\sigma \rho\left(D_d+\frac{\mu_t}{\sigma_f}\right)\cdot(\nabla f_d\cdot\vec{n})d\sigma + \int_\Omega \frac{\rho(f_d-f_{de})}{\tau}d\Omega \quad (4)$$

$$f_{de} = \frac{p}{p_{der}}\cdot f_{der} \quad (5)$$

$$f_g + f_d = constant \quad (6)$$

where f_d is the mass fraction of dissolved gas, f_{de} is the equilibrium mass fraction of non-condensable gas at the local pressure, D_d is the diffusivity of dissolved gas, τ is the gas absorption/release time, and p_{der} is the reference pressure of the equilibrium mass fraction of dissolved gas.

In the full cavitation model, the hydraulic fluid is considered as a fluid mixture including liquid, vapor, and free gas, and the density of the mixture can be calculated as:

$$\frac{1}{\rho} = \frac{f_V}{\rho_V} + \frac{f_g}{\rho_g} + \frac{1-f_V-f_g}{\rho_l} \quad (7)$$

The compressibility of oil is described by the bulk modulus. The relationship between the bulk modulus K and liquid density ρ_l is

$$K = \frac{\rho_{lr}}{\frac{\partial\rho_l}{\partial p}} \quad (8)$$

where ρ_{lr} is the liquid density at the reference pressure. Since the vapor density is assumed to be constant in this paper, the density of the non-condensable gas is calculated as

$$\rho_g = \frac{Wp}{RT} \quad (9)$$

where W is the molecular weight of gas, R is the universal gas constant, and T is the thermodynamic temperature.

The non-condensable gas volume fraction α_g and vapor volume fraction α_V can be calculated as:

$$\alpha_g = f_g\frac{\rho}{\rho_g} \quad (10)$$

$$\alpha_V = f_V\frac{\rho}{\rho_V} \quad (11)$$

The above numerical model is implemented in the commercial CFD software Pumplinx. Unlike traditional centrifugal pumps, the impeller integrated in the APP pressurizes the oil and transports it through the suction passage to the piston chamber with a constantly changing volume. Since it is difficult to accurately simulate the above process using the impeller model alone, the model established in this paper integrates the impeller and APP. Moreover, since this paper mainly studies the cavitation phenomenon in APPs, the established CFD model does not consider the lubrication interface and the leakage between friction pairs. As shown in Figure 3, the fluid domain of the APP with an impeller can be divided into six regions, namely the suction passage, the impeller region, the discharge passage, the valve plate suction chamber, the valve plate discharge chamber, and the piston chamber. To accurately simulate the operating conditions of the APP, this paper used the transient simulation method for numerical simulation.

In the established CFD model, the suction port and discharge port of APP are set as pressure inlet and pressure outlet respectively, the piston chamber is set as the rotation boundary condition, the valve plate suction chamber and valve plate discharge chamber are set to be stationary, the wall boundary is defined as a no-slip boundary, and the contact surface of the adjacent area is set as an interface to ensure the continuity of the oil flow. Since the impeller rotates synchronously with the piston chamber, the impeller is set to rotate relative to each other, and the adjacent surface of the impeller and the suction passage is set as an interface. The boundary conditions of the established model are detailed in Table 2.

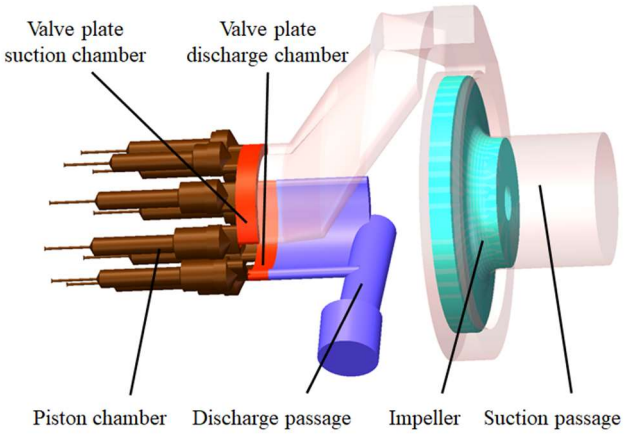


Figure 3: Fluid domain of APP with impeller.

Table 2: Boundary conditions of the CFD model.

Parameters	Values	Units
Density	837	kg/m^3
Dynamic viscosity	0.01155	Pa·s
Saturated vapor pressure	400	Pa
Inlet pressure	0.2	MPa
Outlet pressure	28	MPa
Speed	12000	rpm
displacement of the pump	2	mL/r

2.3 Orthogonal test design

Orthogonal test design is an effective method for studying multiple factors and levers, which can obtain statistically equivalent results to full factorial experiments with a limited number of experiments[17]. In this paper, two parameters are selected as evaluation indicators for measuring impeller effect, one is the impeller inlet and outlet pressure difference Δp, and the other is the average gas volume fraction in the piston chamber during one rotation of the APP per_g. The orthogonal tests are designed to investigate three structural parameters of the impeller: the impeller inlet diameter D_j, the blade outlet width b_2, and the blade outlet angle β_2, and the L9（3^3）orthogonal test results are shown in Table 3. Keeping the other structural parameters of the impeller constant, numerical simulation is performed according to the parameter combination in Table 3. The quantitative changes of Δp and per_g are utilized to evaluate the effects of different parameter combinations on cavitation.

Table 3: Orthogonal test results.

No.	Structural parameters			Δp (MPa)	per_g
	D_j (mm)	b_2 (mm)	β_2 (°)		
1	21	2	20	0.244038	1.1946%
2	21	2.5	40	0.283537	1.0507%
3	21	3	30	0.310530	0.9723%
4	25	2	40	0.257847	1.1411%
5	25	2.5	30	0.286801	1.0390%
6	25	3	20	0.302210	0.9927%
7	27	2	30	0.259146	1.1329%
8	27	2.5	20	0.270655	1.0920%
9	27	3	40	0.299906	0.9995%

The range analysis results of Δp and per_g are shown in Table 4. In Table 4, k_i (i=1,2,3) represents the average result of three structural parameters D_j, b_2, and β_2 at the i-th value, reflecting the average impact of different structural parameters on the results at a certain value. For example, k_1, k_2, and k_3 in the first row to the third row of the first column in Table 4 represent the average Δp of three impellers when D_j= 21mm, 25mm, and 27mm, respectively. The meanings of other parameters in Table 4 are similar. R is the difference between the maximum and minimum values of each structural parameter, reflecting the impact of changes in a single structural parameter on the results.

From Table 4, it can be found that, for both Δp and per_g, the range of b_2 is significantly higher than those of the other two structural parameters. Therefore, it can be concluded that within the range of structural parameters in Table 3, b_2 has the greatest effect on Δp and per_g, followed by β_2, while D_j has the least effect. The optimal combination of three structural parameters determined by the range analysis is D_j=25mm, b_2=3mm, and β_2=30°.

Table 4: Range analysis results.

		D_j	b_2	β_2
Δp	k_1	0.279368	0.253677	0.272301
	k_2	0.282286	0.280331	0.285492
	k_3	0.276569	0.304215	0.280430
	R	0.005717	0.050538	0.013191
per_g	k_1	1.072522	1.156170	1.093081
	k_2	1.057581	1.060560	1.048041
	k_3	1.074780	0.988154	1.063763
	R	0.017199	0.168017	0.04504

To further explore the effects of different structural parameters on cavitation, the effect of a single structural parameter on Δp and per_g is analyzed by keeping any two structural parameters unchanged, based on the structural parameter combination where D_j=25mm, b_2=3mm, and β_2=30°.

3 Results and discussions

3.1 Analysis of the effect of structural parameters

First, under the condition of 12000 rpm (speed)and 28MPa (outlet pressure), b_2=3mm and β_2=30° are kept unchanged, and D_j is set to 21 mm, 23 mm, 25 mm, 27 mm, and 29 mm, respectively. The obtained Δp and per_g are shown in Figure 4. It can be observed that with the increase of D_j, the boosting effect first increases and then weakens. When D_j exceeds 25 mm, the boosting effect begins to weaken. The reason is that the increase of D_j increases the cross-sectional area of the impeller flow passage and reduces the inlet flow velocity, thereby reducing flow friction loss and impact loss, and improving boosting efficiency. However, as D_j further increases, the flow velocity continues to decrease, the kinetic energy of the oil entering the impeller is insufficient, causing problems such as flow separation and vortex, which reduces the boosting effect. In contrast to the trend of Δp, the per_g value first decreases and then increases with the increase of D_j. When D_j exceeds 25 mm, per_g begins to increase. The results show that there exists an optimal value of D_j to optimize the boosting effect and cavitation suppression effect of the impeller.

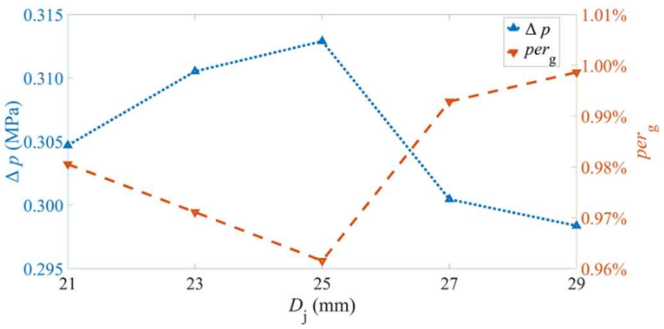


Figure 4: Effects of D_j on Δp and per_g.

Then, keeping D_j=25mm and β_2=30° unchanged, and setting b_2 to 1.5mm, 2mm, 2.5mm, 3mm, and 3.5mm, respectively. The obtained Δp and per_g are shown in Figure 5. It can be seen that with the increase of b_2, the boosting effect of the impeller gradually improves. The reason is that the increase of b_2 increases the cross-sectional area of the impeller and reduces the impeller outlet velocity, thereby reducing friction loss and turbulence loss, and improving the boosting effect. On the contrary, per_g gradually decreases with the increase of b_2. The simulation results indicate that a larger b_2 can be selected when designing the impeller. However, increasing b_2 will increase the radial size of the impeller and increase the weight of the APP. Therefore, the selection of b_2 needs to strike a balance between weight and boosting effect.

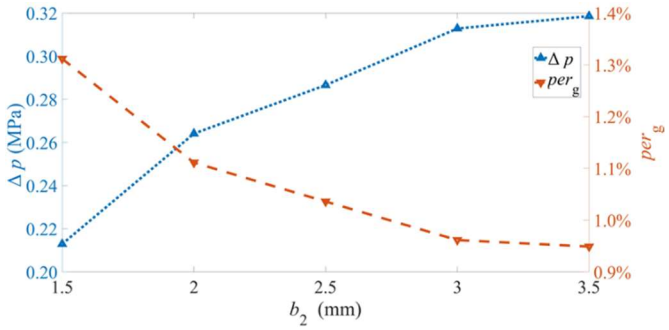


Figure 5: Effects of b_2 on Δp and per_g.

At last, keeping D_j=25mm and b_2=3mm unchanged, settingβ_2 to 10°, 20°, 30°, 40° and 45°, respectively. The obtained Δp and per_g are shown in Figure 6. It can be seen that with the increase of β_2, the boosting effect first increases and then weakens. When β_2 exceeds 30°, the boosting effect begins to weaken. The reason is that the increase of β_2 increases the tangential velocity at the impeller outlet and the theoretical head, that is, the boosting effect is improved. However, continuing to increase β_2 will intensify oil flow separation and vortex, intensify energy dissipation, and reduce the boosting effect. Contrary to the changing trend of Δp, per_g first decreases and then increases with the increase of β_2. When β_2 exceeds 30°, per_g begins to rise. Therefore, there exists an optimal β_2 that can achieve the optimal boosting effect and cavitation suppression effect at the same time.

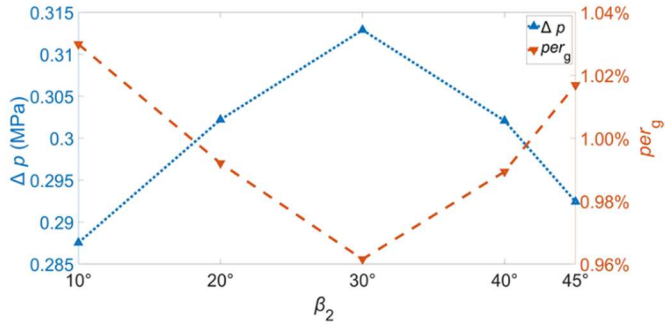


Figure 6: Effects of β_2 on Δp and per_g.

3.2 Analysis of the effect of operating conditions

To further analyze the effects of the impeller structural parameters on cavitation, simulations need to be carried out under different pressure and speed conditions.

3.2.1 Analysis of the effect of outlet pressure

The speed of APP is set to be constant at 12000 rpm, and the outlet pressures are set to 7MPa, 14MPa, 21MPa, 28MPa, and 35MPa, respectively. The obtained Δp and per_g are shown in Figures 7 and 8, respectively. It can be seen from Figure 7 that the change trend of Δp under different outlet pressures is similar, that is, with the increase of D_j, the boosting effect first increases and then weakens. The Δp of the impeller with the same D_j increases with the rise of outlet pressure. Taking the impeller with D_j=29mm, b_2=3mm, and β_2=30° as an example, as the outlet pressure increases from 7MPa to 35MPa, Δp increases from 0.295MPa to 0.299MPa. It can be seen from Figure 8 that the change trend of per_g under different outlet pressures is also similar, that is, with the increase of D_j, per_g first decreases and then increases. The per_g of the impeller with the same D_j decreases with the rise of outlet pressure. Taking the impeller with D_j=29mm, b_2=3mm, and β_2=30° as an example, as the outlet pressure increases, the maximum value of per_g decreases from 1.153% to 0.980%.

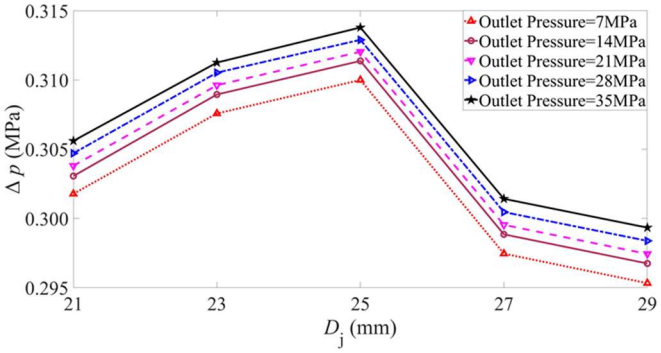


Figure 7: Effects of D_j on Δp under different outlet pressures.

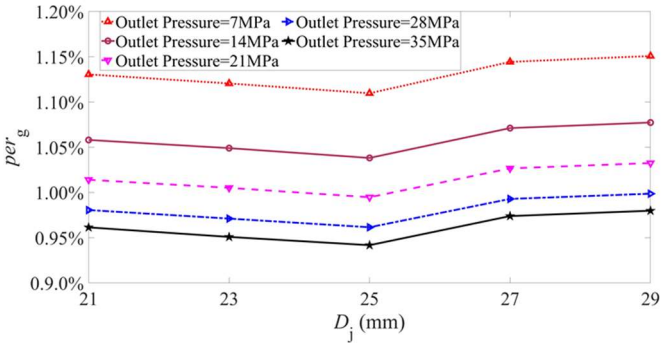


Figure 8: Effects of D_j on per_g under different outlet pressures.

The speed of APP is set to be constant at 12000 rpm, and the outlet pressures are set to 7MPa, 14MPa, 21MPa, 28MPa, and 35MPa, respectively. The obtained Δp and per_g are shown in Figures 9 and 10, respectively. It can be seen from Figure 9 that when the outlet pressure is constant, Δp is positively correlated with b_2. The Δp of the impeller with the same b_2

continues to increase with the rise of outlet pressure. Taking the impeller with D_j=25mm, b_2=1.5mm, and β_2=30° as an example, Δp increases from 0.209MPa to 0.214MPa as the outlet pressure rises from 7MPa to 35MPa. It can be seen from Figure 10 that when the outlet pressure is constant, per_g is negatively correlated with b_2. The per_g of the impeller with the same b_2 continues to decrease with the rise of outlet pressure. Taking the impeller with D_j=25mm, b_2=1.5mm, and β_2=30° as an example, the maximum value of per_g decreases from 1.48% to 1.29% as the outlet pressure rises from 7MPa to 35MPa.

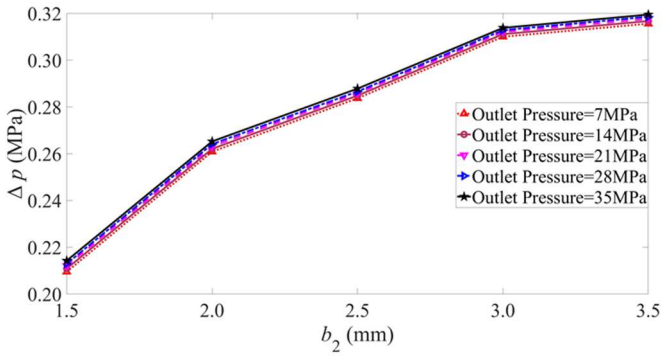


Figure 9: Effects of b_2 on Δp under different outlet pressures.

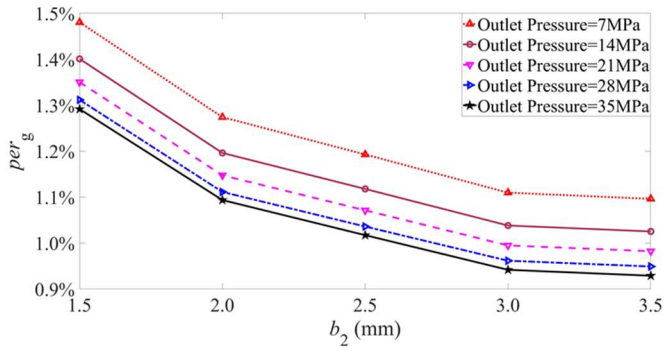


Figure 10: Effects of b_2 on per_g under different outlet pressures.

The speed of APP is set to be constant at 12000 rpm, and the outlet pressures are set to 7MPa, 14MPa, 21MPa, 28MPa, and 35MPa, respectively. The obtained Δp and per_g are shown in Figures 11 and 12, respectively. It can be seen from Figure 11 that the change trend of Δp under different outlet pressures is similar, that is, with the increase of β_2, the boosting effect first increases and then weakens. The Δp of the impeller with the same β_2 increases slightly with the rise of outlet pressure. Taking the impeller with D_j=25mm, b_2=3mm, and β_2=10° as an example, as the outlet pressure increases from 7MPa to 35MPa, Δp rises from 0.285MPa to 0.289MPa. It can be seen from Figure 12 that the change trend of per_g under different outlet pressures is also similar, that is, with the increase of β_2, per_g first decreases and then increases. The per_g of the impeller with the same β_2 decreases with the rise of outlet pressure. Taking the impeller with D_j=25mm, b_2=3mm, and β_2=10° as an example, as the outlet pressure increases, the maximum value of per_g decreases from 1.18% to 1.01%.

In summary, it can be found that with the increase of outlet pressure, the boosting effect and cavitation suppression effect are slightly improved. The reason is that with the increase of APP outlet pressure, the oil compression volume increases, resulting in a decrease in output flow, causing the impeller operating point to shift slightly to the left along the P-Q curve, that is, the boosting effect is slightly enhanced, thereby improving the cavitation suppression effect.

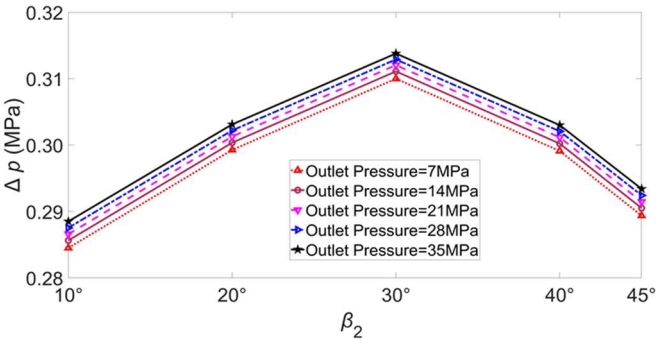


Figure 11: Effects of β_2 on Δp under different outlet pressures.

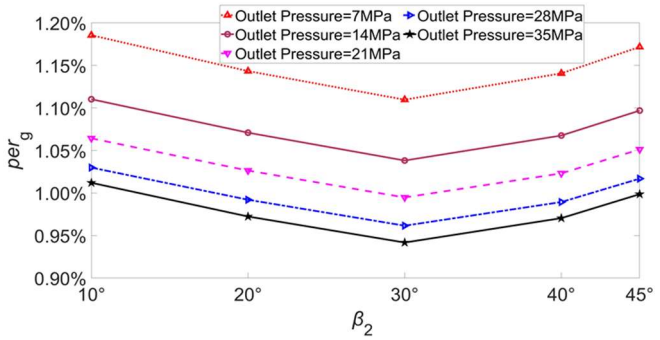


Figure 12: The effect of β_2 on per_g under different outlet pressures.

3.2.2 Analysis of the effect of speed

The outlet pressure of APP is set to be constant at 28MPa, and the speeds are set to 6000rpm, 8000rpm, 10000rpm, and 12000rpm, respectively. The obtained Δp and per_g are shown in Figures 13 and 14, respectively. It can be seen from Figure 13 that the change trend of Δp under different speeds is similar, that is, with the increase of D_j, the boosting effect first increases and then weakens. The Δp of the impeller with the same D_j increases sharply with the rise of speed. Taking the impeller with D_j=29mm, b_2=3mm, and β_2=30° as an example, with the speed increasing from 6000 rpm to 12000 rpm, Δp rises from 0.075 MPa to 0.298 MPa. It can be seen from Figure 14 that the change trend of per_g under different speeds is also similar, that is, with the increase of D_j, per_g first decreases and then increases. The per_g of the impeller with the same D_j decreases sharply with the rise of speed. Taking the impeller with D_j=29mm, b_2=3mm, and β_2=30° as an example, with the speed increasing from 6000 rpm to 12000 rpm, the maximum value of per_g decreases from 1.40% to 1.00%.

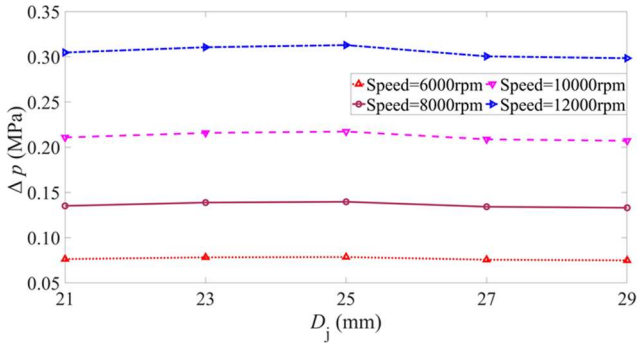


Figure 13: Effects of D_j on Δp under different speeds.

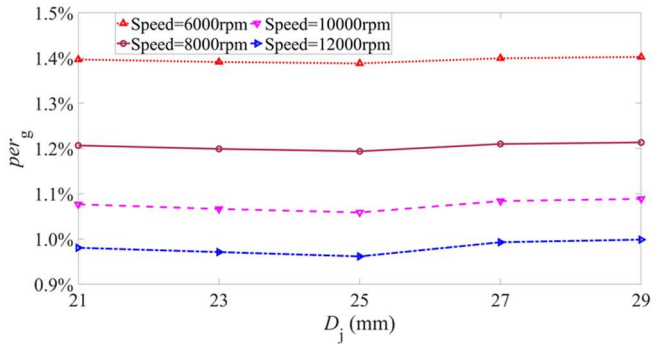


Figure 14: Effects of D_j on per_g under different speeds.

The outlet pressure of APP is set to be constant at 28MPa, and the speeds are set to 6000rpm, 8000rpm, 10000rpm, and 12000rpm, respectively. The obtained Δp and per_g are shown in Figures 15 and 16, respectively. It can be seen from Figure 15 that when the speed is constant, Δp is positively correlated with b_2. The Δp of the impeller with the same b_2 continues to increase with the rise of speed. Taking the impeller with D_j=25mm, b_2=1.5mm, and β_2=30° as an example, with the speed increases from 6000 rpm to 12000 rpm, Δp rises from 0.051 MPa to 0.213 MPa. It can be seen from Figure 16 that when the speed is constant, per_g is negatively correlated with b_2. The per_g of the impeller with the same b_2 continues to decrease with the rise of outlet pressure. Taking the impeller with D_j=25mm, b_2=1.5mm, and β_2=30° as an example, with the speed increasing from 6000 rpm to 12000 rpm, the maximum value of per_g decreases from 1.50% to 1.31%.

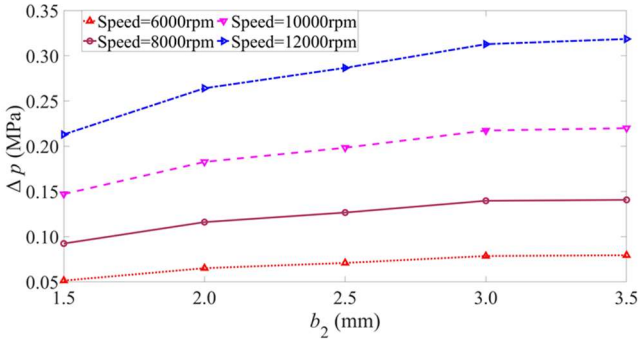


Figure 15: Effects of b_2 on Δp under different speeds.

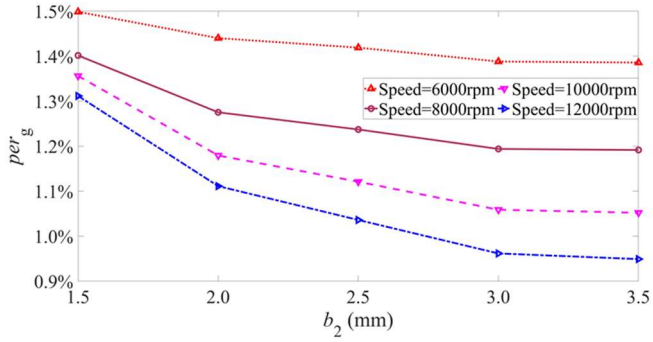


Figure 16: Effects of b_2 on per_g under different speeds.

The outlet pressure of APP is set to be constant at 28MPa, and the speeds are set to 6000rpm, 8000rpm, 10000rpm, and 12000rpm, respectively. The obtained Δp and per_g are shown in Figures 17 and 18, respectively. It can be seen from Figure 17 that the change trend of Δp under different speeds is similar, that is, with the increase of β_2, the boosting effect first increases and then weakens. The Δp of the impeller with the same β_2 increases with the rise of speed. Taking the impeller with D_j=25mm, b_2=3mm, and β_2=10° as an example, as the speed rises from 6000 rpm to 12000 rpm, Δp rises from 0.072 MPa to 0.288 MPa. It can be seen from Figure 18 that the change trend of per_g under different speeds is also similar, that is, with the increase of β_2, per_g first decreases and then increases. The per_g of the impeller with the same β_2 decreases with the rise of speed. Taking the impeller with D_j=25mm, b_2=3mm, and β_2=10° as an example, as the speed rises from 6000 rpm to 12000 rpm, the maximum value of per_g decreases from 1.41% to 1.03%.

In summary, the higher the speed, the better the boosting effect. The reason is that as the speed increases, the output flow increases significantly, which greatly improves the boosting effect. The impeller enables the high-speed APP to maintain low cavitation even under harsh conditions of low speed and high outlet pressure.

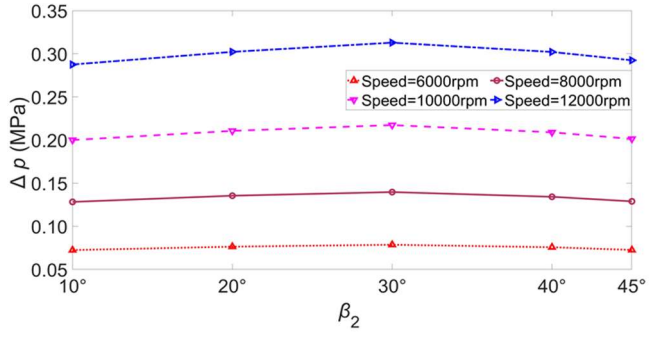


Figure 17: Effects of β_2 on Δp under different speeds.

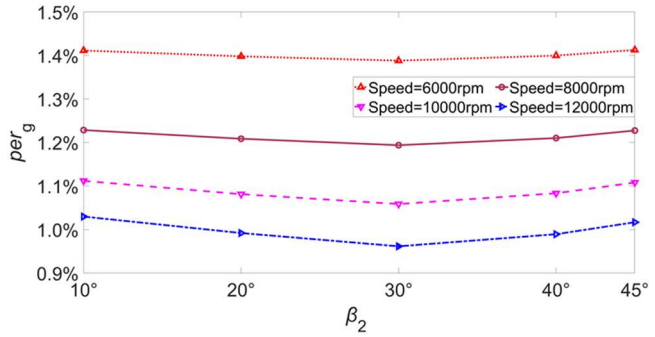


Figure 18: Effects of β_2 on per_g under different speeds.

4 Conclusion

To investigate the influence of impeller structure parameters on the cavitation of high-speed APPs, this paper established a CFD model of high-speed APP with impeller, and conducted an orthogonal test analysis to obtain the optimized impeller structure parameter combination. Then, based on the CFD model, the effects of different structural parameter changes on cavitation were analyzed independently. Furthermore, the effects of impeller structural parameter changes on cavitation under different pressure and speed conditions were analyzed. Simulation results show that

(1) The blade outlet width has the greatest effect on cavitation, followed by the blade outlet angle, while the impeller inlet diameter has the least effect.

(2) There are optimum impeller inlet diameter and blade outlet angle to optimize the impeller cavitation suppression effect. The design of the blade outlet width needs to consider the balance of weight and performance.

(3) The effects of speed on the impeller cavitation suppression effect are much higher than those of pressure.

This paper studies the effects of impeller structural parameters on the cavitation suppression effect in high-speed APPs. The effects of impeller structural parameters on the cavitation phenomenon can be utilized to guide the optimization design of the impeller. However, this paper still has limitations. On the one hand, only three key impeller structural parameters are considered, while other parameters are not considered. On the other hand, only the effect of the impeller is considered, while the coupling effect of the impeller-diffuser is not considered. In the future, we will further focus on the effects of diffuser structural parameters on the cavitation suppression effect and carry out the joint optimization design of the impeller and diffuser.

Acknowledgement

This research was funded by the National Natural Science Foundation of China (Grant No. U2141209), the Aeronautical Science Foundation of China (Grant No. 20182876003 and 20240028076003), and the Natural Science Foundation of Zhejiang Province (Grant No. LQ23E050013).

Corresponding authors: Hao Liu: haoliu7850052@zju.edu.cn; Xiaoping Ouyang: ouyangxp@zju.edu.cn.

Symbols	Parameters	Units
b_1	Blade inlet width	mm
b_2	Blade outlet width	mm
D_1	Blade inlet diameter	mm
D_2	Impeller outlet diameter	mm
D_j	Impeller inlet diameter	mm
d_h	Hub diameter	mm
β_1	Blade inlet angle	°
β_2	Blade outlet angle	°
φ	Wrap angle	°
ρ	Oil density	kg/m^3
v	Oil velocity	m/s
$\Omega(t)$	Control volume as a function of time	/
σ	Surface of the control volume	/
$\vec{n}$	Surface normal of σ	/
μ_t	Turbulent dynamic viscosity	Pa·s
$\vec{v}$	Fluid velocity	m/s
$\overrightarrow{v_\sigma}$	Surface motion velocity	M,/s
p	Static pressure	MPa
f_v	The mass fraction of the vapor	/
D_v	Diffusivity of the vapor	Kg/(m^3·s)
σ_f	Turbulent Schmidt number	/
R_e	The generation rate of vapor	Kg/(m^3·s)
R_c	Condensation rate of vapor	Kg/(m^3·s)
p_V	Liquid vapor pressure	MPa
ρ_l	Liquid density	kg/m^3
ρ_V	Vapor density	kg/m^3
c_e	Coefficients of R_e	/
c_c	Coefficients of R_c	/
f_g	Mass fraction of non-condensable gas	/
f_d	Mass fraction of the dissolved gas	/
f_{deq}	Equilibrium mass fraction of the non-condensable gas at the local pressure	/
D_d	Diffusivity of the dissolved gas	m^2/s
τ	Gas absorption/release time	*s*

$p_{deq,ref}$	Reference pressure for the dissolved gas equilibrium mass fraction	MPa
ρ_{1r}	Liquid density at reference pressure	kg/m^3
W	The molecular weight of the gas	g/mol
R	Universal gas constant	J/ (kg·K)
T	Thermodynamic temperature	K
α_g	Non-condensable gas volume fraction	/
α_V	Vapor volume fraction	/

References

[1] Q. Chao, JH. Zhang, B. Xu, HP Huang, M. Pan. A Review of High-Speed Electro-Hydrostatic Actuator Pumps in Aerospace Applications: Challenges and Solutions. Journal of Mechanical Design, 2019, 141(5): 050801. DOI:10/gh6vjt.

[2] B. Zhao, WW. Guo, L. Quan. Cavitation of a Submerged Jet at the Spherical Valve Plate/Cylinder Block Interface for Axial Piston Pump. Chinese Journal of Mechanical Engineering, 2020, 33(1): 67. DOI:10/gnzd7c.

[3] MY. Hu, Y. Jiang. Research on the High Speed of Piston Pumps Based on Rapid Erecting of Launch Vehicles. Applied Sciences, 2023, 13(12): 7178. DOI:10.3390/ app13127178.

[4] Y. Shi, T. Lin, G. Meng and J. Huang. A study on the suppression of cavitation flow inside an axial piston pump, 2016 Prognostics and System Health Management Conference, Chengdu, China, 2016, pp. 1-5, DOI: 10.1109/PHM.2016.7819843.

[5] ZG. Sun, YD. Li, Ning Liang, Hongmin Zhong. Study on Restraining Cavitation of Axial Piston Pump Based on Structure of Cylinder Block and Valve Plate Triangular Throttling Groove. Shock and Vibration, 2022: e6918936. DOI:10.1155/2022/6918936.

[6] B. Zhang, C. Zhao, HC. Hong, GZ. Cheng, HY Yang, SB. Feng, J. Zhai, WH. Xiao. Optimization of the outlet unloading structure to prevent gaseous cavitation in a high-pressure axial piston pump. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 2022, 236(7): 3459-3473. DOI:10.1177/09544062211043129.

[7] JH. Chen, SR. Guo, YL. Lu. Long Life Design Technology of Civil Aircraft Engine Driven Pump. Chinese Hydraulics & Pneumatics, 2021,45 (08) :160-170.

[8] L. Tao. Study on bootstrap reservoir type pressurized system for civil aircraft hydraulic supply system, 2016 IEEE International Conference on Aircraft Utility Systems (AUS), Beijing, China, 2016, pp. 1117-1121, DOI: 10.1109/AUS.2016.7748226.

[9] SR. Guo, JH. Chen, YL. Lu, Y. Wang, HK. Dong. Hydraulic piston pump in civil aircraft: Current status, future directions and critical technologies. Chinese Journal of Aeronautics, 2020, 33(1): 16-30. DOI:10/gpk79t.

[10] JH. Chen. Self-pressurized study of hydraulic piston pump based on the centrifugal turbine. Mech Eng 2012,(7):18–20.

[11] W. Li XF. Zhao, WQ. Li, WD. Shi, LL . Ji, L. Zhou. Numerical Prediction and Performance Experiment in an Engine Cooling Water Pump with Different Blade Outlet Widths． Mathematical Problems in Engineering, 2017, 4:8945712.1-8945712.11

[12] J. Chen, Y. Zheng, L Zhang, X. Chen, D. Liu, Z Xiao. Influence Analysis of Runner Inlet Diameter of Hydraulic Turbine in Turbine Mode with Ultra-Low Specific Speed. . Energies 2023, 16, 7086. https://doi.org/10.3390/en16207086

[13] Y. Long, H. Wang, R. He, Q. Fu, R. Zhu. Engineering. Effect of Induced Wheel and Impeller Inlet Diameter on the Hydrodynamic Performance of High-speed Centrifugal Pumps. Journal of Applied Fluid Mechanics, 2024, 17(11): 2424-2447. DOI:10.47176/jafm.17.11.2497.

[14] G. Peng, Q. Chen, L. Zhou, B. Pan, Y. Zhu. Effect of Blade Outlet Angle on the Flow Field and Preventing Overload in a Centrifugal Pump. Micromachines, 2020, 11(9): 811. DOI:10.3390/mi11090811.

[15] HC. Ding, T. Chang, FY. Lin. The Influence of the Blade Outlet Angle on the Flow Field and Pressure Pulsation in a Centrifugal Fan. Processes, 2020, 8(11): 1422.

[16] C. Guo, M, Gao, JY. Wang, YT. Shi, SY. He. The effect of blade outlet angle on the acoustic field distribution characteristics of a centrifugal pump based on Powell vortex sound theory. Applied Acoustics,2019,155 (12):297-308 DOI:10.1016/j.apacoust.2019.05.031.

[17] W. Li, W. Pu, LL. Ji, X. Zhang, XR. He. Hydraulic structure optimization of centrifugal pump based on orthogonal test. Journal of Physics: Conference Series, 2024, 2707(1): 012056. DOI:10.1088/1742-6596/2707/1/012056.

[18] YP. Wang, XF. Deng, ML. Geng, RL. Tao, Y. Li, ZC. Zhu. Influence of blade wrap angle on internal flow and wear performance of solid-liquid two-phase centrifugal pump with semi-open impeller. Advances in Mechanical

Engineering, 2024,16(6):1-10. DOI: 10.1177/16878132241258827.

[19] MG Tan, HL Liu, SQ Yuan, Y Wang, W Kai. Effects of Blade Outlet Width on Flow Field and Characteristic of Centrifugal Pumps//ASME 2009 Fluids Engineering Division Summer Meeting. American Society of Mechanical Engineers Digital Collection, 2010: 51-60.

[20] T. Shigemitsu, J. Fukutomi, R. Nasada, K. Kaji. The effect of blade outlet angle on performance and internal flow condition of mini turbo-pump. Journal of Thermal Science, 2011, 20(1): 32-38. DOI: 10.1007/s11630-011-0431-3.

[21] Y. Wang, L. Zhou, S. Zheng. Multi-Objective Optimization Design of Low Specific Speed Centrifugal Pumps Based on Genetic Algorithm. IEEE Access, 2023, 11: 97896-97908. Doi: 10.1109/ACCESS.2023.3311893.

[22] J. Pei, WJ. Wang, SQ. Yuan, JF. Zhang. Optimization on the impeller of a low-specific-speed centrifugal pump for hydraulic performance improvement. Chinese Journal of Mechanical Engineering, 2016, 29(5): 992-1002. DOI: 10.3901/CJME.2016.0519.069.

[23] J. Zhao, J. Pei, J. Yuan, W. Wang. Energy-saving oriented optimization design of the impeller and volute of a multi-stage double-suction centrifugal pump using artificial neural network. Engineering Applications of Computational Fluid Mechanics, 2022, 16(1): 1974-2001. DOI: 10.1080/19942060.2022.2127913.

[24] YL. Chen, JW. Chen, YY. Pan, MY. Yan. Optimization of turbocharging system for aviation piston pump based on entropy production theory. Acta Aeronautica et Astronautica Sinica, 2024,45(04):286-298.

[25] HK. Dong, YL. He, Y. Wang, GY. Kou. Numerical investigation of effect of a centrifugal boost impeller on suction performance of an aircraft hydraulic pump Chinese Journal of Aeronautics, 2022, 35(8): 236-248.DOI: 10.1016/j.cja.2021.07.016.

[26] E. Frosina; D. Buono, A. Senatore. A Performance Prediction Method for Pumps as Turbines (PAT) Using a Computational Fluid Dynamics (CFD) Modeling Approach. Energies 2017, 10, 103. DOI: https://doi.org/10.3390/en10010103

[27] A K. Singhal, M M. Athavale, HY. Li, Y. Jiang. Mathematical Basis and Validation of the Full Cavitation Model. Journal of Fluids Engineering, 2002, 124(3): 617-624. DOI:10.1115/1.1486223.

[28] E. Frosina; G. Marinaro, A. Senatore. Experimental and numerical analysis of an axial piston pump: A comparison between lumped parameter and 3D CFD approaches. Proceedings of the ASME-JSMEKSME 2019 8th joint fluids engineering conference, San Francisco, USA, DOI: https://doi.org/10.1115/AJKFluids2019-5406

Simulative controller evaluation for an electro-hydrostatic excavator actuator with load-holding capability using active hysteresis and throttling control

Felix Figge, Katharina Schmitz

Institute for Fluid Power Drives and Systems (ifas), RWTH Aachen University, Aachen, Germany
E-mail: felix.figge@ifas.rwth-aachen.de, sc@ifas.rwth-aachen.de

Abstract

In recent years, the electrification of mobile machinery has progressed. Due to the change of the primary energy source, the hydraulic system must be more efficient and sustainable. In this context, the use of decentralized electro-hydrostatic actuators (EHA) is highly promising. Still, a widespread use cannot be observed. Reasons for this might be the higher initial component costs, worse dampening, as well as higher expenditure in the implementation of the required controls.

This paper presents a simulative evaluation of an electro-hydrostatic actuator with a hydraulic high-speed unit for use in mobile machinery. The circuit layout was selected based on cost considerations and includes load-holding capabilities. Furthermore, the load-holding valves are a key aspect of the control scheme. They are used to introduce a small active throttling at the discharge side of the cylinder in both pumping quadrants, thereby extending their load pressure range and enabling full operational coverage of the four-quadrants. The developed controller is mainly based on two state machines. It aims to synchronize the switching times of the valves with motor speed to enhance system performance and prevent system excitation. The controller is implemented in Matlab Simulink. For the aspired evaluation, it was coupled with a lumped parameter model of the circuit. The simulated cylinder velocities for several different speed and load combinations are evaluated to determine the control performance. Special focus is placed on the quadrant changes.

Keywords: EHA, four-quadrants, electrification, active throttling, hysteresis control

1 Introduction

In recent years, the electrification of mobile machinery has progressed. A wide range of electrified vehicles, including larger ones, is now available. The change of the prime mover from a combustion engine to an electric motor can be considered the most important change in mobile hydraulics of the past decades. Historically, many mobile machines use a centralized hydraulic pressure supply, which distributes the power via valves and fluid lines to the actuators. While this approach is cost-effective in terms of initial costs, it results in high valve throttle losses. These are mainly a result of the various actuators requiring different pressure levels. The recent focus on energy saving and the limited amount of energy available in machinery with batteries requires the hydraulic system to be more efficient and sustainable.

In this context, the use of decentralized electro-hydrostatic actuators (EHA) is highly promising. The use of a dedicated pump and motor for an actuator avoids the throttle losses from a higher to a lower pressure level. While EHAs are more often considered for industrial applications, widespread use in mobile machinery cannot be observed. Reasons for this might be the higher initial component costs, higher build volume, worse dampening, directly linked to the reduced system losses, as well as higher expenditure in the implementation of the required controls. The higher component costs of EHAs in comparison to a load-sensing system currently mainly arise from the additional prime movers as well as the additional hydraulic units. Furthermore, a decentralized hydraulic power supply with EHAs requires a small build volume to allow for easy integration into existing mobile machinery. An increase in rotational speed reduces the volume of the electric motor and allows the use of a pump with smaller displacement volume, reducing the total build volume of the EHA even further. Therefore, the underlying research project aims to develop a high-speed closed-circuit EHA for the application in mobile machinery.

Closed-circuit EHAs using a single pump are known to suffer from undesirable performance during certain operating conditions. Especially changes between operating quadrants are critical. Often, pressure feedback is used to increase the overall dampening of the system [1,2]. Williamson and Ivantysynova [3] developed a predictive observer to provide sufficient lead time for feed-forward control of pump

displacement in a skid-steer and allow four-quadrant operation. The approach was aimed at dealing with excessive phase lag, rendering feedback control ineffective at handling mode-switching oscillations. They later concluded that the main problem is the interaction between mass and the compressible fluid that acts like a mass-spring-damper [1]. The unequal actuator areas aggravate any excitation because the pump still provides the same volume flow, sending the system back and forth between quadrants [1,2]. Increasing the system dampening is therefore recommended and pressure feedback is implemented. Michel further calculated a maximum allowable acceleration to prevent mode switches [2]. Wang et. al. [4] used electrically actuated proportional valves in combination with a control algorithm to introduce leakage in critical operation conditions for increased dampening and to prevent oscillation. The valves were added in addition to a shuttle valve. The compensation is only active close to critical operation conditions at low loads. They also use the valves to ensure the accumulator is connected to one cylinder side in case of oscillations of the shuttle valve. A mechanical variant of the added leakage in the critical region was presented by Çalışkan et. al. [5]. They used an underlapped shuttle valve, not requiring additional electrical valves, sensors and control algorithms. Imam et. al introduced a new valve type they called "limited throttling valve" to prevent oscillation in the critical region of a circuit using two POCV [6,7]. The valve is designed to restrict the cylinder flow only at low loads near the critical working conditions. Their simulation results show no oscillations in all operating conditions. They also show that the critical zone and oscillation amplitudes can be reduced by applying different charge pressures to both cylinder sides [7]. Costa and Sepehri [8] introduced a new four-quadrant definition based on the cylinder load pressure instead of the pressure difference over the pump. They concluded that this was the main reason for the poor behavior of earlier circuits. They also introduced a new circuit to account for their new quadrant definition. However, Gøytil et al. [9] showed that the proposed circuit can suffer from unwanted mode switches under certain conditions due to dynamic considerations. They suggest using an observer to calculate a modified force estimate to decide the valve states, eliminating the unwanted mode switches. The observer excludes forces introduced by valve switches and acceleration. Their strategy requires a certain accumulator pressure (in their case 30 bar) to prevent cavitation, which is inappropriate for certain circuits because of pressure limits. Their approach is software-based, while previous works primarily focused on hydraulic implementations. Kärnell and Ericson [10] investigated hysteresis control to prevent unwanted mode switches. They concluded that the pressures should be sensed in such a way that the difference is unchanged after a mode switch. Even then, switches can occur, but active hysteresis control reduces that risk if reservoir pressure is sufficient.

The approach presented in this paper uses the load pressure definition by [5,8] and implements active throttling [7,10] to allow full coverage of the four-quadrant plane in an excavator. Further, to reduce mode switches due to small pressure oscillations, active hysteresis is added [10]. The control approach further tries to minimize cylinder velocity spikes during quadrant changes by decoupling the cylinder from the pump by use of the incorporated load-holding valves.

Following, the selected circuit is introduced and its operation boundaries are derived. Afterwards, the developed control scheme is presented and validated by use of a lumped parameter simulation.

2 Circuit

There are many circuits for EHAs known in literature. An overview including a classification was introduced by Ketelsen et. al. [11]. Regarding the underlying project aiming to integrate the EHA into mobile machinery, especially excavators, the following restrictions excluded certain circuit layouts. In general, components with fewer and simpler parts cost less. Therefore, it was decided to use internal gear pumps with fixed displacement, which have fewer and simpler parts than piston pumps. Initially, circuits using more than one pump were considered to allow for a larger variety of layouts. The use of a fixed displacement hydraulic unit requires a variable speed prime mover for flow adjustment. Since variable speed prime movers are expensive, only one should be used. The EHA shall be encapsulated with only electrical and mechanical connections, requiring some sort of internal reservoir to store the differential fluid volume of the cylinder. The application in mobile machinery with uncertain operation conditions requires the capability to operate in all four quadrants. To achieve high efficiency, the EHA shall be capable of recuperating energy. Furthermore, the integration of load-holding valves was deemed to be necessary to allow a full stop of the prime mover and prevent overheating as well as wear on the hydraulic unit. Based on these restrictions, four circuits were selected for closer investigation. The circuits were introduced in [12] and are using a single fixed displacement pump with one exception. The first circuit uses pilot-operated check valves (POCV) to accommodate for the differential volume flow of the cylinder [13] as well as for load-holding [14]. The second circuit uses a cylinder with multiple chambers, allowing for a structural design close to a differential cylinder with equal rod and piston areas [15]. The differential volume flow is thereby omitted. The third circuit uses two pumps with a displacement ratio equal to the cylinder area ratio to provide appropriate volume flows [16]. The last circuit uses electrically actuated valves to handle the differential volume flow [9] as well as for load-holding. This circuit was selected for further investigation and will be regarded in the following sections

2.1 Selected circuit layout

The circuit shown in fig. 1 was selected because of the use of only a single hydraulic unit, the use of a standardized cylinder and the high customizability of the software control. The circuit mainly consists of the variable high-speed prime mover, a four-quadrant internal gear unit, as well as the differential cylinder. An accumulator is required to store the differential volume of the retracted cylinder rod length as well as to provide sufficient charge pressure for the hydraulic unit, especially at high speeds. To allow the differential flow to return to the accumulator, two electric switching valves $Diff_{rod}$ and $Diff_{pist}$ are included. They are called differential valves in the following. The electric actuation allows precise

control over the timing of the connection between the pump ports and the accumulator. The load-holding valves LH_{rod} and LH_{pist} are designed to be proportionally adjusted. The oil flowing back into the accumulator is filtered using two additional check valves CV_{acc} and CV_{filter}. The check valves in parallel to the load-holding valves CV_{LH} and differential valves are required to prevent cavitation of the pump and the cylinder. During the circuit design, it was found that, based on available valves, separate oversized check valves were required to keep the pressure drop low enough to prevent cavitation. This was necessary because of the pressure level of the accumulator, which is limited to 10 bar because of the shaft sealing of the hydraulic unit. While an electric motor with a rotor running in oil can solve this by not needing a shaft seal, it introduces higher shear losses.

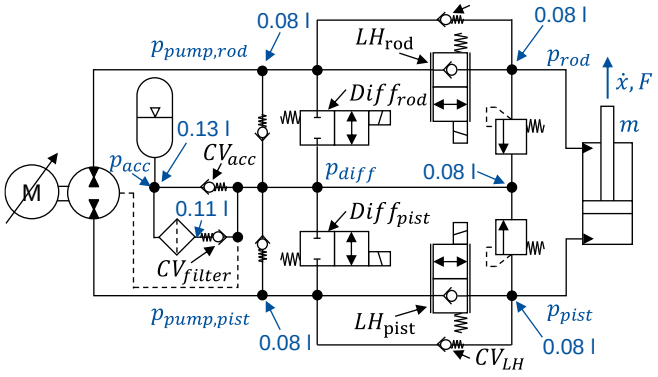


Figure 1: Circuit layout, pressures and fluid volumes

The operation of a cylinder can be divided into the well-known four-quadrants based on the cylinder velocity as well as the cylinder load, as shown in fig. 2. While the cylinder velocity is straightforward, different approaches for the load definition exist. They range from the pressure difference between the pump ports, the pressure difference between the cylinder chambers and the load pressure following eq. (1). In eq. (1) α represents the cylinder area ratio. The main advantage of the load pressure definition is the consideration of all active forces on the cylinder and only these forces, as can be seen in eq. (2). Therefore, the load pressure as defined by [5,8] is used as the second criterion for the quadrant definition. The external force F in fig. 2 is only used for clarification of load direction and does not represent the used criterion.

$$p_{load} = p_{pist} - \alpha \cdot p_{rod} \tag{1}$$

$$\ddot{x}m - F - F_{fric} = p_{pist} \cdot A_{pist} - p_{rod} \cdot A_{rod} = p_{load} \cdot A_{pist} \tag{2}$$

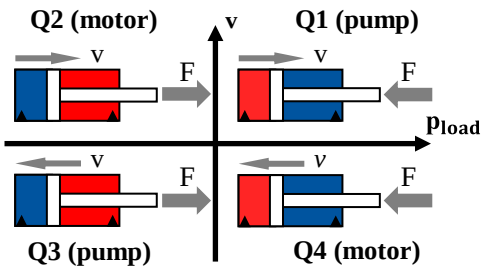


Figure 2: Four-quadrant plane

In fig. 3, the actuation of valves and oil flow in all four quadrants of operation is shown. Active valves are denoted with a green coil symbol and the oil flow with orange arrows. A simplified circuit layout with integrated check valves is used for clarity. Following the quadrant definition from fig. 2 the cylinder moves outwards in Q1 and Q2 (dark grey arrows) and in Q3 as well as Q4 inwards. Q1 and Q3 are the pumping quadrants in which the hydraulic unit is supplying energy into the circuit, marked with a light blue arrow towards the hydraulic unit. In Q2 and Q4 the unit works as a motor and extracts energy. Quadrant changes based on a reversal of cylinder movement (Q1↔Q4 or Q2↔Q3) are controlled by a change of motor speed and can therefore be more easily handled by the control scheme.

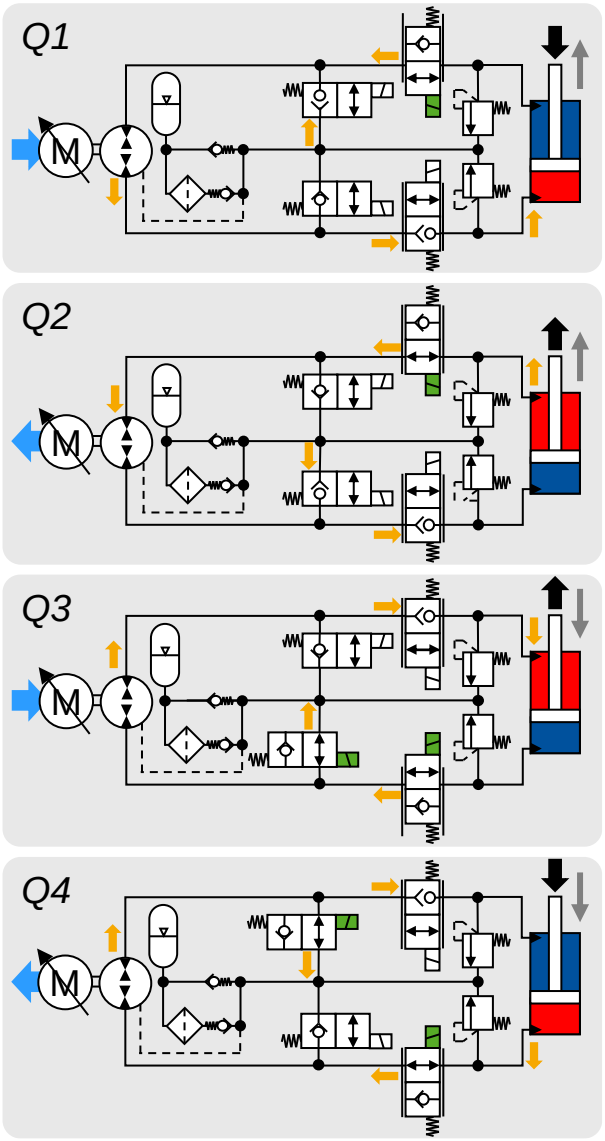


Figure 3: Valve actuation and oil flow in all four quadrants

A change of operating quadrant based on a direction reversal of the external load (Q1↔Q2 or Q3↔Q4) is more difficult. The exact time of direction change cannot be predicted. In both cases, it requires a change of motor speed because of the different cylinder areas. The speed change ideally should happen instantaneously but is limited by motor torque and

inertia. The positions of the load-holding valves need to be adjusted as well. Additionally, for the changes between Q3 and Q4, a change of differential valve actuation is required, as can be seen in fig. 3. To avoid poor system performance due to unwanted system states, the valve actuation must be synchronous to the motor speed change.

Before going into detail about how these challenges were tackled, the circuit behavior is explained in more detail and the existing boundaries of operation are derived.

2.2 Circuit behavior and boundaries

A similar concept to the one in this section was investigated in [6] for a circuit without load-holding valves and differential flow handling by pilot-operated check valves.

First, the effect of the accumulator pressure is investigated. Neglecting all pressure drops across valves as well as all other forces, the accumulator pressure results in an active pushing force on the cylinder rod. The cylinder is moving outwards without energy input from the pump, corresponding to an operation in Q2. The related load pressure results by inserting the accumulator pressure p_{acc} into eq. (1) for both cylinder pressures (eq. (3)).

$$p_{bound,acc} = p_{acc} \cdot (1 - \alpha) \tag{3}$$

$p_{bound,acc}$ is larger than zero for all accumulator pressures larger than zero and thereby represents a new boundary between motor and pump operation. This can be checked for the other quadrant boundaries by considering an external force that pushes the cylinder inwards. The cylinder will only move inwards without assistance from the pump (Q4) if the external force is larger than the outwards pushing force from the accumulator pressure. This in turn would result in a load pressure that is larger than $p_{bound,acc}$, since p_{pist} would rise. The new boundary is shown in fig. 4 as a blue rectangle for an accumulator pressure between 4.6 bar and 10 bar, since the accumulator pressure changes during charging and discharging.

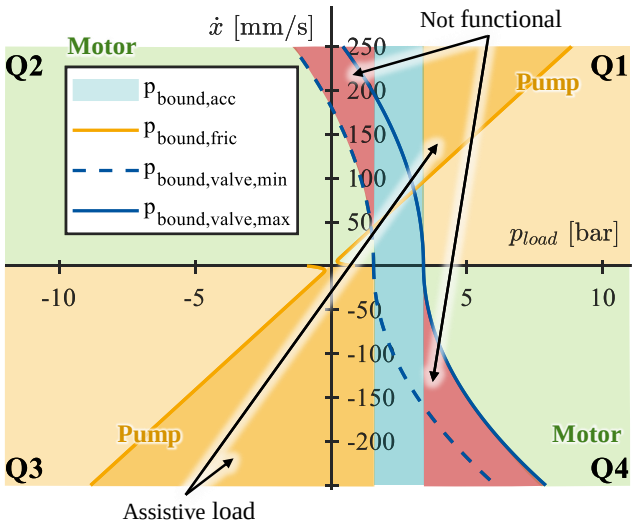


Figure 4: Load pressure boundaries of the circuit

Next, the quadrants Q1 and Q3 are considered, in which the pump provides power. If the pressure drop across the load-holding valves is neglected, stationary movement of the cylinder is assumed and no external force is applied, the load pressure originates only from the cylinder friction (see eq. (2)). The neglection of the pressure drop across the valves is not strictly necessary. A pressure increase on the discharging cylinder side due to the valve resistance would be canceled out in the load pressure calculation by the necessary pressure increase on the inflow side due to the force balance. Except for an increase in pump pressure, this has no further implications. The load pressure required for motion with a specific velocity $p_{bound,fric}$ follows from eq. (4) in Q1 and (5) in Q3 for load-free condition. The corresponding boundaries are drawn in orange in fig. 4 and are calculated by means of the friction model used in the simulation as defined by eq. (6) and the parameters given in tab. 1.

$$p_{bound,fric,Q1} = \frac{|F_{fric}|}{A_{pist}} \tag{4}$$

$$p_{bound,fric,Q3} = -\frac{|F_{fric}|}{A_{pist}} \tag{5}$$

$$F_{fric} = \begin{cases} F_{fric,base} + \mu_{vis} \cdot \dot{x} + F_{fric,mix} & , \ \dot{x} \le \dot{x}_{vis} \\ F_{fric,base} + \mu_{vis} \cdot \dot{x} & , \ \dot{x} > \dot{x}_{vis} \end{cases}$$
$$F_{fric,mix} = \left(F_{fric,static} - F_{fric,base}\right) \cdot \left(1 - \frac{|\dot{x}|}{\dot{x}_{vis}}\right)^4 \tag{6}$$

If the external force is pushing in Q1 and therefore has a negative sign, the load pressure will be higher than in the load-free condition, as stated by eq. (7). Therefore, a load pressure below the load-free condition can only be achieved if the external force assists the motion. The same conclusion can be made for Q3 by using eq. (8) with the only difference that an assistive load results in a higher load pressure due to the reversed movement direction. Both areas are marked in a darker orange shade in fig. 4. The boundary thereby can be used to determine if the external force is assistive, provided the friction force is known with sufficient accuracy. The only implication function-wise is that the load pressure will increase along the boundary line if the velocity is increased and no assistive load is applied. All the velocities inside the quadrant can still be achieved.

$$0 \le -F = p_{load} \cdot A_{pist} - |F_{fric}|$$
$$\Rightarrow p_{load} \ge \frac{|F_{fric}|}{A_{pist}} \tag{7}$$

$$0 \ge -F = p_{load} \cdot A_{pist} + |F_{fric}|$$
$$\Rightarrow p_{load} \le -\frac{|F_{fric}|}{A_{pist}} \tag{8}$$

Lastly, the quadrants with motor operation are investigated. Contrary to the previous case, energy is now provided by the cylinder and not the pump. The volume flow through the load-holding valves can be described by the orifice equation (9), which is further related to the cylinder velocity by eq. (10) considering the currently loaded cylinder side. $A^*(s)$ is the

active flow area depending on the valve position *s*. One can directly conclude that the cylinder velocity will be limited for low pressure differences and therefore small external loads.

$$Q = \alpha_D \cdot A^*(s) \cdot \sqrt{2/\rho} \cdot \sqrt{\Delta p} = K(s) \cdot \sqrt{\Delta p} \qquad (9)$$

$$\dot{x} = \frac{Q}{A_{pist/rod}} = \frac{K(s) \cdot \sqrt{\Delta p}}{A_{pist/rod}} \qquad (10)$$

Therefore, the cylinder velocity will not be controlled by the rotational speed of the pump but by the orifice flow at low loads. The pressure at the pump inlet port will then be close to the accumulator pressure p_{acc} because the connected check valve opens to prevent cavitation. The pump outlet pressure is also close to p_{acc} because the corresponding differential valve is actuated and connects to the accumulator. Further, the unloaded cylinder chamber pressure is also at the same level because the check valve opens to prevent cavitation. Depending on the operating quadrant, the loaded cylinder side changes, resulting in eq. (11) for Q2 and eq. (13) for Q4 to calculate the boundary load pressure $p_{bound,valve}$ following the explanations beforehand. By rearranging eq. (11) and (13) to eq. (12) and (14) a relation between the pressure drop Δp and $p_{bound,valve}$ is derived. The second half of eq. (12) and (14) is equal to $p_{bound,acc}$, resulting in a dependence on the accumulator pressure. By inserting eq. (12) and (14) into eq. (10) a correlation between cylinder velocity and load pressure is derived, which is shown in fig. 4 as blue lines. These lines mark the needed load pressure $p_{bound,valve}$ to achieve the respective cylinder velocity. The solid line results for the maximum and the dashed line for the minimum accumulator pressure. If the load pressure falls "below" the current boundary, taking into consideration the different signs in Q2 and Q4, the velocity will sink along the boundary.

$$p_{bound,valve,Q2} = p_{acc} - \alpha \cdot (p_{acc} + \Delta p) \qquad (11)$$

$$\Delta p_{Q2} = -\left[p_{bound,valve} - (1-\alpha) \cdot p_{acc}\right] \cdot 1/\alpha \qquad (12)$$

$$p_{bound,valve,Q4} = p_{acc} + \Delta p - \alpha \cdot (p_{acc}) \qquad (13)$$

$$\Delta p_{Q4} = +\left[p_{bound,valve} - (1-\alpha) \cdot p_{acc}\right] \qquad (14)$$

In conclusion, an operation in motor mode is only possible above the boundary. The corresponding velocities and load pressures are marked green in fig. 4. By closer inspection, two areas between $p_{bound,acc}$ and $p_{bound,valve}$ remain, marked red in fig. 4, in which an operation of the circuit is not possible. The assistive load is not large enough to achieve the required volume flow through the load-holding valve and the cylinder velocity remains too low. Valves with smaller pressure drops reduce the area but do not remove the boundary completely. This behaviour is unacceptable for the operation in mobile machinery since it would result in a large and prolonged deviation from the control signal.

An extension of Q1, respectively Q3, and therefore the supply of energy from the pump in only these areas is theoretically possible. It requires the throttling of the cylinder discharge to allow the pump to control the cylinder velocity by providing volume flow in the otherwise unloaded cylinder chamber. Without the throttling, the cylinder would accelerate due to the external load. The valve characteristics need to be known precisely for this approach, which they generally are not. Further, they change during operation, for example, due to temperature. Assuming the control supposes a pressure drop that is too high, this would result in a larger volume flow near the boundary than expected. If the load pressure lies inside the red area. the pump would be providing energy by pumping into the cylinder. The higher discharge from the cylinder would result in a higher speed. There are only two ways to slow down the cylinder in this scenario. One would need the hydraulic unit to be working in motor mode and use it to control the cylinder velocity. This contradicts the earlier assumption that the unit is pumping. Secondly, the pressure drop across the valve could be increased to lower the volume flow out of the cylinder. Since the valve characteristics are not known precisely enough, as supposed at the beginning, a closed-loop control would be required. The required position sensor would increase the cost of the EHA and therefore, this approach was discarded.

If an extension of the pumping quadrants further than the red marked areas is allowed, a simpler approach, following the same principle, is possible. The valve is not actuated to match the pressure drop to the currently required volume flow, but to achieve a fixed pressured rop. This was discussed in [10] as a possible way to introduce hysteresis to prevent unwanted quadrant changes and was called "active throttling". No further investigation was carried out. A familiar approach was also introduced in [7] as "selective throttling" with specifically designed valves that increase throttling in the critical region to reduce oscillations. As was discussed, active throttling is needed to allow the investigated circuit with load-holding valves to cover the full operational range and not to introduce hysteresis or reduce oscillations. The fixed pressure drop must be larger than the pressure drop of the load-holding valves at the highest volume flow and full opening. Otherwise, the required volume flow at high cylinder velocity cannot be achieved. The fixed pressure drop results in a larger area in which the hydraulic unit can be operated in pumping mode without loss of control over the cylinder velocity. The new boundaries are calculated by eq. (15) for Q1 and eq. (16) for Q3. Exactly at these boundaries, the external load accelerates the cylinder and the volume flow provided by the pump becomes insufficient. The check valve opens to prevent cavitation, resulting in a pressure level close to p_{acc}. The pressure in the other cylinder chamber at the boundary equals p_{acc} plus the artificial pressure difference $\Delta p_{throttle}$. Again, the accumulator boundary $p_{bound,acc}$ is part of the equations resulting in a maximum and minimum boundary drawn as green lines in fig. 5 for $\Delta p_{throttle}$ set to 22 bar. While a reduction of $\Delta p_{throttle}$ above $p_{bound,acc}$ is possible, this is discarded to allow for better cylinder control.

$$p_{pist} = p_{acc}\,; \quad p_{rod} = p_{acc} + \Delta p_{throttle} \\ \Rightarrow p_{bound,throttle,Q1} = p_{acc} \cdot (1-\alpha) - \alpha \cdot \Delta p_{throttle} \qquad (15)$$

$$p_{pist} = p_{acc} + \Delta p_{throttle}\,; \quad p_{rod} = p_{acc} \\ \Rightarrow p_{bound,throttle,Q3} = p_{acc} \cdot (1-\alpha) + \Delta p_{throttle} \qquad (16)$$

The required control current to actuate the valve is taken from a characteristic diagram including the pressure difference, volume flow and current. The volume flow is calculated from the requested cylinder velocity. The boundary position is

influenced by the accumulator pressure, which could be compensated for via measurement. Still, the characteristic diagram in general will be inaccurate and deviations are unavoidable. By introducing a virtual boundary below $p_{bound,throttle}$ and above $p_{bound,valve}$ a safety gap can be introduced to account for deviations. Additionally, "active hysteresis" [10] is added to this switching boundary so that small changes in load pressure will not result in a change of detected quadrant. The introduced hysteresis Δp_{Hyst} is drawn in violet in fig. 5 with a width of 4 bar. For a positive cylinder velocity, the motor mode in Q2 starts if the load pressure is below -10 bar. Afterwards, the EHA will change back to pump mode (Q1) only if the load pressure increases above -6 bar. Negative load pressures are the result of higher pressures on the rod side as defined by eq. (1).

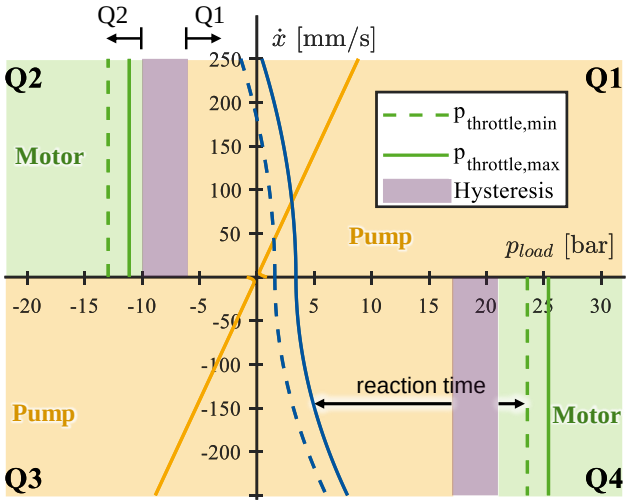


Figure 5: Load pressure boundaries of the circuit with active throttling and hysteresis

The positioning of the virtual boundary, including hysteresis, determines the amount of time before control of cylinder velocity is lost. These reaction times are exemplary marked in fig. 5 for Q4. For Q4→Q3 this timeframe starts when the load pressure falls below the left hysteresis boundary and ends when it reaches the valve boundary, since then the cylinder will start to slow down. The hysteresis boundary is not positioned in the middle because the control was developed for the arm of an excavator. Changes from Q1→Q2, as well as Q3→Q4, are unlikely during normal operation since the mass centre cannot cross the top centre position. The position of the virtual boundary and the width of the hysteresis are therefore dependent on the application and need to consider reaction time against the tendency of unwanted quadrant changes.

3 Control scheme

Based on the load pressure boundaries shown in fig. 5 a control scheme was implemented. The top level of the control is shown in fig. 6. The main component is the EHA controller that outputs the valve modes and required rotational speed depending on the loaded cylinder side. The speed request is limited and then adjusted by pressure feedback to reduce pressure oscillation [2]. Depending on the loaded cylinder side, one of the speed requests is fed into the characteristic field of the hydraulic unit, adjusted to compensate for volumetric losses and then output to the motor controller. The valve controllers output the current to actuate the load-holding valves based on different modes that are explained later. To reduce the number of pressure sensors p_{Diff} is used as the low-pressure level of the hydraulic unit, since the check valve is sized to have a low pressure drop. For clarity, the overload protection was left out, which increases the pressure drop over the load-holding valves during motor operation if the electric motor gets close to its torque limit. This is essential at high speeds to prevent an uncontrolled acceleration of the cylinder, since any acceleration reduces the torque capacity further.

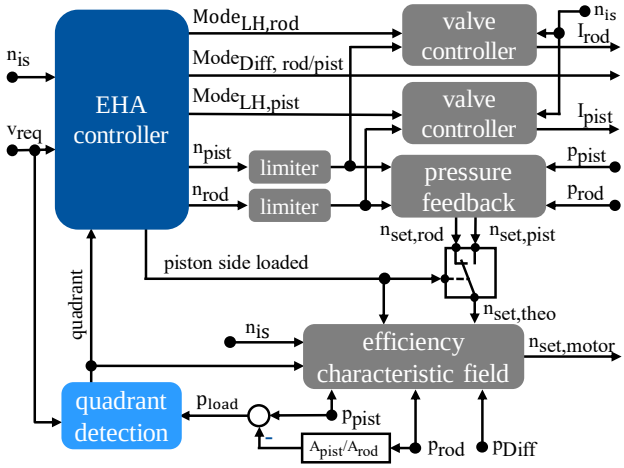


Figure 6: Top level of control scheme

The load pressure boundaries from fig. 5 are mainly implemented in the state machine for quadrant detection, as shown in fig. 7. In addition to the well-known four-quadrants, a fifth is added for the standstill of the EHA.

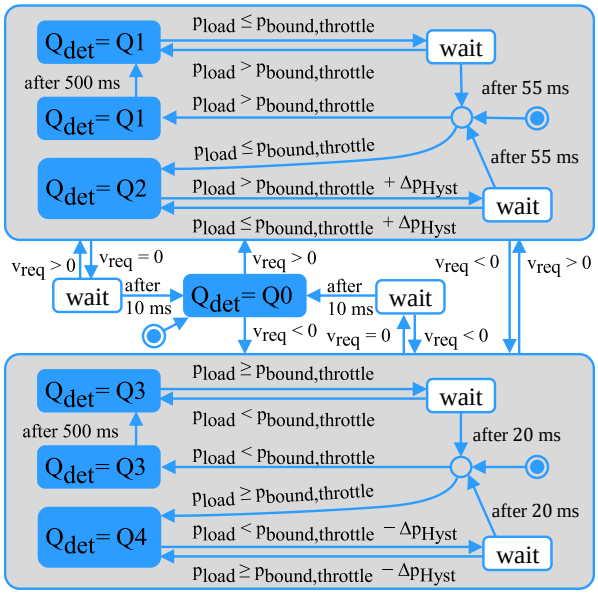


Figure 7: Quadrant detection

Based on the velocity request v_{req}, one of the substates is activated. Inside the substates the hysteresis boundaries for the

load pressure p_{load} are implemented. To further prevent unwanted quadrant changes because of short impulses, additional wait times are introduced before a quadrant is left. A longer wait time of 500 ms is added after entering each of the pumping modes to allow the initial excitation after valve actuation to subside and to prevent a quadrant change. Therefore, the states in which Q1 and Q3 are detected are duplicated. The times are chosen based on iterative testing and might not be optimal. It shall be pointed out that a request to change velocity direction or to stop will still lead to a quick change of quadrant.

Following, the EHA controller state machine is introduced, whose main function is the synchronisation of the valve positions and motor speed. For example, before the pump starts to provide volume flow into the cylinder, the load-holding valve on the discharging cylinder side needs to be opened. There is a delay or dead time before a control signal change takes effect, which must be accounted for to achieve better system performance. The symbols used are explained in more detail in the nomenclature. They comprise the times needed to change valves states, the time delay of the motor $t_{dead,Motor}$, the current rotational speed of the hydraulic unit n_{is} and of course the detected quadrant Q_{det}.

Figure 8 shows the sequences for acceleration and reversal of the cylinder velocity. Based on the detected quadrant Q_{det} different sequences are necessary for acceleration. If the external load is directed against the requested movement direction (Q1 & Q3), first, the load-holding valve blocking the discharge of the cylinder is opened completely ("1") as well as the differential valve on the piston side in case of Q3. After the maximum switching time minus the motor delay has elapsed, the motor signal is applied. After a wait time equal to the motor delay, the valve on the outlet is then changed to throttling mode. The valve is first fully opened to prevent hydraulic tensioning of the cylinder. If an assistive load acts on the cylinder, a quick opening of the load-holding valves would result in a strong acceleration because of the sudden drop of pressure in the cylinder chamber. Afterwards the volume surplus due to the higher cylinder velocity would result in a pressure spike at the motor inlet slowing the cylinder down again. This behaviour was deemed to be unwanted and the displayed approach implemented. By briefly turning the hydraulic unit in the opposite direction the volume between the motor inlet and LH valve is pressurized. The approach could be refined further by adapting the time the pump turns depending on the pressure level. Additionally, the valve current is increased slowly ("opening") to allow for smoother pressure equalization and acceleration of the cylinder. As soon as the valve should start to move, after $t_{dead,LH}$ passed, the pump direction is changed back to the correct one. After stopping the cylinder, which is explained later, it is possible that the rotational unit is still spinning. To prevent the start of the sequences with a high-speed rotation in the wrong direction, a check against a threshold n_{wait} was added.

The reversal of the cylinder movement under the same load direction can be handled well by the controller since it is mainly influenced by the rotational speed of the hydraulic unit. But especially when entering the motoring quadrants Q2 & Q4, a precise synchronisation is required to prevent the earlier described pressure drop because of the motor turning while the valve is still closed. The implemented approach is compatible

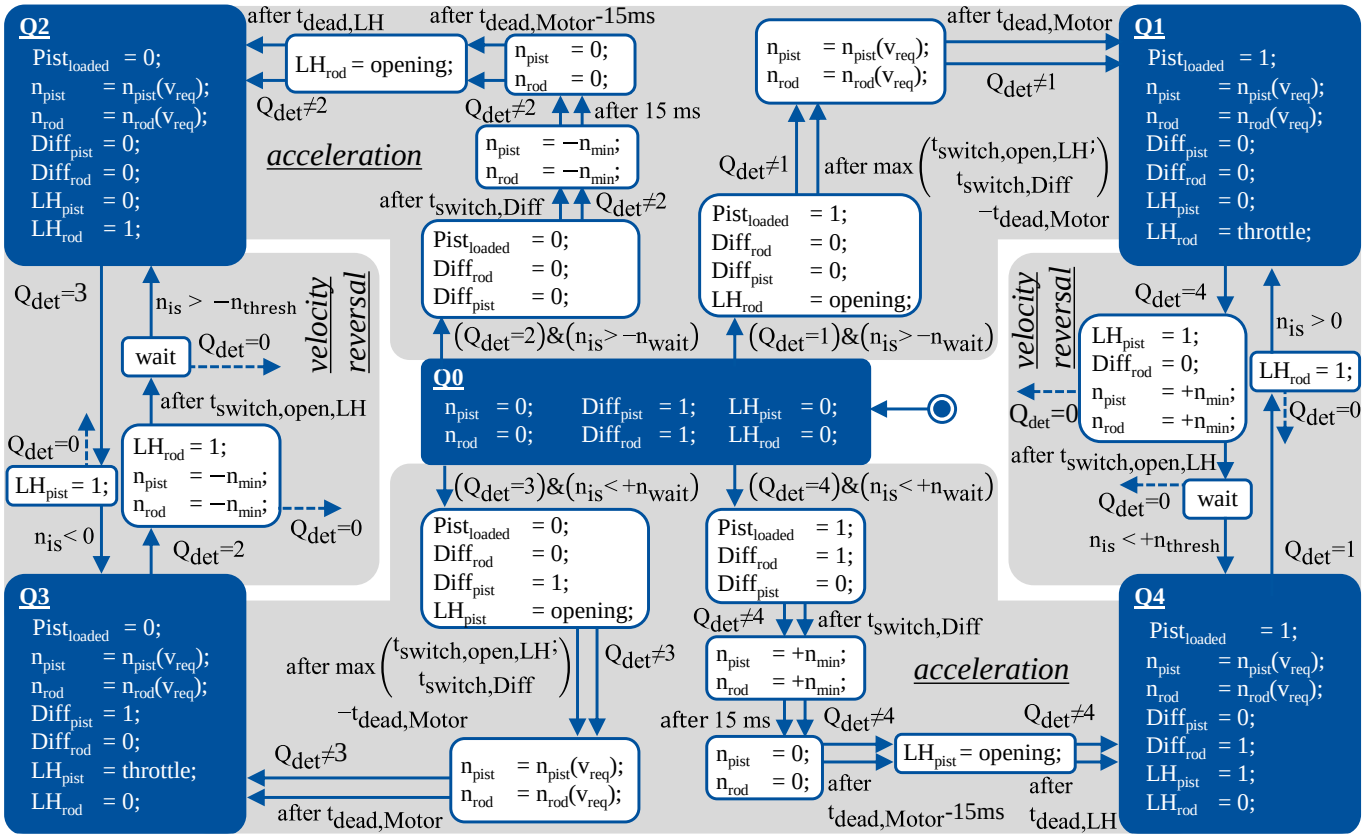


Figure 8: EHA controller – acceleration and reversal of cylinder movement

with either faster valve or motor characteristics as displayed in fig. 9. By first dropping the requested motor signal down to the minimum rotational speed n_{min} the speed is reduced, but a change of direction is prevented. Only after the switching time of the valve has passed, the signal is changed to the value calculated based on the requested speed, ensuring an open valve and preventing a pressure drop due to insufficient flow out of the cylinder. Further, the valve on the previous discharging side during pumping mode is only closed after the rotational speed drops below a certain threshold n_{thresh} ensuring that the pump has changed rotational direction and preventing a blockage of discharge.

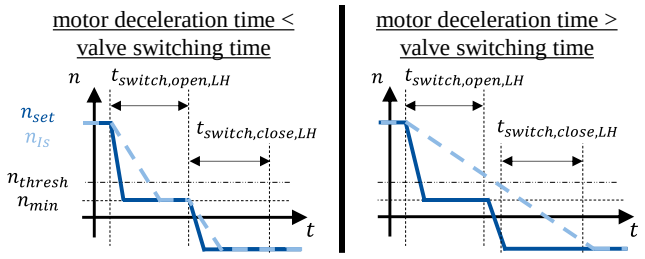


Figure 9: Motor speed request during movement reversal

A trivial value for n_{thresh} would be zero, but a higher value, considering the required switching time, can increase performance by earlier actuation. A change into the pumping quadrants requires less synchronisation because throttling and therefore hydraulic tensioning is desired. Depending on valve and motor characteristics, a more detailed synchronisation might be needed as well.

In fig. 10 the sequences for changing quadrant due to reversal of load direction and stopping are presented. Changes of load direction cannot be predicted and are critical since the loaded cylinder side changes, requiring an adaptation of the pump speed. As explained earlier, the time until a loss of control depends on the virtual boundary, the valve boundary and the throttle boundary. To allow for an adaptation of pump speed during the changes, the load-holding valves are switched to "resistance" mode. The cylinder speed in this mode is controlled by the valve using its characteristic map and the cylinder decoupled from the hydraulic unit.
If the change goes into a motoring quadrant, the valve is slowly opened for smooth pressure equalisation afterwards. To prevent the cylinder from being slowed down during Q3→Q4 the differential valve on the piston side is only closed if the difference between current and requested pump speed is smaller than 15% or after a time limit of 300 ms. The same logic is true for the change Q4→Q3 where the differential valve must open to prevent the reduction in rotational speed, slowing down the cylinder. The change Q2→Q1 only requires a change of the mode of the load-holding valve on the rod side. The time required for the increase of the pump speed depends on the motor performance and current load but is generally short. While the motor speed is too small, the cylinder velocity might be as well.

The implemented approaches will not result in a completely smooth cylinder velocity during changes but shall limit the intensity and duration of deviations as much as possible. The assessment of whether the remaining deviations can be accepted mainly depends on the subjective feel of the operator.

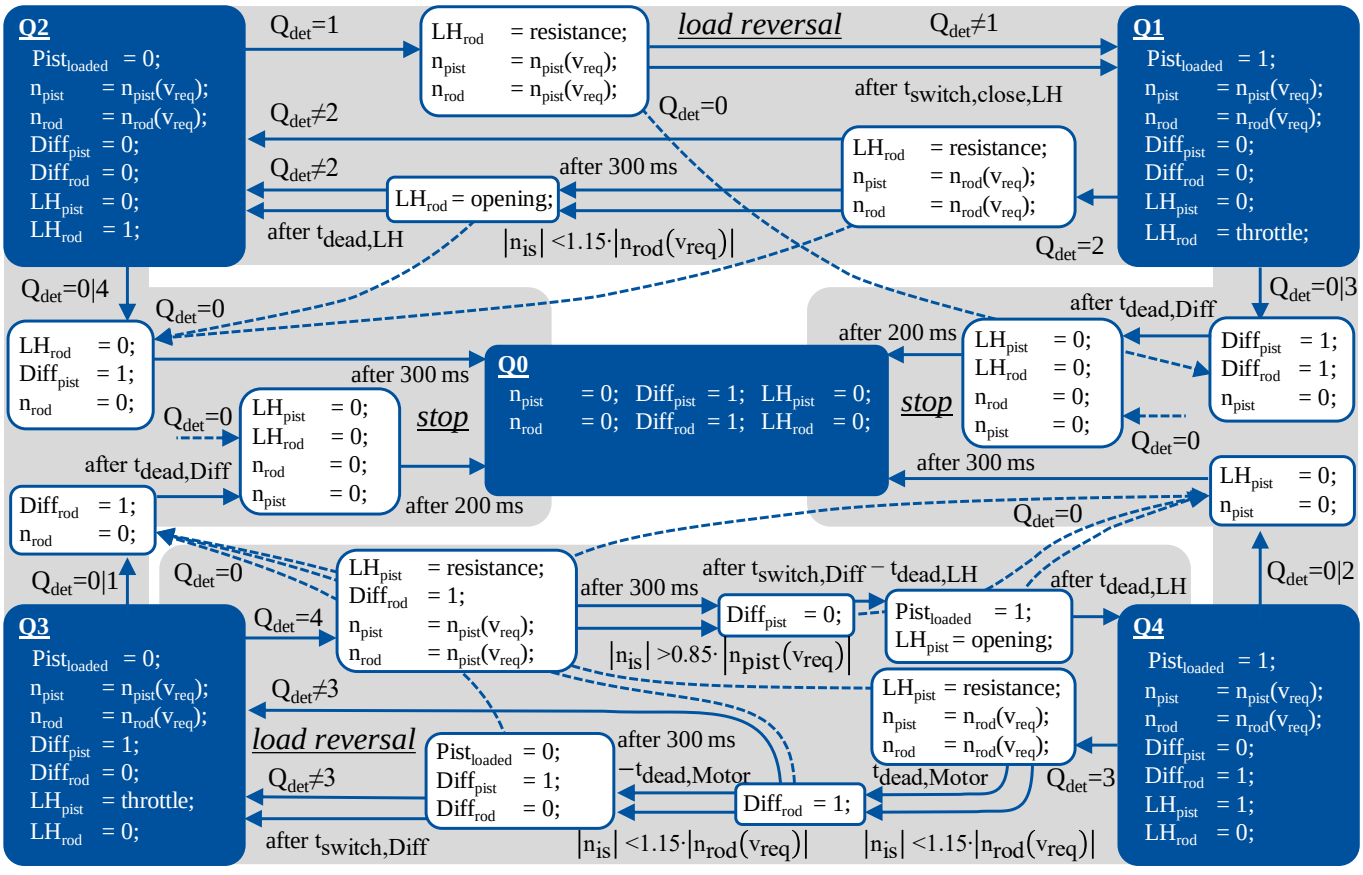


Figure 10: EHA controller – stopping and reversal of load direction

Looking at experimental data and considering the rough operation conditions of excavators, the authors assumed that small and short deviations can be accepted to a degree. This assumption must be experimentally verified in the future, since no meaningful metric could be found.

Lastly, the stopping of the EHA is mainly done by closing the load-holding valves. In case of a pumping operation, both pump ports are connected by additionally opening both differential valves. This results in fast depressurisation and energy cutoff. A stopping signal during any sequence will result in an immediate stop of the EHA. The dashed arrows in fig. 8 denoted with $Q_{det} = 0$ link to the equally marked arrows in fig. 10. On an additional note, changes between quadrants of the same kind (Q1↔Q3 & Q2↔Q4) are handled by first stopping and then accelerating again. The changes Q1↔Q3 occur frequently if the friction force dominates.

4 Simulation

In this chapter, the simulation design is briefly explained as well as the test cases used. Afterwards, the simulation results are introduced and discussed.

4.1 Design

The simulation approach was already described in [12]. The hydraulic circuit was modelled and parameterised in the lumped parameter fluid power simulation program DSHplus and exported as a functional mockup unit (FMU). It was then imported into Matlab Simulink, where the control was implemented. The most important simulation parameters are given in tab. 1. Sizing was done for a 1.8-ton excavator. The differential valves are modelled to have a linear opening characteristic and the load-holding valves are modelled as a first-order lag element. The sum of dead times and switching times is equal to the values taken from data sheets. Volumetric and hydromechanical efficiencies of the pump are based on a measured characteristic field. The cylinder friction is based on existing measurement data. The electric motor is of PMSM type and is modelled as shown in fig. 11. The model was aimed to include the reduction of torque at higher speeds. The maximum motor torque $M_{EM,max}$ as function of the rotational speed is therefore taken from the data sheet of the motor. Depending on the difference between requested speed n_{set} and current speed n, a PI-controller adjusts the applied motor torque M_{EM}. The acceleration is then calculated based on M_{EM}, the current motor load M_{load} as result of pressure difference and the moment of rotational inertia of the motor I_{EM} and pump I_P.

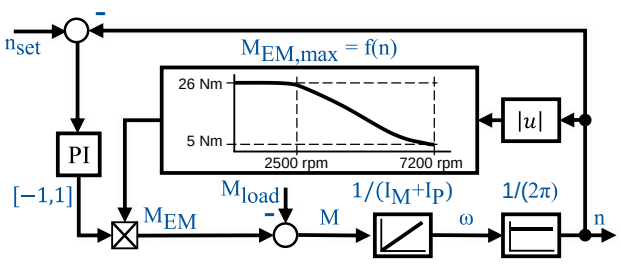


Figure 11: Simulation model of the PMSM

Table 1: Simulation parameters

Parameter	Value
pump displacement volume	5.4 cm^3
cylinder piston diameter	60 mm
cylinder rod diameter	35 mm
cylinder stroke	415 mm
cylinder equivalent mass	1000 kg
accumulator nominal volume	2 l
accumulator pre-charge pressure	4.5 bar
oil bulk modules at high pressures	14,000 bar
$t_{dead,Motor}$	50 ms
$t_{dead,Diff}$; $t_{dead,LH}$	20 ms
$t_{switch,Diff}$	40 ms
$t_{switch,\ open,LH}$	50 ms
$t_{switch,\ close,LH}$	150 ms
rotational acceleration limit	25,200 rpm/s
moment of inertia of motor and pump	0.011 kgm^2
$F_{Fric,static}$	250 N
$F_{Fric,base}$	0 N
$\dot{x}_{vis}$	50 mm/s
μ_{vis}	10 Ns/mm

The simulation allows investigation of the circuit behaviour for the testcases displayed in fig. 12. The velocity reversal case is used to investigate the corresponding quadrant changes by applying a constant force and changing the requested velocity direction. For the velocity reversal testcase, simulations were done for all combinations of force and velocity, since the velocity value correlates to the first request. Secondly, during the force reversal test case, the velocity is kept constant and the applied force is changed linearly. F_{mean} correlates roughly to the position of the hysteresis band. Therefore, the positive value is used for negative velocities and vice versa. It was adjusted to achieve the desired quadrant changes due to friction influences. Simulations were done for each combination of velocity and ΔF. The start position of the cylinder was set to 50 mm for positive velocities and 365 mm for negative velocities.

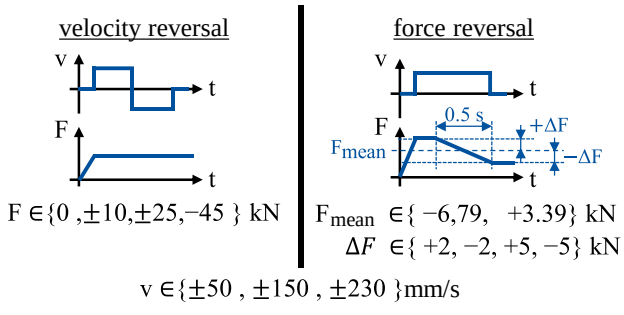


Figure 12: Test cases

4.2 Results

In this section, the simulation results for the introduced testcases of velocity and force reversal are shown and discussed.

4.2.1 Velocity reversal

Figure 13 shows the simulated cylinder velocities for the testcase in which the direction of cylinder movement is changed. The top side shows cases with a positive starting velocity and the bottom side with a negative. The shown timeframe excludes the initial load application. The requested velocity correlates with v_{req} in fig. 6 and therefore the change rate is not limited.

As shown in fig. 1, a negative external force corresponds to a force trying to push the cylinder inwards (Q1 & Q4). Looking at the velocities, one can see that the transitions from Q4→Q1 and Q2→Q3 are smooth and linear as designed. The transitions Q1→Q4 and Q3→Q2 show the designed rest close to zero velocity as shown in fig. 9 before accelerating again, noticeable by the offset between the lines above and below zero velocity. The influence of the reduced torque of the electric motor can be seen. It results in speed deviations for higher loads as well as reduced acceleration.

As explained earlier, at low external load conditions, the friction forces dominate, resulting in a quadrant change Q1↔Q3. Control-wise this is handled by a combination of stopping and accelerating, which clearly can be seen by looking at the 0 kN (black) lines. The velocity is reduced way quicker since the stopping is handled by closing the load-holding valves instead of slowing down the motor. To increase efficiency further, an additional sequence could be added that slows the cylinder down due to motor speed before then changing into the acceleration sequence.
Overall, the performance of the control is rated good. The transitions are smooth and there are only movements in the requested direction. The oscillations after stopping are a result of the quick stopping. The small oscillations during acceleration are assessed to be not noticeable during operation, since the operation conditions of mobile machinery are rough.

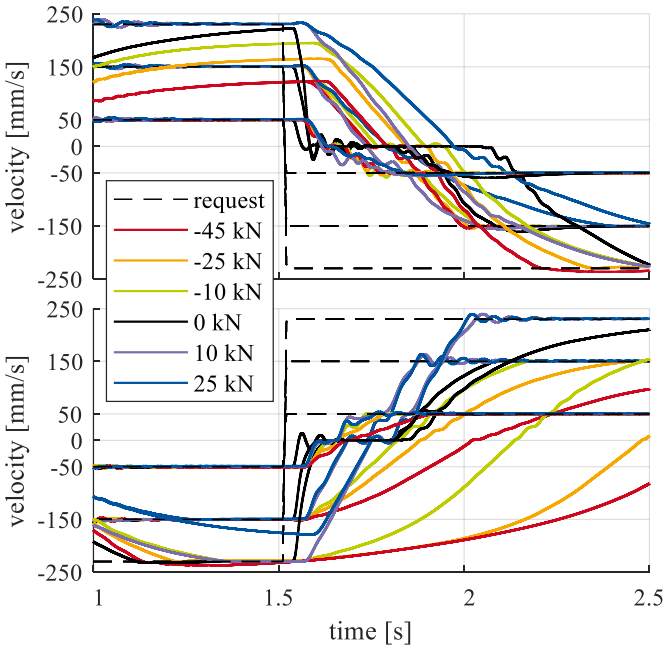


Figure 13: Results for the test case velocity reversal

4.2.2 Force reversal

Figure 14 shows, analogous to fig. 13, the simulated cylinder velocities for the different force combinations of the test case force reversal. The simulation results are sorted by the change of operating quadrant, resulting in four groups. The time when the quadrant change is detected is marked by a vertical line in the same colour as the corresponding velocity plot.

The topmost graph shows the change Q1→Q2. During the acceleration phase, small oscillations can be seen. Also, slower acceleration due to the reduced torque capacity of the electric motor at high speeds is visible as well. During the change, the motor speed needs to decrease to adjust to the smaller volume flow because of the smaller cylinder rod area. Looking at the timeframe in which the change occurs, only small deviations can be seen. The momentum of mass, assistive load and resistive valve mode keep the velocity close to the requested value. The initial overshoot at the higher load change rate is a result of the valve dynamics and the time needed to switch to resistive mode. The valve closes too slowly, so that the assistive load accelerates the cylinder. At lower change rates, the cylinder is still controlled by the pump speed if the rotational speed is reduced, resulting in a slowdown before valve control takes over. Looking at the requested velocities of 230 mm/s an oscillation is clearly visible. The motor speed does not match when the differential valve closes and results in an excitation of the system.

The second graph displays the results for the change Q2→Q1, in which the pump speed needs to increase. After the quadrant change is detected, the cylinder slows down due to friction and throttle pressure buildup. As soon as the pump volume flow matches the required cylinder volume flow, the pressure increases and the cylinder is controlled by the pump speed again. The switch back to pump control results in a small excitation. Higher change rates of load result in a faster slowdown. The pump speed adjusts fast enough to prevent large deviations at low speeds. At higher speeds, the reduced performance and larger speed adjustments increase the required timeframe, resulting in larger deviations.

The changes from Q3→Q4 are shown in the third graph. Analogue to Q2→Q1, the pump speed needs to increase during the change, but switches to motoring mode. At higher speeds, the pump speed does not match, resulting in a small excitation. The initial slowdown is a result of the pump rod side port being connected to the accumulator, removing the pushing force and the time required by the valve to switch to resistive mode. The valve dynamics explain the small overshoot afterwards as well, since the assistive load continues to increase.

Lastly the changes from Q4→Q3 are plotted in the downmost graph. Looking at the cylinder velocities, the worst performance of all quadrant changes can be seen. While changes at slower speeds show only small deviations they increase with increasing speeds. Further, a slower change rate of load initially results in acceleration due to a higher remaining load when the differential valve is opened. The valve must be opened, since otherwise the decrease of pump speed results in a slowdown of the cylinder because of the reduced flow. At -150 mm/s and -230 mm/s and -5 kN load change rate the cylinder slows down because the valve

boundary is reached. Especially at -230 mm/s the pump speed adjustment is too slow.

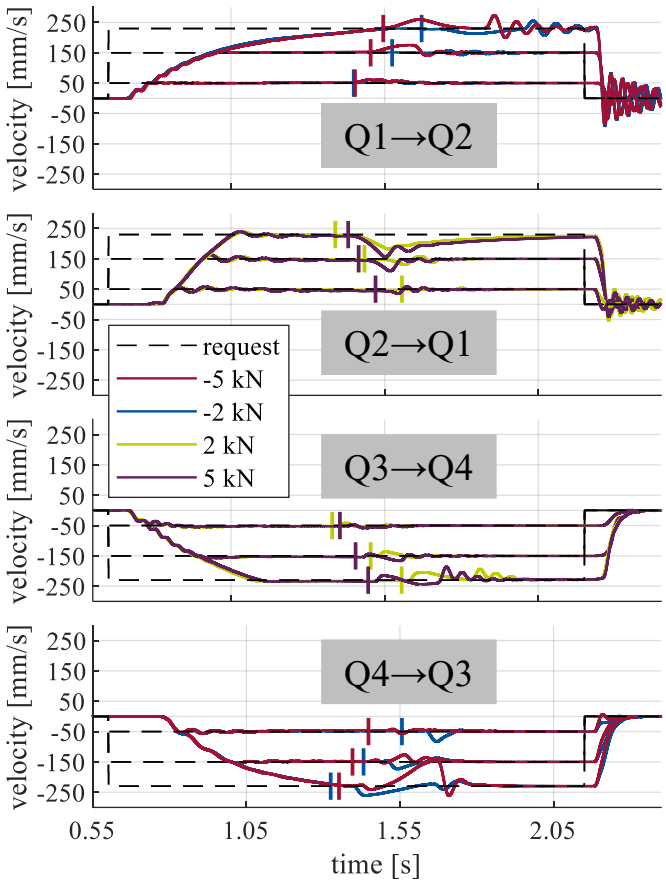


Figure 14: Results for testcase force reversal

Additionally, it can be observed that the valve-based stops result in strong oscillations after a movement in positive direction and a smooth deceleration after a movement in the negative direction. The reason for this is that the cylinder is under hydraulic tension while slowing down during inward movement. In Q4 the pump provides more fluid to the rod side than is required by the cylinder slowing down. The oversupply causes a pressure drop over the differential valve, tensioning the cylinder. The same is true for Q3 where the cylinder slows down faster than the pump because of the closing load-holding valve. The hydraulic tension prevents excitation of the system. In Q2 the pressure on the piston side of the cylinder is close to the accumulator pressure since the check valves are open. Only when the cylinder almost reaches zero velocity, the pump volume flow causes a small pressure buildup. The missing hydraulic tension results in a pressure buildup on the rod side due to the closing load-holding valve. The built-up pressure accelerates the cylinder inwards and results in oscillation. The same is true for Q1. The pressure does not drop to the accumulator pressure and is higher afterwards due to the faster pump speed. The oscillation is smaller as a result.

The duration of the cylinder velocity deviations is short and small, except for the change at higher speeds from Q4→Q3. If one considers that the reversal of external force represents a big change in operational conditions some acceptance for deviations by the operator is likely. For example, this occurs when ground excavation starts after lowering the arm in motor mode. Considering this, the shown velocity graphs are rated to be acceptable. Nevertheless, as already mentioned before, further research and experiments with human operators need to confirm this assessment.

5 Conclusion

In this paper, a hydraulic circuit for an electro-hydrostatic actuator (EHA) aimed at the application in mobile machinery is investigated. The circuit was selected based on cost considerations, the use of a high-speed internal gear pump and standard differential cylinder, inclusion of load-holding capability and wide flexibility of the control.
The boundaries for four-quadrant operation of this circuit are derived and explained. The existing boundaries require an extension of the pumping quadrant to allow full coverage of the four-quadrant plane. This is necessary since the operational conditions in mobile machinery are unpredictable and always changing. The extension is done by active throttling of the outgoing cylinder volume flow using proportional load-holding valves.
Based on the derived physical boundaries, a virtual boundary for the quadrant change is defined and combined with active hysteresis to prevent unwanted quadrant changes. The developed control scheme and controller architecture, mainly comprising two state machines, are presented. The controller aims to synchronise motor speed with valve positioning, allowing for smooth cylinder velocities under changing velocity requests and load directions.
The simulation results show good performance for the reversal of velocity under constant force. During the reversal of external load under constant speed, the deviations are higher but rated to be acceptable for the application in an excavator. The biggest deviations result from changes from Q4→Q3. Further, especially high speeds result in larger deviations because of the required adjustment of pump speed combined with reduced torque capacity of the electric motor.

6 Outlook

Experimental validation in an excavator with a human operator is recommended and planned. In future, the control scheme could be further refined and optimised. For example, an optimised pre-pressurisation during the acceleration sequence in motoring quadrants or a sequence allowing for recuperation during changes between quadrants Q1 and Q3 could be implemented. Additionally, a demonstrator of the EHA will be built and used to experimentally validate the simulation results.

Acknowledgements

The IGF research project “01IF22246N” of the research association Forschungskuratorium Maschinenbau e. V. – FKM, Lyoner Straße 18, 60528 Frankfurt am Main was supported from the budget of the Federal Ministry of Economic Affairs and Climate Action through the DLR within the scope of a program to support industrial community research and development (IGF) based on a decision of the German Bundestag.

Nomenclature

Designation	Denotation	Unit
A_{pist}, A_{rod}	Cylinder areas	m^2
$Diff_{pist/rod}$	Actuation of the switching valves	-
F	External cylinder force	N
F_{fric}	Friction force	N
$LH_{pist/rod}$	Mode of the load-holding valves	-
m	Equivalent mass	kg
n_{is}	Current rotational speed of the motor	min^{-1}
n_{min}	Minimal rotational speed of the pump	min^{-1}
$n_{pist/rod}$	rotational speed calculated by EHA controller if piston/rod side is loaded	min^{-1}
n_{thresh}	Threshold during a change of rotational direction	min^{-1}
p_{acc}	Accumulator pressure	bar
$p_{bound,acc}$	Load pressure boundary introduced by the accumulator pressure	bar
$p_{bound,fric}$	Load pressure boundary resulting from the cylinder friction	bar
$p_{bound,throttle}$	Load pressure boundary resulting from the throttling	bar
$p_{bound,valve}$	Load pressure boundary introduced by the load-holding valve	bar
p_{load}	Load pressure	bar
p_{rod}, p_{pist}	Cylinder pressures	bar
$p_{pump,rod/pist}$	Pressures at the pump ports	bar
$Pist_{loaded}$	Equal to 1, if piston side is loaded	-
POCV	Pilot-operated check valve	-
Δp_{Hyst}	Width of hysteresis	bar
$\Delta p_{throttle}$	Pressure difference while throttling	bar
Q	Volume flow	l/min
$Q0, Q1, Q2, Q3, Q4$	Operating quadrant in the four quadrant plane (Q0 = standstill)	-
Q_{det}	Output of quadrant detection	-
$t_{dead,LH}$	Dead time of load-holding valve	ms
$t_{dead,Diff}$	Dead time of switching valve	ms
$t_{dead,Motor}$	Dead time of motor	ms
$t_{switch,close,LH}$ $t_{switch,open,LH}$	Time to open/close the load-holding valve (incl. dead time)	ms
$t_{switch,Diff}$	Time to switch the position of the switching valve (incl. dead time)	ms
v_{req}	Requested cylinder velocity	m/s
$\dot{x}$	Cylinder velocity	m/s
α	Cylinder area ratio	-

References

[1] Williamson, C. and Ivantysynova, M., Stability and motion control of inertial loads with displacement controlled hydraulic actuators. *Proc. of 6th FPNI-PhD Symp.* West Lafayette, Indiana, June 15-19, 2010, S. 499–514.

[2] Michel, S., Elektrisch-hydrostatische Kompaktantriebe mit Differentialzylinder für die industrielle Anwendung, Dissertation, TU Dresden, Dresden, 2020. DOI 10.25368/2021.86.

[3] Williamson, C. and Ivantysynova, M., Pump mode prediction for four-quadrant velocity control of valveless hydraulic actuators. *Proc. of 7th JFPS International Symposium on Fluid Power.* Toyama, Japan, 15-18 September, 2008.

[4] Wang, L., Book, W. J., and Huggins, J. D., A Hydraulic Circuit for Single Rod Cylinders. In: *Journal of Dynamic Systems, Measurement, and Control*, Vol. 134, No. 1, 2012. DOI 10.1115/1.4004777.

[5] Çalışkan, H., Balkan, T., and Platin, B. E., A Complete Analysis and a Novel Solution for Instability in Pump Controlled Asymmetric Actuators. In: *Journal of Dynamic Systems, Measurement, and Control*, Vol. 137, No. 9, 2015. DOI 10.1115/1.4030544.

[6] Imam, A., Rafiq, M., Jalayeri, E., and Sepehri, N., Design, Implementation and Evaluation of a Pump-Controlled Circuit for Single Rod Actuators. In: *Actuators*, Vol. 6, No. 1, S. 10, 2017. DOI 10.3390/act6010010.

[7] Imam, A., Tolba, M. S., and Sepehri, N., Improving performance of pump-controlled hydraulic circuits for single-rod actuators: conceptual study. In: *Journal of Physics: Conference Series*, Vol. 2616, No. 1, S. 12015, 2023. DOI 10.1088/1742-6596/2616/1/012015.

[8] Koury Costa, G. and Sepehri, N., A Critical Analysis of Valve-Compensated Hydrostatic Actuators: Qualitative Investigation. In: *Actuators*, Vol. 8, No. 3, S. 59, 2019. DOI 10.3390/act8030059.

[9] Gøytil, P. H., Padovani, D., and Hansen, M. R., A Novel Solution for the Elimination of Mode Switching in Pump-Controlled Single-Rod Cylinders. In: *Actuators*, Vol. 9, No. 1, S. 20, 2020. DOI 10.3390/act9010020.

[10] Kärnell, S. and Ericson, L., Hysteresis Control in Pump-Controlled Systems—A Way to Reduce Mode-Switch Oscillations in Closed and Open Circuits. In: *Energies*, Vol. 15, No. 2, S. 424, 2022. DOI 10.3390/en15020424.

[11] Ketelsen, S., et al., Classification and Review of Pump-Controlled Differential Cylinder Drives. In: *Energies*, Vol. 12, No. 1293, 2019. DOI 10.3390/en12071293.

[12] Figge, F. and Schmitz, K., Simulative Comparison Approach of Electro-Hydrostatic Actuator Circuits in Excavators. *Proc. of ASME/BATH 2023 Symposium on Fluid Power and Motion Control.* Sarasota, Florida, USA, 16.-18. October, 2023. DOI 10.1115/FPMC2023-111798.

[13] Rahmfeld, R., Energy saving hydraulic actuators for mobile machines. *Proc. of 1st Bratislavian Fluid Power Symposium.* Casta-Pila, Slovakia, 2–3 June, 1998.

[14] Padovani, D., Ketelsen, S., Hagen, D., and Schmidt, L., A Self-Contained Electro-Hydraulic Cylinder with Passive Load-Holding Capability. In: *Energies*, Vol. 12, No. 292, 2019. DOI 10.3390/en12020292.

[15] Habibi, S. and Goldenberg, A., Design of a New High-Performance ElectroHydraulic Actuator. In: *IEEE/ASME TRANSACTIONS ON MECHATRONICS*, Vol. 5, No. 2, 2000.

[16] Michel, S. and Weber, J., Energy-efficient electrohydraulic compact drives for low power applications. *Proc. of Fluid Power and Motion Control FPMC 2012.* Bath, United Kingdom, 12-14 September, 2012.

Physics-Informed Neural Network for Solving Hydrodynamic Lubrication Characteristics of Piston Pump Slipper Pair

Hanyu Gao[1], Hua Fang[1], Hao Liu[1*], Yan Tong[2], Dayong Yang[2], Xiaoping Ouyang[1*]

[1]State key laboratory of fluid power components and mechatronic systems, Zhejiang University Hangzhou, Zhejiang, China
[2]Shanghai Aircraft Manufacturing Co., Ltd. Shanghai, China
E-mail: henrygao@zju.edu.cn, fhua@zju.edu.cn, haoliu7850052@zju.edu.cn, tongyan@comac.cc, yangdayong@comac.cc, ouyangxp@zju.edu.cn

Abstract

The slipper pair is one of the key friction pairs, and its lubrication characteristics are one of the main difficulties in the study of aviation piston pumps. Traditional numerical models used to simulate the lubrication characteristics of friction pairs have problems such as slow calculation speed and poor convergence. Recently, the emergence of Physics-Informed Neural Networks has provided a new way to improve the simulation and solution of lubrication characteristics. Compared with traditional numerical models based on fluid dynamic lubrication theory, PINN provides a more efficient and flexible solution by embedding deep learning models into solving partial differential equations (PDEs). This paper proposes, for the first time, to integrate the prior knowledge of hydrodynamic lubrication in the friction pairs into neural networks, presenting a PINN-based model for solving the oil film characteristics of friction pairs, along with a corresponding training strategy. Simulation results on the slipper pair show that the PINN method effectively overcomes the limitations of traditional numerical methods, and can accurately predict lubrication characteristics and significantly shorten the computation time. The potential of PINN in modeling and solving lubrication characteristics is of great significance for the study and optimization of the friction pair characteristics in piston pumps.

Keywords: Piston pump; PINN; slipper; friction pair; lubrication; oil film

1 Introduction

Axial piston pumps, recognized for their high power density, output pressure, and efficiency, are widely utilized in aircraft hydraulic systems. The enhancement of aircraft flight performance and the need for weight reduction drive the ongoing improvement of power density in hydraulic systems and aviation piston pumps[1]. Increasing pressure and speed are effective ways to enhance the power density of aviation piston pumps[2,3]. The performance and longevity of aviation piston pumps are primarily influenced by three essential friction pairs: the slipper pair, the piston pair, and the port plate pair. The slipper pair, a vital component that connects the piston and the swash plate, plays a key role in maintaining the stability of the piston's reciprocating and linear composite movements, which enables the pump's oil suction and discharge functions. Optimal oil film characteristics are crucial for the reliable, long-term operation of the slipper pair. However, increases in pressure and speed can cause various issues in the slipper pair's oil film, including insufficient load-bearing capacity, unstable dynamics, lubrication failure, and increased temperature rise—all of which compromise the reliability and service life of the aviation piston pump. Thus, enhancing the oil film characteristics of the slipper pair is a critical strategy for improving the reliability and longevity of aviation piston pumps under high-pressure and high-speed conditions.

The characteristics of the slipper pair's oil film are influenced by the complex interaction of several factors, including materials, design, manufacturing, operating conditions, and environmental influences, with many of these elements mutually constraining each other. Consequently, it is impractical to rely solely on expert experience to optimize the oil film characteristics of the slipper pair, making high-precision modeling and simulation an essential process. Furthermore, modeling and simulation play a crucial role in uncovering formation mechanisms, analyzing evolution patterns, and monitoring or predicting the performance of the slipper pair's oil film.

At present, the modeling and simulation of the oil film in slipper pairs face three primary challenges. The primary challenge is the high-precision modeling difficulty caused by model complexity. Current modeling techniques are largely rooted in fluid lubrication theory and involve complicated interdependencies among various factors such as materials,

design, manufacturing, operational conditions, and environmental aspects, as well as multiple physical domains like fluid dynamics, thermal fields, and solid mechanics. This complexity requires the simultaneous analysis of numerous physical equations, parameters, and constraints, making high-precision modeling difficult. The second challenge is the significant computational load associated with resolving complex partial differential equations. Existing models depend largely on numerical techniques, where the requirement for fine mesh discretization and strict iteration tolerance management makes it difficult to balance accuracy and efficiency. As a result, simulating the oil film of the slipper pair requires considerable computational time. The third challenge arises from the limited generalizability of customized models. Simulation models created for specific piston pump slippers are typically designed for particular structures and operating conditions, which makes them poorly equipped to handle wide variations in operating parameters or design changes. In practical engineering research and application, these restrictions notably impede the optimization efficiency of the slipper pair's oil film characteristics, slowing down the rapid iterative development of aviation piston pumps. Fortunately, the recent introduction of Physics-Informed Neural Networks presents a promising opportunity to tackle these challenges.

Physics-informed Neural Networks are models that incorporate existing knowledge of physical equations by embedding partial differential equations (PDEs) and boundary conditions (BCs) as constraints in the training process. This approach ensures the neural network is aligned with the defined physical principles during training constraints[4,5]. PINN provides a variety of benefits. To begin with, it can seamlessly integrate the governing equations of physical fields directly into the neural network, eliminating the requirement for intricate discrete solution frameworks. This greatly streamlines the process of solving PDEs and makes modeling easier. Once trained, PINN can also deliver solutions throughout the entire domain, bypassing the extensive grid calculations and iterative steps common in traditional numerical methods. As a result, it achieves high-accuracy solutions while significantly lowering computational demands time[6,7]. Moreover, the prediction and solution processes of a trained PINN model are exceptionally fast, avoiding the repeated iterations necessary in numerical computations and reducing computational burdens. Additionally, PINN is adaptable to various physical domains, as the training direction and rules can be easily modified by adjusting the loss function. Once trained, PINN models can be applied to similar problems for solutions, eliminating the need for re-modeling, grid discretization, and iterative calculations, thereby enhancing the generalizability of the solution process.

Currently, PINN is being extensively applied in various modeling and simulation fields, including fluid dynamics[8–10], mechanical modeling[11,12], and thermal characteristic modeling[13–15]. In the area of fluid lubrication, PINN has also shown significant promise. Initial advancements have been made in utilizing PINN to address lubrication-related characteristics in bearings. For instance, Almqvist[16] developed a physics-informed neural network to address initial and boundary value problems for solving one-dimensional Reynolds equations. Li et al. [17] introduced a ReF-nets model for determining the lubrication characteristics of gas bearings. Rom [18] employed PINN to solve two-dimensional Reynolds equations incorporating JFO cavitation modeling. Zhao et al. [19] compared the accuracy of FEM and PINN methods in solving two-dimensional Reynolds equations and analyzed the impact of different network layers and neuron counts on computational precision. Xi et al. [20] utilized PINN to address inverse problems in hydrodynamic lubrication for predicting journal-bearing parameters. These studies collectively underscore the immense potential of PINN in modeling and simulating fluid lubrication. In aviation piston pumps, the structure and state of the oil film in friction pairs are more complex than those in bearing oil films, featuring a more significant number of parameters and constraints, which poses greater challenges for PINN-based solutions. To date, no applications have been reported regarding the application of PINN in the oil film lubrication characteristics of slipper pair of aviation piston pump.

This study pioneers the application of the Physics-Informed Neural Network to the modeling of the oil film in the slipper pair of aviation piston pumps. By substituting the traditional two-dimensional Reynolds equation, typically used in numerical models for pressure distribution calculations, with a PINN-based approach, we greatly improve computational efficiency while preserving model accuracy. The subsequent sections of this paper are organized as follows: Section 2 introduces the numerical solution model for the slipper pair's oil film. Section 3 details the development of the PINN-based oil film model and discusses the training strategies employed. Finally, Section 4 analyzes the performance of PINN in solving pressure distributions and compares the errors between PINN and numerical solutions under various parameters.

2 Slipper Pair Hydrodynamic Lubrication Model

2.1 Slipper Pair Overview

As a type of volumetric hydraulic pump, the axial piston pump primarily generates pressure and outputs high-pressure oil by altering its volume. Figure 1 illustrates a structure diagram of an aviation axial piston pump. The basic working principle is that rotational motion drives the piston to reciprocate within the cylinder bore, resulting in periodic changes in the volume of the piston chamber, which generates low suction pressure and high discharge pressure.

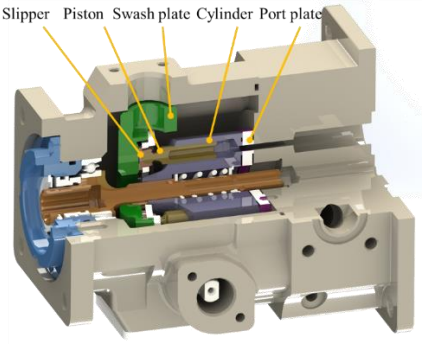


Figure 1: Aviation axial piston pump structure.

The slipper pair, which includes the slipper and the swash plate, forms a plane pair and serves as the primary force-transmitting component in an axial piston pump. Figure 2 illustrates the structure of the slipper pair. The design of the slipper pair is based on the principle of hydrostatic oil film support. High-pressure oil is introduced from the piston chamber into the oil pocket within the slipper pair, creating an oil film on the sealing band of the slipper. During actual operation, the slipper experiences a tilting moment that causes it to tilt at a certain angle. Consequently, the oil film in the slipper pair takes on a wedge-shaped profile, and the shape of this wedge-shaped oil film changes with variations in the slipper's tilting posture.

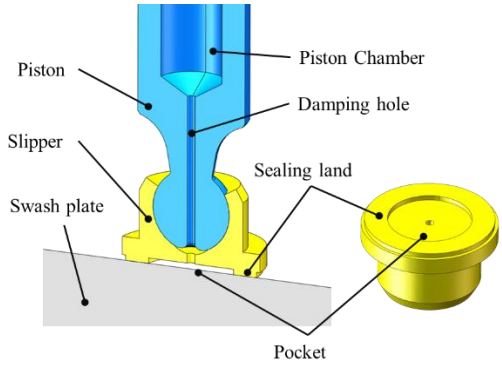


Figure 2: Slipper pair structure.

2.2 Mathematical Model

Firstly, according to the corresponding fluid mechanics assumptions and the continuity equation and NS equation, the two-dimensional Reynolds equation of the oil film of the sliding shoe pair in the polar coordinate system [21] is

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{h^3}{\mu}\frac{\partial p}{\partial r}\right)+\frac{1}{r^2}\frac{\partial}{\partial \theta}\left(\frac{h^3}{\mu}\frac{\partial p}{\partial \theta}\right)=6\left(v_r\frac{\partial h}{\partial r}+v_\theta\frac{\partial h}{\partial \theta}\right) \tag{1}$$

where r is the radial distance, θ is the polar angle, h is the oil film thickness, μ is the viscosity, and v_r and v_θ are the radial and tangential velocities, respectively.

Since the oil film of the slipper pair is wedge-shaped, it is essential to derive the equation for oil film thickness based on the three-point coplanar geometric relationship. As illustrated in Figure 3, points A_1, A_2, and A_3 represent three locations on the outer edge of the slipper oil film with a phase difference of 120°, and the corresponding oil film thicknesses are h_1, h_2, and h_3, respectively. Therefore, the oil film thickness at any point on the bottom surface of the slipper, $h(r,\ \theta)$, can be expressed as:

$$h(r,\ \theta)=\frac{r\cos\theta}{3R}(2h_1-h_2-h_3)+\frac{r\sin\theta}{\sqrt{3}R}(h_2-h_3)+\frac{1}{3}(h_1+h_2+h_3) \tag{2}$$

Where R is the outer radius of the slipper sealing band:

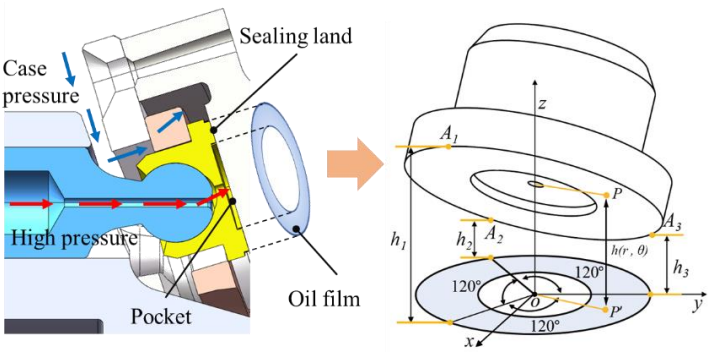


Figure 3: Oil film thickness model of slipper pair.

To reduce computational complexity and avoid instability in numerical calculations, the mathematical model undergoes dimensionless preprocessing. The introduction of dimensionless preprocessing provides a unified benchmark for comparison and eliminates the influence of physical parameters with different orders of magnitude on the subsequent PINN training [22], and helps prevent the imbalance of gradient size during the backpropagation process, thereby enhancing the stability of the training convergence. The equations for oil film pressure and thickness after dimensionless preprocessing are

$$\frac{1}{\bar{r}}\frac{\partial}{\partial \bar{r}}\left(\bar{r}H^3\frac{\partial P}{\partial \bar{r}}\right)+\frac{1}{\bar{r}^2}\frac{\partial}{\partial \theta}\left(H^3\frac{\partial P}{\partial \theta}\right)=6\left(V_r\frac{\partial H}{\partial \bar{r}}+V_\theta\frac{\partial H}{\partial \theta}\right) \tag{3}$$

$$H(\bar{r},\ \theta)=\frac{\bar{r}\cos\theta}{3R}(2H_1-H_2-H_3)+\frac{\bar{r}\sin\theta}{\sqrt{3}R}(H_2-H_3)+\frac{1}{3}(H_1+H_2+H_3) \tag{4}$$

where: the dimensionless oil film pressure is $P=p/p_0$, the dimensionless thickness is $H=h/h_0$, the dimensionless radial distance is $\bar{r}=r/R$, the dimensionless velocities are $V_r=v_r/v_{r0}$ and $V_\theta=v_\theta/v_{\theta 0}$ and the polar angle θ is inherently dimensionless. Assuming that viscosity is constant, the dimensionless viscosity can be regarded as 1.

Meanwhile, the partial derivative of the oil film thickness can also be derived as:

$$\frac{\partial H}{\partial \bar{r}}=\frac{\sin\theta}{\sqrt{3}R}(H_2-H_3)+\frac{\cos\theta}{3R}(2H_1-H_2-H_3)$$
$$\frac{\partial H}{\partial \theta}=\frac{\bar{r}\cos\theta}{\sqrt{3}R}(H_2-H_3)-\frac{\bar{r}\sin\theta}{3R}(2H_1-H_2-H_3) \tag{5}$$

The Dirichlet boundary condition for solving the Reynolds equation is:

$$p(r_0,\ \theta)=p_{in}=1,\ \ p(R,\ \theta)=p_{out}=0 \tag{6}$$

Given that the solution region is annular, the following additional periodic boundary conditions must be specified

$$p(r,\ 0)=p(r,\ 2\pi),\ \ \left.\frac{\partial p}{\partial \theta}\right|_{(r,\ 0)}=\left.\frac{\partial p}{\partial \theta}\right|_{(r,\ 2\pi)} \tag{7}$$

2.3 Numerical Method

The oil film pressure equation governing the slipper pair is a standard elliptical partial differential equation with no analytical solution. This paper employs the finite volume

method to derive the numerical solution. The oil film region for the slipper pair is discretized using polar coordinates, with the discrete grid representation is illustrated in Figure 4. For each discretized area, the volume integral is converted into a surface integral and simplified to the following equation

$$\int_w^e \int_s^n \frac{1}{\bar{r}} \frac{\partial}{\partial \bar{r}} \left(\bar{r} H^3 \frac{\partial P}{\partial \bar{r}} \right) \bar{r} d\bar{r} d\theta + \int_s^n \int_w^e \frac{1}{\bar{r}^2} \frac{\partial}{\partial \theta} \left(H^3 \frac{\partial p}{\partial \theta} \right) \bar{r} d\bar{r} d\theta = \oiint 6 \left(V_r \frac{\partial H}{\partial \bar{r}} + V_\theta \frac{\partial H}{\partial \theta} \right) \quad (8)$$

Eq. (8) is further discretized and organized into the form of linear equation.

$$\begin{aligned} a_p p_p &= a_N p_N + a_S p_S + a_W p_W + a_E p_E + S \\ a_p &= a_N + a_S + a_W + a_E \end{aligned} \quad (9)$$

where:

$$\begin{aligned} a_{\mathrm{n}} &= \bar{r}_{\mathrm{n}} (H^3)_{\mathrm{n}} \frac{\Delta\theta}{(\delta\bar{r})_{\mathrm{n}}} \\ a_{\mathrm{s}} &= \bar{r}_{\mathrm{s}} (H^3)_{s} \frac{\Delta\theta}{(\delta\bar{r})_{\mathrm{s}}} \\ a_{\mathrm{e}} &= \frac{1}{\bar{r}_{\mathrm{e}}} (H^3)_{\mathrm{e}} \frac{\Delta\bar{r}}{(\delta\bar{r})_{\mathrm{e}}} \\ a_w &= \frac{1}{\bar{r}_w} (H^3)_w \frac{\Delta\bar{r}}{(\delta\bar{r})_w} \\ S &= 6 \left(V_r \frac{\partial H}{\partial \bar{r}} + V_\theta \frac{\partial H}{\partial \theta} \right) \bar{r}_{\mathrm{p}} \Delta\theta \Delta\bar{r} \end{aligned} \quad (10)$$

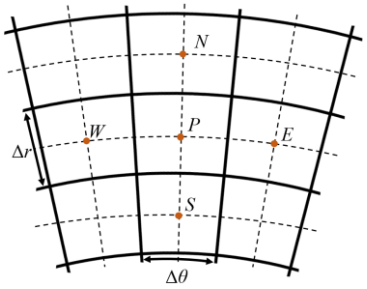


Figure 4: Oil film pressure discrete grid.

The Gauss-Seidel over-relaxation iterative method is employed to solve the discretized equation of the oil film pressure field, and the iterative format of the pressure calculation can be expressed as

$$P_p^{k+1} = \alpha \left(\frac{a_n P_n^k + a_s P_s^{k-1} + a_e P_e^k + a_w P_w^{k-1} + S}{a_p} - P_p^k \right) + P_p^k \quad (11)$$

where α is the over-relaxation coefficient, P_p^k is the pressure at the current iteration, P_p^{k-1} is the pressure from the previous iteration, and P_p^{k+1} is the pressure for the next iteration.

In the process of numerical iteration, the convergence residual is defined as:

$$\left| \frac{p_{\mathrm{p}}^{(k)} - p_{\mathrm{p}}^{(k-1)}}{p_{\mathrm{p}}^{(k)}} \right| \leqslant err \quad (12)$$

When the convergence residual error is below the threshold of 10^{-8}, the current computation is regarded as converged. The pressure values calculated for all nodes together create the pressure field distribution on the bottom surface of the slipper.

The overall flow chart of the numerical solution is illustrated in Figure 5.

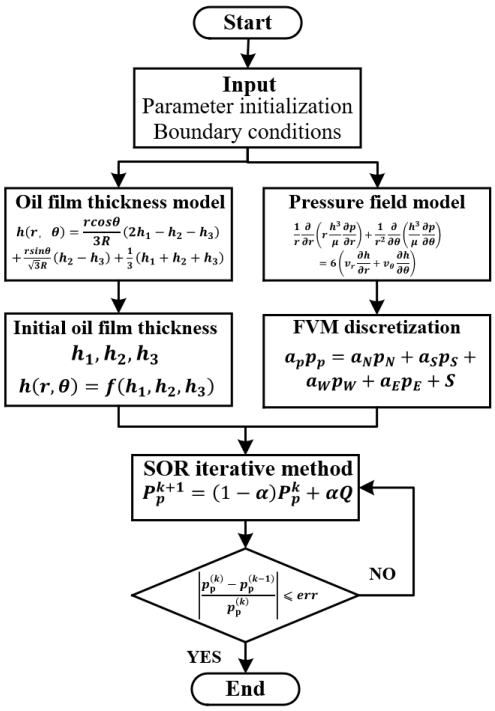


Figure 5: *Flow chart for the numerical solution of oil film pressure.*

3 PINN method

Figure 6 shows a PINN model designed to address the oil film problem in the slipper pair. The PINN model consists of two main components: the neural network block on the left and the physical information constraint block on the right. We chose a standard Artificial Neural Network model (ANN) for the neural network. The inputs to the neural network are the coordinates (r, θ) of the solution domain, and the output is an approximation or prediction of the Reynolds equation solution $P(r, \theta)$.

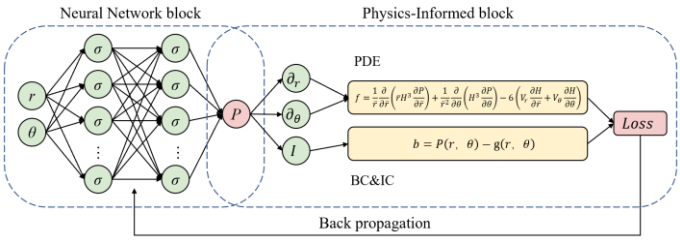


Figure 6: *The PINN Framework for Oil Film Pressure of Slipper Pair*

The objective of training the PINN is to minimize the residuals, which include both boundary condition residuals and PDE residuals. During the training process, the neural network samples across the entire solution domain to generate predicted pressure values for each sampling point. A total of Nb training points are selected on the inner and outer boundaries of the domain (represented by red dots in Figure 7), while Nf sampling points are gathered within the domain (represented by blue dots in Figure 7). The predicted pressure

values at these sampling points are utilized to compute derivatives with respect to the inputs r and θ through automatic differentiation (AD) enabled by the neural network. These results are then input into the physics-informed constraint block to calculate the boundary condition residuals and PDE residuals, which are combined into the loss function. Through iterative training via backpropagation, the loss value is minimized, ultimately leading to the optimal solution and concluding the PINN training process.

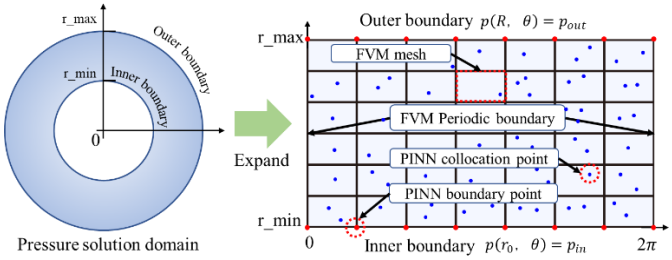


Figure 7: Solution domain and boundary conditions of slipper pair oil film.

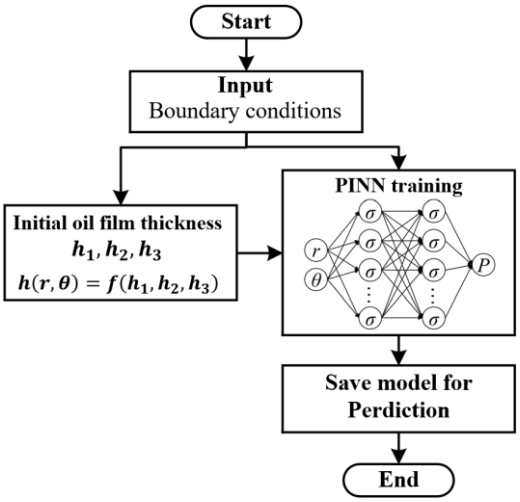


Figure 8: *Flow chart for the PINN-based solution of oil film pressure.*

3.1 Loss function

The loss function determines the training direction of PINN. In this paper, the total loss function $Loss_{mse}$ of the PINN model is the sum of the residuals and PDE residuals of all training points on the inner and outer boundaries, as shown in Formula (14). In each iterative training, the weight and bias parameters of ANN are updated by using iterative optimization algorithms such as Adam or SGD to minimize the total loss $Loss_{mse}$.

The weight coefficient chosen for the loss function heavily affects the optimization direction and effectiveness of the PINN. A loss term with a high weight coefficient indicates it has a substantial impact on the overall loss, thus influencing its optimization during the iterative training phases. Therefore, it is essential to establish an appropriate balance between the weight coefficients of the loss function. In formula (14), we set the weight values of the loss function as β_1=100， β_2=10， and β_3=10, respectively.

$$\begin{cases} loss1 = \frac{1}{Nf}\sum_{i=1}^{Nf}\left\| \frac{1}{\bar{r}}\frac{\partial}{\partial \bar{r}}\left(\bar{r}H^3\frac{\partial P}{\partial \bar{r}}\right) + \frac{1}{\bar{r}^2}\frac{\partial}{\partial \theta}\left(H^3\frac{\partial P}{\partial \theta}\right) - 6\left(V_r\frac{\partial H}{\partial \bar{r}} + V_\theta\frac{\partial H}{\partial \theta}\right)\right\|^2 \\ loss2 = \frac{1}{Nb}\sum_{i=1}^{Nb}\|\hat{p} - p_{in}\|^2 \\ loss3 = \frac{1}{Nb}\sum_{i=1}^{Nb}\|\hat{p} - p_{out}\|^2 \end{cases} \quad (13)$$

$$Loss_{mse} = \beta_1 loss1 + \beta_2 loss2 + \beta_3 loss3 \quad (14)$$

3.2 Network structure and activation function

Choosing hyperparameters for neural networks is a task that significantly depends on empirical knowledge, as their suitability directly affects the model's convergence, accuracy, and computational efficiency. Based on previous research on Physics-Informed Neural Networks, the PINN model in this study utilizes a 6-layer neural network architecture, which comprises four hidden layers with 50 neurons in each layer.

In Physics-Informed Neural Networks, the commonly used ReLU (Rectified Linear Unit) function is not suitable for scenarios that require higher-order derivative calculations due to its non-differentiability at zero. Relevant studies [17] have compared the computational performance of the sigmoid, hyperbolic tangent (tanh), and cosine functions as activation functions, showing that the hyperbolic tangent function demonstrates superior performance in terms of iteration speed and computational accuracy. Therefore, this study selects the hyperbolic tangent function as the nonlinear activation function.

3.3 Weight initialization strategy

In training neural networks, weight initialization significantly affects the network's convergence speed and training efficiency. Commonly used initialization methods include Glorot (Xavier) initialization and He initialization, among others. The choice of initialization method closely relates to the activation function. To ensure compatibility with the hyperbolic tangent activation function, this study uses Glorot uniform initialization to set the neural network's weights.

The core idea of Glorot initialization [23] is to adjust the range of initial weights based on the input and output dimensions to avoid the problems of vanishing or exploding gradients. In Glorot initialization, each weight element is randomly drawn from a uniform distribution within the range [-a, a], where a is determined by the number of neurons in the input layer n_{in} and the number of neurons in the output layer n_{out}. The range of possible weight values can be determined as:

$$\text{Weight} \sim \text{Uniform}\left(-\sqrt{\frac{6}{n_{\text{in}} + n_{\text{out}}}}, \sqrt{\frac{6}{n_{\text{in}} + n_{\text{out}}}}\right) \quad (15)$$

3.4 Optimizer

The choice of optimizer significantly impacts the training efficiency and accuracy of Physics-Informed Neural Networks (PINNs). The loss function of PINNs is often complex, encompassing the residuals of physical equations as well as boundary condition constraints. This paper adopts a combined

optimization strategy. In the early stages of training, when the residuals exhibit significant fluctuations, the Adam optimizer is utilized to manage the highly nonlinear loss function, accelerating the convergence process by adaptively adjusting the learning rate. Once the residual loss stabilizes and the decline becomes gradual, the optimization shifts to the second-order L-BFGS optimizer to fine-tune the update direction and step size for each parameter, further enhancing the accuracy of the results. The L-BFGS optimizer greatly improves the precision of the solution during the later stages of PINN training, meeting the high-accuracy requirements for solving partial differential equations (PDEs). Although its computational cost is relatively high, it remains completely acceptable given the scale of the training task in this study.

Additionally, 10,000 sample points are established in the solution domain, and 400 sample points are placed on the boundary to ensure the adequacy of the training data and accurately meet the boundary conditions.

4 Results And Discussion

In the preceding sections, we introduced the Physics-Informed Neural Network model for solving the Reynolds equation that governs the lubrication of the slipper pair, encompassing the model's principles and detailed configurations. In this section, we validate the model's performance and analyze the influence of various boundary parameters on its effectiveness.

4.1 Parameters and solution environment

This paper utilizes the same hardware and software setup for the numerical calculations and the implementation of the PINN. The environmental configuration details are as follows: the numerical computation is performed using Python 3.11.11, the deep learning framework is Pytorch 2.2.0 running on a Windows 10 system, and the hardware consists of an Nvidia 4090 with 24GB memory and an Intel i9-10900K processor at 3.7GHz.

The solution parameters for the numerical model and the PINN model are presented in Table 1. The parameters of the slipper pair, derived from both the numerical calculation and the PINN model, match the boundary conditions and are scaled identically using dimensionless methods.

Table 1: Parameter setting.

Parameter	Value	Unit
Radial grid number	30	
Angular grid number	120	
Error control	1e-8	
Piston pump rotational speed	12000	rpm
Outer radius of the slipper pair sealing band	3.85	mm
Inner radius of the slipper pair sealing band	2.5	mm
Oil film thickness of the characteristic point （vanilla）	10，12，8	μm
Number of pistons/slippers	9	
Piston pump swashplate angle	17	°
Oil viscosity	0.046	Pa*s

4.2 Model performance

The results of the model are illustrated in Figure 9, where (a), (b), and (c) represent the numerical solution results of the pressure distribution, the PINN solution results, and the absolute value error between the PINN and numerical solutions, respectively. The vertical axis of the figure indicates the dimensionless pressure value, while the horizontal axis corresponds to the dimensionless Cartesian coordinate. (d) shows the oil film thickness distribution. The vertical axis in this section refers to the dimensionless oil film thickness value, and the horizontal axis corresponds to the dimensionless Cartesian coordinate, consistent with the pressure distribution. (e) depicts the change in the loss curve of the PINN model during the training process. Initially, the loss value is high, and the Adam optimizer is employed. The loss value decreases rapidly with oscillation. After 8000 training steps, it drops to the specified accuracy and transitions to a second-order optimizer for smooth fine-tuning. After approximately 12000 training steps, the loss value stabilizes and remains unchanged, which may be regarded as convergence.

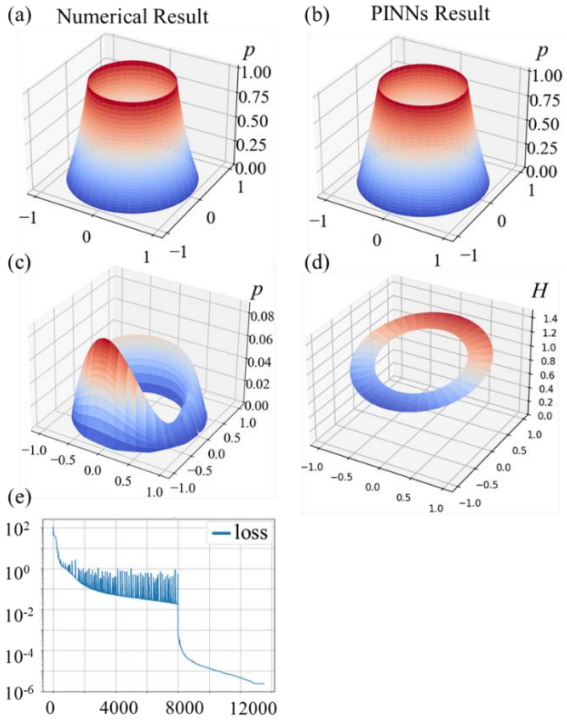


Figure 9: Oil film pressure solution results. (a) Numerical Method Result (b) PINN Method Result (c) Error between PINN and Numerical Solution (d) Oil Film Thickness Distribution (e)PINN Training Loss.

Because the central hole of the slipper is connected to the high-pressure oil in the plunger cavity, and the outer edge pressure of the slipper matches the case pressure, the oil film pressure distribution of the slipper pair exhibits high pressure in the central area that gradually decreases to low pressure at the edges. Figure 9 (a) and (b) illustrate that the results from PINN are highly consistent with the numerical solutions, and the pressure distribution is uniform and continuous. Figure 9 (c) further shows that the solution error of PINNs is very low. A thorough quantitative error analysis follows, comparing

several different error calculation methods, as presented in Table 2. The error remains low within the solution domain. Analyzing the shape of the error region reveals that larger deviations occur in two areas, correlating with the oil film thickness distribution shown in Figure 9 (d), which appears at both sides of the larger oil film overturning angle. The oil film is wedge-shaped, leading to an uneven thickness distribution, with overturning angles along both horizontal axes. This may be attributed to the PINN method's sensitivity to the oil film thickness angle, resulting in a non-uniform distribution of calculation errors.

Table 2: Error Comparison.

Error Type			
Relative Error	Mean Squared Error	Root Mean Squared Error	Mean Absolute Error
6.54E-02	1.21E-03	3.48E-02	2.86E-02

Table 3 displays the numerical solution time along with the average time utilized for PINN training and prediction. It can be seen that the training time of the PINNs method is longer than the simulation time of the numerical method. However, the more significant advantage of the PINN approach lies in its greatly reduced prediction or reasoning time. The prediction time of PINN shows an order of magnitude improvement compared to numerical calculation, being approximately 282 times faster. The trained model can be employed for prediction calculations by retaining the neural network's weight settings, eliminating the need for the iterative process of numerical calculations. Similar research also points out that[24], although the calculation cost of training neural network is high in the early stage, the prediction of pressure distribution after training is several orders of magnitude faster than that of numerical simulation, and the prediction time is relatively stable, which is not easily affected by grid density.

Table 3: Average Computation Durations.

Task	Time		
	Simulation	Training	Prediction
Numerical Method	10.454s	\	\
PINN method	\	240.945s	0.037s

4.3 Influence of oil film thickness

Additionally, we compare the solving performance of the PINN model using different parameters. With all other parameters held constant, we assess the model's effectiveness at various rotational speeds. First, we examine how changes in oil film thickness affect the solution, since oil film thickness is crucial in solving the Reynolds equation. Even slight adjustments in thickness can greatly impact the results of the solution.

We conducted a comparative analysis of solutions for three different oil film thicknesses, as shown in the first row of Figure 10, which presents three distinct oil film thickness distributions: a flat oil film and two wedge-shaped oil films with different tilt directions. The flat oil film represents an ideal state where the slipper and swash plate are perfectly parallel, resulting in a uniform oil film thickness. However, under practical operating conditions, factors such as load variations, installation errors, and mechanical vibrations introduce an inclination between the slipper and swash plate, leading to non-uniform oil film thickness. According to the Reynolds equation, oil film pressure increases as oil film thickness decreases. Consequently, the pressure gradient shows a correspondingly higher variation on the thinner side of the wedge-shaped oil film.

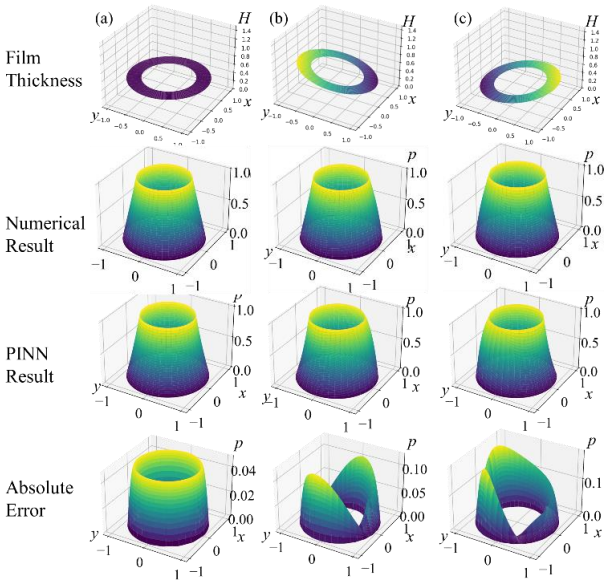


Figure 10: *Solution comparison under different oil film thicknesses.*

By comparing the three states of oil films, we can see that the results from the PINN solution closely align with those from numerical calculations. This indicates that PINN effectively captures how changes in oil film thickness influence pressure distribution. However, the distribution of oil film thickness significantly impacts the pressure distribution within the slipper pair. When the oil film is flat, the calculation error between the two methods is very small and nearly negligible. Under wedge-shaped lubricating film conditions, the pressure distribution demonstrates non-uniform characteristics, where error distribution predominantly manifest in areas with pronounced pressure gradient variations. Different error characteristics are illustrated in Figure 11. This discrepancy may stem from PINN's larger error in regions with a significant pressure gradient change. To enhance calculation accuracy across various oil film thickness distributions, it may be necessary to refine the sampling strategy during training and adjust the density of sampling points in accordance with oil film thickness.

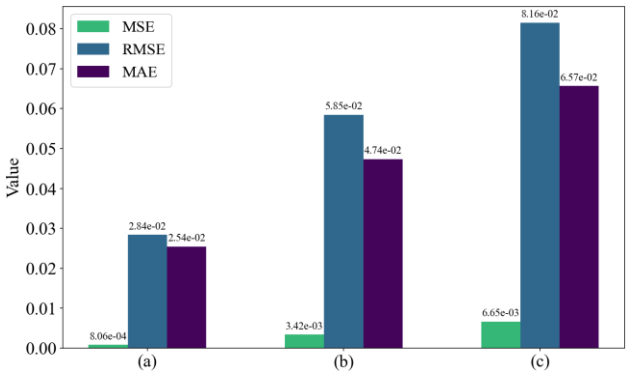


Figure 11: Comparison of pressure solution errors under different oil film thicknesses.

4.4 Influence of pressure boundary

Similarly, we utilized the data and oil film thickness distribution from the solution parameters for the numerical model and the PINN model are presented in Table 1. The parameters of the slipper pair, derived from both the numerical calculation and the PINN model, match the boundary conditions and are scaled identically using dimensionless methods. Figure 12 shows the results under different boundary pressures. The effect of varying the boundary pressure on the internal pressure distribution of the oil film is quite direct. The pressure distribution patterns obtained from the PINN model under different boundary pressures align with those derived from the numerical method. This demonstrates that the PINN model possesses strong generalization capabilities and can adapt to diverse boundary conditions.

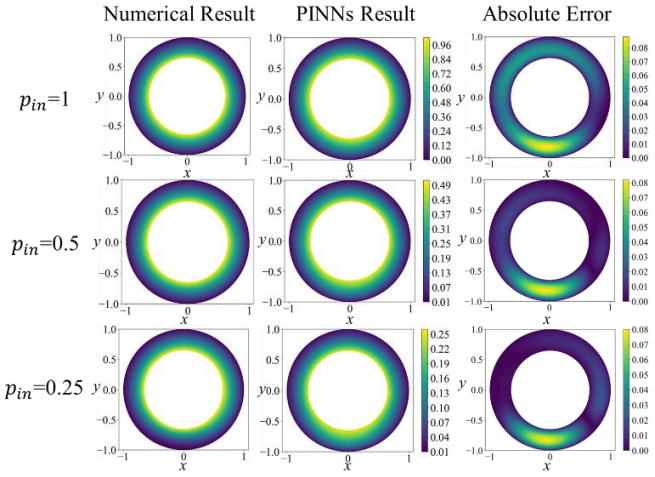


Figure 12: Solution comparison under different boundary pressures.

In the error distribution mode, similar to the previous analysis regarding the impact of oil film thickness, the error associated with smaller oil film thickness shows a larger peak. The error distribution is more confined near the pressure boundary, while the pressure gradient is greater in the middle layer. As the boundary pressure decreases, the absolute value of the error decreases slightly, and the distribution area also contracts. Different types of error pairs are illustrated in Figure 13.

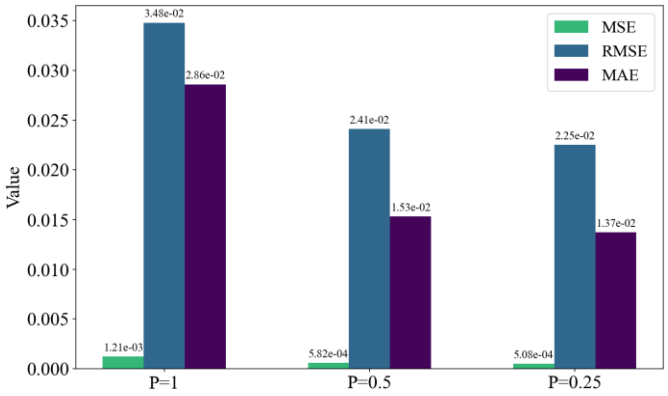


Figure 13:Comparison of pressure solution errors under different pressure boundaries.

4.5 Influence of rotational speed

We further assess the solution performance at various speeds, as shown in Figure 14, and evaluate the oil film distribution at 12000 rpm, 8000 rpm, and 4000 rpm. The error comparison is illustrated in Figure 15. Unlike the alterations made to the pressure boundary conditions, the change in rotational speed has a relatively minor effect on the error. From the error distribution across the three cases, as rotational speed increases, the overall error level rises slightly; however, the degree of change is noticeably smaller than that observed with the pressure boundary, indicating that PINN demonstrates greater stability when addressing the dynamic pressure effect.

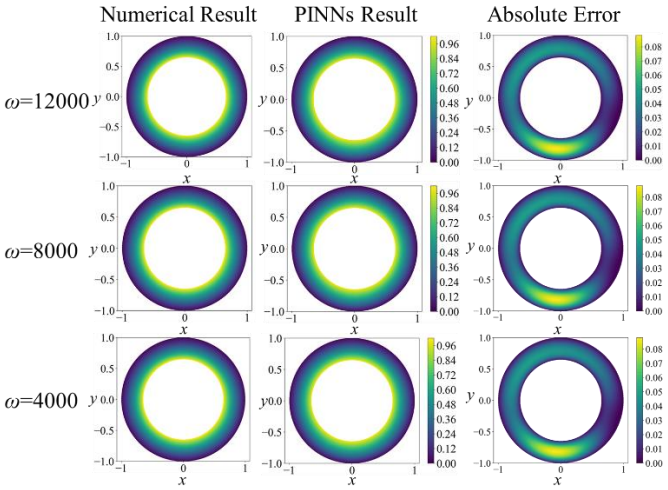


Figure 14: Solution comparison under different rotational speeds.

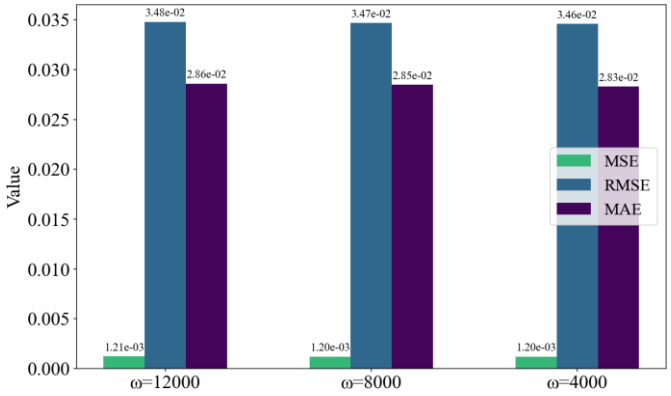


Figure 15: Comparison of pressure solution errors under different rotational speed conditions.

Although the aforementioned related discussions have proved the effectiveness and accuracy of PINN in solving the Reynolds equation of slipper pair, it must be admitted that it still faces challenges in the practical application of lubrication analysis. At present, PINN only solves the lubrication characteristics under specific parameters. Due to the complex motion and force of the slipper in actual operation, these constraints need to be added to the training of the model to improve the generalization of the model. However, more physical constraints are integrated into the loss function as soft penalty terms, which will lead to gradient imbalance between these constraints, and then lead to training failure or inaccurate enforcement of boundary conditions. Future work needs to further improve the structure or optimization strategy of PINN and extend the training model to complex working conditions including squeezing motion.

5 Conclusion

This study proposes a method utilizing PINN to analyze the lubrication characteristics of high-speed piston pump slippers, verifying the potential of PINN in addressing complex hydrodynamic lubrication issues. The results from comparisons with traditional numerical approaches indicate that PINN effectively handles complex nonlinear hydrodynamic lubrication equations, accurately capturing oil film pressure distribution and changes in film thickness during the lubrication of slipper pairs. Analyzing different operational parameters, including speed, boundary pressure, and oil film thickness, demonstrates that the PINN method demonstrates high solution accuracy, adaptability, and robustness. Nonetheless, PINN encounters challenges related to accuracy, convergence speed, and generalization in real-world applications, indicating a need for further optimization of the network structure and refinement of the training strategy to enhance performance, so that it has the potential to be extended to solve the problems including extrusion motion and elastic deformation. Overall, this paper presents an innovative deep-learning technique for modeling and calculating lubrication characteristics in piston pumps, which is anticipated to impact the design and optimization of high-speed piston pumps significantly.

Acknowledgment

This research was funded by the National Natural Science Foundation of China (Grant No. U2141209), the Aeronautical Science Foundation of China (Grant No. 20182876003 and 20240028076003), and the Natural Science Foundation of Zhejiang Province (Grant No. LQ23E050013).

Corresponding authors:

Hao Liu: haoliu7850052@zju.edu.cn;

Xiaoping Ouyang: ouyangxp@zju.edu.cn.

Nomenclature

Designation	Denotation	Unit
r	Polar radius	m
θ	Polar coordinate angle	rad
r_0	Radius of inner sealing belt of slipper	m
R	Radius of outer sealing belt of slipper	m
h	Slipper film thickness	μm
μ	Oil viscosity	Pa*S
v_r, v_θ	Slipper velocity	m/s
h_1, h_2, h_3	Oil film thickness of feature points	μm
P	Dimensionless oil film pressure	
H	Dimensionless oil film thickness	
$\bar{r}$	Dimensionless polar radius	
V_r, V_θ	Dimensionless velocity	
p_{out}	Pressure outer boundary conditions	
p_{in}	Pressure inner boundary condition	

References

[1] J Zhao, Y Fu, J Ma, J Fu, Q Chao, and Y Wang. Review of cylinder block/valve plate interface in axial piston pumps: Theoretical models, experimental investigations, and optimal design. *Chinese Journal of Aeronautics*, 34: (1) 111–134, 2021. https://doi.org/10.1016/j.cja.2020.09.030.

[2] B Zhang, C Zhao, H Hong, G Cheng, H Yang, S Feng, J Zhai, and W Xiao. Optimization of the outlet unloading structure to prevent gaseous cavitation in a high-pressure axial piston pump. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 236: (7) 3459–3473, 2022. https://doi.org/10.1177/09544062211043129.

[3] J Zhang, H Xu, J Chen, W Huang, X Huang, F Lyu, B Xu, M Pan, and Q Su. Modeling and Analysis of the tilt

behavior of the cylinder block in a high-speed axial piston pump. *Mechanism and Machine Theory*, 170: 104735, 2022. https://doi.org/10.1016/j.mechmachtheory.2022.104735.

[4] M Raissi, P Perdikaris, and G E Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378: 686–707, 2019. https://doi.org/10.1016/j.jcp.2018.10.045.

[5] G E Karniadakis, I G Kevrekidis, L Lu, P Perdikaris, S Wang, and L Yang. Physics-informed machine learning. *Nature Reviews Physics*, 3: (6) 422–440, 2021. https://doi.org/10.1038/s42254-021-00314-5.

[6] C Zhao, F Zhang, W Lou, X Wang, and J Yang. A comprehensive review of advances in physics-informed neural networks and their applications in complex fluid dynamics. *Physics of Fluids*, 36: (10) 101301, 2024. https://doi.org/10.1063/5.0226562.

[7] L Lu, X Meng, Z Mao, and G E Karniadakis. DeepXDE: A Deep Learning Library for Solving Differential Equations. *SIAM Review*, 63: (1) 208–228, 2021. https://doi.org/10.1137/19M1274067.

[8] Y Tian, M Woodward, M Stepanov, C Fryer, C Hyett, D Livescu, and M Chertkov. Lagrangian large eddy simulations via physics-informed machine learning. *Proceedings of the National Academy of Sciences*, 120: (34) e2213638120, 2023. https://doi.org/10.1073/pnas.2213638120.

[9] F Lorenzen, A Zargaran, and U Janoske. Potential of physics-informed neural networks for solving fluid flow problems with parametric boundary conditions. *Physics of Fluids*, 36: (3) 037143, 2024. https://doi.org/10.1063/5.0193952.

[10] Y Xiao, L M Yang, C Shu, H Dong, Y J Du, and Y X Song. Least-square finite difference-based physics-informed neural network for steady incompressible flows. *Computers & Mathematics with Applications*, 175: 33–48, 2024. https://doi.org/10.1016/j.camwa.2024.08.035.

[11] E Haghighat, M Raissi, A Moure, H Gomez, and R Juanes. A physics-informed deep learning framework for inversion and surrogate modeling in solid mechanics. *Computer Methods in Applied Mechanics and Engineering*, 379: 113741, 2021. https://doi.org/10.1016/j.cma.2021.113741.

[12] J Wang, Y L Mo, B Izzuddin, and C-W Kim. Exact Dirichlet boundary Physics-informed Neural Network EPINN for solid mechanics. *Computer Methods in Applied Mechanics and Engineering*, 414: 116184, 2023. https://doi.org/10.1016/j.cma.2023.116184.

[13] J Mai, Y Li, L Long, Y Huang, H Zhang, and Y You. Two-dimensional temperature field inversion of turbine blade based on physics-informed neural networks. *Physics of Fluids*, 36: (3) 037114, 2024. https://doi.org/10.1063/5.0190429.

[14] H Bararnia, and M Esmaeilpour. On the application of physics informed neural networks (PINN) to solve boundary layer thermal-fluid problems. *International Communications in Heat and Mass Transfer*, 132: 105890, 2022. https://doi.org/10.1016/j.icheatmasstransfer.2022.105890.

[15] T P Agatho, Pranowo, and A W R Emanuel. Solving the two-dimensional conduction heat transfer problem by using artificial neural network method, 2023 14th Int. Conf. Inf. Commun. Technol. Syst. ICTS, 2023: pp. 228–231. https://doi.org/10.1109/ICTS58770.2023.10330864.

[16] A Almqvist. Fundamentals of Physics-Informed Neural Networks Applied to Solve the Reynolds Boundary Value Problem. *Lubricants*, 9: (8) 82, 2021. https://doi.org/10.3390/lubricants9080082.

[17] L Li, Y Li, Q Du, T Liu, and Y Xie. ReF-nets: Physics-informed neural network for Reynolds equation of gas bearing. *Computer Methods in Applied Mechanics and Engineering*, 391: 114524, 2022. https://doi.org/10.1016/j.cma.2021.114524.

[18] M Rom. Physics-informed neural networks for the Reynolds equation with cavitation modeling. *Tribology International*, 179: 108141, 2023. https://doi.org/10.1016/j.triboint.2022.108141.

[19] Z Yang, L Guo, and P Wong. Application of physics-informed neural network in the analysis of hydrodynamic lubrication. *Friction*, 11: 2022. https://doi.org/10.1007/s40544-022-0658-x.

[20] Y Xi, and R Sun. Inverse problems in hydrodynamics lubrication: parameter identification in the reynold equation by using physics-informed neural networks. *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology*, 13506501241291403, 2024. https://doi.org/10.1177/13506501241291403.

[21] A Schenk. Predicting lubrication performance between the slipper and swashplate in axial piston hydraulic machines. *Open Access Dissertations*, 2014. https://docs.lib.purdue.edu/open_access_dissertations/359.

[22] S Roy, D R Sarkar, C Annavarapu, P Roy, B Lecampion, and D M Valiveti. Adaptive interface-PINNs (AdaI-PINNs) for transient diffusion: applications to forward and inverse problems in heterogeneous media. *Finite Elements in Analysis and Design*, 244: 104305, 2025. https://doi.org/10.1016/j.finel.2024.104305.

[23] X Glorot, and Y Bengio. Understanding the difficulty of training deep feedforward neural networks, 2010. https://www.semanticscholar.org/paper/Understanding-the-difficulty-of-training-deep-Glorot-Bengio/b71ac1e9fb49420d13e084ac67254a0bbd40f83f.

[24] N Hess, and L Shang. Development of a Machine Learning Model for Elastohydrodynamic Pressure Prediction in Journal Bearings1. *Journal of Tribology*, 144: (8) 2022. https://doi.org/10.1115/1.4053815.

Simulation of Hydrostatic Pockets Between the Cylinder Block and Valve Plate of a Piston-type Pump

Haotian Han[1], Thomas Heeger[2], Lizhi Shang[1], and Liselott Ericson[2]

[1]Maha Fluid Power Research Center, Purdue University, West Lafayette, Indiana, USA
[1]E-mail: han592@purdue.edu; shang1@purdue.edu
[2]Fluid and Mechatronic Systems, Linköping University, Linköping, Sweden
[2]E-mail: thomas.heeger@liu.se; liselott.ericson@liu.se

Abstract

Piston-type positive displacement machines are used across diverse applications and operating conditions, making it a critical design challenge to balance the minimization of solid-body contact while maintaining efficiency. This study investigates the potential of hydrostatic pockets between the cylinder block and valve plate to provide dynamically and passively controlled pressure forces, mitigating contact issues at low speeds without excessive losses at high speeds.

Simulations of a baseline pump design revealed persistent solid-body contact under low-speed and high-pressure conditions, indicating the need for enhanced lubrication strategies. Retaining the baseline design, the study examined multiple hydrostatic pocket configurations through simulation, varying their location and quantity. Although the primary focus is on low-speed high-pressure and high-speed high-pressure scenarios, additional operating points at high-speed low-pressure and medium-speed medium-pressure were also considered.

The effectiveness of each design was evaluated on the basis of film thickness, contact pressure, and viscous losses under key operating conditions. The experimental findings from previous studies were used to validate or challenge the conclusions from the simulation results. This paper seeks to deliver a better understanding of the hydrostatic pockets, offering design guidance for optimizing the lubrication management for future piston-type positive displacement machines and informing strategies for improved efficiency and longevity in demanding applications.

Keywords: contact, cylinder block, efficiency, hydrostatic, lubrication, piston-type pump, simulation, valve plate

1 Introduction

Piston-type positive displacement machines are used in various industries, including off-highway vehicles, aviation, automotive, etc. due to their high efficiency and strong reliability over a wide range of operating conditions [1]. However, with the modern trend toward electrification, these machines must now adapt to an even wider range of operating conditions. This expansion has revealed limitations in current axial piston machines, particularly their inability to operate efficiently and reliably at extremely low and high speeds [2].

To combat these challenges, Ericson and Forssell proposed the Floating Piston pumps [3] as shown in Figure 1. This innovative design minimizes the centripetal force acting on the cylinder block, which enables this pump to operate at higher speeds without causing cylinder block unbalance.

Among the three primary lubricating interfaces in Floating Piston pumps: swash-plate/piston-plate, piston-ring/cylinder-bore, and cylinder-block/valve-plate; since the piston plate does not carry radial forces and the friction losses at the piston-ring/cylinder-bore interface are minimal [3], the cylinder-block/valve-plate interface is the most complex [4, 5]. This complexity arises because it must balance the axial force from the displacement chambers and the tilting moment generated by the radial force based on the crowning point location; therefore, the cylinder-block/valve-plate interface requires detailed design to maintain high performance.

At higher rotational speeds, the relative sliding motion between the stationary valve plate and the rotating cylinder block generates hydrodynamic effects within the fluid lubricating interface [6]. These effects lift the cylinder block from the valve plate, reducing solid-body contact. Additionally, the

hydrodynamic effects help balance the net moment acting on the cylinder block, reducing tilting and the risk of metal-to-metal contact [4].

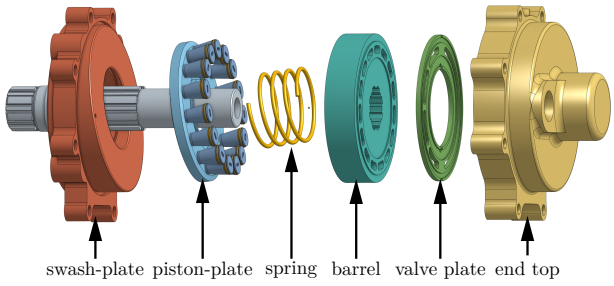


Figure 1: Floating Piston Pump Components [3]

However, at low speeds, the hydrodynamic effects diminish drastically, which leaves the cylinder block under-supported, causing solid-body contact and damage [4]. Moreover, this damage further reduces the lubricating interface's ability to generate hydrodynamic forces and increases leakage, compromising the machine's performance. To address this issue, Achten et al. introduced a novel hydrostatic bearing within the cylinder block sealing land of a floating cup pump [7]. Experimental validation by Achten et al. demonstrated the ability of this design to provide hydrostatic support to the cylinder block, particularly at extremely low speeds when hydrodynamic forces are insufficient, while also improving efficiency [2].

This study uses numerical simulations to explore the working principles and evaluate the effectiveness of hydrostatic pockets between the cylinder block and valve plate in a Floating Piston pump. The focus is on comparing contact severity, power loss, and gap height distribution under low-speed high-pressure and high-speed high-pressure conditions; additionally, the study evaluates performance at other operating points, including high-speed low-pressure and medium-speed medium-pressure conditions. While the study is performed using a Floating Piston pump platform, the findings are applicable to all piston-type hydraulic machines. The findings seek to provide a deeper understanding of the potential of hydrostatic pockets, contributing to the optimization of piston-type positive displacement machines for a wider range of operating conditions.

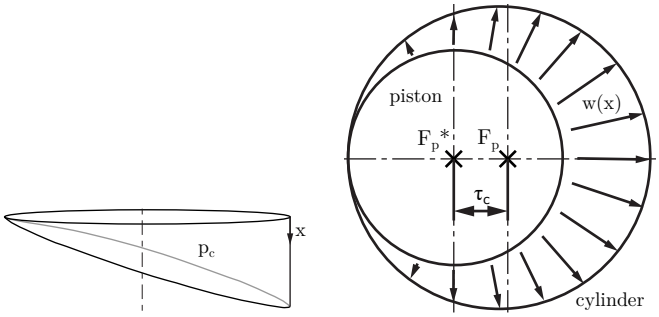
2 Related Literature

This section summarizes literature on Floating Piston pump working principles and the design of the cylinder-block/valve-plate lubricating interface.

2.1 Floating Piston Pump Design

The design of a Floating Piston pump is based on the swashplate-type axial piston pump, where the cylinder block, valve plate, and port block are positioned perpendicular to the drive shaft; however, it incorporates several key design differences. As shown in Figure 1, the Floating Piston pump eliminates slippers, fixing the pistons directly onto an inclined plate, known as the piston plate. Ericson and Forssell have explained the working principles and benefits of this novel piston machine [3].

With the piston heads and sealing rings fixed perpendicularly to the piston plate, displacement chamber pressure forces are directly transferred to the pistons and balanced by the piston-plate/swash-plate lubricating interface. This design prevents side forces on the cylinder block. However, the tilted pistons create a wedged volume in the cylinder bore shown in Figure 2a. The pressurized fluid creates a force imbalance within the displacement chamber, generating a radial force that acts perpendicular to the cylinder bore and on the cylinder block. With a properly positioned cylinder-block/shaft reaction point, the moment generated by this radial force on the cylinder block remains symmetric throughout a revolution and can be balanced [3].



(a) The wedge volume *(b) Force per unit of one piston*

Figure 2: The Pressurized Volume in a Piston Chamber [3]

Moreover, in traditional axial piston machines, the mass of the piston-slipper assembly, as well as the fluid volume, generates centrifugal forces that the cylinder block must balance. These forces create a moment and introduce additional tipping motivation to the cylinder block. Since these forces are proportional to the square of the shaft speed, traditional machines are highly sensitive to rotational speed, creating a barrier to achieving higher operating speeds. In contrast, the Floating Piston pump connects the pistons on a piston plate, balancing their centrifugal forces and eliminating centrifugal forces transmitted from the piston plate to the cylinder block. Additionally, due to the small dead volume and low piston strokes, the centrifugal forces and moments generated by the fluid volume are also minimized [3]. As a result, this optimized design is expected to enable much higher rotational speeds without requiring an overpowering hold-down central spring, which could otherwise cause additional power losses.

2.2 Block Sealing Land Design

The cylinder-block/valve-plate lubricating interface is one of the most critical design aspects of piston-type positive displacement machines. This interface bears the responsibility of providing pressure forces to lift the cylinder block away from the valve plate, sealing high-pressure fluid to prevent leakage into the casing, and balancing moments acting on the cylinder block. These functions are achieved through the thrust effect of rapid gap height changes and the wedge effect result-

ing from converging and diverging surfaces in the circumferential direction. The gap height of the cylinder-block/valve-plate lubricating interface is determined by various factors, including moment and force balances, which are coupled with solid-body deformation. Ultimately, the gap height influences the elasto-hydrodynamic effects, which directly impact power losses in this interface through viscous dissipation, volumetric losses, and solid-body contact friction.

2.2.1 Hydrostatic Compensation Ratio

Traditionally, an analytical starting point of the block sealing land design is based on the compensation ratio ζ, which compares the loading forces F_{load} and the compensating hydrostatic forces F_{comp}, see Eq. 1 and Figure 3.

$$\zeta =: \frac{F_{\text{load}}}{F_{\text{comp}}} \tag{1}$$

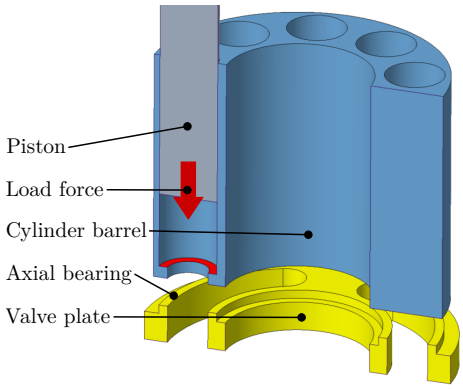


Figure 3: Load Force and Axial Bearing Between the Valve Plate and Cylinder Block [8].

The compensation ratio only considers hydrostatic forces. The load originates from the pressurized chambers pushing the cylinder block against the valve plate, and the compensation force is defined by the area of the hydrostatic bearing [1].

Compensation ratios above 100 % lift the cylinder block off the valve plate, causing undesired leakage from the high-pressure port. Thus, compensation ratios need to be smaller than 100 %. On the other hand, solid contact between the valve plate and cylinder barrel causes wear. The combination of the two requirements leads to compensation ratios just below 100 % [1, 8, 9]. However, note that there are different calculation methods for the compensation ratio [8], and Ernst and Vacca [5] claimed that the hydrodynamic forces in pumps are drastically underestimated, with Vacca and Franzoni [10] stating that typical hydrodynamic forces are in the range of 10 to 20 % of the load forces.

2.2.2 Hydrodynamic Effects

In a traditional cylinder-block/valve-plate sealing land design, the fluid forces in the lubricating interface depend not only on the pressure from the inlet, outlet ports as well as the casing, but also, more importantly, on the elasto-hydrodynamic wedge and squeeze effects between the valve plate and cylinder block surfaces [5, 11]. However, generating pressure from the wedge effect requires sufficient relative sliding velocity between the two opposing surfaces; consequently, at low speeds, the cylinder-block/valve-plate lubricating interface loses its ability to create effective hydrodynamic wedge pressure, often resulting in metal-to-metal contact.

2.2.3 Design with Hydrostatic Pockets

In 2010, Innas introduced a new feature within the lubricating interface incorporating additional chambers in the sealing land, called hydrostatic pockets, along with grooves that connect these pockets to the displacement chamber [7] as shown in Figure 4.

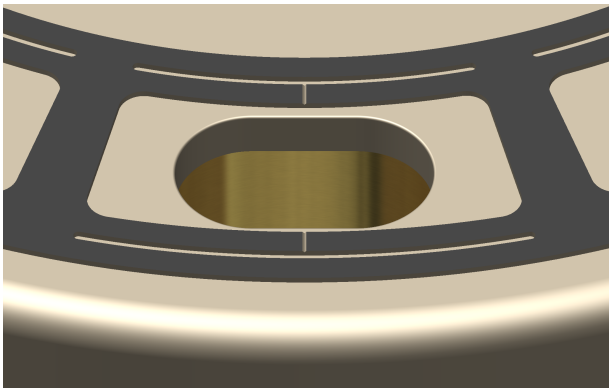


Figure 4: Cylinder Block with Pockets

The pressure in these hydrostatic pockets is determined by multiple flow paths. First, the grooves between the displacement chamber and the hydrostatic pockets act as constant orifices. As the displacement chamber pressurizes, hydraulic oil flows into the pocket and pressurizes it. Simultaneously, the sealing lands between the pocket and the displacement chamber as well as between the pocket and the casing act as variable orifices. These variable orifices are passively controlled by the lubricating film gap height distribution [2] as shown in Figure 5 below.

This design provides several benefits, as suggested by Achten et al. [2]. First, the hydrostatic pockets provide additional pressure forces to support the cylinder block to prevent contact with the valve plate even at extremely low shaft speeds, which leads to the diminishing of contact power losses and surface damage. Additionally, the hydrostatic pockets can dynamically balance the cylinder block; when the cylinder block is tilted, the local gap height becomes extremely small, causing the variable orifices' areas connecting the pockets in the outer edge of the sealing land and the casing to reduce to nearly zero. This effectively cuts off the outlet flow from the outer pockets, and since the constant orifices connecting the pockets and the displacement chambers are independent of the gap height, the inlet flow to the pockets allows the pocket pressure to reach a level similar to that of the displacement chamber, which provides additional force to counteract the cylinder block's moment imbalance. Conversely, when less force

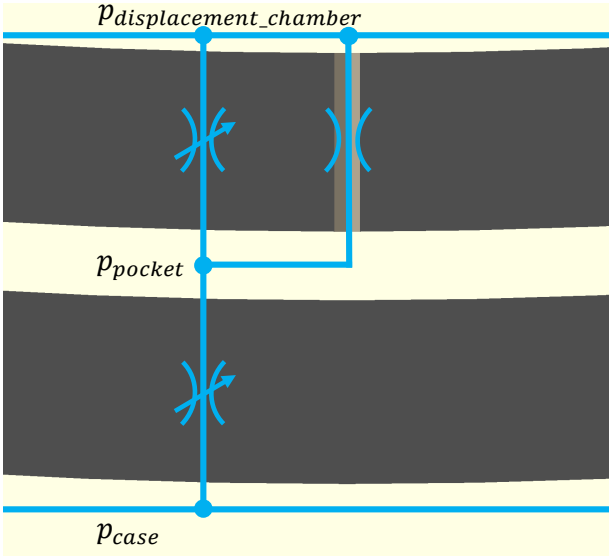


Figure 5: Hydraulic Circuit of Hydrostatic Pocket

is needed, a higher gap height increases the variable orifices' area, balancing the inlet and outlet flow to the pocket and reducing the hydrostatic pressure.

3 Modeling Approach

The modeling approach for simulating the cylinder-block/valve-plate lubricating interface was developed using a multi-physics simulation suite, Multics [12]. Although Multics is capable of simulating all three lubricating interfaces in a piston machine, as well as the dynamic coupling between each module, this study is limited to the cylinder-block/valve-plate interface, and the information flow is shown in Figure 6 [11].

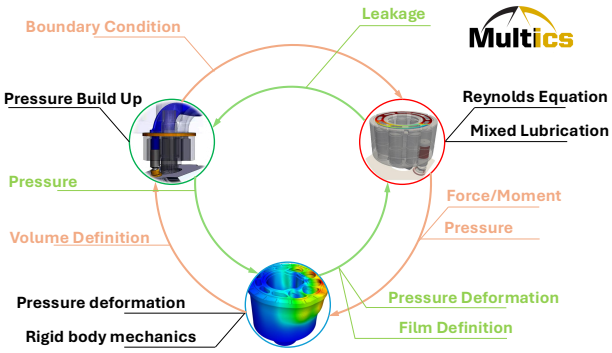


Figure 6: Modeling Information Flow

3.1 Lumped Parameter Model

The pressure within the displacement chambers is simulated using Hopsan [13], which employs a lumped parameter model as shown in Figure 7. The model computes the translational movements of the pistons and solves the continuity equation for each chamber. Pseudo-cavitation and the pressure dependency of the bulk modulus are considered. The openings of the chambers to each port are modeled as variable orifices. The model assumes no leakage and friction. The lumped parameter model is used to compute the pressure profile under consideration of commutation effects for each operating point. Multics then reads in the pressure profile in the displacement chambers as a look-up table based on the shaft rotation angle.

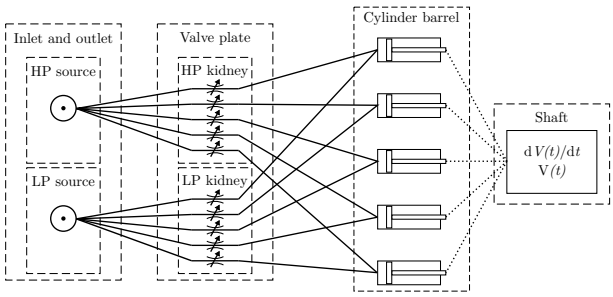


Figure 7: Lumped Parameter Model Structure. [14]

The pressure from the displacement chambers, used as the upstream pressure for the hydrostatic pockets, always connects to the pockets through constant-area orifices, which the area is determined for optimal power efficiency based on Achten et al.'s experimental results [2]. The mass flow rate through the orifices into a given hydrostatic pocket, $\dot{m}$, is calculated using the orifice equation shown in Eq. 2.

$$\dot{m} = C_d A \sqrt{2\rho\Delta p} \tag{2}$$

The hydrostatic pockets are modeled as fluid volumes, and the pressure is calculated using the isothermal pressure build-up equation, as shown in Eq. 3. Since there is no volume derivative for the hydrostatic pockets, the last term can be omitted.

$$\frac{\partial p}{\partial t} = \frac{K}{V}\left(\frac{1}{\rho}\dot{m}_{in} - \frac{1}{\rho}\dot{m}_{out} - \frac{\partial V}{\partial t}\right) \tag{3}$$

The pressures from the inlet ports, outlet ports, casing, displacement chambers, and hydrostatic pockets serve as pressure boundaries for the lubricating interface model. These pressures are also used to apply loads to the cylinder block.

3.2 Lubricating Film Model

There are three lubricating interfaces in a Floating Piston machine. However, since the focus of this study is on the effect of hydrostatic pockets on the cylinder block, and the majority of forces from the piston and piston plate do not transfer to the cylinder block, only the cylinder-block/valve-plate lubricating interface is simulated.

The traditional form of the Reynolds equation is written in terms of the pressure, as shown by Hamrock et al. [15]. However, this formulation is unable to simulate the cavitated regions in the lubricating interface precisely, leading to an incorrect flow estimation. Therefore, Ransegnola proposed the density-based Reynolds equation accounting for cavitation effects and providing accurate flow predictions [12]. In addition, Ransegnola utilized statistical mixed lubrication relations to formulate the universal mixed Reynolds equation, as shown in Eq. 4, which is used in Multics [16–18].

$$\nabla \cdot \left(\phi_p \frac{Kh^3}{12\mu} \nabla \rho \right) = \nabla \cdot (\rho \boldsymbol{v}_m (\phi_R R_q + \phi_c h)) + \nabla \cdot \left(\rho \frac{\phi_s}{2} R_q (\boldsymbol{v}_t - \boldsymbol{v}_b) \right) + \frac{\partial \rho (\phi_R R_q + \phi_c h)}{\partial t} \tag{4}$$

The Reynolds equation predicts the pressure distribution within the lubricating interface and is critical for assessing power loss. The power loss considered in this study arises from the energy dissipation due to the viscous shear of the lubricating fluid with the consideration of the shear loss from high-pressure fluid leakage [19]. Viscous dissipation is a major contributor to power loss at the lubricating interface and is ultimately converted into heat. The governing equation for viscous dissipation in polar coordinates, as used in Multics, is presented below in Eq. 5 and Eq. 6.

$$P_{\text{loss}} = \iint \mu \phi_D \, dz \, dA \tag{5}$$

$$\phi_D = \left(\frac{\partial u}{\partial z} \right)^2 + \left(\frac{\partial v}{\partial z} \right)^2 + \frac{1}{r^2} \left(\frac{4}{3} u^2 + v^2 \right) \tag{6}$$

More importantly, for this study, the pressure distribution can be used to predict the volumetric leakage through the interface, which is crucial for simulating the correct behavior of the hydrostatic pockets. This is because the equivalent variable orifices connecting the hydrostatic pockets to the displacement chamber and casing are determined by the volumetric flow in the lubricating interface.

Additionally, Multics considers both penetration contact and asperity contact. Specifically, the asperity contact arises from the interaction of surface micro-structures due to roughness. Multics calculates the equivalent roughness between surfaces, incorporates material properties, and uses these to generate an asperity contact curve, from which the asperity contact pressure is determined based on the gap height [20].

Lastly, the combined pressure distribution from the lubricating interface and contact model is used for body dynamics calculations as well as solid-body deformation analysis to estimate the accurate gap height in the lubricating interface.

3.3 Body Dynamics Model

The axial forces and moments arising from the pressurized fluid in the displacement chambers and the lubricating interface are used to solve the equations of conservation of linear and angular momentum for the cylinder block body, as shown in Eq. (7) and Eq. (8); encompassing three degrees of freedom, allowing the cylinder block to move in the axial direction and tilt around the cylinder-block/shaft reaction point, or the crowing point, which also balances the block's radial forces and provides the drive torque to rotate the cylinder block [9].

$$m \frac{d\boldsymbol{v}_{IF}}{dt} = \sum \boldsymbol{F}_{l,IF} \tag{7}$$

$$[I_{BF}] \frac{d\omega_{BF}}{dt} = \sum \boldsymbol{M}_{l,BF} \tag{8}$$

The radial forces and moments from the wedge-shaped fluid inside the displacement chamber shown in Figure 8, caused by the tilted piston and piston sealing ring, are calculated using Eq. 9 and Eq. 10 [3].

$$F_{R,i} = \pi R_c^2 \tan \alpha \, p_{c,i} \tag{9}$$

$$F_{R,\text{tot}} = \sum_{i=1}^{z} F_{R,i} \tag{10}$$

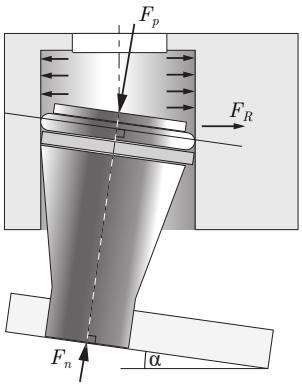


Figure 8: The Radial Force F_R [3]

Moreover, to accurately evaluate contact severity and power loss within the lubricating interfaces, it is essential to account for solid-body deformation [6]. In this simulation model, the half-space method based on the Hertzian contact stress is employed to evaluate the deformation in the cylinder block and valve plate caused by local fluid film pressure and contact pressure. Finally, the resulting cylinder block positions, velocities, and deformation are fed back to assess the influence of squeeze and wedge phenomena on the lubricating interface through the Reynolds equation.

4 Results and Discussion

4.1 Simulation Setup

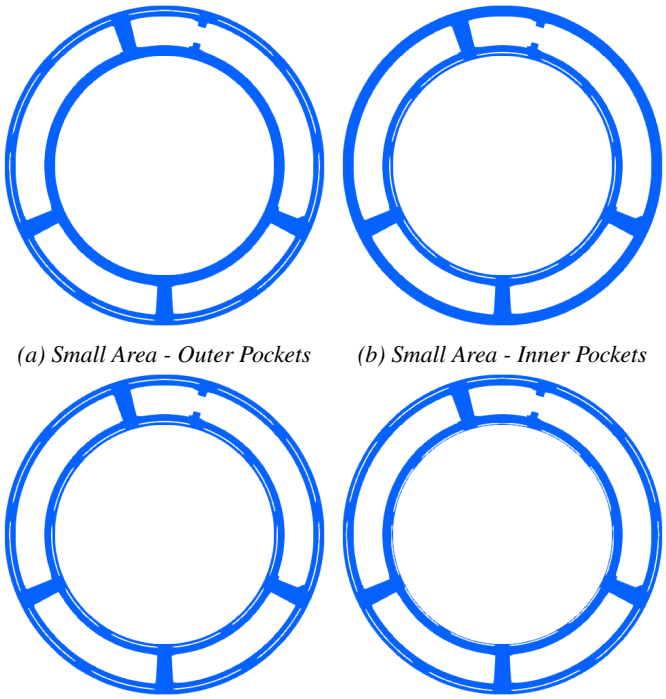
Four groups of operating conditions are tested, as outlined in Table 1: Low-Speed High-Pressure (LSHP), High-Speed Low-Pressure (HSLP), High-Speed High-Pressure (HSHP), and Medium-Speed Medium-Pressure (MSMP). These conditions are defined by the rotational speed and outlet pressure, where low-speed is 300 rpm, medium-speed is 2000 rpm, and high-speed is 6000 rpm, while low-pressure is 20 bar, medium-pressure is 200 bar, and high-pressure is 350 bar.

To evaluate the effectiveness of the hydrostatic pockets, a baseline pump without pockets is simulated alongside configurations where pockets are positioned on the inside (Inner Pockets), outside (Outer Pockets), or both sides (Both Pockets) of the sealing land. Each configuration is further analyzed based on the pocket placement relative to the displacement chamber: one set positioned closer to the chamber (Small

Table 1: Simulated Pocket Designs and Operating Conditions.

Operating Condition	Baseline	Small Area Inner Pockets	Small Area Outer Pockets	Small Area Both Pockets	Large Area Inner Pockets	Large Area Outer Pockets	Large Area Both Pockets
Low-Speed High-Pressure	✓	✓	✓	✓	✓	✓	✓
High-Speed Low-Pressure	✓	✓	✓	✓	✓	✓	✓
High-Speed High-Pressure	✓	✓	✓	✓	✓	✓	✓
Medium-Speed Medium-Pressure	✓	✓	✓	✓	✓	✓	✓

Area) and the other set positioned farther away (Large Area) representing the different areas of the sealing land trapped between the pockets and the displacement chambers, as illustrated in Figure 9.



(a) Small Area - Outer Pockets *(b) Small Area - Inner Pockets*

(c) Small Area - Both Pockets *(d) Large Area - Both Pockets*

Figure 9: Hydrostatic Pockets Location Examples

To create a fair comparison between the outer and inner pockets, the locations are determined by maintaining the lubricating interface area between the hydrostatic pockets and the displacement chambers. This assumption is based on the premise that the additional hydrostatic pressure in the lubricating film generated by the pockets is directly proportional to this area. Lastly, to accurately evaluate the efficacy of the hydrostatic pockets as well as the differences between different configurations, local meshing surrounding the pockets in the lubricating film model is refined.

To ensure the applicability of this study, a sensitivity analysis is performed under the medium-speed medium-pressure operating condition to determine the crowning point location for the remaining simulations. As introduced in the modeling approach section 3.3, the crowning point directly influences the moments applied to the cylinder block, impacting its stability and overall performance. As shown in Figure 10, the viscous power loss and the moments acting on the cylinder block are normalized for ease of comparison.

The green zone in the sensitivity analysis represents the optimal scenario, indicating that solid-body deformation, manufacturing tolerances, wear, and other factors are functioning without causing a large shift at the moments applied to the cylinder block. Meanwhile, the orange zone shows the scenarios where small disturbances may introduce some deviations, such as manufacturing imperfections and wear, but are still within acceptable limits. Finally, the red zone represents the extreme conditions where the cylinder block experiences significant moment imbalance, resulting in instability and further decreased performance. It can be seen that as the moment applied on the cylinder block increases, the viscous power loss for the baseline machine increases exponentially; meanwhile, with the addition of the hydrostatic pockets, the power losses are kept at a constantly low level.

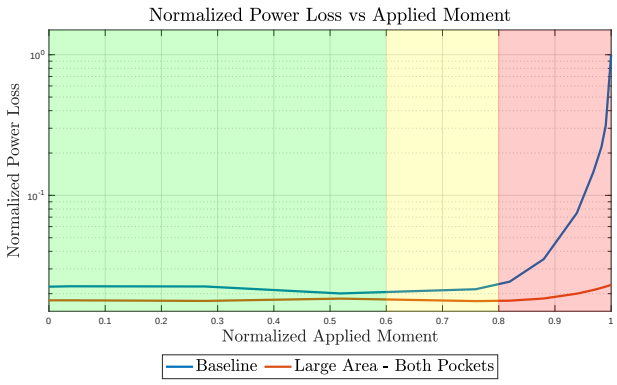


Figure 10: Normalized Power Loss vs. Applied Moment

For this study, simulations were conducted assuming a non-ideal crowning point for several reasons. First, testing the hydrostatic pockets under such conditions ensures this feature will remain effective even in the most non-optimal situations. Second, since the piston-seal/cylinder-bore friction forces and oil inertia forces are not included in these simulations, assuming a non-ideal crowning point would help to increase the robustness of this study. Lastly, while the baseline cylinder block in a Floating Piston pump exhibits mostly balanced moments, the study is made more general when the crowning point is placed sub-optimally due to the complexities of a typical axial piston unit, providing more reliable insights that can be applied to a wider range of applications.

4.2 Low-Speed High-Pressure Operating Condition

As discussed in Section 2.2.2, low-speed operating conditions are a major challenge for the cylinder-block/valve-plate lubric-

ating interface in a piston machine due to the lack of the hydrodynamic wedge effect. Furthermore, the high outlet pressure presents a greater moment imbalance for the cylinder block caused by radial chamber forces. Thus, at this operating condition, the benefits from the addition of hydrostatic pockets become more pronounced. The lowest simulated speed performed is 100 rpm, and the results show similar behavior to the 300 rpm case discussed in this section; however, due to the nature of the simulation, decreasing shaft speed significantly increases wall time.

The maximum gap height of the cylinder-block/valve-plate lubricating interface is shown in Figure 11, while the minimum gap height is shown in Figure 12 over one displacement chamber period, which is 360 degrees divided by the number of pistons. The baseline case results show that the maximum gap height is close to 90 micrometers, while the minimum gap height is below 0.2 micrometers, indicating that the block is severely tilted. With the addition of hydrostatic pockets, not only does the minimum gap height increase and the maximum gap height decrease, but the cylinder block also exhibits a much lower amplitude of tilting oscillation, indicating that the unbalancing motion is significantly reduced.

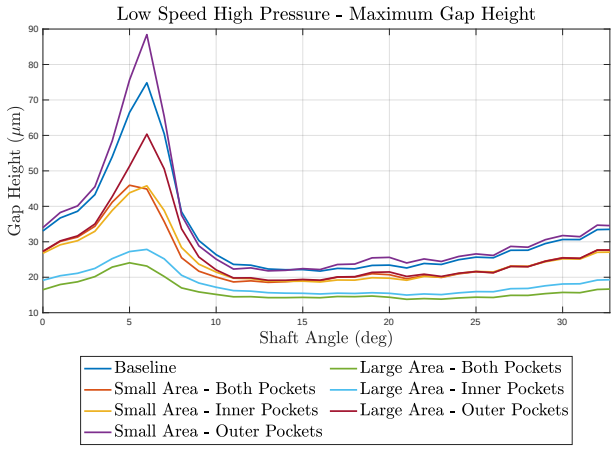


Figure 11: LSHP - Maximum Gap Height, h_{max}

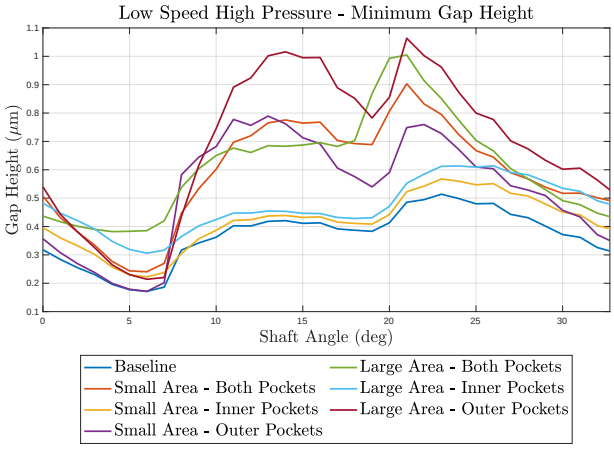


Figure 12: LSHP - Minimum Gap Height, h_{min}

The relief of strong tilting motion is also reflected in the metal-to-metal contact behavior. As shown in Figure 13, the baseline case experiences the highest magnitude of contact pressure, indicating that the lubricating film is unable to balance the moment. It is worth noting that in both the Large Area - Inner Pockets and Large Area - Both Pockets cases, no solid-body contact is observed.

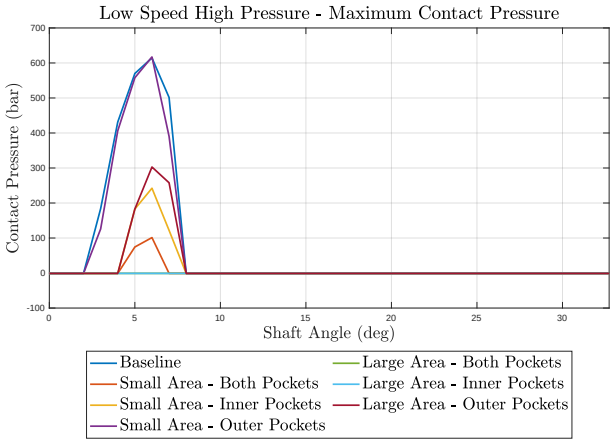


Figure 13: LSHP - Maximum Contact Pressure, p_{cont}

The condition in which the hydrostatic pockets are pressurized is an important aspect of this study. Figure 14 demonstrates that, initially, when the local gap height is large, the pressure in the outer pocket is maintained at approximately the average level between the casing and displacement chamber; and as the pocket transitions into the low gap height region, its pressure increases and approaches the level of the displacement chamber. For the inner pockets, as illustrated in Figure 15, the relatively neutral and stable gap height distribution towards the center of the cylinder block prevents a significant pressure increase. This distinction between inner and outer pockets showcases the dynamic behavior of pressurization in the hydrostatic pockets and their dependency on local film thicknesses.

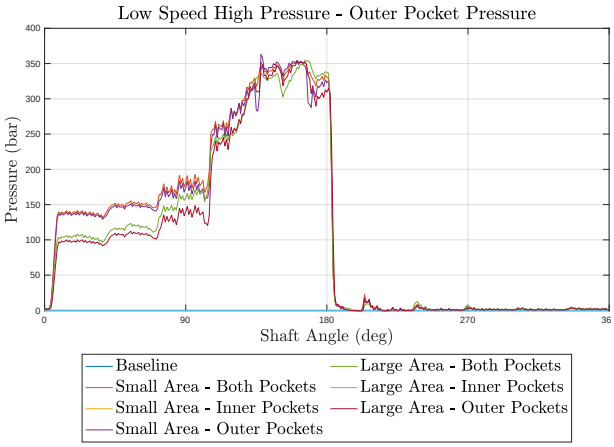


Figure 14: LSHP - Outer Pocket Pressure, p_{out}

This behavior can be further observed through the pressure distribution shown in Figure 16, where the colors represent pressure levels, with red being high pressure and blue being low pressure. The high-pressure delivery ports are located on the left side of the valve plate, and the cylinder block is rotating counterclockwise, as indicated in the figures. During the

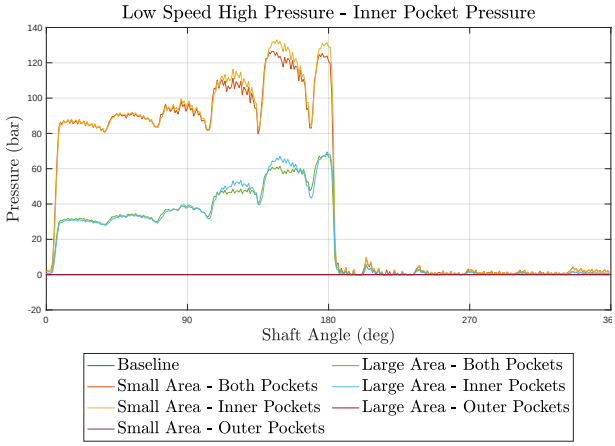
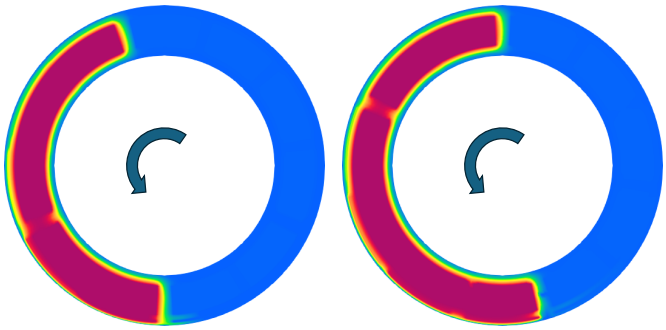


Figure 15: LSHP - Inner Pocket Pressure, p_{in}

high-pressure zone, as the displacement chamber rotates from the high gap height region to the low gap height region, the hydrostatic pocket begins to pressurize as its color approaches the color of the displacement chamber. This dynamic behavior demonstrates that hydrostatic pockets adapt their pressure and provide hydrostatic support depending on the gap height, passively mitigating tilting.



(a) 0 Deg *(b) 16 Deg*

Figure 16: LSHP - Large Area - Both Pockets - Film Pressure

Moreover, as shown in Table 2, the viscous power loss from the cylinder-block/valve-plate lubricating interface is reduced with the addition of the hydrostatic pockets, with the greatest reductions being the cases with the Large Area - Inner Pockets and Large Area - Both Pockets configurations. There are two key findings from the results. First, hydrostatic pockets not only generate additional hydrostatic force within their areas but also help elevate the pressure in the surrounding lubricating film, particularly when the gap heights are low, creating an additional boundary for the lubrication fluid between the pockets and displacement chambers. As shown in Figure 17, the pressure in the pockets and the fluid film approach similar levels as the displacement chamber. Therefore, locating the hydrostatic pockets farther away from the displacement chambers further increases the hydrostatic force when necessary, improving cylinder block balance.

The second point may seem counterintuitive. Although locating the pockets closer to the outer edge of the cylinder block sealing land generates larger balancing moments, the squeeze

Table 2: LSHP - Viscous Power Loss

Pocket Setup	Power Loss Reduction [%]
Baseline	-
Small Area - Inner Pockets	57.1
Small Area - Outer Pockets	-26.6
Small Area - Both Pockets	53.9
Large Area - Inner Pockets	82.5
Large Area - Outer Pockets	41.7
Large Area - Both Pockets	85.4

effect of the bearing results in high pressure near the outer edge, where gap heights are minimal and oscillating. This generated hydrodynamic pressure diminishes the contribution of the hydrostatic pockets, reducing their effectiveness. In contrast, positioning the pockets on the inner side of the sealing land, while having a shorter moment arm, enhances the pressure in the lubricating interface where all hydrodynamic effects are weak. Therefore, locating the hydrostatic pockets on the inner side of the cylinder block is more effective for mitigating block tilting and solid-body contact.

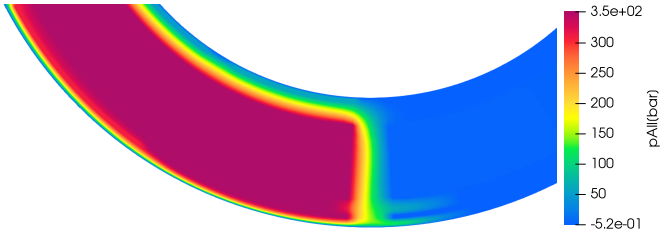


Figure 17: LSHP - Large Area - Both Pockets - Film Pressure Zoomed at 0 Deg

4.3 High-Speed High-Pressure Operating Condition

High-speed high-pressure condition also leads to challenges in lubricating the cylinder-block/valve-plate interface. While elevated rotational speeds enhance hydrodynamic wedge effects, helping balance the cylinder block and reduce viscous and contact losses, the simultaneous presence of high pressures increases the risk of moment imbalances, potentially compromising stability.

Simulation results in Figure 18 show that, under HSHP conditions, the baseline configuration exhibits large film thickness variation, causing increased tilting. The introduction of hydrostatic pockets improves the cylinder block stability by providing additional pressure support to counteract these tilting moments.

More specifically, Figure 19 illustrates how pocket pressure dynamically adjusts to film height variations in order to maintain more uniform gap height distributions and enhance moment balancing.

Table 3 compares viscous power losses across different hydrostatic pocket configurations, consistent with the low-speed high-pressure operating condition, showing that the Large

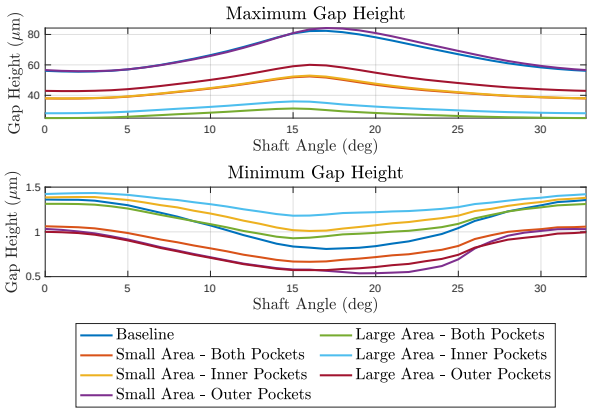


Figure 18: HSHP - Gap Heights, h_{max}, h_{min}

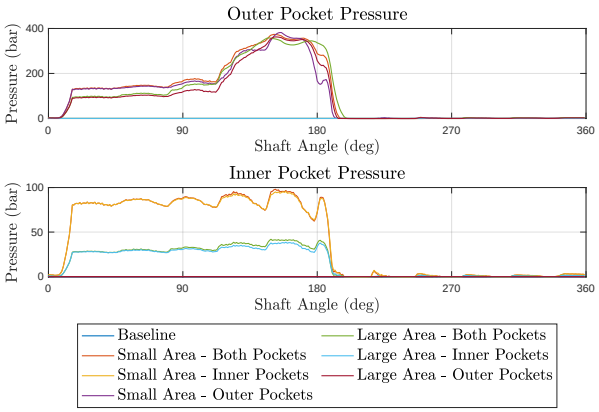


Figure 19: HSHP - Pocket Pressures, p_{out}, p_{in}

Area - Inner Pockets and Large Area - Both Pockets cases achieve the greatest reductions. These results further emphasize the importance of the hydrostatic pocket placements to maximize performance.

Table 3: HSHP - Viscous Power Loss

Pocket Setup	Power Loss Reduction [%]
Baseline	-
Small Area - Inner Pockets	59.8
Small Area - Outer Pockets	-4.5
Small Area - Both Pockets	60.9
Large Area - Inner Pockets	73.3
Large Area - Outer Pockets	50.7
Large Area - Both Pockets	76.6

4.4 High-Speed Low-Pressure Operating Condition

Compared with the LSHP and HSHP operating conditions, the high-speed low-pressure case is inherently more stable due to the strong hydrodynamic edge effects resulted by high relative sliding velocity and low radial chamber forces and moments, as shown in Figure 20.

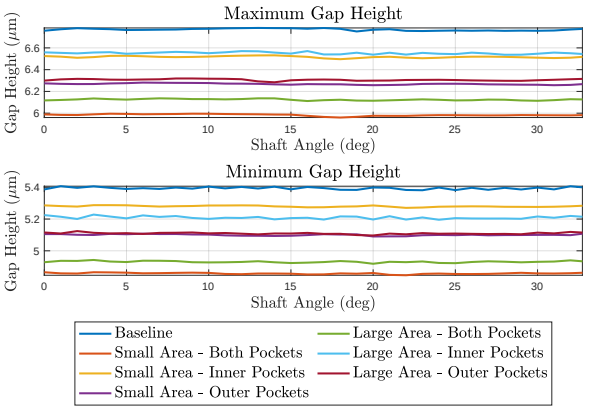


Figure 20: HSLP - Gap Heights, h_{max}, h_{min}

The pressure within the hydrostatic pockets does not increase significantly as shown in Figure 21, indicating that the hydrostatic pockets are not over-pressurized when there is no motivation to create additional balancing moments on the cylinder block.

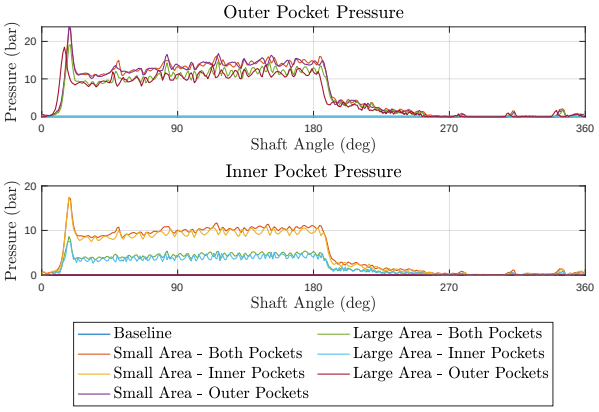


Figure 21: HSLP - Pocket Pressures, p_{out}, p_{in}

Lastly, since there is no excessive moment tilting the cylinder block, the power loss reduction due to the hydrostatic pockets is less significant. However, the slight reduction, as shown in Table 4 suggests that hydrostatic pockets can still provide benefits even when the cylinder block is inherently balanced.

Table 4: HSLP - Viscous Power Loss

Pocket Setup	Power Loss Reduction [%]
Baseline	-
Small Area - Inner Pockets	0.42
Small Area - Outer Pockets	2.82
Small Area - Both Pockets	1.92
Large Area - Inner Pockets	-0.01
Large Area - Outer Pockets	3.44
Large Area - Both Pockets	3.96

4.5 Medium-Speed Medium-Pressure Operating Condition

Unlike some other operating conditions examined, the medium-speed medium-pressure operation conditions do not

typically lead to severe failure of the pump, such as strong metal-to-metal contact or excessive gap height variation; however, since hydraulic machines frequently operate under these conditions, the power efficiency is critical. Table 5 presents the power loss reduction with and without the incorporation of hydrostatic pockets, and the results are aligned with the findings under other operating conditions.

Table 5: MSMP - Viscous Power Loss

Pocket Setup	Power Loss Reduction [%]
Baseline	-
Small Area - Inner Pockets	90.7
Small Area - Outer Pockets	59.1
Small Area - Both Pockets	90.8
Large Area - Inner Pockets	95.1
Large Area - Outer Pockets	90.3
Large Area - Both Pockets	95.9

5 Discussion

A summary of the result comparison between the baseline design and the Large Area - Both Pockets case is presented in Figures 22 and 23. The figures show the largest values of maximum gap height and the smallest values of minimum gap height over one revolution, respectively. The data suggest that across various operating conditions, the maximum gap heights for the Large Area - Both Pockets case are consistently lower than those for the baseline case. Conversely, the minimum gap heights for the Large Area - Both Pockets case are either similar to or higher than those of the baseline case, which effectively prevents solid contact.

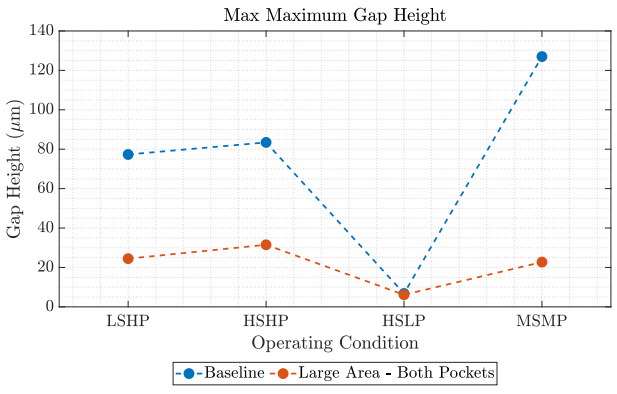


Figure 22: Maximum h_{max}

This study's findings align with Achten et al.'s experimental work on Floating Cup pumps with hydrostatic pocket integration, demonstrating that the addition of hydrostatic pockets can create a more uniform film thickness, reduce viscous power loss, enhance cylinder block stability, and mitigate substantial solid-body contact on the sealing land [2]. This paper examined multiple hydrostatic pocket configurations in addition to Achten et al.'s study, confirming that the pressure distribution in the lubricating interfaces and the pressure in the

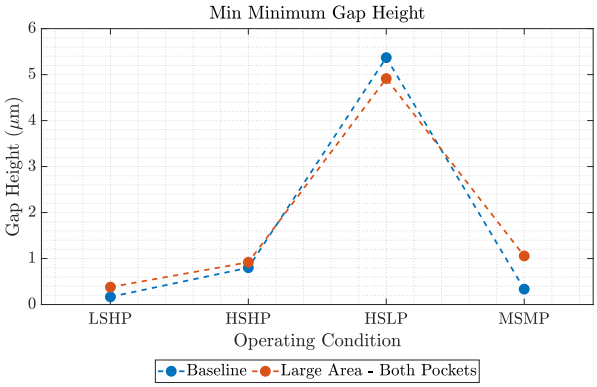


Figure 23: Minimum h_{min}

pockets shows that the hydrostatic pockets provide a dynamic and passive stabilizing effect to the cylinder block.

The placement of hydrostatic pockets is a critical design consideration. Although it may seem intuitive that positioning pockets toward the outer edge would generate larger balancing moments, our results suggest that the squeeze and wedge effects in the low gap height region already generate sufficient pressure, diminishing the pockets' influence. Instead, placing the hydrostatic pockets on the inside of the displacement chambers can supplement lubrication where hydrodynamic effects are the weakest. This finding is consistent across all four operating conditions simulated, and interestingly, in Achten et al.'s experimental study, only inner hydrostatic pockets were employed [2]. When the hydrostatic pockets are positioned farther from the displacement chamber, namely, the Large Area - Inner Pockets and Large Area - Both Pockets cases, the highest reductions in power losses are yielded (Table 3). Furthermore, since the Large Area - Inner Pockets and Large Area - Both Pockets configurations exhibit similar performance in all aspects, the Large Area - Inner Pockets setup alone would be sufficient when considering manufacturing cost-effectiveness.

6 Conclusion

This study showcases the effectiveness of hydrostatic pockets by improving the lubrication performance of piston-type machines, supporting and extending Achten et al.'s [2] experimental findings. Using advanced computational modeling techniques, the impact of the hydrostatic pockets is analyzed under several operating conditions, which resulted in improved gap height uniformity, pressure distribution, contact behavior, and overall efficiency; more generally, the addition of hydrostatic pockets would allow piston-type machines to broaden their range of operating speeds.

Out of all the setups simulated, placing the hydrostatic pockets on the inside of the cylinder block sealing land and farther away from the displacement chambers achieved the best performance in maintaining gap height, preventing metal-to-metal contact, and reducing power loss. The results demonstrate the importance of the placement of the hydrostatic pockets for lubrication performance and provide guidance for designing future piston-type machines.

However, the study also has limitations. Future research will refine the simulation model by including different hydrostatic pocket sizes, piston-seal/cylinder-bore lubricating interface, as well as the influence matrix and thermal deformation to further improve prediction accuracy. Experimental tests on the Floating Piston pump, incorporating hydrostatic pockets are intended to focus on evaluating efficiency across various shaft speeds. Additionally, torque losses at low speeds will be computed to validate the separation of the surfaces, and volumetric losses will be analyzed as they can can correlate with gap heights. Furthermore, a visual inspection of component surface wear will be carried out.

Acknowledgement

Linköping University's contribution was funded by the Swedish Electromobility Centre with grant number 13070 and the Swedish Energy Agency (Energimyndigheten) with grant number P2023-00594. The share of simulations run at Linköping University were enabled by resources provided by the National Supercomputer Centre (NSC), funded by Linköping University.

Nomenclature

Designation	Denotation	Unit
$\boldsymbol{A}$	Orifice area	[m^2]
α	Tilt angle of the piston plate	[rad]
$\boldsymbol{C}_d$	Discharge coefficient	[-]
$\eta_{\text{tot,p}}$	Total power efficiency loss	[-]
$\boldsymbol{F}_{\text{comp}}$	Hydrostatic compensation force	[N]
$\boldsymbol{F}_{\text{load}}$	Load force on the cylinder block	[N]
$\boldsymbol{F}_{l,\text{IF}}$	Sum of forces in inertia frame	[N]
$\boldsymbol{F}_R$	Radial force due to wedge-shaped fluid in displacement chamber	[N]
$\boldsymbol{h}$	Lubricating film thickness	[m]
$\boldsymbol{I}_{\text{BF}}$	Inertia tensor of the block frame	[kg·m^2]
$\boldsymbol{K}$	Bulk modulus	[Pa]
$\boldsymbol{M}_{l,\text{BF}}$	Sum of moments acting on the block frame	[N·m]
$\boldsymbol{m}$	Mass of the cylinder block	[kg]
$\dot{\boldsymbol{m}}$	Mass flow rate	[kg/s]
μ	Dynamic viscosity	[Pa·s]
ω_{BF}	Angular velocity of the block frame	[rad/s]
$\boldsymbol{p}$	Pressure	[Pa]
$\boldsymbol{p}_{\text{in}}$	Inlet pressures	[Pa]
$\boldsymbol{p}_{\text{out}}$	Outlet pressures	[Pa]
$\boldsymbol{P}_{\text{loss}}$	Viscous power loss	[W]
$\Delta\boldsymbol{p}$	Pressure difference across the orifice	[Pa]
$\phi_p, \phi_R, \phi_c, \phi_s$	Mixed lubrication flow factors	[-]
ϕ_D	Dissipation function	[-]
$\boldsymbol{r}$	Radial coordinate in polar form	[m]
$\boldsymbol{R}_c$	Radius of the sealing ring	[m]
$\boldsymbol{R}_q$	Surface roughness parameter	[μm]
ρ	Fluid density	[kg/m^3]
$\boldsymbol{t}$	Time	[s]
$\boldsymbol{u}, \boldsymbol{v}$	Velocity components in the interface	[m/s]
$\boldsymbol{V}$	Volume of hydrostatic pocket	[m^3]
$\boldsymbol{V}_d$	Displacement volume	[m^3]
$\boldsymbol{v}_{\text{IF}}$	Velocity of the inertia frame	[m/s]
$\boldsymbol{v}_m$	Mean velocity of the fluid in the interface	[m/s]
$\boldsymbol{v}_t, \boldsymbol{v}_b$	Surface velocities of the top and bottom plates	[m/s]
ζ	Hydrostatic compensation ratio	[-]
LSHP	Low speed high pressure	
HSLP	High speed low pressure	
HSHP	High speed high pressure	
MSMP	Medium speed medium pressure	

References

[1] J. Ivantysyn and M. Ivantysynova. *Hydrostatic Pumps and Motors, Principles, Designs, Performance, Modeling, Analysis, Control and Testing*. Academia Books International, New Delhi, India, 2001.

[2] P. Achten, R. Mommers, J. Potma, and J. Achten. Experimental investigation of a hydrostatic bearing between barrels and port plates in floating cup axial piston pumps. In *BATH/ASME 2020 Symposium on Fluid Power and Motion Control*, 2020.

[3] Liselott Ericson and Jonas Forssell. A Novel Axial Piston Pump/Motor Principle With Floating Pistons: Design and Testing. In *BATH/ASME 2018 Symposium on Fluid Power and Motion Control, FPMC2018*. American Society of Mechanical Engineers, 2018.

[4] P. Achten and S. Eggenkamp. Barrel tipping in axial piston pumps and motors. In *Proceedings of the 14th International Fluid Power Conference*, volume 144, pages 381–391, 2017.

[5] M. Ernst and A. Vacca. Hydrostatic vs. hydrodynamic components of fluid pressure in the tribological interfaces of axial piston machines. *Tribology International*, 157:106878, 2021.

[6] R. Chacon and M. Ivantysynova. An investigation of the impact of the elastic deformation of the endcase/housing on axial piston machines cylinder block/valve plate lubricating interface. In *Proceedings of the 10th IFK International Conference on Fluid Power*, volume 1, pages 283–294, Dresden, Germany, 2016.

[7] P. A. J. Achten, T. L. v. d. Brink, and G. E. M. Vael. A robust hydrostatic thrust bearing for hydrostatic machines. In *Proc. 7th International Fluid Power Confer-*

ence (IFK), pages 100–112, Aachen, Germany, March 22–24 2010.

[8] Stephan Wegner, Stefan Gels, and Hubertus Murrenhoff. Vergleich analytischer Berechnungsmethoden des Entlastungsgrades im Kolbentrommel-Steuerspiegel-Kontakt in Axialkolbenmaschinen. *O+P Fluidtechnik*, 11-12/2017:60–69, 2017.

[9] N. D. Manring. Tipping the cylinder block of an axial-piston swash-plate type hydrostatic machine. *Transactions of the ASME*, 122:216–221, 2000.

[10] Andrea Vacca and Germano Franzoni. *Hydraulic Fluid Power: Fundamentals, Applications, and Circuit Design*. John Wileys and Sons, 2021.

[11] M. Zecchi and M. Ivantysynova. A novel approach to predict the cylinder block/valve plate interface performance in swash plate type axial piston machines. In *Bath/ASME Symposium on Fluid Power and Motion Control*, Bath, UK, 2012.

[12] T. Ransegnola. *A strongly coupled simulation model of positive displacement machines for design and optimization*. PhD Dissertation, Purdue University, West Lafayette, Indiana, 2020.

[13] Linköping University (Division of Fluid and Mechatronic Systems). Hopsan. https://liu.se/en/research/hopsan, accessed 2025-02-17.

[14] Thomas Heeger. *Design of Electro-Hydraulic Energy Converters : With Focus on Integrated Designs and Valve Plate Rotation*, volume 1971 of *Linköping Studies in Science and Technology. Licentiate Thesis*. Linköping University Electronic Press, Linköping, jun 2023.

[15] B. J. Hamrock, B. J. Schmid, and B. O. Jacobson. *Fundamentals of Fluid Film Lubrication*. CRC Press, 2004.

[16] N. Patir and H. S. Cheng. An average flow model for determining effects of three-dimensional roughness on partial hydrodynamic lubrication. *J. Lubr. Technol.*, 100(1):12–17, 1978.

[17] N. Patir and H. S. Cheng. Application of average flow model to lubrication between rough sliding surfaces. *J. Lubr. Technol.*, 101(2):220–229, 1979.

[18] C. Wu and L. Zheng. An average reynolds equation for partial film lubrication with a contact factor. *J. Tribol.*, 111(1):188–191, 1989.

[19] A. Schenk and M. Ivantysynova. An investigation of the impact of elastohydrodynamic deformation on power loss in the slipper swashplate interface. In *Proc. 8th JFPS International Symposium on Fluid Power*, Okinawa, Japan, 2011.

[20] Si C. Lee and Ning Ren. Behavior of elastic-plastic rough surface contacts as affected by surface topography, load, and material hardness. *Tribology Transactions*, 39(1):67–74, 1996.

Simulation Study on Harmonic Torque Injection for Suppressing Flow Pulsations in Electrified Axial Piston Pumps

Thomas Heeger and Liselott Ericson

Dep. of Management and Engineering, Fluid and Mechatronic Systems,
Linköping University, 581 83 Linköping, Sweden
E-mail: thomas.heeger@liu.se; liselott.ericsson@liu.se

Abstract

Noise is one of the major challenges for the electrification of hydraulic systems for mobile working machinery. Without the masking sounds of a combustion engine, the distinct noise of the hydraulic pump becomes more audible and thus needs to be reduced. Compressible pump flow pulsations caused by imperfect commutation are one of the main noise sources.
Electric motor control offers the possibility to inject torque pulses at the pump's first harmonic frequency in order to reduce common noise. This paper uses lumped parameter simulation to drive a hydraulic pump with different shapes of drive torques, and in different operating conditions. The potential of injecting harmonics into the drive torque for the reduction of flow pulsations is explored. At the right phase and amplitudes, these pulses can basically eliminate flow pulsations at their frequency. Scaling laws for comparisons of different piston numbers are summarised, and the behaviour of a 9 and a 10 piston pump are compared, demonstrating that lower torque amplitudes are required when using smaller piston numbers. Required torque amplitudes are quantified, showing that torque amplitudes in the region multiples of the pump's ideal drive torque are required at medium or high speeds. Furthermore, these torque amplitudes significantly increase losses and increase the requirements for the inverter, making harmonic injection impractical at medium and high speeds. At low speeds, the first pump harmonic can be eliminated, but this is of limited benefit due to the low sensitivity of human hearing at these frequencies.

Keywords: axial piston pump, electric motor, flow pulsations, fluid power, harmonic injection, noise

1 Introduction

Conventional mobile fluid power systems such as excavators and wheel loaders often exhibit low efficiency, with estimations of around 21% in 2012 [1] or around 30% in 2017 [2]. Due to concern about greenhouse gas emissions and the rising energy cost, different approaches to increase energy efficiency are investigated, e.g., using energy recovery and avoiding throttling losses [3]. One mayor trend is electrification, leading to the combination of electric motors with hydraulic pumps [4].

Whilst a lot of research focuses on efficiency, noise is also of large relevance for health, customer acceptance, and legislation [5]. Due to the disappearance of the masking noise of the combustion engine, the noise of the hydraulic pump becomes more audible. The interaction between hydraulic pump and electric motor for noise generation is however not very well researched. A previous paper of the authors showed that the combination of pole numbers and pistons affects the common noise [6]. This highlights the need to consider both electric motor and hydraulic pump for noise emission, and to design them for optimisation of their combined noise.

2 Literature Review

There are several contributors to noise, both in hydraulic and electric machines. For details on the hydraulic noise contributors, refer to the books by Ivantysyin and Ivantysynova [7], and Manring [8] on hydraulic pumps. Amongst others, Heeger et al. [9] described various commutation features used to reduce hydraulic pump noise. However, none of the methods fully eliminates force/moment and flow ripples. It is still not fully understood which ripple is the main contributor to noise, and it can depend on the system installation. However, typically, the pump flow ripple is believed to be a main contributor for hydraulic system noise and thus a main focus. Furthermore, Ericson et al. [10] showed that minimising the flow ripple reduces the other noise contributors as well.

Minav et al. [11] investigated using variable electric motor speed in order to eliminate / reduce an axial piston pump's flow ripple. They considered the kinematic flow ripple, but neglected the compressible flow ripple. Reference speeds to counteract the kinematic flow ripple with an opposing speed ripple were computed. They concluded that the torque capability of the electric motor limits the compensation capability. The required accelerations (and therefore torque) increase with higher speeds, so flow ripple compensation is especially challenging at higher speeds.

O'Shea [12] aimed to overcome this limitation using a low inertia gerotor pump and a feed-forward control algorithm, and the study was limited to reducing the first harmonic. At an average speed of around 400 rpm, significant reductions of the pressure ripple were measured using this strategy. However, similar to Minav et al. [11], limitations for higher frequencies were stated.

Mikota and Manhartsgruber [13] suggested passive control of the pump pulsations using a dedicated compensator, which is located between the prime mover and the hydraulic machine. Simulation showed that a coupling using a lightly damped torsional oscillator can create an antiresonance. If the antiresonance matches the first pump harmonic, pulsations at that harmonic can be reduced. A challenge is that the antiresonance frequency depends on the entire pump drivetrain dynamics, including some parameters such as damping which possess some uncertainty.

Liang et al. [14] developed a controller for a switched reluctance machine to drive a hydraulic pump. They aimed to reduce the torque ripple of the switched reluctance machine, and to avoid pressure overshoots when changing operating points.

Mehta [15] and Manring et al. [16] claimed that torque ripple might be the main cause of noise, referring to unpublished measurements. Mehta [15] further showed a correlation of torque and flow pulsations, under the assumption of incompressible fluids and simplified pressure profiles. For a reduction of torque ripple, swash plate oscillations were suggested. Furthermore, it was shown that tandem pump design significantly reduces the kinematic torque ripple if the offset angle is half a period for even piston numbers or quarter of a period for odd piston numbers.

Heeger et al. [6] summarised the harmonic noise contributors in electrically driven pumps and investigated different pistons numbers in combination with a 12 slot 10 pole PMSM. This work also showed combinations of e-motor torque and pump torque, and opened the question on whether an optimisation for a co-design for reduced torque pulsations is possible.

This paper aims to investigate the relevance of the type of torque ripple an electric machine transmits to a hydraulic pump on the other noise contributors in the system.

In comparison to the works of Minav et al. [11] and O'Shea [12], this contributes detailed pump simulations which consider compressible effects. The behaviour for different piston numbers is investigated based on the example of a 9 and a 10 piston pump, as odd and even piston numbers have different inherent properties. Scaling laws for rotation group geometry, pressure relief groove geometry and inertia are applied to provide a fair comparison between different piston numbers. Furthermore, this work considers the damping properties of a hydraulic volume at the pump outlet. Additionally, torque data from PMSM simulations is fed to the simulation model to demonstrate the feasibility of harmonic cancellation when combining it with a torque ripple at the pump's harmonic frequency. Lastly, different aspects such as electromagnetic losses due to harmonic injection, human hearing sensitivity at low frequencies and the angular position of the pump relative to the electric motor are discussed.

3 Simulation Model

This section gives an overview on the structure and assumptions of the simulation model and the used torque inputs. Furthermore, the scaling of inputs for comparable results for different numbers of pistons z are summarised.

3.1 Pump Model

This study uses a lumped parameter model in Hopsan [17] and is structured as shown in Figure 1. The model is capable of simulating different types of axial piston machines, and this study is done on the example of a machine of floating piston type [18].

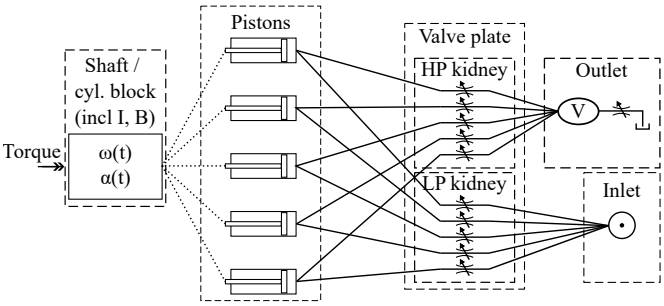


Figure 1: Schematic of Hopsan model structure.

The shaft / cylinder block component includes the torque balance as shown in Equation 1. The lever L_i of each piston's force f_i around the drive axis is computed based on the shaft angle α and the pump's geometry.

$$T = \sum_{i=1}^{n} F_i L_i + I\dot{\omega} + B\omega \tag{1}$$

Each piston's velocity v_i is computed using equation 2.

$$v_i = \omega L_i \tag{2}$$

For each piston, the force balance (Equation 3) and the continuity equation (Equation 4) are solved. The effective bulk modulus β_e is computed under the consideration of the piston's chamber pressure and pseudo cavitation.

$$F = Ap \tag{3}$$

$$\sum q = \frac{dV_{cyl}}{dt} + \frac{V_{cyl}}{\beta_e}\frac{dp}{dt} \tag{4}$$

The inlet and outlet ports are modelled as orifices with variable areas, which are computed based on the rotational angle α. The inlet port is connected to a constant pressure source. The outlet is modelled as a volume, followed by an orifice and tank. The simulation model neglects leakage.

3.2 Estimation of the Friction Coefficient

To estimate the friction coefficient B, a simple surrogate model as shown in Equation 5 is used. The parameters a, b and c are fitted to test data. B summarises the friction of all components.

$$T_{\text{loss}} = a \cdot p + b \cdot n + c \tag{5}$$

Figure 2 shows a comparison of the surrogate model to the measurement data it was fitted to. The R^2 value is 0.916, which is acceptable considering the simplicity of the model and that e.g., mixed lubrication at low speeds is not captured by it.

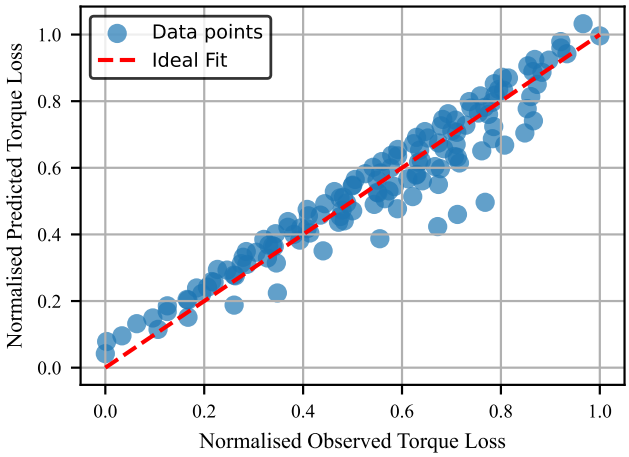


Figure 2: Normalised measured torque loss in comparison to surrogate model.

For the sake of simplicity, this work uses the same coefficients for Equation 5 independent of the number of pistons of the pump.

3.3 Scaling of Rotating Group Geometry

As the combination of piston number and pole pair number is expected to affect noise, pumps with different numbers of pistons are to be simulated. To ensure a fair comparison, the rotating group geometry is designed using the scaling laws presented by Achten [19]. Therefore, the displacement (35 cc/rev), displacement angle (8 deg) and piston pitch ratio (0.22) are kept constant, and the piston area and cylinder block pitch diameter are scaled.

3.4 Scaling of Inertia

The inertia of a cylinder follows Equation 6.

$$J_{\text{cyl}} = \frac{1}{2} m r^2 \tag{6}$$

Assuming that the piston pitch radius R_{pp} is representative for r, and that the mass of the rotating parts scales with R_{pp}^3, yields Equation 7.

$$J_{\text{pump}} \propto R_{\text{pp}}^5 \tag{7}$$

The piston pitch radius R_{pp} can be taken from the scaling of the geometry of the rotating group (see Equation 8), with k being 1 for a regular pump and 2 for a tandem pump.

$$R_{\text{pp}} \propto \frac{1}{z^{\frac{1}{3}} \left(\sin\left(\frac{\pi k}{z k}\right)\right)^{\frac{2}{3}}} \tag{8}$$

Finally, the inertia of pump J_{pump} and the inertia of the electric motor J_{EM} are added in Equation 9. The inertia of the electric motor is independent of the number of pistons. In this study, the electric motor's inertia is roughly 15 times larger than the pump's inertia and therefore dominates Equation 9.

$$J_{\text{tot}} = J_{\text{pump}} + J_{\text{EM}} \tag{9}$$

3.5 Scaling of Pressure Relief Grooves

Pressure relief groove design plays a crucial role for pump noise. Palmberg [20] derived Equations 10 and 11 for the length and final area of pressure relief grooves.

$$l_{\text{g}} = R \sqrt{\frac{2 \alpha_{\text{p}} h \, p_0}{\beta}} \tag{10}$$

$$A_{\text{g}} = \frac{(1+n) D \omega}{C_{\text{q}} z} \sqrt{\frac{\alpha_{\text{p}} \rho}{h \beta}} \tag{11}$$

The bulk modulus β, the density ρ, the orifice coefficient C_{q}, the displacement D, the shaft frequency ω, the exponent in the orifice area function n, and the system pressure p_0 are not (or negligibly) affected by a change of the number of pistons. Assuming a constant dead volume ratio, the clearance volume factor α_{p} also remains constant.

The flow gradient h scales according to Equation 12.

$$h \propto \frac{1}{z} \tag{12}$$

Assuming that the piston pitch radius R_{pp} and the pressure relief groove pitch radius R are identical results in the scaling laws for the groove's length l_{g} and final area A_{g} as shown in Equations 13 and 14.

$$l_{\text{g}} \propto R_{\text{pp}} \sqrt{\frac{1}{z}} \propto \frac{1}{z^{\frac{5}{6}} \left(\sin\left(\frac{\pi k}{z k}\right)\right)^{\frac{2}{3}}} \tag{13}$$

$$A_{\text{g}} \propto \frac{1}{z^{3/2}} \tag{14}$$

4 Results for Sine Shaped Torque Pulsations

For the torque input, the following set-ups were used:

- Constant torque input.
- Constant torque input combined with a sine shaped pulsation as shown in Equation 15.
- Look-up table based torque input, using results from Ansys Motorcad simulations for a 12 slot 10 pole permanent magnet synchronous machine. This is done in Section 5.

$$T = T_{\text{ave}} + T_{\text{amplitude}} \cdot \sin\left(z \cdot \alpha + \phi\right) \tag{15}$$

This section shows results for constant torque inputs as well as for sine shaped torque pulsations following Equation 15. Effects of varying different simulation parameters on speed, flow and pressure pulsations are investigated. Also, pumps with 9 and 10 pistons were considered as the noise emission of pumps with odd and even piston numbers differs significantly.

4.1 Effect of Phase of Torque Pulsation

Previous work by the authors [6] suggested using the combination of the electric motor's and hydraulic pump's torque ripples to reduce the combined ripple. Typically, the first harmonic dominates the flow pulsations' frequency spectrum of axial piston pumps. Therefore, a sine shaped torque pulsation is added to a constant drive torque to investigate the potential for noise reduction. Harmonic injection is used in electric motor control to reduce the motor's torque ripple [21], but in principle the motor control could also be used to artificially create a torque ripple. Different phase shifts can either amplify or reduce the ripples.

4.1.1 Definition of Phase Shift

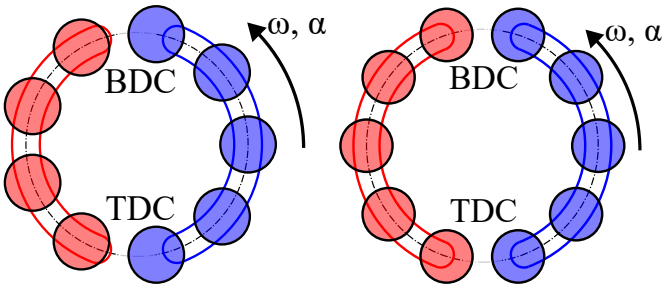
The significance of the phase of the torque pulsation is investigated by testing the baseline case (without torque pulsation), and applying a torque pulsation of 5 Nm amplitude at first fundamental piston frequency. This is done for 20 different phase shifts, differing by 5% of a period. A rotation angle of 0 degrees is defined when the first piston is at the center of its suction stroke as visualised in Figure 3. That means that the commutation to connect to the high-pressure kidney starts after 25% (or 90°) of a period for the 9 piston pump, and after 50% (or 180°) of a period for the 10 piston pump.

4.1.2 Effect on Speed and Flow Pulsations

The phase of the torque pulsation changes the phase of the speed pulsations as well as the amplitude at the first piston frequency, see Figures 4 and 5.

The speed ripple alters the flow ripple significantly.

For the 9 piston pump, Figure 6 shows that flow pulsations are reduced when the torque pulsation is 15% of a period ahead. This results in the peak of the speed ripple (see Figure 4) coinciding with the valley of the flow ripple. For a 10 Nm torque ripple, the first fundamental is reduced by 79% in comparison



(a) α=0 for 9 piston pump. *(b) α=0 for 10 piston pump.*

Figure 3: Definition of 0 degree position.

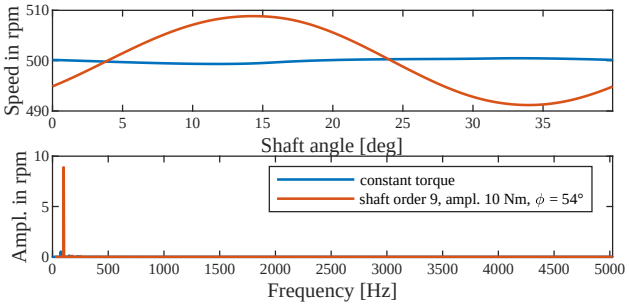


Figure 4: Speed pulsations for constant torque and a 10 Nm amplitude torque ripple with 15% period phase shift for a 9 piston pump operated at 500 rpm and 50 bar.

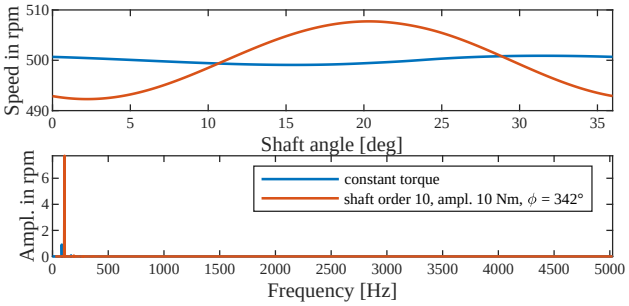


Figure 5: Speed pulsations for constant torque and a 10 Nm amplitude torque ripple with 95% period phase shift for a 10 piston pump, operating at 500 rpm and 50 bar.

to the baseline. The higher harmonics are not significantly affected, which is reasonable considering that the input torque signal did not contain their frequencies.

Figure 7 shows a similar situation for the 10 piston pump. For a phase shift of 95%, the first fundamental is reduced by 42%.

4.1.3 Methodology for Finding Optimal Phase Shift

For a given pump and operating condition, the optimal phase shift is found by simply comparing the first harmonic of the flow ripple for different phase shifts. Exemplary results can be seen in Figure 8.

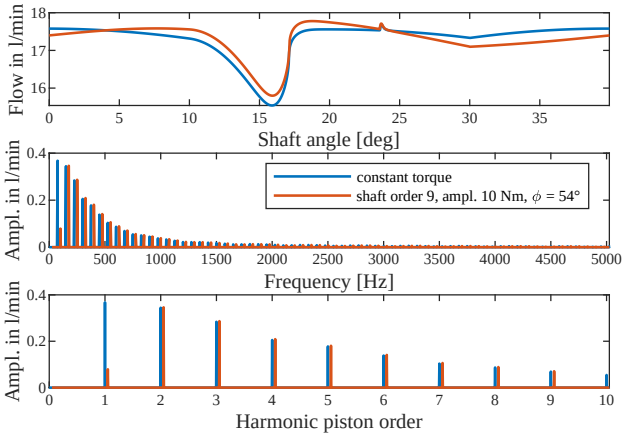


Figure 6: Flow pulsations for constant torque and a 10 Nm amplitude torque ripple with 15% phase shift for a 9 piston pump, operating at 500 rpm and 500 bar.

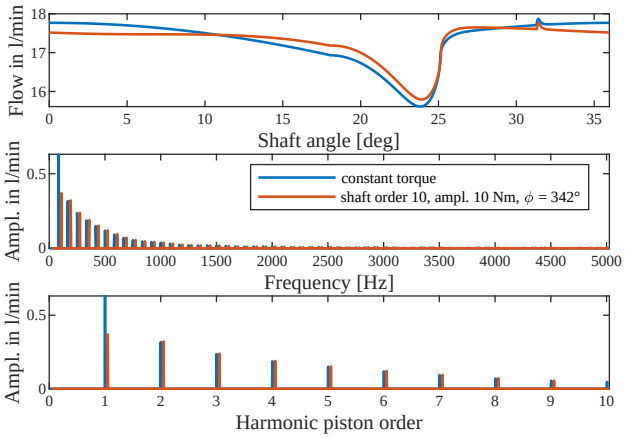


Figure 7: Flow pulsations for constant torque and a 10 Nm amplitude torque ripple with 95% phase shift for a 10 piston pump, operating at 500 rpm and 50 bar.

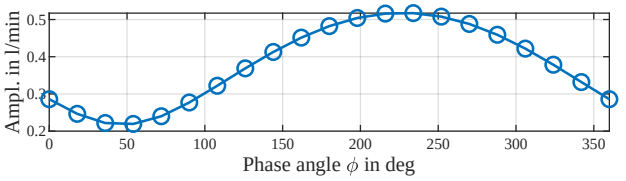


Figure 8: Phase shift dependency of first fundamental harmonic on the phase shift. The example is for a 9 piston pump, 50 bar, 500 rpm, and 5 Nm amplitude.

4.1.4 Effect of Operating Conditions

The previous results showed operation at 50 bar and 500 rpm. However, it is also of interest how the ideal phase changes with changing operating conditions. At this stage, the torque amplitude is fixed to 50 Nm. A parameter sweep over 20 phase shifts is used to identify the ideal phase shift for each operating point. Tables 1 and 2 show that the operating point has some effect on the ideal phase angle, as the commutation is affected both by the pressure and speed level.

Table 1: Ideal Phase Shift for 9 Piston Pump in Percent of a Period (in increments of 5), for a torque amplitude of 50 Nm.

	50 bar	200 bar	350 bar
500 rpm	15	10	10
2000 rpm	10	5	5
4000 rpm	10	0	0

Table 2: Ideal Phase Shift for 10 Piston Pump in Percent of a Period (in increments of 5), for a torque amplitude of 50 Nm.

	50 bar	200 bar	350 bar
500 rpm	95	90	85
2000 rpm	90	85	80
4000 rpm	90	80	75

4.2 Effect of Amplitude of Torque Pulsation

The torque amplitude controls the amplitude of the speed pulsation. A too small amplitude has no significant effect, and a too high amplitude overcompensates for the pump's inherent flow pulsations.

Previously, arbitrary torque amplitudes were used to find good phase shifts. In this chapter, torque amplitudes are optimised for fixed phase shifts.

At 500 rpm and 50 bar, a torque amplitude of 12 Nm with a phase shift of 13% is applied to the 9 piston machine. Figure 9 shows that the first fundamental is almost completely eliminated with a reduction of 99.8%.

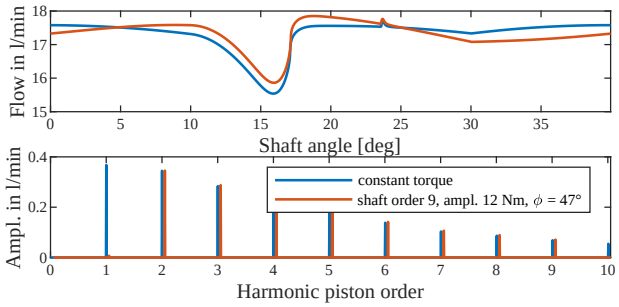


Figure 9: Flow pulsations for 9 piston machine with optimised torque amplitude at 500 rpm and 50 bar.

The 10 piston machine behaves similarly. A torque amplitude of 24 Nm is applied at a phase shift of 93% to the 10 piston machine. As shown in Figure 10, this yields an almost complete elimination of the first fundamental with a reduction of 97%.

A complete elimination of the first harmonic is feasible with a finer tuning of the phase angle and amplitude of the torque pulsation.

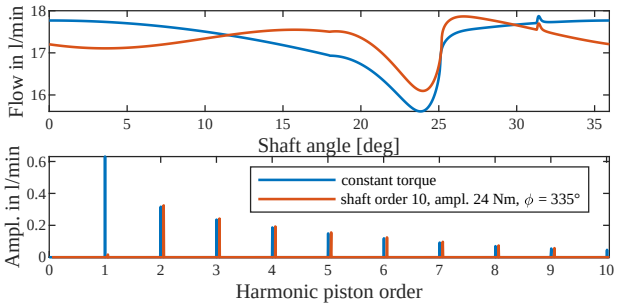


Figure 10: Flow pulsations for 10 piston machine with optimised torque amplitude at 500 rpm and 50 bar.

4.2.1 Effect of Operating Conditions

Tables 3 and 5 show the ideal torque amplitudes to eliminate the first fundamental under consideration of the pump's and the electric motor's inertia. As a comparison, Tables 4 and 6 show the torque amplitudes required when only considering the pump's inertia, i.e., neglecting the electric motor's inertia. A comparison of the tables shows that the electric motor's inertia increases the required torque amplitudes by one order of magnitude. Note that the phase angles in Tables 1 and 2 were used for all tables.

Table 3: Ideal Torque Amplitude for 9 Piston Pump in Nm (considering inertia of pump and electric machine).

	50 bar	200 bar	350 bar
500 rpm	12	28	43
2000 rpm	200	430	650
4000 rpm	820	1670	2540

Table 4: Ideal Torque Amplitude for 9 Piston Pump in Nm when only considering pump inertia (neglecting the inertia of the electric machine).

	50 bar	200 bar	350 bar
500 rpm	1.0	2.0	3.5
2000 rpm	15	32	49
4000 rpm	61	124	189

Table 5: Ideal Torque Amplitude for 10 Piston Pump in Nm (considering inertia of pump and electric machine).

	50 bar	200 bar	350 bar
500 rpm	24	39	52
2000 rpm	370	550	750
4000 rpm	1430	1970	2700

Note that in line with Minav et al. [11] and O'Shea [12], for higher speeds, higher torque amplitude are required for more significant reductions, and the practical limitations due to the high required amplitudes are confirmed. In addition, the torque amplitude increases with increasing pressures.

Table 6: Ideal Torque Amplitude for 10 Piston Pump in Nm when only considering pump inertia (neglecting the inertia of the electric machine).

	50 bar	200 bar	350 bar
500 rpm	2.0	3.0	4.5
2000 rpm	31	46	61
4000 rpm	125	169	228

The 10 piston machine requires higher torque amplitudes for all operating points. This is caused by a higher frequency requiring faster accelerations as well as higher inertia. The effect is more profound at low and medium pressures than at high pressures.

4.2.2 Importance of First Harmonic

Pressure relief grooves are most effective in reducing harmonics of second and higher order. However, the first harmonic typically is only slightly reduced [22], and remains as the major noise contributor. Therefore, a torque pulsation capable of eliminating the first harmonic is a significant contribution to noise reduction.

4.3 Effect of Frequency of Torque Pulsation

Up to this point, torque pulsations of the first fundamental were considered. However, the first fundamental of the inherent torque pulsation of a 12 slot 10 pole permanent magnet synchronous machine is at 30 times the shaft frequency [6]. This chapter investigates the behaviour of the pumps in combination with a torque pulsation according to Equation 16.

$$T = T_{\text{ave}} + T_{\text{amplitude}} \cdot \sin\left(30 \cdot \alpha + \phi\right) \tag{16}$$

In case of the 9 piston pump, this does not coincide with any pump harmonics. Therefore, as visualised in Figure 11, adding the torque pulsations simply leads to additional frequency content of the flow pulsation, independent of the phase of the torque pulsations.

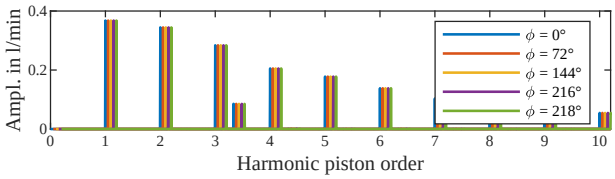


Figure 11: Harmonics at 50 bar and 500 rpm for the 9 piston machine and a torque pulsation of 30 times shaft frequency for different phase angles.

In case of the 10 piston pump, this coincides with the third harmonic. Therefore, as Figure 12 visualises, depending on the phase angle, the third pump harmonic is either reinforced or weakened. Table 7 shows the ideal phase shift for different

operating conditions. As the ideal phase shift differs between operating conditions, there is not one single ideal positioning of the electric motor to the hydraulic pump.

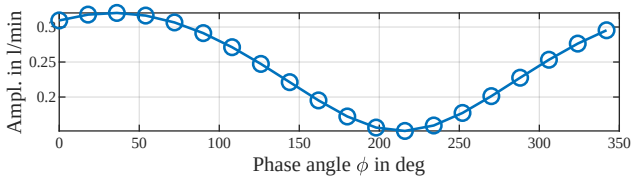


Figure 12: Phase angle dependency of third harmonic of 10 piston machine for a torque pulsation running at 30 times shaft frequency. Example is done for 50 bar, 500 rpm, and 10 Nm amplitude.

Table 7: Ideal Phase Shift for 10 Piston Pump to Reduce the Third Harmonic (in Percent of a Period and in increments of 5)

	50 bar	200 bar	350 bar
500 rpm	60	55	50
2000 rpm	35	25	20
4000 rpm	30	15	0

4.4 Effect of Volume

Up to this point, a volume of 100 ml at the pump outlet was considered.

However, the size of the volume affects the damping, and therefore the flow pulsations are reduced by the volume. This section investigates the influence of the volume size by comparing the system behaviour for a volume of 10 ml, 100 ml, and 1000 ml.

A parameter sweep showed that the ideal phase shift and torque amplitude are not significantly affected by the volume at the pump outlet. Figure 13 shows the flow pulsations before the volume for the same operating conditions and different volumes. Figure 14 shows the flow pulsations for the same operating condition after the volume. The smallest volume (10 ml) has very little damping, but the larger volumes (100 ml, 1000 ml) show significantly reduced flow pulsations after the volume.

Figure 15 shows the ratio of the harmonics before and after the volume, confirming that higher order harmonics decrease more than low-order harmonics.

5 Results for Torque Pulsations of a PMSM

In this section, the interaction of the torque pulsations of a 12 slot 10 pole permanent magnet synchronous machine with the pump setups are investigated. For this purpose, a lookup table with the rotational angle and the instantaneous torque is created using Ansys Motorcad and fed to the Hopsan pump model.

Figure 16 shows the inherent torque curves of the PMSM and

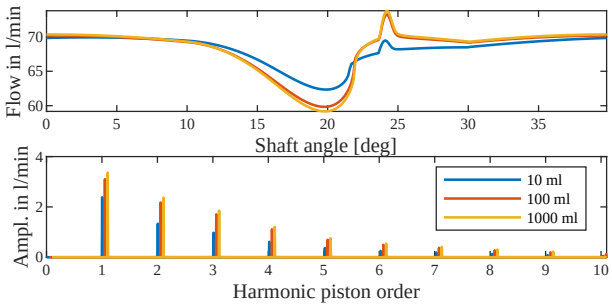


Figure 13: Flow before the volume at 200 bar and 2000 rpm for the 9 piston machine with constant torque for different volumes.

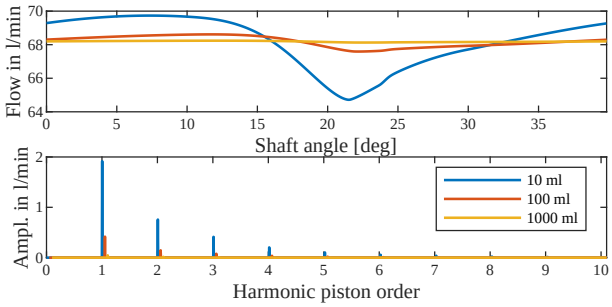


Figure 14: Flow after the volume at 200 bar and 2000 rpm for the 9 piston machine with constant torque for different volumes.

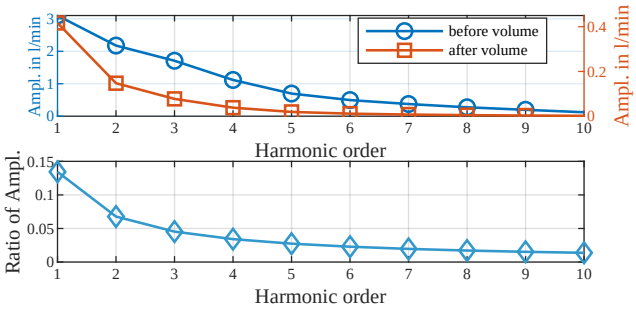


Figure 15: Increased damping of higher harmonics based on 9 piston pump operating at 2000 rpm and 200 bar and 100 ml volume.

the pumps. The average PMSM torque is larger than for the pumps due to friction.

Figure 17 shows the flow pulsations after a 100 ml volume for a constant torque source in comparison to the PMSM with different relative positions to the pump. The PMSM acts very similar to the constant torque source.

Figure 18 shows the flow for different relative positions of the PMSM to a 10 piston pump whose first harmonic is suppressed by a first order torque pulsation. The torque suppression works for either orientation, and only negligible differences for the third harmonic are observed.

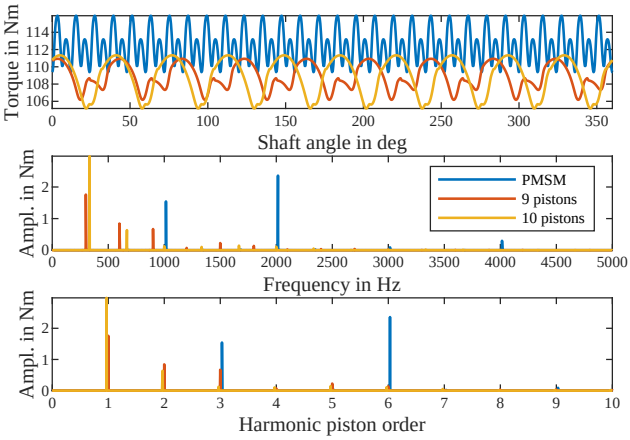


Figure 16: Inherent torque curves 12 slot 10 pole PMSM and 9 resp. 10 piston pump for operation at 200 bar and 2000 rpm. Note that for the PMSM, the 10 piston machine was used to define the harmonic piston orders.

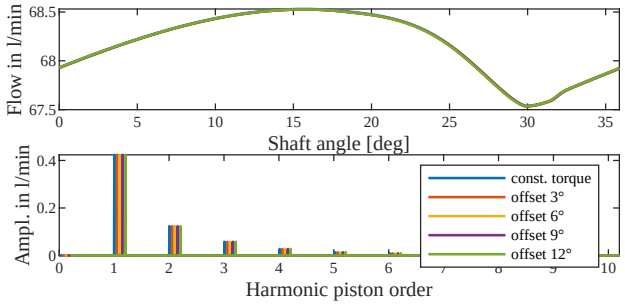


Figure 17: Flow after a 100 ml volume for a 10 piston pump driven by a constant torque or a 12 slot 10 pole PMSM with different positional offsets. The example is for operation at 200 bar and 2000 rpm.

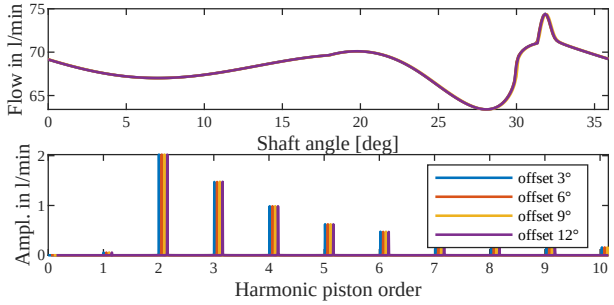


Figure 18: Flow pulsations for different positions of the PMSM relative to a 10 piston pump with eliminated first harmonic for operation at 2000 rpm and 200 bar.

6 Discussion

This section discusses the implications of the results on the choice of number of pistons, the feasibility and usefulness to create the required torque ripples, electromagnetic losses, and refinements to the presented method.

6.1 Effect of the Number of Pistons

In principle, the reduction of the first harmonic works the same independent of the piston number. However, comparing Tables 3 and 5 shows that the 10 piston pump requires higher torque amplitudes than the 9 piston pump. This is caused both by the decreased time a period requiring faster accelerations and by an increased inertia.

Using 10 pistons in combination with a 12 slot 10 pole PMSM offers the possibility to use the electric machine's inherent torque ripple to reduce the pump's third harmonic. However, Table 7 shows that the ideal phase has some spread, and Figure 18 showed that the effect of the angular orientation of the PMSM on the third pump harmonic is negligible.

6.2 Torque Requirement

It is in principle possible to use torque/speed pulsations to eliminate harmonics of the flow pulsations, but medium or high speeds require very high torque amplitudes (as already stated by Minav et al. [11] and O'Shea [12]). In comparison to Minav et al.'s work, the torque amplitudes are significantly higher due to increased inertia, and the consideration of compressible flow. The electric motor's inertia is one order of magnitude higher than the pump's inertia, and is therefore the main driver of the torque requirements. A low number of pistons leads to higher inherent pump pulsations, but they are easier to cancel using a torque pulse as they provide increased acceleration time and decreased inertia.

The considered displacement (35 cc/rev) in combination with the considered maximum pressure (350 bar) yields an ideal drive torque of 195 Nm. Tables 3 and 5 show that for low speeds, only small torque amplitudes are needed. At medium and high speeds, torque amplitudes of some 100 Nm or higher are required. This is expected to exceed the feasible torque ripple for the electric motor and a standard inverter.

6.3 Human Hearing

The human hearing area spans frequencies from 20 Hz to 20 kHz. However, the sensitivity of human hearing is frequency-dependent and e.g. sounds of with los frequencies need higher sound pressure levels to be perceived than sounds at medium frequencies [23].

When operating a 9-piston pump at 500 rpm, the first fundamental frequency is 75 Hz. As the human ear is not very sensitive at these frequency, removing the first harmonic from an audio signal has no audible effect.

6.4 Losses

Losses in PMSM are typically dominated by copper losses and iron losses. Harmonic injection increases the losses of the electric motor [21], which reduces the efficiency and creates additional heat in the electric motor. Note that even the inverter losses increase with increasing switching frequencies, but this will not be quantified in this paper.

6.4.1 Copper Losses

Copper losses follow Ohm's law, see Equation 17.

$$P_{\text{loss,Cu}} = R_{\text{el}} I_{\text{rms}}^2 \tag{17}$$

For surface-mounted PMSM outside the field-weakening region, the current is approximately proportional to the torque. Therefore, copper losses scale approximately with the square of the torque.

Equation 18 describes the current as a function of time, yielding Equation 19 giving an expression of the square of the rms current I_{rms}^2, which is proportional to the copper losses $P_{loss,Cu}$. Equation 19 enables an estimation of the increase of the copper losses due to harmonic injection. The exemplary results in Tables 8 and 9 show a slight increase of copper losses due to harmonic injection at low speeds, but very high increases at medium speeds.

$$I(t) = I_{\text{ave}} + I_{\text{amp}} \cdot \sin(k\omega t) \tag{18}$$

$$I_{\text{rms}}^2 = I_{\text{ave}}^2 + \frac{I_{\text{amp}}^2}{2} \tag{19}$$

Table 8: Increase of copper losses caused by harmonic injection of amplitudes as given in Table 3 for a 9 piston machine.

	50 bar	200 bar	350 bar
500 rpm	9.3%	3.2%	2.4%
2000 rpm	2578%	745%	556%

Table 9: Increase of copper losses caused by harmonic injection of amplitudes as given in Table 5 for a 10 piston machine.

	50 bar	200 bar	350 bar
500 rpm	37.1%	6.1%	3.6%
2000 rpm	8824%	1219%	740%

6.4.2 Iron Losses

Iron losses consist of the hysteresis losses, eddy losses and excess losses which cannot be explained by the two aforementioned loss mechanisms, and are generally estimated according to Equation 20 [24].

$$P_{\text{loss,Fe}} = k_{\text{h}} f B_{\text{stator}}^{\alpha_{\text{h}}} + k_{\text{cl}} f^2 B_{\text{stator}}^2 + k_{\text{exc}} f^{1.5} B_{\text{stator}}^{1.5} \tag{20}$$

The equation shows a sensitivity both to frequency and magnitude of flux variations. Sohail Ahmed [25] made measurements with different harmonics at different frequencies. For a fundamental frequency of 50 Hz, introducing a harmonic with an amplitude of 20% of the base value yielded a maximum loss increase of 85% for a 7th order harmonic and 168% for an 11th order harmonic. For the 9 piston machine at 500 rpm, the ratio is in a similar region, so iron losses in a similar order of magnitude are to be expected. For the 10 piston machine, the torque ratio is higher, thus higher increases of iron losses are to be expected.

6.5 Positioning of PMSM relative to the Pump

For the considered example, Figure 18 implies that the positioning of the PMSM relative to the pump is only of minor importance, as it only affects the third pump harmonic in this case, and the magnitude changes are very small. Furthermore, Table 7 shows that an ideal position for all operating points cannot be found anyway.

For other electric motor configurations, these conclusions may however not be valid as they possess other frequencies and amplitudes.

6.6 Improvements to the Method

The parameter sweeps done in Sections 4.1 and 4.2 show the potential of drastic reductions of the first fundamental harmonic. A co-optimisation of phase angle and torque amplitude could however yield even better results.

7 Conclusion

This paper investigated the idea to intentionally introduce speed pulsations with the electric motor to cancel the first harmonic of the hydraulic pump's flow pulsation in order to reduce common noise.

Due to the large inertia of the electric motor, the angular orientation between the pump and the electric motor does not significantly affect the pump's flow pulsations.

Pressure relief grooves are a typical commutation feature to reduce flow pulsations from hydraulic pumps. However, they only have a small effect on the first order harmonics, which often dominate a pump's frequency spectrum. Motor control can be used to induce torque pulsations at the pump's first harmonic, and using the correct amplitude and phase, the pump's first harmonic can be eliminated. Unfortunately, due to the electric motor's large inertia, very large torque amplitudes in the order of multiples of nominal torque are required to cancel the first harmonic at medium and high speeds. Also, higher piston numbers require larger amplitudes, as higher frequencies require higher accelerations, and as inertia typically increased with higher piston numbers. These large torque amplitudes cause unreasonable levels of currents. Furthermore, electromagnetic losses drastically increase due to the required amplitudes at medium and high speeds. At low speeds, the torque amplitudes are in a reasonable region, and losses are slightly increased. Therefore, the cancellation of the first harmonic is in principle feasible. However, at these speeds, the first harmonic is at frequencies where the human ear is not very sensitive, yielding almost no audible benefit of removing this harmonic.

Therefore, introducing speed pulsations to the electric motor is no practical method of eliminating the first harmonic of hydraulic flow pulsations.

Acknowledgement

This research was funded by the Swedish Electromobility Centre with grant number 13070 and the Swedish En-

ergy Agency (Energimyndigheten) with grant number P2023-00594.

Nomenclature

Designation	Denotation	Unit
α	Shaft angle	rad
α_h	Hysteresis loss exponent	-
α_p	Clearance volume factor	-
β_e, β	(Effective) bulk modulus	N/m^2
ϕ	Phase offset	rad
ρ	Density	kg/m^3
ω	Angular speed	rad/s
A	Area	m^2
A_g	Final groove area	m^2
B	Friction coefficient	Nm/(rad/s)
B_{stator}	Magnetic flux density in the stator	T
C_q	Orifice flow coefficient	-
D	Displacement	m^3/rad
f	Frequency	Hz
F	Force	N
h	Kinematic flow gradient	m^3/s^2
I	Current	A
J	Rotational inertia	kg m^2
k	1 for regular pump, 2 for tandem pump	-
k_{cl}	Classical eddy loss coefficient	Ws/(kg T$^\alpha$)
k_{exc}	Excess loss coefficient	Ws$^{1.5}$/(kg T$^{1.5}$)
k_h	Hysteresis loss coefficient	Ws^2/(kg T^2)
L	Lever of piston force around drive axis	m
l_g	Groove length	m
m	Mass	kg
n	Exponent in groove area function	-
p	Pressure	N/m^2
p_0	System pressure	N/m^2
P	Power	W
r	Radius	m
R	Radius to relief groove	m
R_{el}	Resistance	Ω
R_{pp}	Piston pitch radius	m
T	Torque	Nm
v	Velocity	m/s
z	Number of pistons	-

References

[1] Lonnie J Love, Eric Lanke, and Pete Alles. *Estimating the Impact (Energy, Emissions and Economics) of the US Fluid Power Industry*. Number December. Oak Ridge National Laboratory, Oak Ridge, USA, 2012.

[2] Lauren A Lynch and Bradley T Zigler. Estimating Energy Consumption of Mobile Fluid Power in the United States. Technical Report November, National Renewable Energy Laboratory, 2017.

[3] Damiano Padovani, Pavlos Dimitriou, and Tatiana Minav. Challenges and solutions for designing Energy-Efficient and Low-Pollutant Machines in Off-Road hydraulics. *Energy Conversion and Management: X*, 21(December 2023):100526, 2024.

[4] Thomas Heeger. *Design of Electro-Hydraulic Energy Converters : With Focus on Integrated Designs and Valve Plate Rotation*, volume 1971 of *Linköping Studies in Science and Technology. Licentiate Thesis*. Linköping University Electronic Press, Linköping, June 2023.

[5] Liselott Ericson and Thomas Heeger. The Electrification of Material Moving Machines : An Overview of Opportunities and Challenges Regarding Noise. In *Proceedings of the Interdisciplinary Conference on Innovation, Design, Entrepreneurship, and Sustainable Systems (IDEAS), 24-27 November 2024, Recife, Brazil*. Springer, 2025. ISBN 9783031961724. Paper presented at conference, proceedings to be published.

[6] Thomas Heeger, Martin West, and Liselott Ericson. Harmonic Characterisation of Electrically Driven Pumps. In Liselott Ericson and Petter Krus, editors, *Advancements in Fluid Power Technology: Sustainability, Electrification, and Digitalization. Proceedings of the Global Fluid Power Society PhD Symposium 2024, 17-20 June, 2024, Hudiksvall, Sweden*. Springer, 2025. ISBN 9783031845048. Paper presented at conference, proceedings to be published.

[7] Jaroslav Ivantysyn and Monika Ivantysynova. *Hydrostatic pumps and motors: principles, design, performance, modelling, analysis, control, and testing*. Akademia Books International, New Delhi, 2001. ISBN 8185522162.

[8] Noah D. Manring. *Fluid Power Pumps and Motors: Analysis, Design and Control*. McGraw Hill Book CO, 2013. ISBN 0071812202.

[9] Thomas Heeger, Samuel Kärnell, and Liselott Ericson. Challenges for multi-quadrant hydraulic piston machines. *Energy Conversion and Management: X*, 22(March):100578, 2024.

[10] Liselott Ericson, Johan Ölvander, and Jan-Ove Palmberg. On Optimal Design of Hydrostatic Machines. In *Proceedings of the 6th International Fluid Power Conference (6. IFK), 1-2 April, 2008, Dresden, Germany*, pages 273–286, 2008.

[11] T.A. Minav, L.I.E. Laurila, and J.J. Pyrhönen. Axial Piston Pump Flow Ripple Compensation by Adjusting the Pump Speed with an Electric Drive. In *The Twelfth Scandinavian International Conference on Fluid Power, May 18-20, 2011, Tampere, Finland*, 2011.

[12] Colin O'Shea. Hydraulic flow ripple cancellation using the primary flow source. In *Proceedings of the BATH/ASME 2016 Symposium on Fluid Power and Motion Control (FPMC2016), September 7-9, 2016, Bath, UK*, pages FPMC2016–1783, 2016.

[13] Gudrun Mikota and Bernhard Manhartsgruber. a Novel Pulsation Compensator for Displacement Machines. In *14th International Fluid Power Conference, 19-21 March, 2024, Dresden, Germany*, pages 389–401, 2024.

[14] Jianing Liang, Jin-woo Ahn, and Dong-hee Lee. High Performance Hydraulic Pump System Using Switched Reluctance Drive. In *Proceeding of International Conference on Electrical Machines and Systems 2007, Oct. 8∼11, Seoul, Korea*, pages 1470–1475, 2007.

[15] Viral Mehta. *Torque Ripple Attenuation for an Axial Piston Swash Plate Type*. Phd thesis, University of Missouri - Columbia, 2006.

[16] Noah D. Manring, Viral S. Mehta, Frank J. Raab, and Kevin J. Graf. The Shaft Torque of a Tandem Axial-Piston Pump. *Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME*, 129(3):367–371, 2007.

[17] Linköping University (Division of Fluid and Mechatronic Systems). Hopsan. https://liu.se/en/research/hopsan, accessed 2025-02-17.

[18] Liselott Ericson and Jonas Forssell. A Novel Axial Piston Pump/Motor Principle With Floating Pistons: Design and Testing. In *BATH/ASME 2018 Symposium on Fluid Power and Motion Control (FPMC2018), September 12-14, 2018, Bath, United Kingdom*, pages FPMC2018–8937. American Society of Mechanical Engineers, 2018.

[19] Peter A J Achten. Power Density of the Floating Cup Axial Piston Principle. In *Proceedings of the ASME 2004 International Mechanical Engineering Congress and Expo (IMECE2004), November 13-19, 2004, Anaheim, California, USA*, pages IMECE2004–59006. ASME, 2004.

[20] J O Palmberg. Modelling of Flow Ripple from Fluid Power Piston Pumps. In *Fluid power components systems: 2nd Bath International Fluid Power Workshop, Taunton, Somerset, United Kingdom*, pages 207–227, 1990.

[21] Jianzhen Qu, Juri Jatskevich, Chengning Zhang, and Shuo Zhang. Torque ripple reduction method for permanent magnet synchronous machine drives with novel harmonic current control. *IEEE Transactions on Energy Conversion*, 36(3):2502–2513, 2021.

[22] Andreas Johansson. *Design Principles for Noise Reduction in Hydraulic Piston Pumps*. Phd thesis, Linköping University, 2005.

[23] Eberhard Zwicker and Hugo Fastl. *Psychoacoustics: Facts and Models*. Springer-Verlag Berlin Heidelberg GmbH, second edition, 1999. ISBN 9783540650638.

[24] Zi-Qiang Zhu, Shaoshen Xue, Wenqiang Chu, Jianghua Feng, Shuying Guo, Zhichu Chen, and Jun Peng. Evaluation of Iron Loss Models in Electrical Machines. *IEEE Transactions on Industry Applications*, 55(2):1461–1472, 2019.

[25] Muhammad Sohail Ahmed. *Effect of harmonics on iron losses*. Master thesis, Chalmers University of Technology, 2007.

Measurement of a rotary mechanical pulsation compensator under hydraulic resonance

David Holzer and Gudrun Mikota

Institute of Machine Design and Fluid Power, Johannes Kepler University, Linz, Austria
E-mail: david.holzer@jku.at, gudrun.mikota@jku.at

Abstract

If the pulsation frequency of a hydraulic pump or motor coincides with a natural frequency of the hydraulic system, there is a particular need for pulsation compensation. Under these conditions, a rotary mechanical pulsation compensator is investigated. The compensator consists in a torsional oscillator which is formed by the drivetrain of a radial piston pump and requires a special compensator coupling. At the pump outlet, two different hydraulic layouts are considered, including a closed pipeline and a cylinder chamber connected via a hose. Prior to experiments, simulations are conducted to determine system parameters at which a natural frequency of the hydraulic system matches the compensation frequency. In the experimental setup, the compensator coupling is installed on the pump shaft and driven by a speed-controlled induction motor. The measurements conducted with a closed pipeline show a pulsation reduction by a factor of more than three in the resonance case during a run-up test. Similar results are obtained in a run-up test with a cylinder in its limit stop position. An extension test with constant motor speed shows significant reduction of pressure pulsation. In the present configuration, the compensation effect is narrowband and the design of the coupling causes a dependency between compensation frequency and drivetrain torque.

Keywords: pulsation, radial piston pump, drivetrain, compensator, coupling, resonance

1 Introduction

The noise generated by hydraulic systems, both in industrial applications and in mobile machinery, represents a significant disadvantage. Noise is caused by pressure pulsations in the fluid, which lead to mechanical vibrations. As explained in [1], pressure pulsations arise both from the operating principles of positive displacement machines and from external vibrations acting on hydraulic components. If the pulsation frequency of a pump or motor matches a natural frequency of the system, resonance will occur, which amplifies the pressure pulsation and increases the noise. Under such resonance conditions, the need for countermeasures is most apparent.

To reduce hydraulic pressure pulsation, there are both active and passive concepts. Passive attenuators are significantly more widespread than active solutions and operate based on resistive, reactive or composite working principles. Resistive damping attenuators work over an unlimited frequency range but cause high energy losses in the system. In contrast, reactive compensators have a limited frequency range [2]. These compensators are designed as oscillating systems, with the compensator system's natural frequency tuned to the disturbing pulsation frequency. At this frequency, the compensator oscillates in anti-phase with the pump or motor pulsation. A well-known example of this type of compensator is the Helmholtz resonator. It is based on a hydraulic oscillator consisting of an inductance and a capacity. The passive compensator described in [3] replaces the hydraulic resonator with a mechanical spring-mass oscillator placed in the hydraulic circuit.

The present paper investigates a special coupling installed on the pump shaft, which turns the pump drivetrain into a rotary mechanical pulsation compensator [4] modelled by a two-mass torsional oscillator. This oscillator consists of the mass moment of inertia on the drive side, the mass moment of inertia on the pump side and the stiffness of the designed compensator coupling. Starting from the simulations and design from [5, 6], the compensator coupling was designed for an existing experimental setup. It was fabricated and installed in the test rig. Initial measurements were carried out, which already showed a compensation effect [7].

In order to study the properties of the compensator coupling in the case of hydraulic resonance, measurements were conducted on two different hydraulic layouts and are presented in this paper.

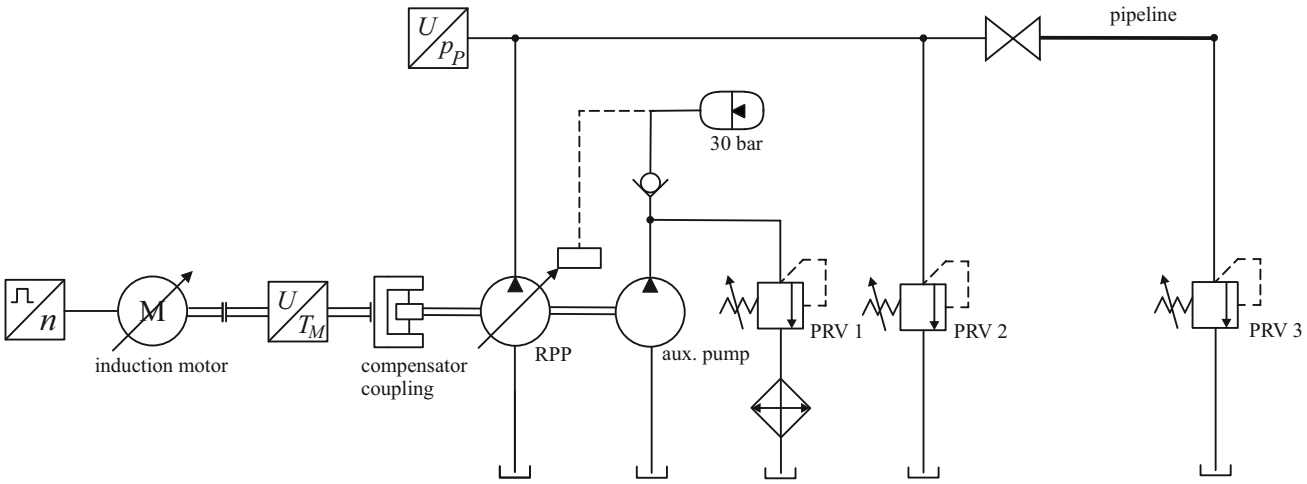


Figure 1: Overall hydraulic layout including the drivetrain, with compensator coupling and a closed pipeline.

2 Experimental test setup

The experimental setup is driven by a speed-controlled induction motor with a rated speed of 1420 rpm and a rated torque of 71.3 Nm. To measure the drivetrain torque, a measuring shaft is installed after the motor. The connections between the motor, the measuring shaft and the compensator coupling are realised by curved tooth couplings. The compensator coupling is installed directly on the pump shaft via a key connection. The structure of the drivetrain is shown in fig. 1 as simplified hydraulic circuit diagram and in fig. 2 as photo.

The radial piston pump used in this test rig is shown in fig. 3. It has seven pistons and variable displacement with a maximum of 32 cm^3 / rev. To adjust the displacement via the stroke ring, an auxiliary gear pump is required as pressure supply. The stroke ring, which is controlled by a servo valve, is set to maximum displacement for all experiments. It is necessary to keep the oil temperature within a narrow range for the tests, as preliminary tests have shown that the dependence between oil temperature and pressure pulsation cannot be neglected. Pressure pulsation increases with higher oil temperatures. Most of the heat is generated by the pressure relief valve in the main circuit (PRV 2 in fig. 1), depending on the set pressure. To effectively cool the system between tests without adjusting the setting of the pressure relief valve PRV2, the cooling system is installed in the return line of the auxiliary pump. The circuit of the auxiliary pump can be operated separately by adjusting the radial piston pump to zero displacement. The pressure of the auxiliary pump is limited by a pressure relief valve (PRV 1 in fig. 1) to around 30 bar, which is near the minimal pressure for the stroke ring adjustment. Actually, the auxiliary circuit is designed for higher pressures, but the lower pressure helps to prevent excessive heating of the oil and reduces the required motor torque.

The pressure in the main hydraulic circuit is limited by the pressure relief valve PRV2 (installed at the outlet of the radial piston pump). To protect the compensator coupling prototype and to prevent overloading the electric motor, the maximum permissible pressure in the main hydraulic circuit is limited to 130 bar.

To investigate the compensator coupling under hydraulic resonance conditions, two different hydraulic systems installed downstream of the shut-off valve are considered. The first one includes a closed pipeline and is shown in fig. 1. The explanation of this hydraulic system is provided in Section 4.1. The second one includes a hydraulic cylinder (see fig. 14).

For measurement data acquisition, the test rig is equipped with an incremental encoder to record the angular motor position, a torque measurement shaft, and a pressure sensor at the pump outlet. All measurements are conducted with a sampling rate of 5 kHz.

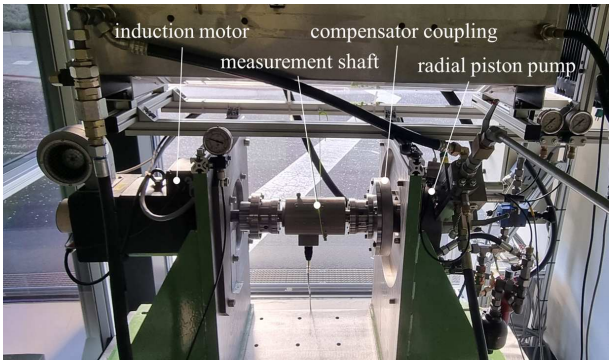


Figure 2: Test rig with compensator coupling.

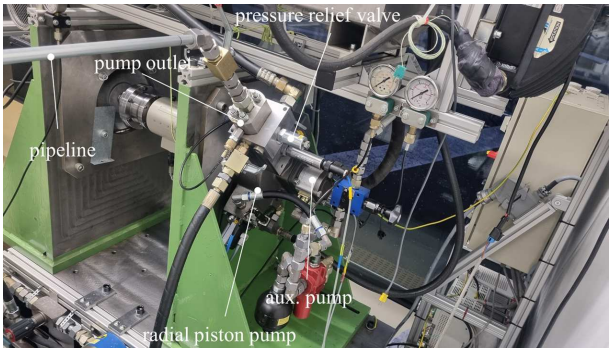


Figure 3: Radial piston pump and outlet block.

Figure 4: Compensator coupling.

3 Compensator coupling

The compensator coupling prototype consists of a spoke wheel and a flange. The assembled compensator coupling is shown in fig. 4. Inner and outer ring of the spoke wheel are connected through radial beams. For calculating the compensation frequency, the drivetrain is modelled as two-mass torsional oscillator. The mass moment of inertia I_M of the motor side contains the inertia of the motor, the measuring shaft including the curved tooth couplings, the flange and the outer ring. The mass moment of inertia I_P on the pump side includes the inertia of the inner ring, the radial piston pump and the auxiliary pump. With a compensation frequency f_C, the required stiffness of the compensator coupling is given by

$$k_T = (2\pi f_C)^2 \frac{I_M I_P}{I_M + I_P}. \tag{1}$$

For the initial design of the compensator, the radial beams were treated as bending beams according to first-order theory. This calculation was performed in [6]. As a result of this consideration, the stiffness of the compensator coupling would only depend on material parameters and dimensions. To understand the actual behaviour of the coupling, measurements have been carried out at eight different system pressures, corresponding to eight different torque values in the drivetrain. The pressure levels were set by adjusting the pressure relief valve PRV2 according to fig. 1. To excite the frequency range under investigation, the motor speed was increased linearly from 850 rpm to 1700 rpm. For the determination of the compensation frequency f_C, the signal of the pump pressure p_P was transferred into the frequency domain using Fourier Transformation. More details of the test procedure are described in section 4. Figure 5 demonstrates the determination of compensation frequencies through four selected measurement results of the pump pressure p_P in the frequency domain. In each signal at least two local minima of the amplitude occur. In each measurement a reduction of the amplitude occurs in the frequency range around 162 Hz, which is not caused by the compensator. This is discussed in more detail in section 4.1. The lower anti-resonance frequency varies between the different experiments. The red line in fig. 5 illustrates the result of experiment 1 from tab. 1, with the compensation frequency of 137.2 Hz at a pump pressure of 60.6 bar. In contrast the purple line shows the result of experiment 8, with a pump pressure of 122.2 bar and a detected compensation frequency of 153.6 Hz.

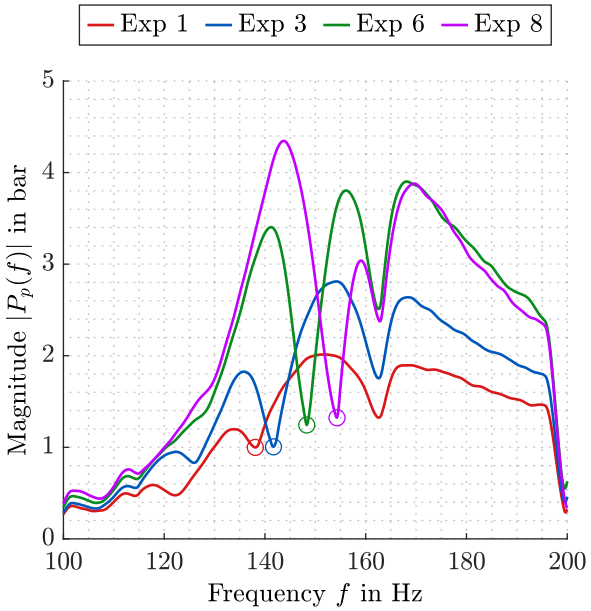


Figure 5: Frequency Spectra of selected measurement results.

The results for all eight different system pressures are summarized in tab. 1, which includes the torque T_M in the drivetrain and the compensation frequency f_C at the pump pressure p_P. It is evident that the compensation frequency increases with increasing torque. This suggests that the modelling with first-order theory is only a rough approximation for the calculation of a compensator coupling, as it cannot account for the dependency on the torque. Second-order theory would include normal stresses that increase stiffness with increasing angular displacement, leading to torque dependence. In the investigated pump pressure range from approximately 60 bar to 120 bar, the compensation frequency increases by 12.7 %.

Table 1: Dependence of the compensation frequency.

Number	Pump pressure p_P	Drivetrain torque T_M	Compensation frequency f_C
Exp 1	60.6 bar	41.4 Nm	138.1 Hz
Exp 2	71.8 bar	46.8 Nm	139.9 Hz
Exp 3	83.7 bar	53.0 Nm	141.7 Hz
Exp 4	94.5 bar	57.3 Nm	144.7 Hz
Exp 5	100.6 bar	61.4 Nm	146.7 Hz
Exp 6	105.8 bar	65.1 Nm	148.3 Hz
Exp 7	117.9 bar	71.2 Nm	151.9 Hz
Exp 8	122.2 bar	74.3 Nm	154.3 Hz

As opposed to the motor, the mass geometry of the pump is not known. To verify the dynamic measurements and estimate the mass moment of inertia of the pump side, static tests and finite element (FEM) analyses were conducted to determine the stiffness k_T. The results are shown in tab. 2. The stiffness values from measurements and FEM analyses match well. Therefore, the mass moment of inertia of the pump side is approximated for the torques given in tab. 1, using the stiffness

calculated from FEM analyses and the measured frequencies. The resulting value is approximately $3.1 \cdot 10^{-3}$ kg m^2.

Table 2: Identification of stiffness k_T.

Static torque	Measured stiffness	FEM stiffness
50 Nm	2540 Nm/rad	2470 Nm/rad
75 Nm	2990 Nm/rad	2920 Nm/rad

4 Hydraulic system with closed pipeline

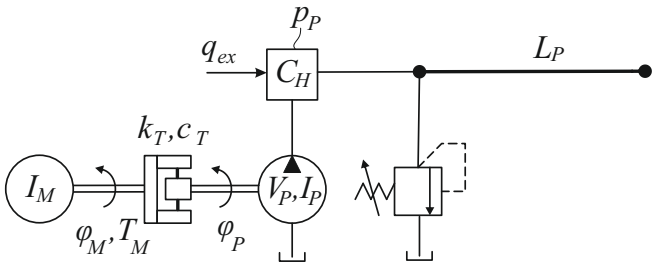


Figure 6: Hydraulic system with pipeline.

To design an experimental setup with a resonance frequency near the compensation frequency, the system is initially simulated in the frequency domain. Figure 6 shows a schematic of the hydraulic system with a closed pipeline including all components considered in the model. The pressure relief valves and the auxiliary pump are neglected. The system contains a closed hydraulic pipeline with a length L_p on the pump outlet. The coupled hydraulic-mechanical model, which is described more closely in [5], can be written as

$$(-\omega^2\mathbf{M} + i\omega\mathbf{C} + \mathbf{K})\,\mathbf{X}(i\omega) = \mathbf{F}(i\omega) \tag{2}$$

with the mass matrix $\mathbf{M}$, the damping matrix $\mathbf{C}$, the stiffness matrix $\mathbf{K}$, the state vector $\mathbf{X}$ and the excitation vector $\mathbf{F}$. In explicit form they read

$$\mathbf{M} = \begin{bmatrix} I_M & 0 & 0 \\ 0 & I_P & 0 \\ 0 & -\frac{V_p}{2\pi} & C_H \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & c_T & 0 \\ 0 & 0 & S_{12} \end{bmatrix},$$

$$\mathbf{K} = \begin{bmatrix} k_T & -k_T & 0 \\ -k_T & k_T & \frac{V_P}{2\pi} \\ 0 & 0 & 0 \end{bmatrix}, \tag{3}$$

$$\mathbf{X} = \begin{bmatrix} \Phi_M(i\omega) \\ \Phi_P(i\omega) \\ P_P(i\omega) \end{bmatrix}, \qquad \mathbf{F} = \begin{bmatrix} T_M(i\omega) \\ 0 \\ i\omega Q_{ex}(i\omega) \end{bmatrix}.$$

The states of the linear model are the rotation angle of the motor Φ_M, the rotation angle of the pump Φ_P and the pressure at pump outlet P_P. T_M describes the moment of the electric motor and Q_{ex} represents the flow rate pulsation caused by the pump mechanism. V_P describes the displacement of the radial piston pump, C_H the hydraulic capacity of the pump outlet block and c_T the torsional damping of the compensator coupling. The function

$$S_{12} = \frac{\sinh(\gamma L_P)}{Z\cosh(\gamma L_P)} \tag{4}$$

with

$$Z = \frac{\sqrt{E\rho}}{r_P^2\pi}\sqrt{-\frac{J_0(ir\sqrt{i\omega/\nu})}{J_2(ir\sqrt{i\omega/\nu})}} \tag{5}$$

and

$$\gamma = \frac{i\omega}{\sqrt{E/\rho}}\sqrt{-\frac{J_0(ir\sqrt{i\omega/\nu})}{J_2(ir\sqrt{i\omega/\nu})}} \tag{6}$$

considers the closed-end pipeline by a dynamic laminar pipeline model as developed in [8]. J_0 and J_2 are Bessel functions of first kind. The parameters for the simulation are given in tab. 3. To account for additional damping in the hydraulic system e.g. at bends or junctions, the kinematic viscosity is increased from 32 m^2 s^{-1} to 100 m^2 s^{-1} for the simulations. The damping of the pump rotor c_T has been estimated in [5]. Only the gap at inner rotor radius is modelled. c_T is increased to account for further sources of damping, e.g. piston $\leftrightarrow$ barrel, slippers or rotor faces $\leftrightarrow$ housing. To ensure that the compensation frequency f_C exactly matches the resonance frequency, the stiffness k_T was calculated using equation (1).

Table 3: Parameters for the experiment with closed pipeline.

I_M	$158.4 \cdot 10^{-3}$ kg m^2	Inertia motor side
I_P	$3.1 \cdot 10^{-3}$ kg m^2	Inertia pump side
V_P	32 cm^3/rev	Displacement volume
V_H	160 cm^3	Volume pump outlet block
E	$1.5 \cdot 10^9$ Pa	Bulk modulus
ρ	835 kg m^{-3}	Density
ν	100 m^2 s^{-1}	Kinematic viscosity
c_T	0.5 Nm/rad s	Torsional damping
k_T	2810 Nm/rad	Torsional stiffness
L_P	3.2 m	Pipeline length
r_P	6 mm	Pipeline inner radius
C_H	V_H/E	Capacity pump outlet

The simulation results, as transfer functions between flow rate excitation and pressure response $G_{PQ}(i\omega) = P_P(i\omega)/Q_{ex}(i\omega)$, are shown in fig. 7. The blue line represents the transmission behaviour without compensator coupling. The pipeline length was chosen such that the hydraulic resonance occurs around 150 Hz. In fig. 7, the frequency response with compensator coupling is shown in red. The simulation predicts a clear reduction in magnitude at the compensation frequency. Below and above this frequency, two secondary resonances arise, where the peaks depend on damping. These simulation results suggest that a well-tuned compensator can significantly improve the system behaviour in the resonance case.

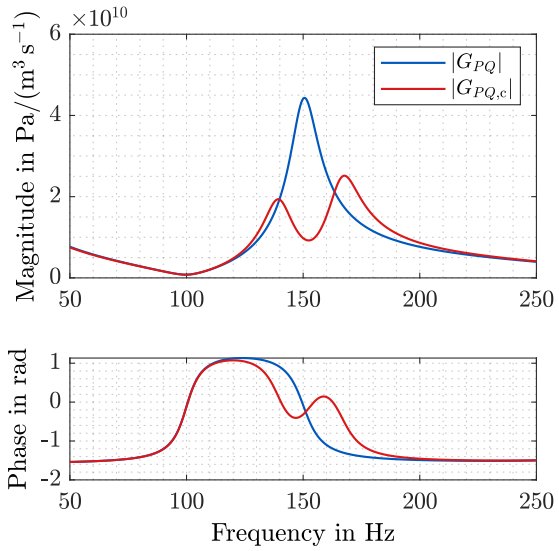


Figure 7: Simulation results with pipeline as transfer function between flow rate excitation and pressure response. Blue: Without compensator, Red: With compensator.

4.1 Run-up experiment with closed pipeline

Based on the simulation results and the investigation concerning the compensation frequency, the hydraulic system in fig. 1 was tuned for the existing compensator. For this purpose, a 3.25 m long pipeline, including all fittings, was used. The pipeline is shown in fig. 8. To maximize the visibility of the compensator effect, high pulsations are required in hydraulic resonance. A lower oil viscosity reduces the damping of the hydraulic system. Therefore, the oil in the system is heated

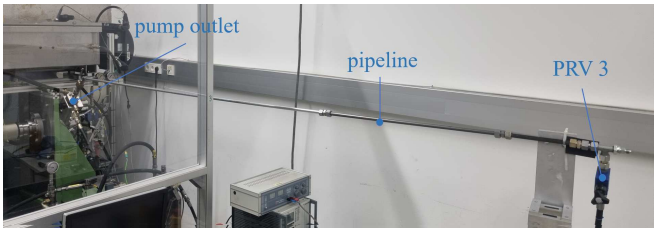


Figure 8: Hydraulic system with closed pipeline.

up to approximately 50 °C. To get a uniform oil temperature in both the pipeline and the tank, oil must flow through the pipeline. To achieve this, a pressure relief valve (PRV 3 in fig. 1) is mounted at the end of the pipeline, which is inactive during the measurements. Additionally, it is essential for venting the pipeline.

To analysis the effect of the compensator coupling, the pressure at the pump outlet is evaluated. For a run-up experiment, fig. 9 shows the pump pressure p_P in the time domain. At the top, the signal of the pressure sensor is displayed, in blue without compensator and in red with compensator. To suppress the influence of noise, a low-pass filter with a cut-off frequency of 1000 Hz was used. For better visibility, the envelope of the reference measurement is shown. In the reference measurement the pressure level was around 1.5 bar lower. Due to the flow rate dependence of the pressure relief valve, the pump pressure increases with higher rotational speed. The diagram in the middle of fig. 9 shows the pressure signal filtered with a bandpass filter with a passband frequency from 50 Hz to 250 Hz. In the measurement without the compensator, a significant reduction of the pressure amplitude is observed after 3.5 seconds. At this point, the reference measurement indicates a pressure pulsation amplitude of approximately 4.2 bar. In contrast, the amplitude of the measurement with the compensator is 1.3 bar. This corresponds to an amplitude reduction by a factor of more than three in the resonance case. For the experiment, the rotational speed of the motor was increased with constant acceleration from 850 rpm up to 1700 rpm, as shown at the bottom in fig. 9. At the resonance case, after 3.5 seconds the motor operates at a speed of 1270 rpm. The frequency of the dynamic flow rate pulsation can be calculated as follows:

$$f_{ex} = \frac{n_{mot}}{60} m_p = 148.2\text{Hz}, \tag{7}$$

with the number of pistons $m_p = 7$ and the motor speed n_{mot} in rpm.

For the analysis in the frequency domain, the pump pressure at the pump outlet is transformed using the Short-Time Fourier Transformation (STFT). Figure 10 shows the result of the reference measurement on the top and of the measurement with compensator at the bottom visualised as 3D waterfall plot. The STFT is applied to analyse how the frequency components of the pressure signal evolve during the measurement, as the excitation frequency changes over time. Therefore, the signal is divided into small overlapping time segments, with each segment transformed into the frequency domain using a Fast Fourier Transform (FFT). At the start, an excitation at 100 Hz is observed, and by the end of the measurement, it has shifted to 200 Hz. This corresponds to the motor speed. Higher-order excitations are also visible. The effect of the compensator is clearly noticeable after 3.5 seconds and at a frequency of around 150 Hz.

To simplify the analysis, only the frequency axis is examined. Figure 11 shows the individual Fourier-transformed signals overlaid, with only the maximum values considered. In contrast to fig. 10, the overlap in the STFT was increased to improve the time resolution and obtain a smoother signal. Furthermore, the data points were interpolated cubically. The maximum compensation effect occurs at 148.3 Hz. The calculated excitation frequency matches the measured compensation frequency. This confirms that the dynamic flow rate pulsation excites the system. As described in [9], the geometric flow rate pulsation would have twice this frequency due to the odd number of pistons. In the frequency range between 140 Hz and 155 Hz, the compensator improves the system behaviour. Compared to the simulation results shown in fig. 7, the frequency range in which the compensator coupling improves the system is more narrowband. When comparing measurement and simulation, it must be considered that the simulation represents a transfer function, while the measurement reflects a system response. In the simulation, a constant excitation

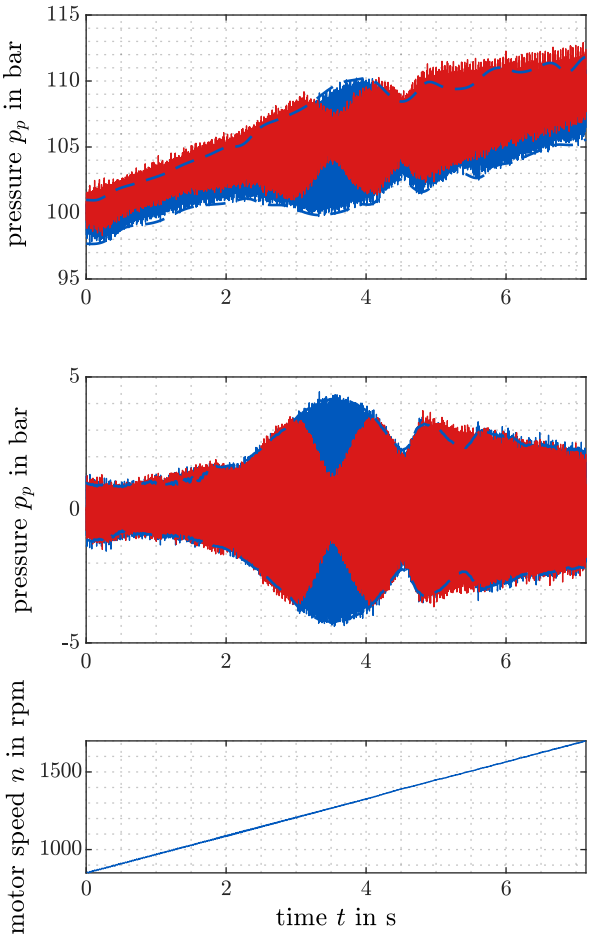


Figure 9: Measurement results of the run-up experiment with closed pipeline, Blue: Without compensator, Red: With compensator, Top: Low-pass filtered signal of the pressure on the pump outlet, Middle: Bandpass filtered pressure signal at the pump outlet, Bottom: Motor rotation speed.

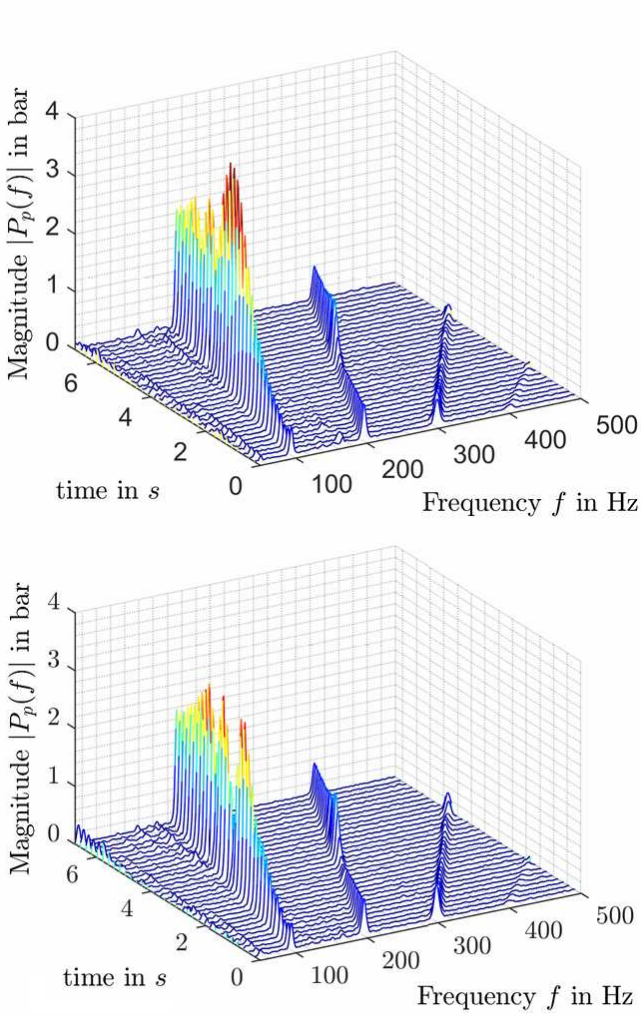


Figure 10: STFT of the pressure at the pump outlet from the run-up measurement with closed pipeline, visualized as a 3D waterfall plot. Top: Without compensator, Bottom: With compensator

amplitude over time was assumed, which does not account for the effect of rotor oscillations on dynamic pressure pulsations. Besides the narrowband characteristics, a difference appears when analysing secondary resonances. The simulations predict two secondary resonances, which leads to increased pulsations caused by the compensator. Neither in the time domain nor in the frequency domain measurements, secondary resonances appear that would degrade the system behaviour. The test setup was designed in such a way that the main resonance dominates the system behaviour. However, additional amplitude reductions are observed at 162 Hz in both measurements and at 173 Hz only in the reference measurement. Compared to the effect of the compensator coupling, these reductions are considered as negligible and are not further investigated.

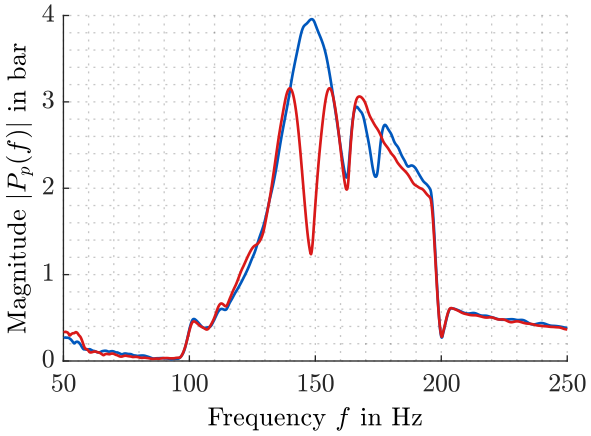


Figure 11: Fourier transformed signal of the pressure at the pump outlet of the run-up measurement with closed pipeline, Blue: Without compensator, Red: With Compensator.

5 Hydraulic system with cylinder

The second investigated scenario involves a cylinder chamber connected to the pump via a hose. The measurements primarily analyse the behaviour of the hydraulic system, with the cylinder in its limit stop position. For this configuration, the frequency response is first analysed through a simulation. Figure 12 illustrates the hydraulic system including all components relevant for the simulation. At the limit stop position, the piston chamber can be modelled as constant hydraulic capacity C_C and the rod side can be neglected.

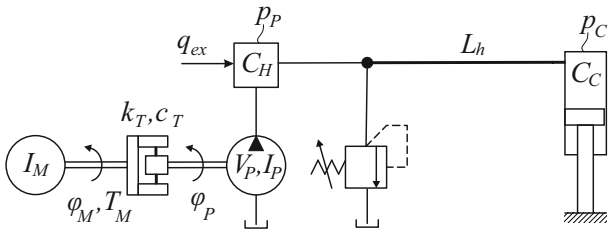


Figure 12: Hydraulic system with cylinder in limit stop position.

A dynamic laminar model is applied to describe the hose, considering the inlet and outlet pressure as the input (p_P and p_C). For the calculation in the frequency domain, this can be written as

$$\begin{bmatrix} Q_H(i\omega) \\ Q_C(i\omega) \end{bmatrix} = \begin{bmatrix} S_{12} & T_{12} \\ T_{12} & S_{12} \end{bmatrix} \begin{bmatrix} P_P(i\omega) \\ P_C(i\omega) \end{bmatrix} \tag{8}$$

with

$$S_{12} = \frac{i\sin(2\gamma' L_h/i)}{Z'(\cos(2\gamma' L_h/i) - 1)}, \tag{9}$$

$$T_{12} = \frac{i}{Z'(\sin(\gamma' L_h/i) - 1)}, \tag{10}$$

$$Z' = \frac{\sqrt{E'\rho}}{r_P^2 \pi} \sqrt{-\frac{J_0(ir\sqrt{i\omega/\nu})}{J_2(ir\sqrt{i\omega/\nu})}} \tag{11}$$

and

$$\gamma' = \frac{i\omega}{\sqrt{E'/\rho}} \sqrt{-\frac{J_0(ir\sqrt{i\omega/\nu})}{J_2(ir\sqrt{i\omega/\nu})}}. \tag{12}$$

With the hose model and the cylinder represented as constant capacity, the system is described by equation (2), using the adapted system matrices **M**, **C** and **K**, according to the state vector **X** and the excitation vector **F**:

$$\mathbf{M} = \begin{bmatrix} I_M & 0 & 0 & 0 \\ 0 & I_P & 0 & 0 \\ 0 & -\frac{V_P}{2\pi} & C_H & 0 \\ 0 & 0 & 0 & C_C \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & c_T & 0 & 0 \\ 0 & 0 & S_{12} & T_{12} \\ 0 & 0 & T_{12} & S_{12} \end{bmatrix},$$

$$\mathbf{K} = \begin{bmatrix} k_T & -k_T & 0 & 0 \\ -k_T & k_T & \frac{V_P}{2\pi} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{X} = \begin{bmatrix} \Phi_M(i\omega) \\ \Phi_P(i\omega) \\ P_P(i\omega) \\ P_C(i\omega) \end{bmatrix}, \qquad \mathbf{F} = \begin{bmatrix} T_M(i\omega) \\ 0 \\ i\omega Q_{ex}(i\omega) \\ 0 \end{bmatrix}. \tag{13}$$

The simulation parameters for the system with the cylinder are listed in tab. 3 and tab. 4. The bulk modulus value for the hose is taken from [10].

Table 4: Model parameters for the experiment with cylinder.

k_T	2600 Nm/rad	Torsional stiffness
E'	$0.6 \cdot 10^9$ Pa	Bulk modulus hose
V_C	$1.8 \cdot 10^{-3}$ m^3	Volume cylinder chamber
L_h	1.25 m	Hose length
r_h	8 mm	Hose inner radius
C_C	V_C/E	Capacity cylinder chamber

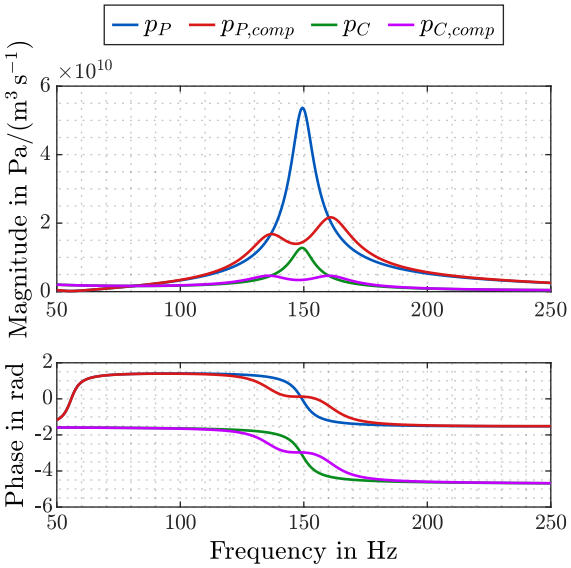


Figure 13: Simulation results with cylinder as transfer functions between flow rate excitation and pressure. Blue & Red: Pressure at the pump outlet without (p_P) & with compensator ($p_{P,comp}$), Green & Purple: Pressure in the cylinder chamber without (p_C) & with compensator ($p_{C,comp}$).

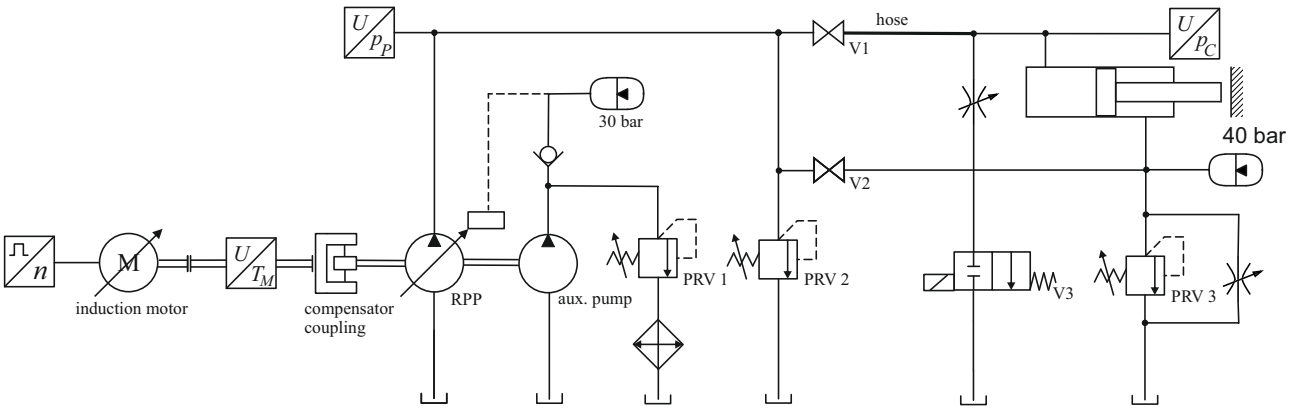


Figure 14: Overall hydraulic layout including the drivetrain, with compensator coupling and a cylinder.

Figure 13 shows the calculated transfer functions between flow rate excitation and pressure response at the pump outlet $G_{P_PQ}(i\omega) = P_P(i\omega)/Q_{ex}(i\omega)$ and in the cylinder chamber $G_{P_CQ}(i\omega) = P_C(i\omega)/Q_{ex}(i\omega)$. The length of the hose was adjusted so that the resonance frequency occurs near 150 Hz. By calculating the stiffness of the compensator coupling with equation (1), the compensation frequency was tuned for the simulation. The effect of the compensator is observed at both the pump outlet and the cylinder chamber.

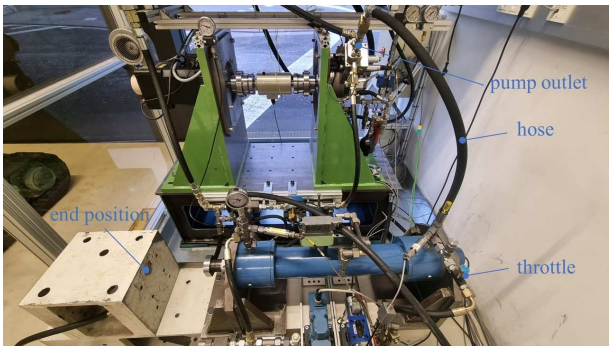


Figure 15: Test rig with cylinder.

Figure 14 illustrates a simplified layout of the hydraulic system used and fig. 15 shows the test rig including the cylinder. Based on the simulations results, the cylinder stroke limitation and the hose length were chosen. The dead volume of the cylinder and the bulk modulus of the hose were not exactly known. With all fittings, the hose length is approximately 1.3 m. To retract the cylinder after the experiment, a hydraulic accumulator is attached to the rod side of the cylinder. The line for filling the hydraulic accumulator is closed during the experiments by the shut-off valve V2. In order to prevent the closed filling line from influencing the system behaviour, the shut-off valve is directly installed on a second outlet of the pump block. The pressure relief valve PRV3 with closed parallel throttle limits the pressure in the hydraulic accumulator. During retraction, the switching valve V3 connects the piston side of the cylinder with the tank. To reduce the influence of the line to the switching valve, a throttle is installed near the cylinder chamber, which is almost closed for the experiments. During the experiments, the valve V1 is open and the valves V2 & V3 are closed. As in the previous experiments, the pressure is limited by the pressure relief valve (PRV 2 in fig. 14) at the pump outlet.

5.1 Run-up experiment with cylinder in limit stop position

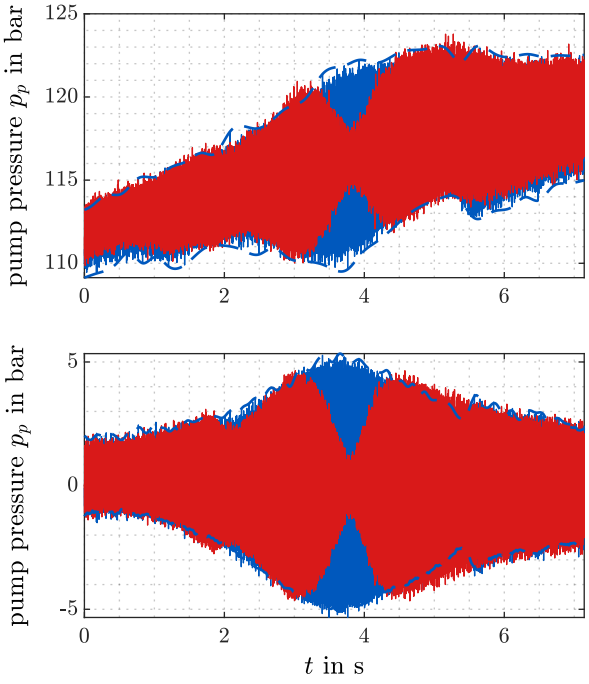


Figure 16: Measurement results of the run-up experiment with cylinder in limit stop position. Blue: Without compensator, Red: With compensator. Top: Low-pass filtered signal of the pressure at the pump outlet, Bottom: Bandpass filtered signal of the pressure at the pump outlet.

In this experiment, the cylinder is already at its limit stop position. The motor accelerates uniformly from 850 rpm to 1700 rpm while the pump generates a dominant excitation in a frequency range of 100 Hz to 200 Hz. Figure 16 shows the measurement result in the time domain. The upper illus-

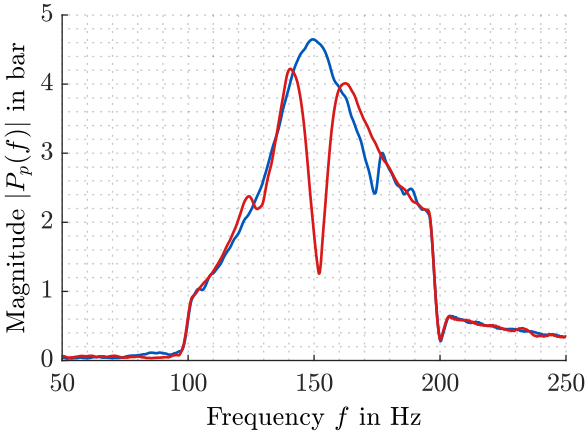


Figure 17: Fourier transformed signal of the pressure at the pump outlet, run-up measurement with cylinder in limit stop position. Blue: Without compensator, Red: With Compensator.

tration displays the low-pass filtered signal (cut-off frequency 1000 Hz) with and without compensator. After approximately 3.7 seconds, the largest notch in the pressure signal occurs. At the bottom of fig. 16 the bandpass filtered signal is shown. The bandpass filter has a passband frequency from 50 Hz to 250 Hz. The measurement without the compensator shows a pressure pulsation amplitude of 4.7 bar, whereas with the compensator, it is 1.1 bar. Figure 17 displays the pressure signal in the frequency domain. As in the experiment with the closed pipeline, significant compensation occurs at approximately 150 Hz. In contrast, secondary resonances are clearly noticeable. The increased pressure pulsation at the secondary resonances can be observed in the bandpass filtered pressure signal in the time domain. After 4.3 seconds the amplitude of the reference measurement is slightly lower than the amplitude of the measurement with the compensator. Significant improvements in system behaviour occur in the frequency range from 142 Hz to 158 Hz.

Figure 18 illustrates the pressure in the cylinder chamber in the time domain. On top the low-passed filtered signals (cut-off frequency 1000 Hz) are shown. No constriction of the amplitude is visible. Through further signal processing in the form of a bandpass filter with a passband frequency from 50 to 250 Hz, the compensator effect becomes apparent, which is shown at the bottom. The Fourier transformed signal in fig. 19 reveals that the second harmonic frequency, at approximately 350 Hz, oscillates significantly higher than the first. In all previous measurements, the amplitudes at the second harmonic frequency were significantly lower than the amplitudes at the first and were therefore not taken into account. The highest pulsation is observed after 5.4 seconds in fig. 18. At this point, the motor drives the pump at 1500 rpm. Equation (7) gives an excitation frequency of 175 Hz. This is consistent with the observations from the frequency analysis.

5.2 Extending experiment to limit stop position

In the previous investigations, the experimental setup used the combination of a speed controlled motor and a pump with con-

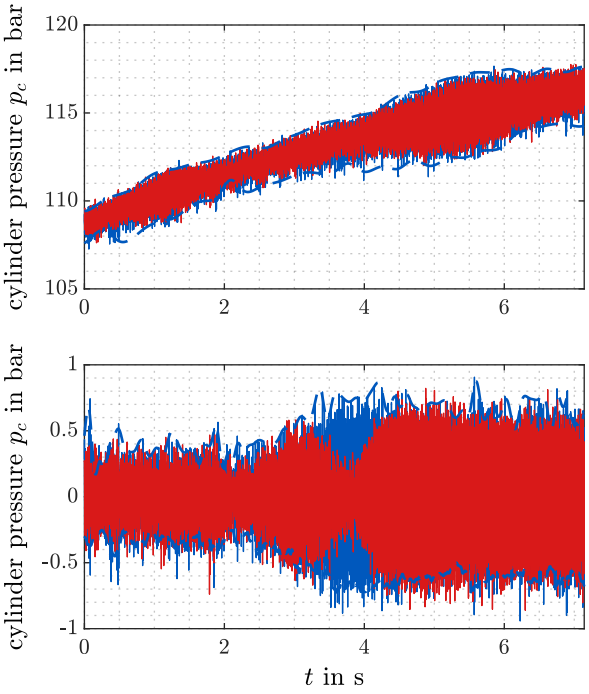


Figure 18: Measurement results of the run-up experiment with cylinder in limit stop position. Blue: Without compensator, Red: With compensator. Top: Low-pass filtered signal of the pressure in the cylinder chamber, Bottom: Bandpass filtered signal of the pressure in the cylinder chamber.

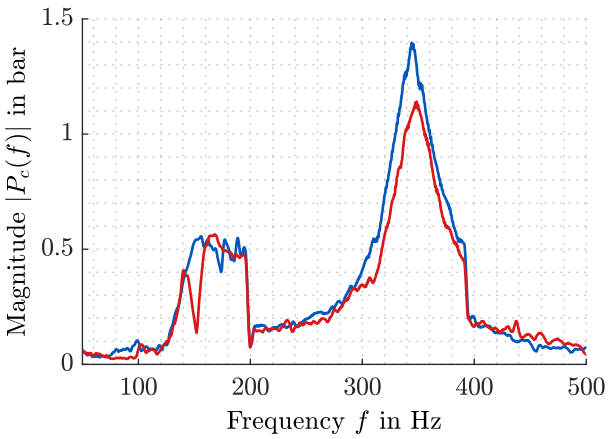


Figure 19: Fast Fourier transformed signal of the pressure in the cylinder chamber, run-up measurement with cylinder in limit stop position. Blue: Without compensator, Red: With compensator.

stant displacement volume. In another experiment, the motor operates at a constant speed of 1280 rpm, resulting in a dominant excitation frequency of approximately 148 Hz. At the beginning of the experiment, the displacement of the pump is set to 0 cm^3/rev and the cylinder is fully retracted. To extend the cylinder to its limit stop position, the displacement of the pump is set to its maximum of 32 cm^3/rev. Figure

20 shows the pressure signal at the pump outlet. During the first 2.8 seconds the cylinder extends and the pressure in the piston chamber increases due to the hydraulic accumulator on the rod side. At the beginning, the compensator has no significant effect on the pressure pulsation. Due to the undesirable dependency between drivetrain torque and compensation frequency, the compensator is misadjusted in that case. As the pressure increases, the system gets closer to the operating range of the compensator. Just before reaching the limit stop position at 2.8 seconds, a reduction in pressure pulsation becomes noticeable. At the end position, where the compensator is well tuned, the pressure amplitude decreases by a factor of two. The measurement confirms that the compensator coupling achieves significant improvements at the correct operating point.

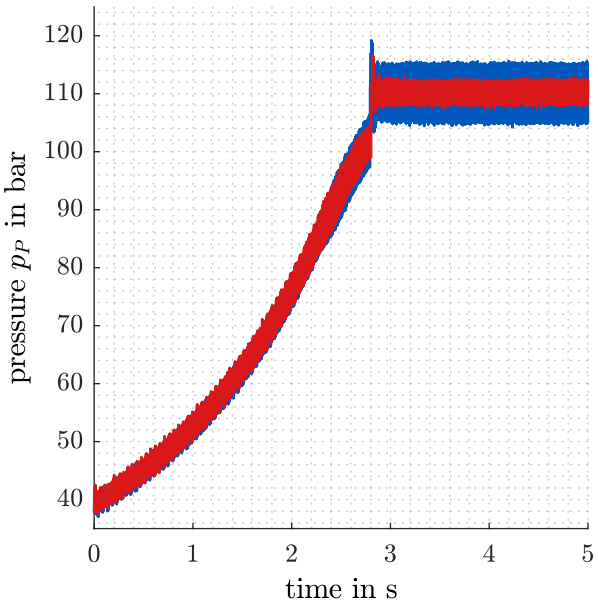


Figure 20: Measurement result from extending cylinder to limit position. Blue: Pressure signal at the pump outlet without compensator, Red: Pressure signal at the pump outlet with compensator, both low-pass filtered at 1000 Hz.

6 Conclusion and outlook

Investigations with different hydraulic systems in hydraulic resonance were made using the compensator coupling in the drivetrain. The measurement results demonstrate that significant improvements can be achieved with this type of compensator at the compensation frequency. In the case of hydraulic resonance, the amplitude of the pressure pulsation was reduced by a factor greater than three. The experiments also showed that the compensator coupling, in combination with the radial piston pump, is narrowband and exhibits an undesirable dependence between torque and compensation frequency. The dependence is particularly unfavorable for a narrowband compensator, whereas it may be negligible for a broadband one. To increase the bandwidth of the compensator, lower inertia on the pump side is required. Therefore, investigations with gear pumps, which exhibit a significantly lower inertia than the used radial piston pump, are planned as a next step. For the next prototypes, new designs are being investigated, whose torsional stiffness should be independent of the operating conditions. Once a broadband rotary mechanical pulsation compensator has been developed, comparative measurements can be carried out with other compensation systems, such as the Helmholtz resonator, to allow a direct comparison of their performance under identical conditions.

References

[1] M Stosiak, I Yatskiv, O Prentkovskis, and M Karpenko. Reduction of Pressure Pulsations over a Wide Frequency Range in Hydrostatic Systems. Machines 13(1), 25, 2025. DOI: 10.3390/machines13010025.

[2] Y Wang, T Shen, C Tan, et al. Research Status, Critical Technologies, and Development Trends of Hydraulic Pressure Pulsation Attenuator. Chin. J. Mech. Eng. 34, 14, 2021. DOI: 10.1186/s10033-021-00532-z.

[3] J Mikota and H Reiter. Development of a compact and tuneable vibration compensator for hydraulic systems. *International Journal of Fluid Power*, 4(1):17–31, 2003. DOI: 10.1080/14399776.2003.10781152.

[4] G Mikota. Vorrichtung zur Reduktion von Druckpulsationen in einem Hydrauliksystem. Austrian Patent AT 526449 B1, 2024.

[5] G Mikota and B Manhartgruber. A novel pulsation compensator for displacement machines. In *Proceedings of 14th International Fluid Power Conference*, pages 395–406, Dresden, Germany, March 19-21, 2024. DOI: 10.13052/rp-9788770042222C30.

[6] G Mikota. Design and Simulation of a Drivetrain Integrated Pulsation Compensator for Pumps. In *BATH/ASME 2024 Symposium on Fluid Power and Motion Control*, Bath, England, September 11-13, 2024. V001T01A008. DOI: 10.1115/FPMC2024-140002.

[7] J Rametsteiner. Entwicklung einer torsionselastischen Tilgerkupplung für eine Radialkolbenpumpe. Bachelor's thesis, Johannes Kepler University, Institute of Machine Design and Fluid Power, Linz, August, 2024.

[8] A F D'Souza and R Oldenburger. Dynamic Response of Fluid Lines. *Transactions of the ASME - Journal of Basic Engineering*, 86:589–598, 1964. DOI: 10.1115/1.3653180.

[9] J Ivantysyn and M Ivantysynova. *Hydrostatische Pumpen und Motoren*. Vogel Verlag und Druck KG, Würzburg, Germany, 1993.

[10] D Will and N Gebhardt. *Hydraulik Grundlagen, Komponenten, Schaltungen*. Springer, 2011. 5. Auflage, S. 21, DOI: 10.1007/978-3-642-17243-4.

Towards Full Electrification: Virtual Thermal Prototyping of a Dual-Mode Electro-Hydrostatic Actuator for a Hybrid Wheel Loader

Baoyan Hu[1,*], Chao Wang[1], and Jian Fu[1]

1. School of Mechanical Engineering and Automation, Beihang University, Beijing, P.R China
E-mail: buaahby@outlook.com, wangchaoo@buaa.edu.cn, fujian@buaa.edu.cn
*School of Mechanical Engineering and Automation, Beihang University, Beijing, P.R China

Abstract

The shift from traditional diesel-powered machinery to more environmentally-friendly alternatives in the non-road mobile machinery sector is driven by strict environmental regulations and the demand for greater efficiency. While hybrid solutions have demonstrated benefits such as improved fuel efficiency and reduced emissions, the electrification of hydraulic actuation systems has received less attention. This study introduces a dual-mode electro-hydrostatic actuator designed to replace the hydraulic lift cylinder in hybrid wheel loaders, addressing challenges in efficiency, safety, and thermal management. The actuator operates in two distinct modes: a typical mode and a safety mode employing counterbalance valves, which enhances load-lowering control and safety through integrated solenoid valves and counterbalance mechanisms. Using co-simulation between AMESim and MATLAB/Simulink, the dynamic performance and thermal behavior of the electro-hydrostatic actuator are evaluated over a 3600 second duty cycle. Simulation results highlight differences in thermal and dynamic performance between the two modes. The safety mode achieves superior load-holding capability with reduced motor current demand but experiences higher temperatures in the cylinder and valves due to discontinuous fluid flow. Conversely, the typical mode encounters greater thermal stress in the pump, particularly at the leakage point. This work provides a detailed analysis of the thermal behavior of the electro-hydrostatic actuator, offering crucial insights for optimizing thermal design and cooling strategies. The study also contributes to the advancement of electrified hydraulic systems and integrating hybrid powertrains with electrified hydraulic systems, paving the way for sustainable, high-performance non-road mobile machinery solutions that meet safety and operational standards.

Keywords: thermal behavior, hybrid wheel loader, electrification, electro-hydrostatic actuator, counterbalance valve

1 Introduction

The transition toward full electrification of non-road mobile machinery (NRMM) is driven by increasingly strict environmental regulations and the rising demand for sustainable technologies[1]. Conventional diesel-powered NRMM confronts challenges such as high emissions and inefficiencies, necessitating innovative solutions. Hybrid powertrains are emerging as practicable alternatives, offering enhanced fuel efficiency, reduced emissions, improved operational performance, lower noise levels, and reduced maintenance costs[2]. Significant research has explored various hybrid powertrains, the focus has primarily been on powertrains[3]. For instance, [4], [5], [6] introduces three hybrid powertrains developed by different original equipment manufacturers (OEMs): 1. series 2. parallel 3. series-parallel. However, OEMs retain initial hydraulic circuits that possess system inherent power losses (SIPL).

Besides, limited attention is given to the electrification of hydraulic actuation systems. Two concepts, individual metering system and secondary control using hydraulic transformer are contributions to the improvement of hydraulic actuation systems. Although individual metering system lowers the energy consumption, SIPL still exists when loads are in different weights[7]. Hydraulic transformer is one of valve-less systems that avoids SIPL, but the increased number of hydraulic components and unchanged hydraulic power supply bring issues of more required space and higher costs[8]. Electro-hydrostatic actuators (EHAs) and electro-mechanical actuators (EMAs) are among the promising technologies to address this gap, providing opportunities to improve hybrid NRMM efficiency and enable progress toward full electrification. One of them,

EHA, its essential concept is to connect an electrical-motor-controlled pump to a hydraulic cylinder, realizing varied transmission of power. The servo valve is not required in this scheme and thus the power loss across it can be decreased sharply[9]. Because EHA has the potential to recover energy[10], many scholars started to research its application on NRMM. [11] and [12] proposes specific and independent electrified working hydraulic system by adopting direct-driven hydraulics or EHAs. In [13], the researchers incline their work towards how to evaluate the match level between novel actuators and NRMM. Nevertheless, these papers related to adopting EHAs only take functional compatibility into account. In fact, key factors such as physical characteristics, dynamic performance, and compliance with safety standards like ISO 8643[14]—which regulates load-lowering control in case of hydraulic line failure happening in earth-moving machinery—must be considered as well before EHA is implemented on NRMM for real. Additionally, integrated EHA systems face unique challenges in managing heat due to limited space for heat exchangers[15], making thermal analysis essential. Energy is dissipated through heat conduction, convective heat transfer, and radiative heat transfer as passive outputs, which are not considered effective dissipation mechanisms[16]. Moreover, elevated temperatures have a significant impact on the performance of the EHA [17]. [18], [19], [20] have studied the temperature and efficiency of different EHAs by building models or doing experiments. Nonetheless, dynamic performance and thermal behavior of the EHA cannot be observed at the same time via their models.

Therefore, a tool is needed to simultaneously monitor the dynamic performance and thermal behavior of the EHA that satisfies safety standards, enabling its evaluation for potential implementation in hybrid NRMM. This paper introduces an innovative prototype of an integrated EHA designed to replace the lift cylinder in a hybrid wheel loader. This EHA is equipped with counterbalance valves (CBVs) that could be seen as a kind of lowering control devices which ISO 8643 requires to enhance load control and safety while a hybrid wheel loader is working in the safety mode. Furthermore, it also offers an alternative operating mode without CBVs by energizing solenoid valves, which is called typical mode as normal EHA scheme. Naturally, the EHA under safety mode requires more hydraulic power to open the CBVs, making it less energy-efficient compared to the typical mode. However, ISO 8643 mandates the inclusion of lowering control devices in wheel loaders to avoid hazard. Therefore, comparing the performance differences between the two operating modes is both relevant and meaningful for ensuring regulatory compliance. EHA with dual modes also possesses varied heat transfer mechanisms, including the impact of convective cooling in the pump and accumulator in safety mode and the elevated thermal stress on components in typical mode. In order to investigate the specific EHA's thermal behavior and dynamic performance under dual operating modes, co-simulation via MATLAB/Simulink and AMESim is deployed.

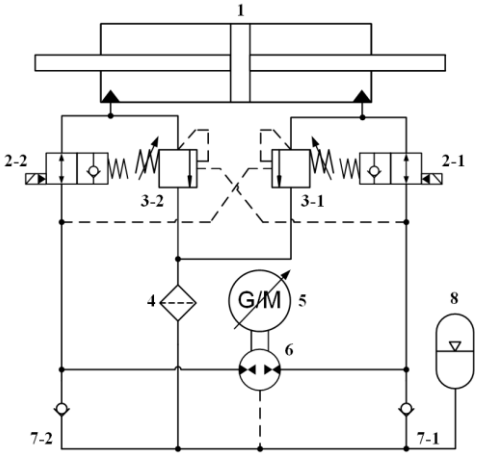
This study not only advances understanding of EHA technology but also provides a pathway for high-performance and sustainable hydraulic actuation systems that meet safety and operational standards, offering valuable guidance for the effective implementation of electrified hydraulic systems in hybrid wheel loaders.

The rest of the paper is organized as follows: Section II analyzes heat transfer in the EHA; Section III presents mathematical thermal models based on the lumped-parameter method (LPM); Section IV shows temperature results from a 3600-second duty cycle; Section V discusses the simulation results; Section VI offers conclusions.

2 Heat transfer analysis

EHA system's configuration is given in 2.1, having great influence on analyzing heat transfer. Thermal network models are illustrated to characterize the heat transfer mechanisms within the EHA system in 2.2.

2.1 EHA system



1-symmetric cylinder 2-solenoid valves 3-counterbalance valves 4-filter 5-PMSM 6-axial piston pump/motor 7-check valves 8-accumulator

Figure 1: Schematic diagram of EHA prototype

Figure. 1 shows the schematic diagram of an integrated EHA prototype. The system is built around a symmetric hydraulic cylinder (1), with one of its rods concealed within the housing. The EHA operates in two modes, triggered by the configuration of the two solenoid valves (2):

Typical Mode: In this typical volume-controlled EHA operation, oil flows directly through both solenoid valves, allowing standard actuation. CBVs (3) act as relief valves for safety. It takes the advantage that the system can realize the more accurate and rapider position control. Compared to safety mode, there is a degradation in system security. This mode serves as the baseline for observing the system's behavior when an EHA with a fundamental structure is directly integrated into a wheel loader.

Safety Mode: Solenoid valves operate as check valves, enabling oil to pass through CBVs. This mode directs the fluid flow into the accumulator (8), through the check valves (7), and then to the suction port of the pump/motor (6). This design prevents oil from flowing directly to the inlet of the pump. During load-lowering control, if a pipeline failure occurs,

CBVs can shut off the flow to prevent the load from descending uncontrollably, thereby mitigating the risk of accidents and significantly enhancing system safety. However, this comes at the cost of reduced control precision and response speed. This mode provides insight into the system's behavior when the EHA is equipped with CBVs.

The hydraulic circuit includes additional components for functionality and efficiency: a filter (4) for fluid cleanliness, an axial piston pump/motor (6), and a permanent magnet synchronous motor (PMSM) (5), which drives the pump via a coupling. Notably, the pump is immersed in the accumulator, providing a compact and integrated design. The accumulator (8) stores energy and helps smooth operation. This configuration offers enhanced flexibility and allows for detailed studies of performance differences between the two operating modes.

2.2 Heat transfer within EHA

The EHA system comprises components involved in heat transfer, including the cylinder (1), solenoid valves (2), CBVs (3), PMSM (5), pump (6) and accumulator (8). Due to dual operating modes, heat transfer mechanisms and thermal network models differ, as shown in fig. 2 and fig. 3. No matter how the pump rotates, in a clockwise direction or a counterclockwise direction, the direction of the heat transfer is always defined by fig. 2 and fig. 3. Circular nodes represent fluid nodes, while rectangular nodes indicate solid nodes. The pump has four nodes, with mechanical losses at the outlet node and churning losses at the leakage node. In typical mode, convective heat transfer happens between the pump shell and accumulator via direct oil contact. The cylinder and solenoid valves are modeled with three nodes each, with mechanical losses distributed to the cylinder inlet node and outlet node, and valve's energy losses at the outlet node. Heat exchange with external environment occurs at the shells of the cylinder, accumulator, and PMSM via convection and radiation. In safety mode, a CBV replaces one solenoid valve, altering heat transfer pathways and facilitating heat transfer from the CBV to the accumulator as depicted in fig. 3.

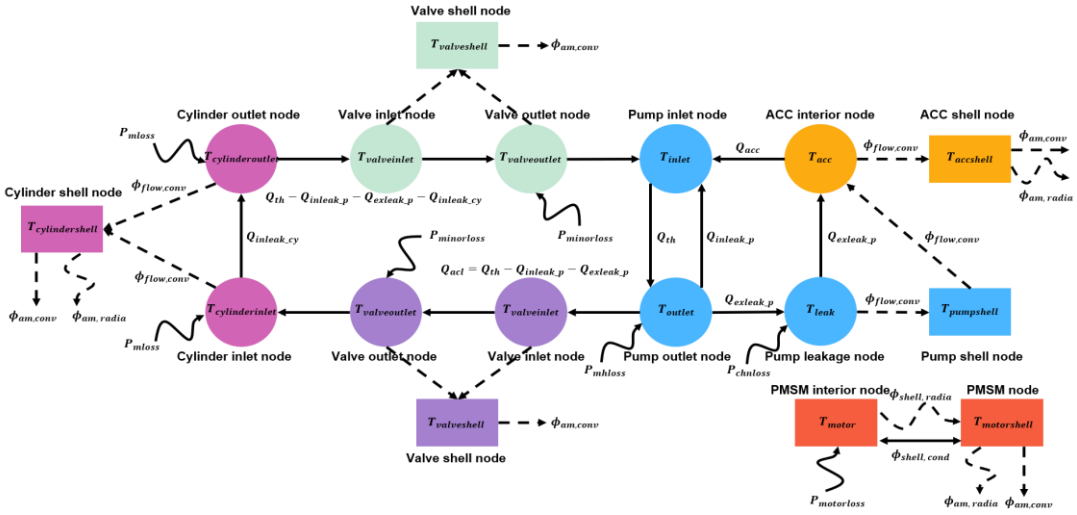


Figure 2: Heat transfer in typical mode based on LPM

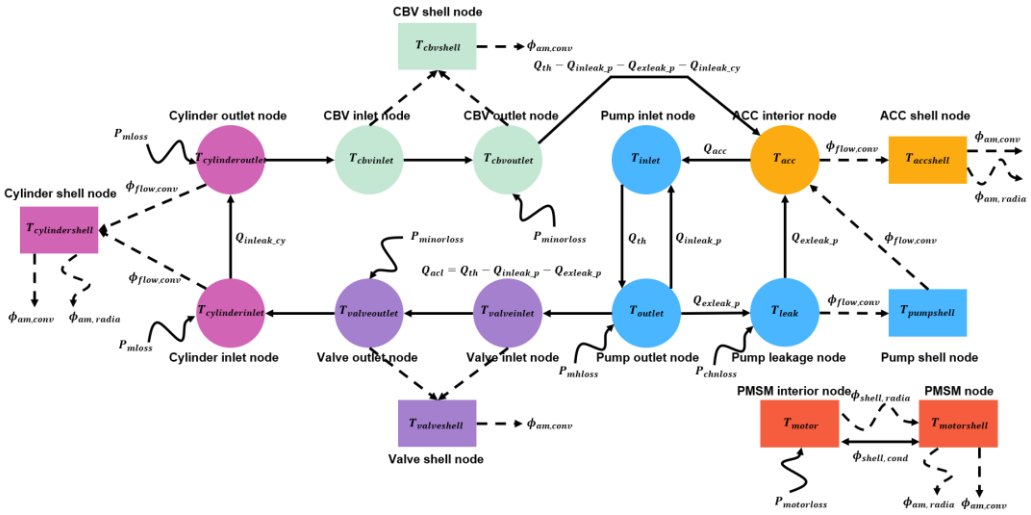


Figure 3: Heat transfer in safety mode based on LPM

3 Thermal modelling

For lumped parameter method and one-dimensional flow, there is a simplified representation of the temperature differential equation. Several authors have presented thermal models of hydraulic systems, for example, referring to [21] for details. In this paper, thermal models are derived from the principles of conservation of energy and mass.

3.1 Basic equation building for different systems

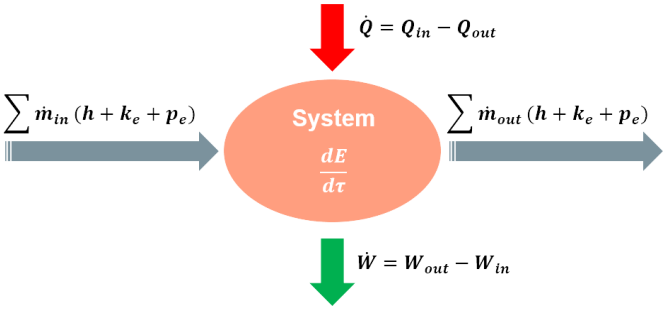


Figure 4: The First law of thermodynamics

Building thermal models must be based on the first law of thermodynamics that is illustrated by fig. 4. There are three primary types of thermodynamic systems categorized by how this system exchanges mass and energy with its surroundings, referring to open systems, closed systems and isolated systems. As shown in fi. 1, cylinder (1), solenoid valves (2), CBVs (3), pump (6) and accumulator (8) belong to open systems, which means that these components exchange energy as well as mass with surroundings. Therefore, an open system should also obey the law of conservation of mass, analyzed through control volume method. PMSM (6) is a kind of closed systems because no mass crosses the boundary of this system, analyzed through control mass method.

For open systems, the temperature differential equation with respect to time has been previously derived in [21]. In this study, the same equation is obtained using an alternative approach. The temperature differential equation is directly presented as follows:

$$\frac{dT}{d\tau} = \frac{1}{c_p m} * \left\{ \begin{array}{c} \sum \dot{m}_{in}\left[c_p(T_{in} - T_{out}) + (1 - \bar{T}\alpha_V)\bar{v}(p_{in} - p_{out})\right] \\ +\dot{Q} - \dot{W}_s + mT\alpha_V v\frac{dp}{d\tau} \end{array} \right\} \quad (1)$$

To enhance the clarity of this equation, the last term within the curly brackets should be extracted separately, as it explicitly represents the effect of pressure variation on temperature change. Apart from this term, the remaining components influence temperature at a given pressure.

Equation (1) is derived based on the assumption that the enthalpy within the control volume is equal to the enthalpy leaving the system. However, this assumption is not practical to all hydraulic components in EHA. Therefore, it is essential to understand the derivation of eq. (1). which is presented in eq. (2):

$$\frac{dT}{d\tau} = \frac{1}{c_p m} * \left\{ \begin{array}{c} \sum \dot{m}_{in}\left[c_p(T_{in} - T) + (1 - T\alpha_V)v_{in}(p_{in} - p)\right] \\ +\sum \dot{m}_{out}\left[c_p(T - T_{out}) + (1 - T\alpha_V)v_{out}(p - p_{out})\right] \\ +\dot{Q} - \dot{W}_s + mT\alpha_V v\frac{dp}{d\tau} \end{array} \right\} \quad (2)$$

For closed systems, the energy equation is quite simpler than that for open systems:

$$\dot{Q} - \dot{W} = \frac{dE_{CM}}{d\tau} \quad (3)$$

Similarly, dE_{CM} is the differential element of the internal energy of this control mass. No mass changes in closed systems, so the term on the right side of the equal sign can be expressed as:

$$\frac{dE_{CM}}{d\tau} = \frac{dU}{d\tau} = \frac{dmu}{d\tau} = m\frac{du}{d\tau} \quad (4)$$

where introduces another thermodynamics property, c_v, which means the specific heat at constant volume. Therefore, the temperature differential equation for the closed system is:

$$\frac{dT}{d\tau} = \frac{1}{c_v m}(\dot{Q} - \dot{W}) \quad (5)$$

3.2 Thermal models of components

Expressions of temperature differential equations on the basis of heat transfer in fig. 2 are as the followings.

The symmetric cylinder (1) in fig. 1 includes 1-inlet node, 1-outlet node and 1-shell node, following the eq. (6-8) respectively:

$$\frac{dT_{cylinderinlet}}{d\tau} = \frac{1}{c_{cylinder} m_{cylinderinlet}} * \left\{ \begin{array}{c} \rho_{oil}\left(Q_{th} - Q_{inleak_p} - Q_{exleak_p}\right) \\ \left[\begin{array}{c} c_{cylinder}\left(T_{valveoutlet} - T_{cylinderinlet}\right) \\ +\left(1 - \frac{\alpha_V\left(T_{cylinderinlet} + T_{valveoutlet}\right)}{2}\right) \\ * \bar{v}\left(p_{valveoutlet} - p_{cylinderinlet}\right) \end{array} \right] \\ -h_{cylindershell}A_{cylinderinlet}\left(T_{cylinderinlet} - T_{cylindershell}\right) \\ +P_{mloss} + m_{cylinderinlet}T_{cylinderinlet}\alpha_V v\frac{dp_{cylinderinlet}}{d\tau} \end{array} \right\} \quad (6)$$

$$\frac{dT_{cylinderoutlet}}{d\tau} = \frac{1}{c_{cylinder} m_{cylinderoutlet}} *$$

$$\begin{Bmatrix} \rho_{oil} Q_{inleak_{cy}} \begin{bmatrix} c_{cylinder}\left(T_{cylinderinlet} - T_{cylinderoutlet}\right) \\ +\left(1 - \dfrac{\alpha_V\left(T_{cylinderoutlet} + T_{cylinderinlet}\right)}{2}\right) \\ * \bar{v}\left(p_{cylinderinlet} - p_{cylinderoutlet}\right) \end{bmatrix} \\ -h_{cylindershell} A_{cylinderoutlet}\left(T_{cylinderoutlet} - T_{cylindershell}\right) \\ +P_{mloss} + m_{cylinderoutlet} T_{cylinderoutlet} \alpha_V \nu \dfrac{dp_{cylinderoutlet}}{d\tau} \end{Bmatrix} \tag{7}$$

$$\frac{dT_{cylindershell}}{d\tau} = \frac{1}{c_{cylindershell} m_{cylindershell}} *$$

$$\begin{bmatrix} h_{cylindershell} A_{cylinderoutlet}\left(T_{cylinderoutlet} - T_{cylindershell}\right) \\ +h_{cylindershell} A_{cylinderinlet}\left(T_{cylinderinlet} - T_{cylindershell}\right) \\ -h_{cylinderamt} A_{cylinderamt}\left(T_{cylindershell} - T_{amt}\right) \\ -\varepsilon A_{cylinderamt} \sigma\left(T_{cylindershell}^4 - T_{amt}^4\right) \end{bmatrix} \tag{8}$$

The solenoid valve (2-1) in fig. 1 involves 2-1-inlet node, 2-1-outlet node and 2-1-shell node. The solenoid valve (2-2) includes 2-2-inlet node, 2-2-outlet node and 2-2-shell node. The modelling principle of two solenoid valves is same, following eq. (9-11):

$$\frac{dT_{valveinlet}}{d\tau} = \frac{1}{c_{valve} m_{valveinlet}} *$$

$$\begin{Bmatrix} \rho_{oil}\left(Q_{th} - Q_{inleak_p} - Q_{exleak_p}\right) \\ \begin{bmatrix} c_{valve}\left(T_{outlet} - T_{valveinlet}\right) \\ +\left(1 - \dfrac{\alpha_V\left(T_{outlet} + T_{valveinlet}\right)}{2}\right) \bar{v}\left(p_{outlet} - p_{valveinlet}\right) \end{bmatrix} \\ -h_{valveshell} A_{valveshell}\left(T_{valveinlet} - T_{valveshell}\right) \\ +m_{valveinlet} T_{valveinlet} \alpha_V \nu \dfrac{dp_{valveinlet}}{d\tau} \end{Bmatrix} \tag{9}$$

$$\frac{dT_{valveoutlet}}{d\tau} = \frac{1}{c_{valve} m_{valveoutlet}} *$$

$$\begin{Bmatrix} \rho_{oil}\left(Q_{th} - Q_{inleak_p} - Q_{exleak_p}\right) \\ \begin{bmatrix} c_{valve}\left(T_{valveinlet} - T_{valveoutlet}\right) + \\ \left(1 - \dfrac{\alpha_V\left(T_{outlet} + T_{valveinlet}\right)}{2}\right) \bar{v}\left(p_{valveinlet} - p_{valveoutlet}\right) \end{bmatrix} \\ -h_{valveshell} A_{valveshell}\left(T_{valveoutlet} - T_{valveshell}\right) \\ +P_{minorloss} + m_{valveoutlet} T_{valveoutlet} \alpha_V \nu \dfrac{dp_{valveoutlet}}{d\tau} \end{Bmatrix} \tag{10}$$

$$\frac{dT_{valveshell}}{d\tau} = \frac{1}{c_{valveshell} m_{valveshell}} *$$

$$\begin{bmatrix} h_{valveshell} A_{valveshell}\left(T_{valveinlet} - T_{valveshell}\right) \\ +h_{valveshell} A_{valveshell}\left(T_{valveoutlet} - T_{valveshell}\right) \\ -h_{valveamt} A_{valveamt}\left(T_{valveshell} - T_{amt}\right) \end{bmatrix} \tag{11}$$

PMSM (5) only has two nodes, 5-interior node and 5-shell node as the following eq. (12) and eq. (13):

$$\frac{dT_{motor}}{d\tau} = \frac{1}{c_{motor} m_{motor}} *$$

$$\begin{bmatrix} P_{motorloss} - \dfrac{\lambda A_{motorshell}\left(T_{motor} - T_{motorshell}\right)}{\delta} \\ -\varepsilon A_{motorshell} \sigma\left(T_{motor}^4 - T_{motorshell}^4\right) \end{bmatrix} \tag{12}$$

$$\frac{dT_{motorshell}}{d\tau} = \frac{1}{c_{motorshell} m_{motorshell}} *$$

$$\begin{bmatrix} \dfrac{\lambda A_{motorshell}\left(T_{motor} - T_{motorshell}\right)}{\delta} \\ +\varepsilon A_{motorshell} \sigma\left(T_{motor}^4 - T_{motorshell}^4\right) \\ -h_{motorshellamt} A_{motorshellamt}\left(T_{motorshell} - T_{amt}\right) \\ -\varepsilon A_{motorshellamt} \sigma\left(T_{motorshell}^4 - T_{amt}^4\right) \end{bmatrix} \tag{13}$$

The axial piston pump/motor (6) is comprised of 6-inlet node, 6-outlet node, 6-leakage node and 6-shell node, as the following eq. (14-17):

$$\frac{dT_{inlet}}{d\tau} = \frac{1}{c_{pump} m_{inlet}} *$$

$$\begin{Bmatrix} \rho_{oil} Q_{inleak_p} \begin{bmatrix} c_{pump}\left(T_{outlet} - T_{inlet}\right) \\ +\left(1 - \dfrac{\alpha_V\left(T_{outlet} + T_{inlet}\right)}{2}\right) \bar{v}\left(p_{outlet} - p_{inlet}\right) \end{bmatrix} \\ +\rho_{oil}\left(Q_{th} - Q_{inleak_p} - Q_{outleak_p} - Q_{inleak_{cy}}\right) \begin{bmatrix} c_{pump}\left(T_{valveoutlet} - T_{inlet}\right) \\ +\left(1 - \dfrac{\alpha_V\left(T_{valveoutlet} + T_{inlet}\right)}{2}\right) \bar{v}\left(p_{lrvalveoutlet} - p_{inlet}\right) \end{bmatrix} \\ +\rho_{oil} Q_{acc} \begin{bmatrix} c_{pump}\left(T_{acc} - T_{inlet}\right) \\ +\left(1 - \dfrac{\alpha_V\left(T_{inlet} + T_{acc}\right)}{2}\right) \bar{v}\left(p_{acc} - p_{inlet}\right) \end{bmatrix} \\ +m_{inlet} T_{inlet} \alpha_V \nu \dfrac{dp_{inlet}}{d\tau} \end{Bmatrix} \tag{14}$$

$$\frac{dT_{outlet}}{d\tau} = \frac{1}{c_{pump} m_{outlet}} *$$

$$\begin{Bmatrix} \rho_{oil} Q_{th} \begin{bmatrix} c_{pump}(T_{inlet} - T_{outlet}) \\ +\left(1 - \frac{\alpha_V (T_{outlet} + T_{inlet})}{2}\right) \bar{v}(p_{inlet} - p_{outlet}) \end{bmatrix} \\ +P_{mhloss} + m_{outlet} T_{outlet} \alpha_V v \frac{dp_{outlet}}{d\tau} \end{Bmatrix} \tag{15}$$

$$\frac{dT_{leak}}{d\tau} = \frac{1}{c_{pump} m_{leak}} *$$

$$\begin{Bmatrix} \rho_{oil} Q_{exleak_p} * \\ \begin{bmatrix} c_{pump}(T_{outlet} - T_{leak}) \\ +\left(1 - \frac{\alpha_V (T_{outlet} + T_{leak})}{2}\right) \bar{v}(p_{outlet} - p_{leak}) \end{bmatrix} \\ +P_{chnloss} - h_{pumpshell} A_{pumpshell} (T_{leak} - T_{pumpshell}) \\ +m_{leak} T_{leak} \alpha_V v \frac{dp_{leak}}{d\tau} \end{Bmatrix} \tag{16}$$

$$\frac{dT_{pumpshell}}{d\tau} = \frac{1}{c_{pumpshell} m_{pumpshell}} *$$

$$\begin{bmatrix} h_{pumpshell} A_{pumpshell} (T_{leak} - T_{pumpshell}) \\ -h_{accpump} A_{accpump} (T_{pumpshell} - T_{acc}) \end{bmatrix} \tag{17}$$

The accumulator (8) contains 8-interior node and 8-shell node as the following eq. (18) and eq. (19). There is a noteworthy change in the equation for one node as the mass of flow entering the accumulator might not be equal to that leaving the accumulator, which means that eq. (2) should be taken into account.

$$\frac{dT_{acc}}{d\tau} = \frac{1}{c_{acc} m_{acc}} *$$

$$\begin{Bmatrix} \rho_{oil} Q_{exleak_p} \begin{bmatrix} c_{acc}(T_{leak} - T_{acc}) \\ +\left(1 - \frac{\alpha_V (T_{leak} + T_{acc})}{2}\right) \bar{v}(p_{leak} - p_{acc}) \end{bmatrix} \\ -\rho_{oil} Q_{acc} \left(\begin{matrix} c_{acc}(T_{inlet} - T_{acc}) \\ +\left(1 - \frac{\alpha_V (T_{inlet} + T_{acc})}{2}\right) \bar{v}(p_{inlet} - p_{acc}) \end{matrix} \right) \\ -h_{accshell} A_{accshell} (T_{acc} - T_{accshell}) \\ +h_{accpump} A_{accpump} (T_{pumpshell} - T_{acc}) \\ +m_{acc} T_{acc} \alpha_V v \frac{dp_{acc}}{d\tau} \end{Bmatrix} \tag{18}$$

$$\frac{dT_{accshell}}{d\tau} = \frac{1}{c_{accshell} m_{accshell}} *$$

$$\begin{bmatrix} h_{accshell} A_{accshell} (T_{acc} - T_{accshell}) \\ -h_{accamt} A_{accamt} (T_{accshell} - T_{amt}) \\ -\varepsilon A_{accamt} \sigma (T_{accshell}^4 - T_{amt}^4) \end{bmatrix} \tag{19}$$

Expressions of temperature differential equations on the basis of heat transfer in fig. 3 are:

The symmetric cylinder (1), solenoid valve (2-1), (2-2) and PMSM (5) in fig. 1 have the same equations as eq. (6-13). Therefore, equations that are different from those in typical mode are given. CBV (3-1) includes 3-1-inlet node, 3-1-outlet node and 3-1-shell node. CBV (3-2) contains 3-2-inlet node, 3-2-outlet node and 3-2-shell node, following eq. (20-22):

$$\frac{dT_{cbvinlet}}{d\tau} = \frac{1}{c_{cbv} m_{cbvinlet}} *$$

$$\begin{Bmatrix} \rho_{oil} \left(Q_{th} - Q_{inleak_p} - Q_{exleak_p} - Q_{inleak_{cy}}\right) * \\ \begin{bmatrix} c_{cbv}(T_{cylinderoutlet} - T_{cbvinlet}) \\ +\left(1 - \frac{\alpha_V (T_{cylinderoutlet} + T_{cbvinlet})}{2}\right) \\ \bar{v}(p_{cylinderoutlet} - p_{cbvinlet}) \end{bmatrix} \\ -h_{cbvshell} A_{cbvshell} (T_{cbvinlet} - T_{cbvshell}) \\ +m_{cbvinlet} T_{cbvinlet} \alpha_V v \frac{dp_{cbvinlet}}{d\tau} \end{Bmatrix} \tag{20}$$

$$\frac{dT_{cbvoutlet}}{d\tau} = \frac{1}{c_{cbv} m_{cbvoutlet}} *$$

$$\begin{Bmatrix} \rho_{oil} \left(Q_{th} - Q_{inleak_p} - Q_{exleak_p} - Q_{inleak_{cy}}\right) * \\ \begin{bmatrix} c_{cbv}(T_{cbvinlet} - T_{cbvoutlet}) + \\ \left(1 - \frac{\alpha_V (T_{cylinderoutlet} + T_{cbvinlet})}{2}\right) \\ \bar{v}(p_{cbvinlet} - p_{cbvoutlet}) \end{bmatrix} \\ -h_{cbvshell} A_{cbvshell} (T_{cbvoutlet} - T_{cbvshell}) \\ +P_{minorloss} + m_{cbvoutlet} T_{cbvoutlet} \alpha_V v \frac{dp_{cbvoutlet}}{d\tau} \end{Bmatrix} \tag{21}$$

$$\frac{dT_{cbvshell}}{d\tau} = \frac{1}{c_{cbvshell} m_{cbvshell}} *$$

$$\begin{bmatrix} h_{cbvshell} A_{cbvshell} (T_{cbvinlet} - T_{cbvshell}) \\ +h_{cbvshell} A_{cbvshell} (T_{cbvoutlet} - T_{cbvshell}) \\ -h_{cbvamt} A_{cbvamt} (T_{cbvshell} - T_{amt}) \end{bmatrix} \tag{22}$$

The axial piston pump/motor (6) is as the following eq. (23-26):

$$\frac{dT_{inlet}}{d\tau} = \frac{1}{c_{pump} m_{inlet}} *$$

$$\begin{bmatrix} \rho_{oil} Q_{inleak_p} * \\ c_{pump}(T_{outlet} - T_{inlet}) \\ \left(+ \left(1 - \frac{\alpha_V (T_{outlet} + T_{inlet})}{2} \right) \bar{v} (p_{outlet} - p_{inlet}) \right) \\ + \rho_{oil} Q_{acc} \left(\begin{matrix} c_{pump}(T_{acc} - T_{inlet}) \\ + \left(1 - \frac{\alpha_V (T_{acc} + T_{inlet})}{2} \right) \bar{v} (p_{acc} - p_{inlet}) \end{matrix} \right) \\ + m_{inlet} T_{inlet} \alpha_V v \frac{dp_{inlet}}{d\tau} \end{bmatrix} \tag{23}$$

$$\frac{dT_{outlet}}{d\tau} = \frac{1}{c_{pump} m_{outlet}} *$$

$$\begin{bmatrix} \rho_{oil} Q_{th} \left(\begin{matrix} c_{pump}(T_{inlet} - T_{outlet}) \\ + \left(1 - \frac{\alpha_V (T_{outlet} + T_{inlet})}{2} \right) \bar{v} (p_{inlet} - p_{outlet}) \end{matrix} \right) \\ + P_{mhloss} + m_{outlet} T_{outlet} \alpha_V v \frac{dp_{outlet}}{d\tau} \end{bmatrix} \tag{24}$$

$$\frac{dT_{leak}}{d\tau} = \frac{1}{c_{pump} m_{leak}} *$$

$$\begin{bmatrix} \rho_{oil} Q_{exleak_p} * \\ c_{pump}(T_{outlet} - T_{leak}) \\ \left(+ \left(1 - \frac{\alpha_V (T_{outlet} + T_{leak})}{2} \right) \bar{v} (p_{outlet} - p_{leak}) \right) \\ + P_{chnloss} - h_{pumpshell} A_{pumpshell} (T_{leak} - T_{pumpshell}) \\ + m_{leak} T_{leak} \alpha_V v \frac{dp_{leak}}{d\tau} \end{bmatrix} \tag{25}$$

$$\frac{dT_{pumpshell}}{d\tau} = \frac{1}{c_{pumpshell} m_{pumpshell}} *$$

$$\begin{bmatrix} h_{pumpshell} A_{pumpshell} (T_{leak} - T_{pumpshell}) \\ - h_{accpump} A_{accpump} (T_{pumpshell} - T_{acc}) \end{bmatrix} \tag{26}$$

Equation (27-28) about the accumulator (8) are expressed as:

$$\frac{dT_{acc}}{d\tau} = \frac{1}{c_{acc} m_{acc}} *$$

$$\begin{bmatrix} \rho_{oil} \left(Q_{th} - Q_{inleak_p} - Q_{exleak_p} - Q_{inleak_{cy}} \right) \\ c_{acc}(T_{cbvoutlet} - T_{acc}) \\ \left(+ \left(1 - \frac{\alpha_V (T_{cbvoutlet} + T_{acc})}{2} \right) \bar{v} (p_{cbvoutlet} - p_{acc}) \right) \\ + \rho_{oil} Q_{exleak_p} \left(\begin{matrix} c_{acc}(T_{leak} - T_{acc}) \\ + \left(1 - \frac{\alpha_V (T_{leak} + T_{acc})}{2} \right) \bar{v} (p_{leak} - p_{acc}) \end{matrix} \right) \\ - \rho_{oil} Q_{acc} \left(\begin{matrix} c_{acc}(T_{inlet} - T_{acc}) \\ + \left(1 - \frac{\alpha_V (T_{inlet} + T_{acc})}{2} \right) \bar{v} (p_{inlet} - p_{acc}) \end{matrix} \right) \\ - h_{accshell} A_{accshell} (T_{acc} - T_{accshell}) \\ + h_{accpump} A_{accpump} (T_{pumpshell} - T_{acc}) \\ + m_{acc} T_{acc} \alpha_V v \frac{dp_{acc}}{d\tau} \end{bmatrix} \tag{27}$$

$$\frac{dT_{accshell}}{d\tau} = \frac{1}{c_{accshell} m_{accshell}} *$$

$$\begin{bmatrix} h_{accshell} A_{accshell} (T_{acc} - T_{accshell}) \\ - h_{accamt} A_{accamt} (T_{accshell} - T_{amt}) \\ - \varepsilon A_{accamt} \sigma (T_{accshell}^4 - T_{amt}^4) \end{bmatrix} \tag{28}$$

4 Simulation results

According to fig. 5, dynamic performance of the EHA is simulated through AMESim as one part of the co-simulation, and PID controller is deployed. In order to estimate thermal behavior, two thermal models are developed in MATLAB/Simulink for each operating mode.

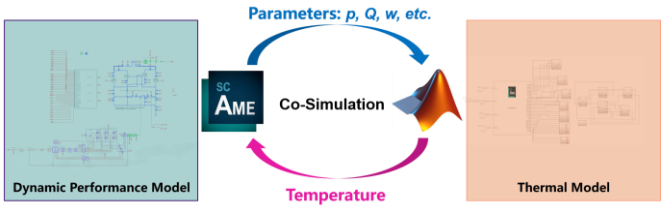


Figure 5: Data exchange within co-simulation

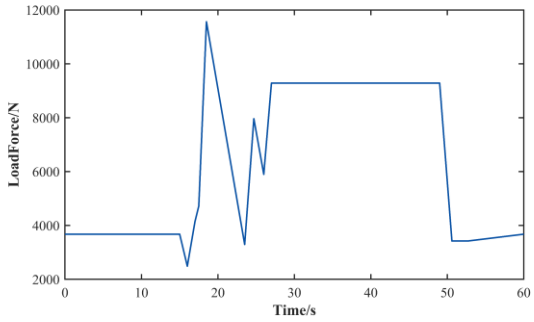
The EHA achieves rod extension and retraction by driving the pump in different rotational directions. Consequently, although the mathematical temperature differential equations are derived, it is crucial that the Simulink model accurately accounts for directional parameters when data is transferred from AMESim.

4.1 Duty cycle

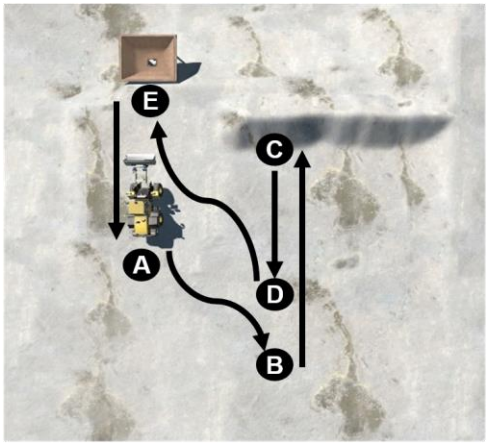
Since this EHA is designed for lifting applications, it consistently sustains force in a single direction, which means the EHA operates in two quadrants. As shown in fig. 6(a), the load force is obtained from Mevea-based simulations[22] and subsequently scaled down and linearized to align with the specifications of the integrated EHA.

Figure. 6(b) illustrates the trajectory of a hybrid wheel loader completing a single and 60-second duty cycle, as viewed from Mevea user interface. This duty cycle represents a typical loading and unloading process. The sequence begins with the wheel loader moving from point A to point B without a load. Upon reaching point C, the loader lowers its bucket, tilts it, and raises the boom to collect material into the bucket. After loading, the loader reverses to point D, then drives forward to point E to unload the material. The cycle is completed when the loader returns to the starting position at point A.

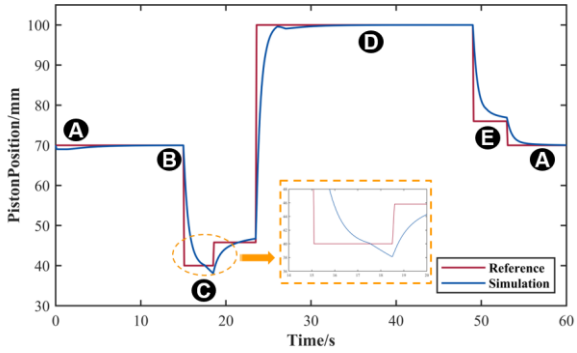
The stroke of the cylinder is also scaled down. Even though models of dual operating modes in AMESim employ three-loop PID controllers, they have varied control parameters and effects due to different working hydraulic components. Figure. 6(c) and (d) presents a comparison of piston position dynamics as a function of time between a reference command and a simulated response. The most notable discrepancy occurs around point C, where the piston continues to retract in fig. 6(c) because of a sharp increase in load force, as depicted in fig. 6(a). This sudden load boom exceeds the hydraulic system's damping to counteract the impact in real-time, resulting in a pronounced deviation in the response. Nevertheless, CBVs play an important role in resisting this impact in safety mode, which can be proved by the curve around point C in fig. 6(d).



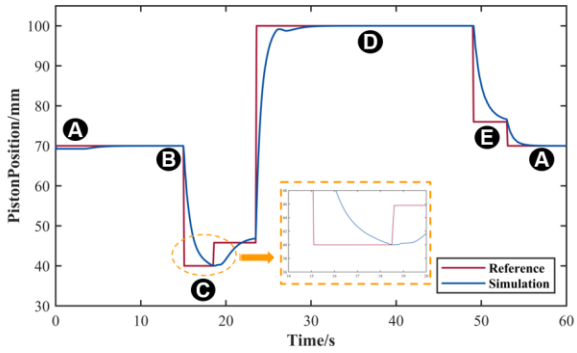
(a) Load in one duty cycle



(b) Trajectory of the hybrid wheel loader



(c) Comparison between reference and simulation results under typical mode



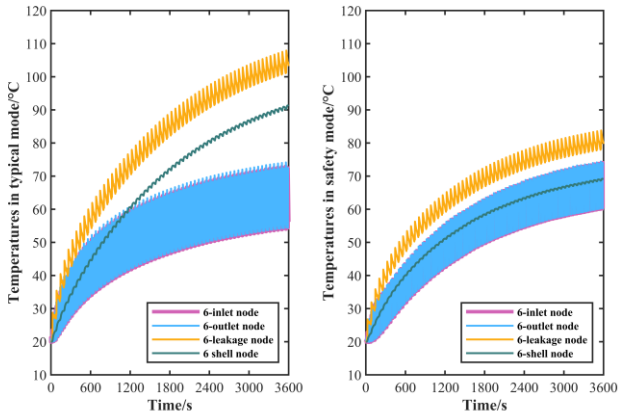
(d) Comparison between reference and simulation results under safety mode

Figure 6: Duty cycle of the hybrid wheel loader under two modes

Initial temperatures of all components are set to 20.00 ℃ (293.15 K), matching the ambient temperature, which remains constant at 20.00 ℃ throughout the process. To intuitively observe the temperature trend, a co-simulation was conducted with a 60-second duty cycle and repeated a total duration of 3600 seconds.

4.2 Temperature results

Temperature in 6-inlet node under safety mode is a little larger than that under typical mode. The temperature of 6-outlet node in safety mode in fig. 7(b) closely follow that in typical mode in fig. 7(a). On the contrary, the temperatures of 6-leakage node and 6-shell node in typical mode are significantly higher than those in safety mode, with the difference becoming more pronounced as time progresses. A critical observation from fig. 7 is that in typical mode, the temperature in 6-shell node curve gradually rises above the 6-outlet node temperature after approximately 1500 seconds. Nonetheless, in safety mode, the 6-shell node temperature consistently remains lower than the 6-outlet node temperature throughout the entire duration.



(a) typical mode *(b) safety mode*

Figure 7: Temperatures in the pump under dual modes

As mentioned in Chapters 2 and 4, the inlet and outlet ports of a pump switch roles as the flow direction changes. This presents two possible approaches for implementation in MATLAB/Simulink. The first approach defines the two pump ports as Port A and Port B, where each port can function as either an inlet port or an outlet port. Consequently, each port requires two distinct states (represented by two separate equations, such as eq. (14) and eq. (15)) that must switch accordingly. However, this method is challenging to implement in software.

The second approach assigns that the inlet port is either Port A or Port B, requiring only parameter switching within a single equation while inheriting temperature from the previous state. This study adopts the second approach, though it results in significant temperature oscillations during switching.

In spite of different modes, temperatures of 5-interior node and 5-shell node are quite similar respectively as fig. 8 shows.

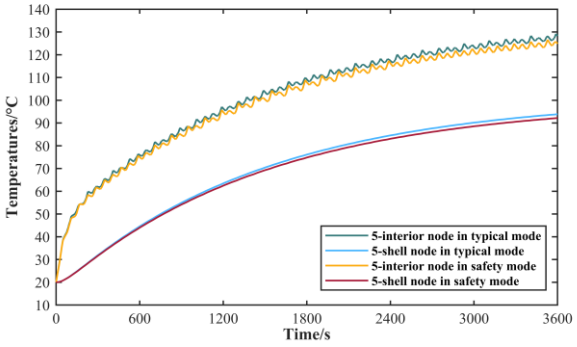


Figure 8: Temperatures in the PMSM under dual modes

Figure. 9 clearly demonstrates that the temperatures of the three CBV (3-2) nodes—the inlet, outlet, and shell nodes—are varied, with 3-2-outlet node temperature being the highest. Conversely, the temperatures of the 3-1-inlet and 3-1-outlet nodes remain nearly identical, exhibiting no distinguishable difference throughout the observed period.

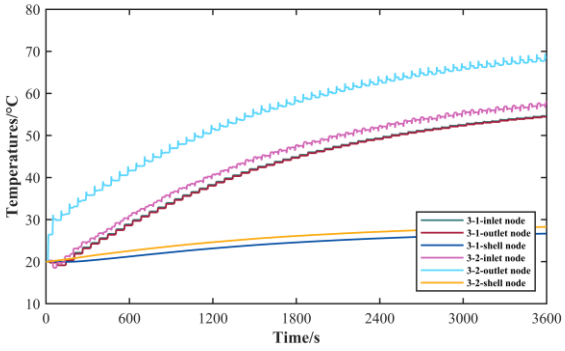


Figure 9: Temperatures in the CBV under safety mode

Table. 1 shows temperatures of all nodes at 3600 seconds for typical mode and safety mode. It is evident that the cylinder (1), solenoid valve (2-1) and (2-2), and exhibit higher temperatures in safety mode, compared to typical mode. In contrast, temperatures at the 8-interior node and 8-shell node are lower in safety mode than those in typical mode.

Table 1 Final Temperatures at the 3600 seconds

Typical mode		**Safety mode**	
Node Name	**Temperature**	**Node Name**	**Temperature**
1-inlet	42.99 ℃	1-inlet	68.29 ℃
1-outlet	37.97 ℃	1-outlet	54.79 ℃
1-shell	36.22 ℃	1-shell	49.08 ℃
2-1-inlet	42.01 ℃	2-1-inlet	72.16 ℃
2-1-outlet	50.56 ℃	2-1-outlet	73.17 ℃
2-1-shell	26.12 ℃	2-1-shell	30.33 ℃
2-2 inlet	46.29 ℃	2-2 inlet	73.08 ℃
2-2-outlet	55.89 ℃	2-2-outlet	73.25 ℃
2-2-shell	27.24 ℃	2-2-shell	30.26 ℃
5-interior	128.26 ℃	5-interior	124.85 ℃
5-shell	93.88 ℃	5-shell	92.20 ℃
6-inlet	54.41 ℃	6-inlet	60.24 ℃
6-outlet	72.44 ℃	6-outlet	74.01 ℃
6-leakage	105.82 ℃	6-leakage	82.06 ℃
6-shell	91.24 ℃	6-shell	69.18 ℃
8-interior	81.13 ℃	8-interior	60.64 ℃
8-shell	72.05 ℃	8-shell	54.35 ℃
		3-1-inlet	74.35 ℃
		3-1-outlet	74.67 ℃
		3-1-shell	30.35 ℃
		3-2-inlet	57.25 ℃
		3-2-outlet	67.70 ℃
		3-2-shell	28.25 ℃

5 Analysis and discussion

There are two possible reasons to explain that the difference in PMSM temperatures between the modes appears surprisingly low: 1. Imprecise simulating values. We haven't done experiments to correct a few of coefficients; 2. In typical mode, PMSM must work in the situation of high torque and low speed when the piston reaches the desired position. In this case, PMSM keeps rotating and generating heat. In safety mode, motor needs more power to drive pump and then open CBVs.

The higher temperatures observed in the cylinder during safety mode can be attributed to the behavior of the CBVs. In safety mode, CBVs prevent fluid flow until the pressure inside the cylinder's chamber reaches a sufficient level to open the valve. On the contrary, in typical mode, fluid flows through the solenoid valve without any restriction. When the piston reaches necessary position, CBVs can assist in maintaining the load without the pump working with small flow rate and high torque. This is the reason why the temperature of pump leakage node in safety mode is lower than in typical mode in fig. 7. Additionally, since the pump is submerged in oil within the accumulator, the pump shell struggles to dissipate heat effectively to the ambient air as the oil temperature increases. However, in safety mode, more fluid enters the accumulator through the CBVs, enhancing convective heat transfer. This results in lower temperature in the pump shell node in safety mode in comparison to the outlet node. Similarly, this enhanced heat transfer explains why the temperatures of accumulator's two nodes are lower in safety mode than in typical mode.

Furthermore, this EHA used for lifting experiences a continuous pushing force, even when the bucket is unloaded, leading to higher pressure in one chamber of the cylinder. The pilot pressure required to open the CBV closely matches the lower pressure in the other chamber. These factors result in discontinuous fluid flow through one of the CBVs when the pilot pressure is insufficiently high. Consequently, varying temperatures are observed at the CBV's inlet and outlet, as illustrated in fig. 9 via curves of temperature rise of the 3-2-inlet node and 3-2-outlet node.

These results indicate a difference in thermal behavior between the two modes, highlighting the influence of mode-specific conditions on system temperature distribution.

6 Conclusion

This study has demonstrated the viability of a dual-mode integrated EHA as a direct replacement for the conventional lift cylinder in hybrid wheel loaders, fully satisfying ISO 8643 safety requirements. Lumped parameter thermal models for both open and closed hydraulic circuits were developed and validated via co-simulation in AMESim and MATLAB Simulink over a representative 3,600 s duty cycle.

Results indicate that engaging the CBVs under safety mode markedly enhances load-holding capacity and enables reliable two-quadrant operation under variable loads. However, intermittent fluid flow in this mode elevates cylinder and solenoid valve temperatures, increasing localized thermal stress, even as convective heat transfer in the pump and accumulator is improved.

In contrast, EHA under typical mode (CBVs disengaged) maintains continuous fluid circulation, which keeps cylinder and valve temperatures lower but shifts the bulk of thermal load to the pump leakage node and the accumulator. Notably, permanent-magnet synchronous motor temperatures remained statistically unchanged between both modes, highlighting the motor's inherent thermal robustness.

Based on these findings, we recommend targeted cooling strategies for each mode: focused heat dissipation for the cylinder and valves during safety-mode operation, and enhanced pump-leakage and accumulator heat rejection in typical mode. Overall, the proposed modeling and simulation framework offers a clear pathway for optimizing component-level thermal management in electrified hydraulic systems, paving the way for future work on transient control integration and advanced heat-exchanger design to further boost efficiency and reliability in non-road mobile machinery.

Future work: The primary objective is to conduct experiments to refine parameters and validate the models. The testrig is being built and is equipped with a real-time simulation machine for Hardware-in-the-Loop (HIL) simulation.

Acknowledgements

This research is supported by international joint doctoral education fund of Beihang University. Also, thanks for the help from Prof. Tatiana Minav.

Nomenclature

Designation	Denotation	Unit
$\dot{Q}$	Rate of heat transfer	W
$\dot{W}$	Rate of work	W
$\dot{m}$	Mass flow rate	kg/s
u	Specific internal energy	J/kg
v	Specific volume	m^3/kg
$\bar{v}$	Average specific volume	m^3/kg
p	Fluid pressure	Pa
ρ	Fluid density	kg/m^3
ε	Emissivity	1
σ	Stefan-Boltzmann constant	$W/(m^2 \cdot K^4)$
λ	Thermal conductivity	W/(m·K)
δ	Thickness	m
k_e	Specific kinetic energy	J/kg
p_e	Specific potential energy	J/kg
α_v	Coefficient of cubic expansion	1/K
E_{CM}	Control mass energy	J

Q	Volumetric flow rate	m^3/s
T	Temperature	K
τ	Time	second
P_{mloss}	Mechanical loss in cylinder	W
$P_{minorloss}$	Energy loss in valve	W
$P_{motorloss}$	Energy loss in Motor	W
P_{mhloss}	Mechanical loss in pump	W
$P_{chnloss}$	Churning loss in pump	W

References

[1] R. Hagan *et al.*, 'Non-Road Mobile Machinery Emissions and Regulations: A Review', *Air*, vol. 1, no. 1, Art. no. 1, Mar. 2023, doi: 10.3390/air1010002.

[2] B. Singh, 'Novel and Ruggedized Power Electronics for Off-Highway Vehicles', *IEEE Electrification Magazine*, vol. 2, no. 2, pp. 31–41, Jun. 2014, doi: 10.1109/MELE.2014.2313157.

[3] M. de F. Ramos, D. A. de L. Brandao, D. P. V. Galo, B. de J. Cardoso Filho, I. A. Pires, and T. A. C. Maia, 'A Study on the Performance of the Electrification of Hydraulic Implements in a Compact Non-Road Mobile Machine: A Case Applied to a Backhoe Loader', *World Electric Vehicle Journal*, vol. 15, no. 4, Art. no. 4, Apr. 2024, doi: 10.3390/wevj15040127.

[4] S. Hui and J. Junqing, 'Research on the system configuration and energy control strategy for parallel hydraulic hybrid loader', *Automation in Construction*, vol. 19, no. 2, pp. 213–220, Mar. 2010, doi: 10.1016/j.autcon.2009.10.006.

[5] M. Ochiai and S. Ryu, 'Hybrid in Construction Machinery', *Proceedings of the JFPS International Symposium on Fluid Power*, vol. 2008, no. 7–1, pp. 41–44, 2008, doi: 10.5739/isfp.2008.41.

[6] N. Zou *et al.*, 'Series-parallel hybrid power loader', CN102653228A, Sep. 05, 2012 Accessed: May 16, 2024. [Online]. Available: https://patents.google.com/patent/CN102653228A/en

[7] M. Axin, *Fluid power systems for mobile applications: with a focus on energy efficiency and dynamic characteristics*. Linköping: Linköping University, Department of Management and Engineering, 2013.

[8] S. Ketelsen, D. Padovani, T. O. Andersen, M. K. Ebbesen, and L. Schmidt, 'Classification and Review of Pump-Controlled Differential Cylinder Drives', *Energies*, vol. 12, no. 7, Art. no. 7, Jan. 2019, doi: 10.3390/en12071293.

[9] S. H. Cho and R. Burton, 'Position control of high performance hydrostatic actuation system using a simple adaptive control (SAC) method', *Mechatronics*, vol. 21, no. 1, pp. 109–115, Feb. 2011, doi: 10.1016/j.mechatronics.2010.09.003.

[10] Y. Shang *et al.*, 'A Novel Electro Hydrostatic Actuator System With Energy Recovery Module for More Electric Aircraft', *IEEE Transactions on Industrial Electronics*, vol. 67, no. 4, pp. 2991–2999, Apr. 2020, doi: 10.1109/TIE.2019.2905834.

[11] Z. Lin, Z. Lin, F. Wang, and B. Xu, 'A series electric hybrid wheel loader powertrain with independent electric load-sensing system', *Energy*, vol. 286, p. 129497, Jan. 2024, doi: 10.1016/j.energy.2023.129497.

[12] Y. Qiao, J. Wu, F. Wang, and B. Xu, 'Electro-Hydrostatic Actuator System With Separate Design to Improve Efficiency for Working Function of Electric Wheel Loader', in *BATH/ASME 2024 Symposium on Fluid Power and Motion Control*, Bath, United Kingdom: American Society of Mechanical Engineers, Sep. 2024, p. V001T01A033. doi: 10.1115/FPMC2024-141117.

[13] D. Fassbender and T. Minav, 'An Algorithm for the Broad Evaluation of Potential Matches between Actuator Concepts and Heavy-Duty Mobile Applications', *Actuators*, vol. 10, no. 6, Art. no. 6, Jun. 2021, doi: 10.3390/act10060111.

[14] ISO Standard No.8643: 2017.

[15] J. Andersson, P. Krus, K. Nilsson, and K. Storck, 'MODELLING AND SIMULATION OF HEAT GENERATION IN ELECTRO-HYDROSTATIC ACTUATION SYSTEMS', *Proceedings of the JFPS International Symposium on Fluid Power*, vol. 1999, no. 4, pp. 537–542, 1999, doi: 10.5739/isfp.1999.537.

[16] S. Michel, 'Prediction of the thermo-energetic behaviour of an electrohydraulic compact drive', Mar. 2016.

[17] S. Fan *et al.*, 'Thermal analysis of key components in EHA system under high temperature condition', in *CSAA/IET International Conference on Aircraft Utility Systems (AUS 2022)*, Aug. 2022, pp. 1179–1184. doi: 10.1049/icp.2022.1837.

[18] T. Minav, L. Papini, and M. Pietola, 'A Thermal analysis of Direct Driven Hydraulics', Mar. 2016.

[19] T. Minav, J. Heikkinen, T. Schimmel, and M. Pietola, 'Direct Driven Hydraulic Drive: Effect of Oil on Efficiency in Sub-Zero Conditions', *Energies*, vol. 12, no. 2, p. 219, Jan. 2019, doi: 10.3390/en12020219.

[20] X. Han, T. Minav, M. Wang, Y. Fu, and M. Pietola, 'Thermal Coupling Simulation of Electro-Hydrostatic Actuator Subjected to Critical Temperature Conditions', *TJFP*, Sep. 2022, doi: 10.13052/ijfp1439-9776.2336.

[21] J. A. Sidders, D. G. Tilley, and P. J. Chappie, 'Thermal-Hydraulic Performance Prediction in Fluid Power Systems', *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, vol. 210, no. 4, pp. 231–242, Nov. 1996, doi: 10.1243/PIME_PROC_1996_210_462_02.

[22] 'Training Simulators for Construction Equipment | Mevea Digital Twins', Mevea. Accessed: May 07, 2025. [Online]. Available: https://mevea.com/solutions/training-simulators/construction-equipment/

An electrohydraulic actuator for bipedal robots with active-passive combination

Junru Hui, Bo Yang, Haiyang Chu, Suyang Xiao, Yingwei Bao, Junhao Chen, Yang Chen, Xiaoping Ouyang* and Mingjie Li*

State Key Laboratory of Fluid Power & Mechatronic Systems, Zhejiang University, Hangzhou, China
E-mail: Junru Hui.junruhui@zju.edu.cn, Bo Yang.yang_bo@zju.edu.cn, Haiyang Chu.22225085@zju.edu.cn, Suyang Xiao.22260255@zju.edu.cn, Yingwei Bao.baoyingwei@zju.edu.cn, Junhao Chen.12425025@zju.edu.cn, Yang Chen.22325144@zju.edu.cn, Xiaoping Ouyang*.ouyangxp@zju.edu.cn, Mingjie Li*. mingjielee@zju.edu.cn
*State Key Laboratory of Fluid Power & Mechatronic Systems, Zhejiang University, Hangzhou, China

Abstract

Bipedel robots have great potential for rehabilitation and service due to their bionic morphology and adaptive mobility. These robots' systems require advanced knee actuation mechanisms, such as dynamic force modulation, precise control, and adaptive impedance properties, to support a variety of movement modes. This study describes an electrohydraulic knee actuator which can work in three modes: active, passive and burst model. It integrates three accumulators, a rotary direct-drive servo valve and a rotary switching valve together into a hydraulic cylinder. The core of the design is to achieve high precision control in active mode, variable stiffness and damping control in passive mode, force burst in burst mode, through coordinating the operation of valves and accumulators. This combination not only supports robot manoeuvres such as walking, jumping and landing by precisely controlling the displacement, stiffness, damping and force of the hydraulic actuators, but also reduces energy consumption effectively.

Keywords: integrated electro-hydraulic actuator, active and passive, rotary switch valve, rotary direct-drive servo valve, bipedel robot

1 Introduction

Bionic-legged robots, with their superior mobility and adaptability to the environment, outperform traditional wheeled or tracked robots in reducing operational blind spots through a gait control mechanism that utilizes discrete contact points. This makes them invaluable in various applications, including exploring hazardous environments, aiding in rehabilitation medicine, and providing intelligent services[1].

The joint driving methods for bipedal robots primarily include motor drive and hydraulic drive. Hydraulic drive offers higher power density and better load characteristics compared to the motor-reducer combination.

The legs of hydraulic legged robots are typically made up of multiple hydraulic actuators connected in series or parallel. The design, manufacturing, and control of these actuators are crucial in determining the leg's motion performance and load-bearing capacity. Hydraulic actuators usually employ linear motion servo cylinders, which excel in low-speed stability, hydraulic stiffness, minimal internal leakage, and are compact in weight and volume.

In the lower limb structure of bipedal robots, the knee joint actuator is pivotal. It assists the leg in flexion and extension movements during walking, supporting the robot's weight and enabling gait control. Moreover, the knee joint actuator must be capable of absorbing impacts when the foot contacts the ground.

Boston Dynamics unveiled the BigDog quadruped hydraulic robot in 2006, featuring a hydraulic actuator that integrates servo valves, servo cylinders, displacement sensors, and force sensors to enable closed-loop force/position control[2]. In 2008, the company upgraded the hydraulic actuator with the lighter MOOG E024 two-stage nozzle flapper electro-hydraulic servo valve and concurrently optimized the cylinder structure. These enhancements resulted in a more compact, smaller, and lighter hydraulic drive unit, thereby enhancing its overall compactness[3].

In China, research institutions such as HIT[4], SDU[5], and BIT[6] have delved into this field, developing hydraulic actuators with structures akin to BigDog's integrated valve-controlled system. These designs incorporate various components within the hydraulic drive unit's cylinder. Despite utilizing numerous small components, the level of integration remains moderate. The servo valve and hydraulic cylinder are

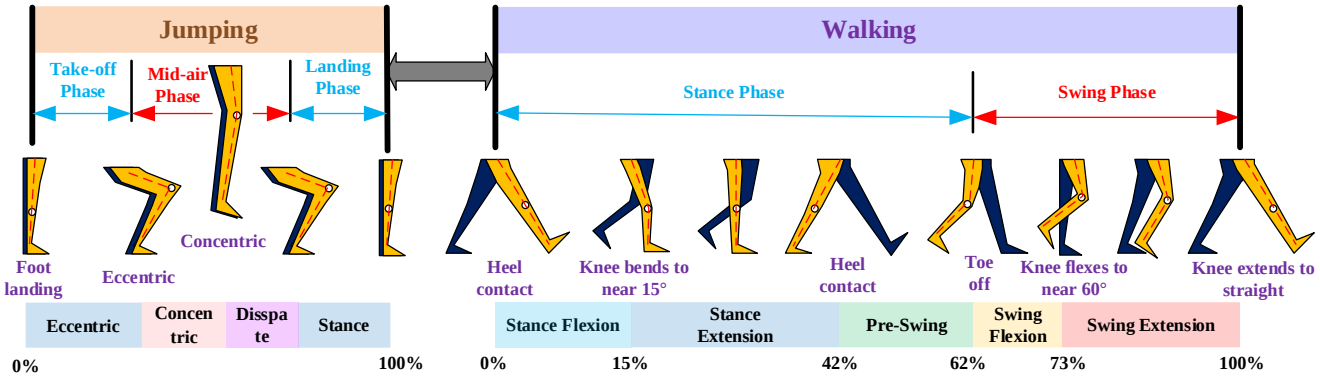


Figure 1 Human Vertical Jumping and Walking Gait Cycles

connected via a plate, which increases the actuator's overall volume. In contrast, the Baby Elephant, introduced by the SJTU team, features a distinct hydraulic actuator structure. It comprises a servo cylinder, a moving slide valve motor, and a pressure sensor, achieving a high degree of multifunctional integration within the hydraulic drive unit[7].

In 2018, MOOG and IIT joined forces to develop highly integrated Integrated Smart Actuators (ISA) using additive manufacturing. These actuators incorporate servo valve structures directly into the actuator bodies, eliminating the need for interfaces, and feature multifunctional built-in controllers for high-performance control[8]. This collaboration has provided valuable references and research insights for many scholars.

However, these are active actuators, which rely on control algorithms to softly absorb the buffering experienced by the actuator. This approach inevitably increases the overall power consumption and decreases the power-to-weight ratio of the entire actuator. Consequently, some researchers are exploring the introduction of passive regulation modes for actuators.

In 2011, researchers at the Polytechnic University of Madrid engineered a bionic agility leg with adjustable flexibility. This innovation merged tandem elasticity with magnetorheological dampers, mimicking the functionality of a horse's leg[9]. Moving forward to 2020, Shandong University revealed a hydraulic actuator capable of passive flexibility. This was achieved by integrating an accumulator within the piston chamber of a hydraulic cylinder, allowing the actuator to modify its stiffness passively upon pressurization, thereby cushioning the impact between the foot and the ground[10]. Separately, the research team from HIT developed the Series Elastic Actuator (SEA) [11] and the Variable Stiffness Actuator (VSA) [12]. These actuators feature an elastic element positioned between the actuator's output and the negative cut. By manipulating the deformation of this elastic element, they attain variable damping and stiffness, producing the necessary output force. Nonetheless, this design enhancement leads to a considerable increase in the actuator's size.

In the design of knee joint actuators, a detailed analysis of the natural human gait cycle was performed. This cycle consists of two main phases: the Stance Phase and the Swing Phase, which account for about 60% and 40% of the walking cycle, respectively, as illustrated in Figure 1.

Biomechanical analysis reveals that the knee joint plays five key roles during the gait cycle. During the Stance Flexion Phase, the knee joint slightly flexes (by approximately 10°-15°), absorbing and dissipating impact energy through elastic deformation. This effectively reduces ground reaction forces on the lower limb joints. In the Stance Extension Phase, the knee joint extends fully to the neutral position (0°), maintaining lower limb stability through bone support and muscle coordination. This ensures efficient mechanical energy transfer during the weight-bearing period.

In the Pre-Swing Phase, the knee joint rapidly flexes (by about 30°-40°), creating enough ground clearance for the lower limb's swing. During the Swing Flexion Phase, the knee flexion velocity decreases progressively through the coordinated control of opposing muscle groups. This restricts the posterior swing angle of the lower leg to a physiological range (about 60°-70°), maintaining gait coordination. Finally, in the Swing Extension Phase, the knee extension velocity shows a decelerating trend, achieving terminal deceleration through damping control. This guides the foot to the optimal landing position, preparing for the next gait cycle[13].

The analysis indicates that during the natural human gait cycle, the knee joint primarily acts as a support mechanism rather than facilitating swinging motions for the majority of the cycle. Additionally, an examination of the human vertical jumping cycle reveals that it consists of three main phases: Take-off, Mid-air, and Landing. The knee joint performs five essential functions for the lower limbs: eccentric knee flexion for energy storage; concentric rapid extension from flexion to release energy and generate force output; and subsequent knee flexion for impact absorption. The analysis shows that during the natural vertical jumping cycle, the knee joint plays crucial roles in energy storage, rapid force generation, and efficient impact absorption [14].

To enhance the compactness and energy efficiency of electro-hydraulic actuators for hydraulic bipedal robots, this study introduces a novel, highly integrated electro-hydraulic actuator that combines active and passive systems, drawing inspiration from the natural human walking gait cycle. The proposed actuator allows for seamless switching between active and passive modes, providing high load capacity with

adjustable stiffness and continuously variable damping. This innovative actuator, incorporating a uniquely designed hydraulic system and novel core components, mimics the functional characteristics of the human knee joint during movement. It combines the versatility of active systems with the energy-saving benefits of passive systems.

The paper's structure is outlined as follows: Section 2 presents a comprehensive overview of the target platform and the specific performance requirements for the actuator. Section 3 provides a detailed explanation of the hydraulic principles underpinning the active-passive electro-hydraulic actuator, including the design of its essential components, the integration of the active and passive systems, and the validation of its performance metrics. In Section 4, a nuanced discussion of the results is offered, followed by concise concluding remarks in Section 5.

2 Overall design and requirements

The currently developed lower limb system for bipedal robots is capable of performing various movements, including walking, squatting, and running. The knee joint is actuated by a linkage mechanism forming triangle ΔABC, with the geometric relationship between the knee joint angle and actuator length illustrated in the left-side of Figure 2.

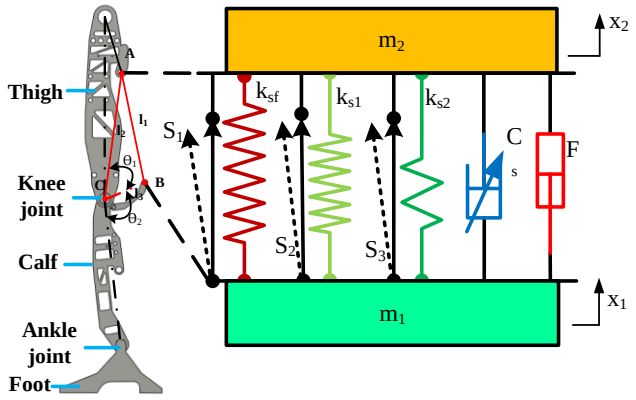


Figure 2 Kinetic Model of "IKPSA"

To meet the structural, load-bearing, and gait cycle requirements of the entire bipedal robot system, the actuator must adhere to the following specifications:

1. The actuator's structure should be compact to fit within the limited installation space in the robot's lower limb, with a stroke of 55mm.
2. Based on the experimental data from the current lower limb part of the robot, the actuator must have a maximum output force exceeding 10,000N at a rated pressure of 21MPa. It should also be capable of operating in both active and passive modes.
3. The actuator should feature high integration and plug-and-play capabilities in both its active and passive modes.

In line with the high integration design objectives and characteristics of the actuator, it has been named the IntelliKnee-Plug Servo Hydraulic Actuator, or "IKPSA" for short.

In response to the system-level requirements, a dynamic model of the actuator has been developed.

$$m_1\ddot{x}_1 + k_s(x_1 - x_2) - F + c(\dot{x}_1 - \dot{x}_2) = 0 \quad (1)$$

$$m_2\ddot{x}_2 + k_s(x_2 - x_1) + F + c(\dot{x}_2 - \dot{x}_1) = 0 \quad (2)$$

The actuator's output force is primarily determined by the fluid pressure and the piston's effective area.

$$F = p_1A_1 - p_2A_2 - F_{\text{visc}} \quad (3)$$

The stiffness adjustment capability of the IKPSA in passive mode, along with its shock absorption and energy source functions in burst mode, are all derived from the specially designed metal bellows accumulator integrated within the IKPSA. This component serves dual purposes as both an energy source and a shock absorption mechanism. To quantify the dynamic behavior of the accumulator, a pressure-flow relationship is based on gas thermodynamics. Assuming adiabatic conditions for rapid actuation scenarios, the pressure change rate in the accumulator can be expressed as

$$\dot{P}_i = \frac{\gamma Q_i P_i}{V_{acc}} \left(\frac{P_i}{P_{pc}}\right)^{1/\gamma} \quad (4)$$

Furthermore, the effective stiffness of the system, governing shock absorption performance, is derived from the linearized force-displacement relationship:

$$k_s = \frac{\gamma A_c^2 P_{pc}}{V_{acc}} \quad (5)$$

3 Actuator Design

3.1 Hydraulic Principle

"IKPSA" is a hybrid hydraulic actuator combining active and passive elements, which constitutes an independent closed-loop hydraulic system. As shown in Figure 3, the major components of IKPSA are a hydraulic cylinder, two relief valves, two 3-way 3-position valves, four switch valves, and three accumulators. It is important to highlight that all four switch valves depicted in the figure are housed within a single Rotary Switch Valve and rotate simultaneously.

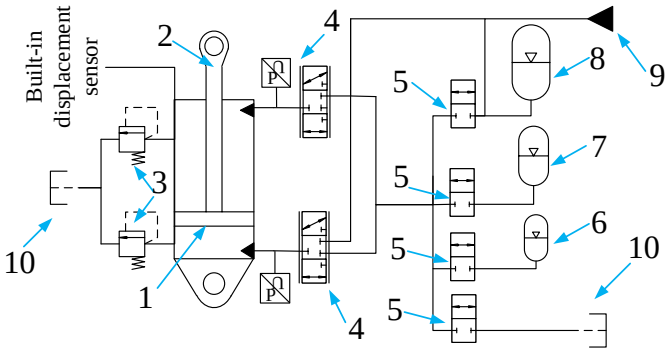


Figure 3 Hydraulic principle of IKPSA(1. Piston, 2. Rod, 3. Relief valve, 4. three -way three -position valve, 5. Switch valve, 6. Accumulator A, 7 Accumulator B, 8 Accumulator C,9 Oil inlet/ constant pressure, 10. Tank.)

The hydraulic system employs an asymmetric cylinder design, with both ends connected to the biped robot's knee joint hinges. Figure 4 illustrates key integration enhancements implemented in this configuration. High-pressure hydraulic fluid (denoted by red arrow H) enters the system through two Rotary Direct-Drive Servo Valves before being distributed to the actuator. These valves feature working ports connected to individual chamber ends of the hydraulic cylinder, enabling independent control of both load chambers. The servo valves' opposite ends connect to a Rotary Switch Valve that coordinates low-pressure fluid distribution (marked by blue L) to the accumulator array.

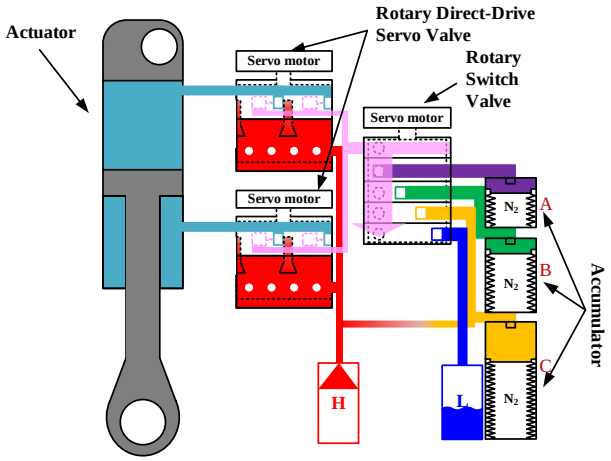


Figure 4 IKPSA schematic diagram

In Active Mode, the Rotary Switch Valve (RSV) maintains direct hydraulic connectivity to the low-pressure port via servo-driven spool alignment, bypassing all accumulators. The dual Rotary Direct-Drive Servo Valves (RDDSVs) function as three-position four-way valves, enabling independent pressure regulation in the cylinder's dual chambers. This configuration provides precise trajectory control of knee flexion-extension during the swing phase.

In Passive Mode, terrain adaptation is achieved by reconfiguring the RSV to allow fluid flow through selectable metal bellows accumulators (A, B), enabling adjustable stiffness of the actuator. The RDDSVs transition to variable-orifice dampers, where spool rotation modulates the overlap area between T-ports and working chambers, enabling real-time damping adjustment. This dual-parameter tunability enhances stability on irregular surfaces.

In addition to "Active Mode" and "Passive Mode," the IKPSA features a "Burst Mode," specifically designed to emulate the knee joint's biomechanical behavior during human jumping motions, thereby enabling hydraulic bipedal robots to perform jumping maneuvers.

Prior to engaging "Burst Mode," accumulator C is pre-charged through the high-pressure port to store energy. Upon activation, the pre-charged accumulator connects directly to the cylinder through fully opened RDDSV T-ports, bypassing servo control, implementing a differential drive principle. Consequently, the piston rod achieves both high force output and rapid extension. This mechanism enables the knee joint to perform explosive extension during jumping motions, providing essential support for the robot's overall jumping capability.

In summary, the IKPSA achieves seamless switching between "Active Mode," "Passive Mode," and "Burst Mode" through the coordinated operation of three rotary control valves and accumulators. Through the design of specialized fundamental components and innovative hydraulic principles, the IKPSA is capable of replicating human knee joint motion states, including stiffness adaptation for various conditions, continuously variable damping adjustment, gait modulation, and high-power burst capabilities. Moreover, the hydraulically-driven IKPSA possesses significantly greater load-bearing capacity compared to both human joints and existing electric motor-based joints. Complemented by its integrated housing design, the IKPSA offers plug-and-play functionality, facilitating easy installation and replacement.

3.2 Rotary Direct-Drive Servo Valve

The knee joint actuator's operational phases involve transitions between active and passive modes, along with motion damping control. To address these requirements, this study developed a Rotary Direct-Drive Servo Valve (RDDSV), a three-position three-way spool rotary direct-drive servo valve characterized by its simple structure and superior performance. Employing a cartridge-style design for integration, the RDDSV's structure primarily consists of a valve sleeve, valve spool, motor, and sensors, as illustrated in Figure 5(a).

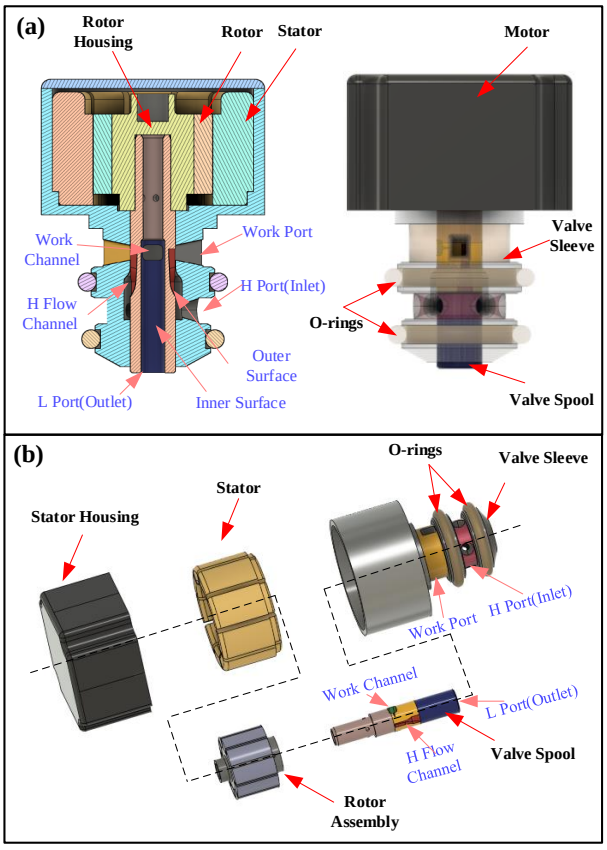


Figure 5 Structural composition of RDDV.

The operational principle of the RDDSV is based on the relative rotation between the valve spool and sleeve, as shown in Figure 5(b). When the valve spool rotates, the cavities formed by the H Flow channel and work channel in the spool align with the circumferential cavities in the valve sleeve, creating various flow path configurations. These flow path combinations determine the direction of fluid energy transfer and the actuator's motion characteristics.

The valve sleeve features an annular groove that serves as the High Pressure inlet, with a circumferential array of ports. Corresponding internal grooves are machined on the sleeve's inner wall, forming cavities that ensure uniform distribution of high-pressure oil across the valve spool surface. A work annular groove, connected to the actuator's load chamber, is equipped with rectangular ports designed to align with corresponding ports on the valve spool. The High Pressure port and Work port are isolated by sealing elements. The valve spool has a hollow tubular structure with both external and internal surfaces. The external surface contains an oil passage (H Flow channel) extending from the H Port to the Work port, formed within the spool's wall thickness without penetrating it, as illustrated in Figure 5(b). The internal surface incorporates axial through-holes that radially traverse the spool, connecting to the Low Pressure port (L Port), thereby facilitating fluid transfer from the actuator's load chamber to the Rotary Switch Valve.

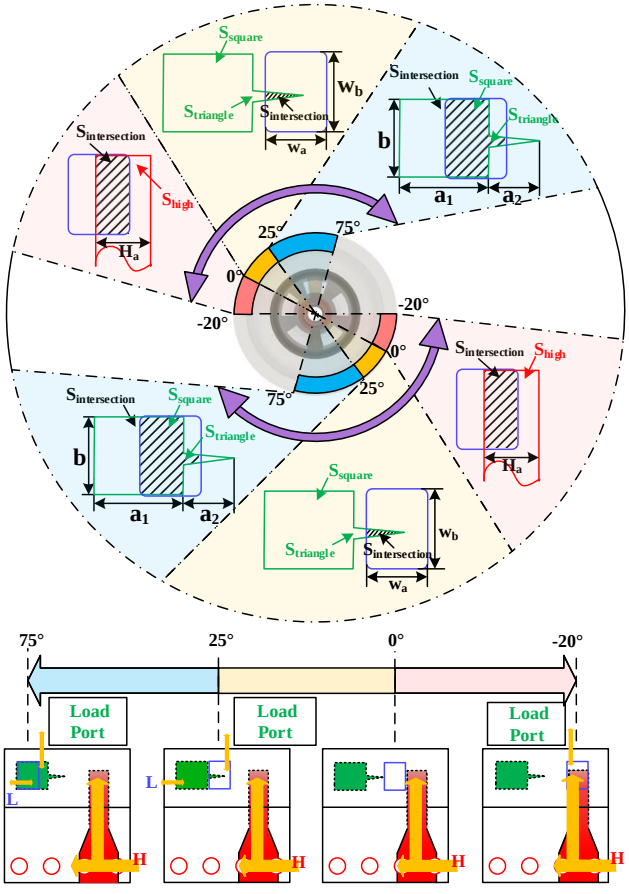


Figure 6 Change in flow area during valve spool rotation.

3.3 Rotary Switch Valve

To achieve these improvements, we developed a directly servo-driven Rotary Switch Valve (RSV). The spool-type rotary valve utilizes motor control algorithms for precise angular positioning, enabling adaptive fluid distribution across pressure gradients. As shown in Figure 8, the cartridge-integrated assembly comprises: a valve sleeve, spool, rotor, stator, rotor seat, bearings, and sensors. This modular architecture supports flexible pressure selection through controlled spool rotation while maintaining operational stability.

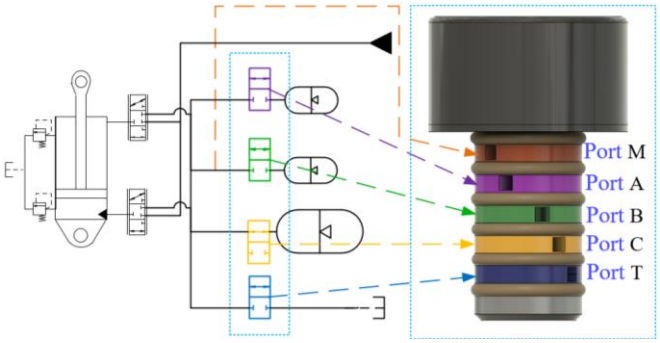


Figure 7 A single RSV control the fluid flow direction.

The valve spool is integrated with the rotor assembly for synchronized rotation, enabling control over various hydraulic circuits to facilitate switching between active and passive modes, servo control in active mode, and damping adjustment in passive mode, as shown in Figure 6. The valve spool's rotational range is 90°. When rotating clockwise between 0° and 20°, high-pressure fluid enters through the sleeve's H Port into the H Flow channel on the spool's external surface, then exits through the sleeve's Work port to enter one load chamber of the actuator, driving piston movement. Simultaneously, the RDDSV connected to the actuator's other load chamber has its spool positioned at a counterclockwise rotation of 25° to 50°, allowing fluid to flow from the load chamber into the sleeve's Work port, pass through the square area of the spool's work channel, and exit through the spool's axial through-hole to connect with the Low Pressure port (L Port). This configuration of both RDDSVs achieves the actuator's active mode setup, enabling active control of fluid flow to drive IKPSA motion.

When the RDDSV spool is in the counterclockwise rotation range of 0° to 25°, it operates in passive mode for damping adjustment. By controlling the overlapping area between the triangular region of the spool's work channel and the sleeve's Work port, a variable-area damping orifice is formed, allowing for adjustment of the IKPSA's passive damping characteristics.

In summary, the RDDSV primarily assumes responsibility for oil circuit closure, active-passive mode switching, and passive damping control. When transitioning to passive mode, the RDDSV switches to the low-pressure port (T-port) by rotating the valve spool, thereby regulating the opening area to modulate the damping force within the passive system.

This mechanism simultaneously facilitates impact absorption, establishing a fundamental design basis for compliant control in the passive system and enhancing overall energy efficiency of the complete system.

As shown in Figure 8(a), the valve sleeve contains four functional selection ports (Ports A, B, C, T) and one main port (Port M). The selection ports are configured to interface with multiple pressure sources - derived from metal bellows

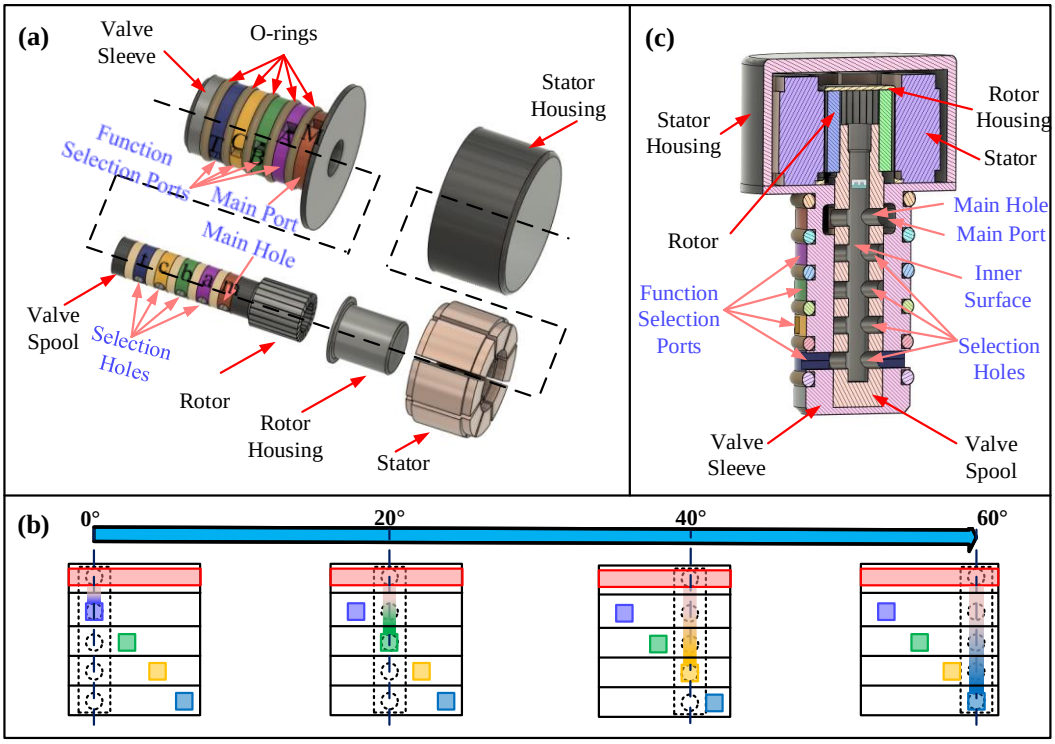


Figure 8 Structural composition of RSV.

accumulators with varying performance characteristics and the system's fluid outlet. Port M connects to the low-pressure (L) ports of both Rotary Direct-Drive Servo Valves (RDDSVs), receiving return flow from the hydraulic cylinder's dual load chambers. This integrated configuration enables system-wide stiffness matching while supporting IKPSA mode switching through coordinated pressure management.

The valve sleeve features helically distributed functional selection ports, spanning a total angle of 60°, with each port positioned at 20° intervals. To balance the hydrodynamic forces on the valve spool, an additional set of four corresponding functional selection ports is symmetrically distributed along the axis, maintaining identical angular arrangement.

The valve spool, located within the valve body, comprises an Inner Surface, eight Selection Holes, and one Main Hole. The Selection Holes and Main Hole facilitate fluid introduction into the Inner Surface and enable intercommunication. The rotor, rotor seat, sensors, and other components are directly connected to the valve spool, forming an integrated structure that rotates in response to controller commands. A magnetic encoder sensor is employed to detect the position of the valve spool relative to the sleeve, enabling the controller to determine the rotational motion.

The operational principle of the Rotary Switch Valve initiates with the controller selecting the desired stiffness level. Based on the position of the selection ports and information provided by sensors, the controller determines the required rotational angle for the valve spool and drives its rotation to align with the selected port. When the spool's selection opening aligns with the chosen port, it establishes a pathway between the internal chamber and the valve body, allowing fluid at the specific hydraulic stiffness level to access the control port.

The rotary rapid switching valve achieves high-speed fluid control through continuous rotational motion, while maintaining a compact structure with minimal spatial footprint.

3.4 Active-passive combined system integration

To enhance the actuator's integration level, an integrated housing was designed to incorporate various interfaces, control valves, accumulators, sensors, and the hydraulic cylinder using a tubeless configuration. Internal channels were engineered to interconnect these components, forming an integrated hydraulic circuit that fulfills the system's design requirements and functional objectives. The controller, mounted on the side of the integrated housing, manages mode switching and knee joint control of the IKPSA. Figure 9 (a)(b) presents the exploded view and cross-sectional diagram of the IKPSA, while Figure 9(c)(d)(e) illustrate schematic diagrams

of each operational mode, respectively. A smulation model was established to analyze the performance characteristics of the IKPSA across various operational modes, with all performance simulations conducted under no-load conditions.

Considering the constraints of the existing lower limb platform, the specific parameters of the IKPSA are presented in the accompanying table.

Table 1 IKPSA's parameters

Parameter	Value
Rated pressure	21.0 MPa
Maximum output force	14000.0 N
Piston diameter	30.0 mm
Rod diameter	22.0 mm
Length of stroke	55.0 mm
Estimated weight	2.5 kg
Accumulator A volume	35.0 ml
Accumulator B volume	20.0 ml
Accumulator C volume	10.0 ml
Spool diameter of RDDSV	5.0 mm
Sleeve Maximum diameter of RDDSV	13.5 mm
Spool diameter of RSV	6.4 mm
Sleeve Maximum diameter of RSV	13.5 mm
Height of actuator	35 mm
Weight of actuator	142 mm

Figure 10 illustrates the temporal relationship between force and displacement of the IKPSA in active mode. Figure 10(a) shows the performance of IKPSA's rods during one cycle of active retraction. The blue line represents displacement, and the magenta color represents force. After the signal is obtained at 0.1s, the piston displacement is gradually retracted to the end of the stroke while the pull force is gradually increased to about 7000N. Figure 10(b) also shows the force versus displacement versus time when the rod of the IKPSA is actively extended during a cycle, with the blue line and magenta color having the same meaning as in graph (a). From the graph, it can be seen that after getting the signal at 0.1s, the piston rod gradually extends and reaches the maximum stroke at 0.86s, and the maximum thrust is able to reach 15000N at 1s.

Figure 11 illustrates the force-time characteristic curve of the IKPSA in burst mode, showing the force variation of the IKPSA during differential drive. A step signal is input to the IKPSA, and upon switching the control valve, the piston begins to move as hydraulic fluid flows in, causing the thrust to increase rapidly. At 0.5 seconds, the thrust force reaches a maximum of approximately 2,500 N, reflecting the maximum output force that the hydraulic cylinder can provide in differential drive mode.

The IKPSA's passive mode operation demonstrates its capacity to resist and absorb external disturbances through controlled mechanical response. Figure 12(a) compares force-displacement curves for two accumulator configurations, revealing distinct stiffness characteristics. Accumulator B's steeper curve indicates enhanced stiffness adjustability compared to Accumulator A, with measurable performance differences during pressurized retraction cycles. The system achieves tunable stiffness between 300 N/mm and 7000 N/mm through accumulator selection.

Moving on to Figure 12(b), this figure illustrates the adjustment of the damping characteristics of the IKPSA by driving the spool of the RDDSV to rotate in passive mode. The angle here is the rotational position of the spool, which controls the flow of liquid through the orifice. As can be seen, the spool rotation in the range of 20° ~ 48° allows a fast and wide range of adjustment of the damping force in the range of 500 ~ 20,000 N, and a fast and infinitely variable adjustment of the damping rate in the range of 40 N·s/mm ~ 2000 N·s/mm.

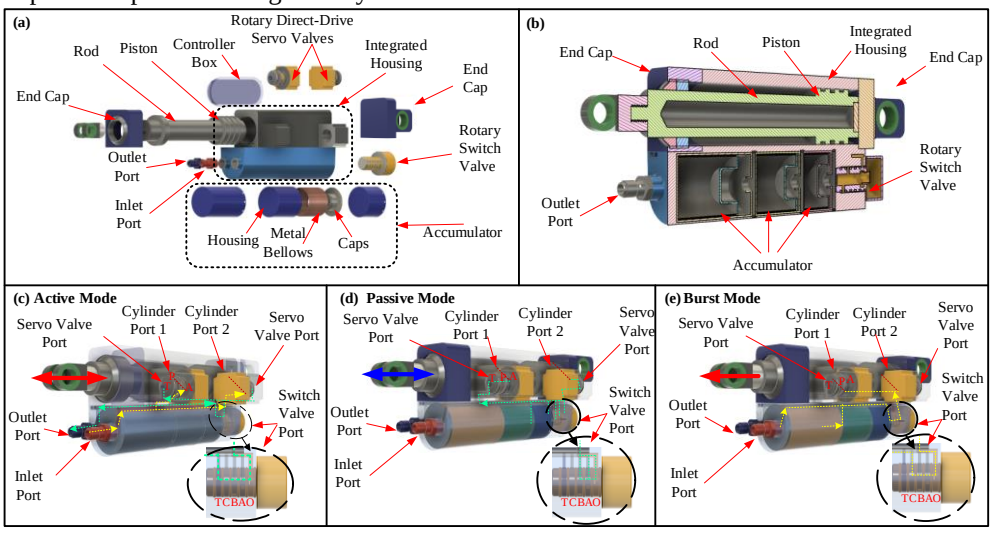


Figure 9 Structural composition of IKPSA and mode switch

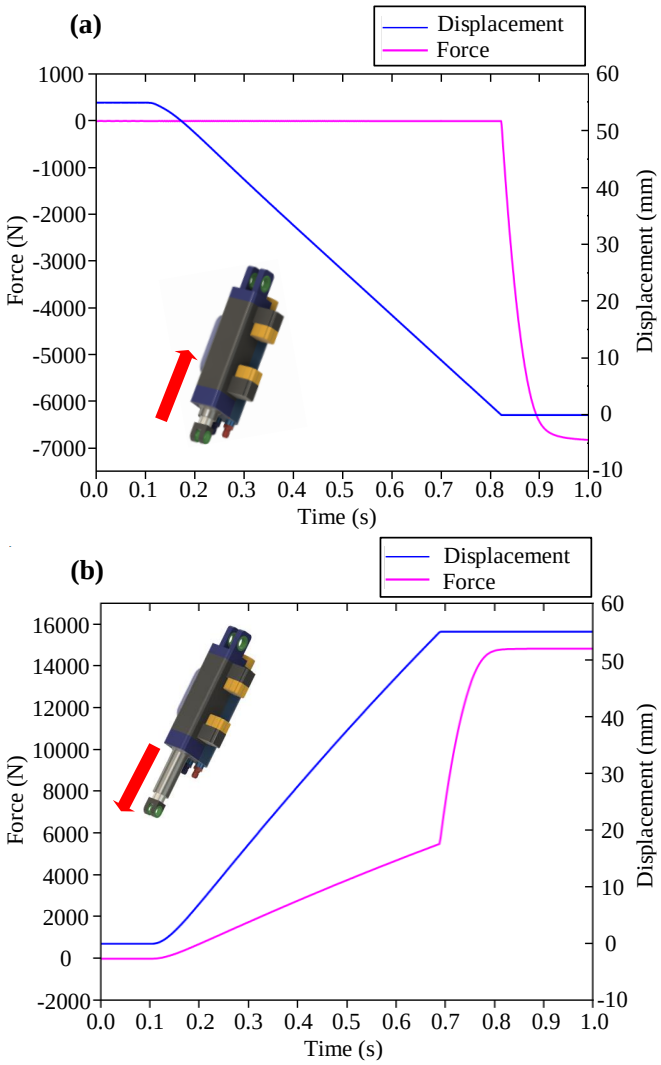


Figure 10 Relationship between force and displacement with time in Active Mode

Simulation data show that IKPSA can support the walking and vertical jumping functions of the biped robot by switching each mode.

4 DISCUSSION

The primary objective of this paper is to design an electro-hydraulic valve-controlled actuator with three modes of operation active, passive, and burst for the existing hydraulic bipedal robot lower limb platform. This design aims to achieve large load capacity, stiffness adjustment, damping adjustment, and high burst performance, inspired by the tasks and functions that human beings undertake during the gait cycle of walking and vertical jumping.

For bipedal robots, especially bipedal robots with large load capacity, in order to reduce the impact between the robot's lower limbs and the unknown environment, it is necessary to ensure that the body has a certain compliance, which is often neglected in the previous actuator design of hydraulic robots.

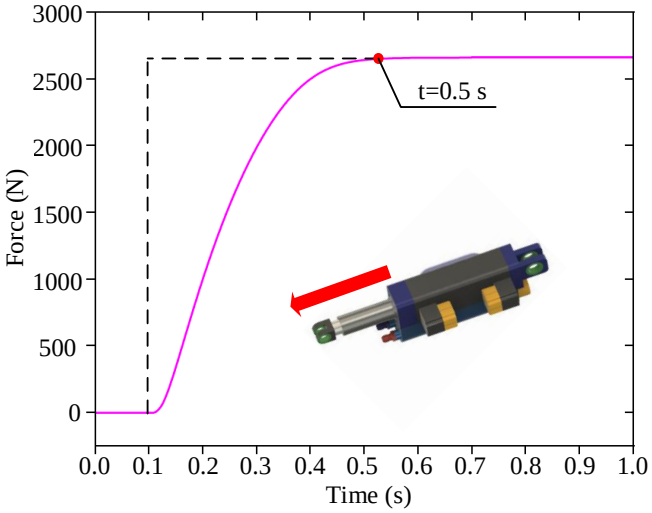


Figure 11 Force-Time Properties in Burst Mode

In IKPSA, through the design of parallel accumulators as well as RDDSVs, we achieve the compliance of the hydraulic knee with an active adjustment. The inherent advantage of hydraulic drives is their ability to carry large loads, but for the hydraulics, increasing the force-to-weight ratio and improving integration is always a goal to be pursued, and although the IKPSA's housing has been designed to maximise the high level of integration, the weight of the IKPSA was able to be further reduced by using a flow channel optimisation algorithm to calculate the flow channel, ditching the process hole technology, and using 3D printing technology, which will be realised in the next work. The IKPSA's modular architecture employs plug-and-play components and standardized interfaces to enhance both operational performance and maintenance efficiency in bipedal robots. This integrated design philosophy incorporates an embedded controller and hydraulic quick-connect couplings, enabling versatile actuator deployment across multiple robotic platforms. A key integration improvement stems from the implementation of cartridge-style rotary control valves with pivot shafts, which streamline hydraulic circuit configuration. The system features dual-functionality components where the Rotary Direct-Drive Servo Valve (RDDSV) simultaneously operates as a precision flow regulator and adaptive damping controller, while a single Rotary Switch Valve (RSV) replaces four conventional on/off valves. When synergized with advanced control algorithms, this optimized configuration demonstrates superior dynamic response characteristics compared to traditional hydraulic systems.

However, it cannot be denied that due to the relatively complex design of the hardware system, higher requirements have been imposed on the final overall machine control and the algorithms for active-passive switching.

5 CONCLUSION

In this paper, a valve-controlled electro-hydraulic actuator (IKPSA) with active-passive combination for bipedal robots is designed to provide different characteristics at different

times, simulating the function of the human knee joint during walking and vertical jump cycles. The hydraulic principle, key basic components, and integrated performance of this actuator are described in detail. The compactness and light weight of the actuator system are enhanced by the integrated housing and the new Rotary Direct-Drive Servo Valve and Rotary Switch Valve to control fluid flow, resulting in a more responsive actuator system. The control system can output commands to the Rotary Switch Valve and Rotary Direct-Drive Servo Valve to control the length, stiffness, damping, and force of the hydraulic cylinder extension.

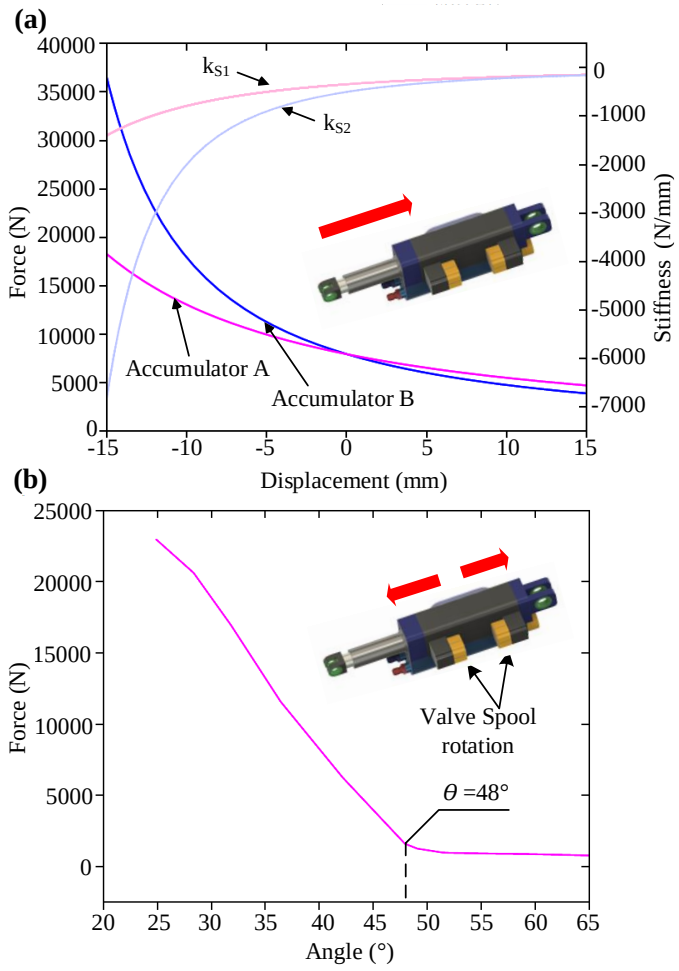


Figure 12 Passive mode (a) Stiffness characteristic curves with different accumulator switching (b) Damping characteristics at spool rotation angle

In IKPSA's active mode, the actuator can actively extend and retract to complete the swing phase of the walking cycle, and at the same time, it is capable of outputting a maximum thrust of 15,000N. In passive mode, the actuator absorbs the impact between the foot and the unknown environment, completes the stance phase of the walking cycle, and completes the landing phase of the vertical jump cycle. It is capable of actively adjusting the actuator stiffness from 1200 N/mm to 300 N/mm and infinitely adjusting the damping from 2000 N·s/mm to 40 N·s/mm. In Burst Mode, energy can be stored to deliver 2,500 N of thrust in 0.41s to complete the Take-Off Phase of the vertical jump.

Nomenclature

Designation	Denotation	Unit
F	Rod Force	N
P_i	Cylinder load chamber pressure	Pa
A_i	Piston area	mm^2
F_{visc}	viscous forces	N
m_i	Masses carried at hinge points	kg
k_s	Actuator stiffness	N/mm
x_i	Hinge point displacement	mm
c	damping ratio	N·s/mm
Q_i	Cylinder load chamber flow	Pa
V_{acc}	Accumulator volume	ml
A_C	Accumulator gas-liquid baffle contact area	mm^2
P_{pc}	Gas precharge pressure	bar
γ	polytropic index	-

Acknowledgement

This work was supported by the National Natural Science Foundation of China under Grant 52405078， National Key Research and Development Program of China under Grant 2024YFF0505300

References

[1] M Chen, Q Li, S Wang, K Zhang, H Chen, and Y Zhang. Single-Leg Structural Design and Foot Trajectory Planning for a Novel Bioinspired Quadruped Robot, *Complexity*, vol. 2021, no. 1, p. 6627043, 2021, doi: 10.1155/2021/6627043.

[2] R Playter, M Buehler, and M Raibert. BigDog, in *Unmanned Systems Technology VIII*, SPIE, May 2006, pp. 896–901. doi: 10.1117/12.684087.

[3] M Raibert, K Blankespoor, G Nelson, and R Playter. BigDog, the Rough-Terrain Quadruped Robot, *IFAC Proceedings Volumes*, vol. 41, no. 2, pp. 10822–10825, Jan. 2008, doi: 10.3182/20080706-5-KR-1001.01833.

[4] Z Jiang, M Li, and W Guo. Running control of a quadruped robot in trotting gait, in *2011 IEEE 5th International Conference on Robotics, Automation and Mechatronics (RAM)*, Qingdao, China: IEEE, Sep. 2011, pp. 172–177. doi: 10.1109/RAMECH.2011.6070477.

[5] X Rong, Y Li, J Ruan, and B Li. Design and simulation for a hydraulic actuated quadruped robot, *J Mech Sci Technol*, vol. 26, no. 4, pp. 1171–1177, Apr. 2012, doi: 10.1007/s12206-012-0219-8.

[6] Bao-ling, Tian, Qing-sheng, Ying, and Ming-hui. Obstacle avoidance technology of bionic quadruped robot based on multi-sensor information fusion, *jbit*, vol. 25, no. 4, pp. 448–454, 2016, doi: 10.15918/j.jbit1004-0579.201625.0402.

[7] J Wang, F Gao, and Y Zhang. High Power Density Drive System of a Novel Hydraulic Quadruped Robot, doi: 10.1115/DETC2014-34804.

[8] V Barasuol *et al.* Highly-Integrated Hydraulic Smart Actuators and Smart Manifolds for High-Bandwidth Force Control, *Front. Robot. AI*, vol. 5, Jun. 2018, doi: 10.3389/frobt.2018.00051.

[9] E Garcia, J C Arevalo, G Muñoz, and P Gonzalez-de-Santos. On the Biomimetic Design of Agile-Robot Legs, *Sensors*, vol. 11, no. 12, pp. 11305–11334, Nov. 2011, doi: 10.3390/s111211305.

[10] Z Hua, X Rong, Y Li, H Chai, B Li, and S Zhang. Analysis and Verification on Energy Consumption of the Quadruped Robot with Passive Compliant Hydraulic Servo Actuator, *Applied Sciences*, vol. 10, no. 1, Art. no. 1, Jan. 2020, doi: 10.3390/app10010340.

[11] D Ren, J Shao, G Sun, and X Shao. The Complex Dynamic Locomotive Control and Experimental Research of a Quadruped-Robot Based on the Robot Trunk, *Applied Sciences*, vol. 9, no. 18, Art. no. 18, Jan. 2019, doi: 10.3390/app9183911.

[12] S-H Hyon, D Suewaka, Y Torii, N Oku, and H Ishida. Development of a fast torque-controlled hydraulic humanoid robot that can balance compliantly, in *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*, Nov. 2015, pp. 576–581. doi: 10.1109/HUMANOIDS.2015.7363420.

[13] J Lee, J Yang, and K Park. The Role of Knee Joint in Passive Dynamic Walking, *Int. J. Precis. Eng. Manuf.*, vol. 26, no. 2, pp. 415–427, Oct. 2024, doi: 10.1007/s12541-024-01084-7.

[14] J H Challis. EXAMINATION OF THE SCALING OF HUMAN JUMPING, *The Journal of Strength & Conditioning Research*, vol. 18, no. 4, p. 803, Nov. 2004.

Health Aware Control Design for Extending Remaining Useful Life of Electro-Hydrostatic Actuator Based Steering

Simo Käki , Vinay Partap Singh , Mikko Huova , and Tatiana Minav

Innovative Hydraulics and Automation, Tampere University, Tampere, Finland
E-mail: simo.kaki@tuni.fi, vinaypartapsingh@tuni.fi, mikko.huova@tuni.fi, tatiana.minav@tuni.fi

Abstract

Electro-hydrostatic actuators (EHAs) are used in aerospace, industrial, and off-road applications due to their efficiency and self-contained operation. However, degradation of critical components, such as the pump, can lead to reduced performance and unexpected failures.
Conventional control strategies often neglect real-time degradation effects, limiting their ability to optimize both actuator performance and remaining useful life (RUL). This study addresses this gap by developing a health-aware control (HAC) framework that integrates degradation modeling with predictive control for EHAs.
The proposed HAC system incorporates a nonlinear Wiener degradation model to estimate internal leakage progression in the pump. The degradation state is estimated online using an extended Kalman filter (EKF) and integrated into a model predictive control (MPC) framework to dynamically adjust control actions based on degradation predictions. The effectiveness of the approach is evaluated through simulations, comparing it to a baseline PID-controller. The choice of MPC and EKF over traditional PID control is justified by their ability to handle nonlinearities, provide predictive capabilities, and offer robust state estimation.
Simulation results show that the HAC system extends the RUL of the gear pump by 3.22 %, compared to the PID controller. The system successfully mitigates degradation effects, while maintaining an average position tracking error of 1.39 mm.
Although the RUL improvement is modest under accelerated degradation conditions, the proposed approach shows strong potential for real-world applications where degradation occurs over extended periods.
Future work will focus on experimental validation using physical test rigs and refining the degradation model to enhance prediction accuracy and control adaptability.

Keywords: electro-hydrostatic actuator, health aware control, hydraulic pump, model predictive control, remaining useful life

1 Introduction

Safety-critical systems in off-road mobile machinery such as steering mechanisms, braking systems, and actuators, play a critical role in ensuring both operator and machine safety during operations in challenging environments. Such systems are subjected to harsh conditions, increasing the likelihood of functional degradation and unexpected failures. Their failure can result in serious consequences, making reliability and timely maintenance essential. These safety implications highlight the need for proactive approaches that ensure sustained functionality even under continuous and demanding operating conditions. By maintaining safe operational states, these systems help prevent accidents, protect assets, and uphold regulatory compliance, particularly in mission-critical applications such as autonomous or remote-operated machines.

In recent years, the focus on sustainability and reliability in engineering has underscored the importance of extending the operational lifespan of key components. This is particularly relevant for off-road machinery, where operational downtime and maintenance costs have significant economic and environmental implications. Functional degradation in such systems often leads to unexpected failures, posing challenges to both performance and safety. Proactive measures such as health monitoring and prognostics are becoming essential tools in mitigating these risks, ensuring that safety-critical systems perform optimally throughout their life-cycle. A core concern for safety-critical systems is the accurate assessment and management of their Remaining Useful Life (RUL), which determines how long a component can operate before performance drops below acceptable thresholds. Degradation is often influenced by how the system is operated, with certain operating

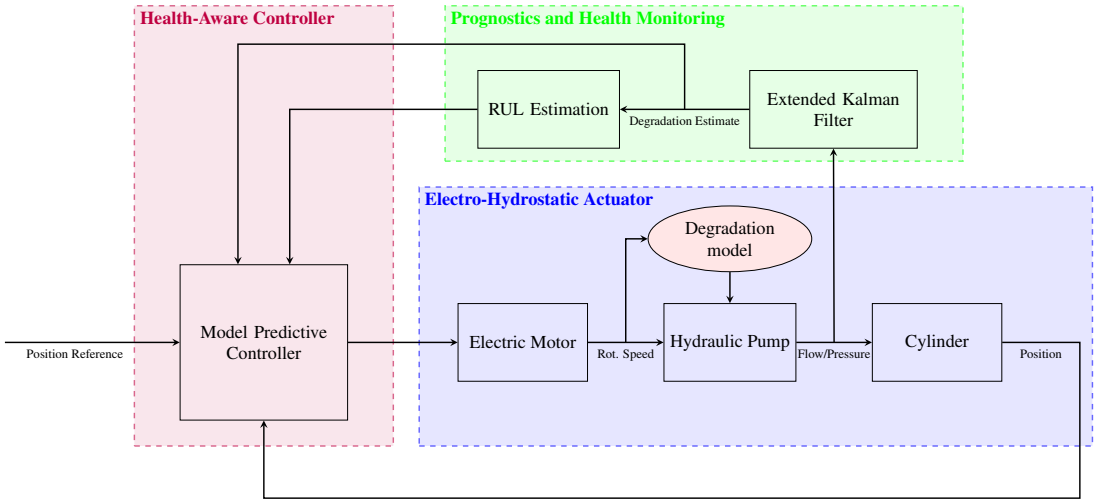


Figure 1: HAC framework for EHA.

points likely to accelerate degradation. Hence, manipulating operating points to manage degradation and reduce the risk of unplanned stoppages is desirable. Therefore health-aware strategies offer the potential to balance performance and reliability, minimizing resource use while maintaining system functionality.

The health-aware control (HAC) paradigm was introduced by Escobet et al. [1,2] as an approach that integrates system health monitoring with control and prognosis to enhance system reliability. In this paradigm, a prognosis module continuously estimates the health status and aging of system components, allowing the controller to modify its objectives accordingly. Unlike fault-tolerant control, HAC proactively adjusts the control strategy even before faults occur, aiming to optimize performance while extending the lifespan of components. Furthermore, while fault-tolerant control can assure acceptable performance when the actuator's health is poor, it may inadvertently accelerate degradation, highlighting the need for integrated health-aware strategies.

Model predictive control (MPC) has emerged as an effective framework for integrating prognosis and control to manage actuator degradation. The robustness and performance of MPC is shown to be better than a traditional control method in an electro-hydraulic system [3]. Unlike traditional PID-controllers, MPC can handle nonlinearities and constraints, making it more suitable for complex systems. For instance, Pereira et al. [4] developed a MPC framework that utilizes prognosis information to distribute control effort among redundant actuators, thereby avoiding breakdowns due to degradation. Similarily, a recent study by Dadash and Björsell [5] utilized MPC with novel methodology to calculate the state-action costs. This approach didn't use preconceived models of machine degradation. Zhang et al. [6] addressed the challenge of balancing reliability and performance in systems with degrading actuators, providing a foundational approach to degradation-aware control and demonstrating significant improvements in maintaining system functionality over extended periods.

To support HAC, accurate degradation modeling is crucial. The study by Letot and Dehombreux [7] provided insights into the use of degradation models for reliability estimation and residual lifetime predictions, showcasing their applicability in designing maintenance schedules that maximize system uptime. Recent advancements by Ma et al. [8] highlight the potential of physics-informed machine learning models to capture complex degradation phenomena in electro-hydrostatic actuators, achieving enhanced prediction accuracy compared to traditional models. Additionally, Zagorowska et al. [9] conducted a comprehensive survey on degradation models, emphasizing their applicability in control design and identifying gaps in existing methodologies. Their findings highlighted the influence of operating conditions on degradation, proposing strategies to mitigate accelerated wear. The combination of HAC and robust degradation modeling methodologies enhances the reliability and sustainability of safety-critical systems, facilitating predictive maintenance and minimizing the risk of unexpected failures.

Electro-hydrostatic actuators (EHAs) are used in modern hydraulic systems due to their efficiency, high power density, and ability to operate without centralized hydraulic circuits [10]. EHAs integrate electrical, mechanical, and hydraulic components, typically consisting of an electric motor, a hydraulic pump, and a hydraulic actuator. Unlike conventional hydraulic systems that rely on centralized pumps and fluid distribution networks, EHAs function as self-contained units, directly converting electrical energy into hydraulic power to drive actuators. This architecture enhances reliability, reduces maintenance needs, and eliminates hydraulic piping, making EHAs particularly attractive for aerospace, off-road machinery, and industrial automation applications [8, 10].

However, EHAs are subject to degradation and faults, particularly due to internal leakage and wear in the hydraulic pump. Previous research has explored real-time health monitoring [11,12] and fault-tolerant control methods [13] to sustain functionality under failure conditions, but these strategies do not actively mitigate degradation. Instead, they focus on ensuring

operability, sometimes at the cost of accelerated wear. This limitation underscores the need for HAC strategies, which integrate degradation estimation and predictive control to extend the RUL while maintaining system performance.

The HAC paradigm enhances system reliability by integrating health monitoring, advanced control strategies such as MPC, and advanced degradation modeling to proactively manage actuator degradation. By dynamically adjusting control actions based on real-time health assessments, HAC optimizes performance while extending component lifespan. Robust degradation models further improve predictive maintenance and failure prevention. These works collectively underscore the growing recognition of the importance of integrating health monitoring and control to enhance system longevity and reliability.

While HAC has been explored in various actuator systems, its application to EHAs remains limited. As a result, control designs often fail to optimize both performance and RUL in real-world conditions. Additionally, there is a lack of frameworks that integrate HAC with accurate degradation modeling to achieve these optimizations in EHA systems.

The primary aim of this research is to develop a health-aware control system for previously proposed EHA-based steering mechanism [14], focusing on extending the RUL of a gear pump. This work addresses the need for integrating degradation modeling, real-time prognostics, and control strategies to prolong the lifespan of the EHA. By doing so, the study seeks to advance current methodologies and provide practical solutions for improving the sustainability and reliability of off-road machinery.

This study develops a HAC framework for EHA with a focus on the degradation of the gear pump. This framework is visualized in fig. 1. The proposed approach integrates real-time degradation modeling into the control strategy, allowing adaptive adjustments to optimize both actuator performance and RUL. By accounting for the progressive wear of the gear pump, the framework ensures reliable operation under varying conditions, enhancing the longevity and efficiency of the EHA system.

The remainder of this paper is organized as follows. Section 2 presents the EHA and pump degradation modeling. Section 3 discusses the implementation of degradation estimation and HAC using a MPC. Section 4 presents simulation results and performance analysis. Finally, Section 5 concludes the paper and discusses future directions.

2 System description and modelling

EHAs offer efficiency and reliability but are susceptible to degradation, particularly in the gear pump, which is the primary focus of this study. Increased internal leakage due to wear reduces performance and safety. First this section presents the EHA system model, capturing its electrical, mechanical, and hydraulic interactions. Then a degradation model for the pump is presented to support health-aware control strategies.

2.1 Electro-Hydrostatic Actuator System Model

The EHA system comprises an electric motor, a fixed displacement hydraulic pump, and a double-acting hydraulic cylinder. A hydraulic diagram of EHA is depicted in fig. 2. The system's dynamics are modeled to capture the interaction between electrical, mechanical, and hydraulic domains. The following EHA model is adapted from literature [8, 10, 13].

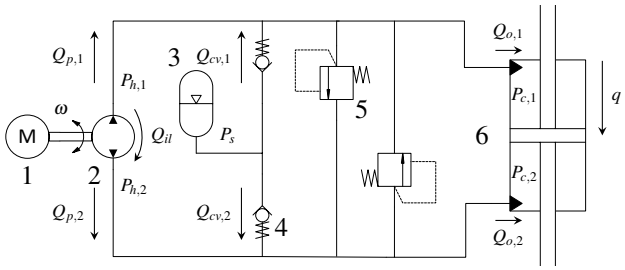


Figure 2: EHA hydraulic diagram. 1. Motor, 2. Pump, 3. Accumulator, 4. Check valve, 5. Pressure relief valve, 6. Linear actuator.

Given that the primary focus of this study is the hydraulic pump, the motor is approximated as a first-order system

$$\dot{\omega} = -\frac{1}{\tau_m}\omega + \frac{1}{\tau_m}\omega_{ref}, \tag{1}$$

where ω is the rotational speed of the motor, ω_{ref} reference rotational speed, and τ_m time constant. This simplification is justified by the high bandwidth of the torque control loop, which makes the transient dynamics of the motor negligible compared to the slower rotational speed and hydraulic system dynamics.

The main subject for investigation is the hydraulic pump. The inlet and outlet flows of the gear pump can be expressed as [8]

$$Q_{p,1} = D_p\omega - C_{il}\left(P_{h,1} - P_{h,2}\right) - C_{el}P_{h,1}, \tag{2}$$
$$Q_{p,2} = -D_p\omega + C_{il}\left(P_{h,1} - P_{h,2}\right) - C_{el}P_{h,2}, \tag{3}$$

where D_p is the pump displacement, $P_{h,1}$, $P_{h,2}$ hose pressures, C_{il} internal leakage coefficient, and C_{el} external leakage coefficient. All internal leakage paths are lumped into single laminar leakage term to minimize the number of parameters that need to be estimated. The external leakage coefficient is assumed to be known.

The hoses connecting the pump to the cylinder chambers are modeled using flow continuity equations

$$Q_{p,1} + Q_{cv,1} = \frac{V_{h,1}}{\beta}\dot{P}_{h,1} + Q_{o,1}, \tag{4}$$
$$Q_{p,2} + Q_{cv,2} = \frac{V_{h,1}}{\beta}\dot{P}_{h,2} + Q_{o,2}, \tag{5}$$

where $V_{h,1}$ and $V_{h,2}$ are hose volumes, β bulk modulus, $Q_{cv,1}$ and $Q_{cv,2}$ flows through the check valves, and $Q_{o,1}$ and $Q_{o,2}$ flows through the orifices.

To prevent cavitation, check valves have been integrated to the model with accumulator as a common pressure source P_s. It

also serves as a source of oil for the system, replacing possible oil losses of the system. The accumulator dynamics are assumed insignificant and thus neglected from the mathematical model of the EHA [15]. The flow through the check valves can be expressed as

$$Q_{cv} = C_q A_{cv} \sqrt{\frac{2}{\rho} \max\{0,\ P_s - P_h\}} \tag{6}$$

where C_q is discharge coefficient, A_{cv} check valve area, and ρ oil density.

The hoses are connected to the cylinder chambers via orifices. The orifice equations can be expressed as

$$Q_o = C_q A_o \sqrt{\frac{2}{\rho} |P_h - P_c|}\, \mathrm{sign}(P_h - P_c), \tag{7}$$

where A_o is the orifice area and P_c is the corresponding chamber pressure. The chamber volumes are

$$V_{c,1} = A_{c,1} q + V_{d,1}, \tag{8}$$
$$V_{c,2} = A_{c,2} (q_{max} - q) + V_{d,2}, \tag{9}$$

where $A_{c,1}$ and $A_{c,2}$ are effective areas of the piston, $V_{d,1}$ and $V_{d,2}$ dead volume of the chambers, and q the displacement of the piston.

The flow continuity equations for the cylinder chambers are

$$Q_{o,1} = \frac{V_{c,1}}{\beta} \dot{P}_{c,1} + A_{c,1} \dot{q}, \tag{10}$$
$$Q_{o,2} = \frac{V_{c,2}}{\beta} \dot{P}_{c,2} - A_{c,2} \dot{q}, \tag{11}$$

where β is the bulk modulus. The force created by the actuator is

$$F_c = A_{c,1} P_{c,1} - A_{c,2} P_{c,2} - B\dot{q}, \tag{12}$$

where B is the viscous friction coefficient. Other friction forces are assumed small and thus neglected.

Finally, the force balance equation

$$F_c - F_l = M\ddot{q}, \tag{13}$$

where F_l is the load force and M is the equivalent load mass. The numerical values of the EHA model are presented in table 1.

The derived model is nonlinear and too complex to be used to design a controller and is thus used as the simulation model of EHA. A linear model derived in [15, 16] gives a good approximation of piston head dynamics from rotational speed of the pump. The transfer function relating the pump angular velocity to the linear displacement of the load is given by

$$\frac{q(s)}{\omega(s)} \approx \frac{\frac{2D_p \beta A}{MV_0}}{s^3 + \left(\frac{B}{M} + \frac{C_L \beta}{V_0}\right) s^2 + \left(\frac{2\beta A^2}{MV_0} + \frac{C_L B \beta}{MV_0}\right) s}, \tag{14}$$

where the total leakage coefficient of the system, consisting of pump internal leakage, pump external leakage, and actuator leakage, is given by $C_L = 2C_{il} + C_{el}$ and V_0 being the mean oil volume in the system.

Table 1: EHA simulation parameter values.

Parameter	Symbol	Value
Motor time constant	τ_m	0.1 s
Pump displacement	D_p	$2.3125 \cdot 10^{-6}$ m^3/rad
External leakage coefficient	C_{el}	$1 \cdot 10^{-13}$ (m^3/s)/Pa
Hose volume	V_h	$4.5239 \cdot 10^{-5}$ m^3
Accumulator pressure	P_s	$2 \cdot 10^6$ Pa
Check valve area	A_{cv}	$1.9635 \cdot 10^{-7}$ m^2
Orifice area	A_o	$2.9225 \cdot 10^{-5}$ m^2
Discharge coefficient	C_q	0.7
Maximum stroke	q_{max}	0.5 m
Chamber dead volume	V_d	$150 \cdot 10^{-6}$ m^3
Viscous friction coefficient	B	1500 N/(m/s)
Bulk modulus	β	$900 \cdot 10^6$ Pa
Oil density	ρ	850 kg/m^3
Effective inertial load	M	650 kg

2.2 Degradation Modeling for Gear Pumps

Gear pumps are critical components in hydraulic systems, where degradation typically manifests as increased internal leakage due to wear and tear. Since precise details of the degradation process are often unavailable, heuristic degradation models can be constructed to represent this behavior [9]. To effectively manage degradation, these models must account for the applied control effort and other influencing factors [17, 18]. For pumps, cumulative degradation models have been commonly used [4, 18, 19], demonstrating that high flow rates accelerate wear and degradation.

In this paper, the degradation of the pump is assumed to depend on rotational speed, with internal leakage increasing as the primary manifestation. This degradation is captured by an increase in the internal leakage coefficient, C_{il}, which serves as a key indicator of the pump's health.

Wiener-process-based models have been widely adopted in control applications due to their ability to handle stochastic variations [20, 21]. A nonlinear Wiener-process model, representing the progression of degradation, is formulated as follows:

$$\psi(t) = \varphi_0 + \int_0^t \mu(\tau, u(\tau), \Theta)\, d\tau + \sigma B(t), \tag{15}$$

where φ_0 is the initial degradation measure, $\sigma > 0$ is the diffusion coefficient, $B(t) \sim \mathcal{N}(0,t)$ standard Brownian motion and Θ parameter vector. The nonlinear part $\int_0^t \mu(\tau, u(\tau), \Theta)\, d\tau$ describes the deterministic degradation path of the actuator coupled with the control action $u(t)$. The form of $\mu(\tau, u(\tau), \Theta)$ is strongly context-dependent and should be carefully determined. Accelerated life test rigs can be utilized to identify the degradation path of the actuator.

This formulation can also be expressed as an ordinary differ-

ential equation (ODE)

$$\dot{\psi}(t) = \mu(t, u(t), \Theta) + \sigma w(t), \tag{16}$$

where $w(t)$ is white noise. This ODE can can be used as the degradation model in the controller design. It is also noted that the degradation speed $\dot{\psi}(t)$ can also depend on the current degradation or degradation at time t. Thus (16) is augmented and presented as

$$\dot{\psi}(t) = \mu(t, \psi(t), u(t), \Theta) + \sigma w(t). \tag{17}$$

To identify the degradation model for gear pumps, data from an accelerated life test rig was utilized [22]. By contaminating the hydraulic oil with particles, controlled wear was induced on a 16 cc/rev gear pump. The particle count used in the test was 23/22/19 in ISO4406:2021 standard [23]. The pump was driven with rotational speed of 1000 rpm for the whole duration of the test. The measured internal leakage data is shown in fig. 3, and the fitted deterministic degradation model is expressed as

$$\dot{\psi} = \Theta |\omega| \psi, \tag{18}$$

where Θ is the fitted parameter. The inclusion of degradation in equation (18) is motivated by higher internal leakage flow wearing the pump down even further.

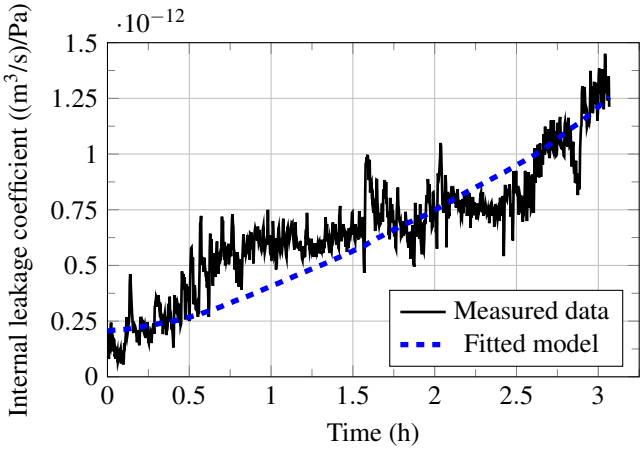


Figure 3: Measured internal leakage coefficient and fitted degradation model in accelerated test conditions.

The stochastic component of (15) is also considered by looking at the residuals of the fitted model in fig. 3. The residuals are shown in fig. 4. The effect of sensor noise to the residuals is assumed to be insignificant. The Wiener process satisfies

$$\mathrm{Var}[r(t)] = \sigma^2 t, \tag{19}$$

from which the diffusion coefficient can be estimated with

$$\hat{\sigma} \approx \sqrt{\frac{1}{n}\sum_{i=1}^{n} \frac{r^2(t_i)}{t_i}}. \tag{20}$$

From the residuals the diffusion coefficient σ is estimated to be approximately $2.819 \cdot 10^{-15}$.

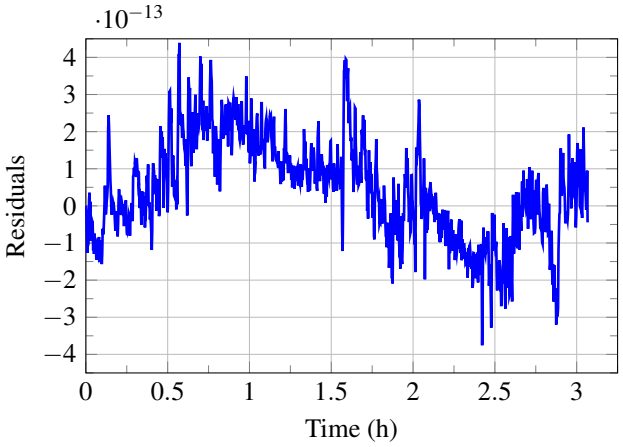


Figure 4: Residuals of degradation model fit.

Finally, the degradation model for the gear pump can be expressed as

$$\dot{\psi} = \Theta |\omega| \psi + \sigma w(t). \tag{21}$$

The internal leakage coefficient can be estimated using extended Kalman filter online.

To determine the RUL of the actuator, the future degradation trajectory is solved from equation (18) by separating variables and integrating both sides. It is assumed that ω stays constant and that the impact of noise is insignificant. Thus the degradation trajectory is

$$\psi(t) = \psi_0 e^{\Theta|\omega|t}, \tag{22}$$

where ψ_0 is the current degradation. By setting this equal to the failure threshold S and solving for time, the RUL is

$$RUL(t) = \frac{\ln S - \ln \psi_0}{\Theta |\omega|}. \tag{23}$$

From (23) it is evident that the RUL will tend to infinity when pumps rotation speed approaches zero. This effect can be mitigated by adding a small constant to the denominator of (23). As ω was assumed constant, the RUL estimate needs to be updated every time ω changes. An estimate of average ω can also be used based on previous usage of the pump. This can be calculated using windowed average for example.

3 Health aware control of EHA

This section presents the implementation of HAC using a MPC strategy to dynamically adjust the control signal and balance system performance with degradation mitigation. While PID-controllers are simple and effective for linear and time-invariant systems, they lack the predictive capabilities and constraint handling required for a health-aware system operating under varying conditions. MPC, on the other hand, offers the ability to anticipate future behavior, explicitly incorporate system constraints, and optimize control actions over a prediction horizon. This predictive and constraint-aware nature makes MPC especially suitable for systems like EHA where maintaining both performance and component longevity is essential. Furthermore, MPC provides superior robustness against disturbances and model inaccuracies, and it can

efficiently handle system nonlinearities, which are inherent in EHA systems. By using MPC, it is possible to systematically trade off between immediate performance needs and the long-term health of critical components, thus maximizing the RUL. In addition to the control strategy, this section also introduces the Extended Kalman Filter (EKF) for degradation estimation.

The EKF is chosen for its capability to estimate unmeasurable degradation states accurately in the presence of system nonlinearities and measurement noise. Unlike simpler estimators, the EKF linearizes the system dynamics around the current estimate, allowing it to maintain a high level of estimation accuracy. This real-time degradation information is crucial for the HAC system to adaptively manage the operation of actuator, ensuring reliable and safe performance. Therefore, the combined use of MPC and EKF provides a comprehensive framework for health-aware control, enabling robust, predictive, and adaptive management of system performance and degradation. By dynamically adjusting control actions based on the estimated degradation state, the system can mitigate wear while ensuring adequate actuator response. This approach reduces the risk of unexpected failures, improves overall system safety and sustainability, and aligns with the growing industry emphasis on reliability and proactive maintenance.

3.1 Extended Kalman Filter for Degradation Estimation

This subsection presents the implementation of an Extended Kalman Filter (EKF) to estimate the degradation state of a gear pump. The degradation is not directly measurable, requiring an estimation approach based on available system measurements. The EKF is chosen for its robustness in providing accurate state estimates in the presence of noise and uncertainties, which is crucial for adaptive control. The EKF algorithm is derived for the degradation model given in (21) and utilizes a linearization method at each time step.

The degradation model is formulated as a nonlinear state-space system:

$$\dot{\psi} = f\left(\psi, \boldsymbol{u}_{kf}\right) + w, \tag{24}$$

$$y_{kf} = h\left(\psi, \boldsymbol{u}_{kf}\right) + v, \tag{25}$$

where $y_{kf} = Q_{p,1}$ represents the measurable output. The function $h(\cdot,\cdot)$ corresponds to equation (2), and the input vector is defined as $\boldsymbol{u}_{kf} = \left[\omega, P_{h,1}, P_{h,2}\right]^{\mathsf{T}}$. The process noise $w \sim \mathcal{N}(0,\Sigma_w)$ and measurement noise $v \sim \mathcal{N}(0,\Sigma_v)$ are assumed to follow zero-mean Gaussian distributions. Thus, the estimation process relies on the measurements of $Q_{p,1}$, ω, $P_{h,1}$, and $P_{h,2}$.

The EKF algorithm consists of two main steps: prediction and update. In the prediction step, the state and covariance are propagated forward in time using the system model. This *a priori* information is denoted by superscript minus ($^-$). In the update step *a posteriori* information is denoted by superscript plus ($^+$). Since the system is nonlinear, the model is linearized at each step using a first-order Taylor series expansion. The Jacobians of the system dynamics and measurement function are computed as:

$$f_\psi\left(\psi, \boldsymbol{u}_{kf}\right) = \frac{\partial}{\partial \psi} f\left(\psi, \boldsymbol{u}_{kf}\right) = \Theta\left|\omega\right|, \tag{26}$$

$$h_\psi\left(\psi, \boldsymbol{u}_{kf}\right) = \frac{\partial}{\partial \psi} h\left(\psi, \boldsymbol{u}_{kf}\right) = P_{h,2} - P_{h,1}. \tag{27}$$

The prediction step updates the state estimate and covariance as:

$$F_\psi = f_\psi(\hat{\psi}^+[k-1], \boldsymbol{u}_{kf}[k-1])T_s, \tag{28a}$$

$$\hat{\psi}^-[k] = \hat{\psi}^+[k-1] + f\left(\hat{\psi}^+[k-1], \boldsymbol{u}_{kf}[k-1]\right)T_s, \tag{28b}$$

$$\Sigma_{\tilde{\psi}}^-[k] = F_\psi \Sigma_{\tilde{\psi}}^+[k-1] F_\psi^{\mathsf{T}} + \Sigma_w. \tag{28c}$$

where T_S is the sampling time of the filter.

In the update step, the measurement residual and Kalman gain are computed as:

$$H_\psi = h_\psi(\hat{\psi}^-[k], \boldsymbol{u}_{kf}[k]), \tag{29a}$$

$$\hat{y}_{kf}^-[k] = h\left(\hat{\psi}^-[k], \boldsymbol{u}_{kf}[k]\right), \tag{29b}$$

$$K = \Sigma_{\tilde{\psi}}^-[k] H_\psi^{\mathsf{T}} \left(H_\psi \Sigma_{\tilde{\psi}}^-[k] H_\psi^{\mathsf{T}} + \Sigma_v\right)^{-1}. \tag{29c}$$

The estimated degradation state is then updated using:

$$\hat{\psi}^+[k] = \hat{\psi}^-[k] + K\left(y_{kf}[k] - \hat{y}_{kf}^-[k]\right). \tag{30}$$

The covariance matrix update follows the Joseph form, which ensures numerical stability and preserves symmetry:

$$\Sigma_{\tilde{\psi}}^+[k] = \left(I - K H_\psi\right)\Sigma_{\tilde{\psi}}^-[k]\left(I - K H_\psi\right)^{\mathsf{T}} + K \Sigma_v K^{\mathsf{T}}. \tag{31}$$

The use of the Joseph form prevents loss of positive definiteness in the covariance matrix due to numerical errors. This formulation is crucial when dealing with small numerical updates in long-term estimation tasks.

By applying the EKF, the degradation state ψ is continuously estimated based on the available system measurements. This enables real-time tracking of the pump's internal condition. The proposed estimation method ensures robustness against measurement noise and modeling uncertainties, improving the reliability of the health-aware control system.

3.2 Model Predictive Control for EHA Piston Head Position and Degradation Control

Model Predictive Control (MPC) is employed to achieve two primary objectives: regulating the piston head position and managing the degradation of the gear pump. The MPC framework optimizes the control inputs while considering constraints and performance criteria [4, 6].

The internal model is derived by augmenting the linear models (1) and (14) with the degradation model (21). The prediction

model is

$$\dot{\boldsymbol{x}} = \begin{bmatrix} -a_3\ddot{q} - a_2\dot{q} + b_1\omega \\ \ddot{q} \\ \dot{q} \\ -\frac{1}{\tau_m}\omega + \frac{1}{\tau_m}\omega_r \\ \Theta\psi\,|\omega| \end{bmatrix} \tag{32}$$

$$\boldsymbol{y} = C\boldsymbol{x}, \tag{33}$$

where $\boldsymbol{x} = \begin{bmatrix} \ddot{q} & \dot{q} & q & \omega & \psi \end{bmatrix}^\mathsf{T}$ and

$$a_2 = \frac{\beta A_1^2}{MV_0} + \frac{C_L B_u \beta}{MV_0},\ a_3 = \frac{B_u}{M} + \frac{C_L\beta}{V_0},\ b_1 = \frac{2D_p\beta A_1}{MV_0}. \tag{34}$$

The controlled values are $\boldsymbol{y} = \begin{bmatrix} q & \psi \end{bmatrix}^\mathsf{T}$. The augmented model, also called degradation-incorporated state space (DISS) [6], is continuous and nonlinear. To be used in MPC-framework, the prediction model is discretized using the implicit trapezoidal rule.

MPC tries to find the optimal control sequence $\boldsymbol{u}_k$ by minimizing the cost function

$$J = \sum_{j=1}^{n_y}\sum_{i=1}^{N_p}\left\{\frac{w_j}{s_j}\left(r_j[k+i|k] - y_j[k+i|k]\right)\right\}^2, \tag{35}$$

where k is the current control interval, N_p prediction horizon, n_y number of controlled variables, $y_j[k+i|k]$ predicted value of the jth controlled variable at the ith prediction horizon step, $r_j[k+i|k]$ its reference trajectory, s_j scale factor, and w_j tuning weight.

The position reference values are kept constant for the duration of prediction horizon, if the reference trajectory is unknown. The degradation reference trajectory can be estimated from historical degradation data from same type of actuators which have successfully completed the mission. If acceptable historical data is available, it can be used as the reference trajectory. Additionally, a model based on (22)

$$r_\psi(t) = ae^{bt} + c, \tag{36}$$

where a, b, and c are fitted parameters, can be fitted to the data and used as a reference trajectory.

The weights in (35) are usually treated as tuning parameters and chosen by hand. But for controlling the degradation, a more sophisticated approach must be used. Firstly the weight for degradation should not be higher than the weight for piston head position. Secondly, increasing the weight for degradation will reduce the reference following performance of piston head position. Thus the weight for degradation is chosen with

$$\begin{cases} w_\psi(t) &= \delta w_q \\ w_\psi(t) &= 0, \quad w_\psi(t) < 0 \\ w_\psi(t) &= w_{\psi,max},\ w_\psi(t) > w_{\psi,max} \end{cases}, \tag{37}$$

where

$$\delta = \frac{T_{req} - RUL}{T_{req}}. \tag{38}$$

The scaling parameter δ needs to be ensured to stay between zero and one. This approach increases the degradation weight if the RUL of the pump is not enough to complete the required mission.

4 Results and Discussion

The proposed HAC system was evaluated using a nonlinear simulation model of the EHA, where degradation was introduced as an increasing internal leakage coefficient in the gear pump. The system was tested under varying operating conditions to analyze its tracking performance, degradation progression, and RUL extension. The HAC system was compared against a baseline PID-controller. The comparison highlights the advantages of MPC and EKF in handling nonlinearities, providing predictive control, and offering robust state estimation, which contribute to the observed improvements in RUL.

The PID-controller was tuned with pole placement method. The linear model (14) includes an integrator, and thus the velocity form of PID-controller was used. The discrete time PID-controller algorithm is [24, 25]

$$\begin{aligned} \Delta u[k] &= u[k] - u[k-1] \\ &= K_p\left(e[k] - e[k-1]\right) + K_i T_s e[k] \\ &\quad - \frac{K_d}{T_s}\left(y[k] - 2y[k-1] + y[k-2]\right) \end{aligned} \tag{39}$$

where $e[k] = r[k] - y[k]$. Note that the derivative uses only the controlled variable y. The parameters of the PID-controller are presented in table 2.

The HAC-system was constructed according to section 3.2. The weight value for position error in (35) was chosen as 10. For degradation error, the weight was chosen according to equations (37) and (38), where the maximum was set to 5. This ensured that the position error is always kept small and degradation is a secondary objective. Static parameters of the MPC are listed in table 3.

Table 2: PID-controller parameters.

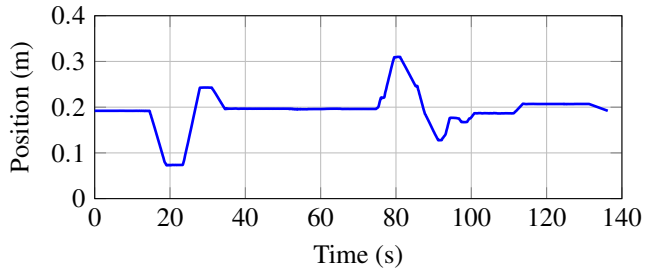
Parameter	Symbol	Value
P-gain	K_p	103.9 rad/(s m)
I-gain	K_i	33 710 rad/(s^2 m)
I-time	t_i	0.003 081 s
D-gain	K_d	15.27 rad/m
D-time	t_d	0.1470 s
Sampling time	T_s	0.05 s

Table 3: HAC parameters.

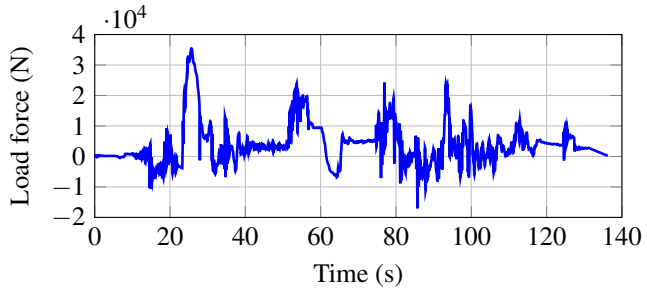
Parameter	Symbol	Value
Prediction horizon	N_p	100
Control horizon	N_c	40
Position error weight	w_q	10
Manipulated variable rate weight	$w_{\Delta u_k}$	1
Sampling time	T_s	0.05 s

The reference position and load force for the EHA are from Mevea virtual environment simulations [14] and are shown in fig. 5. The reference trajectories are cycled through 10 times.

The degradation speed of the pump is greatly accelerated to test the proposed HAC system. This leads to the pump breaking down after 1000 s.



(a) Reference position values.



(b) Piston load force values.

Figure 5: Position and load reference trajectories for simulation.

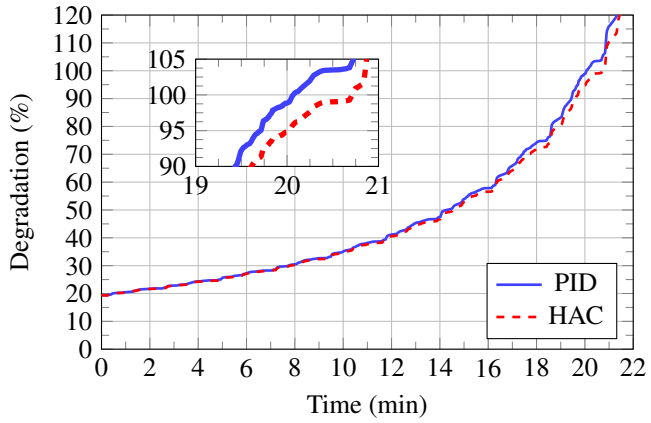


Figure 6: Simulated degradation trajectories of PID-control and HAC in accelerated degradation conditions.

Figure 6 shows how the degradation evolved with the two controllers. As the pump degradation was greatly accelerated, no major differences accumulated between the control methods. The EHA under PID-contol lasted 1203.894 s before the pump reached the breakdown degradation. The HAC system in turn lasted 1242.660 s, which was an increase of 3.22 %.

While the observed RUL extension of 3.22 % may appear modest, it holds significant practical implications. For example, in off-road mobile machinery that operates 8 hours daily, a 3.22 % RUL increase may translate into several additional operating cycles before failure. This can delay costly unscheduled maintenance, increase uptime, and improve planning for component replacement. In critical systems, even minor improvements in RUL can help avoid cascading failures that might compromise safety. Over an extended operational lifespan, incremental improvements accumulate, enhancing overall system reliability and efficiency. Additionally, extended RUL contributes to operational safety by mitigating the risk of sudden component failure, which can prevent hazardous scenarios.

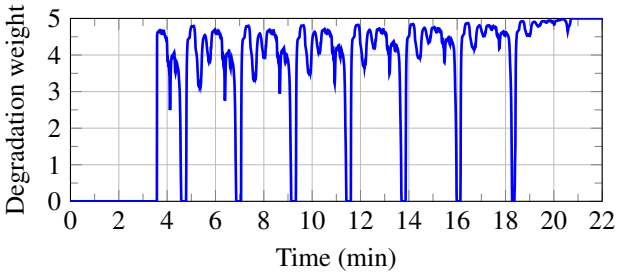


Figure 7: Degradation error weight values.

The HAC adjusted the weight of the degradation based on the RUL estimate. The RUL estimate was smoothed with moving average filter applied to the rotational speed of the pump, with window length of 20 s. Degradation weight is plotted in fig. 7. Initially, with minimal degradation, the HAC prioritized position control exclusively. As degradation increased, the HAC system adjusted pump usage to extend the RUL. When the load on the pump drops, so that the RUL estimate exceeds the requested useful life of 10 cycles, the weight also drops to zero. This strategic trade-off between degradation mitigation and position accuracy highlights a crucial consideration in real-world scenarios, where preserving equipment lifespan may necessitate temporary reduction in control precision. In systems where small deviations in actuator position remain within acceptable performance bounds, this trade-off is justified by the benefit of delayed component wear. However, for applications with tighter precision requirements, further tuning might be needed.

The control error of PID-control and HAC was compared with

$$\text{difference} = |\text{HAC error}| - |\text{PID error}|\,. \tag{40}$$

As seen from fig. 8, the difference in control errors for the position were mostly under 0.5 cm. The difference was closest to zero when the degradation weight from fig. 7 was zero. As the pump degraded, HAC system started to use the pump less to extend the RUL causing the control error to rise. This is can also be seen from fig. 8 during the last 5 min of the simulation.

The results demonstrate that the proposed HAC system effectively mitigates degradation effects and extends the RUL of the gear pump. Under accelerated degradation conditions, the observed RUL extension was comparable to that reported by Zhang et al. [6], who achieved an 8.3 s increase. However, while both studies show improvements in component lifespan, Zhang et al. focused on a different application area, whereas this study specifically addresses EHAs and the degradation of the gear pump.

The degradation model successfully captured the increasing

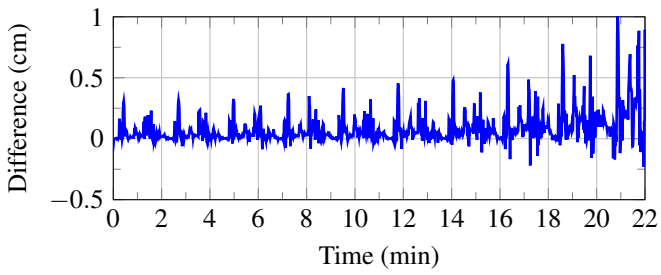


Figure 8: Difference in control errors of HAC and PID.

trend of internal leakage, validating its suitability for estimating wear progression in the gear pump. However, the test data used to calibrate the model was limited in scope and did not represent the full operational range. The simulation relied on accelerated degradation conditions, which may not fully replicate the usage patterns and variability encountered in real-world operations. This introduces uncertainty in how well the simulated results generalize to practical scenarios. Furthermore, the use of a nonlinear Wiener process with static parameters assumes a fixed degradation rate throughout the simulation. This may not accurately reflect the dynamic and often nonlinear nature of wear progression under varying loads, temperatures, or hydraulic fluid conditions. Wang et al. [21] emphasize the value of adaptive Wiener processes that update model parameters in real time to better capture evolving failure modes.

While the increase of 3.22 % may seem modest, it has significant practical implications. In real-world applications, even a small extension in RUL can lead to substantial improvements in operational safety and cost efficiency. For instance, extending the RUL by 3.22 % can reduce the frequency of maintenance interventions, thereby lowering maintenance costs and minimizing downtime. Over the lifespan of the equipment, small incremental improvements in RUL can accumulate, resulting in substantial long-term benefits. This is particularly important in industries where equipment availability is critical. Additionally, improved RUL can enhance operational safety by reducing the likelihood of unexpected failures, which can lead to hazardous situations.

5 Conclusions

This paper presented a HAC approach for an EHA system, aiming to extend the RUL of a gear pump. The proposed method integrates degradation modeling, real-time prognostics, and MPC to balance system performance and component longevity.

The degradation model captures internal leakage progression using a nonlinear Wiener process, estimated online with an EKF. The estimated degradation state is incorporated into the control design, allowing dynamic adjustment of control actions based on RUL predictions. The effectiveness of the HAC approach was evaluated through simulations, compared against a baseline PID-controller.

The results show that the HAC system successfully mitigates degradation effects, increasing RUL by 3.22 % under accelerated conditions. While the improvement is modest due to the accelerated scenario, the approach shows strong potential for real-world applications with gradual degradation. Additionally, the HAC maintains acceptable position control accuracy while reducing degradation impact.

The method relies on certain assumptions, such as constant rotational speed during RUL estimation and simplified electric motor dynamics, which may affect reliability. Deploying the system in real-world applications may also introduce challenges, including computational demands for real-time optimization and the need for robust controller tuning under varying conditions.

Future work will focus on experimental validation using physical test rigs, further refinement of the degradation model, and assessment of scalability and robustness across different system architectures. In addition, investigating the scalability and robustness of the control framework for diverse system architectures would enhance its industrial applicability.

Overall, the HAC strategy offers a promising pathway to extend the life of safety-critical actuators while preserving performance. With further development, it could become an essential component of advanced control systems that enable real-time decision-making and predictive maintenance, supporting sustainability and reliability goals in off-road applications.

Acknowledgment

We acknowledge the financial support of the Finnish Ministry of Education and Culture through the Intelligent Work Machines Doctoral Education Pilot Program (IWM VN/3137/2024-OKM-4).

References

[1] T. Escobet, J. Quevedo, V. Puig, and F. Nejjari. Combining health monitoring and control. In *Diagnostics and Prognostics of Engineering Systems: Methods and Techniques*, pages 230–255. IGI Global Scientific Publishing, 2013.

[2] T. Escobet, V. Puig, and F. Nejjari. Health Aware Control and model-based Prognosis. In *2012 20th Mediterranean Conference on Control & Automation (MED)*, pages 691–696, 2012.

[3] Thomas Sendelbach and Tobias Leutbecher. Model Predictive Control of Electro-Hydraulic Systems with Multiple Degrees of Freedom. In *14th International Fluid Power Conference*, Dresden, Germany, 2024. River Publishers.

[4] Eduardo Bento Pereira, Roberto Kawakami Harrop Galvão, and Takashi Yoneyama. Model Predictive Control using Prognosis and Health Monitoring of actuators. In *2010 IEEE International Symposium on Industrial Electronics*, pages 237–243, July 2010. ISSN: 2163-5145.

[5] Amirhossein Hosseinzadeh Dadash and Niclas Björsell. A framework for designing a degradation-aware controller based on empirical estimation of the state–action cost and model predictive control. *Journal of Manufacturing Systems*, 76:599–613, October 2024.

[6] Jianchun Zhang, Tianyu Liu, and Jianzhong Qiao. Solving a reliability-performance balancing problem for control systems with degrading actuators under model predictive control framework. *Journal of the Franklin Institute*, 359(9):4260–4287, June 2022.

[7] Christophe Letot and Pierre Dehombreux. Degradation models for reliability estimation and mean residual lifetime. VUB, Brussels, Belgium, May 2009.

[8] Zhonghai Ma, Haitao Liao, Jianhang Gao, Songlin Nie, and Yugang Geng. Physics-Informed Machine Learning for Degradation Modeling of an Electro-Hydrostatic Actuator System. *Reliability Engineering & System Safety*, 229:108898, January 2023.

[9] Marta Zagorowska, Ouyang Wu, James R. Ottewill, Marcus Reble, and Nina F. Thornhill. A survey of models of degradation for control applications. *Annual Reviews in Control*, 50:150–173, January 2020.

[10] Min Gu, Shuai Wu, Chunfang Li, Zongxia Jiao, and Tao Yang. Multi-discipline simulation of electro-hydrostatic actuator with Modelica. In *CSAA/IET International Conference on Aircraft Utility Systems (AUS 2018)*, pages 1499–1504, June 2018.

[11] Anik Kumar Samanta, Shrinivas Kulkarni, and Nitin Hande. Predictive Maintenance for Axial Piston Pumps: A Novel Method for Real-Time Health Monitoring and Remaining Useful Life Estimation. In *14th International Fluid Power Conference*, Dresden, Germany, 2024. River Publishers.

[12] Ajinkya Pawar, Andrea Vacca, and Manuel Rigosi. Prediction of Housing Wear-in in External Gear Machines Considering Deformation Effects. In *ASME/BATH 2023 Symposium on Fluid Power and Motion Control*, page V001T01A034, Sarasota, Florida, USA, October 2023. American Society of Mechanical Engineers.

[13] Syed Abu Nahian, Dinh Quang Truong, Puja Chowdhury, Debdatta Das, and Kyoung Kwan Ahn. Modeling and fault tolerant control of an electro-hydraulic actuator. *International Journal of Precision Engineering and Manufacturing*, 17(10):1285–1297, October 2016.

[14] Vinay Partap Singh, Mikko Huova, and Tatiana Minav. Energy Efficient and Redundant Steer-by-Wire for Articulated Non-road Mobile Machines. *International Journal of Fluid Power*, pages 291–324, October 2024.

[15] Saeid R. Habibi and Gurwinder Singh. Derivation of Design Requirements for Optimization of a High Performance Hydrostatic Actuation System. *International Journal of Fluid Power*, January 2000. Publisher: Taylor & Francis.

[16] Y. Chinniah, R. Burton, and S. Habibi. Failure monitoring in a high performance hydrostatic actuation system using the extended Kalman filter. *Mechatronics*, 16(10):643–653, December 2006.

[17] Ahmed Khelassi, Didier Theilliol, Philippe Weber, and Dominique Sauter. A novel active fault tolerant control design with respect to actuators reliability. In *2011 50th IEEE Conference on Decision and Control and European Control Conference*, pages 2269–2274, December 2011. ISSN: 0743-1546.

[18] J.M. Grosso, C. Ocampo-Martinez, and V. Puig. A service reliability model predictive control with dynamic safety stocks and actuators health monitoring for drinking water networks. In *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, pages 4568–4573, December 2012. ISSN: 0743-1546.

[19] Juan M. Grosso, Carlos Ocampo-Martinez, and Vicenç Puig. Reliability–based economic model predictive control for generalised flow–based networks including actuators' health–aware capabilities. *International Journal of Applied Mathematics and Computer Science*, 26(3):641–654, September 2016.

[20] P. Zhang, W. Jiang, X. Shi, and S. Zhang. Remaining Useful Life Prediction of Gear Pump Based on Deep Sparse Autoencoders and Multilayer Bidirectional Long–Short–Term Memory Network. *Processes*, 10(12), 2022.

[21] Han Wang, Xiaobing Ma, and Yu Zhao. An improved Wiener process model with adaptive drift and diffusion for online remaining useful life prediction. *Mechanical Systems and Signal Processing*, 127:370–387, July 2019.

[22] Simo Käki, Vihtori Sova, Luca Romagnuolo, Abid Abdul Azeez, Tatiana Minav, and Simone Bulleri. Experimental Method to Estimate Remaining Useful Life of Gear Pumps. In *The 11th International Conference on Fluid Power Transmission and Control (ICFP 2025) Conference Proceedings*, pages 160–166, Hangzhou, China, April 2025.

[23] ISO 4406-2021, January 2021.

[24] Dale E. Seborg, Thomas F. Edgar, Duncan A. Mellichamp, and Francis J. Doyle. *Process Dynamics and Control, 4th Edition*. Wiley, Hoboken, N.J, fourth edition edition, 2017.

[25] Karl J. Åström and Tore Hägglund. *Advanced PID control*. ISA - The Instrumentation, Systems, and Automation Soc, Research Triangle Park, NC, 2006.

Experimental Investigation of Control Strategies for an Asymmetric Cylinder with Two Individually Controlled Pumps

Samuel Kärnell[1], Axel Jemsson[1], Simon Johansson[1], and Liselott Ericson[1]

[1]Division of Fluid and Mechatronic Systems, Linköping University, Linköping, Sweden
E-mail: samuel.karnell@liu.se

Abstract

This paper investigates various control strategies for an electro-hydraulic actuator (EHA) equipped with two speed-controlled pumps, enabling simultaneous control of speed and pressure. The study examines the design and implementation of several control structures, including PI-controllers with feed-forward and different decoupling techniques, fuzzy logic controllers, optimal state feedback, and reinforcement learning. Initially, all control strategies were tested in a simulation environment, followed by validation in a laboratory test rig. The results demonstrate that even relatively simple control strategies can achieve satisfactory performance. Velocity feed-forward combined with a decoupled PI-controller for pressure showed strong performance and was easy to tune. Incorporation of fuzzy logic and virtual states could improve the performance further, but this came with additional efforts in design and tuning.

Keywords: Electro-hydraulic actuator, electrification, control

1 Introduction

As a way to reduce the CO2 emissions, the electrification of mobile machinery is ongoing. This provides incentives to improve energy efficiency as well as allowing for new system designs. Solutions with electro-hydraulic actuators (EHAs) are commonly discussed as a way forward. However, there are plenty of ways of realising EHAs, especially for asymmetrical cylinders since the differential flow can be handled in several ways. To get an idea of this, see [1]. This paper focuses on a concept that uses two individually controlled hydraulic units to take care of the differential flow. The layout of the concept and how it works in different modes of operation is illustrated in Fig. 1. As the figure shows, the hydraulic pumps can work in both pump and motor modes. However, for simplicity, from here on they will be referred to as pumps.

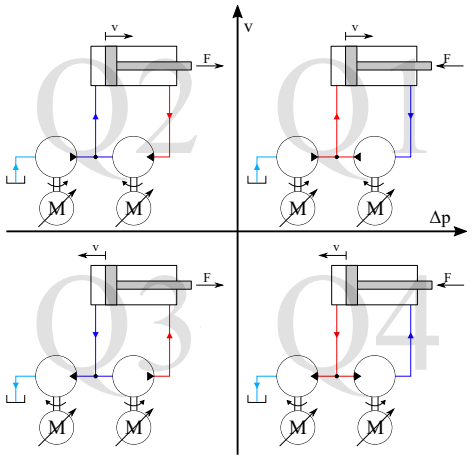


Figure 1: The concept studied in this paper and how it operates in different modes.

1.1 State of the art

A feature of the concept that is presented here is that it enables simultaneous pressure and speed (or position) control. However, the concept is not novel; a variant was patented in the early 90s [2]. At that time, it was proposed to have the pumps connected to the same shaft and to use the displacement setting for the control. The concept has regained attention in recent years, but now it is pumps with their own speed-controlled electric motors that are primarily getting attention.

The system is a multiple-input multiple-output (MIMO) system. The inputs are the signals to the electric motors (typically speed or torque) and the outputs are actuator speed (or position) and pressure (typically the low-pressure side of the actuator). Most work related to this concept is based on a specific kind of decoupled controller [3–6], but other simpler approaches have also been proposed [7]. Still, literature on practical implementation of controllers is limited even though examples can be found [8].

1.2 Contributions

In this paper, several control strategies are presented and evaluated in a physical test rig. The paper is based on the master's thesis by Jemsson and Johansson [9] and can be considered a condensed version.

2 System description

A detailed circuit diagram of the hydraulic system used for evaluation is shown in Fig. 2. The system was developed and built by Gnutti Carlo Sweden AB. Check valves and pressure relief valves are installed to prevent cavitation and overpressure. Electrically controlled load holding valves are implemented to unload the pumps when the cylinder is intended to be stationary. The pressure source shown in the figure was realised by an additional supply pump, providing about 5.5 bar.

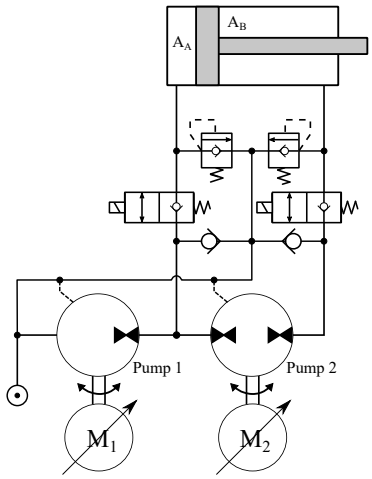


Figure 2: Detailed circuit diagram for the analysed system.

The circuit was implemented on the stick cylinder of an excavator arm in a lab environment. Weights were used as load. The pump and valve unit was placed on the side. The rig is shown in Fig. 3 and relevant system parameters are presented in Table 1.

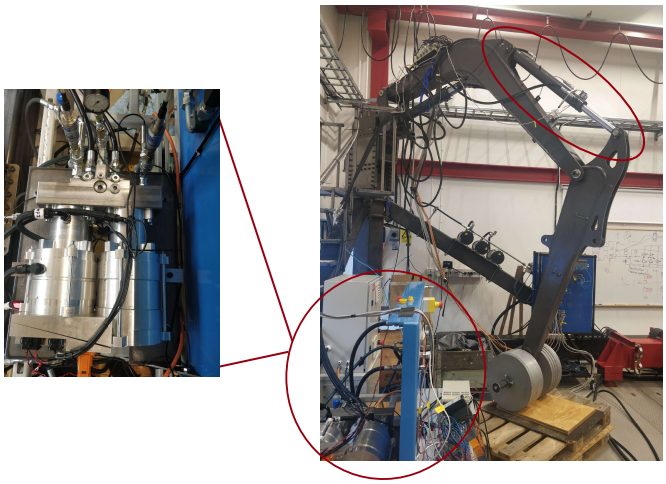


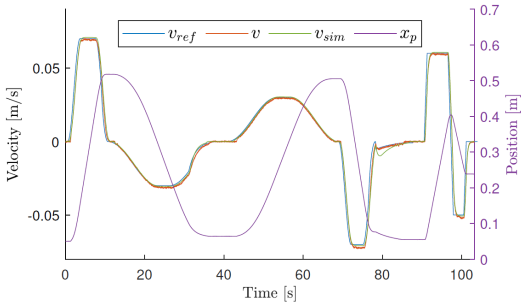
Figure 3: The test rig developed by Gnutti Carlo Sweden AB.

3 Modelling

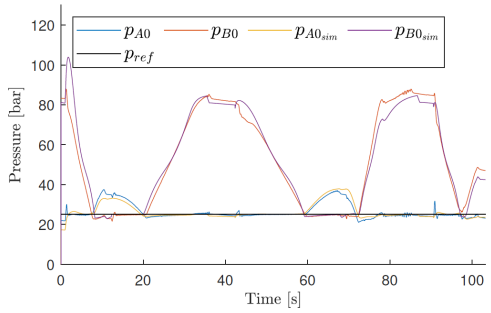
A Simscape model of the system was created to allow for design, tuning, and testing of controllers before evaluation in the rig. Model parameters were adapted to measurement data. The model fit is exemplified in Fig. 4.

Table 1: System parameters.

	Symbol	Value	Unit
Cylinder stroke	$x_{\max}$	0.6	m
Cylinder linearization point	x_0	0.3	m
Cylinder area A side	A_A	0.0071	m^2
Cylinder area B side	A_B	0.0020	m^2
Hose volume A side	V_{Ah}	0.002	m^3
Hose volume B side	V_{Bh}	0.002	m^3
Displacement pump 1	D_1	14	cm^3/rev
Displacement pump 2	D_2	7	cm^3/rev
Tank pressure	p_t	5.5	bar
Maximum stroke	$x_{\max}$	0.6	m
Bulk modulus	β_e	1	MPa



(a) Piston velocities.



(b) Pressures measured at the cylinder.

Figure 4: Comparison of Simscape model and measurements.

In addition to the Simscape model, a linear model, which most of the controllers were based on, was formulated using the following equations:

$$m_{eq}\dot{v}_p = p_{A0}A_A - p_{B0}A_B - Bv_p - F, \tag{1}$$

$$\dot{p}_{A0} = \frac{\beta_e}{V_{A0}}\left(D_1 n_{p1} + D_2 n_{p2} - v_p A_A\right), \tag{2}$$

$$\dot{p}_{B0} = \frac{\beta_e}{V_{B0}}\left(-D_2 n_{p2} + v_p A_B\right). \tag{3}$$

p_A and p_B are the pressures on the A and B sides respectively, and index 0 represents a linearisation point. The equations above can be rewritten in state-space form as follows:

$$\dot{x}_M = A_M x_M + B_M u_M, \quad y_M = C_M x_M \tag{4}$$

where

$$x_M = \begin{bmatrix} v_p & p_{A0} & p_{B0} \end{bmatrix}^T, \quad u_M = \begin{bmatrix} n_{p1} & n_{p2} \end{bmatrix}^T,$$

$$A_M = \begin{bmatrix} -\frac{B}{m_{eq}} & \frac{A_A}{m_{eq}} & -\frac{A_B}{m_{eq}} \\ -\frac{A_A \beta_e}{V_A} & 0 & 0 \\ \frac{A_B \beta_e}{V_B} & 0 & 0 \end{bmatrix},$$

$$B_M = \begin{bmatrix} 0 & 0 \\ \frac{D_1 \beta_e}{V_A} & \frac{D_2 \beta_e}{V_A} \\ 0 & -\frac{D_2 \beta_e}{V_B} \end{bmatrix},$$

$$C_M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

4 Control strategies

The following control strategies have been evaluated:

- PI-controller with feed-forward (section 4.1)
- Decoupled PI-controller with feed-forward (section 4.2)
- Fuzzy logic PI-controller with feed-forward (section 4.3)
- Decoupled PI-controller with virtual states (section 4.4)
- Linear-quadratic-integral control (section4.5)
- Linear-quadratic regulator with virtual states (section4.6)
- Feed-forward with reinforcement learning (section4.7)

All controllers aim to control cylinder speed and pressure on the low-pressure side. They will only be briefly described here. For a more comprehensive description of the controllers, see [9].

4.1 PI-controller with feed-forward

This can be considered the simplest controller of the ones presented in this paper. It utilises the relationships between pump speeds and cylinder velocity in the continuity equations **??** and **??**, neglecting pressure dynamics. The structure of the controller is shown in Fig. 5. In [7], it was shown that the first pump should be used to control the pressure when working in Q2 and Q3, whilst the second should be used when working in Q1 and Q4. However, that introduces a switch case which can cause oscillations. Therefore, the alternative to use the first pump to always control the pressure has also been investigated. Using the second pump has not been considered since that is more prone to instability in certain modes.

The pressure controller is a PI-controller, and the velocity controller is simply a gain of 1. The W-matrix is set to W_{M1} when

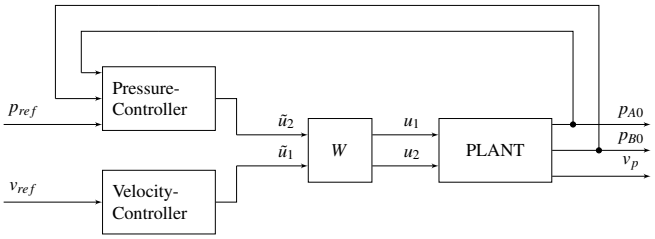


Figure 5: PI-controller with feed-forward.

p_{A0} is controlled and W_{M1} or W_{M2} when p_{B0} is controlled (depending on whether the first or second pump is used for the pressure control).

$$W_{M1} = \begin{bmatrix} \frac{A_A - A_B}{D_1} & 1 \\ \frac{A_B}{D_2} & 0 \end{bmatrix}, \quad W_{M2} = \begin{bmatrix} \frac{A_A - A_B}{D_1} & 0 \\ \frac{A_B}{D_2} & -1 \end{bmatrix}.$$

4.2 Decoupled PI-controller with feed-forward

In contrast to the controller described in Section 4.1, this controller partially decouples the pressure from the velocity by utilising the pressure build-up in the continuity equations. This has an impact on the W-matrix, which now depends on the volumes and is therefore updated based on the cylinder position. The W-matrix becomes as follows:

$$W = \begin{bmatrix} \frac{A_A - A_B}{D_1} & \frac{\frac{A_B}{A_A} V_A + V_B}{\beta_e D_1} \\ \frac{A_B}{D_2} & -\frac{V_B}{\beta_e D_2} \end{bmatrix}.$$

The structure of the controller is shown in Fig. 6.

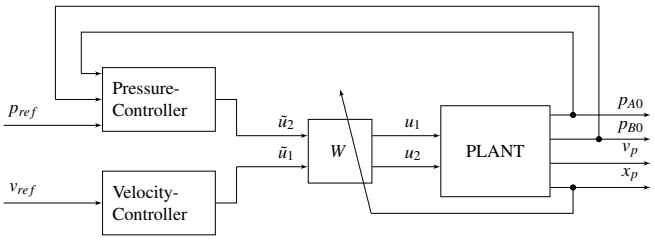


Figure 6: Decoupled PI-controller.

4.3 Fuzzy logic PI-controller with feed-forward

This controller is an extension of the decoupled PI-controller presented in Section 4.2. It uses so-called fuzzy logic to adjust the gains of the PI-controller. The core of the fuzzy logic lies in the rule table, which specifies how the output should be based on the input. In this case, the pressure error (e) and the derivative of the error (de) are inputs (called fuzzy variables). The rule table developed for the fuzzy logic is shown in Table 2. **Z** represents zero, **VS** represents very small, **S** represents small, **M** represents medium, and **B** represents big. Based on these different error inputs, the rule table determines the output values for both the proportional gain and the integral gain. These output values are also abbreviated, where **VS** means very small, **S** means small, **M** means medium, **B**

means big, and **VB** means very big. For information on how the outputs from the table are translated into controller gains, see [9].

$\|e\|\backslash\|de\|$	Z	VS	S	M	B
Z	(S,VB)	(S,VB)	(VS,VB)	(VS,VB)	(VS,VB)
VS	(S,VB)	(S,VB)	(S,VB)	(VS,B)	(VS,B)
S	(M,B)	(M,B)	(M,B)	(S,M)	(S,M)
M	(B,VS)	(B,VS)	(B,S)	(M,S)	(M,S)
B	(VB,S)	(VB,S)	(VB,S)	(VB,S)	(B,S)

Table 2: Rule table.

The structure of the controller is shown in Fig. 7.

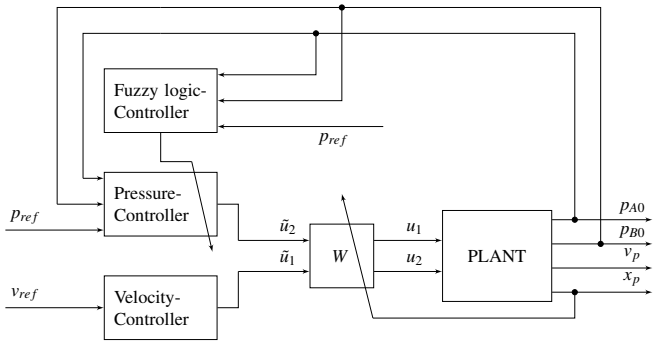


Figure 7: Decoupled fuzzy PI-controller.

4.4 Decoupled PI-controller with virtual states

This controller is based on the one presented in [3]. It utilises virtual states to facilitate full decoupling. The structure of the controller is shown in Fig. 8.

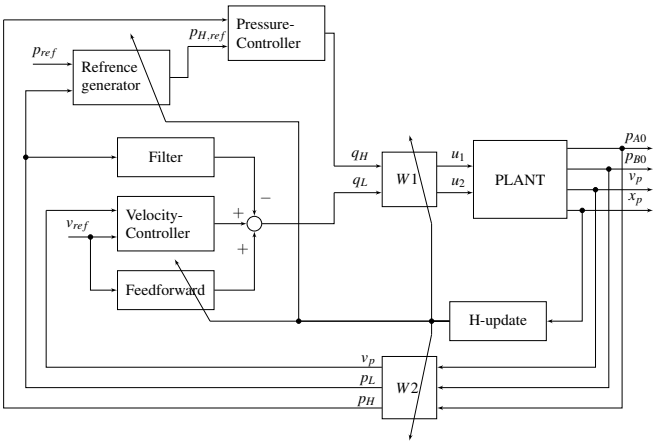


Figure 8: Decoupled controller with virtual states.

The W-matrices are:

$$W_1 = \begin{bmatrix} D_1 & D_2\left(1+\frac{1}{H}\right) \\ D_1 & D_2\left(1-\frac{1}{\alpha}\right) \end{bmatrix}^{-1}, \quad W_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -\alpha \\ 0 & 1 & H \end{bmatrix},$$

where $H = \rho/\alpha$, $\rho = V_B/V_A$ and $\alpha = A_B/A_A$.

4.5 Linear-quadratic-integral control

Modern control often relies on optimal state feedback controllers, with the linear-quadratic regulator being a commonly used example. Here, two variants were implemented and evaluated. One of them was based on so-called linear-quadratic-integral control (LQI), in which the error is integrated. Its structure is shown in Fig. 9.

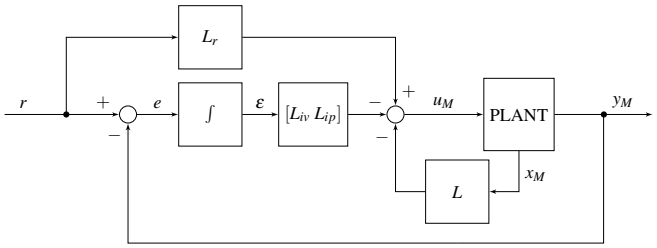


Figure 9: LQI controller.

4.6 Linear-quadratic regulator with virtual states

The other optimal state feedback controller that was tested was a linear-quadratic regulator (LQR) that utilises virtual states, inspired by the controller presented in Section 4.4. This to account for disturbances. It is shown in Fig. 10.

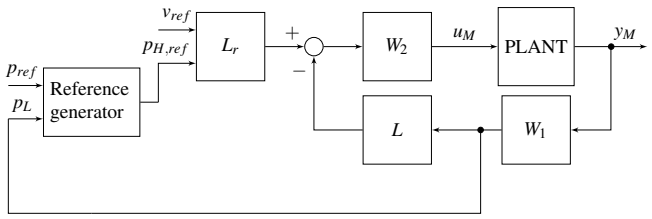


Figure 10: LQR controller.

4.7 Feed-forward with reinforcement learning

The last controller to be evaluated was based on reinforcement learning (RL) together with feed-forward. Its structure is shown in Fig. 11. The reinforcement used the so-called twin-delayed deep deterministic (TD3) policy gradient algorithm. It was trained on the Simscape model. The feed-forward was the same as in the PI-controller presented in section 4.1.

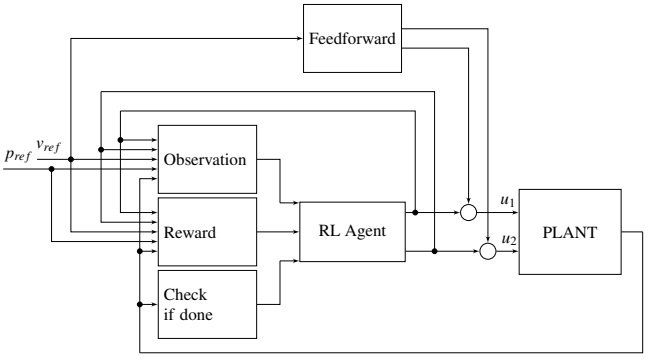


Figure 11: RL-based controller.

5 Results

The different control strategies were evaluated in the test rig described in section 2 using a drive cycle that forces the EHA to work in all four modes of operation. Results for the controllers described above are presented below in Fig. 12 to 19.

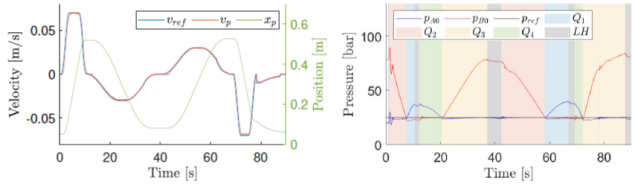


Figure 12: PI-controller (using pump 1 for pressure control).

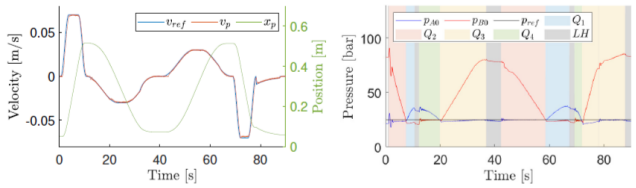


Figure 13: PI-controller.

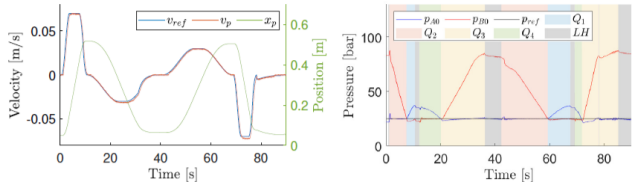


Figure 14: Decoupled PI-controller.

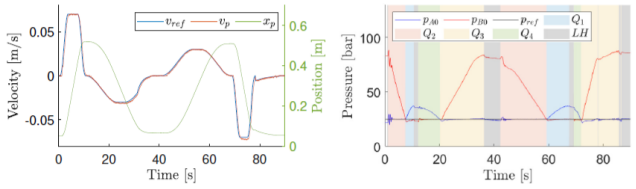


Figure 15: Decoupled fuzzy logic PI-controller.

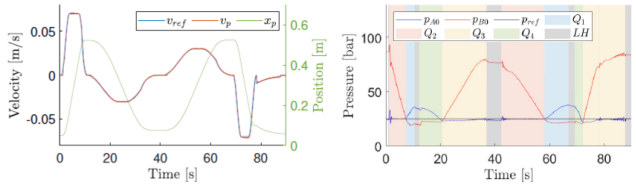


Figure 16: Decoupled PI-controller with virtual states.

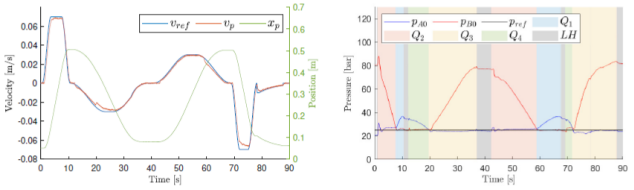


Figure 17: LQI controller.

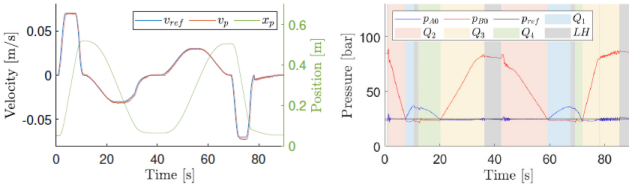


Figure 18: LQR controller with virtual states.

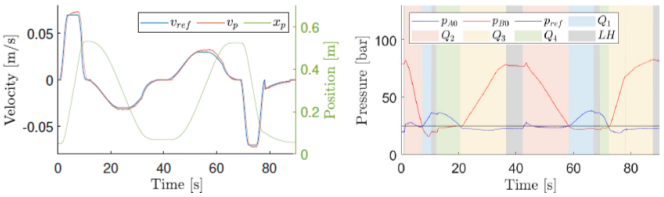


Figure 19: Reinforcement learning-based controller.

It can be seen that all controllers perform quite similarly. They manage to track the velocity and pressure fairly well. To better show how they differ, the root mean square error (RMSE) for the velocity and pressure is presented in Table 3. It also shows the RMSE values for a step response (again, see [9] for details). In addition, a somewhat subjective measure of "tuneability" is included, where "A" is easy and "F" is difficult.

	PI (M1)	PI	Dec.	Fuz.	Vrt. st.	LQI	LQR	RL
V. [mm/s]	3.1	2.7	3.2	3.2	2.1	3.3	3.2	4.1
P. [bar]	1.22	1.53	1.04	0.85	2.19	1.20	1.12	2.86
V. step [mm/s]	9.9	9.7	9.5	9.6	7.3	9.9	9.5	12.0
P. step [bar]	7.48	8.95	1.41	1.20	1.49	2.79	1.32	3.48
Tuneability	B	B	A	D	D	D	C	F

Table 3: RMSE values for different controllers. The three best controllers in each category are colour-coded, where gold is the best, silver is second, and bronze is third.

When examining the table, it can be noted that the controller using virtual states has the best velocity tracking, and that the fuzzy logic has best pressure tracking. However, they are considered to require considerably higher tuning effort than e.g. the decoupled PI-controller, which in fact performs quite similar. Even the simple feed-forward controllers perform somewhat well, with the exception of the pressure control during the step response.

6 Discussion

Satisfactory results were achieved with all controllers. This means that the simpler controllers can likely be considered good enough in many applications. However, the electric motors that were used in the test rig had slow dynamics. For faster systems, it is likely that larger differences would emerge. It should also be mentioned that the result do not say too much about the full potential of the controllers since the tuning was performed manually. Nonetheless, it gives an insight into the level of performance achievable with reasonable tuning and modelling effort.

Furthermore, an advantage of the simpler controllers is their predictability. It is easy to understand what they are doing and why. This makes them more pleasant to run in the rig from an operability point of view.

Another thing to note is that all controllers, except the simplest feed-forward controller, relied on position and/or velocity feedback. Eliminating the need for such feedback can be considered an advantage since fewer sensors are required.

7 Conclusions

Feed-forward with a decoupled PI-controller can achieve good performance with low effort. Better performance can be achieved by using fuzzy logic to set the controller gains, but at the cost of increased controller complexity and tuning effort. Similarly, decoupling using virtual states also improves performance but requires additional tuning.

Acknowledgements

This research was funded by the Swedish Energy Agency (Energimyndigheten) with grant number P2023-00594. Gnutti Carlo Sweden AB have supported with hardware for the rig.

References

[1] Søren Ketelsen, Damiano Padovani, Torben O Andersen, Morten Kjeld Ebbesen, and Lasse Schmidt. Classification and review of pump-controlled differential cylinder drives. *Energies*, 12(7):1293, 2019.

[2] Jorg Dantlgraber. Hydraulic system for a differential piston type cylinder, January 19 1993. US Patent 5,179,836.

[3] Søren Ketelsen, Torben Ole Andersen, Morten Kjeld Ebbesen, and Lasse Schmidt. A self-contained cylinder drive with indirectly controlled hydraulic lock. *Modeling, Identification and Control*, 41(3):185–205, October 2020.

[4] Viktor Donkov, Søren Ketelsen, Torben Andersen, Morten Kjeld Ebbesen, and Lasse Schmidt. Improving the performance of self-contained cylinder drives with load-holding capabilities. In *Scandinavian International Conference on Fluid Power (SICFP21)*, pages 50–61, 2021.

[5] Wei Zhao, Mohit Bhola, Morten K Ebbesen, and Torben Ole Andersen. A novel control design for realizing passive load-holding function on a two-motor-two-pump motor-controlled hydraulic cylinder. *Modelling, Identification, and Control*, 44(3):125–139, 2023.

[6] Wei Zhao, Mohit Bhola, Morten Kjeld Ebbesen, and Torben Ole Andersen. A motor-controlled hydraulic cylinder with cylinder pressure control and load-holding function-an experimental study. In *Fluid Power Systems Technology*, volume 87431.

[7] Samuel Kärnell and Liselott Ericson. Control of an asymmetric cylinder with two individually controlled pump/motors. volume ASME/BATH 2023 Symposium on Fluid Power and Motion Control of *Fluid Power Systems Technology*.

[8] Wei Zhao, Morten Kjeld Ebbesen, Michael Rygaard Hansen, and Torben Ole Andersen. Experimental assessment of energy consumption: a comparative study between motor-controlled and valve-controlled cylinders in knuckle boom crane operations. *Frontiers in Mechanical Engineering*, 11, 2025.

[9] Simon Johansson and Axel Jemsson. Investigation of control strategies for an electro-hydraulic actuator with dual variable-speed pumps: An applied study on an excavator arm with load-holding capabilities. Master's thesis, Linköping University, 2024. LIU-IEI-TEK-A–24/04784–SE.

Automatic Sequential Actuation of Multiple Cylinders in Mobile Cranes Using a Single Electro-Hydraulic Converter

Timofei Komarov , Victor Zhidchenko , and Heikki Handroos

LUT Mechanical Engineering, LUT University, Lappeenranta, Finland
E-mail: timofei.komarov@lut.fi, victor.zhidchenko@lut.fi, heikki.handroos@lut.fi

Abstract

The electrification of actuation systems of mobile cranes presents challenges in adaptating electro-hydraulic (EHA) or electro-mechanical (EMA) actuators. While the implementation of EHAs promises efficiency improvement, conventional hydrostatic systems require a dedicated electro-hydraulic energy converter (EHC) for each actuator. This study proposes a solution to actuate multiple actuators sequentially using a single EHC. The solution comprises two main parts: a hydraulic system featuring switching the connection between the EHC and the cylinders; and a control algorithm for automated crane movement. Performance was evaluated through simulation and compared with two conventional systems: one using a load-sensing system with proportional directional valves, and the other employing independent closed-loop circuits with a dedicated EHC for each cylinder.

Keywords: Automation, Mobile machines, Energy Efficiency, Electro-Hydraulics, Sequential Motion

1 Introduction

In recent decades, the development of mobile machinery has been focused on several key areas. One major area is automation, which is ongoing on several levels. Commercially available solutions pursue the increase of productivity and simplification of machine operation [1], while the remote operation remains in the research phase [2]. Automatic movement has been studied in both vehicle driving [3] and manipulator control [4]. However, the complexity and variability of the work environment remain to be the main challenge on the way to machine automation [5].

Another research direction is electrification of the machine powertrain, driven by the increase of energy efficiency and the reduction of carbon emissions [6]. The substitution of conventional valve-controlled hydraulics with electro-hydraulic actuators was found perspective for increasing energy efficiency of the mobile machines [7]. The common solution is to implement independent closed hydraulic circuits for each machine's actuator [8]. However, this approach requires an substantial number of electro-hydraulic actuators (5-8) for machine operation, posing significant challenges in implementation.

This article revises the approach to actuating multiple cylinders using a single electro-hydraulic converter (EHC) [9]. Previous research proposed switching between cylinders using fast on/off valves, and simulation demonstrated inconvenience of the solution for manual operation: every switch between cylinders caused high pressure oscillations.

To address this limitation, this study introduces path-planning, a switching algorithm, and automatic motion control. The system was evaluated using simulation and compared with a load-sensing system (LS-system) with proportional directional valves (DCVs) and a system with two independent EHC-based circuits.

2 Description of the System Under Study

A PATU655 mobile log crane is a hydraulic manipulator with a typical mechanical structure, presented in fig. 1. The crane comprises a pillar, a lift boom, and a jib (tilt) boom interconnected with rotational joints, and an extension boom connected to the jib by a linear joint.

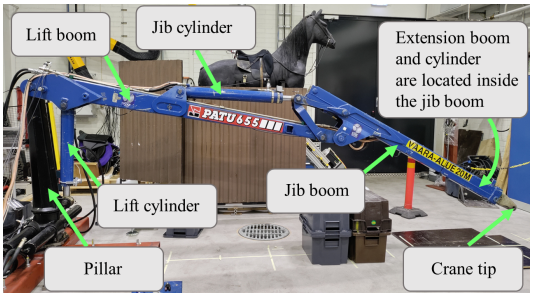


Figure 1: Mobile hydraulic log crane PATU 655, Laboratory of Intelligent Machines at LUT University, 2023 [10].

2.1 Electro-Hydraulic System Architecture

Figure 2 illustrates the proposed actuation system of a mobile log crane. System parameters are listed in tab. 1. The system's fluid power source is an EHC 1 comprising a Parker F11-019 pump [11] and a Parker GVM210-050 electric motor [12]. The crane actuators are hydraulic cylinders 3 and 4. Parker DS162CV on/off valves 2 [13] switch the connection between the EHC and the cylinders. The control system measures pressure in the cylinder chambers and EHC ports using pressure sensors 5. Pressure-relief valves 6 ensure overload protection for the system.

Operating an asymmetric cylinder with an EHC presents a challenge due to the need to compensate for differences in volumetric flow in cylinder chambers. Insufficient flow can cause cavitation, while excessive flow leads to a pressure rise. To address this issue, various solutions and system configurations have been studied [14]. In this research, a low-pressure line (LP-line) maintains a constant pressure of 5 bar by directing flow from the boost pump 8 or to the tank via a Parker DS163BN 3/2 directional control valve 7 [13].

Table 1: System parameters.

Parameter	Value
Lift cylinder piston diameter, mm	100
Lift cylinder rod diameter, mm	56
Lift cylinder stroke, mm	535
Tilt cylinder piston diameter, mm	90
Tilt cylinder rod diameter, mm	56
Tilt cylinder stroke, mm	780
Pump displacement, cm^3/rev	19

In the present research, only the movements of the lift and tilt cylinder are considered. However, the proposed system architecture can be expanded to a larger number of actuators.

3 Automatic Motion Algorithm

The automatic motion of a log crane is achieved through several steps:

1. Planning a path between the given start and end points considering crane kinematics limitations and obstacles within the discretized crane workspace.
2. Building a smooth trajectory along the obtained path.
3. Automatic movement of the hydraulic cylinders along the trajectory.

Considering movement of only the lift and tilt cylinders, the crane workspace is planar. However, the proposed automation approach can be expanded to spatial cases.

3.1 Path-planning Algorithm

A* path-planning algorithm provides a direct and simple solution for finding the shortest path in a weighted graph [15]. The implementation of this algorithm requires workspace discretization, and previous studies have used square cells in Cartesian coordinates to represent possible crane tip positions constrained by the crane kinematics [10].

However, linear tip movements in Cartesian coordinates require simultaneous control of both the lift and tilt booms, which is not possible with the proposed actuation system. To address this issue, the crane workspace was discretized with a step size 0.05 m using cylinder positions S1 and S2 instead of Cartesian coordinates (X,Z). The dependency between the tip position and the cylinder positions is determined by crane's geometry.

Constraining the A* algorithm to move only along coordinates S1 and S2 (excluding diagonal movements) allow crane operation using one cylinder at a time. Hence, the path-planning algorithm provides a sequence of cylinder movements to move between the given start and end points.

3.2 Trajectory Generation Algorithm

Cylinder trajectories are generated based on the sequence of cylinder motions provided by the A* path-planning algorithm. The crane movement consists of two successive phases: cylinder actuation and switching between active cylinders.

The trajectory of each cylinder is formulated using a jerk-bounded quintic polynomial method [16]. This approach ensures smooth trajectories between two given points with minimized movement duration, yet secures velocity, acceleration, and jerk constrains. The limits considered in this study are listed in the tab. 2.

Table 2: Actuators limits.

Parameter	Value
Lift cylinder maximum velocity, m/s	0.1
Lift cylinder maximum acceleration, m/s^2	0.2
Lift cylinder maximum jerk, m/s^3	0.2
Tilt cylinder maximum velocity, m/s	0.15
Tilt cylinder maximum acceleration, m/s^2	0.2
Tilt cylinder maximum jerk, m/s^3	0.2

Switching between cylinders can cause pressure oscillations in the system [9]. In static conditions, the pressure in cylinders is primarily determined by the applied load and depends on the crane configuration. Before switching, the pressure at the EHC ports equals to that of the previously operated cylinder. When the valves 2 open to connect the EHC to the next cylinder, the difference in pressure leads to uncontrolled flow through the valves, resulting in pressure oscillations.

To mitigate the problem, a switching algorithm has been de-

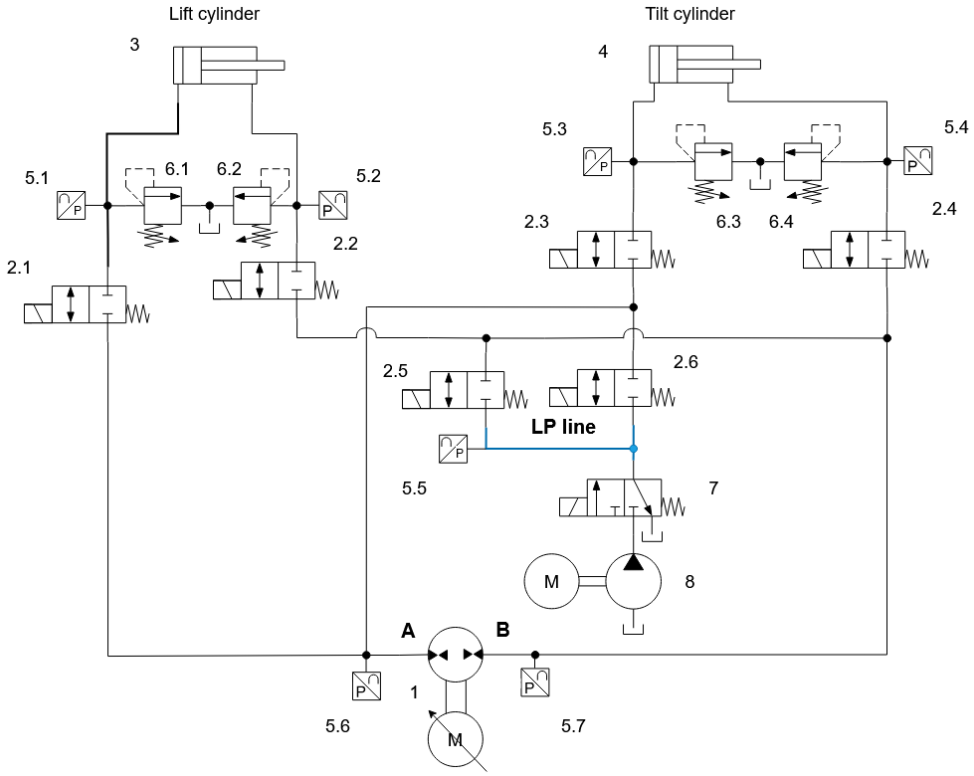


Figure 2: A diagram of the electrohydraulic actuation system featuring sequential control of multiple actuators using a single EHC

veloped. Firstly, the cylinders have to fully stop before switching. Secondly, a 0.5-second delay is introduced, when the EHC is temporarily disconnected from all cylinders.

During this pause, the algorithm identifies the chamber of the next cylinder with the load-holding pressure and sets it as the reference pressure for the corresponding EHC port. At the same time, the second EHC port is connected to the low-pressure line through valves 2.5 or 2.6 and to the low-pressure chamber of the next cylinder through valves 5.2, 5.3, or 5.4. After that, the EHC controls pressure at its ports to match that of the next cylinder. Once the pressure is stabilized, the corresponding on/off valves open, connecting the EHC and the cylinder.

Hence, the trajectory generation algorithm produces reference position, velocity, and acceleration of the cylinders, and regulates the switching order of the on/off valves.

3.3 Movement Control Algorithm

Cylinder movement control is based on the Virtual Decomposition Control (VDC) approach, which calculates the actuator forces needed to achieve the required motion by calculating the manipulator dynamics [17]. This versatile algorithm has been successfully implemented in hydraulic manipulators, demonstrating high accuracy [4].

The reference rotation velocity of the EHC, ω_{ref}, is calculated based on the required pump volumetric flow Q_{ref} and closed-loop proportional force and velocity control:

$$\omega_{ref} = \frac{Q_{ref}}{V_P} + k_F \cdot (F_r - F_{meas}) + k_v \cdot (v_{ref} - v_{meas}) \qquad (1)$$

where V_P is a pump displacement, m^3/rev; F_r and F_{meas} are actuator reference and measured forces, respectively, N; v_{ref} and v_{meas} are cylinder reference and measured velocities, respectively, m/s; and k_F and k_v are proportional coefficients for feedback force and velocity control, manually tuned for suitable performance.

The required pump volumetric flow rate Q_{ref} is determined using the dynamic continuity equation:

$$Q_{ref} = \begin{cases} \frac{\dot{F}_r \cdot x_{meas}}{B_e} + A_A \cdot v_{ref}, & \text{if } G > 0 \\ \frac{\dot{F}_r \cdot (x_{max} - x_{meas})}{B_e} + A_B \cdot v_{ref}, & \text{otherwise} \end{cases} \qquad (2)$$

where B_e is the effective bulk modulus of oil in a chamber, Pa; x_{max} is a cylinder stroke, m; and A_A and A_B are piston-side and rod-side cylinder areas, respectively, m^2; $G \in -1, 1$ corresponds to the direction of the gravity force estimated from the crane position and determines the cylinder chamber with the load-holding pressure. For the lift cylinder, G is not calculated because of its unidirectional load.

The force balance equation is used to estimate the generated cylinder force F_{meas}:

$$F_{meas} = p_A \cdot A_A - p_B \cdot A_B - k_V \cdot v_{ref} \qquad (3)$$

where p_A and p_B are the pressures in the piston-side and rod-side cylinder chambers, respectively, Pa; and k_V is the coefficient of cylinder viscous friction, N/(m/s).

The control system coefficients were found empirically and are listed in the tab. 3.

Table 3: Control system parameters.

Parameter	Value
Lift cylinder k_F	$2 \cdot 10^{-5}$
Lift cylinder k_v	10
Tilt cylinder k_F if $G \geq 0$	$2 \cdot 10^{-5}$
Tilt cylinder k_v if $G \geq 0$	10
Tilt cylinder k_F if $G < 0$	$-1 \cdot 10^{-5}$
Tilt cylinder k_v if $G < 0$	10

4 Simulation

The system performance was analyzed using a simulation in MATLAB/Simulink. The multibody dynamics of the crane was modeled using an iterative Newton-Euler dynamic formulation [18] and verified against the physical crane at LUT University (fig. 1). Since the reference crane is not equipped with a grapple, load swing was not considered, and logs were not explicitly modeled. Instead, a point mass of 180 kg, attached to the crane tip, represented the payload of a single log.

The electro-hydraulic system model incorporates the crane cylinders, the EHC, and on/off valves. The cylinder models neglect the leakages but include a friction model developed in previous research [18]. The EHC model accounts for the first-order dynamics of the electric motor, with a 50 ms rise time, and incorporates a laminar pump leakage model based on the manufacturer data [11]. Estimation of consumed power comprises the calculation of mechanical energy delivered to the pump and accounts for the motor efficiency based on the charts from the datasheet [12]. The on/off valves are modeled using the first-order dynamics and flow characteristics derived form the datasheet [13].

Pressure-relief valves were excluded, as system operation during overload conditions is not considered in this study. To avoid unnecessary model complexity, the LP-line, boost-pump, and 3/2 valve were simplified to a constant pressure source of 5 bar.

The Simulink model incorporates the previously described movement control algorithm and valve-switching logic. Path-planning and trajectory-building algorithms are implemented in MATLAB script.

4.1 Reference Systems

The proposed electro-hydraulic system is compared with two state of the art approaches: an LS-system with DCVs; and a system with an independent EHC for each cylinder. The implementation of the modified A* algorithm for both reference systems to find the shortest path was investigated in a previous study [10]. The movement velocity was adjusted to limit the cycle duration to 11-13 seconds, which represents typical operator performance [19].

4.1.1 LS-system with DCVs

Figure 3 illustrates the hydraulic diagram of the LS-system with DCVs. It uses the same EHC 1 for power supply as in the proposed electro-hydraulic system. The DCVs 4 are Bosch Rexroth 4WRPEH 6 with a rated flow of 40 lpm [20].

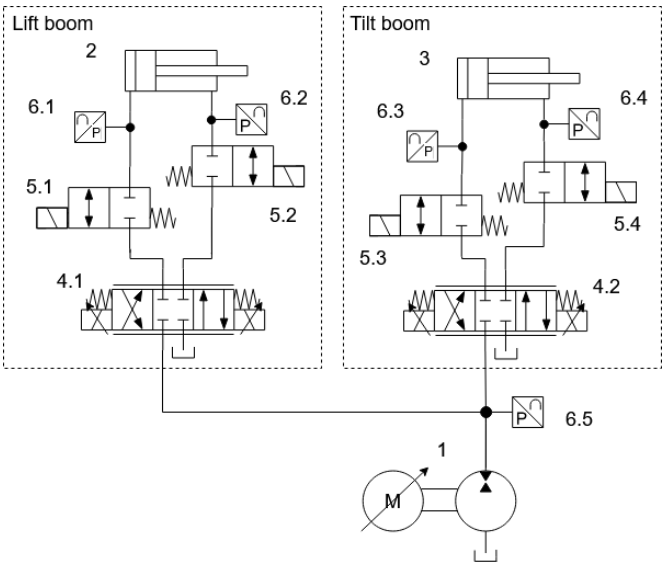


Figure 3: Hydraulic diagram of the LS-system with DCVs

The LS-system maintains the supply pressure at a level greater than the highest pressure in the cylinders by a predefined offset, commonly of 35 bar. Conventionally, this function is controlled hydro-mechanically with a variable-displacement LS-pump [21]. However, in this system, pressure regulation is performed by adjusting the EHC's rotation velocity based on real-time cylinder pressures measured by pressure sensors 6.1-6.4.

An open-loop model-based control is implemented to calculate the EHC reference rotation velocity ω_{ref}:

$$\omega_{ref} = (p_{LS} + \Delta p_{LS} - p_{max}) \cdot k_{LS} + \frac{v_{lift} \cdot A_{lift} + v_{tilt} \cdot A_{tilt}}{V_P} \quad (4)$$

where p_{LS} is the measured supply line pressure, Pa; Δp_{LS} is the pressure offset, Pa; p_{max} is the maximum cylinder pressure, Pa; $k_{LS} = 5$ is a proportional feedback coefficient; A_{lift} and A_{tilt} are cylinder areas depending on the movement direction, m^2.

The cylinder motion control combines three components: an open-loop model-based control U_{MB}; a closed-loop proportional position control; and a closed-loop PD velocity control:

$$U_{DCV} = k_x \cdot (x_{ref} - x_{meas}) + k_{vP} \cdot (v_{ref} - v_{meas}) + k_{vD} \cdot \frac{\mathrm{d}(v_{ref} - v_{meas})}{\mathrm{dt}} + U_{MB} \quad (5)$$

where U_{DCV} is the DCV control voltage, V; $k_x = 30$ is a closed-loop position proportional coefficient; $k_{vP} = 20$ and $k_{vD} = 10$ are closed-loop PD velocity control coefficients.

The model-based control component U_{MB} is defined as:

$$U_{MB} = \frac{Q}{C_V \cdot \sqrt{p_{LS} - p}} \quad (6)$$

where $Q = \dot{x}_{ref} \cdot A$ is the volume flow rate required to achieve the reference cylinder velocity $\dot{x}_{ref}$, m^3/s; and C_V is a valve flow coefficient.

4.1.2 System with independent EHCs

A reference closed-loop circuit with an independent EHC for each cylinder is presented in fig. 4. The illustrated system controls a single cylinder, so the considered crane requires two such systems to actuate lift and tilt cylinders. The model of the system and the control algorithm were developed in the previous research [22]. The on/off valves 5.1 and 5.2 were not modeled because this research does not consider system behavior during overloading conditions.

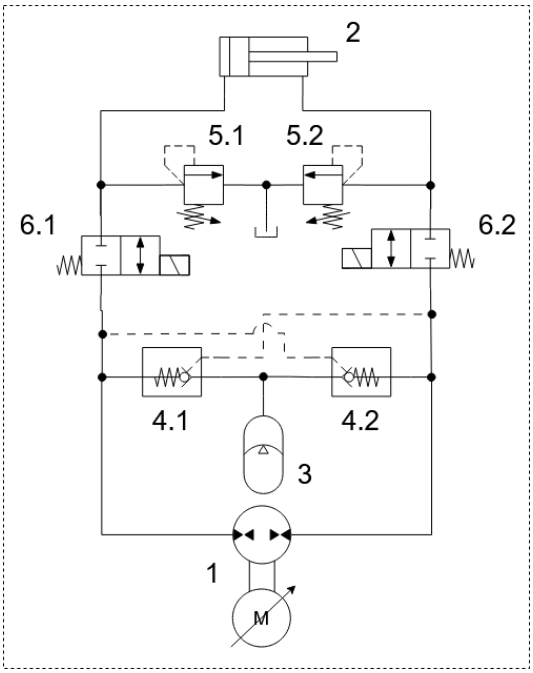


Figure 4: Diagram of the hydraulic system with independent EHCs [22]

4.2 Simulation Conditions

The simulations replicated typical log crane operations: logs loading from the pile into the trailer, and unloading them to the ground. Figure 5 illustrates the crane, an obstacle, and start and end points considered in the simulation. The logs are supposed to be grabbed close to their center of mass for ensuring the balance. Therefore, the brown circles describe positions of the centers of mass of the logs, even though normally the logs are pointed orthogonally to the road. The simulation conditions included a vertical obstacle representing the log trailer bunks.

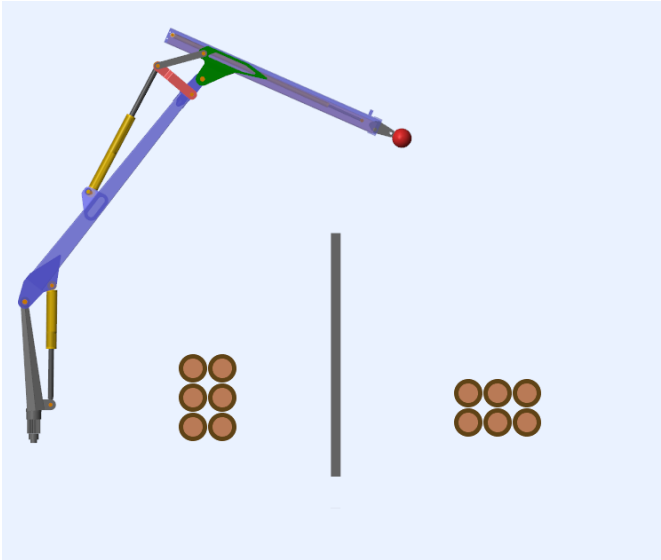


Figure 5: Diagram of the simulated working environment. The gray vertical line represents the obstacle, and brown circles depict log positions

The set of points included 6 positions in the pile and 6 positions in the truck, resulting in 36 combinations of start and end points. For each combination of the points, two shortest paths were found: one with the proposed electro-hydraulic system, and one with reference systems.

5 Results

Figure 6 presents a boxplot comparing the ratio of total energy consumption per cycle during a single simulation of the proposed system (E) to that of the LS-system with DCVs (E^{DCV}) and the system with independent EHCs (E^{EHC}). Summary statics is presented in the tab. 4.

Table 4: Summary statistics of the ratio of total energy consumption per cycle of the proposed system (E) compared to the LS-system with DCVs (E^{DCV}) and the system with independent EHCs (E^{EHC}).

Parameters		$\frac{E}{E^{DCV}}$	$\frac{E}{E^{EHC}}$
Loading logs	75th percentile	31.38 %	75.10 %
	Median	28.33 %	68.50 %
	25th percentile	26.10 %	61.98 %
Unloading logs	75th percentile	35.55 %	48.87 %
	Median	32.37 %	42.10 %
	25th percentile	25.31 %	32.74 %

Figure 7 illustrates the boxplot of a ratio of cycle durations

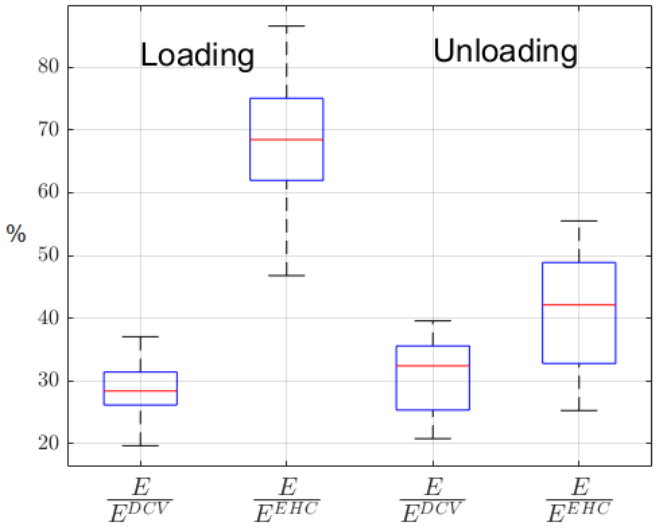


Figure 6: Ratio of total energy consumption per cycle of the proposed system (E) to that of the LS-system with DCVs (E^{DCV}) and the system with independent EHCs (E^{EHC})

during operation of the proposed system (T) to those of the LS-system with DCVs (T^{DCV}) and the system with independent EHCs (T^{EHC}). Table 5 contains the corresponding summary statistics.

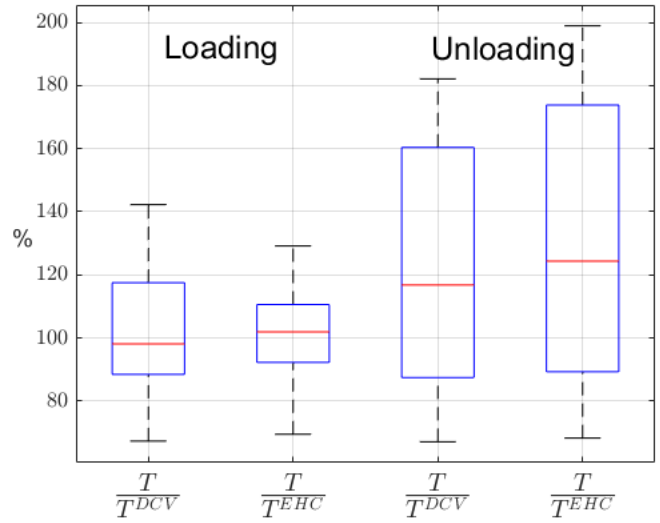


Figure 7: Ratio of cycle durations of the proposed system (T) to those of the LS-system with DCVs (T^{DCV}) and the system with independent EHC (T^{EHC})

An example of the shortest paths found during a simulation is presented at fig. 8. The green and pink circles illustrate the start and final points, respectively, and the green lines represent the starting crane configuration. The trajectory building algorithm transforms the paths into reference commands for actuation systems. Figure 9 demonstrates reference trajectories and actual crane tip positions obtained in a simulation for each actuation system.

Table 5: Summary statistics values of the ratio of cycle durations of the proposed system (E) to those of the LS-system with DCVs (E^{DCV}) and the system with independent EHCs (E^{EHC}).

Parameters		$\frac{T}{T^{DCV}}$	$\frac{T}{T^{EHC}}$
Loading logs	75th percentile	117.50 %	110.57 %
	Median	98.11 %	101.90 %
	25th percentile	88.41 %	92.21 %
Unloading logs	75th percentile	160.37 %	173.78 %
	Median	116.75 %	124.31 %
	25th percentile	87.43 %	89.27 %

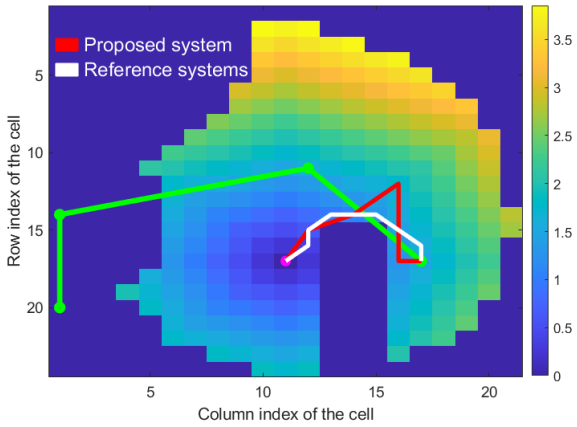


Figure 8: An example of the paths found by the developed path-planning algorithm for the proposed and reference systems

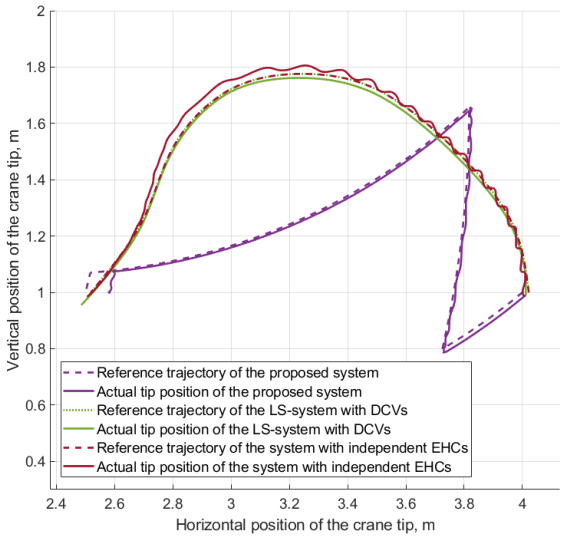


Figure 9: Trajectories generated based on the found paths, and actual crane tip movements when following them

Conclusion

A solution featuring a single EHC for the automatic sequential control of hydraulic cylinders in the log crane was developed in this study. Simulations demonstrated significant reduction in energy consumption with median values reaching 28.33 % and 68.50 % for log loading cycles and 32.37 % and 42.10 % for log unloading cycles compared to the reference LS-system with DCVs and the system with independent EHCs, respectively.

The decrease in energy consumption was achieved due to two key factors. First, controlling only a single actuator at a time simplifies trajectory building and cylinder position control, allowing the EHC to operate in its high-efficiency range. Second, when cylinders are disconnected, they don't consume energy for load holding, unlike systems using DCVs or independent EHCs.

The cycle durations of the proposed system were close to ones of the reference systems during log loading cycles. However, during unloading, the cycle time increased, with median values reaching 116.75 % and 124.31 % compared to the LS-system with DCVs and the system with independent EHCs, respectively. Moreover, the 75th percentile of the duration ratio reached 160.37 % and 173.78 % relative to the reference systems. This significant cycle slowdown is primarily attributed to the high correlation between the number of switches between the cylinders and cycle duration. Every switch introduces a fixed 0.5 s delay for pressure stabilization, and since the cylinders have to stop for switching, movement speed is reduced due to additional acceleration and deceleration phases.

The proposed system provides a wide range of opportunities for further study. First, optimizing the path-planning algorithm to minimize switches between the cylinders could improve work cycle duration. Second, the approach of actuating multiple cylinders with the fewest possible EHCs allows for a wide variety of potential modifications and combinations within actuation systems. Third, incorporating the extension boom introduces redundancy, which can be used for energy efficiency improvement and control optimization. Finally, further investigation into the implementation of the VDC in EHC-based systems is needed for maximizing the system performance and accuracy.

References

[1] The perfect cut. ctl - the cut-to-length method. Ponsse Plc, Ponssentie 22, 74200 Vieremä, 2023.

[2] Aamir Mahmood, Sarder Fakhrul Abedin, Mattias O'Nils, Mats Bergman, and Mikael Gidlund. Remote-timber: An outlook for teleoperated forestry with first 5g measurements. *IEEE Industrial Electronics Magazine*, 17(3):42–53, 2023.

[3] Jouko Kalmari, Juha Backman, and Arto Visala. Coordinated motion of a hydraulic forestry crane and a vehicle using nonlinear model predictive control. *Computers and Electronics in Agriculture*, 133:119–127, 2017.

[4] Janne Koivumäki and Jouni Mattila. High performance nonlinear motion/force controller design for redundant hydraulic construction crane automation. *Automation in Construction*, 51:59–77, 2015.

[5] Volker Waurich, Benjamin Beck, Jürgen Weber, and Frank Will. Levels of digitalization for construction machinery on the connected and automated construction site. In Johannes Fottner, Konrad Nübel, and Dominik Matt, editors, *Construction Logistics, Equipment, and Robotics*, pages 157–165, Cham, 2024. Springer Nature Switzerland.

[6] Antti Lajunen, Panu Sainio, Lasse Laurila, Jenni Pippuri-Mäkeläinen, and Kari Tammi. Overview of powertrain electrification and future scenarios for non-road mobile machinery. *Energies*, 11(5), 2018.

[7] T. A. Minav, C. Bonato, P. Sainio, and M. Pietola. Efficiency of direct driven hydraulic drive for non-road mobile working machines. In *2014 International Conference on Electrical Machines (ICEM)*, pages 2431–2435, 2014.

[8] Daniel Hagen, Damiano Padovani, and Morten K. Ebbesen. Study of a self-contained electro-hydraulic cylinder drive. In *2018 Global Fluid Power Society PhD Symposium (GFPS)*, pages 1–7, 2018.

[9] Teshale Bekele, Rafael Åman, and Heikki Handroos. The control of multiple actuators using single iehec pump/motor. In *10th International Fluid Power Conference (10. IFK)*, 03 2016.

[10] Victor Zhidchenko, Timofei Komarov, Antoine Williot, Nathan Bauer, and Heikki Handroos. A method for planning the trajectory of mobile hydraulic crane booms with a focus on energy efficiency. In *2023 IEEE 21st International Conference on Industrial Informatics (INDIN)*, pages 1–7, 2023.

[11] Parker Hannifin, Mobile Controls Division. *Catalogue HY17-8249/US - Hydraulic Motor/Pump Series F11/F12*, 2004.

[12] Parker Hannifin, Hydraulic Pump and Power Systems. *GVM Global Vehicle Motor Power Series. Permanent Magnet (PMAC) Motor for Mobile Systems*, 2022.

[13] Parker Hannifin, Hydraulic Valve Division. *Catalog 3105/USA. Directional Control Valves. Series DS162*.

[14] Zhongyi Quan, Long Quan, and Jinman Zhang. Review of energy efficient direct pump controlled cylinder electro-hydraulic technology. *Renewable and Sustainable Energy Reviews*, 35:336–346, 2014.

[15] Peter E. Hart, Nils J. Nilsson, and Bertram Raphael. A formal basis for the heuristic determination of minimum cost paths. *IEEE Transactions on Systems Science and Cybernetics*, 4(2):100–107, 1968.

[16] Sonja Macfarlane and Elizabeth Croft. Jerk-bounded manipulator trajectory planning: Design for real-time applications. *Robotics and Automation, IEEE Transactions on*, 19:42 – 52, 03 2003.

[17] Wen-Hong Zhu. *Virtual decomposition control: toward hyper degrees of freedom robots*, volume 60. Springer Science & Business Media, 2010.

[18] Iuliia Malysheva, Heikki Handroos, Victor Zhidchenko, and Alexander Kovartsev. Faster than real-time simulation of a hydraulically actuated log crane. In *2018 Global Fluid Power Society PhD Symposium (GFPS)*, pages 1–6, 2018.

[19] Pedro La Hera and Daniel Ortiz Morales. What do we observe when we equip a forestry crane with motion sensors? *Croat. J. For. Eng.*, 40(2):259–280, August 2019.

[20] Bosch Rexroth AG. *RE 29121*, 2022.

[21] Henrik Pedersen, Torben Andersen, and Michael Hansen. Load sensing systems - a review of the research contributions throughout the last decades. In *Proceedings 4th IFK Workshop*, pages 125–139, 01 2004.

[22] Timofei Komarov, Victor Zhidchenko, and Heikki Handroos. Effect of electrification on the energy efficiency of boom trajectories of semi-autonomous mobile cranes. In *Advancements in Fluid Power Technology: Sustainability, Electrification, and Digitalization*, 08 2024.

Potential of Hydraulic Waste Heat Recovery for Cabin Comfort

Fabian Lagerstedt[1,2], Samuel Kärnell[2], Marcus Rösth[3], and Liselott Ericson[2]

[1]Huddig AB, Hudiksvall, Sweden
E-mail: fabian.lagerstedt@huddig.se, samuel.karnell@liu.se, marcus@rosth.com, liselott.ericson@liu.se
[2]Division of Fluid and Mechatronic Systems, Linköping University, Linköping, Sweden
[3]Hudiksvalls Hydraulikkluster, Hudiksvall, Sweden

Abstract

Electrification of mobile machines requires focus on reducing unnecessary energy consumption, which is essential to avoid the need for oversized energy storage. Waste heat recovery of already existing heat flows in the vehicle, such as the working hydraulic system, is a promising way of reducing the energy consumption needed for keeping the cabin warm in the winter. This paper presents a model of a heat recovery concept from the working hydraulic system to the cabin, different from the ones typically seen in the literature. The model is created with the open-source tool OpenModelica together with the DLR Thermofluid Stream library. Focus of the model was to analyze how the dynamic characteristics of the working hydraulic system affects the cabin temperature, and to evaluate the waste heat performance. The results show that the cabin temperature fluctuates close to detectable variations without any interference of active temperature control, and also indicates promising cabin heating performance even under lower operation intensities.

Keywords: Heat recovery, working hydraulics, ITMS, mobile machine, electric vehicle

1 Introduction

Transitioning from internal combustion engine (ICE) and diesel hybrid into fully electric mobile machines typically means losing a major heat source and energy storage. Increasing the energy efficiency of the vehicle is one way to tackle these issues. One widespread issue with electric vehicles is high energy consumption for cabin heating during the winter, which has been addressed recent years by implementation of waste heat recovery and heat pump systems in integrated thermal management system architectures [1] [2].

Due to the lack of available waste heat on electric passenger to heat the cabin alone, waste heat recovery solutions on cars involves boosting the performance of a heat pump [3]. For mobile machines the waste heat conditions are different. From previous work [4] it was concluded that the available waste heat exceeds the cabin demand during high-utility excavating and transport cycles, shown in fig. 1. Based on those results, waste heat recovery from the working hydraulic system gained attention for further studies.

In contrast to other fluid circuits on a vehicle, the working hydraulic system in a mobile machine shows highly dynamic characteristics. This can cause challenges in the context of heat recovery. Especially when the recovered heat is targeted to a cabin, where arbitrary temperature fluctuations is not considered desirable. Battistel et.al. showed by experiments in climate chambers that a temperature deviation as small as 0.38°C in any direction from the base temperature 24°C was detected with 75% accuracy, and that +0.8°C and –0.91°C was detected with 100% accuracy [5].

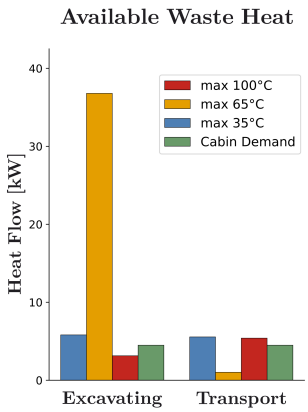


Figure 1: Thermal load or available waste heat during high-utility excavating or transportation in –15°C, divided into temperature categories—compared with cabin heating demand in green. Originates from previous work [4].

Heat recovery from working hydraulic systems has recently been investigated by [6], where the dynamic challenge was

specifically addressed. They used an Organic Rankine Cycle to convert the waste heat into electric power, and the results showed that the flow fluctuations did not have as high impact as the authors anticipated.

In this study, a concept to recover heat from the power losses generated in the working hydraulic system—in order to heat the cabin on an electric mobile machine—was evaluated through model-based simulation. The concept is based on the traditional principle of cabin heat management, where a glycol-water mix—here called transport medium—recovers heat from the ICE and possibly an auxiliary heater and delivers it to a heating core in the cabin. A corresponding version for an electric mobile machine involves multiple heat sources, and is illustrated in fig. 2. Here only one of them is considered, but on the other hand the heat source with the most uncertainties.

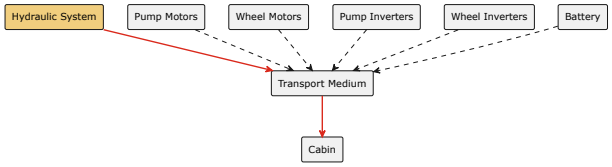


Figure 2: A diagram illustrating the focus area of the paper—the heat loss path from the hydraulic system to the cabin, highlighted with solid red arrows—in the context of an interconnected heat recovery system with several additional potential heat sources marked with dashed arrows.

1.1 Contributions

This paper focus on the working hydraulic system—and specifically the possibilities and challenges of recovering hydraulic heat losses to heat the cabin. A model of a hydraulic heat recovery system is presented, and the simulation results are analyzed and discussed. Both the dynamic behavior typically associated with hydraulic systems in combination with heat recovery, and the heating performance of the cabin are considered.

2 Modeling

The model was created with Modelica using standard components within the DLR Thermofluid Stream (TFS) library [7, 8], and the Modelica Standard Library. The TFS components used to model the fluid circuits are described in detail in [8]. Modelica Standard Library components were used for the control of the model. The selected parameters used within the model components and referred to in this chapter are listed in tab. 1. The simulations were performed using the open-source simulation tool OpenModelica. The main purpose of the model was to evaluate the impact of dynamic heat losses on the cabin air temperature—driven by oil temperature and mass flow—from the working hydraulic system of a mobile machine, when part of the heat flow is harvested and used to heat the cabin. The model shown in fig. 3 is the studied practical implementation of the heat loss path illustrated in fig. 2, and consists of the three main fluid circuits: hydraulic circuit, transport medium circuit, and cabin air supply circuit. A fourth circuit representing an ambient chiller, together with a hysteresis control logic is modeled to limit the hydraulic oil reservoir temperature to 60–70°C.

Table 1: Selected parameters used in the model components and simulations, some of which are subject to a sensitivity analysis, where the value listed in the table represents the selected base value.

Variable	Symbol	Value	Unit
Measured mean flow rate [4]	q_{mean}	2.54×10^{-3}	m^3/s
Measured mean heat flow [4]	$\dot{Q}_{mean}$	35.7×10^{3}	W
Intensity factor	X_I	1	–
Amplitude factor (heat)	$X_{A_{heat}}$	1	–
Amplitude factor (flow)	$X_{A_{flow}}$	1	–
Phase angle offset (heat)	φ_{heat}	0	rad
Heat loss frequency	f_h	0.1	Hz
Transport medium mass flow rate	$\dot{m}_{tm}$	0.186	kg/s
Cabin air mass flow rate	$\dot{m}_c$	0.11	kg/s
Overall Conductance (hydraulic)	UA_h	300	W/K
Overall Conductance (cabin)	UA_c	125	W/K
Reservoir volume (hydraulic)	V_h	0.34	m^3
Reservoir volume (transport)	V_{tm}	0.001	m^3
Cabin volume	V_c	1	m^3
Ambient temperature	T_{amb}	-15	°C
Simulation stop time	t_{stop}	3600	*s*
Simulation interval		0.1	*s*

The hydraulic circuit is modeled to represent the low-pressure return line of the working hydraulic system, transporting the heat losses to the ambient chiller and back to the hydraulic reservoir with the volume V_h. The heat losses from the working hydraulic system are modeled as a lumped model consisting of a pump and a conduction element. The pump is following a mass flow reference, and the conduction element produces heat according to a heat flow reference. The reference signals, y, in both cases are sine waves, here defined as eq. 1, with an offset value C_q and C_Q according to eq. 2 and 3. X_I is used as an intensity factor between 0–1 representing the degree of operating utility, where 1 is the highest utility representing 100%. The sine wave shape is chosen as model input because of its similarity with how multiple hydraulic cylinder stroke positions moves back and forth during an excavating cycle, creating continuous alternating flow rates and heat losses, while maintaining the simplicity of the model. The flow rate q_{mean} and heat flow $\dot{Q}_{mean}$ are based on measured average values during a high-utility excavating cycle from previous work [4]—here referred to as intensity factor $X_I = 1$—and can be found in tab. 1. The amplitude A of the sine waves are controlled by the use of amplitude factors $X_{A_{heat}}$ and $X_{A_{flow}}$, according to eq. 4. The frequency f_h is always the same for the two sine waves. The phase angle of the heat flow sine wave φ_{heat} can be adjusted in the positive direction relative to the flow sine wave, to model a transport delay of the oil temperature increase. The conduction element uses the sine wave heat flow reference directly as input, while the pump component creating the flow in the hydraulic circuit needs a P-controller from mass flow rate into an angular velocity control signal. Before reaching the ambient chiller, the hydraulic fluid is passing through a heat exchanger to harvest

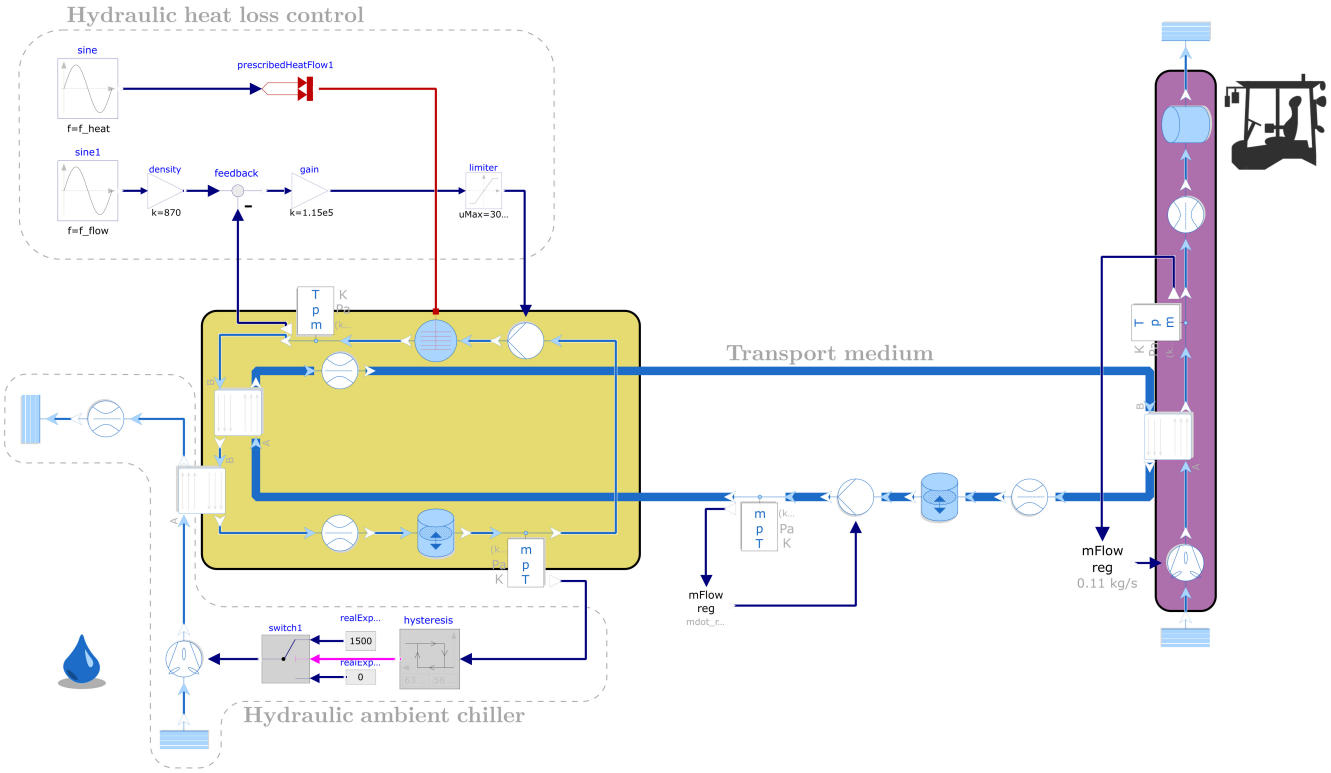


Figure 3: Modelica model of the studied hydraulic heat loss recovery system, composed of a hydraulic fluid circuit highlighted in the yellow zone, a transport medium circuit highlighted in thick blue lines, and a cabin air supply circuit highlighted in the purple zone. Additionally, a hydraulic ambient chiller circuit highlighted with dashed lines makes sure to maintain the oil reservoir temperature within 60–70°C. The fluid circuits are modeled with TFS standard components explained in [8]. The control of the TFS components are modeled with the Modelica Standard Library.

heat losses to the cabin via a transport medium.

$$y = A\,sin(2\pi\,f_h\,t + \varphi) + C \tag{1}$$

$$C_q = X_I\,q_{mean} \tag{2}$$

$$C_Q = X_I\,\dot{Q}_{mean} \tag{3}$$

$$A = X_A\,C \tag{4}$$

The transport medium circuit in this concept has the purpose of collecting heat losses from different subsystems of a vehicle and transporting the heat to the cabin, as shown in fig. 2. In this model only one subsystem, the hydraulic system, is modeled. Similar to the hydraulic circuit a flow is generated with a pump component, controlled to maintain a fixed mass flow rate of glycol-mixed water, $\dot{m}_{tm}$. The heat is collected with a heat exchanger component based on the εNTU-method, parametrized with overall heat transfer coefficient U and the conductive surface area. Here, these two heat exchanger properties are combined as the overall conductance UA and can be found in tab. 1. A second heat exchanger is transferring the heat from the transport medium to the cabin air. A reservoir component with volume V_{tm} is acting as flow source and sink.

The cabin air supply circuit is modeled with a fan maintaining a constant mass flow rate of air, $\dot{m}_c$, from a source with a preset temperature value of T_{amb}. A volume component of the size V_c is used to model the actual cabin. The outlet of the cabin volume is the point of where the temperature dynamics are evaluated in chapter 3, and is treated as the actual temperature in the cabin. The air media used is a dry air model.

The media models used in the hydraulic circuit and the transport medium circuit are both table based. The model for glycol-mixed water used in the transport medium circuit is the Glycol47 model available in the TFS library. The hydraulic oil model is the table of assumed values in tab. 2.

Table 2: Table-based media model of a hydraulic oil used in this model. The table structure is reused from the glycol-water model Glycol47 in the TFS library.

Temp °C	Density kg/m^3	Heat Capacity J/kgK	Conductivity W/mK	Viscosity Pa s	Vapor Pressure Pa
-40	896	1700	0.13	0.27	0.0001
-20	888	1800	0.13	0.18	0.0005
0	880	1900	0.13	0.11	0.001
20	872	2000	0.13	0.061	0.005
40	864	2100	0.13	0.04	0.01
60	856	2200	0.12	0.027	0.05
80	848	2300	0.12	0.021	0.1
100	840	2400	0.12	0.017	0.5
120	832	2500	0.11	0.014	1
140	824	2600	0.11	0.012	2
150	820	2650	0.10	0.011	3

3 Results

This section presents the results from the simulations made with the presented model. The results are divided into the two focus areas dynamic response and heating performance. Figures 4– 9 illustrate different sensitivity analyses of the cabin temperatures as oscillating peak-to-peak values versus model parameters. The cabin heating performance is shown in fig. 10 and 11.

3.1 Dynamic Response

Figure 4 shows three different sensitivity analyses. The green and orange lines show the resulting peak-to-peak temperatures when only one of the two sine wave amplitudes is varied while the other remains at zero. The blue line shows how the peak-to-peak temperature of the cabin air changes when both input amplitudes are changed simultaneously. It can be seen that as long as the amplitude factor X_A for both sine wave signals increases equally, the peak-to-peak temperature of the cabin air does not change significantly until the amplitude reaches its higher end point. The same point is also selected as base value for the other sensitivity analysis below.

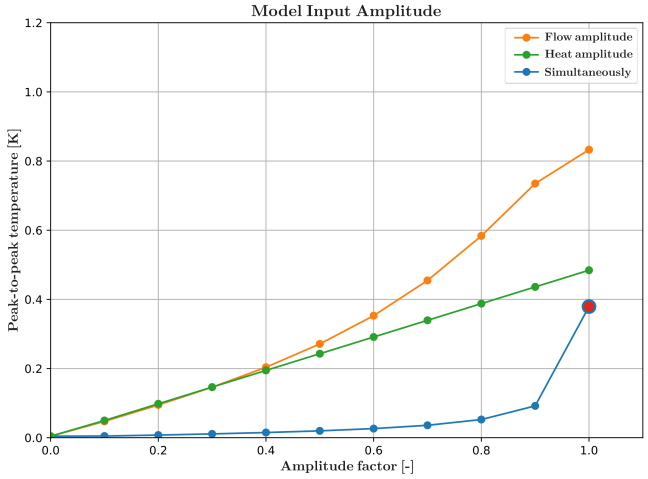


Figure 4: Sensitivity analysis of cabin air peak-to-peak temperature on the vertical axis, with amplitude factors $X_{A_{heat}}$ and $X_{A_{flow}}$ as input parameters on the horizontal axis. Green and orange lines show the cabin peak-to-peak temperature when the heat loss and flow rate amplitudes are changed separately, with the other amplitude factor set to zero. The blue line shows the peak-to-peak temperature when the two amplitude factors are changed simultaneously. The red dot marks the base value selected for the model.

Figure 5 shows how the sine wave frequency affects the peak-to-peak value of the cabin air temperature. It is observed that low frequencies have a significant impact on the peak-to-peak temperature.

Figure 6 shows how the phase angle offset between the flow sine wave and the heat loss sine wave affects the peak-to-peak temperature of the cabin air. It can be noted that the selected base value of the phase angle offset in the model is $\varphi_{heat} = 0$ rad.

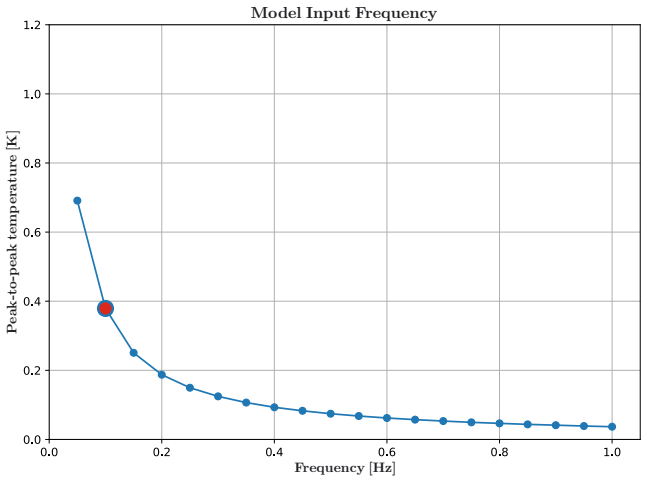


Figure 5: Sensitivity analysis of cabin air peak-to-peak temperature on the vertical axis, with heat loss frequency f_h as input parameter on the horizontal axis. The red dot marks the base value selected for the model.

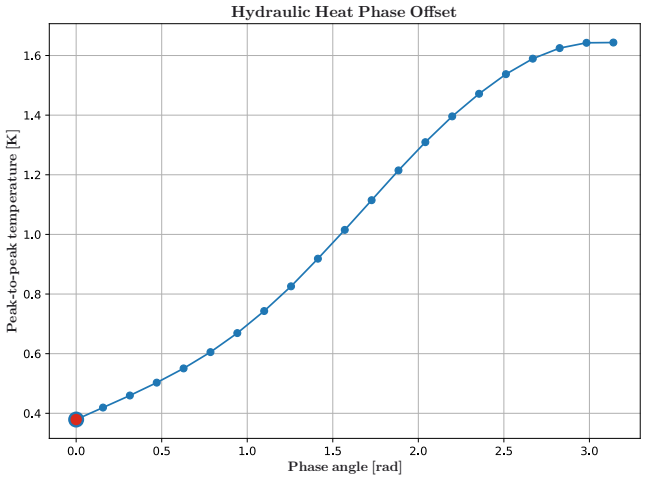


Figure 6: Sensitivity analysis of cabin air peak-to-peak temperature on the vertical axis, with heat flow phase angle offset φ_{heat} as input parameter on the horizontal axis. The red dot marks the base value selected for the model.

The transport medium mass flow rate sensitivity is shown in fig. 7, and the cabin air mass flow rate in fig. 8. It is observed that changing the mass flow rate has opposite effects on the peak-to-peak temperatures.

Finally, fig. 9 shows how the overall conductance *UA* affects the temperature peak-to-peak values for both the hydraulic and the cabin heat exchanger. It is noted that the cabin heat exchanger is more sensitive to changes in overall conductance than the hydraulic heat exchanger.

3.2 Heating Performance

Figure 10 shows the heat-up time of the cabin for different hydraulic tank volumes after a cold start at –15°C, with the other model parameters set to their base values. The cabin air

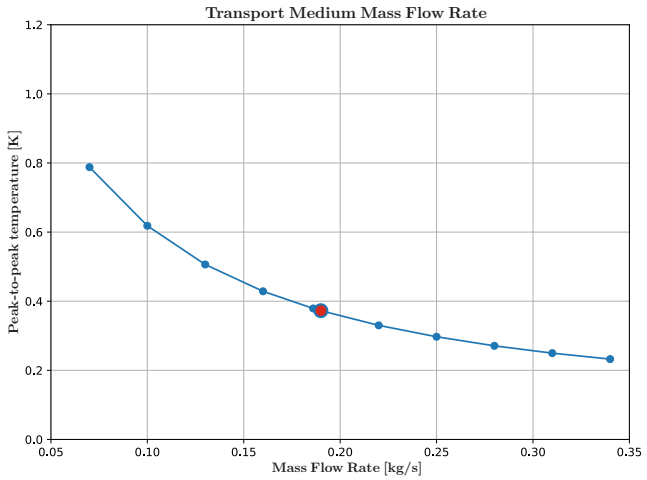


Figure 7: Sensitivity analysis of cabin air peak-to-peak temperature on the vertical axis, with varying transport medium mass flow rate $\dot{m}_{tm}$ as input parameter on the horizontal axis. The red dot marks the base value selected for the model.

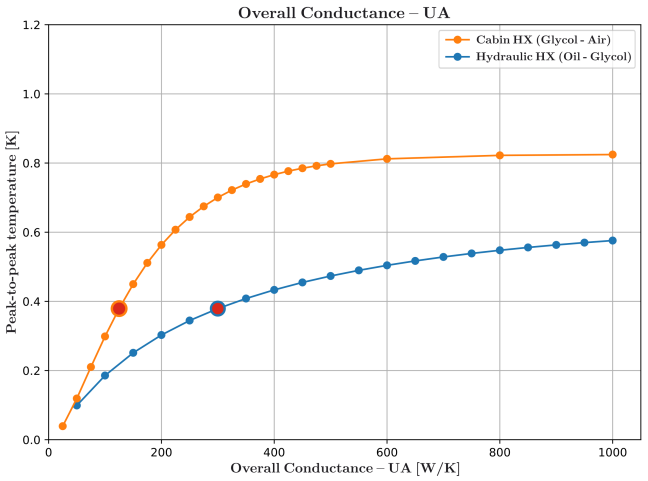


Figure 9: Sensitivity analysis of cabin air peak-to-peak temperature on the vertical axis, with varying overall conductance of the hydraulic and cabin heat exchanger UA_h and UA_c as input parameters on the horizontal axis—one heat exchanger at a time, with the other at base value. The red dots mark the base values selected for the model.

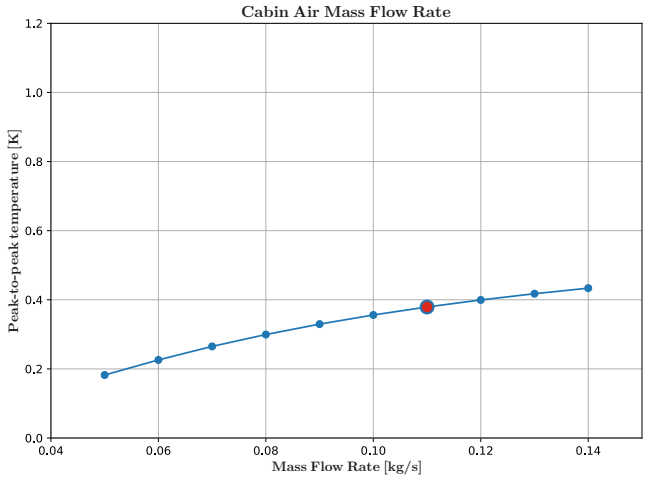


Figure 8: Sensitivity analysis of cabin air peak-to-peak temperature on the vertical axis, with varying cabin air mass flow rate $\dot{m}_c$ as input parameter on the horizontal axis. The red dot marks the base value selected for the model.

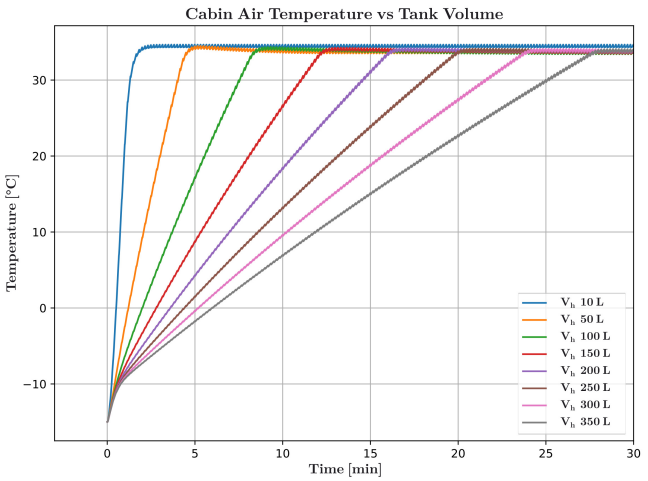


Figure 10: Cold start performance of the simulated cabin air temperature is shown on the vertical axis starting from –15°C, with simulation time on the horizontal axis. Each curve represents a different value of the hydraulic tank volume V_h.

temperature is limited by the maximum oil temperature, which in turn is limited by the ambient chiller. The tank volume of 10 liters gives a heat-up time for the cabin that is 12 times faster than for a tank volume with 350 liters. It can also be noted that the heat up time for the cabin from –15°C to 20°C with the 350 liter hydraulic tank is approximately 18 minutes.

Figure 11 shows the cabin air temperature during a cold start as a result of different values of the intensity factor X_I, with the other parameters set to their base values. Even though the graph ends after 60 minutes, the temperatures keep increasing up to 33°C for all intensity factors except for $X_I = 0.1$.

4 Discussion

The purpose of this paper was to study the possibilities and challenges of recovering waste heat from a working hydraulic system to warm the intake air supply of a cabin by utilizing a heat exchanger on the hydraulic low-pressure return line. Additionally, to investigate and to learn more about important parameters to consider if a such heat recovery concept would be implemented in a final product. Due to the dynamic nature of working hydraulic systems, and the challenge to practically measure these losses in real time, an attempt to create a model and simulate a heat recovery system was presented and studied.

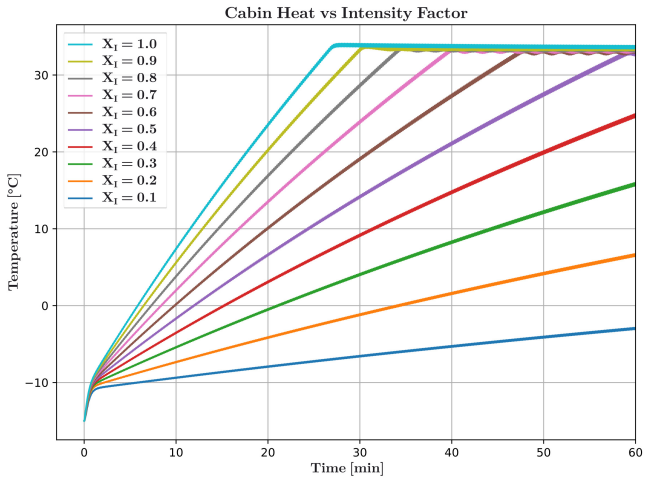


Figure 11: Cold start performance of the simulated cabin air temperature is shown on the vertical axis starting from –15°C, with simulation time on the horizontal axis. Each curve represents a different value of the intensity factor X_I, corresponding to excavating utility levels from 10–100%.

Overall, the findings suggest that harnessing heat from an operational hydraulic system to warm a cabin is a promising strategy. Even though the sensitivity analyses show us that unknown variables such as hydraulic amplitudes, frequency and phase offset affects the peak-to-peak temperature variation in the cabin significantly, there are ways to mitigate these variations such as active temperature control. Additionally, the model used in this paper is quite ideal, with the primary focus of catching the dynamic response. Masses and passive cooling entities exposed to ambient temperatures that occurs in a physical system will likely act as a filter on the dynamic heat flow. Looking back at fig. 2, the working hydraulic system is only one out of several other subsystems, meaning that a heat flow with dynamic behavior can possibly be mixed with more continuous heat flows before reaching the cabin. An active control of the cabin heating and ventilation system could possibly deal with fluctuations in the magnitude of what has been presented here. Finally, the peak-to-peak temperature with the current set of model parameters are almost equal to the human-detectable temperature deviations studied by [5], which further suggests that mitigating measures can be needed.

The heat recovery possibilities and cabin heating performance shown in fig. 10 are promising, however, additional parameter adjustment and validation is recommended to confirm their reliability. On the other hand, by making a simple comparison of the rate of temperature increase of the hydraulic oil tank between the simulation results from V_h = 350 L, and a previous measurement on a system with a similar tank volume, the simulated result is only 12% faster. The simulation result is based on the highest intensity factor—which was also the case in fig. 1—and is not the most common operating point. Assuming that intensity factor X_I = 0.5 is the most common operating point in average, the time it takes for the cabin to reach 20°C is almost 40 minutes, as can be seen in fig. 11. With that in mind, it can be motivated to look at downsizing the hydraulic tank volume to faster reach operation temperatures. For example, three cylinders required for the most basic digging motion requires a differential volume of roughly 35 liters, even though the hydraulic tank holds hundreds of liters. Due to the variety of hydraulic functions and operating scenarios of a machine, it might not be possible to downsize the tank. However, there could still be a reason to introduce a tank solution such that only a fraction of the total stored volume is utilized and heated until more is needed. Essential functionality such as air bleeding and oil filtration might however prevent or complicate such solutions to be implemented.

From the results it seems like the heat losses from the hydraulic system alone can sufficiently heat the cabin, which might be the case in some specific operating modes. There are, on the other hand, numerous situations where a mobile machine does not use the hydraulic system in the continuous manner required to produce the amount of heat presented in this paper—for example, transportation along a road from one work site to another, or idle-like tasks such as moving and holding heavy objects. This is one of the motivations behind an interconnected heat recovery system like the one illustrated in fig. 2. However, the working hydraulic system allows the direct heat recovery principle for electric mobile machines to cover the cabin heat demand during hydraulic operation, even under lower intensities. During low intensity transportation the available waste heat might not cover the cabin demand, which (with a bit of imagination) is indicated in fig. 1. In some of those cases—given that the machine was hydraulically operated before—the hydraulic tank could serve as a thermal storage device. In contrast to electric passenger cars, the amount of available waste heat is not always the issue for electric mobile machines—it is how to recover it.

5 Conclusions

Working hydraulic heat recovery may not always be exclusively sufficient to heat the cabin, but it offers a good heat source despite its dynamic characteristics. Technical solutions to allow the hydraulic oil heat-up time to be shortened can improve the heating performance in the cabin.

Acknowledgement

This research was funded by the Swedish Energy Agency (Energimyndigheten) with grant number P2023-00596.

References

[1] Zhen Tian, Bo Gu, Wenzhong Gao, and Yuan Zhang. Performance evaluation of an electric vehicle thermal management system with waste heat recovery. *Applied Thermal Engineering*, 169:114976, March 2020.

[2] Gabriele Leoncini, Ronan Mothier, Benoît Michel, and Marc Clausse. A review on challenges concerning thermal management system design for medium duty electric vehicles. *Applied Thermal Engineering*, 236:121464, January 2024.

[3] Sagar Singh, Mark Jennings, Sunil Katragadda, Judy Che, and Nenad Miljkovic. System design and analysis methods for optimal electric vehicle thermal management. *Applied Thermal Engineering*, 232:120990, September 2023.

[4] Fabian Lagerstedt, Samuel Kärnell, Marcus Rösth, and Liselott Ericson. Analysis of opportunities for integrated thermal management on battery powered mobile machines. In *Global Fluid Power Society PhD Symposium, GFPS 2024*, 2024. submitted.

[5] Laura Battistel, Andrea Vilardi, Massimiliano Zampini, and Riccardo Parin. An investigation on humans' sensitivity to environmental temperature. *Scientific Reports*, 13(1):21353, December 2023.

[6] Zhengzhao Gu, Kewen Feng, Lei Ge, and Long Quan. Dynamic modeling and optimization of organic Rankine cycle in the waste heat recovery of the hydraulic system. *Energy*, 263:125673, January 2023.

[7] Dirk Zimmer. Robust object-oriented formulation of directed thermofluid stream networks. *Mathematical and Computer Modelling of Dynamical Systems*, 26(3):204–233, 2020.

[8] Dirk Zimmer, Michael Meißner, and Niels Weber. The dlr thermofluid stream library. *Electronics*, 11(22), 2022.

An Elastohydrodynamic Lubrication (EHL) Model for Radial Piston Motors

Jinhwan Lee[1], Lizhi Shang[1], and Pierre BERNARD[2]

[1]Maha Fluid Power Research Center, Purdue University, Lafayette, IN, USA
E-mail: lee4775@purdue.edu, shangl@purdue.edu
[2] Poclain Hydraulics Industrie, Verberie, France
E-mail: pierre.bernard@poclain.com

Abstract

This study presents an Elastohydrodynamic Lubrication (EHL) simulation model developed for a multi-lobe radial piston motor and its experimental validation. Since those types of radial piston motors are frequently operated under high torque and low-speed conditions, the severe wear and excessive power loss challenge their lubricating interface design. However, understanding these lubrication interfaces has been limited by difficulties in experimental measurement under high pressure in small gaps and the complexity of the multi-physics involved. To address these challenges, the proposed model solves a density-based Reynolds equation to predict tribological behaviors of fluid films around a stepped piston with the consideration of multi-body dynamics, port and throttling loss, and elastic deformations. The simulation results indicate that the lubrication regimes and asperity contact pressure strongly rely on the chamber pressure and motor rotational speed. The roller – bushing interface undergoes mixed lubrication regimes, and severe asperity contact occurs in piston – cylinder interfaces, indicating boundary lubrication during high load conditions. In addition, the interaction between the fluid films demonstrated an asymmetric deformation of the piston, significantly influencing the hydrodynamic pressure and film thickness distributions in the roller – bushing interface. Also, the consideration of tilting of a piston was essential to predict hydrodynamic and asperity contact pressures for piston – cylinder interfaces. As a result, the power loss analysis identified the main power loss contributors; the upper piston – cylinder interface, and throttling loss. This paper presents the simulation, the governing equations, the numerical solver, the force balance calculation, and the fluid and structure interactive scheme. The paper also compares the simulated friction torque and contact area against the experimental measurement. The developed model can provide insights into fluid film behaviors and assess the effect of design parameters such as clearance, dimensions, and material properties on tribological characteristics of radial piston motors.

Keywords: Radial piston motors, elastohydrodynamic lubrication, deformation, power loss.

1 Introduction

A multi-lobe radial piston motor is essential in various industrial applications that require high-torque actuation, especially a hydraulic transmission in mobile hydraulics for construction and agricultural machinery. Radial piston motors convert the flow energy from pressurized inlet fluid into mechanical work through the movement of pistons. In this process, forces are transferred through a series of components—cam, roller, bushing, piston, and cylinder bore wall—via lubricating interfaces. The hydrostatic pressure applied to the piston is balanced by the normal force from the cam, causing the piston to exert force on the cylinder wall and generate torque. More details on the working principle of radial piston motors can be found in [1,2].

Tribological behaviors of the lubricating interfaces can substantially affect the overall efficiency and reliability of radial piston motors. There are multiple lubricating interfaces for a typical stepped piston assembly as described in fig. 1. The cam – roller interface is a nonconformal rolling contact, and the roller – bushing interface is a conformal contact. Also, the connected upper and separate lower bore interfaces are piston – cylinder interfaces that do not have reliable pressure build-up mechanisms. Radial piston motors typically operate at low speeds and high loads, and these operating conditions have the potential to result in excessive power loss due to increased solid friction from asperity contact and mechanical failure from surface damage.

Many experimental research works focused on improving the overall performance of the lubricating interfaces by reducing frictional coefficient and surface damage. Olsson et al [1] used a simulated test bench to investigate the effect of lubricant properties on tribological behaviors of a piston – cylinder interface, assuming a rectangular contact area. Pettersson et al [3] showed that the textured surfaces of a roller – piston interface could improve friction coefficient variation. Lewis [4] studied the effect of surface layer material on

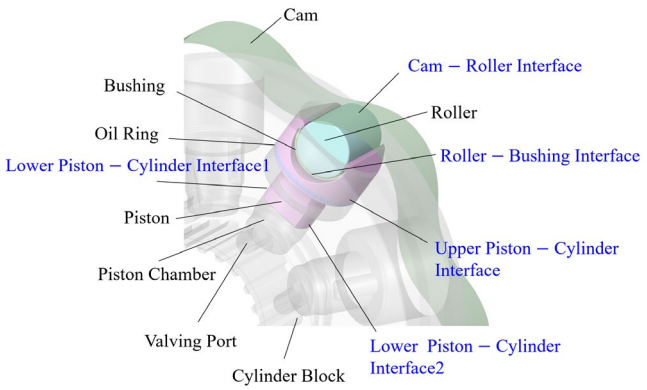


Figure 1: Lubricating Interfaces.

friction coefficient and scuffing damage with different lubricant supply conditions and demonstrated lower friction coefficient with cloth material than sintered bronze. Nilsson et al [5] studied seizure mechanisms of a radial piston motor under actual motor operation conditions using a test configuration that can measure hydromechanical losses. They demonstrated that seizure could initiate from a roller – piston interface and piston – cylinder interface with increased hydromechanical loss, severe wear, and scuffing. A drawback with the above-mentioned works is that their results may not exactly represent the motor operation condition because the lubricating interfaces were investigated in isolation with simplified test setups in which lubricating conditions might be different [1,3,4]. Also, experimental studies may need bulky and high-cost test rigs that cannot be easily configurable for different machine designs. Above all, it is difficult to directly measure important physical quantities such as hydrodynamic pressure, asperity contact pressure, and film thickness. To address those difficulties, many numerical models were proposed to understand the behaviors of lubricating interfaces in hydraulic machines. Dahlén et al [6] proposed a full film lubrication model for the roller – bushing interface and valving interface in a radial piston motor by solving a Reynolds equation. Isaksson et al [7, 8] proposed a full film and mixed lubrication model for the roller – piston interface with a hydrostatic groove, considering the piston deformation and surface roughness. Zhang et al [9] proposed the unsteady EHL model for a roller – piston interface and showed a slotted hole could largely increase the film thickness. Zhang et al [10] proposed a modified lubrication model for a roller – piston interface with a focus on thermal analysis, demonstrating a significant temperature rise from viscous and asperity friction. Li et al [11] proposed a CFD simulation for a valving interface with leakages and volumetric efficiency analysis. Besides simulation models for radial piston motors, more advanced coupled simulation models have been developed for other types of hydraulic machines. Ransegnola et al [12] proposed a simulation model that considered multiple lubricating interfaces in a swashplate-type axial piston machine to investigate piston and slipper spin. Mukherjee et al [13] developed a coupled thermal simulation model to study the effect of thermal behaviors on the efficiency and film thickness in a swashplate-type axial piston machine. However, a coupled EHL simulation model for a radial piston motor that considers various multiphysics such as transient chamber pressure, lubricating interfaces, asperity contact, rigid body dynamics, and deformation has never been developed. Also, previous numerical studies for a radial piston motor neglected the piston dynamics in the actual operations such as tilting which greatly affected the asperity contact pressure and area in this work. Moreover, the interaction between the roller – bushing interface and piston – cylinder interface has never been accounted for due to complexity despite their combined significant effect on the deformation and dynamics of the piston. This explains why although a considerable amount of experimental work and simulation models have been made for the lubricating interfaces in radial piston motors, designing those lubricating interfaces remains a challenge for engineers with a limited understanding of their behaviors.

The purpose of this study is to develop and validate a coupled EHL simulation model that can assess various design

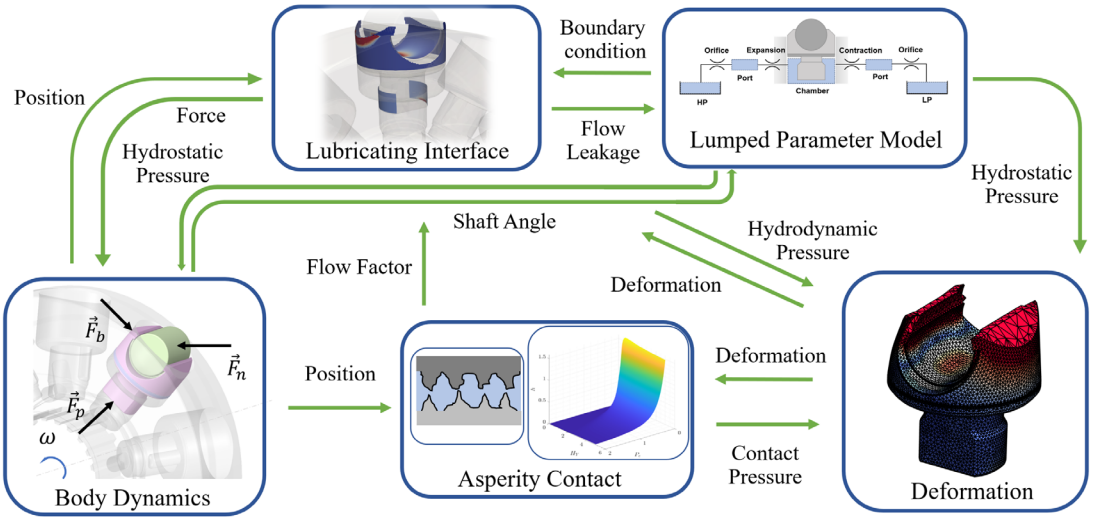


Figure 2: Simulation domains.

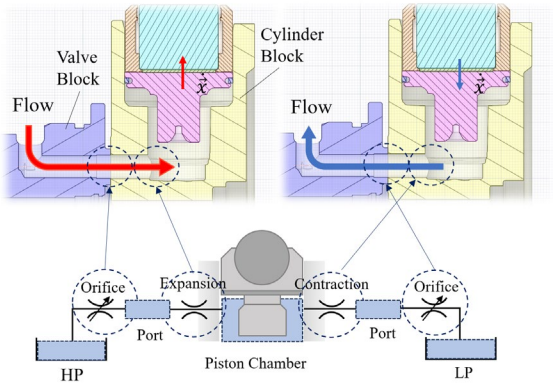


Figure 3: Hydraulic circuit for the LP model.

parameters and provide physical insights into tribological behavior considering actual motor operation. Universal Reynolds equations for the roller – piston interface and piston – cylinder interfaces are fully coupled with the lumped parameter model, rigid body dynamics, and deformation caused by hydrodynamic and asperity contact pressures. The physical quantities exchanged between different physical domains are illustrated in fig. 2. In an effort to validate the developed simulation model, experimental measurement was performed for the roller – bushing interface, and all lubricating interfaces including piston – cylinder interface. A commercially available MSE08 from Poclain Hydraulics Industrie with a maximum displacement of 1248 cc and power of 41 kW was chosen as a reference machine. The working fluid is ISO VG46 and an isothermal condition of 50 °C is considered for typical operating conditions of the machine. The simulation was run for the shaft speed ranging from 30 RPM to 80 RPM, inlet pressure ranging from 50 bar to 350 bar, and outlet pressure set to 20 bar.

2 Numerical models

2.1 Lumped parameter (LP) model

The LP model is crucial to the entire simulation because it calculates the piston chamber pressure, which other domain solvers rely on, and the throttling loss, another major source of the power losses. The cross-section view of a piston with the valving interface and its equivalent hydraulic circuit are presented in fig. 3. As the cylinder block rotates, the port is alternatively connected to the high-pressure inlet and low-pressure outlet. When the port connects to the high-pressure inlet, the piston moves upward, increasing the piston chamber volume; when it connects to the low-pressure outlet, the piston moves downward. The port connection is where the throttling loss occurs, and any flow leakage or friction loss through the valving interface was not considered in the simulation.

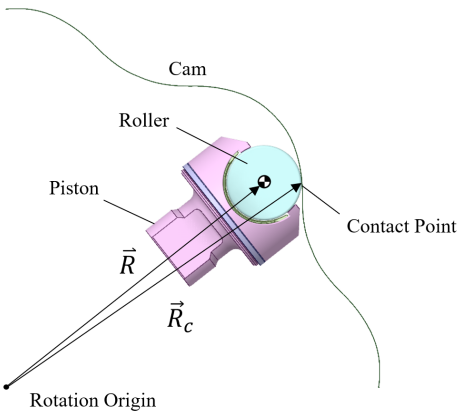


Figure 4: Piston and roller kinematics.

While computationally cost effective and reasonably accurate model to calculate pressure for a control volume such as a piston chamber and port is pressure build-up equation (eq. (1)) where p, K and V represent volume averaged pressure, isothermal bulk modulus, and chamber volume.

$$\frac{dp}{dt} = \frac{\mathrm{K}}{V}\left(\Sigma Q - \frac{dV}{dt}\right) \tag{1}$$

Also, the volume flow Q in eq. (1) depends on opening area A of the valving port and the pressure difference Δp, and this can be modeled as flow with a restrictive orifice (eq. (2)). C_f and ρ represent orifice discharge constant and averaged density of the fluid.

$$Q = C_f A \sqrt{\frac{2\Delta p}{\rho}} \tag{2}$$

In addition to the throttling loss, expansion and contraction of the flow path may provide additional pressure drop, Δp as modeled in eq. (3). κ and f, L, D, and v represent minor loss coefficient, Darcy friction coefficient, port length, port diameter, and averaged flow velocity respectively.

$$\Delta p = \left(\kappa + f\frac{L}{D}\right)\frac{\rho v^2}{2} \tag{3}$$

Since the orifice opening area and volume change are function of a motor shaft angle, eq (1), (2) and (3) should be solved simultaneously with body dynamics.

2.2 Lubricating interface

The hydrodynamic pressure of the lubricating interfaces can be calculated using the density-based universal Reynolds equation for a mixed lubrication suggested by Ransegnola [14] where ϕ_P, ϕ_S, ϕ_C and ϕ_R represents pressure flow factor, shear flow factor, contact flow factor, and roughness flow factor in eq (4). The equation considers the effect of asperity contact on hydrodynamic pressure by introducing the concept of averaged flow using flow factors. The pressure flow factor ϕ_P represents increased resistance from asperity for pressure-driven Poiseuille flow and shear flow factor ϕ_S represents added shear flow caused by asperity [15,16]. Also, contact flow factor ϕ_C and roughness flow factor ϕ_R are used to relate averaged film thickness with local film thickness for compressible flow [14,17]. Those flow factors can be pre-calculated for nondimensional film thicknesses assuming Gaussian and isotropic surface roughness, and the results can be found in [14].

$$\nabla\cdot\left(\phi_P\left(\frac{Kh^3}{12\mu}\nabla\rho\right)\right)=\nabla\cdot\left(\rho v_m\left(\phi_R R_q+\phi_C h\right)\right)+$$
$$\nabla\cdot\left(\rho\phi_S R_q\left(\frac{\vec{v}_t-\vec{v}_b}{2}\right)\right)+\frac{\partial\rho\left(\phi_R R_q+\phi_C h\right)}{\partial t} \quad (4)$$

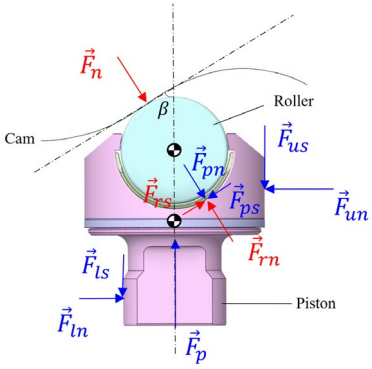


Figure 5: Free body diagram for roller and piston.

2.3 Asperity contact

In mixed or boundary lubrication regimes, the asperity contact pressure provides additional load support for a lubricating interface. Lee et al [18] provided the curve-fitted analytic model that can predict averaged asperity contact pressure for given material properties and fluid film thickness, considering the elastic and plastic behavior of surfaces (eq. (5)). The surface roughness was measured for all lubricating surfaces, and the equivalent surface roughness were used for eq. (5) and flow factor calculations for each lubricating interface. In addition, the measured surface profile of the roller from partial crowning which is known to prevent high stress at the edges was implemented to provide an additional gap height to the edges.

$$\ln\left(\frac{\bar{h}}{R_q}\right)=\sum_i\left(\vec{r}_G^{\,T}G_i\vec{H}\right)\bar{P}^i \quad (5)$$

2.4 Kinematics and rigid body dynamics

As shown in fig. 4, the rotational velocity of the roller can be obtained by assuming no slip at the contact point, which means the velocity at the contact point should be zero (eq. (6)). The calculated rotational velocity was imported as a look-up table which is a function of a cylinder block rotation angle.

$$\vec{V}_c=\vec{V}_r+\vec{\omega}_r\times\left(\vec{R}_c-\vec{R}\right)=\mathbf{0} \quad (6)$$

The free-body diagram for the piston and roller is shown in fig. 5. The forces applied to the roller are represented in red and the forces applied to the piston are represented in blue. The pressure force $\vec{F}_p$ is calculated from the piston chamber pressure in the LP model. The normal force $\vec{F}_n$ represents the reaction force from the contact point and can be approximated from force balance using contact angle β without considering friction from the piston-cylinder interfaces. $\vec{F}_{rs}$ and $\vec{F}_{rn}$, are shear and normal force applied to the roller by the roller – bushing interface, and $\vec{F}_{ps}$ and $\vec{F}_{pn}$ are shear and normal force applied to the piston by the same lubricating interface. Similarly, $\vec{F}_{us}$ and $\vec{F}_{un}$ are shear and normal force applied to the piston by upper piston – cylinder lubricating interface, and $\vec{F}_{ls}$ and $\vec{F}_{ln}$ are shear and normal force applied to the piston by the lower piston – cylinder lubricating interfaces. Also, the cylinder block was set to rotate at a constant shaft speed. Since the motion in axial direction is constrained by a thrust bearing and the valving interface, only planar motions were considered for the roller and piston with Newton's second law (eq. (7,8)).

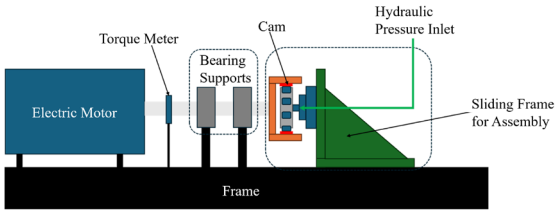


Figure 6: Schematic diagram for experimental setup.

$$\sum\vec{F}=\frac{d}{dt}\vec{P} \quad (7)$$

$$\sum\vec{M}=\frac{d}{dt}\vec{L} \quad (8)$$

2.5 Solid deformation

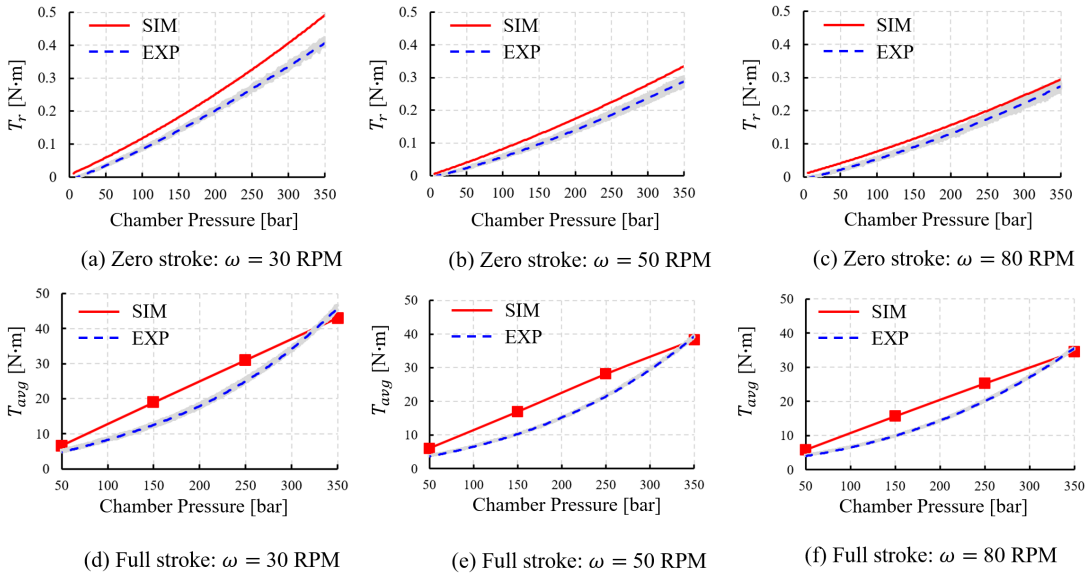
In general, deformation considerably affects hydrodynamic pressure and film thickness distribution in an elastohydrodynamic lubrication. Moreover, the roller – bushing interface and the piston – cylinder interface have a mutual effect on the overall deformation of the piston. One of the most successful methods to consider deformation for a Reynolds equation is the influence matrix method [19]. By solving a linear static elastic equation offline using a finite element method, the relationship between the deformation gap $h_{def,i}$ for i-node and reference pressure p_{ref} on j-face in fluid film domains is found in matrix $[IM]_{i,j}$ as eq. (9). Then, the matrix operation is performed in the main coupled simulation for combined pressure p_j which includes hydrostatic, hydrodynamic, and contact pressure for j-face in fluid films. To generate the influence matrix for each solid body, the mechanical properties of standard alloy steel were used for the roller, piston, and cylinder block, and the equivalent mechanical property was used for PEEK-based bushing.

$$h_{def,i}=[IM]_{i,j}\left(p_j/p_{ref}\right) \quad (9)$$

3 Experimental Validation

3.1 Experimental setup

The friction torque from the simulated lubricating interfaces was compared to measurement using an experimental setup with a torque meter (fig. 6). To isolate frictions from other sources of power loss, the electric motor



(a) Zero stroke: $\omega = 30$ RPM (b) Zero stroke: $\omega = 50$ RPM (c) Zero stroke: $\omega = 80$ RPM

(d) Full stroke: $\omega = 30$ RPM (e) Full stroke: $\omega = 50$ RPM (f) Full stroke: $\omega = 80$ RPM

Figure 7: Friction torque comparisons.

rotates the cam through a shaft supported by two bearing supports while pressurized fluid is supplied to the inlet of the fixed cylinder block. During the measurement, the chamber pressure was controlled to increase slowly. This experimental setup is different from actual motor operating conditions where cylinder block ports are alternatively connected to valve block ports and pressurization and depressurization of chamber pressure occur. Additionally, ISO VG46 was used as the working fluid, and the inlet temperature of the cylinder block was controlled to 50 °C by an external cooling system. For all measurements, the torque meter was calibrated to eliminate the friction torque from the bearing supports, and the chamber pressure was monitored.

3.2 Friction torque comparison

First, a zero-stroke cam without a wavy profile was used to prevent reciprocating motion of the piston where only friction from the roller – bushing interface was measured. The measurement inevitably includes the friction from roller – cam interface, but the amount may be negligible when assuming pure rolling. The measured friction torque T_s from the shaft can be related to simulated roller friction torque T_r using eq. (10, 11) where P, N, ω_s, and ω_r represent the power loss, number of rollers, shaft speed, roller speed, and D_c and D_r represent the diameter of the cam and roller. Similar to the experiment, the inlet pressure and resultant force applied were increased gradually for different shaft speeds during the simulation. The measurement and simulation results showed a relatively good match in trends (fig. 7 (a, b, c)) although the growing deviation can be observed for high chamber pressure at 30 RPM, which can be attributed to nonlinear asperity contact behavior at this condition.

$$P = T_s\omega_s = NT_r\omega_r \tag{10}$$

$$\omega_r = (D_c/D_r)\omega_s \tag{11}$$

Second, the full-stroke cam with a wavy profile was used to measure the friction from all lubricating interfaces including the roller – bushing interface and piston – cylinder interfaces. Unlike the case for the zero-stroke cam with the gradual pressure increase, the simulation was run for 12 operating conditions due to high computational cost. Averaged friction torque over a period of the reciprocating motion of the piston is shown in fig. 7 (d, e, f). The power loss from the piston – cylinder interfaces was much greater than that from the roller – bushing interface and mostly due to asperity contact friction. The simulation results only matched well with the measurement for specific conditions, indicating different trends in the friction torque with increasing chamber pressure. Experimental results suggested apparently quadratically increasing frictional torque while the friction torque from the simulation linearly increased.

4 Results and discussion

Unlike the experimental set up where the chamber pressure is slowly increased or maintained, the simulation results in this section considers transient chamber pressure and valving port timing by solving the lumped parameter model.

4.1 Roller – bushing interface

The fluid films are strongly coupled through dynamics and deformation and this coupling significantly affects

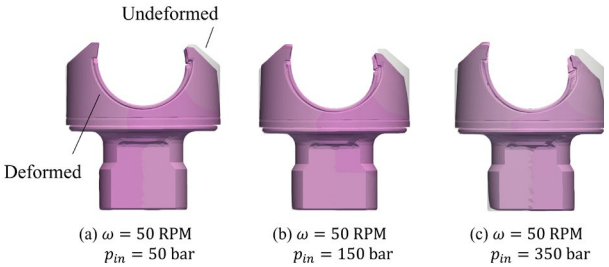


Figure 8: Deformation of piston (exaggerated).

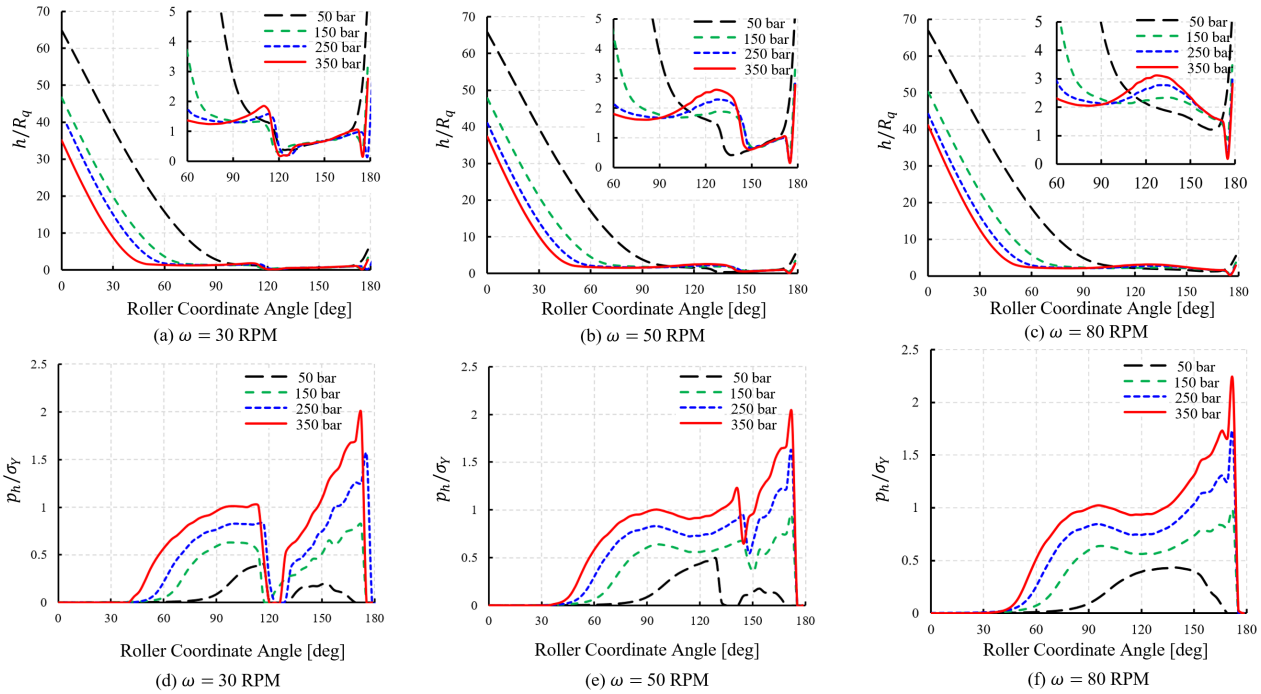


Figure 9: Roller – bushing interface results at the midplane.

hydrodynamic and asperity contact pressures and fluid film thickness distribution. This is especially important when the contact angle β is at the maximum and the piston exerts a maximum force on the cylinder wall, and as a result, the piston deforms asymmetrically (fig. 8). The hydrodynamic pressure and film thickness distribution of the roller – bushing interface at this moment are shown in fig. (9, 10). In fig. (9), the x-axis represents the roller coordinate angle at the mid-plane of the interface, and the y-axis represents the corresponding dimensionless hydrodynamic pressure and film thickness. When the radial piston motor is under low-duty operating conditions such as inlet pressure of 50 bar and shaft speed of 80 RPM, the minimum dimensionless film thickness is greater than 1, indicating a partial lubrication regime [20]. However, for all shaft speeds with the inlet pressure of 150, 250, and 350 bar, the asymmetric deformation of the piston bore reduced the film thickness significantly near the edge (180°), involving the hydrodynamic pressure peak and asperity contact. This behavior matched well with the experimental observation made in the previous study by Nilsson et al [5] which demonstrated severe wear near the edge. Moreover, when the shaft speed decreased from 80 RPM to 50 RPM, the loss of hydrodynamic pressure induced more asperity contact before the roller coordinate angle of 150° (fig. 9 (e) and fig. 10 (b, e)). This behavior is even more pronounced when the shaft speed is 30 RPM, showing a complete loss of hydrodynamic pressure around 120° for all inlet pressure conditions (fig. 9 (d) and fig. 10 (a,d)). The film thickness remained relatively constant due to the severe solid contact, and this eliminated the primary hydrodynamic pressure build-up mechanism, physical wedge from the geometrical nominal clearance. These results demonstrated the tribological behavior of the roller – bushing interface is highly sensitive to the operating conditions, and high loading and low shaft speed can pose wear or seizure as observed in [5].

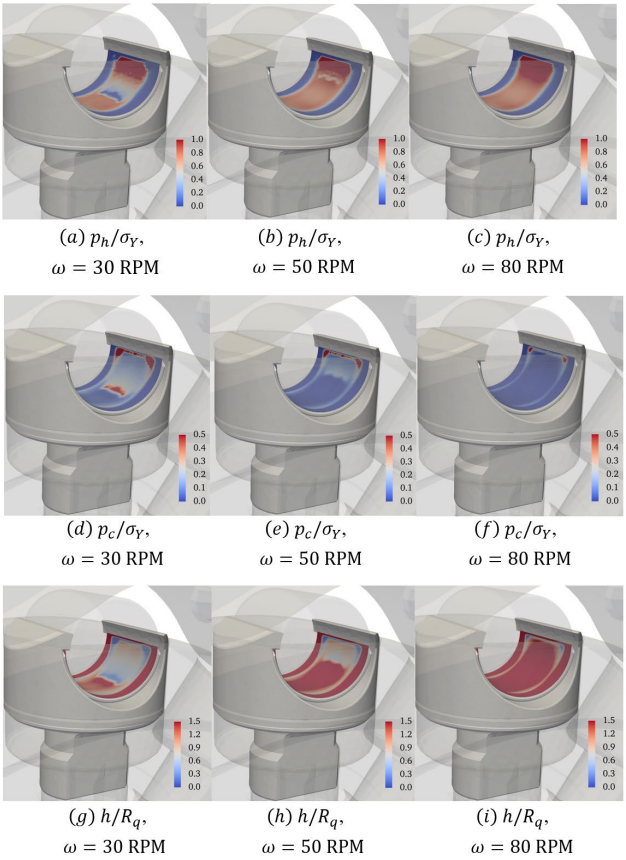
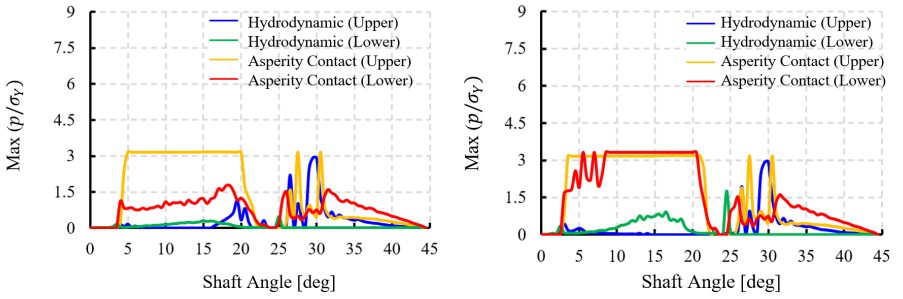
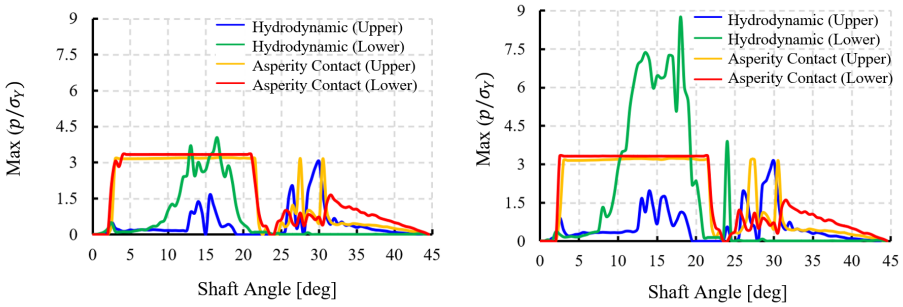


Figure 10: Roller – bushing interface for $p_{in} = 250\ bar$.

4.2 Piston – cylinder interface



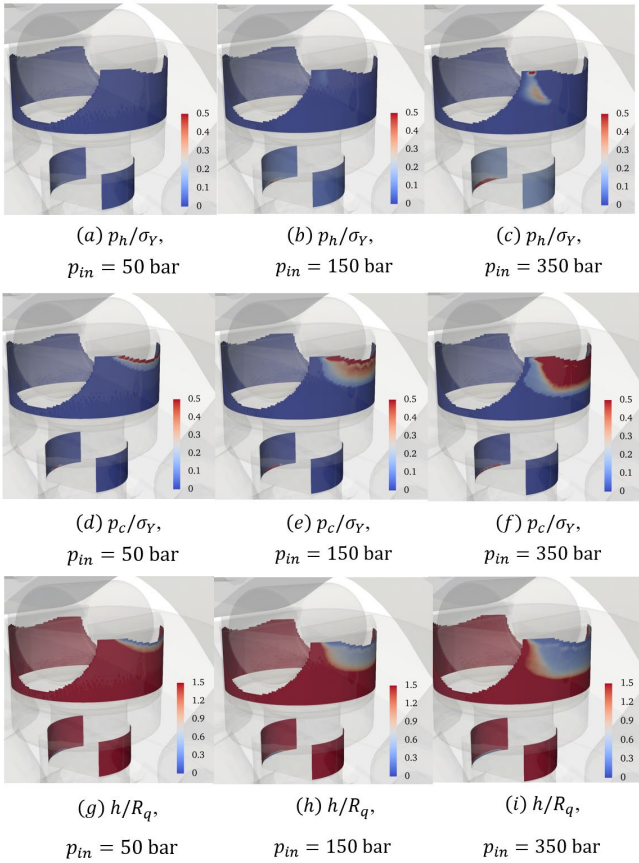
(a) $\omega = 50$ RPM, $p_{in} = 50\ bar$ (b) $\omega = 50$ RPM, $p_{in} = 150\ bar$



(c) $\omega = 50$ RPM, $p_{in} = 250\ bar$ (d) $\omega = 50$ RPM, $p_{in} = 350\ bar$

Figure 11: Maximum dimensionless hydrodynamic and contact pressure.

When the contact angle is not zero, the geometrical nominal clearance between a piston and cylinder bore can induce tilting of the piston which can largely influence the hydrodynamic and contact pressure distribution on the piston – cylinder interfaces. To investigate those behaviors, the maximum value of hydrodynamic and asperity contact pressures over each interface domain for a period of reciprocating motion are shown for the shaft speed of 50 RPM with different inlet pressures (fig. 11). From 0° to 22.5° shaft angle, the piston chamber is connected to a high-pressure inlet, and the piston starts to move upward, tilting in the clockwise direction. This tilting of the piston offers a physical wedge only for the lower lubricating films, and the upper fluid film should rely on squeeze motion to build up hydrodynamic pressure. Therefore, the asperity contact pressure of the upper fluid film rapidly increased to hardness value and supported most of the loading for all inlet pressure conditions (fig. 12 (d, e, f)). It should be noted that the hydrodynamic pressure only temporarily existed up to the shaft angle of 13° due to increased squeeze motion, and the pressure spikes from the shaft angle of 13° to 22.5° originated from a few isolated cells which cannot contribute much to the load support (fig. 12 (c)). On the other hand, the lower fluid films can build up hydrodynamic pressure from the physical wedge, showing a significant increase in the high loading (fig. 11 (d) and fig. 12 (c)). However, the asperity contacts also occurred for lower films because the hydrodynamic alone cannot support the loading, and this asperity contact is highly dependent on the loading (fig. 12 (d, e, f)). When the piston chamber is connected to the low-pressure outlet of 20 bar which corresponds to 22.5° to 45° shaft angle, the piston moves downward. The counterclockwise tilting of the piston provides the physical wedge only to the upper fluid film. Therefore, the maximum



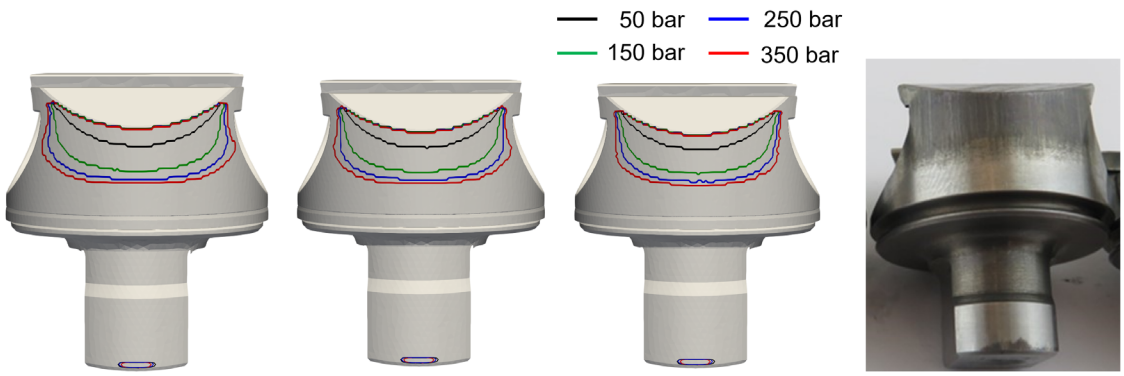
(a) p_h/σ_Y, $p_{in} = 50$ bar (b) p_h/σ_Y, $p_{in} = 150$ bar (c) p_h/σ_Y, $p_{in} = 350$ bar

(d) p_c/σ_Y, $p_{in} = 50$ bar (e) p_c/σ_Y, $p_{in} = 150$ bar (f) p_c/σ_Y, $p_{in} = 350$ bar

(g) h/R_q, $p_{in} = 50$ bar (h) h/R_q, $p_{in} = 150$ bar (i) h/R_q, $p_{in} = 350$ bar

Figure 12: Piston – cylinder interface for $\omega = 50$ RPM

hydrodynamic pressure of the upper fluid film was higher when the piston moved downward than when the piston moved upward despite the much smaller loading. On the other hand, the lower fluid films cannot build up hydrodynamics pressure,



(a) $\omega = 30$ RPM (b) $\omega = 50$ RPM (c) $\omega = 80$ RPM (d) Piston Wear

Figure 13: Asperity contact area.

and most of the loading is supported by asperity contact pressure. The region of high asperity contact pressure (>100 bar) during the reciprocating motion is illustrated in fig. 13 (a, b, c) for various shaft speed and inlet pressure conditions. The predicted area was compared to the wear area of a piston under real motor operating conditions and showed a good match in the location.

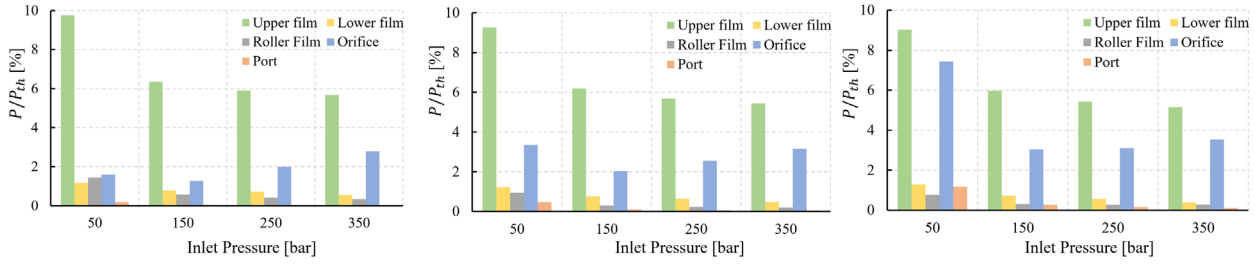
In terms of the power loss calculation for the piston – cylinder interfaces, the viscous friction was almost negligible compared to asperity contact friction. Also, it should be noted that the difference in trends between the simulation and measurement as shown in fig. 7 is based on the constant solid friction coefficient used in the simulation model. However, the experimental results demonstrated an increasing friction coefficient, observed in a typical Stribeck curve where the friction coefficient increases when the lubrication regime changes from mixed lubrication to boundary lubrication.

4.3 Power loss analysis

The ratio of power loss to theoretical power P/P_{th} (eq. 12) was calculated to identify power loss contributions as shown in fig. 14.

$$P/P_{th} = P/(\Delta p \cdot V_{disp} \cdot \omega_s) \qquad (12)$$

The power loss from the upper fluid film accounts for more than 45 % to 70 % of the total power loss depending on operating conditions. This result is expected from the above discussion that the upper fluid film is subject to severe asperity contact. When the inlet pressure increased from 50 bar to 150 bar, the increased hydrodynamic pressure from squeeze motion decreased the power loss ratio. Also, the lower fluid films demonstrated negligible power loss ratio in most cases because they can build up hydrodynamic pressure during the upstroke and are subject to a much lower load than the upper fluid film. The power loss ratio from the roller — bushing interfaces is relatively small because it can always build up hydrodynamic pressure to a certain degree, and benefit from the low solid friction coefficient of the PEEK-based bushing to steel interface. Besides the lubricating fluid film, the developed model can predict the power loss from the orifice and port from the LP model. Those additional losses represent the pressure drop of the flow caused by viscous dissipation, occurring at connections in the cylinder block. The port loss ratio was almost negligible for most of the cases considered while it showed a maximum value of 1 percent in low inlet pressure of 50 bar and high shaft speed of 80 RPM condition. The throttling loss represented as an orifice, however, strongly depended on the inlet pressure and shaft speed. When the inlet pressure changed from 50 bar to 150 bar, the throttling power loss ratio decreased for all shaft speeds. This may be because the small pressure difference between the inlet and piston chamber might lead to the incomplete filling of the piston chamber, increasing the duration of the pressure drop. However, as the inlet pressure increased from 150 bar to 250 bar and 350 bar, the throttling loss ratio increased. If the pressure inlet is high enough to avoid incomplete filing, the additional pressure increase can contribute to higher flow velocity across the orifice area which induces higher viscous dissipation.



(a) $\omega = 30$ RPM (b) $\omega = 50$ RPM (c) $\omega = 80$ RPM

Figure 14: Power loss analysis.

5 Conclusion and outlook

A novel simulation method was developed to investigate the piston lubricating interfaces in a radial piston motor. The simulation model can account for the lumped parameter model, body dynamics, fluid films, asperity contact, and deformation in a fully coupled way. The consideration of those multiphysics was found to be essential in determining tribological behaviors of the lubricating interfaces. The power loss analysis was conducted for all lubricating interfaces and hydraulic components in the lumped parameter model. The upper piston — cylinder interface was a major contributor, followed by the throttling (orifice) loss for all operating conditions considered. In summary, the developed model can be used to access the design parameters including the nominal clearance, piston geometry, cam profile, and valving port area and timing. Also, the model can provide insights into the lubricating behavior of the fluid films around a piston in a radial piston motor.

As a next step, a more realistic model for the solid friction coefficient needs to be developed because the current model overpredicts the friction torque from piston — cylinder interface for the medium inlet pressure range. Even though the new model may allow more accurate friction torque and power loss calculations, the effect of solid friction coefficient on other physical quantities, such as hydrodynamic and asperity contact pressures, and film thickness distribution in the lubricating interfaces may be minimal. Since the friction forces are relatively small compared to the hydrostatic force on the piston or the normal force on the roller, a similar body dynamic will be applied, and the current results for tribological behaviors would still be valid.

Nomenclature

Designation	Denotation	Unit
ρ	Density of a working fluid	kg/m^3
μ	Viscosity of a working fluid	$Pa \cdot s$
κ	Minor loss coefficient	$-$
v	Flow velocity magnitude	m/s
t	Time	s
p	Pressure	Pa
f	Darcy friction factor	$-$
V	Piston chamber volume	m^3
Q	Volume flow rate	m^3/s
P	Power loss	W
N	Number of rollers	$-$
L	Port length	m
D	Port diameter	m
A	Orifice Area	m^2
K	Bulk viscosity	Pa
h	Local fluid film thickness	m
ϕ_S	Shear flow factor	$-$
ϕ_R	Roughness flow factor	$-$
ϕ_P	Pressure flow factor	$-$
ϕ_C	Contact flow factor	$-$
ω_s	Shaft speed	rad/s
ω_r	Roller angular speed	rad/s
σ_Y	Minimum yield strength of two surfaces	Pa
v_m	Mean velocity of two surfaces	m/s
p_{ref}	Reference pressure	Pa
p_j	Pressure on jth face	Pa
p_{in}	Inlet pressure	Pa
p_h	Hydrodynamic pressure	Pa
V_{disp}	Piston displacement per shaft revolution	m^3
T_r	Friction torque of a roller	$N \cdot m$
R_q	Equivalent surface roughness	m
$\vec{R}$	Roller position	m
$\vec{P}$	Linear momentum	$kg \cdot m/s$
$\vec{M}$	Moment	$N \cdot m$
$\vec{L}$	Angular momentum	$kg \cdot m^2/s$
$\vec{H}$	Hardness parameter	$-$
$\vec{F}$	Force	N
D_r	Roller dimeter	m
D_c	Cam diameter	m
C_f	Orifice discharge coefficient	$-$
$h_{def,i}$	Deformation gap of ith node	m
$\bar{h}$	Averaged fluid film thickness	m
$\vec{\omega}_r$	Roller angular velocity	rad/s
$\vec{v}_v$	Bottom surface velocity	m/s
$\vec{v}_t$	Top surface velocity	m/s
$\vec{r}_G^T$	Isotropic index parameter	$-$
$\vec{V}_r$	Roller velocity	m/s
$\vec{V}_c$	Contact point velocity	m/s
$\vec{R}_c$	Contact position	m
$\bar{P}^i$	Contact pressure polynomial	$-$
$[IM]_{i,j}$	Influence matrix	$-$

References

[1] H. Olsson, J. Ukonsaari, Wear testing and specification of hydraulic fluid in industrial applications, *Tribology International*, 36(11): 835-841, 2003.

[2] C. Zhang, H. Tan, Y. Fang, X. Zhang, Y. Yang, Y. Duan, M. Han, S. Cui, B. Xu, J. Zhang, Deformation pre-compensated optimization design of cam ring for low pulsation hydraulic motors, *Journal of Zhejiang University SCIENCE A* , 24(2) 130–145, 2023.

[3] U. Pettersson, S. Jacobson, Textured surfaces for improved lubrication at high pressure and low sliding speed of roller/piston in hydraulic motors, *Tribology International*, 40(2): 355-359, 2007.

[4] R. Lewis, Friction in a hydraulic motor piston/cam roller contact lined with PTFE impregnated cloth, *Wear,* 266(7–8): 888-892, 2009.

[5] D. Nilsson, B. Prakash, Investigation into the seizure of hydraulic motors, *Tribology International,* 43(1–2): 92-99, 2010.

[6] L. Dahlén , H. Olsson, Analysis of two sliding contacts inside a radial piston hydraulic motor, *Proceedings of the JFPS International Symposium on Fluid Power 2002*, (5-2): 537-542, 2002.

[7] P. Isaksson, D. Nilsson, R. Larsson, Elasto-hydrodynamic simulation of complex geometries in hydraulic motors, *Tribology International*, 42(10): 1418-1423, 2009.

[8] P. Isaksson, D. Nilsson, R. Larsson, and A. Almqvist, The influence of surface roughness on friction in a flexible hybrid bearing, *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology*, 225(10):975-985, 2011.

[9] X. Zhang, J. Zhang, B. Xu, Z. Yang, Q. Zhao, H. Zhang, The effect of slotted hole on minimum oil film thickness of piston in radial piston hydraulic motor, *Proceedings of the ASME/BATH 2021 Symposium on Fluid Power and Motion Control*, 2021.

[10] C. Zhang, X. Zhang, P. Dong, H. Zhang, Z. Zheng, J. Zhang, B. Xu, Composite thermal oil film lubrication model for hybrid journal bearings, *Tribology International* 194, 2024.

[11] C. Li, T. Jiang, C. Liu, H. Xu, G. Shi, Investigation of the leakage in the flow distribution pair of radial piston hydraulic motors through CFD analysis and experiments, *Flow Measurement and Instrumentation*, 96, 2024.

[12] T. Ransegnola, L. Shang, Andrea Vacca, A study of piston and slipper spin in swashplate type axial piston machines, *Tribology International*, 167, 2022.

[13] S. Mukherjee, L. Shang, A. Vacca, Numerical analysis and experimental validation of the coupled thermal effects in swashplate type axial piston machines, *Mechanical Systems and Signal Processing*, 220, 2024.

[14] T. Ransegnola, A strongly coupled simulation model of positive displacement machines for design and optimization, Doctoral Thesis, Purdue University, 2020.

[15] N. Patir and H. S. Cheng, An average flow model for determining effects of three-dimensional roughness on partial hydrodynamic lubrication, *Journal of Tribology*, 100(1): 12-17, 1978.

[16] N. Patir and H. S. Cheng, Application of Average Flow Model to Lubrication Between Rough Sliding Surfaces, *Journal of Tribology*, 101(2):220-229, 1979.

[17] W. Chengwei and Z. Linqing, An average Reynolds equation for partial film lubrication with contact factor," *Journal of Tribology*, 111(1):188-191, 1989.

[18] Si C. Lee and N. Ren, Behavior of elastic-plastic rough surface contact as affected by surface topography, load, and material hardness, *Tribology Transactions*, 39(1):67-74, 1996.

[19] M. Pelosi, M. Ivantysynova, The impact of axial piston machines mechanical parts constrain conditions on the thermo-elastohydrodynamic lubrication analysis of the fluid film interfaces, Internation Journal of Fluid Power, 14(3):35-51, 2013.

[20] B.J Hamrock, S.R. Schmid, & B.O. Jacobson, Fundamentals of Fluid Film Lubrication (2nd ed.). CRC Press. 2004.

Particle Dampers for Pump Noise Mitigation: Experimental Investigation of the Effects on Airborne and Structure-Borne Noise and Sound Perception on an Axial Piston Pump

Michael Lenz , Peter Zuppke , and Jürgen Weber

Chair of Fluid-Mechatronic Systems, TU Dresden, Germany
E-mail: michael.g.e.lenz@tu-dresden.de; juergen.weber@tu-dresden.de

Abstract

The reduction of noise emissions is a central target in positive displacement pump design, and the mitigation of structure borne noise is a main approach toward that goal. Particle damping offers a passive means of vibration mitigation without input energy requirement nor compromising the pump function and efficiency. The principle was applied to an axial piston pump and tested in sound power and structure borne noise measurements. A hollow encasing, which envelops most of the pump's sound emitting surface, is mounted, simultaneously providing particle damping and vibration decoupling properties. Finally, damping cavities are introduced modifying the pump's swash plate and end casing. Design considerations and experimental performance of the particle dampers with regards to sound power, sound quality, structure borne noise, and modal behaviour are presented and discussed. The mounted particle damped encasing provides a marked reduction of emitted sound power and improved sound quality. Meanwhile, no systematic acoustic improvement was observed for the particle dampers in the swash plate and end casing, which may be attributable to the pump's modal behaviour.

Keywords: NVH, pump noise mitigation, particle damping, sound perception / psychoacoustics

1 Introduction

The emitted noise of machines in operation is a central criterion for user satisfaction and acceptance. Exposure to high sound pressure level noise will lead to discomfort and possibly adverse health effects. Annoyance from noise is not solely a function of the integral sound pressure level, but may also be determined by the spectral distribution of sound pressure as well as swelling and transient effects. Also, personal sensitivity of auditors to sound characteristics will vary. The study of sound quality perception, as the correlation between objectively determinable characteristics of sound pressure signals and individual human sensation, is part of the field of psychoacoustics (e.g. [1, 2]).

Positive displacement pumps are frequently the dominant noise source of hydraulic drive systems. This is especially true for modern electrified drives, due to the disappearance of the prominent noise contribution of combustion engines and its masking effect. Pumps contribute to noise emissions by radiating airborne sound, as well as by inducing fluid and structure borne noise into adjacent components. Noise and vibration mitigation is therefore a central design goal for displacement pumps.

An effective approach is to tackle vibrations at their source, i.e. to optimize the drive geometry for minimal vibration excitation due to pressure build-up and flow pulsations, e.g. [3–7] . However, a base level of vibration excitation remains unavoidable due to the working principles of displacement pumps, and these mitigation measures are limited by the potential of additional leakage and resulting reduced pump efficiency. As an example, relief grooves in axial piston pump valve plates provide for a smoothed pressure build-up, but over a threshold length cause cross-porting and resulting internal leakage [3, 5, 8].

Active suppression of fluid [9] and structure borne [10] oscillations can yield significant attenuation effects, but may require costly additional electronic and machine components, which also potentially increase the sound emitting surface significantly as well as overall system complexity. In contrast, passive mitigation strategies do not require steering energy and can typically be realized with comparatively simple components. A classical example is resonance avoidance by tuning the natural frequencies of the pump and the attached hydraulic system so that they don't overlap with pump displacement frequencies and their higher orders; however this approach is much more complicated for variable speed modes of opera-

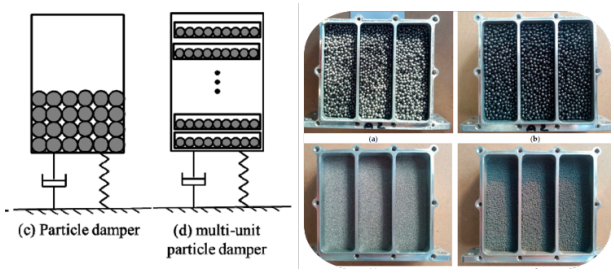


Figure 1: Working principle of particle damping, from [12]

tion which are gaining relevance in the context of electrification, as the spectrum of excitation frequencies becomes broadband. Other passive measures typically operate on the periphery of the pump and include the use of elastic hoses and motor pump unit bushings for vibration isolation or decoupling, silencers and Helmholtz resonators, acoustic insulation and barriers [11].

Particle dampers represent a further emerging means of structure borne noise mitigation. Figure 1 shows the basic setup: hollow spaces in the vibrating structure are partially filled with loose particles. Kinetic vibration energy is transmitted from the walls into the particles, resulting in chaotic or semi-structured [13] relative movement of the particle bed. Due to friction and inelastic impacts between the particles and between them and the walls, energy is dissipated, providing the damping effect. Damping performance is dependent on vibration frequencies and amplitudes, but typically effective over a wide range. [13–15] provide extensive discussion of the basic principles, modelling and influencing parameters of particle damping. Practical applications are reported on in [16, 17], achieving a dramatic reduction of vibration amplitudes.

As a first step towards the integration of particle dampers in displacement pumps, an experimental study was performed with bolt-on damper casings on a 92cc axial piston pump [18]. The casings design was based on an extensive experimental and simulative study of the pump's vibration behaviour, as well as guidelines on particle damper design and performance influencing parameters which were drawn from literature. Figure 2 shows the experimental setup and main sound mitigation result, characterized by the difference in the A-rated sound power level L_{WA}.

The main conclusions drawn from the study were as following:

- A moderate, but systematic sound power level reduction of 0.7 dB(A) on average was achieved for the optimal configuration of casings and particles.
- Steel balls of 2 mm in diameter consistently yielded the best results among the studied particle types.
- Sound power reduction resulted from a combination of damping and decoupling effects, as evidenced by the fact that already empty casings led to a reduction in certain frequency bands, which was increased significantly after filling in particles.
- Insufficiently stiff damper casings resulted in resonant vibrations of the dampers and drastically increased sound emission. The first generation printed plastic casings were replaced with metal ones, which resolved the issue.

For further study of particle damper performance, two approaches were followed, which will be reported on in the following. On the one hand, the concept of mount-on dampers was enhanced with the design of a shell enclosing most of the pump and featuring damper cavities. On the other hand, integrated damper cavities were manufactured in a pump end casing and swash plate.

2 Particle damped shell

2.1 Design

Observing the lessons learned from the bolt-on casings study, the shell was conceptualised to enclose as much of the pump's sound emitting surface as possible. Openings were left for the port connectors, the mounting flange, the control valve and the swivel angle sensor. The inner wall of the shell was placed at a distance of 15 mm from the pump surface, thus determining the thickness of the damping particle layer. While a higher value would be desirable, this was decided for as a compromise between damping effect and space constraints. The cavity

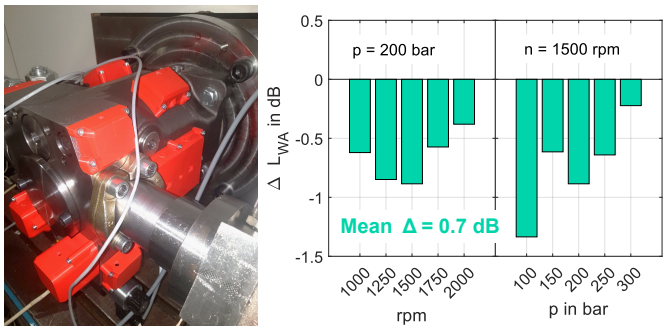


Figure 2: Bolt-on damper casings and achieved sound power level reduction [18]

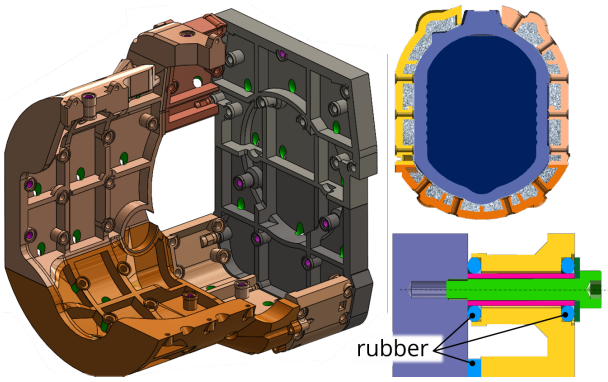


Figure 3: Design of enclosing shell and concept for elastic mounting

under the shell was subdivided into compartments with principal dimensions of approximately 40 mm following the conclusions drawn from [19], observed already in [18]. The subdivision is especially important in the vertical direction, so as to avoid the effect of particle movement being inhibited by the gravity of the particles above. Finally, due to manufacturing constraints the shell was divided into 8 separate parts to be bolted together. The resulting design is shown in Figure 3, with some parts omitted so as to show the inner wall structure.

Particular attention was paid to avoiding resonant vibrations of the shell, which had proved highly problematic in [18]. The solution there had been to employ metal made casings instead of printed plastic, raising their first natural frequencies above 5 kHz, so resonant vibrations of the casings where outweighed with relation to sound power by the particle damping and decoupling effects. Due to the complex and larger scale geometry, this solution was not possible for the shell in this study, with only a 3D printer for plastic material available for the task. To compensate as best as possible by stiffening the structure, a high wall thickness of 10 mm was set, and a carbon fibre reinforced ABS material was chosen for maximum elastic modulus. The compartment walls also serve a secondary function as a reinforcing rib structure. Despite these efforts, a cursory finite element modal analysis using estimated material parameters and boundary conditions found natural frequencies of bending modes starting below 230 Hz, far below a value where resonant vibration could be expected to contribute negligibly to sound power. Problematically low natural frequencies of the thin, multi-part shell structure had to be accepted.

As a further approach to mitigate this problem, the interfaces between pump and shell were devised to be highly elastic, aiming for a decoupling effect in the transmission of pump vibrations towards the shell. As shown in Figure 3, the bolt interfaces were designed such that forces are transmitted through rubber (NBR) O-rings, while still allowing for fastening of the bolts through a steel sleeve (pink). Furthermore, NBR foam tape was also applied to the compartment wall interfaces, on the one hand to provide an elastic mounting, and furthermore to seal the compartments, in order to avoid particle spill.

2.2 Experimental setup and procedure

The pump was installed on a hydraulic test rig in an anechoic chamber, shown in Figure 4. Two types of tests were performed:

On the one hand, sound power measurements at stationary operating points according to the provisions set out in ISO 3744 / DIN 45635-1/26, using 6 microphones on a cuboid enclosing surface limited by 2 reflecting planes. A low reflection line terminator (LRLT) at the end of the high pressure line was included in order to avoid reflections of pressure pulsations from the downstream hydraulic circuit. Apart from airborne sound, structure borne noise signals were measured by means of a laser vibrometer for surface velocity and by piezoelectric accelerometers, as listed in Table 2, and illustrated in Figure 5. Pressure pulsations in the suction and high pressure line were also measured, as well as mechanical and thermal con-

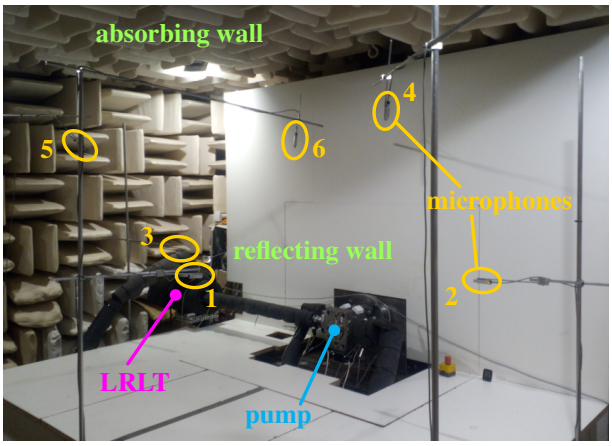


Figure 4: Sound power test rig and structure borne noise signals measured on the pump and damper shell

Table 1: Operating points for sound power measurements

	rev. per minute	high pressure / bar
base set	1000, 1250, ... 2000	200
expanded set	500, 750, ... 2000 500, 1500, 2000	200 50, 100... 300

Table 2: Measured and derived quantities

Symbol	Denotation	Unit
L_{WA}	Sound power level, A-rated	-
v_{EC}	Surface normal velocity on end casing	m/s
$a_{EC,i}$	Triaxial acceleration on end casing surface below shell	m/s^2
a_{PCi}	Normal acceleration on pump casing side wall below shell	m/s^2
a_{PCo}	Normal acceleration on shell surface over pump casing side wall	m/s^2
$H_k(i\Omega)$	Frequency response at point k	m/s/N

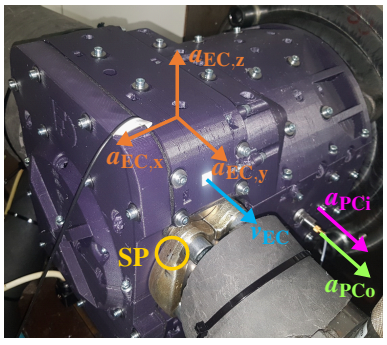


Figure 5: Structure borne noise signals measured on the pump and damper shell

trol and status observation signals from the test rig. Two sets of operating points were run through; a base set for assessment during parameter variation, and an expanded one for in-depth assessment of the most interesting configurations. The operating points are laid out in Table 1. All measurements were carried out at maximum swivel angle and a temperature of $(40 \pm 2)\,^{\circ}\mathrm{C}$, with two runs for each operating point.

Furthermore, frequency responses were determined on the still pump and shell configurations. Measurement points were defined on spots deemed highly relevant for structure vibrations, based on the operational vibration measurement reported in [18]. At these points k, using the laser vibrometer, the normal surface velocity response $v_k(t)$ to impulse hammer strikes (to the strike point SP on the suction line connection in Figure 5) was measured. The frequency response is the quotient of its Fourier transform and that of the measured impact force $f(t)$,

$$H_k(i\Omega) = \frac{V_k(i\Omega)}{F(i\Omega)}$$

The frequency response is useful for assessing changes in the transmission behaviour, here in particular damping and decoupling, between configurations. The main drawback is that the method only captures frequency dependencies, while it is known that particle dampers can exhibit significant dependency on vibration amplitudes as well.

2.3 Performance

Sound power reduction

After a reference measurement of the regular pump, the shell was mounted. Sound power measurements were performed with the shell empty and with the cavities filled to an increasing degree with 2 mm steel balls, which had shown the best results in [18]. The result is shown in Figure 6. In this diagram, the dots refer to individual measurements, while the bars mark the average for each operating point and configuration. As is evident, the addition of the empty shell increases the emitted sound power greatly for all operating conditions, repeating the result for the plastic damper casings in [18]. The cause can be attributed to resonant vibrations of shell parts as well as the increase in overall sound emitting surface. The former means that the intended decoupling effect was not achieved. Either the rubber coupling elements in Figure 3 are still too stiff, or manufacturing imprecisions and gravity caused direct contact between the shell and the pump casing or the screw sleeves, resulting in a mechanical short circuit. Adding particles mostly compensates for this effect, resulting in a sound power level of approximately the reference value. While it can be stated that the particles result in a clear damping effect, the result in comparison to the reference is unsatisfactory.

As a remedy to the apparently overly stiff attachment of the shell to the pump, the shell was mounted without bolt connections to the pump. Instead, the pump parts connect by bolts to each other, but to the pump only through the NBR foam tape on the cavity and outer wall interfaces of the shell, see Figure 3. Again, sound power measurements were carried out for the empty shell and steel ball filling, with the result on sound power levels depicted in Figure 7.

While the empty shell still leads to a substantial sound power level increase at several operating points, in particular at lower speeds, the effect is much less drastic for this configuration. Meanwhile, a clear damping effect is again observable with particles filled into the cavities, resulting in significant sound power level reductions. Remarkably, the degree of filling has little influence on the mitigation effect, and the results at 50% filling are no worse on average than at 80 or 90%, and not even much worse at 25%. This appears particularly relevant as a full filling of the shell adds a non-trivial 11.5 kg in particle mass.

For a clearer picture of how this result is obtained, the third octave band spectral distribution of sound power level is shown in Figure 8. The multiple lines of same colour again refer to individual measurements and are shown to give a sense of the replicability of the results shown. The comparison between the reference configuration and the empty shell clearly shows a decoupling effect in the higher frequency bands, starting from the one around 2 kHz. Sound power is markedly reduced in all bands above. At the same time, the empty shell produces a significant increase of sound power in several lower frequency bands up to the one at 2 kHz. This is plausibly attributed to resonant vibration at natural frequencies in this range, which are not fully decoupled.

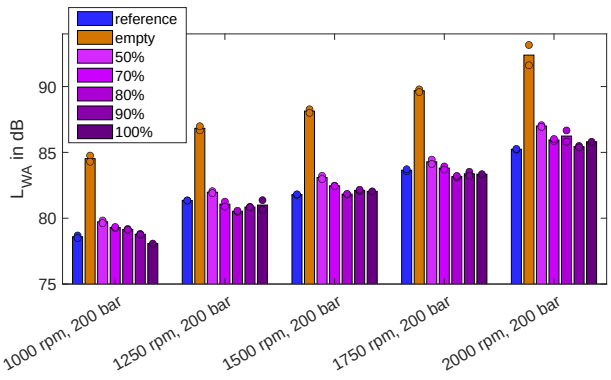


Figure 6: Sound power level for mounted shell filled with steel balls

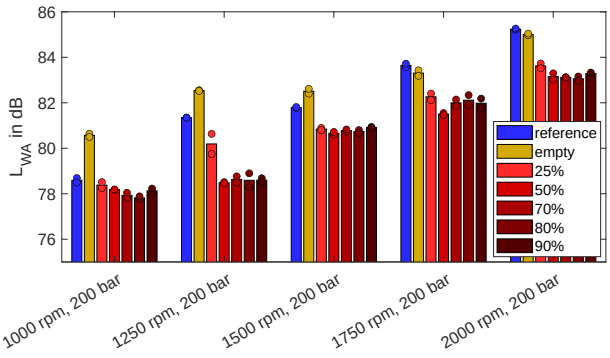


Figure 7: Sound power level for loosely attached shell filled with steel balls

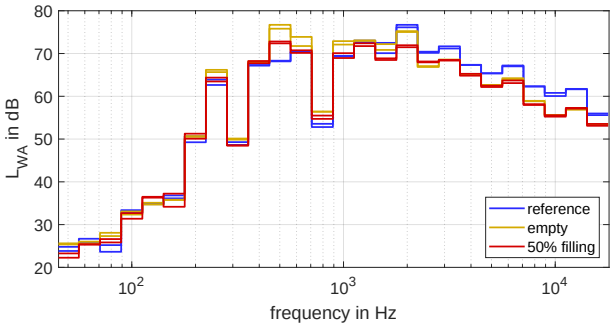


Figure 8: Third octave band spectrum of sound power level for loosely attached shell filled with steel balls at the operating point 1500 rpm, 200 bar

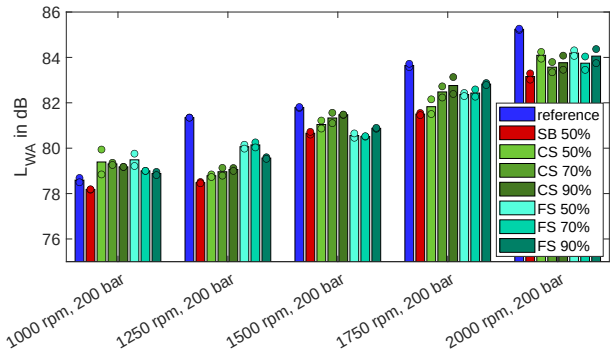


Figure 9: Sound power level for steel balls and sand particles as damping agent; SB.. steel balls; CS.. coarse sand; FS.. fine sand

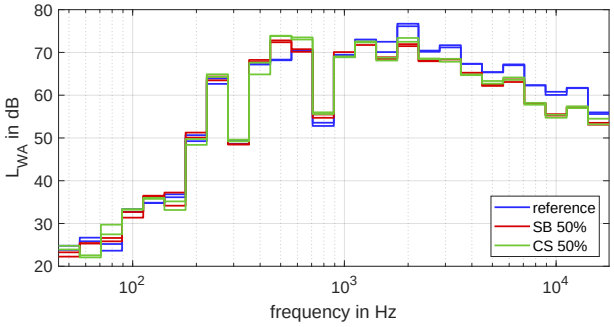


Figure 10: Third octave band spectrum of sound power level at the operating point 1500 rpm, 200 bar, comparing steel balls and coarse sand performance

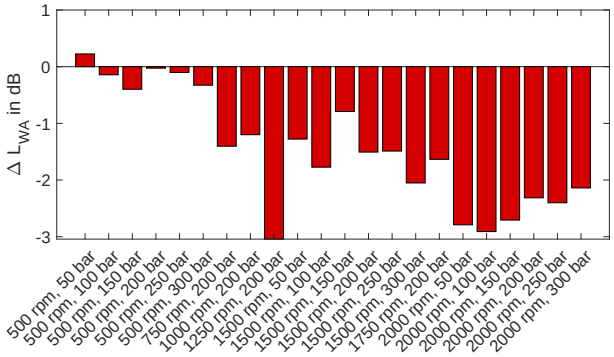


Figure 11: Difference of reference sound power level and the favoured damping configuration of 50% steel ball filling across all examined opterating points

The 50% steel balls filling mitigates these additional sound power contributions significantly, and eliminates it nearly completely in most bands. It also has a strong damping effect on the 2 kHz band, which is the most critical in the reference configuration. The result is the observed overall significantly reduced sound power. At the same time, the sound power spectrum is shifted towards lower frequencies. While the effect is shown for just one operating point here, it repeats qualitatively for all others.

All further studies were done with the loose shell attachment.

While in [18], the steel balls had proven to have the best damping effect compared to cast iron blasting agent of sizes 0.3...2 mm, as well as to irregularly shaped tin drops of 2-4 mm size, the study of the shell performance included an assessment of sand as a damping particle as a possible cost-effective alternative. Two variants were tested: a coarse gravel-sand with 1-2 mm grain size, and fine sand of 0.2-0.6 mm. Figure 9 shows the effects on sound power level. The sound power mitigation effect is reduced for all sand configurations in comparison to the steel balls at most operating points. However, there is still an appreciable and systematic reduction of sound power level with respect to the reference configuration, with the exception of the operating point at 1000 rpm. The exemplary spectrum in Figure 10 shows that the difference between steel and sand particles is mostly the latter having a lesser effect on the additional sound power caused by the shell in the lower frequency bands. This behaviour is found for fine sand as well across operating points.

Finally, Figure 11 presents the obtained sound power level reductions with the preferred damping variant. Across operating points, the average reduction is of -1.5 dB, with a clear tendency for greater reductions at higher revolution speeds. For the operating points at 2000 rpm the reduction is of -2.5 dB.

Sound perception

The above analysis demonstrated how the particle damper shell affects the spectral distribution of the pump's sound emissions. Is it widely known that individual sound perception, i.e. the assessment of sound quality and comfort of sounds, is related in various ways to their spectral distribution [1], while at the same time it is not always a clear function of L_{WA}.

Therefore, an effort was made to assess differences in sound perception of the reference and damped configuration, which may go beyond a comparison of sound power levels. While the research project did not allow for a fully fledged psychoacoustic assessment, which would require extensive randomized sound perception assessments with test subjects, comparative compilations of sounds were shared among colleagues and project partners. There was unanimous agreement that in all cases, the sound from the damped shell configuration is appreciably more comfortable to listen to. Remarkably, even the noise at 500 rpm / 200 bar is generally considered more comfortable for the damper shell configuration, despite the sound power level not being appreciably reduced as per Figure 11.

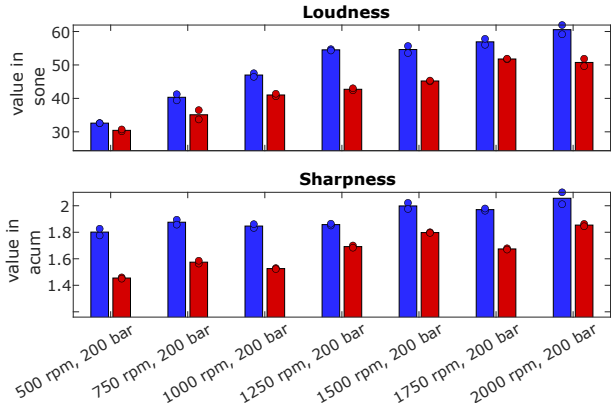


Figure 12: Psychoacoustic functions evaluation for microphone 1 signals

Sounds are made available to the reader on `https://tud.link/zv4ums`. Each sound file consists of an 8 s record of the reference configuration operation noise, followed by 8 s of that with the 50% steel ball filled shell. For each operating point there is a recording available for both microphone 1 - facing the pump end casing front - and microphone 2 - facing the pump laterally from its left (from positive y in Figure 4). The authors hope for a pleasant hearing experience. It has proved valuable to use as high quality headphones or speakers as available.

For a tentative quantitative assessment of sound perception, the psychoacoustic loudness and sharpness functions were computed[1] for the signals from microphone 1. The result is shown in Figure 12 .

Both values are indeed reduced for all operating points.

The loudness N is a measure intended to capture the average subjectively perceived noise volume more precisely than (A-rated) sound power or pressure levels [1] do. For a sine tone at 1 kHz, the relation to the sound pressure level L_p is [2]

$$N = 2^{\frac{L_\mathrm{p}-40\mathrm{dB}}{10\mathrm{dB}}}$$

With this formula, the loudness differences in Figure 12 are converted to apparent sound pressure level differences

$$\Delta L_\mathrm{p} = \frac{10\mathrm{dB}}{\ln 2} \ln \frac{N_\mathrm{SB50\%}}{N_\mathrm{reference}}$$

In Figure 13 these hypothetical level differences are compared to those computed directly from the measured signals. The difference is significant at lower pump speeds, where meagre or even negative measured level reductions contrast with more pronounced apparent reduction values when derived from loudness. Meanwhile, at higher speeds starting from 1500 rpm the difference between the values is negligible.

The sharpness value quantifies the high frequency contents of a sound, as high frequency noise is generally perceived as particularly unpleasant. There is a clear reduction in sharpness

[1] Computations according to ISO 532-1 / DIN 45631 (Zwicker method) for loudness and DIN 45692 for sharpness

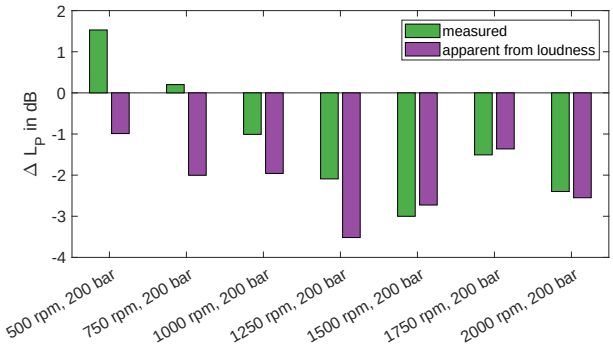


Figure 13: Measured and apparent sound pressure level differences at microphone 1

for all operating points, which corresponds to the observations on the sound power spectra in Figure 8.

Similarly to the results in Figure 13, the reduction is most pronounced at the 500 rpm operating point. Both may help explain the above observation of improved sound quality despite the unchanged sound power level for this speed.

Structure borne noise

The acceleration and surface velocity signals are evaluated for a deeper insight into the mechanisms of sound mitigation by the shell. Figure 14 shows the spectra of the 2 signals which were measured on the pump surface for the reference configuration and then on equivalent spots on the shell. Again, 2 spectra are shown for each configuration, representing individual measurements. The capital letters denote the Fourier transforms of the signals listed in Table 2. The frequency resolution was set to $\Delta f = 10\,\mathrm{Hz}$ for readability. The spectra exhibit the characteristic peaks at the pump order frequencies of $f_i = i \cdot rpm \cdot N_\mathrm{pistons}$, with $N_\mathrm{pistons} = 9$.

The comparison between the reference case and the empty shell corroborates the conclusions from the sound power tertiary band spectra: The empty shell leads to significantly increased vibration amplitudes in the frequency range <1 kHz, $A_\mathrm{PC,o}$, and <2 kHz for V_EC. This effect is mostly reversed after particles are added. At higher frequencies, already the empty

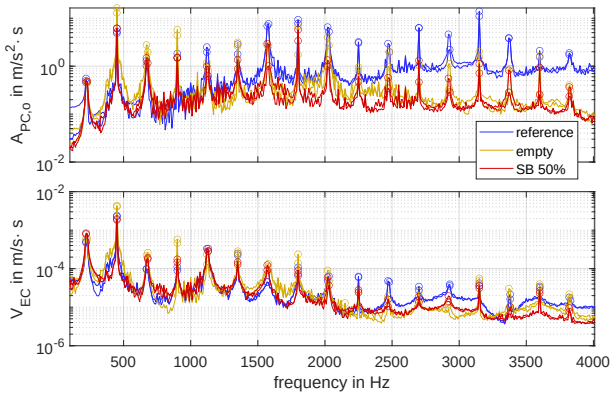


Figure 14: Vibration spectra on the sound emitting surface for 1500 rpm, 200 bar

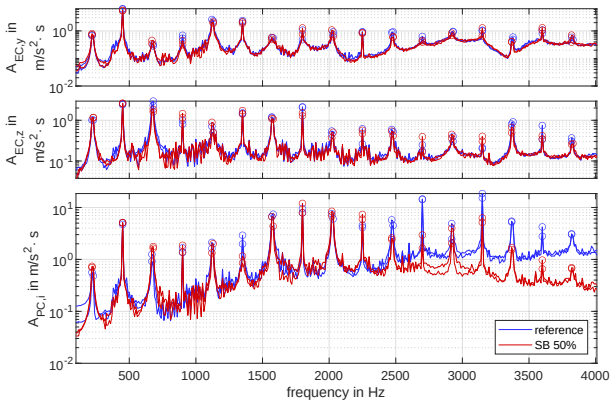


Figure 15: Vibration spectra on the pump surface for 1500 rpm, 200 bar

shell leads to a general reduction of the broadband spectrum and the peaks, starting from 1 kHz on the pump casing, and 2.2 kHz on the end casing. The particles provide a further amplitude reduction at most high frequency peaks of $A_{PC,o}$. This is true most clearly for the 9th order peak at 2.025 kHz, corresponding to the significant sound power reduction in that band in Figure 8.

The sound power reduction and spectral shift results also pose the question if they are purely due to decoupling effects from the elastically mounted shell and damping of its resonant modes, or if the pump vibration itself is damped. This is assessed in Figure 15, showing the spectra of those acceleration signals which were measured on the pump surface under the shell. The two represented components of A_{EC} on the end casing do indeed appear to be mostly indifferent to the particles. Some minor amplitude reduction is observed at several isolated peaks.

Meanwhile, $A_{PC,i}$ on the pump casing side wall shows significant and consistent amplitude reduction starting from the 11th order at 2.475 kHz, as well as at the 6th order(1.35 kHz) with the particle damped shell installed.

Frequency response

Measured frequency response functions on measurement points on the pump side walls are presented in Figure 16. The

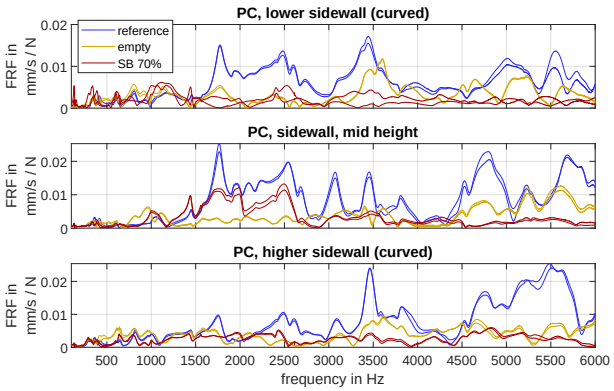


Figure 16: Frequency response measured on pump casing

decoupling effect is well visible with the empty shell leading to near erasure of the original resonant peaks above 1.5 kHz. However, the results with particles are difficult to interpret, and do not correspond well with the observations from the operational measurements.

3 Modified end casing and swash plate

Parallel to the running tests with the shell, a modified swash plate and end casing were manufactured to include cavities for particle damping. The main design motivation here was to exploit available space for cavity placement, which these parts offered so as to allow for a considerable cavity-to-volume fraction. The design concept and implementation on the test rig are depicted in Figure 17.

Cavity covers were made from 5 mm thick steel material, in order to increase stiffness and avoid sound emissions from plate vibrations as much as possible. For test rig experiments, the swash plate had to be installed with a particle filling which could not be feasibly varied, while that was possible for the end casing cavities. Sound power measurements were carried out analogous to Section 2.2. Figure 18 shows the results for various degrees of filling of the end casing cavities.

The pump with the modified, particle damped swash plate and empty end casing cavities has a slightly increased sound power emission compared to the reference configuration. The addition of particles leaves sound power largely unaffected. This

Figure 17: Swash plate and end casing with integrated particle dampers

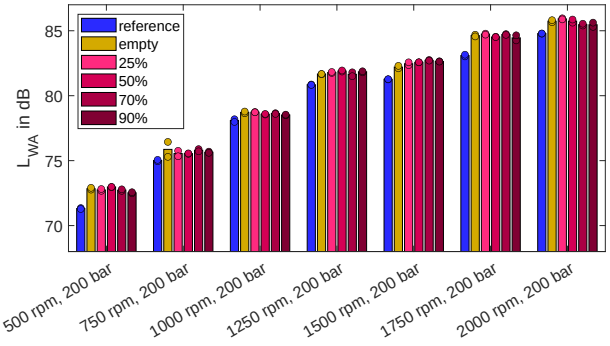


Figure 18: Sound power with integrated particle dampers

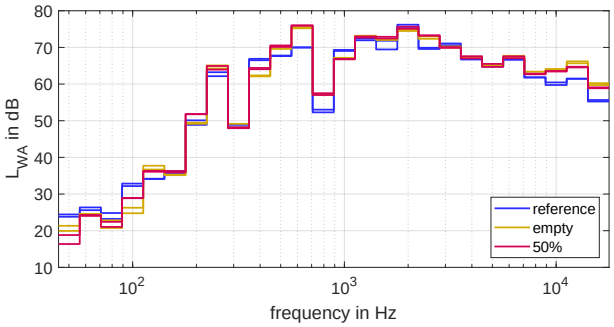


Figure 19: Sound power spectrum with integrated particle dampers at 1500 rpm, 200 bar

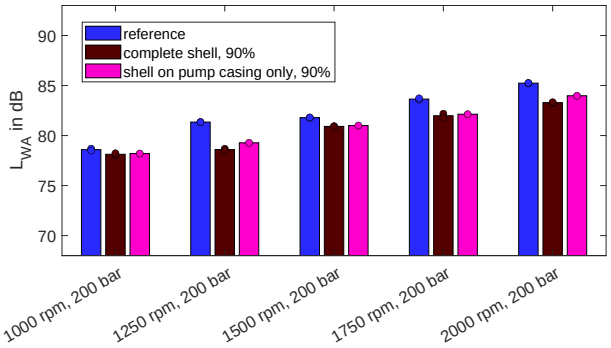


Figure 20: Sound power level comparing the damping shell enclosing the whole pump or the pump casing only

is also true for its spectral distribution, as is demonstrated in Figure 19. Differing from the experiences with the shell as well as with bolt-on casings made of plastic and metal in [18], in this case the particles do not compensate for the additional noise emissions of the empty configuration as compared to the reference.

4 Conclusion and outlook

Two approaches for noise mitigation of displacement pumps through particle damping were implemented and tested.

The elastically mounted and particle damped enclosing shell made of reinforced plastic was able to achieve considerable reductions of sound power depending on the operating point, with a mean value of -1.5 dB and the biggest improvements at high revolution speeds of 2000 rpm with a mean -2.5 dB. This was achieved through a combined mechanism of vibration decoupling between the pump and shell and particle damping. The application of the shell also shifts the spectral distribution of sound power towards lower frequencies, and an improvement of individually perceived sound quality was generally observed among test subjects, although statistically significant listening tests were not performed. The improved sound quality even holds for operating points where the sound power level was unchanged. The application of psychoacoustic valuation functions to the recorded sound signals supports this assessment. The evaluation of structural vibration signals improves the understanding of the mechanisms at work, and demonstrates the decoupling effect, as well as particle damping of the shell and also the pump structure itself. The latter applies to the pump casing at higher frequencies starting around 2.5 kHz, while the vibration of the end casing under the shell was not affected appreciably. Frequency response analysis corroborates the findings on the decoupling mechanism, but proved to be of limited use for assessing the particle damping effect, since it cannot capture its amplitude dependency, and likely the excitation magnitude is too small to yield appreciable damping.

The attempt for a simple improvement of an existing design by introducing damper cavities in available material volume on the other hand proved unsuccessful. Considering the subsequently obtained results from the damper shell regarding spectral distribution of air and structure borne noise, it appears that the sound mitigation effect stems near completely from the dampers on the pump casing. In fact, a measurement series with an incomplete shell leaving the end casing uncovered and enclosing only the pump casing is documented in Figure 20, showing only a minimal difference in emitted sound power between the two configurations. It appears therefore plausible that integrated dampers in the pump housing - which would however require more extensive design modifications - could perform better, although the results with modified pump parts presented here underscore the need for careful preliminary analysis.

For a possible future technical application of the shell, on the one hand room for optimization remains with regards to particle size, cavity design and elastic mounting. The shell as presented here was designed for handy laboratory use, and a technical solution will likely involve fewer parts. At the same time, additive manufacturing might not be feasible and a suitable production method for the required numbers will need to be found and the design adapted accordingly.

Acknowledgement

The presented research activities were carried out in the course of the project "Geräuschminderung in Verdrängereinheiten", funded by Forschungskuratorium Maschinenbau e.V. (FKM-No. 7052400). The authors would like to thank FKM and the Fluid Power Research Fund at VDMA for funding and support.

References

[1] Eberhard Zwicker and Hugo Fastl. *Psychoacoustics: facts and models (2nd, upd. ed.)*. Springer, 1999.

[2] Ville Pulkki and Matti Karjalainen. *Communication Acoustics*. John Wiley & Sons, Ltd, 1 edition, 2015.

[3] Maria Pettersson, Kenneth Weddfelt, and Jan-Ove Palmberg. Methods of Reducing Flow Ripple from Fluid Power Piston Pumps – A Theoretical Approach. *SAE transactions*, 100:158–167, 1991.

[4] Dirk Becher. *Untersuchungen an einer Axialkolbenpumpe mit Ringsystem zur Minderung von Pulsationen*. PhD thesis, TU Dresden, 2003.

[5] Ganesh Kumar Seeniraj and Monika Ivantysynova. Impact of Valve Plate Design on Noise, Volumetric Efficiency and Control Effort in an Axial Piston Pump. In *Proceedings of ASME IMECE2006*, pages 77–84. American Society of Mechanical Engineers Digital Collection, 2006.

[6] Ganesh Kumar Seeniraj and Monika Ivantysynova. A Multi-Parameter Multi-Objective Approach to Reduce Pump Noise Generation. *International Journal of Fluid Power*, 12(1):7–17, January 2011.

[7] Paul Kenneth Kalbfleisch and Monika Ivantysynova. Computational Valve Plate Design in Axial Piston Pumps/Motors. *International Journal of Fluid Power*, pages 177–202, September 2019.

[8] Ganesh Kumar Seeniraj, Zhao , Minming, and Monika Ivantysynova. Effect of Combining Precompression Grooves, PCFV And DCFV on Pump Noise Generation. *International Journal of Fluid Power*, 12(3):53–63, January 2011.

[9] Takao Nishiumi, Takayoshi Ichiyanagi, and Tetsuya Kuribayashi. Development of a novel helmholtz hydraulic silencer for attenuating the pressure ripple from a fixed displacement pump with variable rotational speed. *Proceedings of the 11th IFK, Aachen*, 2018.

[10] Taeho Kim and Monika Ivantysynova. Active vibration control of swash plate-type axial piston machines with two-weight notch least mean square/filtered-x least mean square (lms/fxlms) filters. *Energies*, 10(5):645, 2017.

[11] George E. Totten and Victor J. De Negri. *Handbook of Hydraulic Fluid Technology*. Taylor & Francis Group, Milton, United Kingdom, 2011.

[12] Hang Ye, Yanrong Wang, Bin Liu, and Xianghua Jiang. Experimental study on the damping eect of multi-unit particle dampers applied to bracket structure. *Applied Sciences, vol. 9 (14), p. 2912*, 2019.

[13] Niklas Meyer. *Design of Particle Dampers for Passive Vibration Attenuation*. PhD thesis, TU Hamburg, 2023.

[14] Tobias Ehlers. *Auslegung partikelgedämpfter Strukturbauteile für die Additive Fertigung*. Dissertation, Gottfried Wilhelm Leibniz Universität Hannover, 2023.

[15] Braj Bhushan Prasad, Fabian Duvigneau, Elmar Woschke, and Daniel Juhre. Application and damping mechanism of particle dampers. *PAMM*, 22(1):e202200116, 2023.

[16] Fabian Duvigneau, Sebastian Koch, Elmar Woschke, and Ulrich Gabbert. An effective vibration reduction concept for automotive applications based on granular-filled cavities. *Journal of Vibration and Control*, 24(1):73–82, 2018.

[17] Florian AM Vogel, Sebastian Berger, Ekrem Özkaya, and Dirk Biermann. Vibration suppression in turning $TiAl_6V_4$ using additively manufactured tool holders with specially structured, particle filled hollow elements. *Procedia Manufacturing*, 40:32–37, 2019.

[18] Michael Lenz, Peter Zuppke, Ahmed Shorbagy, and Jürgen Weber. Design strategy for noise reducing particle dampers on hydraulic pumps. *Maha Fluid Power Conference, Lafayette*, 2024.

[19] Thomas Künneke and Detmar Zimmer. Funktionsintegration additiv gefertigter Dämpfungsstrukturen bei Biegeschwingungen. *Additive Fertigung von Bauteilen und Strukturen*, pages 61–74, 2017.

An Inline Hydro-Mechanical Transmission (i-HMT) with a Rotatable Valve Cam to Improve Efficiency

Perry Y. Li , Evan Sand , Zihao Chen , and Dani Putra

Dept. of Mechanical Engineering, Univ. of Minnesota, Minneapolis, MN55455, USA
E-mail: { lixxx099,sandx114,chen6991}@umn.edu, daniputra0123@gmail.com

Abstract

A hydro-mechanical transmission (HMT) is a continuously variable transmission (CVT) that transmits power both hydraulically and mechanically, resulting in higher efficiencies than hydrostatic transmissions. While typical HMTs require a parallel mechanical transmission consisting of planetary and coupling gears, the inline configuration of the HMT (i-HMT), with a two-shafted pump, eliminates the need for such gearing making them less expensive, complex, and bulky. A potential drawback of the i-HMT is the compressibility-related throttling loss due to non-ideal valve timing. To reduce these losses and to improve efficiency, the concept of rotating the valve cam to adjust the valve timing for the different operating conditions is introduced. This paper focuses on the design, construction and testing of an experimental prototype. Experimental results comparing a stock eccentric cam and an optimized rotatable cam demonstrates that power efficiency can be improved by as much as 7% for a broad range of operating conditions.

Keywords: Hydromechanical transmission, pre-compression, de-compression, valve timing, valve cam.

1 INTRODUCTION

A hydro-mechanical transmission (HMT), like hydrostatic transmissions (HST), can vary its transmission ratio continuously. Traditional parallel HMT architectures, such as the output-coupled HMTs in Fig. 1-top, consist of a pair of hydrostatic pump/motors in parallel with a set of mechanical coupling and planetary gears. Since a portion of the power is transmitted via the more efficient mechanical path, an HMT is more efficient than a HST [1].

The large sizes and high costs associated with parallel HMTs are major barriers to their use. The Hondamatic (see e.g. [3]), shown in Fig. 1-bottom and Fig. 2, is an alternate, inline implementation of the HMT. Since planetary and coupling gears are not required, the inline HMT (i-HMT) is smaller and less complex. The Hondamatic transmission has been deployed in some commercial all-terrain vehicles (ATVs) since the 2000's. The i-HMT consists of:

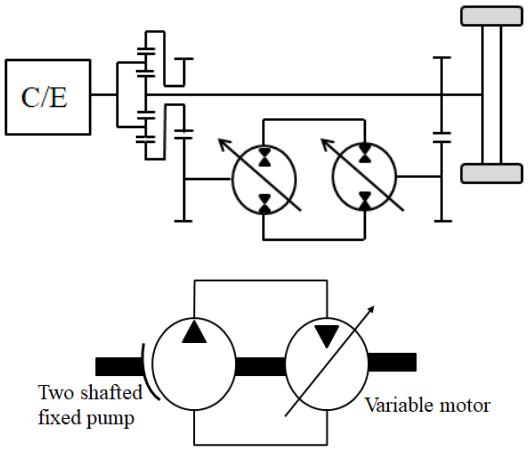


Figure 1: Schematics of Output-coupled HMT. Top: Traditional implementation [2]; Bottom: In-line implementation.

1. A rotatable wobble plate (labeled as shoe plate in Fig. 2) acting as the input for the pump.
2. Opposing axial piston pump and motor barrels fixed to a shared central shaft acting as the output.
3. A variable-angle motor swashplate that is rotationally fixed to the transmission case.
4. A system of radial distributor valves and valve cams connecting the pump and motor cylinders hydrostatically to high-pressure and low-pressure rings. The valve cam for the pump rotates with the wobble plate and pump housing; the one for the motor is fixed.

The wobble plate of the pump is the mechanical input, and the shared central shaft is the output shaft. The two-shafted pump (as the wobble plate and barrels can rotate independently) splits power between the mechanical and hydraulic channels,

assuming the role of a $1:1:1$ planetary gear set on the left side of Fig. 1-top. Therefore, the Hondamatic i-HMT is an output-coupled input-split HMT.

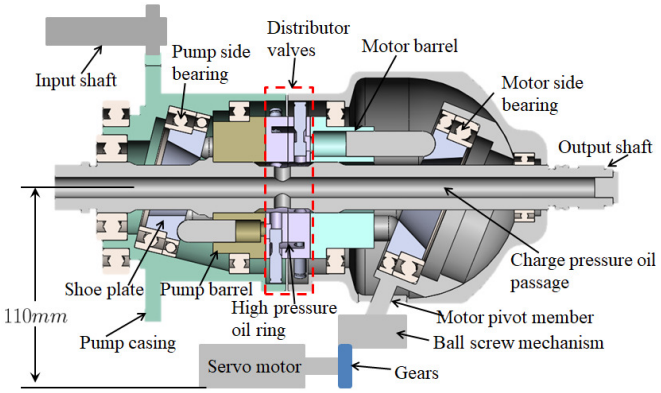


Figure 2: Cross-section view of the stock Hondamatic i-HMT [2].

From Fig. 1-b, using the fact that pump speed is the relative speed between the wobble plate (input shaft) and the barrel (output shaft), and flow continuity between the pump and motor, the ideal transmission ratio can be derived as:

$$i_{ideal} = \frac{\omega_{in}}{\omega_{out}} = \frac{T_{out}}{T_{in}} = \frac{D_p + D_m(t)}{D_p} \tag{1}$$

where ω_{in} and ω_{out} are the input and output shaft speeds, T_{in} and T_{out} are the input and output torques, D_p is the fixed pump displacement and $D_m(t)$ is the variable motor displacement. As $D_p \approx 6cc$ and $D_m(t) \in [0, 18cc]$, $i_{ideal} \in [1,4]$. When $D_m(t) = 0$, $i_{ideal} = 1$ as the transmission is locked up hydraulically. The input shaft, wobble plate, barrels of the pump and motor, and output shaft rotate as a rigid body. Power is transmitted purely mechanically. See [2] for a more detailed description and modeling of the functionality of the iHMT.

The performance of a stock Hondamatic i-HMT was experimentally evaluated in [2]. A possible power loss is throttling in the distributor valves due to the non-ideal fluid precompression and de-compression achieved with the stock eccentric circular valve cam. For instance, at zero motor displacement ($D_m = 0$), when the net motor flow is zero, the dead volumes in the motor piston chambers are repeatedly exposed to high and low pressures leading to wasteful repeated fluid compressions and de-compression cycles.

To address these compressibility losses, a second valve mode in which the motor piston chambers are closed to both high and low pressures was proposed in [4] specifically at lock up ($D_m = 0$). A rotatable valve cam concept was introduced as a secondary control input in [5] to vary the advance in valve timing. This approach applies to all motor displacements not just $D_m = 0$. The present paper describes the application of this latter concept to the Hondamatic - including the design, construction and experimental testing of a prototype.

The idea of rotating the port plate to adjust the displacements of piston machines has been proposed previously, e.g. recently in [6] and as a means of varying the transformation ratio of a hydraulic transformer [7]. The concept of using valve timing to control displacement is the premise of digital displacement machines in the partial stroke such as in [8, 9], typically using electric driven valves. This paper differs from these previous works in that the valve cam of the distributor valves is rotated in conjunction with a swash-plate to control both the displacement and the pre-compression.

The rest of the paper is organized as follows. Section 2 explains the valve timing in the stock Hondamatic i-HMT. In Section 3, the concept and the design of rotary cam are presented. Section 4 presents the cam design procedure. Section 5 presents the experimental test-stand and the testing procedures. Experimental results are presented in 6. Concluding remarks are presented in Section 7.

2 Valve Timing in the Stock i-HMT

The Hondamatic i-HMT uses a set of radial distributor valves instead of port plates to connect the piston chambers of the pump and the motor to high or low pressures (Fig. 3). When the radial spool valves move outwards, the piston chambers are exposed to the center low-pressure ring; and when the spool valves move inwards towards the center, the chambers are exposed to the high-pressure ring. The chambers are closed to both pressures in a small intermediate range of spool positions. As the radial valves rotate with the barrel, the distance between a fixed outer cam ring and the center of the rotating motor barrel determines the valve opening to the high or low pressure. In the stock Hondamatic, the cam is an eccentric circle, so the motor piston chambers are exposed to high pressure during the intake (motoring) half of the cycle and to low pressure during the exhaust half. The chambers are closed to both the high and low pressures prior to the top and bottom dead centers (TDC/BDC), presenting opportunities for pre-compression and de-compression.

Ideally, the decrease in fluid chamber volume during *pre-compression* would increase the fluid pressure from the low-pressure level during exhaust to exactly the high-pressure level when the valve is open to high-pressure. This would eliminate the throttling loss due to the pressure difference across the valve when it opens. Similarly, the increase in fluid chamber volume during *de-compression* would ideally decrease the fluid pressure from the high-pressure level during motoring to exactly the low-pressure level when the valve is open to low-pressure.

With the stock eccentric cam, the angles of the start and end of pre-compression (θ_{ps}, θ_{pe}) and de-compression (θ_{ds}, θ_{de}) are constant regardless of the pressure in the high-pressure ring or of the swashplate angle α_{sw}. The volume change during pre-compression and de-compression are:

$$\Delta V_p = \frac{A_p D}{2} \tan(\alpha_{sw}) \cdot (\cos(\theta_{ps}) - \cos(\theta_{pe})) \tag{2}$$

$$\Delta V_d = \frac{A_p D}{2} \tan(\alpha_{sw}) \cdot (\cos(\theta_{de}) - \cos(\theta_{ds})) \tag{3}$$

where α_{sw} is the swashplate angle, A_p and D are the piston area and the pitch cylinder diameter respectively. With the stock cam, $\theta_{pe} = 0^\circ$ and $\theta_{de} = 180^\circ$.

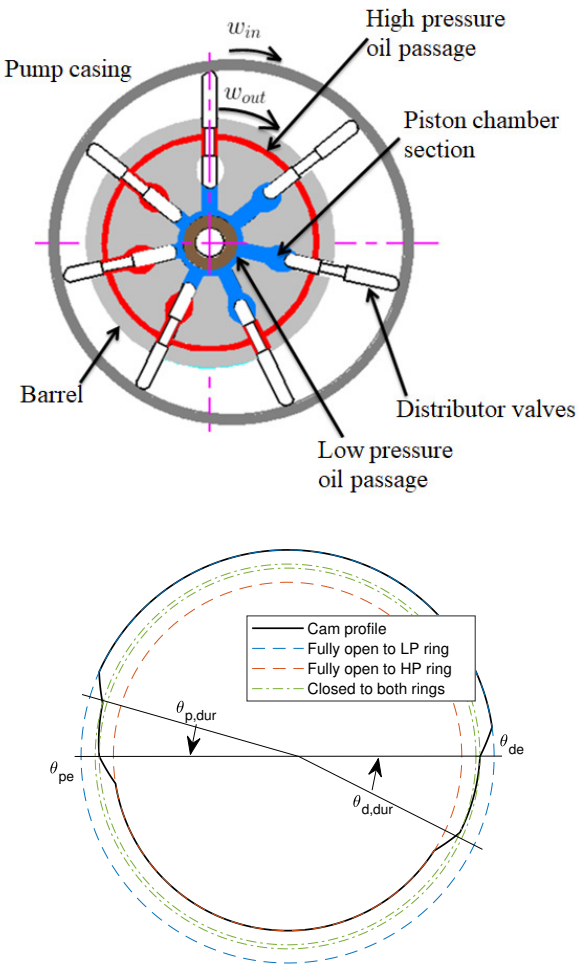


Figure 3: Top: (Exaggerated) Schematic of the stock eccentric cam profile; Bottom: Conceptual (non-circular) cam profile incorporating pre-compression and de-compression ($\theta_{pe} = 0$). Piston barrel is assumed to rotate counter-clockwise in bottom figure.)

On the other hand, the desired amount of volume change to avoid throttling depends on and increases with these quantities:

1. the dead volume in the piston chamber at the end of pre- and de-compression, given by:

$$V_{dead}(\theta_{pe}) = V_0 - \frac{A_p D}{2}\tan(\alpha_{sw})\cos(\theta_{pe}) \tag{4}$$

   where V_0 is the dead-volume at zero swashplate angle ($\alpha_{sw} = 0$).

2. high-pressure level which in turn depends on input torque

3. compressibility of the fluid which depends on the amount of dissolved air content and pressure.

As these quantities change with the operating conditions of the transmission, the stock eccentric circular cam's fixed set of pre-compression and de-compression angles $\theta_{ps}, \theta_{pe}, \theta_{ds}, \theta_{de}$ cannot achieve the ideal pre-compression and de-compression for all pressures and motor displacements. In particular, as α_{sw} decreases, the dead-volume $V_{dead}(\theta_{pe})$ in (4) increases, thus requiring more volume change during pre-compression. However, the actual volume change ΔV_p in (2) provided by the fixed cam decreases (Fig. 4). The situation is slightly better for de-compression as both dead-volume and ΔV_d decrease.

3 Valve Timing Advance via a Rotatable Valve Cam

To cope with the varying required pre-compression and de-compression volume changes, an additional control degree-of-freedom is introduced by allowing the valve cam to rotate. Let the fixed pre-compression and de-compression durations as determined by a physical cam design be:

$$\theta_{p,dur} := \theta_{pe} - \theta_{ps} \tag{5}$$
$$\theta_{d,dur} := \theta_{de} - \theta_{ds} \tag{6}$$

By rotating the cam by $\Delta\theta(t)$ in the opposite direction of barrel rotation, valve opening and closing occur earlier so that:

$$\theta_{pe} = -\Delta\theta(t), \quad \theta_{ps} = -\Delta\theta(t) - \theta_{p,dur}$$
$$\theta_{de} = -\Delta\theta(t) + 180^\circ, \quad \theta_{ds} = -\Delta\theta(t) + (180^\circ - \theta_{d,dur})$$

From (2)-(3), advancing the valve timing up to 90° increases the magnitude of the volume changes during pre-compression and de-compression ΔV_p and ΔV_d. The fluid volume to be pre-compressed is increased and the volume to be decompressed is reduced. In this way, the valve timing advance of $\Delta\theta(t)$ can be adjusted to be close to the optimal pre-compression and de-compression as the swashplate angle, high-pressure level, and fluid compressibility change.

In addition to its effect on ΔV_p and ΔV_d, advancing valve timing also decreases the effective displacement of the motor. It is because the motoring stroke is shortened as the high-pressure valve is closed prior to TBC, and the portion of the cycle prior to TDC is pumping instead of motoring, thus decreasing the net flow input from the high pressure over one cycle. Indeed, the effective fractional displacement relative to a motor in which the chamber is open to high-pressure between $0 - 180^\circ$ is:

$$\begin{aligned}\bar{D}_m(\alpha_{sw}, \Delta\theta) &= \frac{\tan(\alpha_{sw})}{2\tan(\alpha_{sw,max})} \cdot [\cos(\theta_{pe}) - \cos(\theta_{ds})] \\ &= \frac{\tan(\alpha_{sw})}{2\tan(\alpha_{sw,max})} \cdot \left[\cos(\Delta\theta) + \cos(\Delta\theta + \theta_{d,dur})\right]\end{aligned} \tag{7}$$

which shows that if $\Delta\theta > 0$, a larger swashplate angle α_{sw} is needed to maintain the same effective displacement when $\Delta\theta = 0$. Nevertheless, judicious combinations of α_{sw} and $\Delta\theta$ can achieve a range (albeit smaller) of desired motor displacements (hence i-HMT transmission ratios) while maintaining close-to-optimal pre-compression and de-compression.

With valve timing advance, the valves are open and closed as the pistons are moving. To minimize the throttling loss, the valves should be open and closed as quickly as possible.

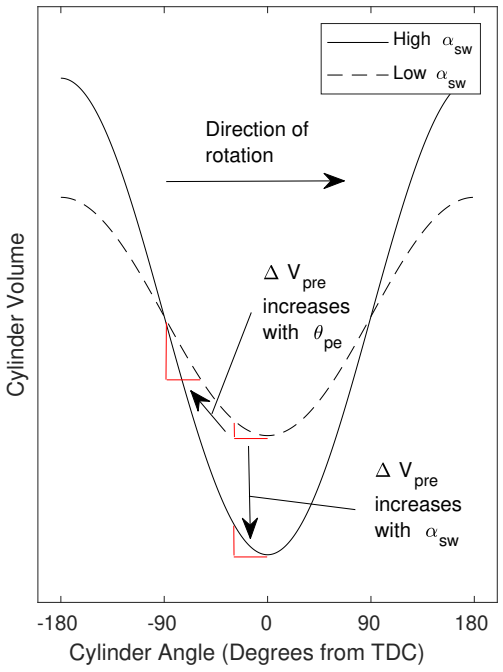


Figure 4: Effects of swashplate angle and cam angle on the change in cylinder volume during pre-compression: ΔV_{pre} increases with both $|\theta_{pe}|$ and α_{sw}.

This is achieved by the short transitions between valve opening and closing as illustrated in the cam profile in Fig. 3-bottom. The maximum rate is limited by the allowable pressure angle for the cam. In the stock eccentric circular cam, because the valves are open and close at TDC and BDC where the pistons are nearly stationary, the rate of valve opening is not as important.

4 Cam profile design

4.1 Design Strategy

For a given valve cam and operating pressure, settings for the cam angle θ_{pe} and swashplate angle α_{sw} may be chosen to achieve a prescribed motor displacement while exactly optimizing either pre-compression or de-compression but not both. A solution, suggested in [5], is to use passive check valves on the low-pressure side, at the expense of added complexity. Because of the effect of entrained air on the bulk modulus of the working fluid, the exact timing of pre-compression is more important: the bulk modulus of the working fluid is much lower at the end of de-compression (at low pressure) than at the end of pre-compression, so a non-optimal timing would not affect the cylinder pressure at the end of de-compression as much as it would affect the pressure at the end of pre-compression. For this reason, our strategy is, for a given cam profile (e.g. as specified by $\theta_{p,dur}$ and $\theta_{d,dur}$), to control the cam and swashplate angles to compensate for pre-compression while the cam profile is designed to minimizes throttling loss at the end of de-compression.

Based on this approach and the geometry of the motor, the optimal motor settings are calculated as in [5], but with the addition of:

1. Orifice leakage flows past the valve spool between the cylinder and both the high and low pressure rings
2. Laminar leakage through the piston-cylinder bore gap

Example relationships between the desired transmission ratio i_{ideal} and the cam timing advance angle $\Delta\theta$ and swashplate angle α_{sw} can be found in Fig. 11.

The analysis of pre-compression and de-compression requires a model of fluid compressibility. Our assumption is that the hydraulic fluid is a mixture of pure hydraulic oil with a constant bulk β and a fixed amount of entrained air that undergoes adiabatic compression/expansion. The volume fraction [1] α of the entrained air is assumed to be either 1% or 5% at atmospheric pressure.

Let p_h and p_t be absolute pressures in the high-pressure ring and in the tank, the optimal change in volume during pre-compression at full displacement is:

$$\Delta V_p = V_{dead}(\theta_{pe}) \cdot \left(\frac{1}{r_v} - 1 \right) \tag{8}$$

where $V_{dead}(\theta_{pe})$ is the cylinder volume at the end of de-compression and r_v is the volume ratio of the working fluid between p_h and p_t:

$$r_v = \frac{e^{-(p_h - p_t)/\beta} + \alpha \left(\frac{p_h}{p_t} \right)^{-1/\gamma}}{1 + \alpha} \tag{9}$$

where α is the volume fraction of the entrained air and $\gamma = 1.4$ is the polytropic index for the adiabatic process.

4.2 Optimization of Cam Parameters

To design the cam profile, it was first necessary to optimize the cam parameters $\theta_{p,dur}$ and $\theta_{d,dur}$, which are the angular durations of pre-compression and de-compression, respectively.

The motor settings are designed to compensate for pre-compression exactly for any operating condition. This causes valve-timing for de-compression to be non-optimal. Thus, the cam parameters are optimized to minimize the efficiency loss from non-optimal de-compression. To this end, an objective function was constructed which:

1. takes as inputs the cam parameters $\theta_{p,dur}$ and $\theta_{d,dur}$
2. iterates over the range of operating conditions possible with the test stand, and at each condition:
   (a) Prescribes motor settings to exactly optimize pre-compression
   (b) Evaluates the cylinder pressure at the end of de-compression using a model of pressure dynamics

[1] Ratio of air volume to pure liquid volume at low pressure

(c) Computes the resulting energy loss per cycle from non-optimal de-compression, using the principle of hydraulic effort [10]

(d) Computes the effect of the energy loss on efficiency "efficiency loss" by dividing by motor input power

(e) Adds a penalty if the pressure at the end of de-compression is below atmospheric pressure, or if the target transmission ratio tolerance cannot be achieved

3. Returns an average of the losses and penalties at all operating points for the cam parameters.

This objective function was then optimized using the `surrogateopt` function in MATLAB to find the cam parameters which would minimize the average impact of non-optimal de-compression on efficiency across all feasible operating conditions.

For $\alpha = 5\%$, the optimal parameters are: $\theta_{p,dur} = 5.82^\circ$ and $\theta_{d,dur} = 10.89^\circ$. Optimized cam parameters for $\alpha = 1\%$ were also obtained. Since testing for this case was limited, it will not be discussed.

4.3 Design of Cam Profile

Having selected the angular durations of pre-compression and de-compression, a spline profile was designed for the each cam. The following design constraints were used:

1. The profile must pass through the pre-compression and de-compression control points i.e. the valves must open and close to the HP and LP rings at the prescribed angles.
2. The pressure angle of the cam surface on the valve spools must not exceed 30°.
3. The valve spools must not separate from the cam surface at any operating condition.
4. The force at the cam-valve spool interface must not exceed 100 N at any operating condition.

Piecewise radial periodic cubic splines were used to design the profiles, using MATLAB's `perspline` function [11]. Control points were placed heuristically based on these constraints, and a resulting spline is shown in Fig. 5.

4.3.1 Cam Force Analysis

The forces at the cam-valve spool interface were evaluated to avoid excessive contact force and separation. The following forces were considered:

1. Inertial force of spool acceleration
2. Inner pressure of LP ring
3. Centrifugal force
4. Spikes in inward steady-state flow force on spool at the end of de-compression (the worst case for each speed)

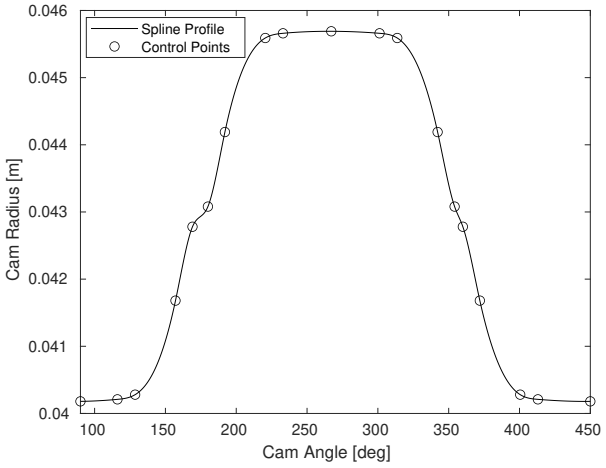


Figure 5: Cam radius as function of angle ($\alpha = 5\%$ case).

Contact force is a function of the motor speed and the cam angle. The contact force is between $19 - 39$ N for the stock cam and between $10 - 56$ N for the new cams across all operating conditions.

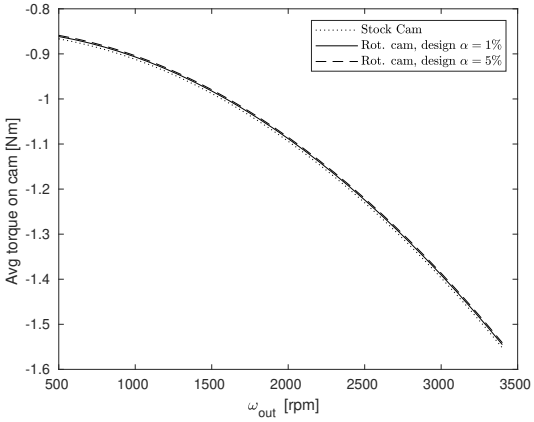


Figure 6: Torques from valve spools on cam for original cam and both rotatable cams, considering Coulomb friction and pressure angle effects.

4.3.2 Cam Torque Analysis

To verify that the torque applied to the new rotatable cams by the valve spools was not excessive, the torque was analyzed for both the stock cam and the new cams. The torque analysis considered torque from radial spool force based on pressure angle effects and Coulomb friction with $\mu = 0.12$. Torques from one spool were computed as functions of angle and motor speed, superimposed with an angle offset for all nine cylinders of the motor, and averaged over one rotation to find the average torque as a function of motor speed. As shown in Fig. 6, the magnitudes of torque for both rotatable cams are very slightly less than the torque on the original cam at all speeds. The torque did not exceed 2 Nm for any of the three cams. Based on this, it was determined that the effect of this torque on power loss was negligible, and that holding torque should

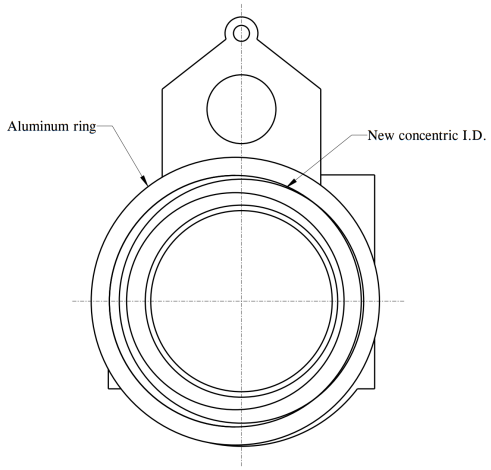


Figure 7: CAD model of modified motor case

not be a major concern in the actuation mechanism for cam rotation.

4.4 Mechanical Design

To design a method of actuating the rotatable cam in the assembled Hondamatic i-HMT in the test stand, a CAD model was created of the motor case based on physical measurements.

The stock motor case has an eccentric inner diameter which holds the circular cam ring. It was necessary to create a concentric inner diameter in the same place, in which the new cams could be rotated about the axis of the central shaft. A concentric cut could not be made without breaking through the case walls, so an aluminum ring was brazed onto the case before machining the concentric ID to match the wall thickness of the original motor case. The modified case is shown in Fig. 7.

To actuate the rotatable cams, an arm was designed to attach to the cams via a dowel and a threaded stud, passing through a slot in the motor case. The end of the actuation arm was fitted with a swiveling nut, which travels along a lead screw driven by a stepper motor. A photograph of the lead screw actuation is shown in Fig. 8. The motor swashplate was also actuated by a lead screw driven by a stepper motor, so the cam angle and swashplate angle could be adjusted together according to the prescribed operating conditions.

5 Experimental Testing

The mechanical input and output of the test stand used for performance testing were a pair of 28 cc Sauer/Danfoss S-42 variable-displacement hydrostatic pump/motors. They were connected in a regenerative configuration, meaning that the pump outlet was connected to the 20.68 MPa (3000 psi) high-pressure supply to re-capture the output power of the transmission. A 2.76 MPa (400 psi) charge pressure for the pump and motor was maintained using a pressure-reducing valve. The low-pressure ports of the hydrostatic pump/motors were kept at 2.76 MPa (400 psi) with a pilot-operated relief valve. A picture and a schematic of the test stand are shown in Fig. 9.

Inside the transmission case, the input shaft of the Hondamatic i-HMT is driven by gears with a 15:13 ratio as shown in Fig. 10. Charge fluid for the Hondamatic i-HMT was provided in a separate circuit (not shown in Fig. 9) by a small pump from the test stand case to the charge inlet. A constant 6.89 MPa (100 psi) charge pressure was maintained using a pressure-reducing valve.

These measurements were taken at the input/output shafts:

1. Input torque, with a Futek TRS300 torque cell on the input shaft
2. Input speed, with an RLS AM8192B magnetic encoder mounted on the input shaft
3. Output torque, with a Futek TRS300 torque cell on the input shaft
4. Output speed, with the S-42 pump's built-in encoder.

The input speed and output torque were regulated by adjusting the displacements of the test stand's hydrostatic motor/pumps at the input and output, respectively. Since only steady-state conditions are of interest, simple P-I controllers were used to regulate the input speed and the output torque.

As performance is sensitive to oil temperature, a temperature sensor was placed in the oil-sump to ensure that data were taken only when the temperature was between $35-40^{\circ}$C.

Two stepper motors were used to adjust the cam and swashplate angles of the Hondamatic motor. Prior to performance testing, the optimal target angles were obtained by adjusting slightly from the analytical value to achieve the desired ratio more precisely and to improve on the efficiency. The changes are shown in Fig. 11-top. During testing, commands were given to two stepper motor controllers based on these modified settings. Because the stepper motors are undersized, the cam and swashplate angles were adjusted only when the transmission was at rest.

The test stand was operated in these ranges of conditions:

- Design entrained air content: 5%
- Input speed range: 1000-3000 rpm
- Output torque range: 10-100 Nm
- Target transmission ratio: 1 to 4
- Input power: 1 kW to 25 kW
- Operating pressure: 10 MPa to 50 MPa

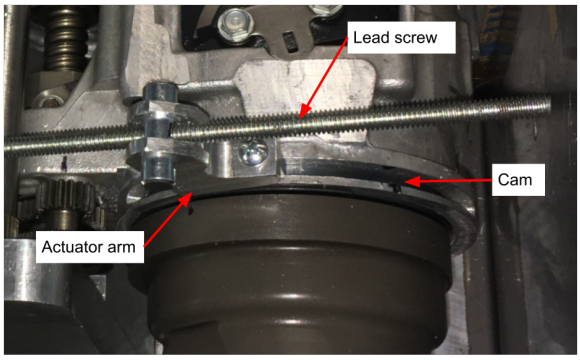


Figure 8: Lead screw actuation of rotatable cam.

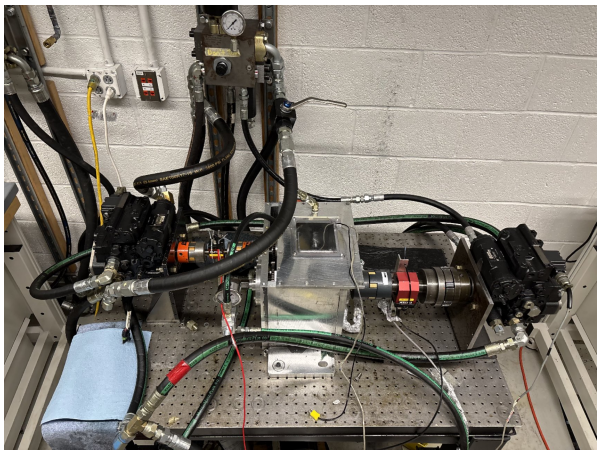


Figure 9: Picture and schematic of test stand used for performance testing.

At each condition, steady-state input and output torques and speeds were recorded and time-averaged to evaluate the achieved speed ratio and power efficiency. Sample input speed and output torque tracking performances are shown in Fig. 12.

6 Results

The efficiency (output power/input power) contours of the i-HMT with stock cam and the optimized 5% cam as functions of input speed and torque are shown in Fig. 13. The efficiencies between the stock and modified cams are directly compared in Figs. 14 and 15. Note that speed ratios in these figures correspond to the ratios between the input speed of the test stand (not of the iHMT) and the output speed. For both the stock and 5% cams, efficiency tends to be lower at low output torque and increases as output torque increases. The peak efficiencies achieved by the stock cam is about 75% whereas the 5% cam achieves a peak of about 80%. For both cams, the peak efficiency for ratio=1 is less than other ratios. When output torque is greater than or equal to 20Nm, the efficiency is higher with the modified cam than with the stock cam. Ignoring the exceptionally low efficiency of the stock cam at 40Nm and ratio = 1, the increase can be as much as 7% (absolute).

Comparing Figs. 14 and 15, it is apparent that at 1000-2000 RPM, the efficiencies of both cams do not vary with speeds significantly. This is confirmed in the surface plots in Fig. 16 where the efficiency surfaces at 1000 RPM and 2000 RPM are practically the same.

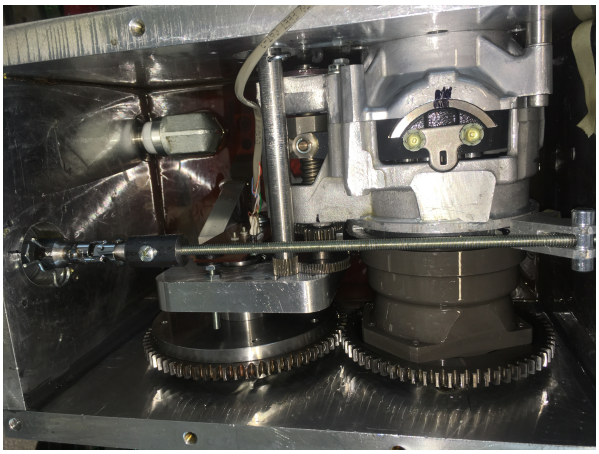


Figure 10: i-HMT inside the test stand transmission case.

In Fig. 17, the efficiencies of both cams are plotted against input torque for all the data points. Whereas Figs. 14-15 and 16 show efficiency as a function of output torque and ratio, Fig. 17 indicates that efficiency is mainly a function of input torque. As input torque, input gear ratio (15:13), and system pressure $p_h - p_t$ (difference between high and low pressure) are directly related by:

$$T_{in} = \frac{13}{15}\frac{D_p}{2\pi}(p_h - p_t)$$

this suggests that efficiency is dictated mainly by the system pressure. In particular, efficiency is highest at about 30 Nm which corresponds to 27 MPa. The maximum working pressure of the Hondamatic is 50-60 MPa.

7 Conclusions

A rotatable valve cam with an optimized profile has been proposed as a method of reducing the compressibility losses in the variable displacement motor of an inline hydromechanical transmission (i-HMT). This introduces an additional control degree-of-freedom for adjusting the motor's valve timing to specifically control pre-compression at different operating conditions so as to avoid throttling losses due to unequal pressure when the valve opens. The approach has been reduced to design, prototyped and experimentally tested by modifying the valve cam of an stock Hondamatic i-HMT. Experimental results show that an optimized valve cam for oil with 5% entrained air does improve efficiency over the stock cam by as much as 7% over a broad range of operating conditions. Although this paper focuses on improving the efficiency of an i-HMT, it should be applicable to other axial piston machines as well.

References

[1] B. Carl, M. Ivantysynova, and K. Williams, "Comparison of operational characteristics in power split con-

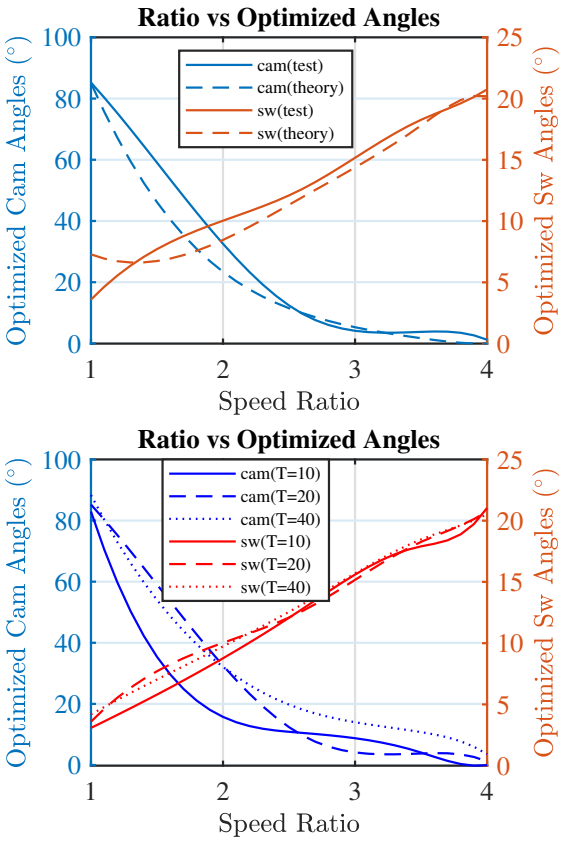


Figure 11: Optimal cam and swashplate angles for the 5% cam and 2000 RPM. Top) Analytical and optimized angles for $T_{out} = 20Nm$. Bottom) Optimized angles $T_{out} = 10Nm, 20Nm, 40Nm$.

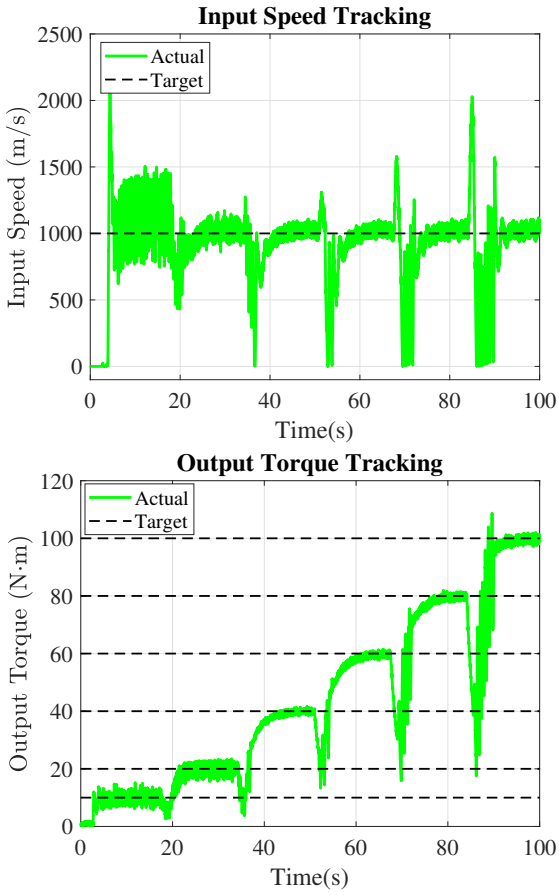


Figure 12: Sample input speed and output torque tracking results.

tinuously variable transmissions," *SAE Technical Paper Series - 2006-01-3468*, 2006.

[2] X. Hu, C. Jing, and P. Y. Li, "Modeling and experimental testing of the hondamatic inline hydro-mechanical transmission (iHMT)," *ASME J. of Dynamic Systems, Measurement, and Control*, vol. 142, no. 5, 2020.

[3] K. Nagatomo, "On hydro-mechanical hybrid transmission," in *Proceedings of the JFPS International Symposium on Fluid Power*, vol. 1993, pp. 273–278, The Japan Fluid Power System Society, 1993.

[4] J. Barkei and P. Y. Li, "Improving the efficiency of a compact inline hydro-mechanical transmission (iHMT) at lock-up," *Proc. of ASME/BATH 2021 Sym. on Fluid Power and Motion Control*, October 2021.

[5] E. D. Sand and P. Y. Li, "Incorporating a rotatable valve cam to improve the efficiency of a hydraulic motor in an inline hydro-mechanical transmission (i-HMT)," in *Proc. of Bath/ASME Sym. on Fluid Power and Motion Control*, vol. 86335, p. V001T01A029, 2022.

[6] L. Ericson, S. Kärnell, and M. Hochwallner, "Experimental investigation of a displacement-controlled hydrostatic pump/motorby means of rotating valve plate," in *The 15th Scandinavian Int. Conf. on Fluid Power, SICFP 17, Linköping, Sweden*, vol. 144, pp. 19–27, 2017.

[7] G. E. Vael, P. A. Achten, and Z. Fu, "The INNAS hydraulic transformer the key to the hydrostatic common pressure rail," in *SAE Tech. Paper*, 2000.

[8] M. Ehsan, W. Rampen, and S. Salter, "Modeling of digital-displacement pump-motors and their application as hydraulic drives for nonuniform loads," *J. Dyn. Sys., Meas., Control*, vol. 122, no. 1, pp. 210–215, 2000.

[9] M. B. Rannow, P. Y. Li, and T. R. Chase, "Discrete piston pump/motor using a mechanical rotary valve control mechanism," in *Proc. of the 8th workshop on digital fluid power, Tampere, Finland*, pp. 83–94, 2016.

[10] J. Barkei and P. Y. Li, "Hydraulic effort and the efficiencies of pump and motors with compressible fluid," *Proceedings of ASME/BATH 2020 Symposium on Fluid Power and Motion Control*, September 2020.

[11] J. Bastien, "Periodical cubic interpolation." `https://www.mathworks.com/matlabcentral/fileexchange/46580-periodical-cubic-interpolation`, May 2014.

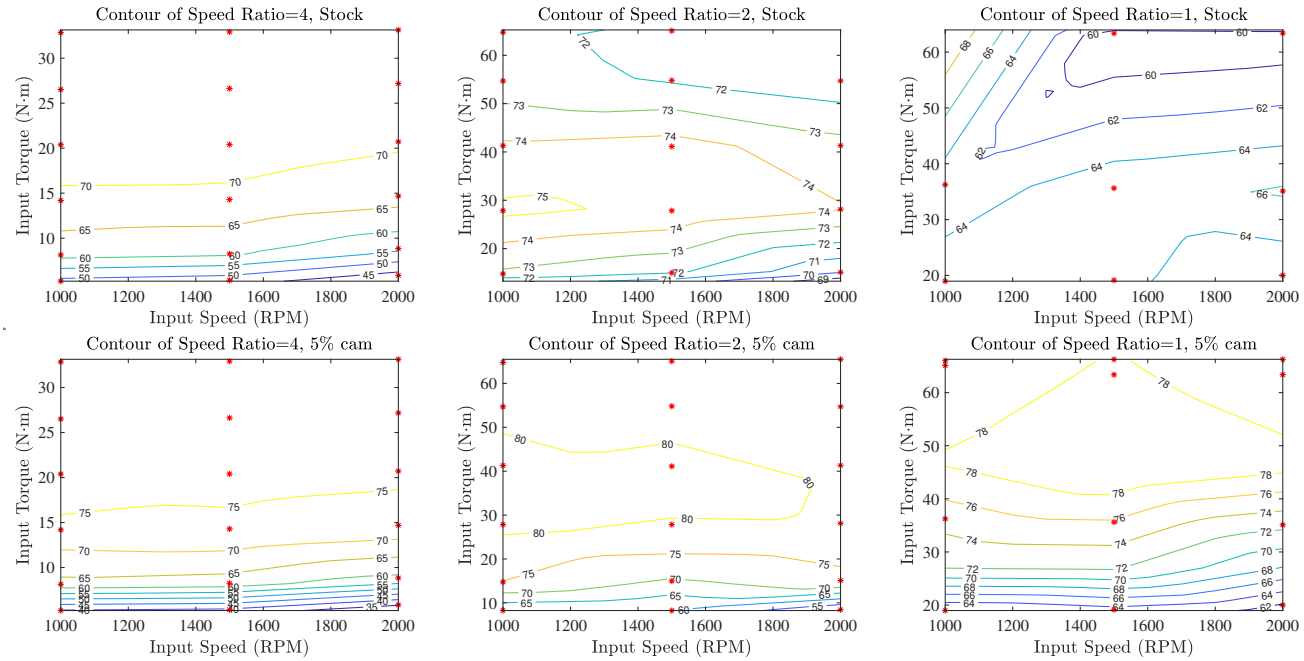


Figure 13: Efficiency of stock and 5% cam at ratio=4, 2, and 1

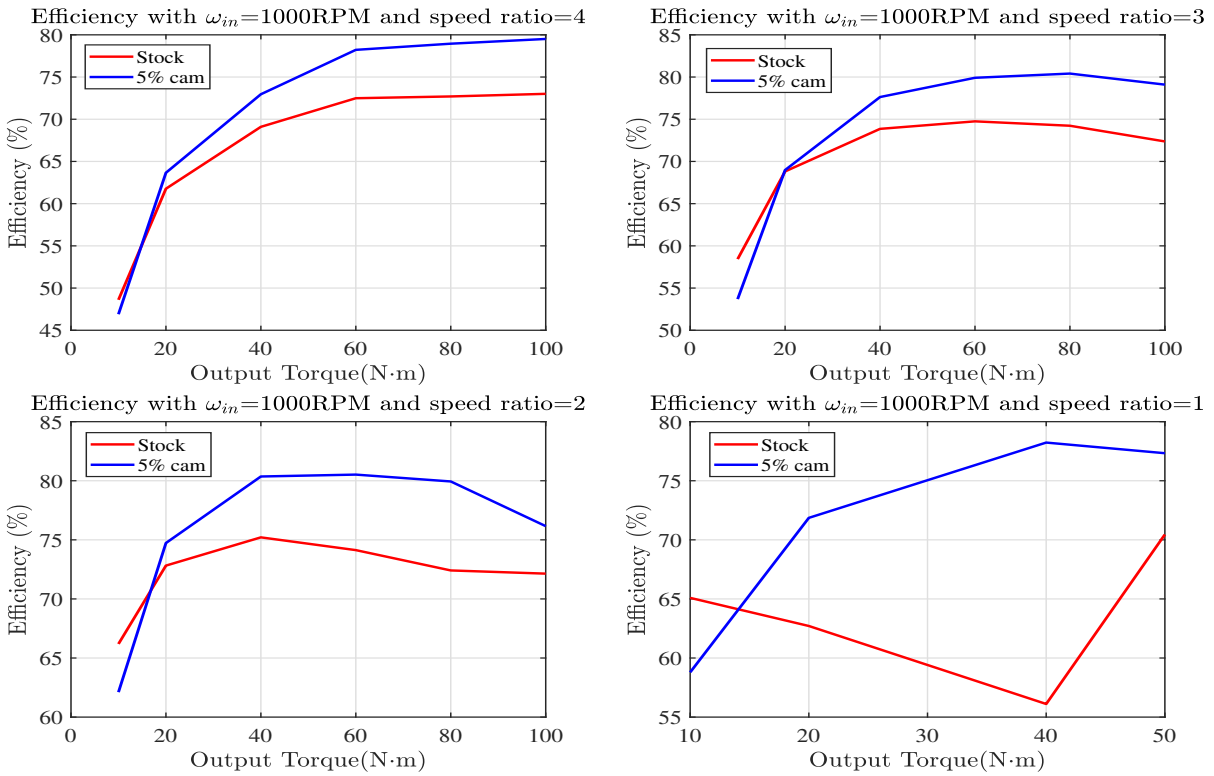


Figure 14: Efficiency of stock and 5% cams versus output torque at ratio = 4,3,2,1 and input speed = 1000 RPM.

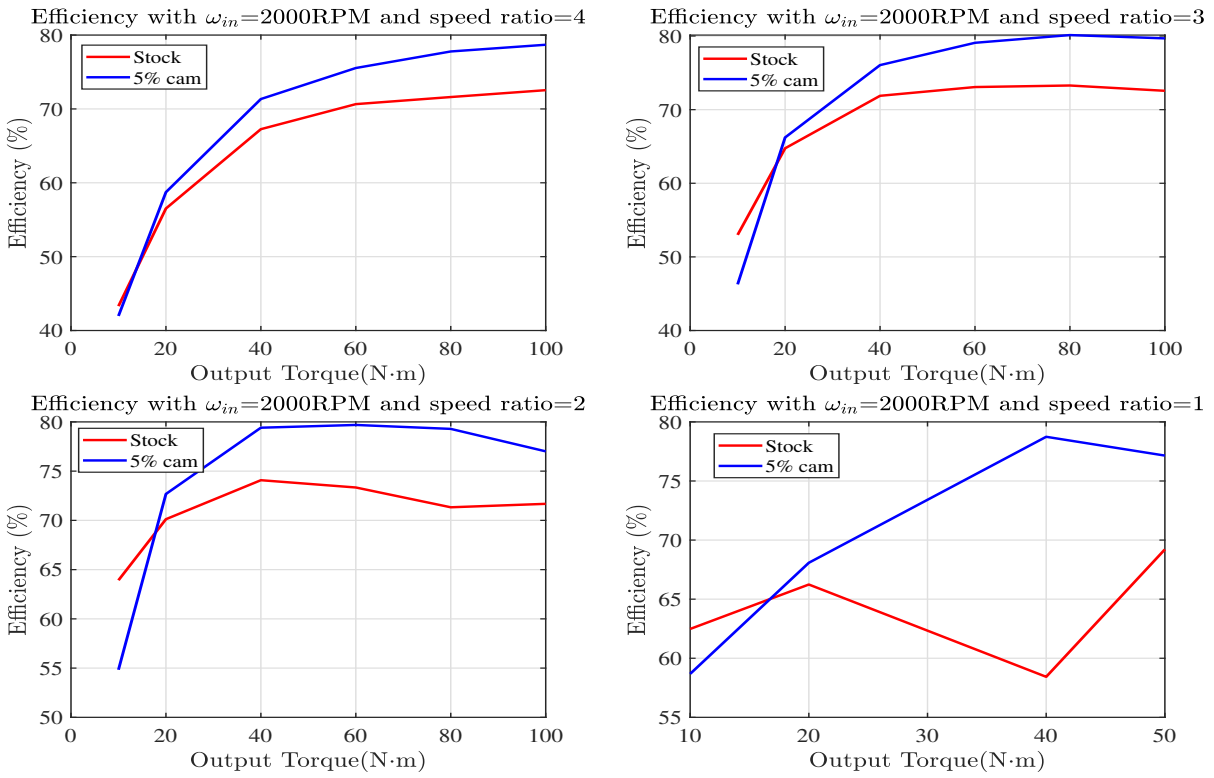


Figure 15: Efficiency of stock and 5% cams versus output torque at ratio = 4,3,2,1 and input speed = 2000 RPM.

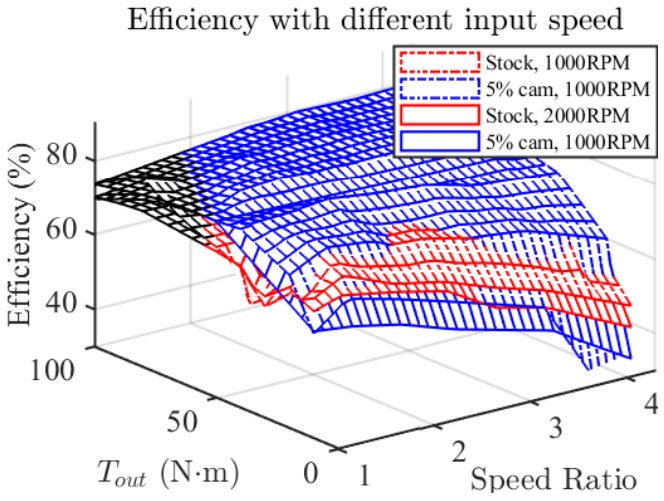


Figure 16: Efficiencies at different speeds.

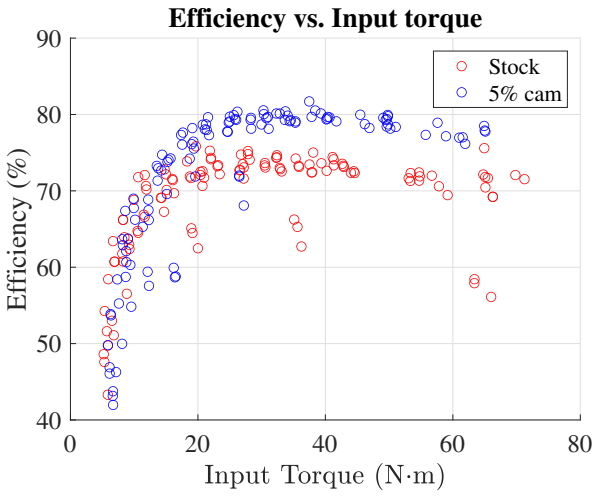


Figure 17: Efficiencies vs input torques (all cases).

Extreme Variable-Speed Lubrication Dynamics of Cylinder Block/Valve Plate Interface Based on Neural Network Method

Dingchong LYU[1], Jian FU[1,*], Chao ZHAO[2], Baoyan HU[1], Siyuan CHEN[1], Shoujun ZHAO[3]

[1]School of Mechanical Engineering & Automation, Beihang University, Beijing, China
[2]School of Naval Architecture and Ocean Engineering, Dalian Maritime University, Dalian, Shandong, China
[3]Beijing Institute of Precision Mechatronics and Controls, Beijing, China
E-mail: fujian@buaa.edu.cn

Abstract

High-integrated pump-controlled servo actuators are revolutionizing aerospace hydraulics with superior power-to-weight ratios and dynamic responses. Yet, hydraulic pumps endure extreme operational stresses under high-frequency aerodynamic loads, particularly on the cylinder block and valve plate interface. This critical friction pair governs reliability, efficiency, and lifespan through flow distribution and lubrication. Traditional modeling struggles with multi-physics coupling (turbulent microflows, mixed friction, micro-asperity contact), requiring computationally prohibitive simulations. This study pioneers a physics-informed neural network (PINN) framework to decode dynamic lubrication characteristics under extreme conditions. A high-fidelity numerical model integrates Reynolds-based compressible fluid dynamics and contact mechanics. Physics-constrained PINN training leverages numerical datasets from representative power spectra, with architecture optimized for fluid-solid coupling laws. Experimental validation demonstrates the model's precision in capturing mixed lubrication regimes and transient flow fields under variable speeds/pressures. The PINN achieves way faster computation than numerical methods with < 5% prediction error, enabling real-time analysis. By synergizing first-principles physics with neural operators, this work establishes an intelligent hybrid modeling paradigm for aerospace hydraulic systems, offering transformative solutions for next-generation pump design, condition monitoring, and performance optimization under extreme operational spectra.

Keywords: Electro-hydrostatic actuator, hydrodynamics, numerical simulation, physics-informed neural networks

1 Introduction

Compared to traditional valve-controlled hydraulic systems, highly integrated pump-controlled servo actuators, with their superior power-to-weight ratios and dynamic response characteristics, are increasingly being adopted in aerospace applications [1,2]. In typical aerospace systems such as aircraft flight control surfaces and vector engine servo mechanisms, aerodynamic surfaces are subjected to high-frequency alternating loads. This subjects aerospace servo motor pumps to complex operating conditions involving repeated start-stop cycles, directional reversals, and variable speeds. The high-speed, high-pressure, and dynamically varying operating environments impose significant challenges on internal hydraulic pump components, particularly friction pairs [3–5]. Among these, the cylinder block/port plate pair serves as a critical interface determining the reliability, lifespan, efficiency, and dynamic performance of axial piston pumps. This interface fulfills multiple functions, including flow distribution, load support, lubrication, and sealing. Investigating the operational mechanisms of the cylinder block/port plate pair under high-frequency alternating conditions is essential for enhancing interface performance and represents a vital step toward optimizing the design of high-performance aerospace servo motor pumps [6–8]. The working characteristics of this interface involve complex interactions such as hydrodynamic effects of hydraulic fluid, mixed/boundary friction, micro-elastic deformation of contact surfaces, and thermal deformation. High-fidelity thermo-elasto-hydrodynamic (TEHD) coupling modeling of these phenomena is highly intricate and computationally demanding [9].

Traditional numerical methods rely heavily on mesh quality for accuracy and efficiency. For complex problems, the time required to generate high-quality computational grids often constitutes a substantial portion of the total solution time, significantly increasing computational costs. Artificial neural networks (ANNs) exhibit strong capabilities in approximating arbitrary functions, with deeper networks demonstrating enhanced approximation capacities. Consequently, solving

partial differential equations (PDEs) using deep neural networks has emerged as a promising frontier in machine learning, offering an alternative to conventional numerical discretization methods [10]. The application of traditional neural networks in hydraulic system modeling and fault diagnosis dates back to the 1990s, with their core advantage lying in data-driven nonlinear mapping for approximating complex input-output relationships. Karkoub et al. [11] pioneered the use of neural networks in 1999 for predicting the power consumption characteristics of axial piston pumps, marking a breakthrough for data-driven approaches in hydraulics. Subsequent studies expanded their applications: in fault diagnosis, multilayer perceptrons achieved classification of typical failures such as cylinder leakage and port plate wear using vibration spectra and pressure signals [12,13], in performance prediction, Hess et al. [14] employed neural networks to forecast elasto-hydrodynamic lubrication pressures in sliding bearings, while Peric et al. [15] modeled dynamic behaviors of external gear pumps. However, purely data-driven methods face notable limitations: first, their generalization capability heavily depends on the coverage of training data, rendering them inadequate for handling diverse hydraulic operating conditions. Second, the lack of explicit representation of underlying physical mechanisms (e.g., fluid dynamics and contact mechanics) leads to severe reliability degradation in untrained scenarios [14]. Guo et al. [12] emphasized that traditional neural networks for drilling pump fault diagnosis require thousands of labeled datasets, incurring prohibitive experimental costs in practice.

To address these limitations, physics-informed neural networks (PINNs) integrate governing equations into network architectures, enabling synergistic fusion of physical principles and data features. The core innovation lies in embedding physical conservation laws, such as the Navier-Stokes equations and Reynolds lubrication equations [16], as residual constraints within loss functions to form a "physics-regularized" mechanism. Eivazi et al. [17] encoded RANS equations into PINNs, significantly improving turbulent flow prediction efficiency. Ramos et al. [18] further demonstrated their robustness against numerical oscillations in solving microscale Reynolds equations. In hydraulic applications, this hybrid strategy demonstrates unique advantages: Wang et al. [19] and Dong et al. [20] incorporated axial piston pump flow ripple equations as physical constraints, achieving fault identification accuracy at industrially practical levels. Liu et al. [21] developed hierarchical PINNs that synchronously evaluated seal degradation and remaining lifespan by integrating rotor dynamics with sensor data. Yin et al. [22] established a PINNs framework coupling Reynolds equations and contact mechanics to resolve cross-scale challenges in friction pair characterization under seawater lubrication. Ma et al. [23] combined PINNs with proper orthogonal decomposition, enhancing computational efficiency of high-dimensional electro-hydrostatic actuator (EHA) models by two orders of magnitude. Exploring the application of PINNs to solve complex dynamic characteristics of servo motor pump cylinder block/port plate pairs under representative power spectra constitutes a novel and compelling research direction.

This study adopts PINNs to predict dynamic lubrication characteristics of cylinder block/valve plate interfaces under high-frequency conditions. The model integrates the Reynolds equation and contact mechanics into PINNs and is validated with high-fidelity data. It accurately predicts oil film pressure and friction losses, supporting intelligent design of aerospace hydraulic pumps.

2 Preliminaries

The framework integrates hydrodynamic theory, mixed friction characterization under variable speeds, and PINNs embedding conservation laws and experimental data to predict lubricating behavior on the cylinder block/valve plate interface.

2.1 Numerical characterization of dynamic lubrication characteristics

The interface governs lubrication and friction in axial piston pumps via hydrodynamic effects (modeled by Reynolds equations) and solid-contact interactions. The cylinder block's tilt generates a wedge-shaped oil film, producing hydrodynamic pressure to reduce friction. However, transient mixed lubrication arises under variable loads/speeds, requiring simultaneous characterization of fluid-film dynamics (pressure distribution, viscosity effects) and asperity contact forces (boundary friction, wear mechanisms). Geometric parameters like tilt angle and film thickness critically influence regime transitions between hydrodynamic and boundary lubrication.

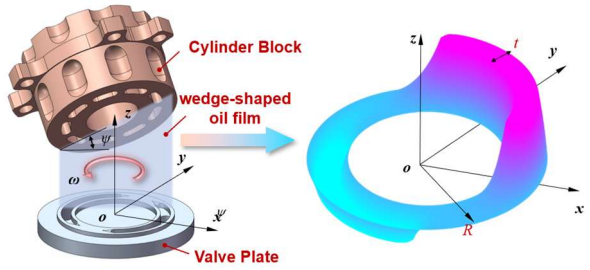


Figure1: Schematic diagram of cylinder block and valve plate pair.

To balance accuracy and computational efficiency, key assumptions streamline the Reynolds equation: (1) Neglect gravity, electromagnetic forces, and fluid inertia, (2) Assume constant viscosity, no wall slip, and isothermal conditions, (3) Approximate circumferential flow as translational motion. These simplifications focus on the dominant factors, such as film thickness dynamics and pressure gradients, while maintaining the predictive capability for mixed lubrication behavior. The reduced model enables efficient analysis of transient friction and lubrication characteristics under variable-speed conditions.

2.2 Transition of mixed lubrication states under variable-speed conditions

In aerospace applications, the inability to sustain hydrodynamic lubrication under rapid speed transitions exacerbates surface interactions, rendering complete fluid lubrication models inadequate. The friction behavior of distribution pairs in high-integrated servo actuators cannot be

characterized solely by low-speed regimes or static coefficients. Instead, a dynamic adaptation of the Stribeck framework is essential to map friction transitions across ultra-wide speed spectra (-20,000~20,000 rpm) and capture transient mixed lubrication effects. This adaptation enables precise modeling of friction losses and wear mechanisms in pump-controlled actuators, addressing the unique demands of aerospace systems where reliability and precision are mission-critical.

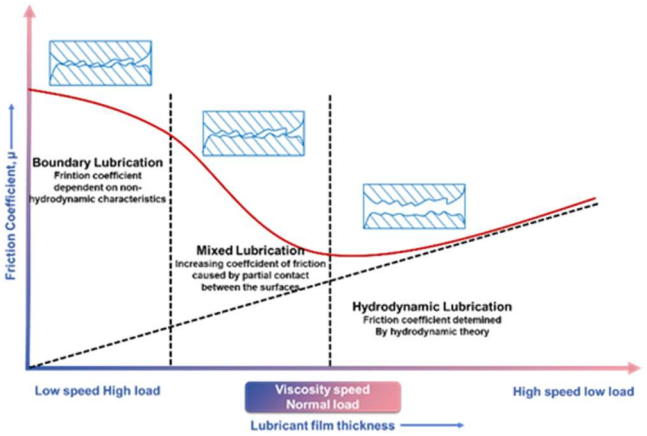


Figure2: Stribeck curve diagram.

2.3 Physics informed neural network prediction

Unlike purely data-driven methods, PINNs synthesize measurement signals and physical residuals through dual loss functions. This capability enables reliable extrapolation in aerospace applications, such as simulating hydraulic pump component behaviors under extreme variable-speed conditions, where experimental data acquisition is costly or impractical. Their interpretable outputs, such as transient pressure fields in dynamic lubrication, bridge theoretical insights with real-world performance validation while maintaining computational efficiency.

Numerical challenges persist in modeling cylinder block/valve plate pairs due to fluid-solid coupling under transient lubrication. Traditional methods (e.g., Runge-Kutta, FEA) suffer from high computational costs and poor convergence under wide-speed ranges, exacerbated by nonlinear hydrodynamic-contact interactions during rapid speed transitions.

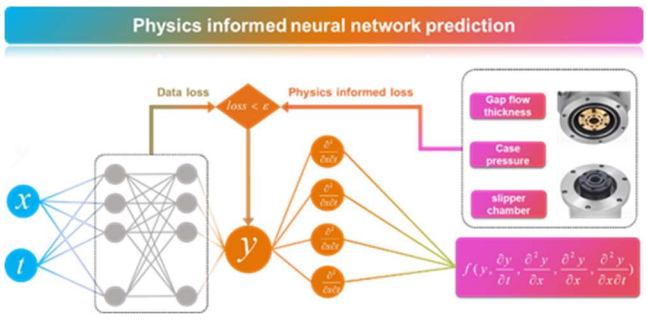


Figure3: Schematic diagram of PINN model.

Implementation trains PINNs using theoretical numerical solutions as ground truth. Inputs include spatiotemporal coordinates (e.g., φ, λ) and operational parameters (e.g., ω, ψ), while outputs predict pressure fields. The loss function integrates PDE residuals, periodicity, and boundary constraints, enabling millisecond-level inference for real-time aerospace mixed lubrication analysis post-training.

3 Neural network design

The construction of PINNs is fundamentally based on governing physical equations. By embedding physical laws as intrinsic constraints, PINNs can significantly reduce the costs associated with data acquisition and storage, demonstrating distinct advantages in addressing complex physical systems.

A Multilayer Perceptron (MLP) with hierarchically structured neurons is employed to execute mathematical operations. Loss functions are formulated based on the residuals derived from the mixed lubrication model of the cylinder block/valve plate interface. Through iterative training, a fully differentiable surrogate model is obtained, whose outputs are differentiable with respect to the inputs. This capability allows for seamless integration with traditional numerical methods while preserving the inherent benefits of neural networks. Fig. 4 [18] illustrates the overall structure of the hybrid lubrication characteristics prediction framework for the cylinder block/valve plate interface based on PINNs.

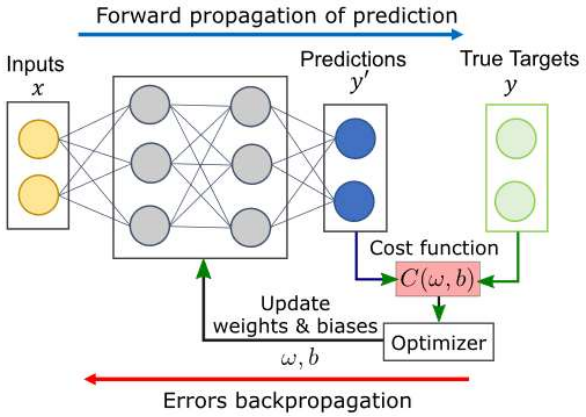


Figure4: Overall structure of the prediction framework. [18]

3.1 Governing equations related to fluid dynamics

In the realm of theoretical research, the characteristics of complete fluid-film lubrication and complete contact friction have been extensively explored and are relatively well-established. However, under the unique aerospace operating conditions pertinent to this study, the characterization equations for the lubrication and friction properties of the friction pair interface remain a subject of ongoing discussion.

3.1.1 Hydrodynamic lubrication characterization

The formulation of dimensionless equations plays a crucial role in streamlining the representation of governing fluid dynamics equations. It not only standardizes the relationships between variables but also minimizes the influence of unit scaling. By reducing variables to a common order of magnitude, the dimensionless form ensures that numerical values remain within a computationally feasible range, avoiding extremes that could compromise simulation accuracy.

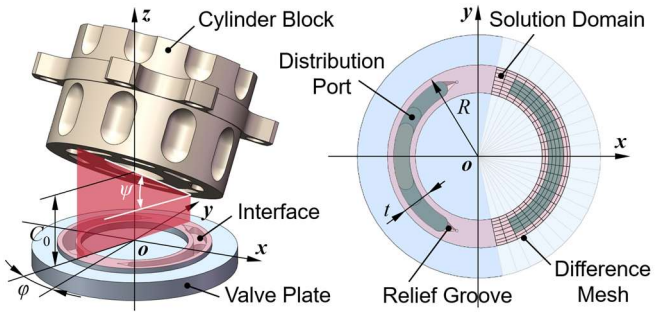


Figure5: Geometry structure of the interface. [8]

In this study, the dimensionless Reynolds equation is employed to characterize the lubrication behavior at the cylinder block/valve plate friction pair, a critical interface in the performance of highly integrated pump-controlled servo actuators. This equation is expressed as follows:

$$\frac{\partial}{\partial\varphi}\left(PH^3\frac{\partial P}{\partial\varphi}\right)+\frac{\partial}{\partial\lambda}\left(PH^3\frac{\partial P}{\partial\lambda}\right)=\Lambda\frac{\partial(PH)}{\partial\varphi}+2\Lambda\frac{\partial(PH)}{\partial T} \tag{1}$$

$$\Lambda=\frac{6\mu\omega}{P_a}\left(\frac{R}{C_0}\right)^2 \tag{2}$$

Where φ and λ denote the dimensionless circumferential and radial coordinates, respectively. T is the dimensionless operating time. P and H represent the dimensionless oil pressure and oil film thickness, respectively. Λ denotes the compressibility number. The tilting angle ψ of the cylinder block and the initial oil film thickness C_0 are incorporated into the characterization of H for the bidirectional coupling solution with rotor dynamics. The force balance analysis of the cylinder block assembly is described in [8].

$$H=\frac{C_0+R\sin\psi\cos(\varphi+\varphi_0)}{R} \tag{3}$$

C_0 is primarily influenced by the oil pressure P, the spring preload force F_{sp}, and the external force exerted by the cylinder block assembly.

3.1.2 Contact friction characterization

The frictional stress in the contact region of the friction pair can be characterized using Hertzian elastic contact theory. Although the contact surface between the cylinder block and valve plate is planar, due to the tilting behavior of the cylinder block, the contact region can be approximated as the interaction between a flat surface and a point (spherical) contact. When a sphere and a flat surface are in contact, the Hertzian stress is expressed as follows:

$$\sigma_{\max}=\sqrt[3]{\frac{6}{\pi^3}\frac{1}{R^2}\left(\frac{N}{\frac{1-\mu_1^2}{E_1}+\frac{1-\mu_2^2}{E_2}}\right)} \tag{4}$$

$$R=\frac{\left(\dot{x}^2+\dot{y}^2\right)^{\frac{3}{2}}}{\dot{x}\ddot{y}-\dot{y}\ddot{x}} \tag{5}$$

Where $\sigma_{\max}$ represents the maximum contact stress, a denotes the contact width, N is the external load, R is the radii of the contacting cylinder, E_1 and E_2 are the Young's moduli, and μ_1 and μ_2 are the Poisson's ratios of the two contacting cylinders.

3.1.3 Integration of contact/non-contact region

In the mixed/boundary friction state, viscous friction dominates in the non-contact region, while coulomb friction prevails in the contact region. Both frictional forces are influenced by the pressure P. The relationship between the two types of frictional torques and the clamping force P is expressed by eq. (6) and (7).

The distribution ratio coefficient ξ between the two types of friction was defined as a specific function to fit the experimental curves in our previous study, as shown in fig. 7.

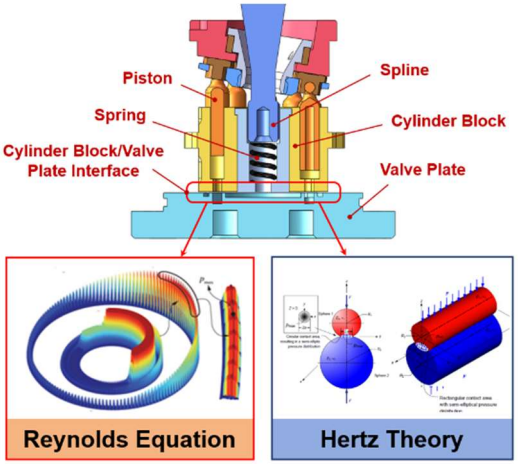


Figure6: Fluid dynamics modeling architecture.

$$T_m=T_v(P)+T_c(P) \tag{6}$$
$$P=P_v+P_c=(1+\xi)P_v$$

$$T_v=P_aR^2C_0\int_{\frac{R-t}{R}}^{1}\int_0^{2\pi}\left(\frac{\Lambda}{6}\frac{1}{H}-\frac{H}{2}\frac{\partial P_v}{\partial\varphi}\right)\mathrm{d}\varphi\mathrm{d}\lambda \tag{7}$$

T_m, T_v, and T_c denote the torques corresponding to mixed friction, viscous friction, and Coulomb friction, respectively. P_v and P_c represent the pressures assigned to viscous and Coulomb friction, respectively.

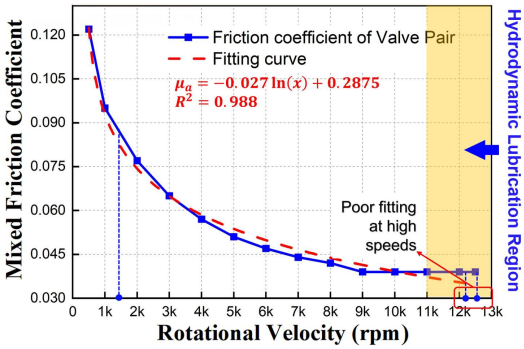


Figure7: Experimental determination of the distribution ratio coefficient ξ.

Table1: *Comparison of circumferential periodicity condition treatments.*

Model Type	Circumferential Periodicity Solution Strategy	Approaches	Key Expressions
Numerical model	Modify the configurations of second-order difference coefficient matrices $D_{\varphi\varphi}$ and $D_{\lambda\lambda}$, incorporating the node values (u_1, u_{X-1}) into each other's difference stencils.	Constructing lost functions	$\left.\frac{\partial^2 u}{\partial \varphi^2}\right\|_{i=1,j} = \frac{u_{X-1,j} - 2u_{1,j} + u_{2,j}}{h_\varphi^{\,2}}$
PINN model	Ensure the initial value $p(\varphi = 0, \lambda)$ and terminal value $p(\varphi = 2\pi, \lambda)$ of the pressure distribution remain equal in the circumferential coordinate φ.	Modifying difference matrix	$L_{\text{cir}} = \frac{1}{N_{\text{cir}}} \sum_{i=1}^{N_{\text{cir}}} \left(u_{\text{n}}(\varphi_1) - u_{\text{n}}(\varphi_{\text{X}})\right)^2$

3.2 Boundary conditions and circularity condition

The equation used to characterize the mixed lubrication characteristics of the cylinder block/valve plate interface serves as the fundamental theoretical basis for PINNs. In addition to satisfying this equation, it is also necessary to consider the geometric structure of the frictional pair's working interface, pressure boundary conditions, and the distribution of contact and non-contact regions.

3.2.1 Circumferential periodicity condition

The working region is an annular zone containing two kidney-shaped grooves. A dimensionless numerical solution method truncates this annular region and approximates it as a rectangular solution domain to facilitate solving via difference matrix operations [8]. Consequently, the neural network model must first satisfy the circumferential periodicity condition of the solution domain. Specifically, in the circumferential coordinate φ direction, the initial value $p(\varphi = 0, \lambda)$ and terminal value $p(\varphi = 2\pi, \lambda)$ of the pressure distribution must remain equal at all times [18], as shown in eq. (8):

$$p(\varphi = 0, \lambda) = p(\varphi = 2\pi, \lambda) \tag{8}$$

It is worth noting that the issue of circumferential periodicity conditions also arises during numerical solution procedures. However, circumferential periodicity can be directly enforced by modifying the configuration of difference matrices without explicitly setting boundary conditions. The fundamental form of the elliptic partial differential equation can be expressed as follows:

$$-c_{ij}\left(\frac{u_{i,j-1} - 2u_{ij} + u_{i,j+1}}{h_\varphi^2} + \frac{u_{i-1,j} - 2u_{ij} + u_{i+1,j}}{h_\lambda^2}\right) + a_{ij}u_{ij} = f_{ij} \tag{9}$$

$$\left[-Diag_c\left(D_{\varphi\varphi} + D_{\lambda\lambda}\right) + Diag_a\right]u = Diag_f \tag{10}$$

Where c, a and f are coefficient values of the elliptic partial differential equation, distributed in main diagonal matrix forms denoted as $Diag_c$, $Diag_a$ and $Diag_f$ are the column vectors of the unknown function values, while h_φ and h_λ represent the unit lengths in the circumferential and radial directions, respectively. The product vectors of matrices $D_{\varphi\varphi}$ and $D_{\lambda\lambda}$ with the column vector u correspond to the second-order partial derivative column vectors of u in the circumferential and radial directions.

Since the column vector u is arranged according to the row-wise node distribution of the rectangular solution domain, the matrix configurations of $D_{\varphi\varphi}$ and $D_{\lambda\lambda}$ can be directly formulated without considering circumferential periodicity. Taking $D_{\varphi\varphi}$ as an example, as shown in eq. (11):

$$D_{\varphi\varphi} = h_\varphi^{-2}\begin{bmatrix} -2 & 1 & & \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & -2 \end{bmatrix}, N \times N \tag{11}$$

Where N is the total number of nodes in the solution domain. When incorporating circumferential periodicity, the function values of the first node in each row and the second-to-last node in the same row are mutually assigned as difference elements, expressed as:

$$\left.\frac{\partial^2 u}{\partial \varphi^2}\right|_{i=1,j} = \frac{u_{X-1,j} - 2u_{1,j} + u_{2,j}}{h_\varphi^{\,2}} \tag{12}$$

$$\left.\frac{\partial^2 u}{\partial \varphi^2}\right|_{X-1,j} = \frac{u_{X-2,j} - 2u_{X-1,j} + u_{1,j}}{h_\varphi^{\,2}} \tag{13}$$

Where X is the number of row nodes and Y is the number of column nodes. Based on the above expressions, the form of $D_{\varphi\varphi}$ can be transformed into the structure shown in eq. (14), where the left side is a $Y \times Y$ identity matrix and the right side is an $(X\text{-}1) \times (X\text{-}1)$ matrix. The same logic applies to $D_{\lambda\lambda}$. This matrix transformation ensures circumferential periodicity mathematically, effectively converting the rectangular solution domain back to a ring-shaped region physically. A comparison between the numerical method and neural network algorithm for handling circumferential periodicity conditions is summarized in Table 1.

$$D_{\varphi\varphi} = h_\varphi^{-2}\begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} \otimes \begin{bmatrix} -2 & 1 & & 1 \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ 1 & & 1 & -2 \end{bmatrix} \tag{14}$$

3.2.2 Pressure boundary conditions

The setting of pressure boundary conditions. In the two kidney-shaped groove regions, one side corresponds to the high-pressure zone (suction port), where the pressure remains consistent with the load-end pressure p_{out}, while the other side corresponds to the low-pressure zone (discharge port), where the pressure aligns with the system pressure p_{in}. Additionally, the inner and outer boundaries of the annular solution domain maintain consistency with the pressure within the hydraulic pump chamber p_{c}, as shown below:

$$p(\varphi, \lambda \in R_{\text{in}}) = p_{\text{in}} \tag{15}$$

$$p(\varphi, \lambda \in R_{\text{out}}) = p_{\text{out}} \tag{16}$$

$$p(\varphi, \lambda = r_1) = p(\varphi, \lambda = r_2) = p_{\text{c}} \tag{17}$$

Where R_{in} and R_{out} are subsets defining the ranges of the suction and discharge port kidney-shaped groove regions, respectively, and r_1 and r_2 represent the inner and outer radii of the solution domain.

3.2.3 Boundary conditions for contact/non-contact region

Under wide-range variable-speed and variable-pressure operating conditions, the cylinder block/valve plate interface struggles to form a stable oil film for support, resulting in a complex lubrication/friction state where contact and non-contact zones coexist. Neither a standalone hydrodynamic model nor a contact mechanics model is sufficient to describe such intricate lubrication characteristics.

In our prior research, to couple hydrodynamic and contact mechanics models for characterizing mixed/boundary lubrication states, we utilized friction test data from the cylinder block/valve plate interface to generate mixed friction coefficient-rotational speed curves. The proportion of viscous friction to contact friction was determined based on friction coefficients obtained under specific test conditions [8]. Consequently, mixed friction/lubrication behavior was primarily characterized by a proportionality factor ξ, as shown in eq. (18). However, this proportionality factor holds purely mathematical significance—it neither physically explains the mechanism of mixed friction/lubrication nor exhibits general applicability.

$$F_t = F_v(p) + F_c(p) = (1+\xi)F_v(p) \tag{18}$$

Where F_t, F_v, and F_c represent the residual clamping force, hydrodynamic support force, and contact support force, respectively.

To address these limitations, we revisited the coupling methodology of the two friction models. By observing wear patterns on the friction pair specimen surfaces under specific speed/pressure conditions, we determined the distribution of contact/non-contact zones. This distribution was then used to define the solution domains for the two friction models, and mixed friction coefficients were solved based on mechanical equilibrium equations, as illustrated in fig. 8. Detailed implementation steps are provided in Chapter 4. The governing equations are as follows:

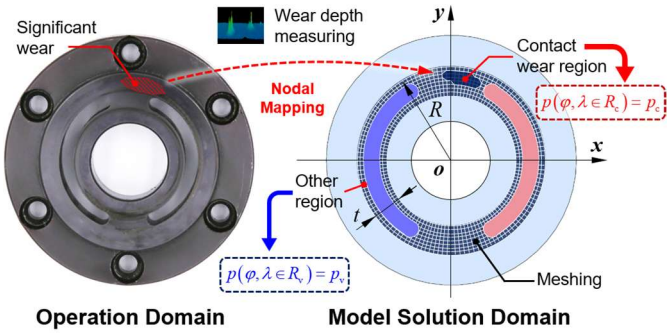


Figure8: Characterizations of contact/non-contact region.

$$p(\varphi,\lambda\in R_v) = p_v \tag{19}$$

$$p(\varphi,\lambda\in R_c) = p_c \tag{20}$$

$$F_t = F_v(p) + F_c(p) = \iint_{\varphi,\lambda\in R_v} p_v \mathrm{d}\varphi\mathrm{d}\lambda + \iint_{\varphi,\lambda\in R_c} p_c \mathrm{d}\varphi\mathrm{d}\lambda \tag{21}$$

Where R_v and R_c are subsets defining the experimentally observed non-contact and contact zones, respectively, and p_v and p_c denote the pressure distributions calculated by hydrodynamic and contact mechanics theoretical models.

3.3 Network loss function

Based on the mixed lubrication model equations for the cylinder block/valve plate pair interface and the boundary/circumferential conditions outlined in Section 3.2, the loss function for PINNs can be constructed as the sum of residuals from the governing differential equations, circumferential periodicity errors, pressure boundary condition errors, and contact zone distribution errors, as shown in eq. (22)~(26). By minimizing the mean squared error (MSE), measured signals are set as learning objectives for training the neural network.

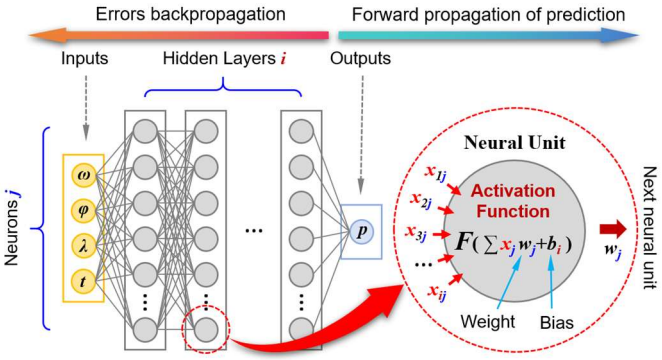


Figure9: Neural unit of the PINN model for prediction.

$$L(w,b) = L_t + L_{cir} + L_b + L_m \tag{22}$$

$$L_t = \frac{1}{N_t}\sum_{i=1}^{N_t} f_t^2 \tag{23}$$

$$L_{cir} = \frac{1}{N_{cir}}\sum_{i=1}^{N_{cir}} \left(u_n(\varphi_1) - u_n(\varphi_X)\right)^2 \tag{24}$$

$$L_b = \frac{1}{N_t}\sum_{i=1}^{N_t} \left(u_n - u_t\right)^2 \tag{25}$$

$$L_m = \frac{1}{N_m}\sum_{i=1}^{N_m} \left(u_n - u_t\right)^2 \tag{26}$$

Where w and b represent the weights and biases, respectively. L_t denotes the residual of the theoretical physical equations, L_{cir} the circumferential periodicity error, L_b the pressure boundary condition error, and L_m the contact zone distribution error. f_t corresponds to the residual of the theoretical differential equations, u_n is the predicted value from the neural network, and u_t is the numerically solved value from the theoretical equations.

The architecture of PINNs is designed such that the number of neurons in the first layer matches the input variables, which include the circumferential coordinate φ, radial coordinate λ, rotational velocity ω, and operation time t. The final layer has neurons equal to the output variables: the pressure value p of

the relevant domain node. The framework and training process of PINNs for solving the theoretical model equations under dynamic conditions are illustrated in fig. 10. The pressure distribution prediction can be generated by arranging the predicted nodal pressures into a matrix.

3.4 Theoretical evaluation of predictive capability

To verify the success of the network training and ensure that the results obtained by PINNs under this convergence criterion align with real physical scenarios, it is necessary to compare them with the results from the theoretical numerical model. The theoretical model of mixed lubrication is solved using an iterative method based on the configuration of difference-elliptic partial differential equations. The specific solution procedure can be referred to in Reference [8] and will not be reiterated here. The parameter settings for the PINN model in this study are listed in Table 2.

Table2: *Parameter settings for PINN model.*

Parameters	Value
No. of hidden layers	3/4/5
No. of neurons per layer	32/48/64
Boundary collocation points	1,000
Interior collocation points	8,200
Numerical solution data points	8,200
Training epochs	2,000

4 Experimental verification

This chapter systematically validates the hybrid lubrication model and PINNs framework through experimental studies. By establishing a controlled test platform and standardizing operating conditions, friction data under variable speeds and pressures are collected to train and test the neural network.

4.1 Experimental setup and preparations

The experimental setup employed in this study is a high-speed and high-pressure piston pump/tribological pair test platform [6,7], as illustrated in fig. 10. The test apparatus primarily consists of a loading cylinder, piston, and anti-rotation pin. The loading cylinder is equipped with four oil ports: (1) The loading port is connected to the upper chamber of the loading cavity to simulate the hydraulic clamping force of the port plate pair. (2) The draining port is linked to the lower chamber to replicate the casing pressure in a piston pump. (3) The inlet and outlet ports are connected to the kidney-shaped ports on the port plate friction pair, generating hydraulic separation forces. Prediction accuracy is rigorously evaluated by comparing model outputs against both numerical simulations and direct experimental measurements, ensuring the proposed method captures real-world mixed.

4.2 Wear test calibration

Wear tests were conducted on hard-soft and hard-hard material pairs. The valve plate surfaces used identical hard materials (DLC coating), while cylinder block specimens were made of HMn manganese brass or CrN-coated variants.

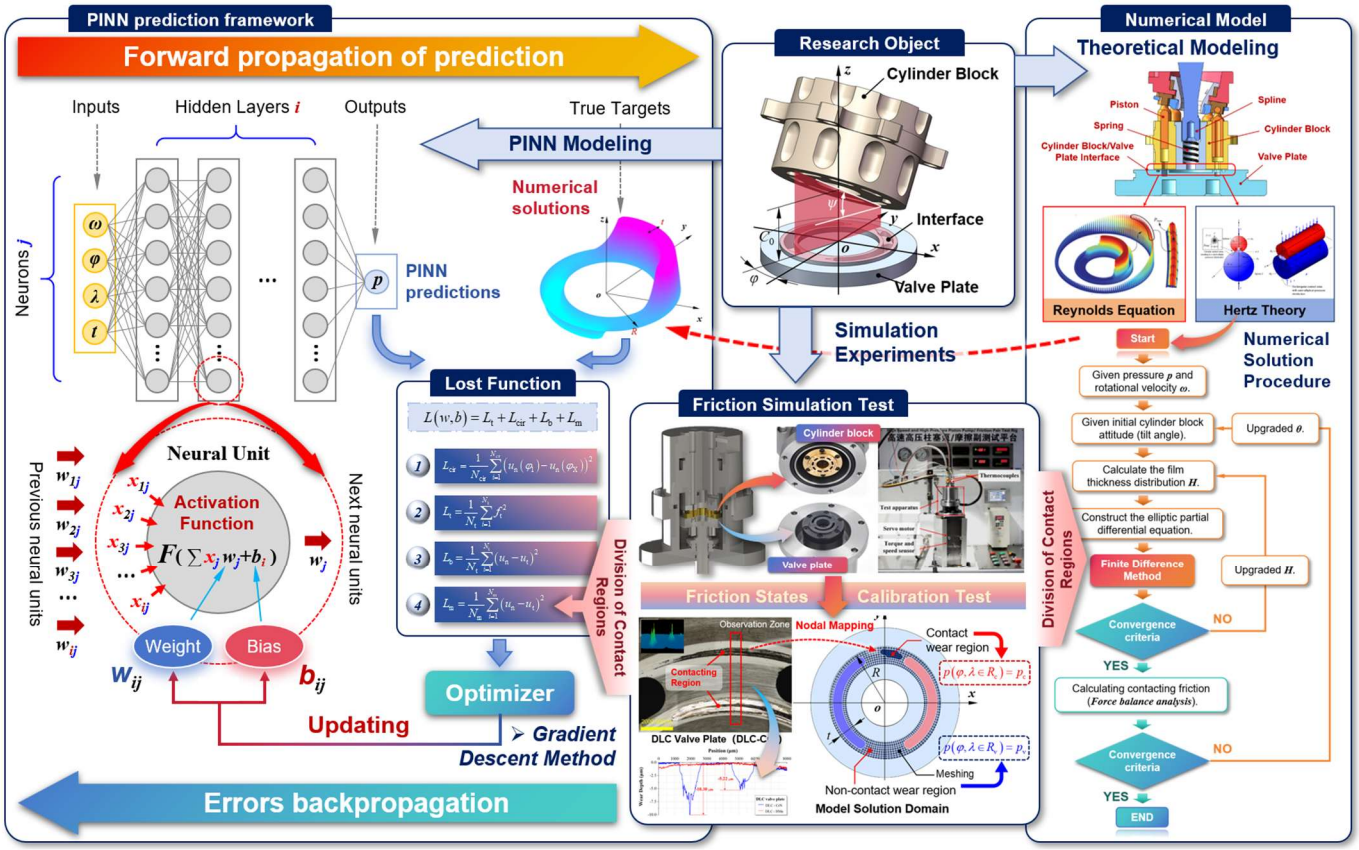


Figure10: Lubrication characteristic prediction framework with the proposed PINN model.

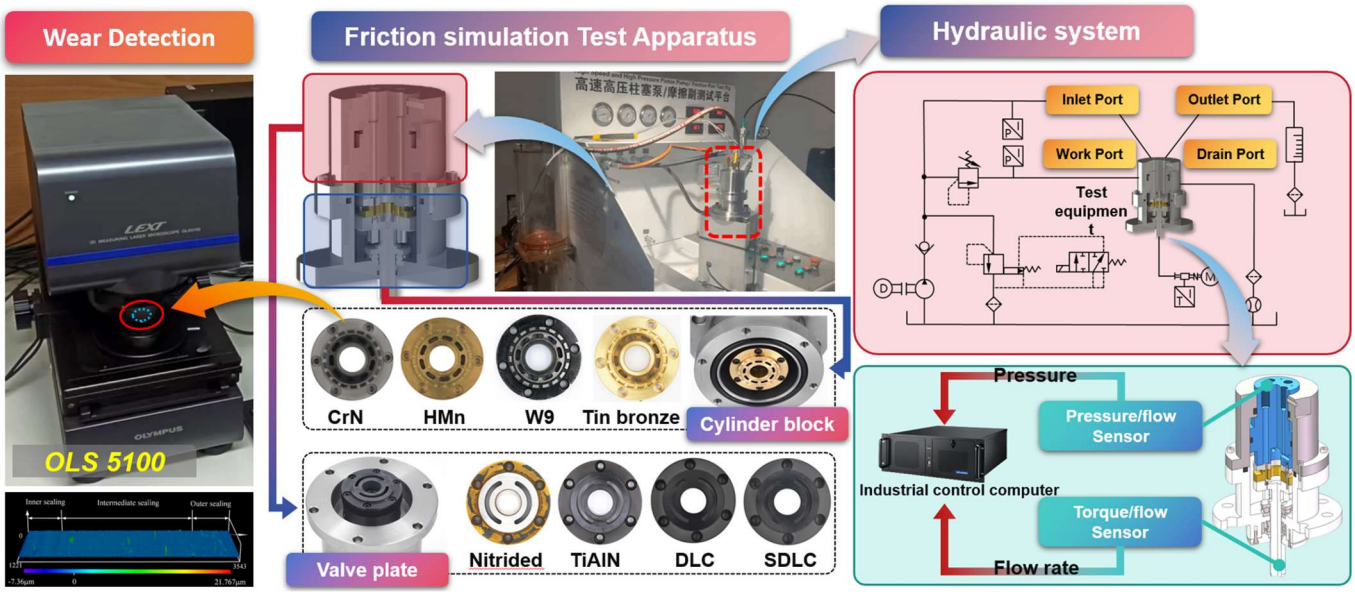


Figure11: High-speed and high-pressure piston pump/friction pair test platform.

Tests were performed under identical initial conditions (clamping force: 2.5 kN, system pressure: 7 MPa), with rotational velocity progressively increased from 0 to 10,000 rpm. Post-test analysis using a 3D laser measuring microscope (OLS5100) revealed interfacial wear patterns. Contact wear locations were identified, wear scar depths quantified, and the contact zone subset R_c was defined in the hybrid lubrication model based on measurements for experimental calibration. Fig. 12 shows concentrated wear near the dead zone, spanning 15–17% of the circumferential extent.

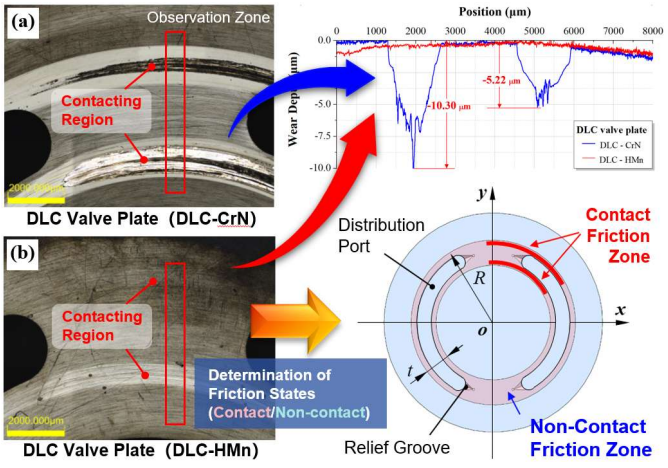


Figure*12: Determination of friction states based on experimental observation.*

Table3: *Experimental conditions (DLC-coated valve plate).*

Group	Specimens	Hardness	Conditions	Duration
1	CrN-coated	HV2200	2.5 kN/7MPa	5min
2	HMn	HV168	2.5 kN/7MPa	5min
3	CrN-coated	HV2200	3 kN/14MPa	5min
4	HMn	HV168	3 kN/14MPa	5min

4.3 Evaluation of prediction precision

By comprehensively considering the test conditions of the experimental platform, computational capabilities of the theoretical model, and lubrication/friction characterization methods, the mixed friction coefficient was selected as the observation metric. Under consistent initial operating parameters, experimental curves, numerical simulation results, and PINN-predicted curves of the mixed friction coefficient for the cylinder block/valve plate pair within the 0-10,000 rpm range were compared, as shown in fig. 13. The average error between the PINN prediction and the theoretical value is 4.132%. However, the maximum error compared to the experimental value is 17.716% due to the inherent errors in the theoretical model. Future studies should enhance PINN accuracy by increasing the incorporation of experimental measurement data (fig. 15).

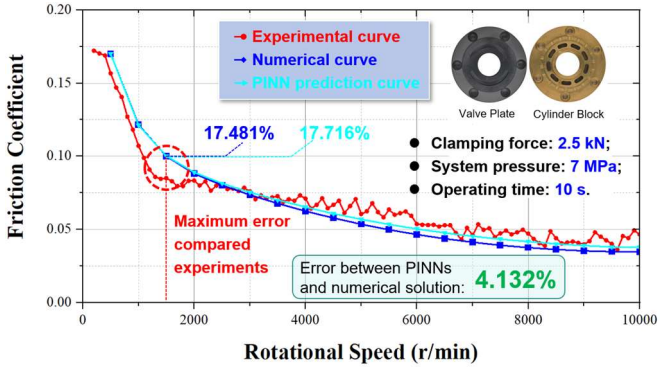


Figure13: Evaluation of predictive capability (mixed friction coefficient-speed curve).

5 Discussion

The study of lubrication under extreme conditions faces limitations in theoretical models (computational power) and experimental setups (hardware). PINNs allow efficient simulations of lubrication characteristics across extreme condition ranges.

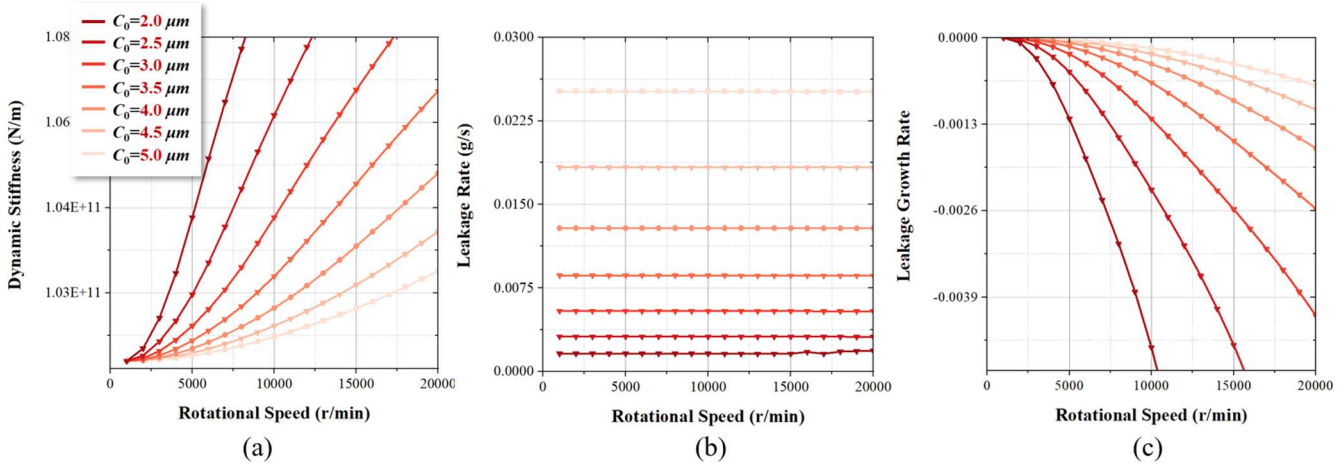


Figure14: Lubrication characteristic prediction: stiffness and internal leakage.

5.1 Dynamic lubrication analysis under extreme variable-speed condition

Lubrication characteristics were predicted over a wide speed range (0-20,000 rpm). Dynamic stiffness and internal leakage were selected as key indicators of lubrication performance and assessed at constant pressures (35 MPa load, 2 MPa system) over 10 s. As shown in fig. 14, dynamic stiffness increases monotonically with speed, while leakage decreases. Both parameters exhibit larger changes at smaller film thicknesses. Starting with an initial film thickness of 2 μm, increasing the speed by 5,000 rpm results in a 5×10⁹ N/m increase in dynamic stiffness.

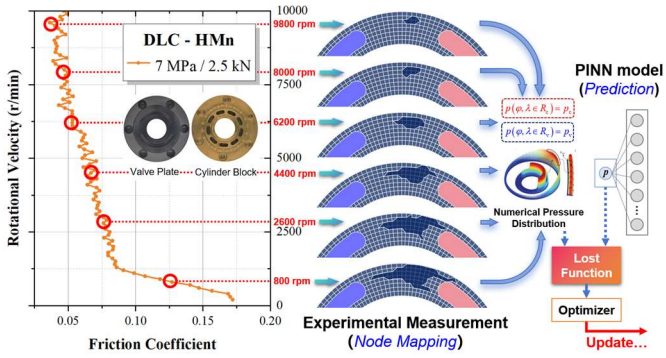


Figure15: Future perspectives: experimental calibration of variable-speed friction states.

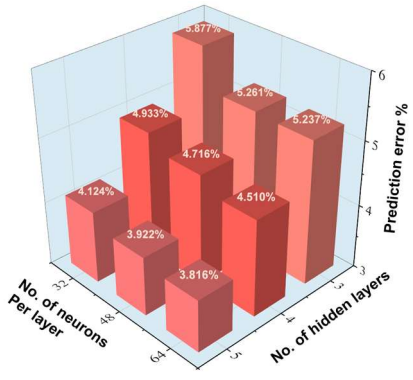


Figure16: Comparisons of prediction error.

5.2 Impact of network parameters on prediction

The impact of network depth and width on the prediction accuracy was investigated under extreme conditions (20,000 rpm, 35 MPa loading pressure, 2 MPa system pressure, 10 s duration), with 2,000 epochs of training. Results show that increasing the number of hidden layers generally led to a significant reduction in prediction errors. For instance, with 32 neurons per layer, the error decreased by 1.753% when increasing from 3 to 5 layers, compared to a 0.64% reduction when increasing neurons from 32 to 64 per layer (fig. 16). These findings suggest that network depth plays a more critical role in enhancing prediction accuracy than width, likely due to its ability to better capture the complex interactions in mixed lubrication regimes. They also highlight the need for topology-specific tuning in multi-physics applications.

6 Conclusion

This study pioneers a PINN model to decode extreme lubrication dynamics in aerospace hydraulic pumps. By synergizing Reynolds-based compressible fluid models, contact mechanics, and neural operators, the framework resolves multi-physics coupling (dynamic lubrication, asperity contact) with significant high efficiency. Validated experimentally, the PINN model achieves fast real-time analysis and <5% error in predicting mixed lubrication transitions and transient flow fields under variable speeds/pressures. This intelligent paradigm bridges first-principles physics with AI-driven modeling, offering transformative potential for next-gen pump design and condition monitoring under ultra-wide operational spectra.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study is supported by the Chinese national project of Electro-Hydrostatic Module and Electro-Hydrostatic Actuator.

References

[1] MARÉ J C, FU J. Review on signal-by-wire and power-by-wire actuation for more electric aircraft. *Chinese Journal of Aeronautics*, 2017, 30(3): 857-870.

[2] FU J, MARE J C, YU L, et al. Multi-level virtual prototyping of electromechanical actuation system for more electric aircraft. *Chinese Journal of Aeronautics*, 2018, 31(5): 892-913.

[3] SHOUJUN Z, JUN H, YUQIANG Z. The study on the dynamic capability of an electro-hydrostatic actuator to drive a large inertia load. In: *2016 IEEE International Conference on Aircraft Utility Systems*, 2016: 836-841.

[4] FU Y, LI Z, AN G, et al. Electrohydraulic pump development status and key technology review. *Machine Tool & Hydraulics*, 2012, 40(1): 143-149+160.

[5] ALLE N, HIREMATH S S, MAKARAM S, et al. Review on electro hydrostatic actuator for flight control. *International Journal of Fluid Power*, 2016, 17(2): 125-145.

[6] FU J, LYU D, KANG J, et al. Tribological characteristics of hard-to-hard matching materials of cylinder block/valve plate interface in electro-hydrostatic actuator pumps. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 2023, 237(16): 3692-3705.

[7] LYU D, FU J, CI K, et al. Experimental Research on Friction Characteristics of DLC Coated Cylinder Block/Valve Plate Interface in Aerospace Integrated Motor Pumps. In: *2023 Asia-Pacific International Symposium on Aerospace Technology Proceedings. Springer Nature*, 2024: 1141-1156.

[8] LYU D, FU J, ZHAO S, et al. Development of Boundary/Mixed Friction Models for Cylinder Block/Valve Plate Interface in a Novel Aerospace Electro-Hydrostatic Module. In: *BATH/ASME Symposium on Fluid Power and Motion Control. ASME Press*, 2024: FPMC2024-140246.

[9] ZHAO J, FU Y, MA J, et al. Review of cylinder block/valve plate interface in axial piston pumps: Theoretical models, experimental investigations, and optimal design. *Chinese Journal of Aeronautics*, 2021, 34(1): 111-134.

[10] LI J, ZHANG W, YUE J, et al. A deep learning method for incompressible fluid equations and coupled problems. Mathematica *Numerica Sinica*, 2024, 46(2): 232-252.

[11] KARKOUB M A, GAD O E, RABIE M G. Predicting axial piston pump performance using neural networks. *Mechanism and Machine Theory*, 1999, 34(8): 1211-1226.

[12] GUO J, YANG Y, LI H, et al. A parallel deep neural network for intelligent fault diagnosis of drilling pumps. *Engineering Applications of Artificial Intelligence*, 2024, 133: 108071.

[13] ROJEK M, BLACHNIK M. A Dataset and a Comparison of Classification Methods for Valve Plate Fault Prediction of Piston Pump. *Applied Sciences*, 2024, 14(16): 7183.

[14] HESS N, SHANG L. Development of a Machine Learning Model for Elastohydrodynamic Pressure Prediction in Journal Bearings. *Journal of Tribology*, 2022, 144(8): 081603.

[15] PERIC B, ENGLER M, SCHULER M, et al. Using Neural Networks as a Data-Driven Model to Predict the Behavior of External Gear Pumps. *Processes*, 2024, 12(3): 526.

[16] REN H J, YUAN X X, CHEN J Q, et al. Neural network prediction of Reynolds stress anisotropy tensor under wide-speed conditions. *Chinese Journal of Theoretical and Applied Mechanics*, 2022, 54(2): 198-212.

[17] EIVAZI H, TAHANI M, SCHLATTER P, et al. Physics-informed neural networks for solving Reynolds-averaged Navier–Stokes equations. *Physics of Fluids*, 2022, 34(7): 075117.

[18] RAMOS D J, CUNHA B Z, DANIEL G B. Evaluation of physics-informed neural networks (PINN) in the solution of the Reynolds equation. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 2023, 45(11): 568.

[19] WANG Z, ZHOU Z, XU W, et al. Physics informed neural networks for fault severity identification of axial piston pumps. *Journal of Manufacturing Systems*, 2023, 71: 421-437.

[20] DONG C, TAO J, SUN H, et al. Innovative fault diagnosis for axial piston pumps: A physics-informed neural network framework predicting pump flow ripple. *Mechanical Systems and Signal Processing*, 2025, 225: 112274.

[21] LIU X, CHENG W, XING J, et al. Hierarchical Physics-Informed Neural Network for Rotor System Health Assessment. *IEEE Transactions on Automation Science and Engineering*, 2024, 21(3): 1-14.

[22] YIN F, LUO H, NIE S, et al. Physics-Informed machine learning for tribological properties prediction of S32750/CFRPEEK tribopair under seawater lubrication. *Tribology International*, 2024, 199: 109965.

[23] MA Z, LIAO H, GAO J, et al. Physics-Informed Machine Learning for Degradation Modeling of an Electro-Hydrostatic Actuator System. *Reliability Engineering & System Safety*, 2023, 229: 108898.

New rotary-actuated proportional valves for humanoid robot

Franc Majdič, Dominik Božič, Jan Bartolj, Ana Trajkovski

Laboratory for fluid power and controls, Faculty of Mechanical Engineering, University of Ljubljana, Slovenia
E-mail: franc.majdic@fs.uni-lj.si

Abstract

Despite the popularity of electric drives, hydraulic power transmission is used very frequently because it has many advantages. Among the most important are the high-power density and the simple and effective overload protection. Imitating nature and humans often leads to new ideas and better machines, which is why we decided to build a human-like hydraulic leg. We were unable to find suitable, sufficiently small and efficient hydraulic valves on the market for controlling humanoid legs. To control the hydraulic legs of the humanoid robot, we decided to design a hydraulic proportional valve with the smallest possible mass and size. Despite its compactness, it must allow large flows for fast robot movements, and at the same time we want to ensure low internal leakage. Based on the principle of the rotary valve, we designed and dimensioned the proportional valve and optimised it for installation and production. To improve sealing and simplify production, we introduced the use of seals on the spool. We carried out numerical simulations of the basic properties and compared them with measurements on the manufactured prototype. Several specimens of the newly developed valve made of aluminium and brass were produced. At the same time, we tested the use of oxide layers on the aluminium surface, which can be used to increase the surface hardness. The newly developed valve represents a new approach in the control of hydraulic components and is suitable both for controlling humanoid robots and for controlling other devices in industrial and mobile hydraulics.

Keywords: hydraulics, humanoid robots, rotary proportional valves, valve seals, servo motor control

1 Introduction

Hydraulics is an indispensable part of drive technology in many areas. Hydraulic components are categorized into four groups according to their function: primary, secondary or execution, auxiliary and control components. Primary hydraulic components are pumps that convert drive energy into hydraulic energy. Executive components convert the energy of the fluid back into mechanical work in various forms. This group includes hydraulic cylinders and hydraulic motors. Auxiliary hydraulic components include filters, coolers and heaters, tanks, pipes, hydraulic accumulators, sensors, etc. The control components include hydraulic valves for controlling the direction, speed and power of the executing units. Valves are categorized into different groups depending on their function. We recognize flow, check, pressure and directional control valves. They are used to control the execution components, which in our case are hydraulic cylinders for the movement of a humanoid robot [1]. In the case of humanoid robots, other forms of actuation of execution components can also be used to move individual parts of the robot. These include electrical energy for driving electrical actuators and compressed air energy in combination with pneumatic cylinders, which are developing rapidly in medicine for patient rehabilitation [2]. In this work, we have opted for a hydraulic actuator because it has advantages over an electric actuator. The most important are the higher power density, the protection against overload and the good adaptation to dynamic loads. Hydraulic drives are currently also used by humanoid robots from one of the best-known American robot developers [3]. However, when controlling larger flows, there is a risk of high flow forces in classic sliding and longitudinal slide valves, which are predetermined by the use of two-stage control. This is problematic for humanoid robots, as such valves are larger and heavier, which means that their installation has an impact on the overall size and mass of the robot.

Due to the desire to reduce the size and mass of valves, in this work we wanted to investigate and find a solution in the field of sliding proportional rotary valves, which have a lower influence of the flow forces on the spool and consequently do not require a large control force for the direct control of large flows [4].

2 Theoretical background

Continuously operating directional control valves are hydraulic components for controlling the direction and flow of hydraulic fluid in and out of actuators. They enable the direction and speed of movement of the actuator's drive part, e.g. the piston rod of a cylinder, to be changed. They work according to the principle of generating a pressure difference that arises when the flow gap is reduced. Depending on their geometric shape and mode of operation, directional control valves can be categorized as longitudinally moving and rotary valves. Longitudinal valves are widely used in all areas of hydraulics, while rotary valves are only used for special purposes in hydraulics [6].

2.1 Longitudinally movable slide directional control valves

In general, each hydraulic valve controls at least one fluid line, but in most cases valves are used to control several different lines simultaneously. This need has led to the development of many different designs of proportional, longitudinally movable spool valves, which differ considerably in the number of ports, the number of spool positions and the connections between the individual ports in the individual positions. They also differ in the type of control, as they can be controlled manually or with a proportional solenoid.

The main difference between classic conventional and proportional valves is the desire to control the flow. Conventional valves do not allow this possibility, as their purpose is only to allow full flow or to interrupt it completely. These valves cannot assume intermediate positions. This also leads to a tendency to minimize the pressure difference that occurs when flowing through the valve, as this represents a loss of energy.

With proportional valves, on the other hand, it is precisely the pressure difference created by partially closing the valve that makes it possible to control the flow through the valve. Therefore, the nominal flow rate of proportional valves is specified with a pressure difference of 7 MPa [7] over the entire valve when the valve is fully open, which in most cases can be simplified to the flow rate via a single active port at a pressure difference of 3,5 MPa due to the symmetry of the valve design.

Proportional, longitudinally displaceable directional control valves (Figure 1) have a displaceable spool with notches inside the housing that moves along its axis and connects various combinations of channels in the housing with the notches [8]. Different principles are used to move the spool in the housing. The most commonly used control method is with a solenoid (direct) or indirectly with oil pressure from the control line, but we also know manual control via a control lever and a combination of electromagnetic and indirect (hydraulic) control, which is mostly used for valves with oil flow rates above 70 l/min.

With proportional valves, the oil flow is usually controlled via the edge of the notch on the spool. These can have different shapes, which makes it possible to utilize different valve characteristics simply by changing spool with different notches.

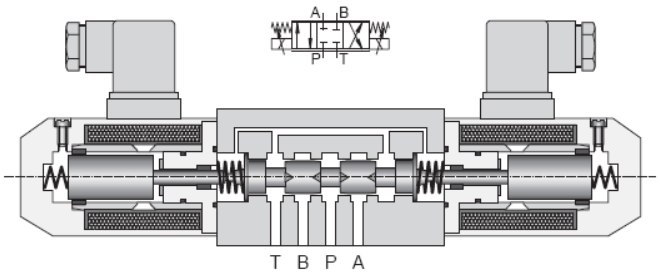


Figure 1: Proportional valve with direct electromagnetic control in cross-section and its symbol [8]

With proportional valves, the oil flow is usually controlled via the edge of the notch on the spool. These can have different shapes, which makes it possible to utilize different valve characteristics simply by changing spool with different notches. By selecting the appropriate shape of the notches and their number, we can ensure very precise changes in the oil flow. To determine the effects of different notches, CFD analyses are now used to simulate the fluid flow through different valve openings, which is detailed in modern research papers comparing different groove shapes [9]. In addition to more precise flow control, the geometric shape of the grooves and their distribution on the spool can also reduce the flow forces of the fluid on them. A reduction in the influence of the flow forces can also be achieved by additional geometric optimization of the entire spool [10].

Reducing the flow forces on the spool has an effect on the control force required to move the spool and also on the stiffness of the return springs that stabilize the valve in the neutral position.

Due to the sharp control edges over which they control the fluid, seals cannot be used with longitudinally moving spools that would ensure a good seal in the case of a closed valve in a spool with blocked ports. In these cases, sealing is only guaranteed by the precision of the spool and the bore in the housing, as the leakage of the valve depends on the radial gap. Leakage also increases with wear during the service life of the valve, so that valve leakage is also one of the indicators of wear [11].

2.2 Rotary slide directional control valves

In rotary valves, a rotational movement is used for control instead of a translational movement. The shape and design of the spool are therefore slightly different. Typical rotary valves have grooves located on different sides of the spool along the circumference, connecting individual ports together at a given rotation. In this case, certain grooves are connected to each other by a bore in the spool so that a cross position can be realized. An example of how a rotary spool works is shown in Figure 2.

Rotary valves are primarily used for special purposes, which is why they are specially developed for individual applications. One such valve is used, for example, to control the movements

of a humanoid robot, for which a patent with the designation US4966192A [12] has been granted.

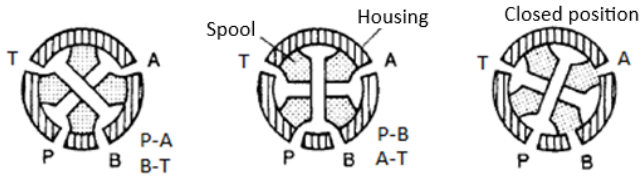


Figure 2: Functionality of a rotary valve [6]

This valve has a large number of control openings on the housing and on the spool, which supply different amounts of fluid to the individual working lines at different rotations of the piston. Among the patents we find a number of different designs of rotary valves that take into account certain advantages of the rotary design. A study was also carried out with a CFD analysis of a 2/2 rotary valve in cartridge design [4]. The aim of this study was to analyze the influence of the flow forces on the control torque and the pressure differences of the valve at flow rates of up to 1000 l/min. By implementing the cartridge design of the valve, they were also able to realize the two control edges in the form of a droplet. With this design, they achieved optimised flow characteristics depending on the angle of rotation of the control spool. They found that the torque of the fluid flow changes with the pressure difference across the valve and the angle of rotation of the control spool. However, under the selected conditions (1 MPa pressure difference when the angle of rotation is changed to full opening, with a flow rate of 1000 l/min), it does not exceed 2 Nm.

Based on the established facts, we decided to develop a new rotary valve to control the leg movements of a humanoid robot.

3 Design of rotary sliding valve

The basic idea of the new valve is based on the angular dependence of the spool position in the housing. A hole is drilled through both components, which runs radially through the center of the spool (Figure 3).

Figure 3: 11 mm diameter spool, with 4 mm diameter holes, 48 mm length and grooves for X-sealing rings

When fully open, such a valve design provides only one channel through the body and the spool without changing the flow cross-section of the channel, which minimizes losses in fluid flow. The inlet opening is located on one side of the housing and the outlet opening on the other. The valve is actuated by the rotation of the piston, which changes the instantaneous flow cross-section of the opening at the junction of the piston and housing. In this case, the rotation of the piston from the fully open position creates a constriction at the circumference of the spool, which generates a greater pressure drop through the valve, as in the case of a diaphragm or a valve with a longitudinally movable piston. In this way, the effects of the flow forces are avoided, which occur in the classic longitudinally moving valves, especially on the flat annular surfaces, where they cause uncontrolled movement of the control spool, which can be compensated for by the control system together with solenoid.

In the case of the newly designed valve, we avoid flat surfaces, from which we conclude that the required control performance of the entire valve is somewhat lower, even at higher flow rates. To control the legs of the humanoid robot, we opted for a 4/3-way valve with blocked outlets in the neutral position. Based on this decision, the basic concept was developed into a rough model of a 4/3-way valve that requires four connections for operation, namely the fluid supply P, the working connections A and B and the return connection T. For the desired valve, we need to ensure a simultaneously open connection between the oil supply P and the working line A and simultaneously between the working line B and the return T. The pressure supply and return lines must be open at the same time. As the pressure supply and return cannot be interchanged without an additional valve, the number of channels is increased to four so that each working line has one channel leading from the pressure supply and the other to the return. Each of the openings opens crosswise, so that a simultaneous connection of the P-A and B-T lines is achieved in the parallel position or in the crosswise P-B and A-T line connection, which corresponds to the classic 4/3 directional control valves.

When developing the rough model, we also followed the recommendations for the design of hydraulic components. One of these is that the fluid supply is directed to the central part of the valve so that it is surrounded on both sides by areas with lower pressure ratios. We have also considered the use of symmetry in the loading of the components and modelled the rough model of the valve based on these recommendations. The choice of the diameter of the spool valve is also related to the size of the bore through the spool, taking into account the overlap at the same time. In order to achieve the blocked spool in the neutral stable position, we must ensure complete closure or interruption of all connections between the individual ports. This means that there must be at least one additional hole of solid material on the circumference of the spool between the individual holes, which closes the inlets of all holes in the housing in the neutral position.

From these findings and correlations, a mathematical relationship can be defined between the spool diameter, the diameter of the individual bore on the spool, and the size of the overlap. In this case, we also assume that the holes are offset by 90 degrees along the circumference, resulting in a neutral position at an angle of 45 degrees, which we assume to be the starting point or neutral position. Therefore, the control angle of the valve from the neutral position to the parallel position corresponds to a rotation of 45 degrees and the rotation angle to the transverse position corresponds to -45 degrees from the neutral position. In order to achieve the described function of the valve, we must therefore divide the entire circumference of the control spool into 8 equal parts. However, since the bore is located on a curved surface, its diameter is not equal to the

length of the arc of the circle it occupies. Therefore, we can connect the diameter of the spool hole with the side length of a regular octagon inscribed in a circle, if we take into account the zero overlap. Taking into account the desired overlap, we can calculate the diameter of the holes from the knowledge of the other dimensions. Figure 4 shows the cross-section of the rotary valve.

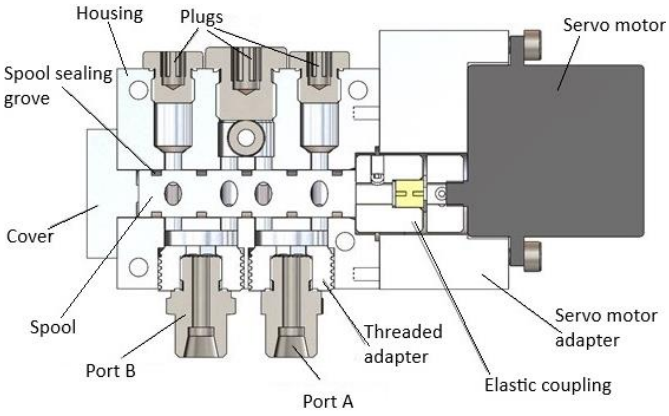


Figure 4: Cross-section of the entire rotary valve

By changing the movement of the spool valve from a back and forth movement along its axis to a rotation around the same axis, we achieved that the spool does not move longitudinally over sharp edges. Based on this realization, we can add grooves for seals on the spool to better seal the entire valve. As the spool rotates, it is advisable to use dynamic seals, which is why we opted for X-sealing rings. In order to achieve a good seal, the individual chambers (areas of the spool and the housing) must be separated from each other, which is why we need five seals in our case. When using seals, there is a problem with the initial assembly of the spool into the housing, although the seals do not run over sharp edges during valve operation. To solve this problem, we decided to further optimize the shape of the valve so that all the control edges of the housing are accessible. Figure 5 shows the final block design with three rotary valves for controlling a humanoid leg.

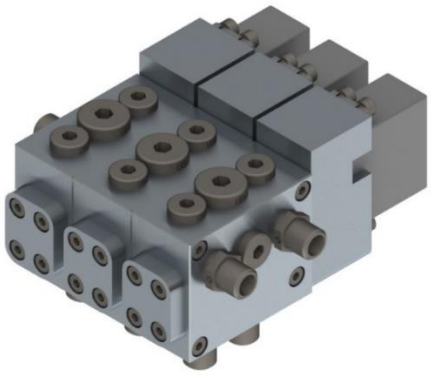


Figure 5: Final model of a valve in a block design for controlling one leg of a humanoid

Technical parameters of the rotary valve: Maximum operating pressure: 350 bar, maximum flow rate 60 l/min at pressure difference of 100 bar.

4 Numerical simulations

With the help of CFD simulations, we can simulate the flow of a fluid through the desired openings. In this case, we need to divide the free volume into small particles generated by volume meshing. As the number of particles increases, the result usually approaches the true value. However, as the number of analyzed particles increases, the number of nodes in which the values are calculated using the finite element method also increases. This leads to a tendency to have as few particles as possible, which is why a minimum number of particles in a line at a single cross-section is recommended for flow analyses. The simulation assumes that ideal conditions prevail. The system is in a stationary state, the fluid used is incompressible, the properties of the fluid do not change with time, the system is thermally insulated and has a constant temperature. The simulation also does not take surface roughness into account, as the surfaces are assumed to be ideally smooth. For all simulations, the hydraulic oil ISO VG46 was used, which has a density of 876 kg/m^3 at 40 °C and a kinematic viscosity of 46 mm^2/s [5]. In order to speed up the CFD simulations and reduce the number of calculations required, we carried out a geometric simplification of the model. This removed geometric details that could affect the need for denser meshing in places where the results are not crucial and could also cause larger errors in the results. The analyzed area is then determined and the entire area is meshed.

Figure 6 shows the numerically calculated pressure difference across the cross-section of the control channel at an oil flow rate of 5 l/min and a spool rotation angle of 15°.

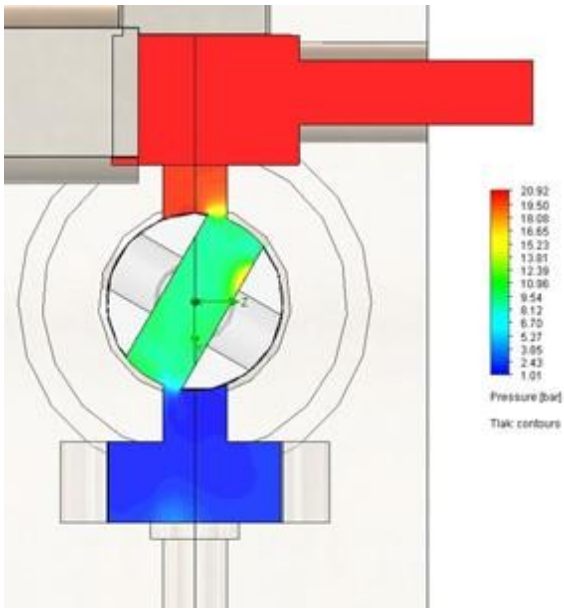


Figure 6: Numerically calculated pressure profile in cross-section with partially open valve

We wanted to use numerical simulations to obtain useful data on the theoretical values of the valve characteristics. These include the Δp/Q and Q/θ characteristics. With the Δp/Q characteristic curve, we determined the maximum flow capacity of the valve at a certain pressure difference across the valve, which is also used to determine the nominal flow rates of hydraulic valves. With the flow characteristic as a function

of the angle of rotation of the control spool, we were able to determine the fluid flow through the valve at different angles of rotation of the control spool or different degrees of valve opening from the data at the same pressure difference.

The pressure difference through the valve as a function of the flow rate was performed when the spool was rotated at an angle (θ) of 45° (full valve opening at port P-A and B-T). On the inlet side of the P and B port, we set the flow rates in the range from 0 to 80 l/min with a step of 5 l/min. On the outlet side, we set the ambient pressure as a boundary condition.

5 Experimental work

The designed rotary directional valve consists of standardised and non-standardised parts. All non-standardised parts were manufactured on a CNC milling machine and a CNC lathe. Machine reamers were used to precisely produce the bore for the control piston in the housing. This ensured the cylindricity of the entire bore, which together with the piston must ensure the lowest possible leakage, as the valve can leak not only via the sealing rings but also via the circumference of the control spool if the gap between the spool and housing is too large. A machine reamer was also used to produce the bores with control edges on the spool and housing, with which we wanted to ensure the best possible precision in the production of the control edges. We produced several prototypes of the valve, two with a brass housing and spool and two made of aluminium, which were surface-protected with an oxide layer [13, 14]. Figure 7 shows the rotary valve in a disassembled state.

Two types of tests were carried out with the newly developed rotary valve. The first measurements were short-term and mainly related to the properties of the valve in the steady state, while the second tests related to the permanent testing of the valve.

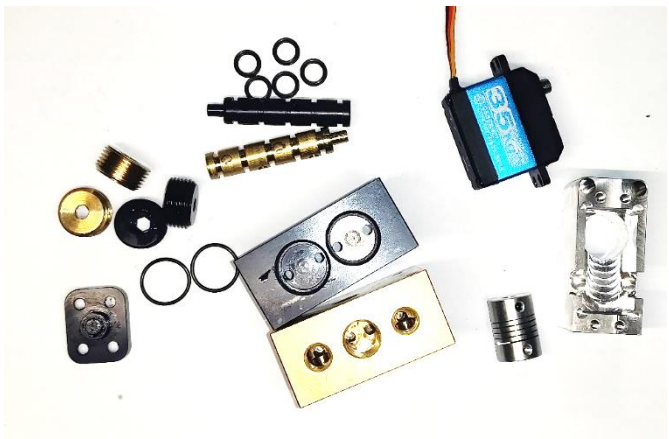


Figure 7: Manufactured non-standardized valve components with servomotor, seals and flexible coupling

Figure 8 shows the hydraulic schematic of the test rig for testing the new rotary 4/3 proportional directional control valve. The hydraulic pump with a variable displacement of up to 100 cm^3/rev (6) is driven by a 20 kW electric motor. The system pressure is set to 300 bar via a pressure relief valve (7). The flow rate can be preset with a manually adjustable flow valve (5) and measured with a flow sensor (4). This is followed by a temperature (3) and pressure sensor (2.1) and the newly developed valve – the test item (1). A pressure sensor (2.2) and a return filter (8) are also located on the A connection of the test item.

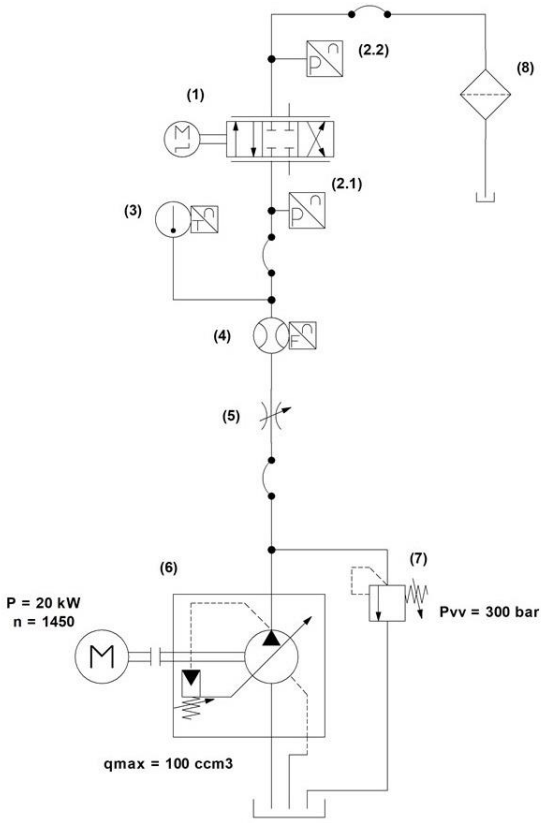


Figure 8: Hydraulic diagram of the valve connection for measuring its properties

With the long-term testing of a 4/3-way rotary valve, we wanted to find out what happens after a long period of operation in which the valve constantly switches between one extreme position and the other. To do this, we monitored the leakage of the valve in the closed position after specified time intervals. To carry out the tests, the rotary valve was connected to a new test station according to the diagram in Figure 9. The tested rotary valve (1) was connected to a hydraulic power pack with a gear pump (7) with a displacement of 8 cm3 and an electric motor with an output of 4 kW at a nominal speed of 1450 rpm. The hydraulic power pack (7) generated an oil flow of 12 l/min. A pressure relief valve (4) was installed between the pump (7) and the valve (1) to ensure that the pressure of the entire system did not exceed 200 bar. The working ports of the tested valve (1) were connected to a flexible hose into which we inserted a nozzle with a bore diameter of 0.6 mm (2), replacing the execution unit that would generate the working pressure on the working line. Both return connections were connected with flexible hoses and a T-port and passed through a water heat exchanger (6), which we used to cool the oil. From the heat exchanger, the oil travelled through another flexible line to the return filter on the power pack (7). Due to the choice of small connections, all flexible pipes had a size of 1/4". A pressure gauge (5) was connected upstream of the valve inlet

connection to check the system pressure at the P connection of the tested valve (1).

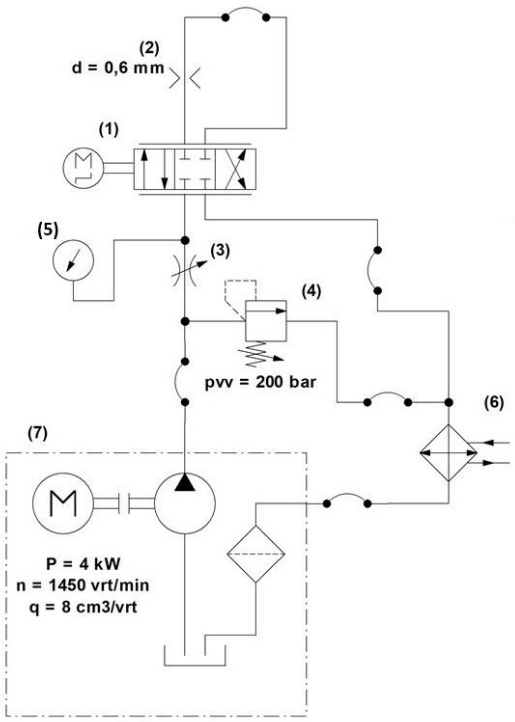


Figure 9: Hydraulic diagram of the valve connection for long-term test

During the endurance test of the valve, the main pressure relief valve was set to 200 bar on the test site. Using a nozzle with a hole diameter of 0.6 mm, we achieved a working pressure of 180 bar by setting the manual flow valve with the valve fully open. We interrupted the flexible line connecting the working lines for a certain time or the number of cycles performed and measured the internal leakage at working ports A and B with the valve closed. The amount of leaked oil was measured using measuring syringes with a measuring range of up to 20 ml. At the same time, we monitored the events and possible leaks at other points on the valve. The entire test was carried out at an oil temperature of 60 ºC, which was ensured by a heat exchanger.

6 Results

The numerically calculated Q-φ characteristic curve of the new rotary valve, which corresponds to the Q-s characteristic curve of other longitudinally moving valves, with the difference that it takes into account the angle of rotation φ instead of the axial movement of the spool, is shown in Figure 10. The result of the flow measurement as a function of the angle of rotation and the pressure difference is shown in Figure 11. Due to the limited performance of the hydraulic power pack at a pressure difference of 200 bar and 150 bar, not all results were achieved. We also measured the maximum torque required to overdrive the valve (Figure 12). In the tests, the greatest resistance occurred when the spool was moved from the fully closed position.

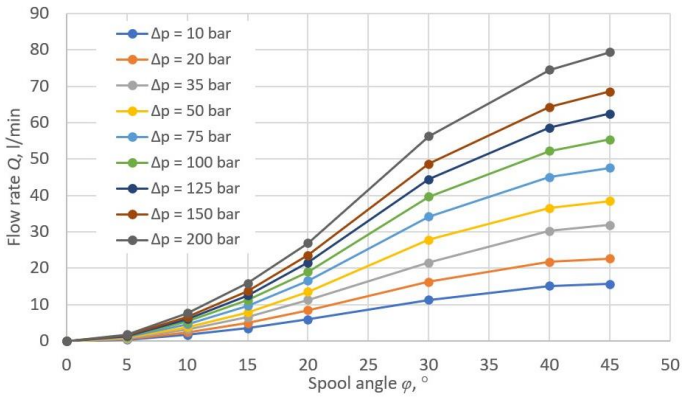


Figure 10: Numerically calculated *Q*- *φ* characteristics of the rotary-controlled valve

From the graph (Figure 11), we can see the similarity with the graph of the Q-φ characteristic curve obtained by numerical simulations (Figure 10). However, we also note some deviations due to the inaccuracy of the desired and actual position of the spool rotation. We found this anomaly in the measurements, but to eliminate this problem, we would need a servomotor with a better position reader that allows for higher accuracy.

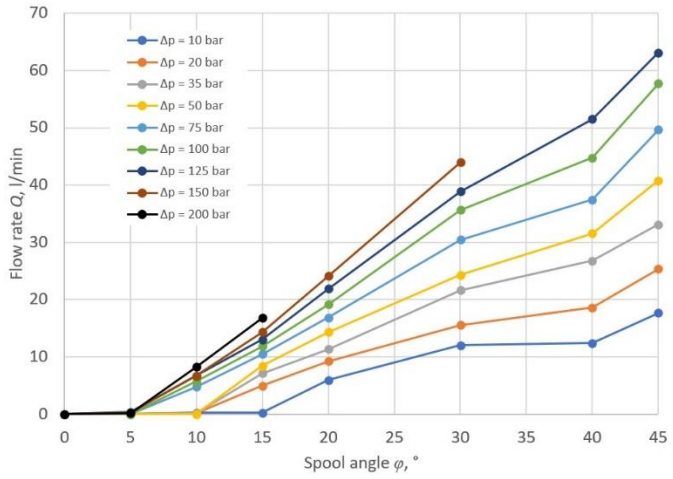


Figure 11: Measured Q-φ curves of the rotary-controlled valve

The graph (Figure 12) shows an almost linear increase in the required torque, which reaches a value of 0.1 Nm at a pressure difference of 0 bar. The result (Figure 12) also shows that a servomotor with a torque of 1.6 Nm is required to control the valve at a pressure difference of 300 bar. From this it can be concluded that this torque is a result of the friction of the seals on the housing and the spool, as no flow forces occur in the valve when it is fully closed and therefore cannot cause any torque. Part of this torque when the valve is fully closed can also be caused by the force that acts on the spool via the two P control edges in the housing and presses it against the opposite wall, as there is a certain amount of leakage between the spool and the housing.

We wanted to determine the service life of the valve by carrying out durability tests. The results showed that the newly

developed valve completed 733,000 switching cycles before failing. The tests also showed that the time for each complete (100 %) switching cycle was around 0.6 s, which means that the valve can switch at a frequency of 1.6 Hz. The measured switching time of the existing valve design (0.6 s) allows the robot to run slowly.

During the durability tests, we encountered some unexpected problems as some purchased components failed.

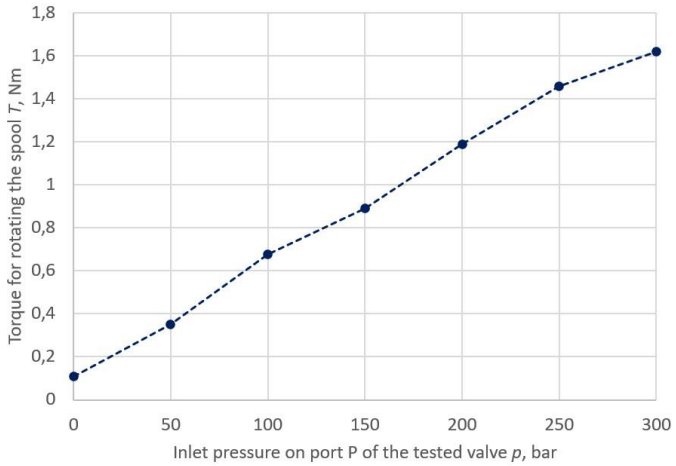


Figure 12: Measured spool torque as a function of the pressure difference

After 296,000 switching cycles, the servomotor failed due to wear of the gearbox in the power transmission. However, as we cannot influence the quality of the servomotors, we replaced the servomotor and continued the tests as we were interested in the lifetime of the main parts of the valve, i.e. the housing and the spool, and their wear. After 544,000 cycles, it turned out that the valve was no longer working due to a break in the flexible coupling between the spool and the servomotor. In this case too, we replaced the broken part and continued the tests. As a result, after the next 3,000 switching cycles, the spool moved axially to a position where the seals reached the control edges on the housing. This resulted in damage to the seals, causing the valve to leak outwards. The problem was solved by replacing the seals and reinstalling them with assembly inserts. After replacing the sealing rings, no new problems occurred until the end of the test. The test was concluded with the failure of the control spool, which broke in the shaft connection with the flexible coupling due to the exceeded dynamic continuous load (Figure 13).

Figure 13: Breakage of the control spool at the contact with the flexible coupling at 733000 switching cycles

During the durability test of the rotary valve, we measured the internal leakage at regular intervals (Figure 14). The last leakage measurement (after the failure) could not be carried out because the valve could no longer be turned.

We were very surprised by the results of the leakage measurements at the beginning of the test, as the measured leakage of the valve is very low. The diagram also shows the point at which the seal was replaced due to the problem described. After replacing the seals, a sudden increase in the leakage value can be seen, which is most likely due to major damage to the seals during reassembly. However, the valve showed no leakage at any other point, which surprised us, as only one sealing ring ensures the seal on both sides between the return zone and the environment.

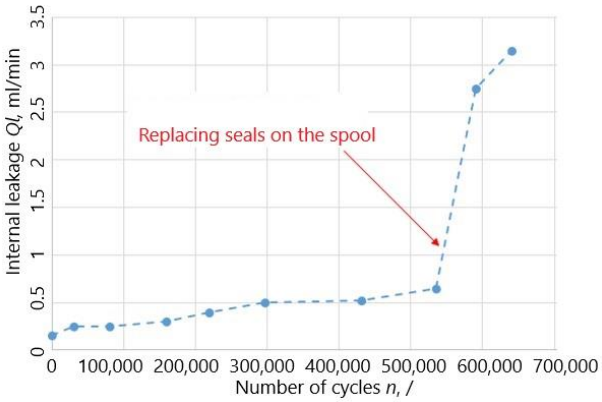


Figure 14: Measured internal leakage of the valve during the durability test

After completing the durability tests, we visually inspected the wear of the control edges on the spool and housing and also noted other areas of wear, which became lighter in colour after wear due to the application of a black oxide layer on the surface of the aluminium parts.

We observed the control edges on the control spool and compared them with the edges on the housing. A change can also be seen on the X-sealing rings. The second seal from the left in Figure 15 shows that only part of the seal has remained in the groove. We conclude that this seal has caused a sudden jump in the values in the diagram. It is interesting to note that the valve achieves a very good seal despite the almost completely damaged seal. Figure 16 shows the wear of the control edge on the spool, which was considerably greater than the wear on the housing.

Figure 15: Spool after 733,000 cycles with visible damage to seals and control edge

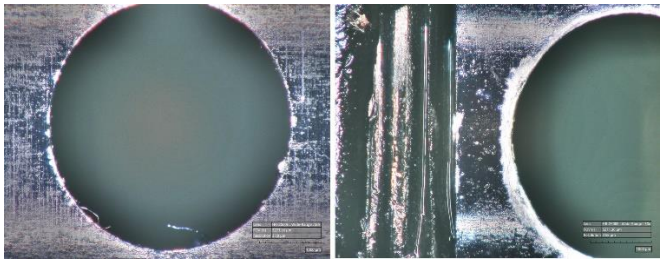


Figure 16: Control edge wear on the housing (left) and on the spool (right).

7 Discussion

From the analysis of the results, we conclude that we have succeeded in developing a new hydraulic rotary valve with a nominal flow rate of 33 l/min according to the criterion of a pressure difference of 35 bar.

With a nominal pressure of the hydraulic system for driving the legs of a humanoid robot of 200 bar, a flow rate of almost 60 l/min can be achieved through the valve. To achieve a higher flow rate, one would either have to increase the nominal working pressure of the hydraulic system of the humanoid robot or increase the valve capacity by increasing the diameter of the control edges in the valve. This would also mean increasing the size of the valve, which is not desirable.

The measurement results agree with the results of the numerical simulations. It is surprising that the measured values are very close to the values of the numerical analyses. Only when comparing the Q-φ curves (numerical and measured) is there a partial deviation, which is due to the inaccuracy of the servomotor positioning. To eliminate this anomaly, we would have to use servomotors with very precise positioning, which are more expensive than the motor used. As the control is realised with a feedback loop, we expected better positioning of the completed rotation from the chosen motor.

The durability test of the manufactured valve prototype showed that aluminium performed well as a material for the piston and valve body, contrary to expectations. We are convinced that the applied oxide layer on the surface, which significantly increased the surface hardness, contributed significantly to this.

Even after a fairly large number of cycles, the wear did not lead to excessive and unacceptable internal leakage of the valve. This was significantly lower than the internal leakage of longitudinally moving spool directional control valves [15] at the end of the endurance test. The decision to incorporate seals into the valve assembly itself certainly contributed to the surprising leak tightness results. This required the development of a solution for the valve fastening, which had to be further improved to enable multiple successful seal changes during the service life of the valve.

The durability tests also highlighted the need to be careful when selecting bought-in valve components, as some were found to fail earlier than the main valve components tested, contrary to expectations.

The life of the valve could be extended by optimizing the spool shaft at the point where the failure occurred during the durability test, as we did not pay enough attention to this detail in the spool design and the impact of the notch effect on the dynamic fatigue strength.

In the further development of valves, it would be useful to consider and test the effects of the spool material on the service life of the valve itself. It would also be worth investigating how the choice of seals affects the sealing and also the torque required to control the valve.

After comparing the results of the newly developed proportional rotary valve with the results of existing sliding, longitudinally moving proportional valves at a similar nominal flow rate, we can conclude that the newly developed valve allows the same flow rate with smaller dimensions. We also found that the flow rate of classic valves stabilizes or even decreases with increasing pressure difference across the valve, while the flow rate of the newly developed valve increases continuously at full opening.

If we compare the size of the newly developed 4/3-way rotary proportional valve with that of the classic, longitudinally moving spool valve, we see that the new valve has only 33 % of the volume of the classic valve with similar operating characteristics (flow/pressure) (Figure 17).

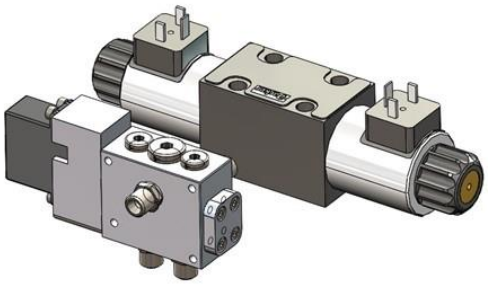


Figure 17: Size comparison between the newly developed valve (front left) and a conventional valve

Figure 18 shows humanoid hydraulic legs, each with three hydraulic cylinders (covered for protection) and three hydraulic rotary valves, each mounted on the support structure of the legs (top).

Figure 18: Humanoid legs with 2x3 rotary valves on top of the structure

8 Conclusions

By researching solutions for proportional spool valves, we have succeeded in developing a new hydraulic proportional spool valve with a rotary design which has significant advantages over conventional proportional valves in terms of internal leakage, size, control and manufacture.

By using seals on the spool, we were able to reduce the internal leakage by almost 100 times compared to conventional valves on the market. To implement the seals, we also had to develop a special device for assembling and disassembling the valve, as the sharp edges otherwise make it impossible to install seals. We also produced a few examples of the newly developed valve from aluminium and brass because it was easier and cheaper to manufacture. We have also tested the use of oxide layers on the aluminium surface, which can increase the surface hardness.

Increasing the surface hardness extends the service life of the entire valve, as the reduced wear, particularly of the control edges, means that critical internal leakage occurs later on, at which point the valve can no longer be used.

The properties of the newly developed valve were calculated numerically using fluid dynamics (CFD) and measurements were carried out on a prototype valve. It was found that the measurements were surprisingly similar to the simulated values, both in terms of the shape of the curves and the values. There were some problems with the measurements of the Q-φ characteristic due to the insufficiently precise servomotor, but despite these problems we were able to achieve acceptable results.

It also turned out that despite the simple design, the new rotary proportional valve completed 733,000 switching cycles before failing the durability test, which could be improved in the future by optimizing the spool shaft.

The newly designed valve represents a new approach to the control of hydraulic components and is useful for controlling humanoid robots and other devices in industrial and mobile hydraulics.

Nomenclature

Variables used in equations are presented in the table below.

Designation	Denotation	Unit
φ	spool rotational angle	°
d	nozzle diameter	mm
n	rotational speed	rpm
n	number of cycles	/
p	pressure	Pa
Δp	pressure difference	Pa
Q	flow rate	l/min
Ql	internal leakage	ml/min
q	pump displacement	cm^3

References

[1] M Bešter. Prototip hidravličnih humanoidnih robotskih nog: diplomsko delo. Ljubljana, 2021.

[2] Q Liua, J Zuoa, C Zhua, S Q Xie. Design and control of soft rehabilitation robots actuated by pneumatic muscles: State of the art, *Future Generation Computer Systems* 113, 620-634, 2020.

[3] K Khokar, P Beeson, R Burridge. Implementation of KDL Inverse Kinematics Routine on the Atlas Humanoid Robot, *Procedia Computer Science 46*, 1441 – 1448, 2015.

[4] I Okhotnikov, S Noroozi, P Sewell, P Godfrey. Evaluation of steady flow torques and pressure losses in a rotary flow control valve by means of computational fluid dynamics, *International Journal of Heat and Fluid Flow 64*, 89-102, 2017.

[5] D Lovrec. Vpliv indeksa viskoznosti različnih hidravličnih tekočin, *Ventil 4*, 272-281, 2020.

[6] D Will, N Gebhardt, H Ströhl. Hydraulik, *Springer-Verlag, Heidelberg*, 2007.

[7] P Chapple. Principles od hydraulic system design, *Oxford*, 2003.

[8] D Scholz. Proportional hydraulics, *Festo didactic*, 1995.

[9] Y Ye, C Yin, X Li, W Zhou, F Yuan. Effects of groove shape of notch on the flow characteristics of spool valve, *Energy Conversion and Management 86*, 1091-1111, 2014.

[10] R Amirante, E Distaso, P Tamburrano. Sliding spool design for reducing the actuation forces in direct operated proportional directional valves: *Experimental validation, Enery conversion and management 119*, 399-410, 2016.

[11] R Ewald, J Hutter, D Kretz, F Liedhegener, W Schenkel, A Schmitt, M Reik. Proportional and servo valve technology, 1986.

[12] S Potter, J A Saunders, Boston dynamics: *Rotary directional control valve*, US4966192A, 1990.

[13] M Mehdizade, A R Eivani, M Soltanieh. Characterization of the anodic oxide layer deposited on severely deformed and aged AA6063 aluminum alloy, *Jurnal of Materials Research and Tehnology 15*, 68-85, 2021.

[14] K H Rashida, A A Khadom. Sulfosalicylic/oxalic acid anodizing process of 5854 aluminum-magnesium alloy: Influence of sealing time and corrosion tendency, *Results in Chemistry 4*, 100289, 2022.

[15] J Bartolj. Razvoj 4/3 proporcionalnega hidravličnega ventila za izdelavo s postopkom 3D tiska kovin, *diplomsko delo*, Ljubljana, 2022.

Experimental and analytical investigation of a highly integrated electro-hydraulic pump motor unit

Lukas Matias[12], Andrew Plummer[1], and Nigel Johnston[1]

[1] Centre for Power Transmission and Motion Control, University of Bath, UK
[2] Domin Limited, Bristol, UK
Email: lwnm20@bath.ac.uk; lukas.matias@domin.com

Abstract

Hydraulic pumps driven by electric motors (ePumps) sit at the core of the electrification of hydraulic systems with applications in electro-hydrostatic actuation for aerospace, automotive and construction machinery. ePumps are rarely designed as an integrated unit, but use readily available pumps and electric motors, coupled via a drive shaft. Aerospace and automotive applications require high power density and compactness to minimize space and weight requirements on the vehicle. This paper presents the experimental and analytical investigation of a novel design integrating a radial piston pump and brushless DC motor into a single very compact unit. ePump volumetric and mechanical losses are investigated experimentally, and analytical modelling approaches are used to assess pump loss results. Methods for testing and modelling are integrated into a framework, outlining a comprehensive approach for the investigation of losses within ePumps. The breakdown of measured losses for such a unit presents a challenge however, as the pump shaft is not easily accessible in order to measure torque. One approach is to infer frictional losses in the pump by experimentally determining overall power loss and subtracting volumetric pump loss. Overall power loss is the difference between electrical input power and fluid output power, and so in this approach the electric machine and power electronics losses are not accounted for. To separate electrical machine losses from frictional losses in the pump, a measurement of the pump shaft torque is required. A non-invasive method of measuring the drive shaft torque via the reaction torque on the motor stator is proposed and was successfully proven out. The torque results have yielded credible and repeatable results, enabling the analysis of power loss fractions. The comparison between analytical and experimental results allowed to identify medium speeds and higher pressures as favourable operating conditions of the unit, as per design intent. The successfully validated ePump loss analysis framework will provide a guide for a detailed breakdown of losses in highly integrated electric pumps, facilitating improved designs with high energy-efficiency and a high level of compactness.

Keywords: ePump, integrated electro-hydraulic pump motor, radial piston pump, brushless DC motor, electrification

1 Introduction

Highly integrated ePumps are a key enabler in the electrification of hydraulic systems. Applications range from off-highway vehicles [1] [2] to automotive brake-by-wire (BBW) systems [3] and aerospace actuation [4]. The electrification of hydraulic systems has become a key research area to facilitate the transition towards low carbon emissions from the transport sector as part of initiatives such as the European Green Deal [5], calling for a 90% reduction in transport emissions by 2050. Fixed displacement pumps coupled with variable speed drives are the most common drive configuration for decentralized electro-hydraulic actuators [6], enabling more energy-efficient hydraulic systems. Typically, the electric motor and pump are separate components, and the integrated design can offer the benefits of a reduced package size, a lightweight design and the opportunity to coordinate the design of pump and motor for optimized performance.

A commercial design was first presented by Voith Turbo, integrating an internal gear unit within an asynchronous electric motor and promoting low noise emissions as well as up to 50% less space requirements compared to conventional solutions [7]. Today, the system is not available anymore. In 2007, Wustmann and Helduser first introduced a design concept for an external gear pump centrally integrated within

a switched reluctance motor to enable a higher level of integration over the common shaft coupling [8]. The design eliminated shaft seals and offered the opportunity of electric motor cooling to enhance thermal durability. However, in simulation the design only proved feasible for low pressure applications and featured a large amount of unused structural space. The external gear pump based concept was picked up again by Zappaterra et al. in 2022 [9]. An optimization study considering both pump and electric motor parameters was presented, and a prototype experimentally validated.

A gerotor pump based design, with the outer gear acting as the electric motor rotor and directly driven by the motor was presented in 2017 by Gamez-Montero et al. [10]. The feasibility of the design was experimentally proven with limitations in pump speed (test only up to 500 rpm), but expected benefits in noise.

In a similar approach, Fiebig et al. presented a vane pump based design [11] in 2014. The vane pump sits in the centre of a surrounding brushless direct-current (BLDC) motor, achieving a 40% reduction in volume compared to conventional designs, according to the authors. The analysis work centred on proving the dynamic control benefits of such a system. In 2017 and 2021, the research group followed up with an experimental proof-of-concept of the vane pump design, validating its control performance and the ability to build pressure and deliver flow [12] and [13].

An axial piston machine concept was first published by Wang and Zhou [14] in 2016. Following an analysis of preferable integration structures, the pump-in-centre approach was again adapted with the cylinder barrel acting as the electric rotor and directly driven by the motor. A first experimental proof-of-concept was successful. A similar design approach was taken by Zhu et al. in 2020, with the focus on an aerospace application [15]. Sarode et. al. presented a permanent magnet synchronous motor (PMSM) drive axial piston machine for off-highway applications in 2023 [16]. The numerical work included the optimization of the electric motor for a set of design specification and considered typical design trade-offs in electric motors such as current density and weight.

A radial piston-based design (“TwinTorq motor”) was patented by KersTech, Inc. and combines a multi-lobed, inside-impinged hydraulic motor with an electric ring motor around the pump circumference for electric vehicle applications [17]. KersTech, Inc. promotes a high efficiency motor with 25% to 45% efficiency improvements shown in simulation over conventional units. In 2019, Bohach et al. investigated the suitability of a radial ball piston pump design in a numerical study [18]. The same research group published an analysis of suitable system topologies with a focus on the electric machine design, eventually proposing an axial flux PMSM [19]. The design of this machine was discussed in [20]. In academia, a radial piston pump design concept tightly integrated with an axial flux PMSM was presented by Nishanth et al. [21] in 2022, with the cam and cylinder block sitting inside the stator. A numerical model of the pump was discussed by Nahin et al. in [22]. Li et al. numerically studied a similar design in 2019 [23].

Most of the discussed designs are targeted at off-highway applications and are generally designed for high power applications with power ratings > 5 kW. This paper presents a novel highly integrated, additively manufactured (AM) radial piston pump unit, driven by a BLDC motor for low power automotive and aerospace applications. As discussed in [24], the application of AM technologies provides the opportunity for further improved power density, and integrated sensing and control within the unit to enable robust control and health monitoring. The presented unit not only integrates motor and pump but also features onboard power electronics and closed-loop speed control of the pump with the opportunity of sensing the motor temperature for condition monitoring.

A common challenge with highly integrated electric pumps is the breakdown of losses in the unit as part of a performance characterization. This is key in the understanding of critical design parameters. As most of the presented design solutions are shaftless, a direct pump shaft torque measurement is not possible. Whilst mechanical losses in the unit may be inferred from volumetric and total unit losses, a clear separation cannot be made. This study presents an ePump loss analysis framework, including a non-invasive torque measurement approach, to further increase the understanding of highly integrated electric pumps. This will enable improved designs for future iterations and facilitate a design process with the aim of an optimized design in terms of size, weight and efficiency.

The paper outlines the design of the prototype demonstrator unit for the purpose of this study. That is followed up with an introduction of the process and methods used for the characterization of highly integrated ePumps. The test set-up and procedure are described in detail with the addition of a pump torque test add on. An analytical model is introduced and validated against experimental data. The loss distribution within the pump is discussed and favourable operating conditions for each of the components are identified.

2 Highly integrated pump-motor unit

As a demonstrator for this study, an early prototype was used, with key specifications listed in Table 1. The ePump is designed as a fixed displacement, inside-impinged radial piston pump, directly integrated into a brushless DC motor as the primary driver. The chosen motor type is beneficial for this type of application as it is a simple design, allows for high performance speed control, and offers higher power densities. The radial piston pump architecture is favorable, as it allows a compact axial integration with the motor. The unit diameter is mainly driven by the motor, with the pump adding simply some stack length. The cam profile can be chosen freely to realize a single, or multilobed profile, allowing the the number of piston strokes per revolution to be changed. The unit uses a proven arrangement of the rotor in oil and the stator in air. They are separated by a non-magnetic pressure tube, removing the need for dynamic rotary seals. This comes at the downside of having an increased non-magnetic “airgap”, reducing the magnetic loading of the motor. Directly on top of the stator sits a printed circuit board, integrating motor driver, control and power electronics. The motor driver uses field-oriented control to achieve position (angle) and speed control of the rotor, allowing for high precision and dynamic control. The pump

and motor share a common rotor, reducing bearings and enabling an ultra-compact assembly. A schematic of the unit is shown in Figure 1. The pump features 7 pistons and a double lobed cam. This ensures a smooth operation by 14 piston strokes per revolution, minimizing flow ripple. The pump inlet is part of the AM pintle, designed to achieve a low flow resistance to minimize the risk of cavitation.

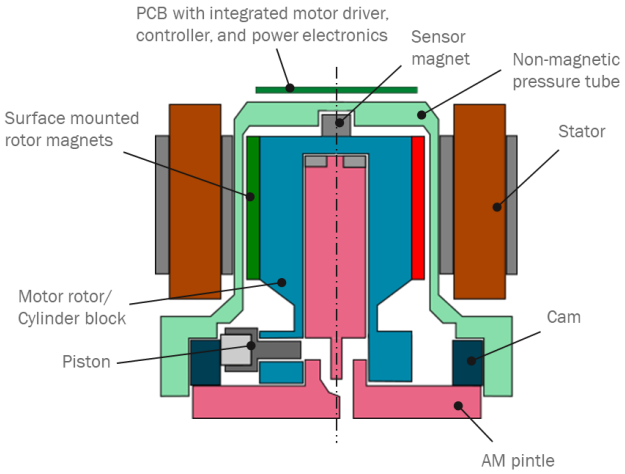


Figure 1: ePump CAD cross-section.

The key loss sources in the unit can be separated into volumetric losses and mechanical losses. Volumetric losses occur mainly as leakage between piston and piston bores, as well as between the pintle and cylinder block interface. Mechanical losses appear mainly as viscous friction in those same interfaces. Additionally, the positive force balance at the individual pistons exerts an outwards force towards the cam profile. The friction between the needle roller bearing sitting at the top of the piston and the cam results in frictional torque, opposing the rotation of the pump shaft. Furthermore, bearing friction between the interfaces of cylinder block and pressure tube as well as cylinder block and pintle is present. Lastly, churning losses occur in the flooded rotor cavity.

Table 1: ePump specifications.

Specification	**Value**
Displacement	0.55 cc/rev
Peak installed motor power	1.2 kW
Supply voltage	48 V

2.1 General method

Typically, steady-state hydraulic pump performance is evaluated based on steady state machine testing as outlined in ISO 4409:2019 [25]. Whilst discussions are ongoing on the suitability of methods outlined in the standard [26] [27], for the purpose of this work, the accepted standard will be used in the following. These tests will produce information about volumetric and mechanical losses, the effective outlet flow rate and effective inlet mechanical power for a range of operating conditions (pressure and flow). Based on those, statements about leakage and friction in the unit can be made. This understanding then feeds back into the design process where critical design parameters such as clearances, sealing land lengths, cam profile and spring rates are being defined.

Volumetric losses are considered as flow loss and are defined as the difference between the theoretical flow rate Q_{th} and the measured pump outlet flow rate Q_P:

$$Q_L = Q_{th} - Q_P \tag{1}$$

The theoretical flow rate is defined as the product of measured pump speed ω and effective pump displacement V_D:

$$Q_{th} = \omega V_D \tag{2}$$

The effective pump displacement was determined experimentally, following the approach as suggested in ISO 8426 [28]. The volumetric power loss can then be quantified to

$$P_{L,vol} = Q_L \Delta p \tag{3}$$

Mechanical loss is considered as additional pump shaft torque and is defined as the difference between the pump shaft torque (measured) T_P and the theoretical torque T_{th} that is dependent on the pressure difference across the pump Δp and the effective pump displacement:

$$T_L = T_P - T_{th} \tag{4}$$

With

$$T_{th} = \Delta p V_D \tag{5}$$

The mechanical power loss is therefore:

$$P_{L,mech} = \omega T_L \tag{6}$$

The overall power loss of the unit can be determined via the total electrical input power (measured) $P_{in,el} = I_{DCBus} V_{DCBus}$ and total hydraulic output power (measured) $P_{out,hyd} = Q_P \Delta p$ as

$$P_L = P_{in,el} - P_{out,hyd} \tag{7}$$

This finally allows the losses introduced by the electric motor and power electronics stage to be determined as the difference in losses between the combined ePump and pump:

$$P_{L,el} = P_L - (P_{L,vol} + P_{L,mech}) \tag{8}$$

As stated above, to quantify total and mechanical losses, information about the pump shaft torque is required. Due to the shaftless design, the pump cylinder block and the motor rotor are a single part (which is running in oil), and access to the rotor for mounting of an inline torque transducer is not possible. There is the option to infer electric motor torque which is assumed to be equivalent to the pump shaft torque from multiplying the measured quadrature axis current I_Q with the motor torque constant K_T:

$$T_{current} = I_Q K_T \tag{9}$$

There are however severe disadvantages of this approach. Firstly, the torque constant is often derived via simulation and

cannot be easily experimentally validated. Secondly, this approach does not allow electrical and mechanical losses to be distinguished. Previous tests have shown that motor calibration procedures heavily influence how much current is drawn for a given load on the pump. As a result, this method is not seen as appropriate for a detailed ePump performance characterization. As an alternative, a non-invasive approach to measure torque is proposed, described in Section 2.2.

With information about the pump torque $T_P = T_{EM}$ (equal to the electric motor torque), pump speed $\omega_P = \omega_{EM} = \omega$ (equal to the electric motor speed), DC bus supply current I_{DCBus} and supply voltage V_{DCBus}, electric motor and power stage losses can be stated as follows:

$$P_{L,el} = (I_{DCBus} V_{DCBus}) - (T_{EM} \omega_{EM}) \qquad (10)$$

If the operating envelope is defined according to the motor power limit envelope, this will give a very clear picture of electric motor performance in the region of interest for the ePump application. Figure 2 illustrates the proposed framework.

The results from the experimental tests will then allow pump and motor models to be validated. These models will serve as design tools for subsequent design iterations.

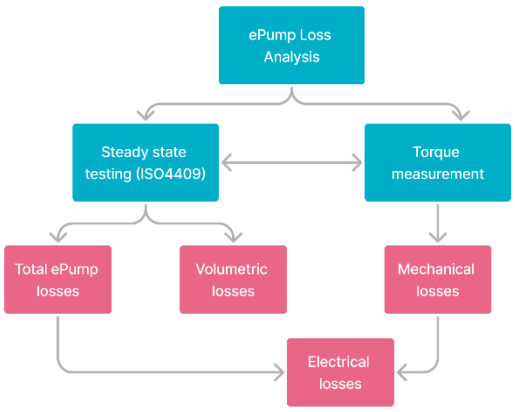


Figure 2: ePump loss analysis framework.

2.2 Non-invasive torque measurement method

Directly measuring torque in a highly integrated ePump is challenging. In this design, the motor rotor/ pump cylinder block runs in oil and is separated from the stator and power electronics via a pressure tube made from non-magnetic material. Direct access to the rotor for an inline torque sensor would require considerable changes to the design and a dynamic seal for a feed-through shaft, potentially compromising results.

As an alternative, it is proposed to measure the reaction torque on the stator via a load cell arrangement. Figure 3 shows a schematic of the set-up. The stator inner diameter sits on the pressure tube outer diameter with sufficient clearance to move freely. A mounting bracket connects to the stator and then via a rod assembly to a load cell. Ball joints on both sides of the rod ensure the elimination of side forces, that could compromise the measured force. The reaction torque leads to a force F_{LC} into the load cell, the torque $T_{reaction} = F_{LC} l$ is then determined with the known lever arm l.

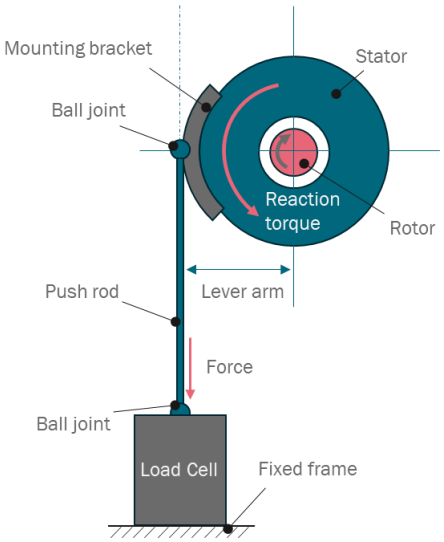


Figure 3: Schematic drawing of the load cell set-up.

2.3 Analytical loss model

For an initial analysis of the pump performance trends, a simple analytical model (constant coefficients) based on the work of Wilson, McCandlish and Dorey [29] [30] is used. The delivery flow rate of the pump is calculated as

$$Q = \omega V_D - c_s \frac{V_D \Delta p}{\mu} - \frac{\omega V_D \Delta p}{B}(V_r + 1) - c_\omega \omega \qquad (11)$$

The first term is the ideal pump flow rate. The second term represents laminar leakage flow with the slip coefficient c_s. The third term considers compressibility losses with the ratio of dead volume V_{dead} and displacement volume V_D as $V_r = V_{dead}/V_D$ and B as the bulk modulus. The fourth term is an addition to account for a speed dependence of flow losses (with c_ω as the speed coefficient) as observed in experiments and is not found in the original model descriptions.

The shaft torque is calculated as

$$T = \Delta p V_D + c_v \omega \mu V_D + c_f \Delta p V_D + T_c \qquad (12)$$

The first term is the ideal torque, and the second term considers viscous friction in the unit with the viscous friction coefficient c_v. The third term represents coulomb friction, scaled by the coulomb friction coefficient c_f. The fourth term is a constant torque component T_c and appeared in Wilsons model but was omitted in the McCandlish and Dorey models due to its usually small magnitude. It was found helpful for this analysis and is therefore included in the model.

3 Experimental validation

3.1 Power loss test set-up

The tests were conducted on a pump performance test rig, following the suggested layout as per ISO 4409 [25] and shown in Figure 4. The pump inlet is fed by a hydraulic power unit (HPU), with a pressure reducing valve allowing to set the pump inlet pressure. Tests are conducted in open circuit, with the pump outlet pressure being manually set by the pressure relief valve for each test run. The ePump speed is set via the internal pump controller. The set-up compromises flow meters, pressure transducers and temperature sensors for both inlet and outlet. The case is permanently connected to the tank line. As both inlet and outlet flow measurements are available, the drain flow is not recorded. The shaft torque is measured indirectly via the proposed load cell set-up (see Section 3.2). The motor speed is measured via a hall effect sensor, integrated into the ePump assembly. Specifications of the installed instrumentation are summarized in Table 2.

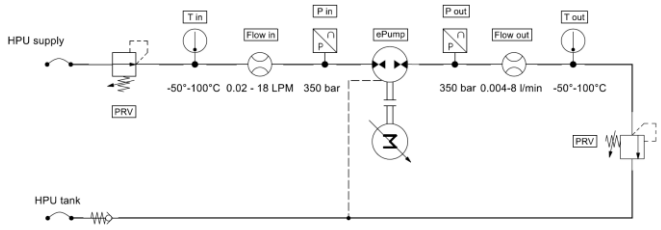


Figure 4: Pump test rig hydraulic circuit diagram.

For steady-state testing, the outlet pressure is set so that the following data points for differential pressure across the pump are generated: [22, 48, 70, 97, 122, 148, 174] bar. The outlet pressure deviation is less than 1% for each speed setting. The inlet pressure is set to approximately 7 bar. For each pressure setting, the pump speed command is varied between 25 rad/s and 175 rad/s in 25 rad/s steps. Each pressure and speed combination is held for about 25 s.

Table 2: Instrumentation.

Sensor	Model
Flow meter (inlet)	VSE Flow VS 0.2 Circular Gear Flowmeter, 0.02 – 18 LPM
Flow meter (outlet)	KOBOLD DZR-4 Gear Flowmeter 0.004 – 8 LPM
Pressure transducer (inlet)	TE Connectivity EB100 350 bar 120 Hz Bandwidth (-3dB)
Pressure transducer (outlet)	TE Connectivity EB100 350 bar 120 Hz Bandwidth (-3dB)
Temperature sensor (inlet)	RS PRO PT100 RTD 3 wire sensor -50 C to +200 C
Temperature sensor (outlet)	RS PRO PT100 RTD 3 wire sensor -50 C to +200 C
Load cell	Tedea Huntleigh Model 614 (50 kg range)

3.2 Torque measurement

To measure the torque, the anti-rotation feature keeping the stator in place during normal operation is removed and substituted by a 3D printed bracket connected to the load cell via a pushrod (Figure 5).

To validate the non-invasive torque testing method, staircase pressure ramps with quasi-steady state conditions for each step are used. For a fixed pump inlet pressure and speed, the outlet pressure is varied so that the following data points for differential pressures across the pump are generated: [4, 28, 54, 78, 103, 128, 153, 128, 103, 78, 54, 28, 4] bar. Tested pump speeds are 50, 100 and 150 rad/s. Each pressure step is held for approximately 25 s.

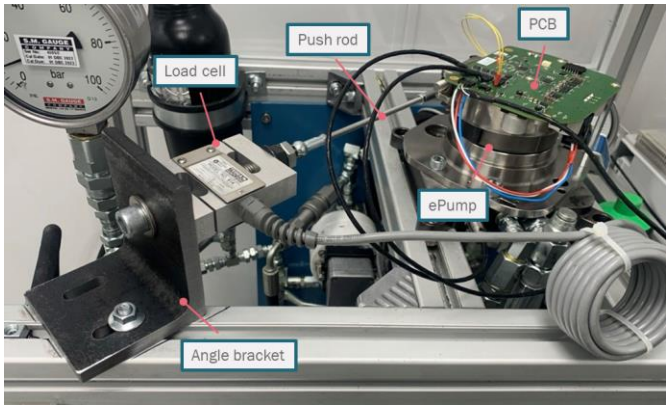


Figure 5: Image of the load cell set-up.

4 Results and discussion

4.1 Torque measurement results

Figure 6 and Figure 7 show the relationship between measured and normalized torque and pump pressure difference for 50 rad/s and 150 rad/s pump speed respectively. For commercial reasons, the torque data is normalized with respect to the maximum measured torque value for each pump speed. The results show the expected linear relationship between pump torque and differential pressure across the pump. A hysteresis effect can be observed between going up the pressure ramp and coming back down again, more distinct for 150 rad/s. The torque is lower on the downwards part of the ramp.

The hysteresis effect was further investigated and compared to the calculated torque based on the simulated torque constant and measured q-axis motor current. The deviation of the "downwards" torque with respect to the torque on the

"upwards" part of the ramp seems to grow with pump speed and is inversely proportional to the load. The motor current based torque shows a very similar trend although more distinctive and with higher magnitude (approximately double percentage deviation) in the torque results. The staircase pressure ramp tests start at 4 bar differential pressure, going up to 153 bar differential pressure and back down, with an oil temperature increase on the rig within approximately +1°C. The temperature profile within the pump is likely to be significantly different, with oil in the key clearances likely to be warmer at the end of the test, leading to lower viscosity and therefore lower torque, which is an effect that would be stronger at higher speeds.

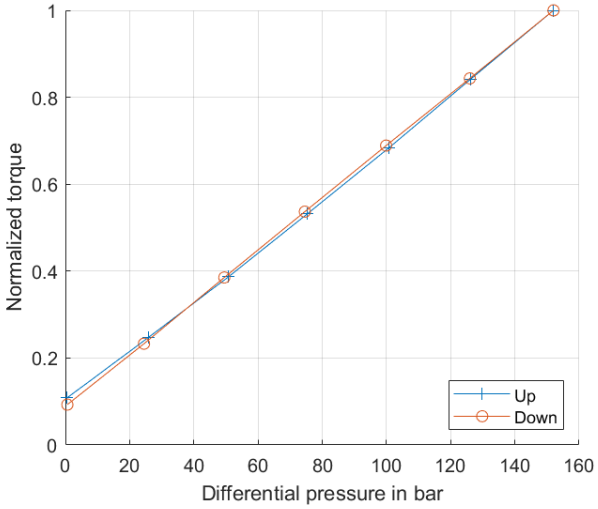


Figure 6: Torque – differential pressure relationship for 50 rad/s.

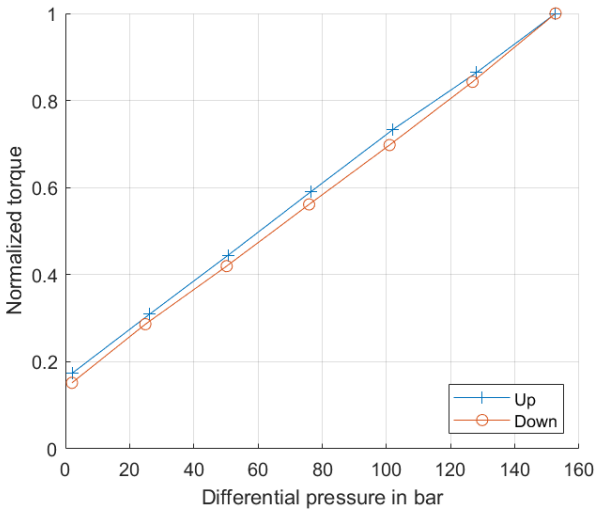


Figure 7: Torque – differential pressure relationship for 150 rad/s.

For comparison of different speeds and against the theoretical pump torque, a linear fit is applied. Figure 8 shows the results. As expected, the absolute torque for a given pressure difference is higher for higher speeds. This is due to a significant portion of the friction in the pump being viscous friction and therefore speed dependent.

It was suspected that friction between the stator inner diameter and pressure tube outer diameter might affect the measurements. The linearity and low hysteresis of the torque results, as well as the high similarity of the gradient in the measured torque vs. ideal torque in Figure 8 show that the impact of friction is not significant. The torque measurement method has yielded repeatable and creditable results.

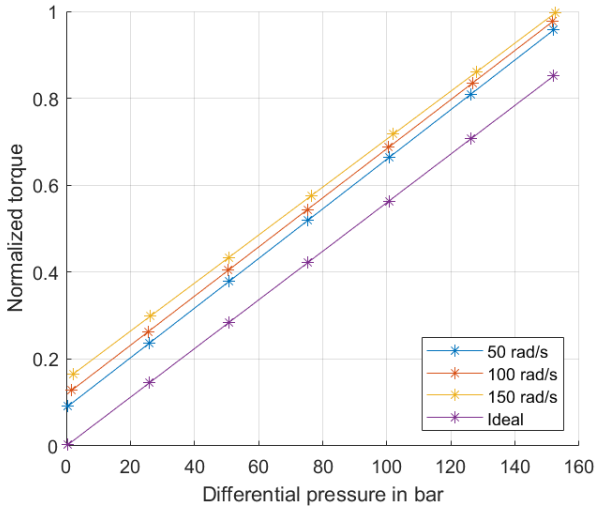


Figure 8: Torque – differential pressure relationship for a set of pump speeds.

4.2 Unit level loss results

The individual pump loss components are expressed as fractions of the total pump power loss in the following equations. The volumetric power loss fraction therefore becomes:

$$P_{L,vol,pump} = \frac{P_{L,vol}}{P_{L,vol} + P_{L,mech}} \tag{13}$$

The mechanical power loss fraction is then expressed as:

$$P_{L,mech,pump} = \frac{P_{L,mech}}{P_{L,vol} + P_{L,mech}} \tag{14}$$

Figure 9 shows the volumetric power loss fraction for three distinct pressure levels. The experimental data points are marked on the figure with an additional best fit curve to indicate trends. As expected, volumetric losses increase with higher differential pressures across the pump, due to increased leakage flow. With increasing speed, the proportion of leakage losses to outlet flow decreases. In combination with increasing viscous friction losses, the volumetric loss fraction is decreasing with increasing speed. For low pump speeds and high pump pressures, the volumetric loss is therefore the dominant loss component in the pump. The mechanical loss fraction (see Figure 10) shows the inverse pressure and speed

dependencies. The mechanical loss fraction decreases with increasing pressure. Higher pressures increase the outlet power whilst increases in friction due to pressure are much less in comparison, therefore making flow losses more distinct. As expected, the mechanical loss fraction increases with speed because of higher viscous friction and increased load on the piston bearing against the cam in the pump. For low differential pressures and high speeds, the mechanical loss is the dominant loss term in the pump.

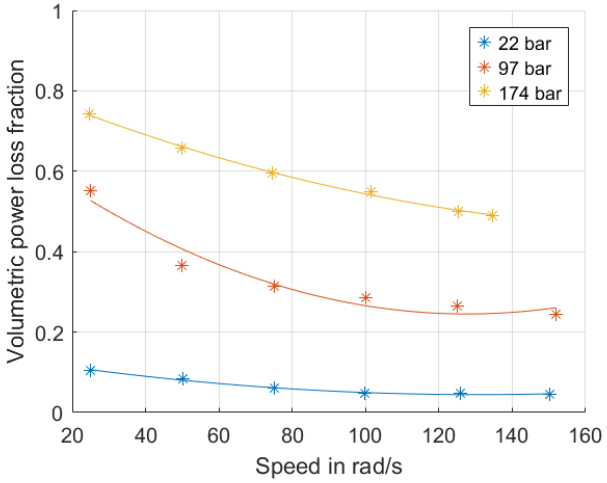


Figure 9: Volumetric loss as a fraction of pump power loss.

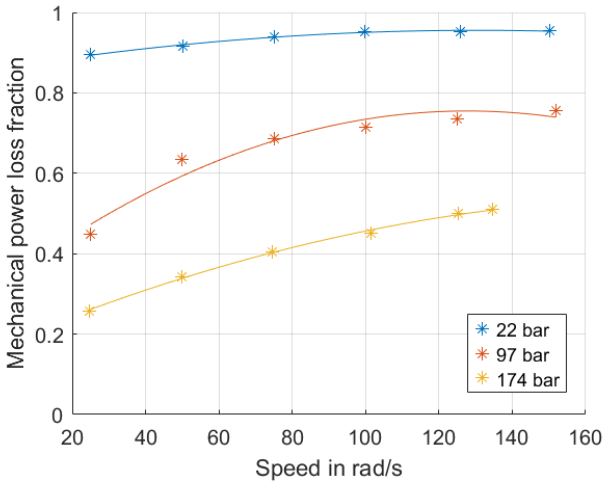


Figure 10: Mechanical loss as a fraction of pump power loss.

To facilitate understanding of loss mechanisms at a unit level, volumetric, mechanical and electrical losses are related to the total unit level losses. The volumetric loss fraction at a unit level is defined as

$$P_{L,vol,unit} = \frac{P_{L,vol}}{P_{L,el} + P_{L,vol} + P_{L.mech}} \quad (15)$$

The mechanical loss fraction at a unit level is then expressed as

$$P_{L,mech,unit} = \frac{P_{L,mech}}{P_{L,el} + P_{L,vol} + P_{L.mech}} \quad (16)$$

And the electrical loss fraction at a unit level finally becomes

$$P_{L,el,unit} = \frac{P_{L,el}}{P_{L,el} + P_{L,vol} + P_{L.mech}} \quad (17)$$

Figure 11 and Figure 12 show the contributions of the separate loss components (volumetric, mechanical, and electrical) to total unit level losses for a differential pressure of 70 bar and 148 bar respectively. The experimental data points are marked on the figure with an additional best fit curve to indicate trends.

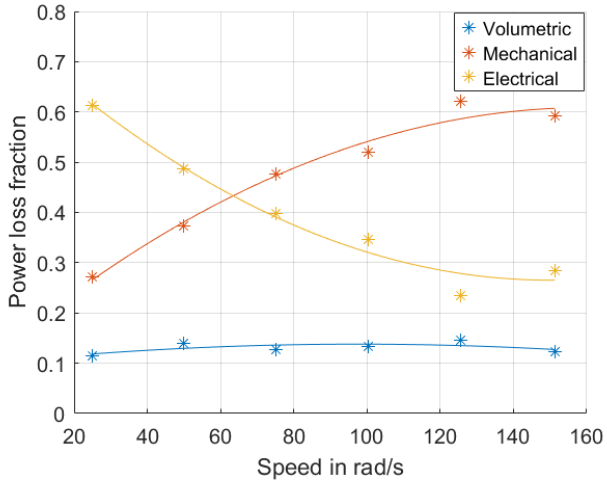


Figure 11: Volumetric, mechanical and electrical power loss as a fraction of total unit loss at 70 bar differential pressure.

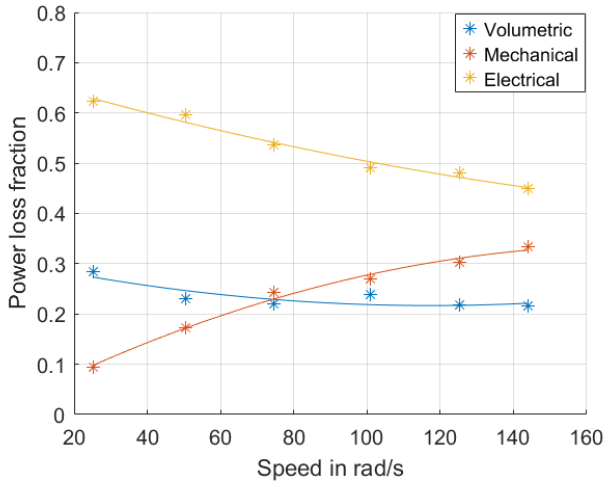


Figure 12: Volumetric, mechanical and electrical power loss as a fraction of total unit loss at 148 bar differential pressure.

For 70 bar differential pressure, the volumetric loss fraction remains near constant at an average of 14%. This is because the speed dependency of the flow loss is low and less distinct at lower pressures. The mechanical loss exhibits a clear speed dependency, with a higher loss fraction of up to 59% at higher speeds. This agrees with overall loss trends and previous observations, resulting from higher friction at higher pump

speeds. Electrical losses are showing an inverse trend to the mechanical losses and are dominant in the lower speed region with a loss fraction above 40% for speeds up to 75 rad/s but decreasing towards approximately 28% for the highest speed.

At 148 bar differential pressure, the volumetric loss fraction average increases to 23%, which is 9% higher than for 70 bar, reflecting the pressure dependence of flow loss. The graph shows a slight downward trend for the higher pressure. This is in line with the previously observed trend within the pump of a higher negative speed dependent gradient for higher differential pressures. The mechanical loss fraction scales with speed and reaches a maximum of approximately 33%. The electrical loss fraction remains consistently above 45% for the full speed envelope, being the dominant loss component at 148 bar differential pressure. This is due to the higher pressure requiring a higher torque at the pump shaft, which subsequently leads to a higher motor current. Power loss in electric motors typically scales quadratically with motor current as resistance loss in the windings.

It is evident from these results that the power loss fractions shift both with speed and pressure. The mechanical loss fraction shows a clear speed dependency whereas the electrical and volumetric losses exhibit a pressure dependence. This is most obvious for the electrical losses that become the dominant loss component for 148 bar differential pressure.

The analysis of power loss fractions, enabled by the load cell set-up to determine pump shaft torque allows dominant loss components in different operating modes to be identified. This provides a guide for a detailed breakdown of losses in highly integrated electric pumps, facilitating improved designs.

4.3 Model validation

To support the performance characterization of the pump, the model of Section 2.3 is compared to experimental data. Figure 13 shows a comparison of calculated and measured volumetric and mechanical efficiency (normalized) for three distinct differential pressures across the pump. For commercial reasons, the efficiency results are normalized with respect to the highest respective efficiency value in the data.

To derive a good fit, the analytical model coefficients (see Table 3) were adjusted until a good agreement between model and test results was found.

Table 3: Model coefficients.

Parameter	Value
Slip coefficient $\boldsymbol{c_s}$	$3.5 \cdot 10^{-8}$
Speed coefficient $\boldsymbol{c_\omega}$	$0.55 \cdot 10^{-8}\ \frac{\mathrm{m}^3}{\mathrm{rad}}$
Viscous friction coefficient $\boldsymbol{c_v}$	$5.3 \cdot 10^{4}$
Coulomb friction coefficient $\boldsymbol{c_f}$	0.14
Constant torque $\boldsymbol{T_c}$	0.065 Nm

The model captures the pressure dependence of the volumetric efficiency reasonably well. This is as expected, considering that the main contribution to flow losses in the model is accounted for by the calculation of a total leakage flow, which is inherently pressure dependent. The added speed coefficient proves beneficial to represent the speed dependence of flow losses. In the originally proposed model by Wilson, and later McCandlish and Dorey, the compressibility term is proportional to speed, but negligible in comparison to the leakage loss. Without the speed coefficient, the model therefore assumes nearly constant flow losses with increasing speed – considering that the power output is rising with speed, this leads to higher efficiencies for higher speeds. This indicates that there are speed dependent loss mechanisms present in the tested pump, which need further investigation. Both the speed and pressure dependence of mechanical losses are captured well by the model, indicating that a good set of coefficients was identified.

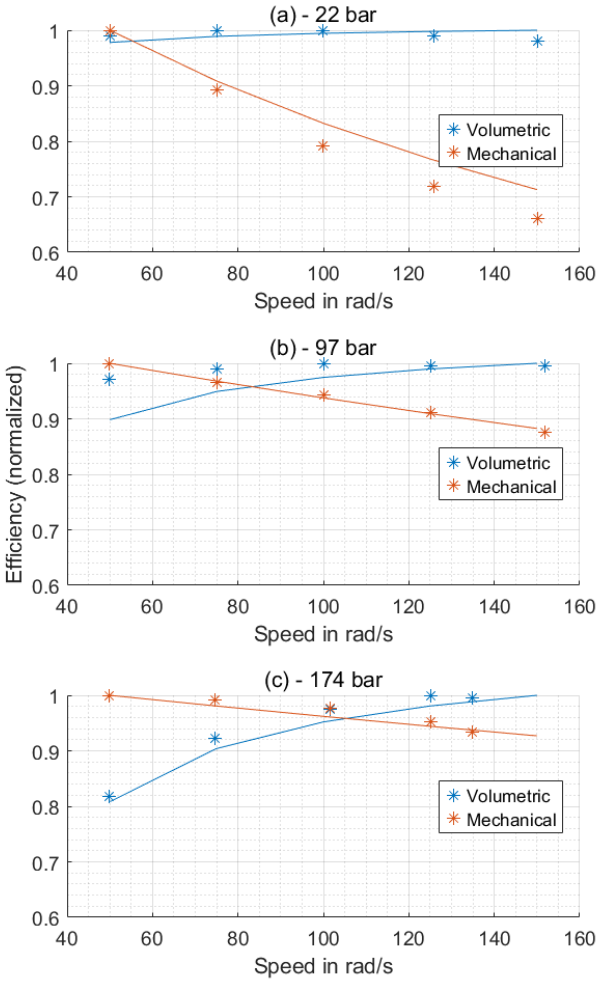


Figure 13: Comparison of experimental (marker) and modelled (line) volumetric and mechanical efficiency (normalized) for (a) 22 bar, (b) 97 bar, and (c) 174 bar differential pressure.

The reasonable agreement between model and test data validates the simplified approach to loss modelling in the pump. The results support common trends in volumetric and mechanical losses within piston type pumps and help identify

regions of high pump operating efficiency. Generally, high pressures are favourable for small mechanical losses, whereas low pressures are preferable for small volumetric losses. Both mechanical and volumetric losses are speed dependent, however this trend is more distinct on the mechanical side due to viscous friction. Overall, the unit under test favours medium speeds and higher pressures as operating conditions as per design intent.

5 Conclusion and outlook

In this study, a novel ePump design and loss analysis framework, including a non-invasive torque measurement approach was presented. This will further enhance the understanding of highly integrated electric pumps. The ePump design is based on a radial piston pump architecture, tightly integrated with a BLDC motor. The design is shaftless, reducing bearing and sealing requirements, and enabling an ultra-compact assembly. The unit not only integrates motor and pump but also features onboard power electronics and closed-loop speed control of the pump with the opportunity of sensing motor temperature for condition monitoring A double lobed cam facilitates 14 piston strokes per revolution, minimizing flow ripple. The loss analysis framework enables the separation of volumetric, mechanical and electrical losses within the unit. Based on this, favourable operating conditions for the respective subsystems can be identified. This will facilitate improved designs for future iterations and allow for design optimizations in terms of size, weight and efficiency.

The proposed torque measurement method was successfully proven. A hysteresis effect was observed in the results with a slight deviation in the torque value between moving up and down a staircase pressure ramp. The measured torque scaled linearly with pressure as expected. Overall, repeatable and credible results could be obtained and enabled the analysis of power loss fractions. This has allowed to identify dominant loss components in different operating modes. The results show that power loss fractions shift both with speed and pressure. The mechanical loss fraction exhibits a clear speed dependency whereas volumetric losses are pressure dependent. The electrical losses demonstrate both pressure and speed dependencies The volumetric loss fraction remains below 30% for all tested operating points. Whilst the mechanical loss fraction reaches up to 59% for the lower pressure setting, mechanical losses stay below 33% for the higher pressure. For the lower pressure, the electrical losses reach up to 61% for the low-speed region but are the overall dominant loss term for the higher pressure operating point with a fraction above 45% for the full speed envelope.

The simple mathematical model was successfully validated and showed reasonable agreement with the experimental data. This validates the simplified approach to loss modelling in the pump. Common piston pump loss trends were observed in the results and high efficiency regions could be identified. Overall, the unit under test favours medium speeds and higher pressures as operating conditions per design intent.

Future work will investigate the hysteresis effect in the torque measurement set up. A more detailed loss model, including physical design parameters such as clearances, piston and cam geometry will be developed and used to improve the understanding of the observed loss mechanisms. Work is ongoing in modelling electric motor losses for a comparison with the measured electrical losses. Additionally, IEC60034-2 will be considered as a framework to break down the electric motor losses. For this study, an early prototype was used as a demonstrator. It is intended to further refine and validate the discussed approach with more units.

Overall, the proposed ePump loss analysis framework was successfully validated with existing hardware and allowed a detailed break-down of losses in the unit under test. This has allowed identification of favourable operating regions and can provide guidance for future design iterations.

References

[1] D. Beltrami, P. Iora, L. Tribioli, and S. Uberti, ‘Electrification of Compact Off-Highway Vehicles—Overview of the Current State of the Art and Trends’, *Energies*, vol. 14, no. 17, p. 5565, Sep. 2021, doi: 10.3390/en14175565.

[2] A. Lajunen, P. Sainio, L. Laurila, J. Pippuri-Mäkeläinen, and K. Tammi, ‘Overview of Powertrain Electrification and Future Scenarios for Non-Road Mobile Machinery’, *Energies*, vol. 11, no. 5, p. 1184, May 2018, doi: 10.3390/en11051184.

[3] D. Li, C. Tan, W. Ge, J. Cui, C. Gu, and X. Chi, ‘Review of Brake-by-Wire System and Control Technology’, *Actuators*, vol. 11, no. 3, p. 80, Mar. 2022, doi: 10.3390/act11030080.

[4] L. Tom, M. Khowja, G. Vakil, and C. Gerada, ‘Commercial Aircraft Electrification—Current State and Future Scope’, *Energies*, vol. 14, no. 24, p. 8381, Dec. 2021, doi: 10.3390/en14248381.

[5] European Commission, ‘COMMUNICATION FROM THE COMMISSION - The European Green Deal’, European Commission, Brussels, Dec. 2019.

[6] S. Ketelsen, D. Padovani, T. Andersen, M. Ebbesen, and L. Schmidt, ‘Classification and Review of Pump-Controlled Differential Cylinder Drives’, *Energies*, vol. 12, no. 7, p. 1293, Apr. 2019, doi: 10.3390/en12071293.

[7] Voith Turbo GmbH & Co KG, ‘Motor/pump hybrid system EPAI for high and medium-pressure applications’. 2012. Accessed: Feb. 23, 2025. [Online]. Available: http://www.hypower.com.hk/pdf/Voith_Turbo/890_e_g1891_en.pdf

[8] W. Wustmann, S. Helduser, U. Schuffenhauer, H. Kuss, and N. Michalke, ‘Fully Integrated Electric-Hydrostatic Drive Based on a Gear Pump And a Switched Reluctance Motor’, in *Proceedings of the Tenth Scandinavian International Conference on Fluid Power, SICFP’07, May 21-23, 2007, Tampere, Finland*, Tampere, Finland, 2007.

[9] F. Zappaterra, A. Vacca, and S. D. Sudhoff, ‘A compact design for an electric driven hydraulic gear machine capable of multiple quadrant operation’, *Mechanism and Machine Theory*, vol. 177, p. 105024, Nov. 2022, doi: 10.1016/j.mechmachtheory.2022.105024.

[10] P. Gamez-Montero *et al.*, ‘GeroMAG: In-House Prototype of an Innovative Sealed, Compact and Non-

Shaft-Driven Gerotor Pump with Magnetically-Driving Outer Rotor', *Energies*, vol. 10, no. 4, p. 435, Mar. 2017, doi: 10.3390/en10040435.

[11] W. Fiebig, D. Ignacy, C. Marek, and K. Hubert, 'A vane pump integrated with an electric motor', in *Proceedings of the 9th International Fluid Power Conference, 9. IFK, March 24-26, 2014, Aachen, Germany*, Aachen, Germany, 2014.

[12] W. Fiebig, P. Cependa, P. Jedraszczyk, and H. Kuczwara, 'Innovative Solution of an Integrated Motor Pump Assembly', in *ASME/BATH 2017 Symposium on Fluid Power and Motion Control*, Sarasota, Forida, USA: American Society of Mechanical Engineers, Oct. 2017, p. V001T01A040. doi: 10.1115/FPMC2017-4277.

[13] M. P. Ciurys and W. Fiebig, 'Experimental Investigation of a Double-Acting Vane Pump with Integrated Electric Drive', *Energies*, vol. 14, no. 18, p. 5949, Sep. 2021, doi: 10.3390/en14185949.

[14] T. Wang and Z. Zhou, 'A Compact Hydrostatic-Driven Electric Generator: Design, Prototype, and Experiment', *IEEE/ASME Trans. Mechatron.*, vol. 21, no. 3, pp. 1612–1619, Jun. 2016, doi: 10.1109/TMECH.2015.2504491.

[15] D. Zhu, Y. Fu, X. Han, and Z. Li, 'Design and experimental verification on characteristics of electro-hydraulic pump', *Mechanical Systems and Signal Processing*, vol. 144, p. 106771, Oct. 2020, doi: 10.1016/j.ymssp.2020.106771.

[16] S. Sarode, L. Shang, A. Vacca, and S. Sudhoff, 'Design methodology for an integrated axial piston-type electrohydraulic unit', *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, p. 09544062231185517, Jul. 2023, doi: 10.1177/09544062231185517.

[17] E. Lester and M. Miles, 'TwinTorq Electric-Hydraulic Motor.E. Lester, M. Miles'. KersTech, Inc Accelerating Vehicle Technology, Beaverton, Oregon. Accessed: Feb. 23, 2025. [Online]. Available: http://kerstech.com/twintorq/

[18] G. R. Bohach, Nishanth, E. Severson, and J. D. Van De Ven, 'Modeling and Optimization Study of a Tightly Integrated Rotary Electric Motor-Hydraulic Pump', in *ASME/BATH 2019 Symposium on Fluid Power and Motion Control*, Longboat Key, Florida, USA: American Society of Mechanical Engineers, Oct. 2019, p. V001T01A010. doi: 10.1115/FPMC2019-1626.

[19] F. Nishanth, G. Bohach, J. V. De Ven, and E. L. Severson, 'Design of a Highly Integrated Electric-Hydraulic Machine for Electrifying Off-Highway Vehicles', in *2019 IEEE Energy Conversion Congress and Exposition (ECCE)*, Baltimore, MD, USA: IEEE, Sep. 2019, pp. 3983–3990. doi: 10.1109/ECCE.2019.8912685.

[20] F. N. U. Nishanth, G. Bohach, M. M. Nahin, J. Van De Ven, and E. L. Severson, 'Design of an Axial Flux Machine With an Integrated Hydraulic Pump for Off-Highway Vehicle Electrification', in *2020 IEEE Energy Conversion Congress and Exposition (ECCE)*, Detroit, MI, USA: IEEE, Oct. 2020, pp. 1772–1779. doi: 10.1109/ECCE44975.2020.9235731.

[21] F. Nishanth, G. Bohach, M. M. Nahin, J. Van De Ven, and E. L. Severson, 'Development of an Integrated Electro-Hydraulic Machine to Electrify Off-highway Vehicles', *IEEE Trans. on Ind. Applicat.*, vol. 58, no. 5, pp. 6163–6174, Sep. 2022, doi: 10.1109/TIA.2022.3189609.

[22] M. M. Nahin, G. R. Bohach, F. N. U. Nishanth, E. L. Severson, and J. D. Van De Ven, 'Dynamic Modeling and Design of a Radial Hydrostatic Piston Pump for Integrated Pump-Motor', in *ASME/BATH 2021 Symposium on Fluid Power and Motion Control*, Virtual, Online: American Society of Mechanical Engineers, Oct. 2021, p. V001T01A028. doi: 10.1115/FPMC2021-68788.

[23] J. Li, X. Meng, P. Dong, and S. Zhao, 'Design and analysis of an axial flux permanent magnet motor for the direct drive radial piston pump', *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 233, no. 19–20, pp. 7077–7088, Oct. 2019, doi: 10.1177/0954406219869970.

[24] A. Plummer, 'Integrated Hydromechatronic Actuators', in *2023 IEEE International Conference on Mechatronics (ICM)*, Loughborough, United Kingdom: IEEE, Mar. 2023, pp. 1–6. doi: 10.1109/ICM54990.2023.10101880.

[25] *BS ISO 4409:2019 - Hydraulic fluid power — Positive displacement pumps, motors and integral transmissions — Methods of testing and presenting basic steady state performance*, 2019.

[26] P. Achten *et al.*, 'Measuring the Losses of Hydrostatic Pumps and Motors: A Critical Review of ISO4409:2007', in *ASME/BATH 2019 Symposium on Fluid Power and Motion Control*, Longboat Key, Florida, USA: American Society of Mechanical Engineers, Oct. 2019, p. V001T01A007. doi: 10.1115/FPMC2019-1615.

[27] C. Schänzle and P. F. Pelz, 'Meaningful and Physically Consistent Efficiency Definition for Positive Displacement Pumps - Continuation of the Critical Review of ISO 4391 and ISO 4409', in *ASME/BATH 2021 Symposium on Fluid Power and Motion Control*, Virtual, Online: American Society of Mechanical Engineers, Oct. 2021, p. V001T01A027. doi: 10.1115/FPMC2021-68739.

[28] *BS ISO 8426:2008 - Hydraulic fluid power — Positive displacement pumps and motors — Determination of derived capacity*, 2008.

[29] W. E. Wilson, 'Rotary-Pump Theory', *Journal of Fluids Engineering*, vol. 68, no. 4, pp. 371–384, May 1946, doi: 10.1115/1.4018484.

[30] D. McCandlish and R. E. Dorey, 'The Mathematical Modelling of Hydrostatic Pumps and Motors', *Proceedings of the Institution of Mechanical Engineers, Part B: Management and engineering manufacture*, vol. 198, no. 3, pp. 165–174, Aug. 1984, doi: 10.1243/PIME_PROC_1984_198_062_02.

A Multi-Domain Thermal Model for Simulating Performance of High-Pressure Gerotor Pumps

Kai Ping Qwah, Ajinkya Pawar, Andrea Vacca, *Matthias Krug, *Michael Gunkel, *Christian Kretzer

Maha Fluid Power Research Center, Purdue University, 1500 Kepner Drive, 47905 Lafayette, IN, USA
E-mail: A. kqwah@purdue.edu, B. pawar10@purdue.edu, C. avacca@purdue.edu

*Thomas Magnete GmbH, Innomotion Park 3, 57562 Herdorf, Germany
E-mail: A. Matthias.Krug@thomas-magnete.com, B. Michael.Gunkel@thomas-magnete.com, C. Christian.Kretzer@thomas-magnete.com

Abstract

Gerotors are widely utilized as positive displacement units in low-to-medium pressure applications, such as automotive engine oil and fuel systems. Latest technological require higher operating pressure, and Gerotors are still a good candidate. Advanced simulation models can accelerate progresses in Gerotors. However, designing a Gerotor pump for high-pressure operation necessitates advanced modeling approaches that account for thermal effects, which are often overlooked in conventional isothermal models. Current methodologies, including 3D CFD, lumped parameter (LP), and coupled LP-lubricating interface models, fail to capture intricate temperature variations that influence frictional losses and leakages.

To address this gap, this paper presents a multi-domain numerical model that integrates thermal effects into the simulation of a gerotor pump operation. Utilizing a fully coupled LP-CFD-FEM framework, this approach is implemented via the proprietary Multics multi-physics simulation tool developed by the authors' research group. The model incorporates mass and energy conservation equations in the fluid domain using the LP methodology. Thermal Reynolds and energy equations are employed to simulate the lubricating interface, capturing thermoelastohydrodynamic (TEHL) effects, while solid components' heat transfer and thermal deformation are evaluated to establish a comprehensive coupling with lubricating interfaces and tooth space volumes (TSVs).

The proposed model provides detailed estimations of operating temperatures across pump's outlet, TSVs, lubricating interfaces, and solid components. Comparative analysis against isothermal models reveals significant differences in predicted frictional losses and leakages, emphasizing the critical role of thermal effects in gerotor pump performance. The results show a 5% rise in total efficiency prediction when thermal effects are considered compared to isothermal simulation. This work shows the importance of thermal considerations in a gerotor design for high pressure applications advancing the state of the art in gerotor pump modeling and simulation.

Keywords: Gerotor, coupled LP-lubricating interface, thermoelastohydrodynamic, frictional losses and leakages

1 Introduction

Gerotor pumps are commonly used in low-pressure fluid power and automotive applications. Their popular usage stems from its variety of benefits including low cost, compactness, robustness, and low noise when in operation. **Figure 1** illustrates a typical assembly for a Gerotor unit, which comprises of an inner gear, outer gear, shaft, casing, and lateral compensation plates.

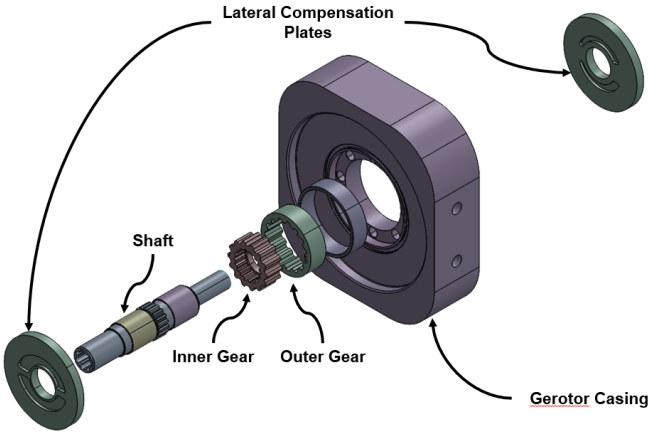


Figure 1: *Typical Gerotor Assembly.*

Figure 2 demonstrates the working of a Gerotor pump. The TSV expands as it connects to the low-pressure suction port, which allows for suction of flow into the TSV. Conversely, the TSV contracts as it connects to the high-pressure delivery port, which allows for delivery of flow from the TSV.

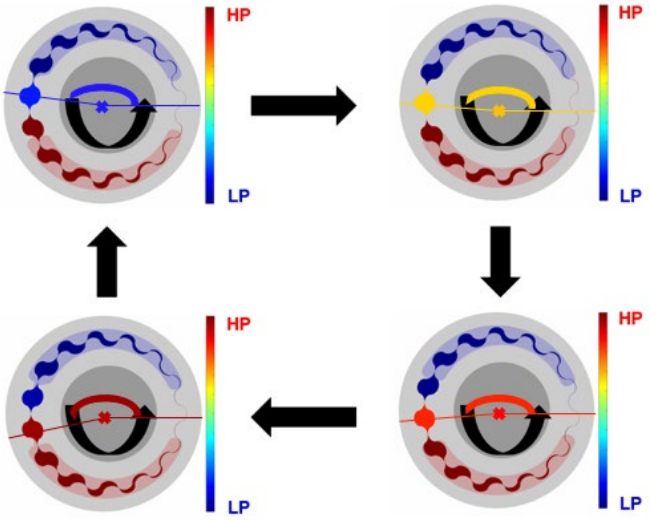


Figure 2: Principle of operation of a Gerotor.

A Gerotor pump has two primary lubricating interfaces as shown in **Figure 3**, which are the radial gap and the lateral gap.

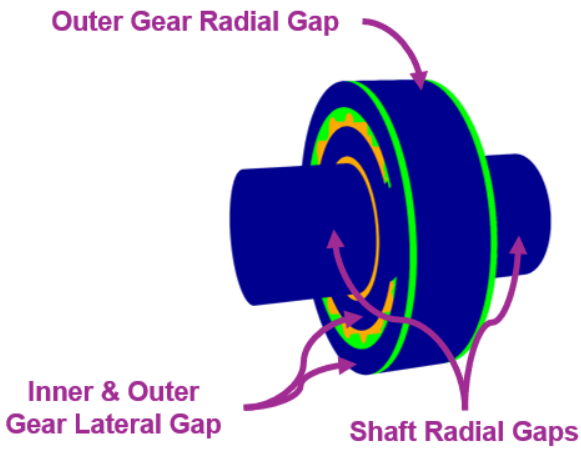


Figure 3: Lubricated interfaces of a Gerotor.

The outer radial gap supports the loads acting on the outer gear by virtue of hydrodynamic pressure developed in the oil. Similarly, the inner shaft radial gap supports the loads acting on the inner gear. The lateral gap is formed at the gear lateral side, and in common gerotor designs it is defined by the housing of the unit.

Gerotors are commonly used in the automotive industry, functioning as engine oil and fuel injection pumps. The pump operates under relatively low-pressure condition (p<30 bar) compared to the typical pressure levels experienced by other positive displacement machines (PDMs) ($p = 200\ bar$ to $300\ bar$). At high pressure applications, gerotors exhibit high inefficiencies due to pressure-driven leakages occurring at the lubricating interfaces as well as the clearances between moving components. Additionally, lubricating interfaces must be designed to support higher loads. Mixed or boundary lubrication can occur at higher pressures resulting in asperity contact which leads to high amounts of frictional shearing and wear on these surfaces.

To address excessive leakages at the lateral gap, a compensation mechanism in the form of thrust plates can be used [1]. Such mechanism is not common for Gerotors, but it is widely used in gear type Positive Displacement Machines [PDMs] such as external or internal gear machines. The lateral plate compensation design is a challenging aspect as the gear side loads can vary with operating speeds, with undercompensated designs showing higher leakages, while over-compensation can lead to higher frictional losses. The linear and angular micro-motion as well as deformation of compensating plates play a critical role in the overall power losses at lateral gap interface especially at high pressures. Recent developments in simulation models address these challenges by considering physical aspects of micro-motion and body deformation while modeling the lubricating interface performance and estimate power losses in pressure compensated designs. Thiagarajan et al. used a coupled LP-tribological model pressure-compensated design in an external gear machine (EGM) pump comprised of lateral bushings to analyze the designs of symmetric [2] as well as an asymmetric axial balancing configuration [3].

Pippes eta al. [4] studied the lubricating interfaces of crescent-type internal gear pump (CIGP) involved in compensation dynamic, which demonstrate the importance of compensation methods in controlling the flow leakages while ensuring an acceptable level of frictional losses. Pan et al. [5] proposed a model-based total efficiency prediction approach for a compensated CIGPs to study the energy losses within the unit and provide a solid foundation for predicting total efficiency for a compensated CIGP unit. However, there are various limitations with the above study, including the isothermal assumption and the neglection of deformation.

Examining the studies of compensated EGP and IGP highlights the importance of developing more accurate models to mirror the performances of actual units and demonstrate more accurately the volumetric efficiencies gained from a pressure-compensated unit in reference to an uncompensated unit.

Notable past efforts have engaged in modelling the fluid dynamic behavior of Gerotor units using LP or CFD approach. The first method to account for clearances between gears and subsequently the effects of fluid leakages as well as fluid compressibility effects was the lumped parameter model developed by Manco et al. [6], where the fluid volume is divided into lumped chambers and flow connections. Gamez et al. [7] proposed a leakage model in a 3-D CFD simulation which establishes a new boundary condition of a virtual wall to allow for simulation of the teeth contact at the interteeth radial clearance. Castilla et al. [8] introduced a novel methodology to achieve more realistic simulation of flow through channels and clearances of a new-born mini Gerotor prototype pump using OPENFOAM. Altare and Rundo [9] presented an extensive study on a reference Gerotor unit using a CFD model through the commercial software PumpLinx (now Simerics) to study the incomplete filling in Gerotor pumps with an accompanying experimental validation to verify the reliability of the CFD code. Bae et al. [10] analyzed the internal flow characteristics in the optimized Gerotor generated using a commercial CFD code, ANSYS-CFX.

Despite the popularity of computational fluid dynamics (CFD) based methods, a major disadvantage of these methods, besides the computational cost arising from the refined discretization of the fluid domain to capture complex geometries of the gerotor teeth, is their inability to reflect contacts and micro-motions. The contact behaviors between the teeth strongly affect their torque efficiency. Radial micromotions of the rotor also play a relevant role in determining the tooth tip sealing ability of the rotors. Pellegri et al. [11, 12, 13] demonstrated a comparison between CFD and lumped parameter (LP) model able to predict rotor micro-motion, showing how the rotor position predicted by the lumped parameter model could allow a good match between the results of all the models compared to experiments. More recently, Mistry et al. [14] developed a fully coupled LP-lubricating interface model that utilizes LP approach within the TSVs, solves Reynolds equation for the lubricating interfaces, and elasto-hydrodynamic lubrication (EHL) for evaluation of contact points between the gears. However, these models were limited by the isothermal consideration of the unit and did not consider any compensation mechanism due to low operating pressures. Due to larger power losses in Gerotors operating efficiently at high operating pressures there will be significant thermal effects on the unit that could affect the load-carrying capacity of the fluid significantly, resulting in smaller film gap heights at the lubricating interfaces and consequently a larger shearing at these interfaces, which ultimately leads to higher frictional losses and leakages uncaptured in an isothermal model.

To address these concerns, there has been an increasingly strong focus towards the development of thermal interaction models to capture thermal effects in positive displacement units. The first non-isothermal fluid model was introduced by Ivantysynova et al. [15] which addressed the variation of fluid properties in the fluid film at the piston/cylinder interface of an axial piston pumps. Since then, multiple studies [16, 17] have been proposed for thermoelastohydrodynamic (TEHL) analysis of various lubricating interfaces in piston pumps. Recently, Mukherjee et al. [18] proposed a comprehensive fully coupled simulation comprising of the thermal lumped-parameter fluid dynamics, thermoelastohydrodynamic lubrication (TEHL) for all lubricating interfaces, as well as the heat transfer, body dynamics, and solid deformation across these interfaces and control volumes (CVs) and illustrated that most power losses associated with the PDM are dominated by the lubricating interfaces. Pawar et al. [19] extended this methodology for the case of an external gear pump simulation, to study increasingly complex isothermal simulation to thermal simulation considering pressure and thermal deformation of gears, bushing and housing and analyzing their overall behavior in terms of EGP performance of a commercial unit. This study revealed the significance of thermal effects in accurate estimation of power losses. Since previous studies focused upon low pressure gerotor applications where power losses and thermal effects are not significant, there has not yet been a fully coupled TEHL model developed for Gerotor units.

This paper proposes a novel fully coupled thermal model for Gerotor under high-pressure applications ($> 50\ bar$). Additionally, the reference unit is pressure compensated by means of a compensation plate which is treated as a floating body which can move and deform, adding a new complexity to the lateral gap interface simulation compared to previous studies on gerotor pump analysis. Considering the state-of-the-art isothermal, strongly coupled model from Mistry, Z. [14] as baseline, this study, extends it to account for the thermal effects on the outer gear journal bearing radial gap and gears lateral gaps, as well as evaluates thermal behaviors of TSVs and the solid components including inner and outer rotor, compensation plate and housing. The analysis of the lateral compensation plate considering body dynamics, pressure and thermal deformation is a novel contribution for the Gerotor pump simulation and will be crucial in a more accurate and reliable evaluation of the lateral film leakage and frictional losses.

The remaining parts of the paper will be comprised of three sections. Section 2 will describe the details behind the methodology behind the fluid-structure interaction (FSI) – thermal coupled modeling, including the overall structure of the model, pertinent fluid dynamics, lubricating interface and thermal-related equations, and a summary of the overall TEHL model. Section 3 describes the simulation operating conditions and provides insights into the comparison between an isothermal and a thermal simulation of a reference Gerotor unit. Section 4 summarizes the results and provides a conclusion to the overall work.

2 Fluid-Structure Interaction (FSI)-Thermal Coupled Modelling Methodology

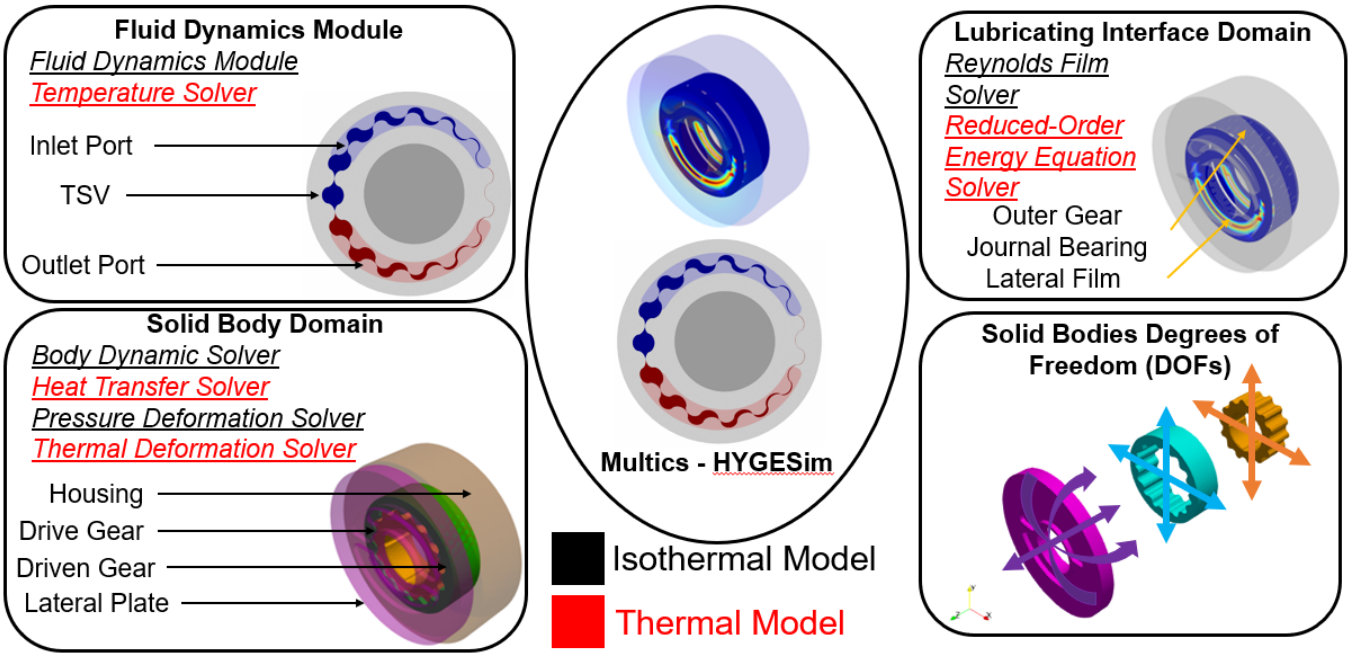


Figure 4: Gerotor Isothermal & Thermal Multi-Domain, Multi-Physics Simulation Model [Source: [14], [19]]

Figure 4 provides an overview of the Gerotor simulation model. The simulation model is comprised of the Preprocessor Module, as well as the multi-domain, multi-physics simulation tool consisting of fluid, solid body and lubricating interface domain solvers. CAD drawings are fed to the Geometric module which provides the input parameters to the multi-domain, multi-physics simulation tool, which includes the lumped parameter solution of the TSV fluid properties in the fluid domain, the evaluation of the behavior of the radial and lateral lubricating interfaces using the Reynolds film solver in the Lubricating Interface Domain, and the evaluation of the lateral compensation plate, inner and outer gear micro-motions through the Body Dynamics Solver as well as the analysis of deformation behaviors of these bodies through pressure deformation solver, Influence Matrix (IM), in the Solid Body Domain. Such tool is like what is presented in [14]. However, to include thermal effects, the tool has been extended to include thermal solvers, which are shown in **Figure 4** highlighted in red. Each domain in this thermal model will additionally include thermal solvers, such as Temperature Solver in the Fluid Domain, Reduced-Order Energy Equation Solver in the Lubricating Interface Domain, and Heat Transfer and Thermal Deformation Solver in the Solid Body Domain. A more detailed description for each solver will be provided in the following subsections.

Thermal simulation of Gerotor involves a fully coupled thermal lumped-parameter simulation of Gerotor. This simulation encompasses a variety of thermal effects on the unit. Firstly, a thermal lumped parameter model of the Gerotor is designed with the consideration of temperature variations during the suction and delivery of fluid within the TSVs due to convection effects at the teeth tips between TSVs and the conduction of heat from the inlet and outlet ports to the TSVs. Additionally, heat convection at the lateral films between compensation plate and Gerotor, and at the journal bearing film between the casing and Gerotor, as well as the heat conduction between these bodies are considered to provide a thermal coupling on the fluid and the solid bodies surrounding it. To evaluate the heat conduction and heat convection of the solid bodies (drive, driven gear, lateral plate, casing), a proper set of thermal boundary conditions are applied to the external surfaces of the bodies.

2.1 Fluid Domain

In the Fluid Domain, fluid dynamics solver is used to evaluate the fluid dynamics behavior by discretizing the pump domain into multiple CVs or TSVs. The choice of an LP model for this module allows for fast computations and permits easy interfacing with the other sub-modules. The layout of the connections can also be represented with the hydraulic schematic shown in **Figure 5**.

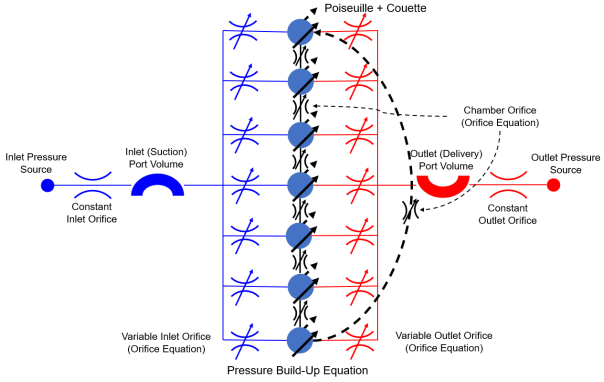


Figure 5: Representation of the control volumes (CV) and of the internal orifice connections.

This LP methodology is extended to a thermal pressure build-up equation Eq. (1) and lumped temperature equation (Eq. 2). These equations are solved for each control volume, or TSV, to compute the pressurization and temperature rise inside the pump. (i in Eq. (1) and Eq. (2) dictates the pressurization and temperature rise of each TSV).

Pressure Build-Up Equation:

$$\frac{\partial p_i}{\partial t} = \frac{K_T}{V}\left[\frac{1}{\rho}\sum \dot{m}_i - \frac{dV}{dt}\right] + K_T \alpha_p \frac{\partial T}{\partial t} \tag{1}$$

$\frac{K_T}{V}\frac{1}{\rho}\sum \dot{m}_i$ represents the time rate of change of fluid pressure due to incoming fluid time rate of volume flow, $-\frac{K_T}{V}\frac{dV}{dt}$ represents the time rate of change of fluid pressure due to outgoing fluid time rate of volume flow, and $K_T \alpha_p \frac{\partial T}{\partial t}$ represents the effects of fluid temperature build-up on the fluid pressure build-up.

Temperature Build-Up Equation:

$$\frac{\partial T_i}{\partial t} = \frac{1}{\rho c_p V}\left[\dot{Q} + \sum \dot{m}_i (h_i - h)\right] + \frac{\alpha_p T}{\rho c_p}\frac{\partial p}{\partial t} \tag{2}$$

$\frac{\dot{Q}}{\rho c_p V}$ represents the time rate of change in fluid temperature due to heat transfer across fluid, $\frac{\sum \dot{m}_i (h_i - h)}{\rho c_p V}$ represents the time rate of change in fluid temperature rise due to changes in fluid enthalpy, and $\frac{\alpha_p T}{\rho c_p}\frac{\partial p}{\partial t}$ represents the effects of fluid pressure build-up on the fluid temperature build-up.

The mass flow entering and leaving each TSVs ($\dot{m}_i$ in Eq. (1) & Eq. (2)) is evaluated using two different approaches: for the connection between each TSV and suction/delivery ports the orifice equation with variable flow coefficient α [15] has been used to account for both laminar and turbulent regimes. The mass flow rate equation is as follows:

$$\dot{m} = \frac{p_i - p_P}{|p_i - p_P|} \cdot \rho|_{p=\overline{p_{i,P}}} \cdot \alpha \cdot A_{i,P} \cdot \sqrt{\frac{2.(p_i - p_P)}{\rho|_{p=\overline{p_{i,P}}}}} \tag{3}$$

Here, the orifice flow coefficient α is a function of Reynolds Number (Re), as described in [20].

Each TSVs are connected via inner-teeth clearances known as tooth tip gap clearances, and these clearances are modelled as orifices with the assumption of laminar flow conditions. This is a reasonable assumption to model real flow conditions due to the microscopic scale in which the orifices are defined in. The modified Couette-Poiseuille equation derived by Rituraj et al. [21] is used to solve for the mass flow rate through the tooth tip gap clearances, which is dependent on geometrical gear properties such as equivalent radius of curvature, R_{eq} and gap height, h_t as well as the Couette flow term:

$$\dot{m} = \rho b \pi h_t \left[\frac{0.0456 b h_t^{2.4836} \Delta p}{\mu R_{eq}^{0.4836}} + \omega_{outer} r_{outer} - \omega_{inner} r_{inner}\right] \tag{4}$$

To solve for the equivalent radius of curvature R_{eq}, Eq. (5) is used:

$$\frac{1}{R_{eq}} = \frac{1}{R_1} + \frac{1}{R_2} \tag{5}$$

2.2 Lubricating Interface Domain

The lubricating films are modelled using the Reynolds Equation Solver. This solver solves for the pressure and density distribution in these film interfaces using the Universal Mixed Thermal Reynolds equation (Eq. 6).

$$\nabla_2 \cdot \left(\phi_P \left(\frac{K_T h^3}{12\mu}\nabla_2 \rho\right)\right) = \nabla_2 \cdot \left(\rho \vec{v}_m \left(\phi_R R_q\right) + \phi_c h\right) + \nabla_2 \cdot \left(\rho \phi_s R_q \left(\frac{v_t - v_b}{2}\right)\right) + \frac{\partial \rho \left(\phi_R R_q + \phi_c h\right)}{\partial t} - \nabla_2 \cdot \left(\phi_p \nabla_2 T \frac{K_T \alpha h^3}{12\mu}\right) \tag{6}$$

where

$\nabla_2 \cdot \left(\phi_P \left(\frac{K_T h^3}{12\mu}\nabla_2 \rho\right)\right)$: Poiseuille Term (Flow driven by pressure gradient)

$\nabla_2 \cdot \left(\rho \vec{v}_m \left(\phi_R R_q\right) + \phi_c h\right)$: Density Wedge Term (Flow due to change in fluid density with respect to position)

$\nabla_2 \cdot \left(\rho \phi_s R_q \left(\frac{v_t - v_b}{2}\right)\right)$: Squeeze Term (Flow due to changes in fluid film thickness with time)

$\frac{\partial \rho (\phi_R R_q + \phi_c h)}{\partial t}$: Local Expansion Term

$-\nabla_2 \cdot \left(\phi_p \nabla_2 T \frac{K_T \alpha h^3}{12\mu}\right)$: Fluid Volumetric Expansion due to Temperature Change

The use of ∇_2 refers to the gradient of variables in a 2-dimensional space (x-y).

2.2.1 Universal Mixed Thermal Reynolds Equation

Eq. (6) was derived initially by Ransegnola et al [22] to solve for both cavitated and non-cavitated regions on a lubricating interface through a single equation, but the equation was limited by simplifying assumption that the model is isothermal. Mukherjee et al. [23] modified this equation to include effects of thermal expansion of lubricant pressure.

An important property for a lubricating interface is the film gap height relative to the roughness profile of the bodies, which is crucial in determining the extent in which the film reaches a point of asperity contact with the bodies surrounding it. Lee and Ren [24] proposed an approach to estimate the contact forces based on the roughness profile of the bodies, which is done by relating the gap height information of the film to the contact pressure in the regime of asperity contact. This mixed lubrication modelling allows for evaluation of viscous and asperity friction, illustrated in Eq. (7) proposed by Manne

et al. [20], resulting in a more accurate evaluation of power losses due to lubricating interfaces.

$$\vec{F}_{friction} = \int_{\Omega_S} \left(\frac{\vec{v}_c - \vec{v}_P}{h} + \frac{h\vec{\nabla}P}{2\mu_f} + \left(\mu_{Asp} p_{Cont} \vec{v}_D\right) \right) dA \quad (7)$$

where

$\int_{\Omega_S} \left(\frac{\vec{v}_c - \vec{v}_P}{h} + \frac{h\vec{\nabla}P}{2\mu_f} \right) dA$: Hydrodynamic Film Shear Force

$\int_{\Omega_S} \mu_{Asp} p_{Cont} \vec{v}_D \, dA$: Asperity Contact Shear Force

Another important attribute to the lubricating interface film gap height is the volumetric losses occurring between the film and its surrounding bodies. If these film gap heights are sufficiently large, a significant amount of volumetric flow losses could be incurred. To account for these losses, mass flow rate at the film boundaries is evaluated using Eq. (8), where s_i represents the boundary that defines the physical connection between a given lubricating interface and a lumped parameter volume.

$$\dot{m}_i = \int_{s_i} \left[-\phi_P \frac{K_T h^3}{12\mu_f} \vec{\nabla}\rho + \rho\vec{v}_m\left(\phi_R R_q + \phi_C h\right) + \rho\phi_S R_{eq} \left(\frac{\vec{v}_l - \vec{v}_b}{2} \right) \right] \cdot \vec{ds} \quad (8)$$

where

$\int_{s_i} -\phi_P \frac{K_T h^3}{12\mu_f} \vec{\nabla}\rho \cdot \vec{ds}$: Poiseuille Term

$\int_{s_i} \rho\vec{v}_m\left(\phi_R R_q + \phi_C h\right) \cdot \vec{ds}$: Couette Term

$\rho\phi_S R_q \left(\frac{\vec{v}_t - \vec{v}_b}{2}\right)$: Squeeze Term

2.2.2 Reduced Order Energy Equation

To determine the temperature distribution of a lubricating interface, a reduced order conservation of energy equation is used as shown in Eq. (9).

$$\frac{\partial\left(\rho c_p T\right)}{\partial t} + \nabla_2 \cdot \left(\rho c_p \vec{v}_f T\right) = \nabla_3 \cdot \left(\lambda_f \nabla_3 T\right) + \phi_D \quad (9)$$

where

$\frac{\partial(\rho c_p T)}{\partial t}$: Time Rate of Change in Heat

$\nabla_2 \cdot \left(\rho c_p \vec{v}_f T\right)$: Heat Convection Term

$\nabla_3 \cdot \left(\lambda_f \nabla_3 T\right)$: Heat Diffusion Term

ϕ_D: Heat Dissipation Term

Velocity in the convection term represents the fluid velocity field resulting from a combination of Couette and Poiseuille effects, as illustrated in Eq. (10).

$$\vec{v}_f = \phi_C \left(\vec{v}_b + (\vec{v}_t - \vec{v}_b)\frac{z}{h} \right) + \frac{\phi_P}{2\mu_f}(z^2 - zh)\vec{\nabla}p \quad (10)$$

$\phi_C \left(\vec{v}_b + (\vec{v}_t - \vec{v}_b)\frac{z}{h} \right)$: Couette Term

$\frac{\phi_P}{2\mu_f}(z^2 - zh)\vec{\nabla}p$: Poiseuille Term

Eq. (9) is converted into 2D Eq. (11) using the approach proposed by Mukherjee et al. [23] based on the Lobatto Point Collocation Method (LPCM) [25, 26] which considers a polynomial approximation of the temperature distribution across the film thickness.

$$\{C\}\frac{\partial\left(\rho c_p \{\hat{T}_k\}\right)}{\partial t} + \{C\}\nabla_2 \cdot \left(\rho c_p \vec{v}_f \{\hat{T}_k\}\right) = \{C\} \cdot \nabla_2 \cdot \left(\lambda_f \nabla_2 \{\hat{T}_k\}\right) + \frac{4\lambda_f}{h^2}\{E\}\{\hat{T}_k\} + \phi_D \quad (11)$$

$\{C\}\frac{\partial(\rho c_p \{\hat{T}_k\})}{\partial t}$: LPCM-Modified Poiseuille Term

$\{C\}\nabla_2 \cdot \left(\rho c_p \vec{v}_f \{\hat{T}_k\}\right)$: LPCM-Modified Heat Convection Term

$\{C\} \cdot \nabla_2 \cdot \left(\lambda_f \nabla_2 \{\hat{T}_k\}\right) + \frac{4\lambda_f}{h^2}\{E\}\{\hat{T}_k\}$: LPCM-Modified Heat Diffusion Term

ϕ_D: LPCM-Modified Heat Dissipation Term

Heat dissipation denoted by ϕ_D arises due to viscous and asperity contact friction losses as shown in Eq. (12), and results from the double inner product of the shear stress tensor $[\tau]$ and the fluid velocity gradient $\nabla \cdot \vec{v}_f$

$$\phi_D = [\tau] : \nabla \cdot \vec{v}_f + \frac{dA}{dV}\left(\mu_{Asp} p_{Cont} |\vec{v}_t - \vec{v}_b|\right) \quad (12)$$

where

$[\tau] : \nabla \cdot \vec{v}_f$: Viscous Dissipation Term

$\frac{dA}{dV}\left(\mu_{Asp} p_{Cont} |\vec{v}_t - \vec{v}_b|\right)$: Friction Dissipation Term

2.3 Solid Body Domain

As is illustrated in **Figure 8**, the solid body domain modelling consists of Body Dynamics Solver, Pressure Deformation Solver, Heat Transfer Solver, and Thermal Deformation Solver.

2.3.1 Rigid Body Domain

Body Dynamics Solver is used to study the allowable degrees of freedom (DOFs) of the floating bodies (i.e. inner gear, outer gear, lateral compensation plate). Linear and angular behavior of the rigid body motions are evaluated using Newton's 2nd Law. Loads acting on the bodies from pressures of all the TSVs and lubricating interfaces as well as the possible contact and frictional forces are considered when evaluating the motion of bodies. In the present work, these loads are evaluated for the inner gear, outer gear and the lateral compensation plate of the Gerotor unit. The position and motion evaluated by the body dynamics solver allows for accurate estimation of the behavior of lubricating interfaces as well as tooth-tip gap heights.

2.3.2 Elastic Pressure Deformation

Pressure Deformation Solver solves pressure deformation using a finite element method (FEM) under reference loading conditions and uses the influence matrix approach (Eq. (13)) to estimate the body deformation and its influence on the change in gap height at lubricating interfaces. The current work involves the influence matrix pressure deformation of the lateral compensation plate of the reference Gerotor unit. The influence matrix approach implements the use of inertia relief methodology to simulate an unconstrained lateral plate in a static analysis using inertial forces, a method that was similarly used by Pelosi, Ivantysynova et al. [16].

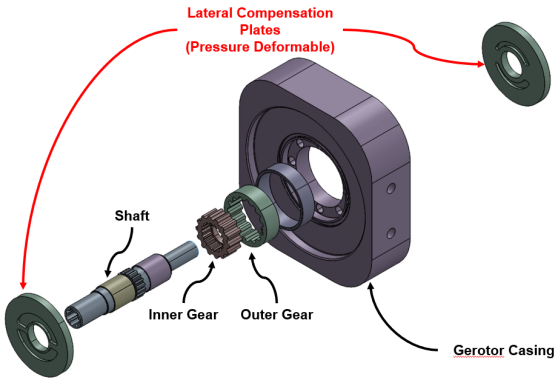


Figure 6: Reference Gerotor unit, highlighting lateral compensation plates as deforming bodies.

$$\Delta h_{eq} = [H]\{p\} \tag{13}$$

In the current work, pressure deformation is evaluated for the lateral compensation plate of the Gerotor to study the effects of deformation on the behavior of the lateral film, as is shown in **Figure 6**..

2.3.3 Heat Transfer

Temperature distribution of solid bodies comprising of the gears, lateral compensation plate, and the casing of the Gerotor are determined by the Heat Transfer Solver. The current work utilizes a 3-dimensional steady state heat transfer equation analysis, as shown in Eq. (15), to estimate the temperatures of all solid bodies in the Solid Body Domain. As discussed by Mukherjee et al. [23], the thermal transience of solid components is much slower than the thermal transience of fluid film lubricating interfaces physically due to the much higher heat capacity of the solid components and therefore modeling it is computationally expensive. To overcome this issue, a steady-state heat transfer equation is used for temperature estimation.

$$\nabla \cdot (\lambda_S \nabla_3 T) = 0 \tag{15}$$

Eq. (15) is a standard heat conduction equation between solid bodies. Eq. (15) is solved using FEM on a linear tetrahedral mesh for characteristic time periods. In the present simulation, the solid bodies of the Gerotor are evaluated for every revolution. This can be perceived as a limitation to Eq. (15), but it is sufficiently accurate in the evaluation of solid body temperatures, since the rate of change in temperatures of solid bodies are much lower than that in the fluid films between these bodies. Solid components temperature is then determined using Eq. (16):

$$T^n = T^{n-1} + \alpha_{relax}(T^n_{sol} - T^{n-1}) \tag{16}$$

where

T^i: Current Temperatures of Solid Bodies

T^{i-1}: Previous Temperatures of Solid Bodies

T^n_{sol}: Current solution of Eq. (15)

α_{relax}: Relaxation Factor ($\alpha_{relax} < 1$).

As the simulation approaches convergence or steady-state solution, temperature difference term, $(T^n_{sol} - T^{n-1})$ in Eq. (16) approaches zero. Relaxation factor α_{relax} is instrumental to couple the steady-state heat transfer equation with other transient system of equations to ensure a smooth transition and avoid sudden temperature gradients in the boundary conditions for reduced-order energy equation (Eq. (11)) in the fluid domain.

2.3.4 Thermal Deformation

Temperature rises occurring due to power losses in a Gerotor unit result in thermal deformation of solid components. In general, solid component deformation, termed thermal warping, can prominently affect the gap height distribution of the lubricating film interfaces in all PDMs, which results in TEHL effects. Thermal deformation is modelled by evaluating the thermal strain using a steady-state solid temperature distribution, evaluated from the steady-state heat transfer equation (Eq. (17))

$$\epsilon_T = \alpha_s \Delta T \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{17}$$

and then relaxing this distribution to find the thermal stress and body deformations, solved using Eqs. (18) to (20).

$$\sigma = 2\mu_L \epsilon + \lambda_L \epsilon_T I \tag{18}$$

$$\nabla_3 \cdot \sigma + f = 0 \tag{19}$$

$$\mu_L = \frac{E}{2(1+\nu)}, \lambda_L = \frac{\nu E}{(1+\nu)(1-2\nu)} \tag{20}$$

3 Results

The reference Gerotor unit is simulated under 2 different operating conditions with the same operating speed of 1000 RPM and operating pressure ranging from an inlet pressure of 0 bar to an outlet pressure of 100 bar. ISOVG-46 is considered as the operating fluid with kinematic viscosity, ν at 40 °C. Nominal dimensions of the unit are considered for the simulations. The two different operating conditions are described as follows. The first case is an isothermal simulation conducted considering pressure deformation of the lateral plate. The second case is a thermal simulation with the same lateral plate considering both temperature and pressure deformations.

The novelty in these results is reflected in both the isothermal and thermal simulations. Isothermal simulation introduces a pressure-compensated lateral plate design under extreme operational conditions. Thermal simulations couples both the

novelty of a lateral compensated Gerotor pump implementation with the thermal simulation of the model.

Figure 7 illustrates the film and mixed boundary conditions for the evaluation of lateral plate and casing temperatures.

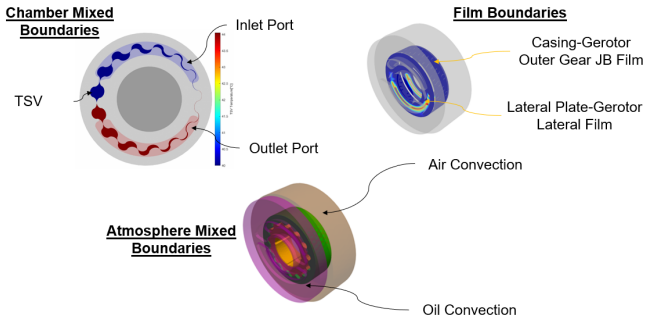


Figure 7: Film and Mixed Boundary Conditions for Lateral Plate and Casing temperature evaluation.

Table 1 shows the important parameters of the reference unit.

Table 1: Parameters of reference Gerotor

Gerotor Parameters	Values
Displacement	3.365 cc/rev
Gear Material	Low-Alloy Steel

3.1 Operational Settings of Isothermal & Thermal Simulation with Lateral Plate Pressure Deformation

Table 2 describes the operating conditions of the isothermal simulation of the unit.

Table 2: Operating Properties and Considerations for Isothermal Simulation

Operating Properties	Values
Shaft Speed	1000 RPM
Inlet Pressure	0 bar
Outlet Pressure	100 bar
Initial, Inlet & Outlet Temperatures	40 °C

Table 3 describes the operating conditions of the thermal simulation of the unit.

Table 3: Operating Properties and Considerations for Thermal Simulation

Operating Properties	Values
Shaft Speed	1000 RPM
Inlet Pressure	0 bar
Outlet Pressure	100 bar
Initial & Inlet Temperatures	40 °C
Outlet Temperature	45.8 °C

In a thermal simulation, TSV temperature follows a much longer transience than TSV pressure, and this is especially notable when thermal simulation is conducted with an outlet temperature set to inlet temperature. To accelerate this thermal transience, a separate thermal pure LP simulation is conducted to evaluate a steady-state outlet port temperature which is then used as the outlet port temperature. The steady-state temperature of 45.8 °C was evaluated after a total computational time of about 50 hours.

Additionally, a comparison of TSV pressures for both isothermal and thermal simulation is shown in **Figure 8**, which illustrates a similar pressure profile but with a slightly steeper pressurization in the thermal simulation, due to the influence of TSV temperature rise as it connects to the outlet port.

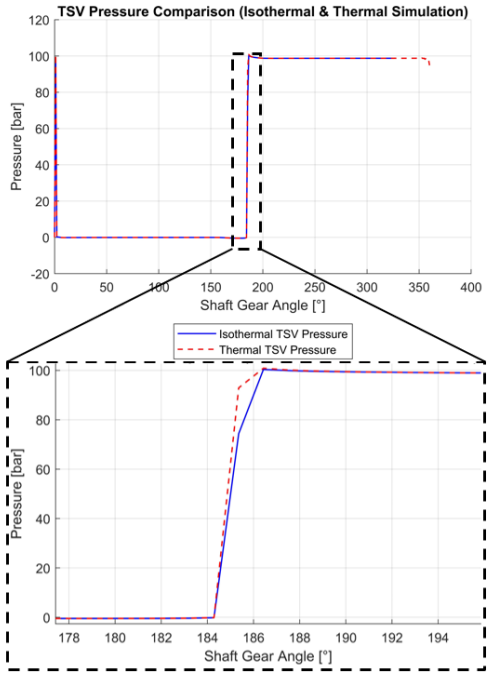


Figure 8: TSV Pressure Comparison for Isothermal & Thermal Simulation (top), highlighting also the steeper pressurization of thermal simulation (bottom).

Setting up thermal simulation requires a sufficiently accurate evaluation of the heat convection coefficient of various surfaces of the solid bodies. Zecchi [27] proposed an equation to approximate the heat flux associated with each region through viscous dissipation, as illustrated in (Eq. 21) below:

$$q \approx \frac{\mu h_{or}\left[\frac{\omega\left(R_{or,out}+R_{or,out}\right)}{2h_{or}}\right]^2}{2} \tag{21}$$

Where μ is the fluid viscosity, ω is the rotational speed, $R_{or_{out}}$ and $R_{or_{in}}$ are respectively the outer and inner radii of the surface and h_{or} is the fluid film thickness the heat flux associated with the region. In general, h_{or} represents the clearance existing in a particular region, such as the clearances within the lateral and outer gear journal bearing films.

Upon evaluating the heat flux, the convection coefficient is determined using the mixed boundary condition equation, shown below also by Zecchi [27]:

$$h \approx \frac{q}{(T_S - T_\infty)} \tag{22}$$

Where h is the heat convection coefficient associated with the heat exchange at the surface, T_S is the surface temperature, and T_∞ is the temperature of the fluid surrounding the surface.

$T_S - T_\infty$ is evaluated as the difference between the surface temperatures (40 °C) denoted by T_S and the fluid temperature (42.9 °C) denoted by T_∞.

Table 4 displays the heat convection coefficients that were considered for all the thermal boundary conditions using Eq. (22).

Table 4: Heat Convection Coefficients of Mixed Boundary Conditions

Mixed Boundary Conditions	Heat Convection Coefficients [W/m²/K]
Tank Drain	348.5
Air	50
TSVs	2175.4
Outer Gear Journal Bearing	935.1
Inner Gear Lateral Film	418.7
Outer Gear Lateral Film	2130.1

3.2 Isothermal & Thermal Simulation with Lateral Plate Pressure Deformation Results

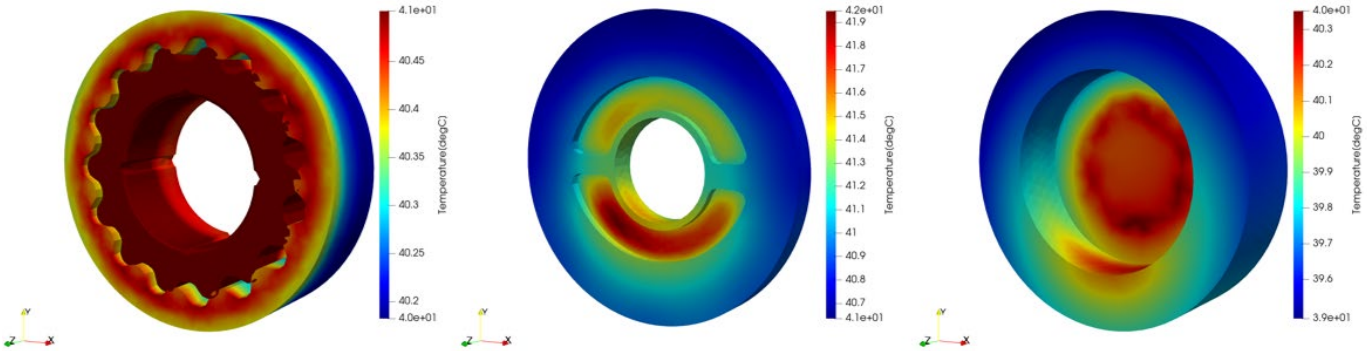


Figure 9: Gerotor (Left), Lateral Plate (Middle) & Casing (Right) Temperature Distribution.

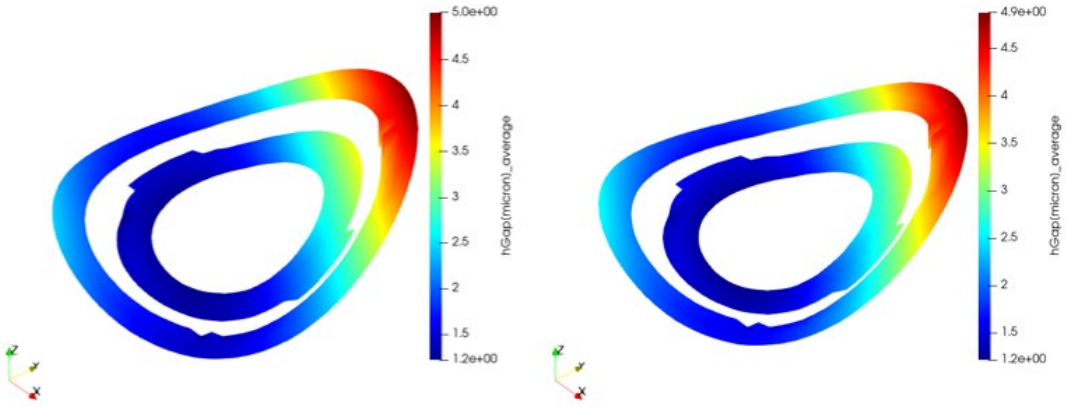


Figure 10: Isothermal (Left) & Thermal Simulation (Right) Time-Averaged Lateral Film Gap Height Comparison

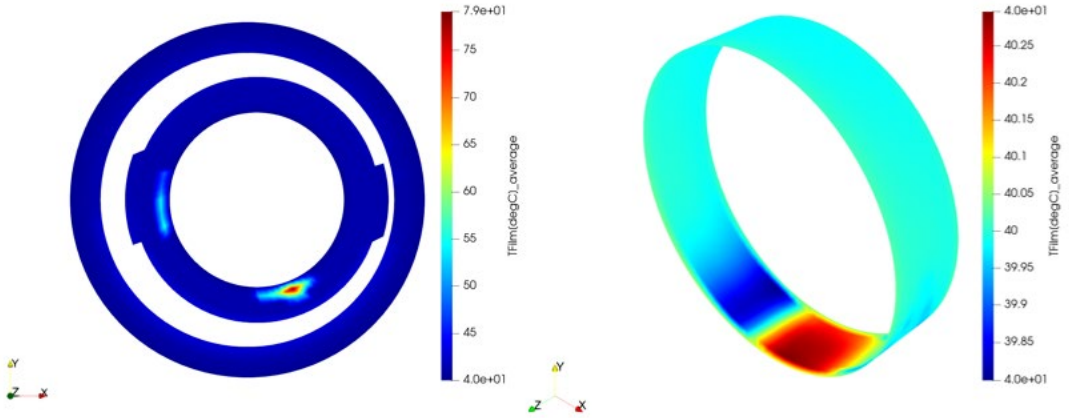


Figure 11: Lateral Film (Left) & Outer Gear Journal Bearing (Right) Time-Averaged Temperature Distribution

As can be observed from **Figure 9** and **Figure 11** heating spots are located where asperity contact occurs, and the temperature rise is much smaller in the solid bodies (40 °C to 42 °C) compared to the lateral film (40 °C to 79 °C).

Table 5 highlights the difference in the efficiencies and power loss percentage in both isothermal and thermal simulation. Power loss is evaluated by using the following equation:

$$\eta_{power\ loss} = \frac{Q_{outlet}\Delta p}{T_{shaft} + T_{inner} + T_{outer}} \cdot 100\ \% \qquad (22)$$

Table 5: Difference in Average Volumetric, Hydromechanical & Total Efficiencies between Isothermal and Thermal Simulations

Average Efficiency Properties	Difference between Isothermal & Thermal Simulation, $\eta_{thermal} - \eta_{isothermal}$
$\boldsymbol{\eta_{vol}}$	+ 3.8 %
$\boldsymbol{\eta_{hm}}$	+ 4.2 %
$\boldsymbol{\eta_{total}}$	+ 5.3 %
$\boldsymbol{P_{loss}}$	- 5.3 %

As is shown in **Table 5**, there is a slight increase in both the volumetric and hydromechanical efficiencies in the thermal model compared to the isothermal model. This can be attributed to difference in the volumetric losses across primarily the lateral film and the tooth tip leakages, as is illustrated in **Table 6**. Additionally, since it is also shown that the difference in the lateral film leakage loss is much higher than the tooth tip leakage loss, we can deduce that the largest contributor in the difference in the volumetric efficiency between an isothermal and thermal simulation is derived from the percentage differences in the lateral film leakage loss.

Table 6: Percentage Difference in Volumetric Loss between Isothermal and Thermal Simulation

Average Volumetric Loss Properties	Percentage Difference between Isothermal & Thermal Simulation, $\frac{Q_{loss,thermal} - Q_{loss,isothermal}}{Q_{loss,isothermal}}$
Lateral Film Leakage Loss	- 51.5 %
Tooth Tip Leakage Loss	- 1.1 %

Additionally, **Figure 10** shows that the overall gap height distribution shows a slightly larger region of more uniformly smaller gap heights from the thermal simulation. This is due to the temperature distribution of the film, as illustrated in **Figure 11**, which shows concentrated spots of high temperature on the film resulting in a reduction of fluid viscosity and consequently the load carrying capacity of the lateral film.

Additionally, **Table 7** demonstrates the much longer computational time required by thermal simulation compared to an isothermal simulation, which creates a trade-off between the higher degree of model accuracy offered and the more expensive computation required by thermal simulation.

Table 7: Average Computational Time for Isothermal and Thermal Simulation

Cases	Simulation Time
Isothermal (Lateral Plate Deformation)	3-4 days
Thermal (Lateral Plate Deformation + Thermal)	7-9 days

4 Conclusion

Current study serves to present a comparative analysis between an isothermal and thermal simulation model of a Gerotor unit, and to also emphasize the effects of deformation

of the lateral pressure compensation plate as well as the effects of thermal considerations of all solid body domains on the pump fluid dynamics, frictional losses, leakages, and efficiency behaviors. A set of isothermal and thermal cases were considered and compared to investigate the significance of thermal effects on a fully coupled, lumped parameter Gerotor model with CFD-defined radial and lateral lubricating interfaces.

Comparative analysis of these models shows that the average volumetric efficiency is higher in the thermal model, but the overall hydromechanical efficiency is much lower in the same model when compared to an isothermal model. As such, the total efficiency and the corresponding power loss in the unit for a thermal and isothermal model are similar, indicating that, while the overall efficiency of the Gerotor unit can be approximated with an isothermal model, the individual frictional and leakage losses are sufficiently disparate to warrant the use of a thermal model to more accurately account for volumetric and frictional losses.

For future work, the thermal model will be used for design optimization purposes to maximize the total efficiency of the Gerotor unit.

Nomenclature

Designation	Denotation	Unit
	Symbols	
α_{relax}	Temperature Relaxation Factor	K^{-1}
α_p	Isobaric Expansion Coefficient	K^{-1}
α	Mass Flow Discharge Coefficient	$-$
Δh_{eq}	Deformation Film Height	m
$[H]$	Influence Matrix Operator	$m \cdot Pa^{-1}$
p	Pressure	Pa
T	Temperature	K
K_T	Bulk Modulus	Pa
$\dot{m}_i$	Mass Flow Rate	kg/s
c_p	Specific Heat Capacity	$JK^{-1}kg^{-1}$
ρ	Density	kg/m^3
h	Fluid Film Thickness	m
h_t	Tooth Tip Gap	m
b	Face Width	m
ω	Angular Velocity	$rad \cdot s^{-1}$
r	Radius	m
μ	Fluid Viscosity	$Pa \cdot s$
R_{eq}	Root Mean Square Roughness	m
ϕ_P	Pressure Flow Factor	$-$
ϕ_c	Contact Flow Factor	$-$
ϕ_s	Shear Flow Factor	$-$
ϕ_R	Roughness Flow Factor	$-$
ϕ_D	Energy Dissipation	J/m^3
μ_{Asp}	Asperity Friction Coefficient	$-$
p_{Cont}	Contact Pressure	Pa
$\vec{v}_m$	Mean Surface Velocity	m/s
$\vec{v}_t$	Top Surface Velocity	m/s
$\vec{v}_b$	Bottom Surface Velocity	m/s
$\vec{v}_D$	Relative Surface Velocity	m/s
λ_f	Fluid Thermal Conductivity	$W/m/K$
λ_S	Solid Thermal Conductivity	$W/m/K$
ϵ_T	Thermal Strain	$-$
ϵ	Strain	$-$
f	Constraint Force per Unit Volume	N/m^3
μ_L, λ_L	Lame's Parameters	Pa
ν	Poisson's Ratio	$-$
E	Elastic Young's Modulus	Pa
Ω_S	Surface Area of Influence	m^2
$[C]$	Legendre Polynomial Matrix	$-$
Q_{outlet}	Actual Volume Flow Rate	m^3/s
Δp	Difference between Inlet & Outlet Port Pressure	Pa
T_{shaft}	Shaft Torque	Nm
T_{inner}	Inner Gear Torque	Nm
T_{outer}	Outer Gear Torque	Nm
$Q_{loss,thermal}$	Isothermal Simulation Volumetric Flow Loss	m^3/s
$Q_{loss,isothermal}$	Thermal Simulation Volumetric Flow Loss	m^3/s
η_{vol}	Volumetric Efficiency	%
η_{hm}	Hydromechanical Efficiency	%
η_{total}	Total Efficiency	%
P_{loss}	Percentage of Power Loss	%

Abbreviations	
LP	Lumped Parameter
TSVs	Tooth Space Volumes
PDMs	Positive Displacement Machines
EGM	External Gear Machine
CIGP	Crescent-type Internal Gear Pumps
CFD	Computational Fluid Dynamics
EHL	Elasto-Hydrodynamic Lubrication
TEHL	Thermo-Elastohydrodynamic Lubrication
CVs	Control Volumes
FSI	Fluid Structure Interaction
IM	Influence Matrix
Re	Reynolds Number
LPCM	Lobatto Point Collocation Method
DOFs	Degree of Freedom
FEM	Finite Element Method

References

[1] Kammerl, L., Dautry, E., & Kretzer, C. (2025, February 20). Gerotor Pump And Motor/Pump Unit.

[2] Thiagarajan, D., & Vacca, A. (2014), "A Numerical Procedure to Design the Optimal Axial Balance of Pressure Compensated Gear Machines," 8th FPNI Ph. D Symposium on Fluid Power, Lappeenranta, Finland.

[3] Thiagarajan, D., Dhar, S., & Vacca, A. (2015), "A novel fluid structure interaction-EHD model and optimization procedure for an asymmetrical axially balanced external gear machine," Tribology Transactions, 58(2), 274-287.

[4] R. Pippes, C.Mayer., (2002), "Gap-compensated high-pressure internal gear pumps", 3rd International Fluid Power Conference, Volume 2, pp.203. IFK3, 2002, Aachen, Germany.

[5] Pan, D., Vacca, A., (2023), "Total efficiency prediction of crescent-type internal gear pump considering floating balancing components", ASME/BATH 2023 Symposium on Fluid Power and Motion Control.

[6] Manco, S., Nervegna, N., Rundo, M., Armenio, G., Pachetti, C., and Trichilo, R., 1998, "Gerotor Lubricating Oil Pump for IC Engines," SAE Trans., 107(3), pp. 2267–2283.

[7] Gamez-Montero, P. J., Castilla, R., Del Campo, D., Ert€urk, N., Raush, G., Codina, E., (2012), "Influence of the Interteeth Clearances on the Flow Ripple in a Gerotor Pump for Engine Lubrication," Proc. Inst. Mech. Eng., Part D, 226(7), pp. 930–942.

[8] Castilla, R., Gamez-Montero, P. J., Raush, G., Codina, E., (2017), "Method for Fluid Flow Simulation of a Gerotor Pump Using OpenFOAM," ASME J. Fluids Eng., 139(11), p. 111101.

[9] Altare, G., Rundo, M., (2016), "Computational Fluid Dynamics Analysis of Gerotor Lubricating Pumps at High-Speed: Geometric Features Influencing the Filling Capability," ASME J. Fluids Eng., 138(11), p. 111101.

[10] Bae, J. H., Kwak, H. S., San, S., Kim, C., 2016, "Design and CFD Analysis of Gerotor with Multiple Profiles (Ellipse–Involute–Ellipse Type and 3-Ellipses Type) Using Rotation and Translation Algorithm," Proc. Inst. Mech. Eng., Part C, 230(5), pp. 804–823.

[11] Pellegri, M., Vacca, A., Devendran, R. S., Dautry, E., and Ginsberg, B., 2016, "A Lumped Parameter Approach for GEROTOR Pumps: Model Formulation and Experimental Validation," Tenth International Fluid Power Conference, Dresden, Germany, Mar. 8–10.

[12] Pellegri, M., and Vacca, A., 2017, "Numerical Simulation of Gerotor Pumps Considering Rotor Micro-Motions," Meccanica, 52(8), pp. 1851–1870.

[13] Pellegri, M., (2018), "A Study on the Operation of Gerotor Type Units Considering Fluid Structure and Mechanical Interaction Effects", (dissertation).

[14] Mistry, Z., Manne, V., Vacca, A., Dautry, E., Petzold, M., 2020, "A Numerical Model for the Evaluation of Gerotor Torque Considering Multiple Contact Points and Fluid-Structure Interactions," Volume 1 – Symposium, 409–418.

[15] M. Ivantysynova (at that time Berge), "An investigation of viscous flow in lubricating gaps. (In Slovak)," SVST Bratislava, Czechoslovakia, 1983.

[16] Rituraj, R., Vacca, A., & Morselli, M. A. (2020), "Thermal modelling of external gear machines and experimental validation.", Energies, 13(11), 2920.

[17] Mukherjee, S., Shang, L., Vacca, A., "A Fast Approach for Coupled Fluid-Thermal Modelling of the Lubricating Interfaces of Axial Piston Machines," Volume 1 - Symposium, 327–340.

[18] Mukherjee, S., Shang, L., Vacca, A., "Numerical analysis and experimental validation of the coupled thermal effects in swashplate type axial piston machines," Mechanical Systems and Signal Processing, 220, 111673.

[19] Pawar, A., Vacca, A., Rigosi M., (2024), "Comparative Analysis of External Gear Machine Performance Considering Deformation and Thermal Effects".

[20] Manne, V. H., Vacca, A., & Singh, K. (2023). A curve-fit traction coefficient relation of mixed EHL line contact

with hydraulic fluid and steel surfaces. *Tribology Transactions*, *66*(2), 364–380. https://doi.org/10.1080/10402004.2022.2163947

[21] Rituraj F, Vacca A., (2019), "Modelling and Validation of Tooth Tip Leakages in Gerotor Pumps," Proceedings of the ASME-JSME-KSME 2019 Joint Fluids Engineering Conference AJKFLUIDS

[22] Ransegnola, T., Zappaterra, F., Vacca, A., (2022), "A Strongly Coupled Simulation Model for External Gear Machines Considering Fluid-Structure Induced Cavitation and Mixed Lubrication," Applied Mathematical Modelling, 104, 721–749.

[23] Mukherjee, S., Vacca, A., Shang, L., Sharma, A., (2023), "A thermal modeling approach for the piston/cylinder interface of an axial piston machine under asperity contact," Meccanica 58 (10) 1929–1957.

[24] Lee, S. C., Ren, N., (1996) "Behavior of Elastic-Plastic rough surface contacts as affected by surface topography, load, and material hardness.", Tribology Transactions, 39(1), 67–74.

[25] Elrod H.G., Brewe E., (1991) "Efficient numerical method for computation of the thermohydrodynamics of laminar lubricating films", Tribol. 113 (3) 506–511.

[26] Zhang S., Hassini M.A., Arghir M., (2018) "Accuracy and grid convergence of the numerical solution of the energy equation in fluid film lubrication: Application to the 1D slider", Lubricants 6 (4).

[27] Zecchi, M. (2013). (dissertation). A novel fluid structure interaction and thermal model to predict the cylinder block.

Energy Efficient Semi-Automatic Control of Loader Cranes using Reinforcement Learning

Amy Rankka[1,2] and Alessandro Dell'Amico[1]

[1]Division of Fluid and Mechatronic Systems, Linköping University, Linköping, Sweden
[2]Hiab AB, Hudiksvall, Sweden
[1]E-mail: amy.rankka@liu.se, alessandro.dellamico@liu.se

Abstract

Operator assistance functions for working machines such as loader cranes are often developed with the aim of increasing the productivity of an unskilled operator or of increasing the safety of operating the machine. For a kinematically redundant system, such as the boom system of a loader crane where three joints are used to control a motion in a two dimensional plane, the operator assistance functions can also be developed with an energy efficiency objective. In this paper, the semi-automatic function Crane Tip Control (CTC) for loader cranes is considered.
The energy required to perform a CTC motion depends on the pressure levels and flow demand of each active cylinder and how this is translated into pump pressure and flow. Previous research has shown that the current CTC strategy is not optimal with regards to energy efficiency. In this work, CTC controllers for limited working areas are developed where the strategy is decided by a neural network trained by a Reinforcement Learning approach with an energy minimizing objective on a simulation model of the crane.
A dynamic programming (DP) algorithm that can find near-optimal solutions for given trajectories offline has been developed in previous research. The energy performance of the new controllers with neural networks, NN controllers for short, are benchmarked on a number of test cases by comparing the solutions to those of the DP algorithm, denoted DP controllers. The results show that energy consumption of the new controllers are only up to 2.2 % higher than the consumption of the DP controllers. The methodology is thus promising, but future work is required to expand it to cover the complete working envelope of the CTC function.
Keywords: loader cranes, energy efficiency, semi-autonomous operation, neural network controller, redundancy resolution

1 Introduction

This paper is part of a research project aiming at reducing the energy consumption of loader cranes through automation. The loader crane of consideration here is a kinematically redundant system where in order to follow a given trajectory in its three dimensional working area there can be an infinite amount of solutions to how to actuate the different cylinders. The redundancy makes it possible to optimize the crane movements with regards to different objectives. As a first step of the project, the semi-automatic function Crane Tip Control (CTC) is investigated. With CTC, the operator controls the movement of the crane tip in a two dimensional plane with one lever for horizontal and one for vertical movements.

In a previous study [1], a dynamic programming (DP) offline optimization was developed with the objective of minimizing the energy consumption of CTC operation on given trajectories. The results show that the energy consumption of the crane on a specified drive cycle can be significantly reduced by optimizing the CTC strategy for energy efficiency. The results also show that for some trajectories, the energy optimized CTC outperforms manual operation, increasing the incentive of investigating semi-automatic, or fully automatic, control as a general solution for reducing the energy consumption of crane operation.

The difference in energy consumption between different strategies of moving the cylinders of the crane are mainly dye to three properties of the system; the total flow that needs to be supplied to the cylinders (even though the crane tip distance is the same, the total traveled distance of the cylinders can differ), the load setting pressure (due to the geometry, cylinder dimensions and friction, different cylinder require different pressure levels for lifting the payload in different crane poses) and the losses from simultaneous operation (in a load sensing system with one pump, the flow to all cylinders active

simultaneously is supplied at the same pressure even though the cylinder pressures might differ).

1.1 Problem Formulation

For a CTC controller to be of practical use it must be able to calculate a solution to any given position and speed reference in real-time. The aim of this study is to investigate whether a neural network controller trained with reinforcement learning could be part of a good solution for a real-time CTC.

Previously, a neural network CTC controller has successfully been developed with the objective of moving the crane tip as fast as possible, see [2]. The controller has been evaluated on a real crane and can operate within the limit of the normal sample rate of the embedded control system.

1.2 System Description

This study is applied to a loader crane, see fig. (1) for a schematic view, with three hydraulic cylinders used for moving the crane tip in the two-dimensional plane with axes x and y; the first boom, "1b", cylinder, the second boom, "2b", cylinder and the extension, "ext", cylinder. The crane also has a cylinder for rotating around the base, but it is not considered here due to that CTC only operates in the plane. The first and second boom cylinders control a revolute joints with angles denoted α_{1b} and α_{2b} respectively. The extension cylinder works as a prismatic joint with a position denoted l_{ext}. The position of the three joints together with the dimensions of the crane defines the crane pose. The crane tip position is given in Cartesian coordinates in the plane ($[x_{tip}$,$y_{tip}]$).

The cylinders of the crane are controlled by a pressure compensated load sensing hydraulic system connected to a variable displacement pump. The speed commands to each cylinder, denoted u_{1b}, u_{2b} and u_{ext}, can be directly translated to spool position commands to the directional valve.

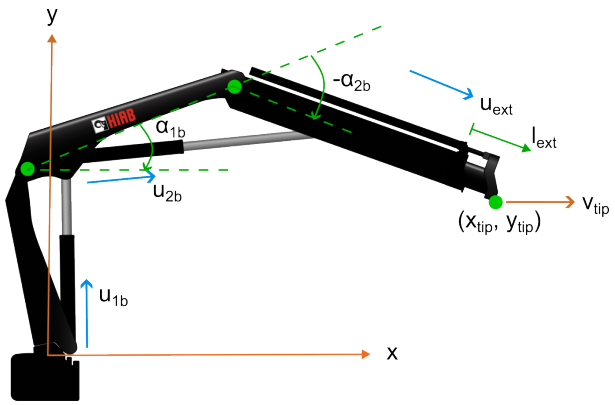


Figure 1: Crane with cylinders, joints and crane tip position.

1.3 Limitations

For a complete CTC solution, it should be possible to control the motion of the crane within its complete working area (which is limited by the stroke of the cylinders) in real-time. The controller should be able to take in any reference crane tip velocity between its limits, and provide a solution for any given state of the crane in real-time for every payload up to the limit of how much the crane can lift. The state space to sample training starting points will thus need to cover all the crane tip positions of the working area 2D plane, the range of cylinder positions that result in these crane tip positions, the range of velocities and the range of payloads.

In this study, which should be seen as an initial proof of concept, the scope is limited to cover one payload, one velocity and one trajectory at a time.

1.4 Prior Art

Solutions that enables control of the end effector in Cartesian coordinates are offered by different manufacturers for their machines. For this study, the Hiab CTC feature for loader cranes, [3], is used as a reference. The strategy for this CTC controller is to maximize the lifting capacity rather than minimizing energy consumption.

A common way of solving the inverse kinematics of the redundant crane boom system is the pseudo-inverse jacobian method, described for example in [4]. The drawback with this approach is that only locally optimal solutions can be found, and in [5] it is shown that the performance is significantly worse than for a dynamic programming algorithm that can find a near-optimal solution offline.

In [6], a real-time controller with an energy minimization objective for a three degree of freedom robotic arm is developed that performs pretty close to the global optimal solutions on trajectories with a length of 10 s. The method is called Predictive Minimization of Kinetic Energy. The motion of the arm is controlled by two tasks, precision tracking by the pseudo-inverse solution and minimization of the kinetic energy by a velocity vector pointing in the direction where the kinetic energy integral is minimized for the chosen trajectory. The optimal control is predicted for a window and continuously updated as the simulation is rolled out. To only minimize the kinetic energy is too much a simplification for the crane with a load sensing hydraulic system, where a significant part of the losses comes simultaneous operation of cylinder with different pressure levels.

As already mentioned, a neural network CTC controller with the objective of moving the crane tip as fast as possible has been developed by reinforcement learning in [2]. The method developed in that work has been applied to the crane used as application in this study with good results. The initial idea was to complement the training of that network by adding an energy minimization term to the objective function. However, the result is not really the desired outcome. When comparing controllers trained with that objective function with different weights on the energy consumption, it became clear that the more the controller cared about the energy consumption, the slower it became.

The issue of lower energy consumption correlating to lower speed can also be in the results of the work in [7], where automatic log collecting of a forwarder is controlled by a neural network trained with reinforcement learning. Compared to a policy trained without an energy objective, the policies trained

with an energy objective show 61 to 78 % decrease in energy consumption and 11 to 36 % increase of cycle time.

In order to avoid this balancing between energy consumption and productivity, in this paper the network is only trained to provide the strategy for CTC, the velocity is controlled by a separate controller.

2 Controller Development

In this work, an energy efficient CTC controller is developed in two steps. First, a velocity controller is implemented that calculates the reference velocities for the three boom cylinders required to follow a certain reference crane tip velocity for a given strategy. Then, a neural network is set up and trained as a strategy controller using reinforcement learning in a simulation environment. When later used in comparison with the solutions from the dynamic programming algorithm, a complete new CTC controller (strategy and velocity controllers combined) will be denoted "NN controller" for short, and the solutions from the DP algorithm "DP controller".

2.1 Controller Structure

The CTC controller is divided into two parts, a neural network that decides the strategy, denoted "CTC strategy controller", and a conventional feed forward controller that handles the velocity reference following given a certain strategy, denoted "CTC velocity controller".

For a solution deployed on a real crane, the reference velocity is given by the operator. Due to the geometry of the crane, the maximum velocity that the crane tip can achieve varies varies with the crane pose. In order not to command the controller to perform an impossible movement, the reference velocity is limited before it is set as input to the velocity controller. The highest possible crane tip velocity for the current state can be calculated according to eq. (1), where J is the jacobian matrix of the crane tip position evaluated at the current crane pose.

$$v_{tip,max} = J(\alpha_{1b}, \alpha_{2b}, l_{ext}) * \begin{pmatrix} \dot{\alpha}_{1bmax} \\ \dot{\alpha}_{2bmax} \\ \dot{l}_{extmax} \end{pmatrix} \tag{1}$$

With the reference velocity given, the strategy controller only needs to decide how one of the cylinders should be operated in order for the resulting movement to be as good as possible with regards to some objective. The operation of the other two cylinders is then uniquely determined by the reference velocity. In this work, the strategy determines the reference to the extension cylinder with the objective of minimizing the energy consumption.

The network for the strategy controller, which is displayed in a schematic view in fig. (2), has an input layer with LeakyReLU activation, two hidden layers also with LeakyRelu activation and an output layer with hardtanh activation in order to get the output between [-1,1], where -1 corresponds to maximum negative speed command, u_{ext}, to the cylinder and 1 to maximum positive speed command.

LeakyReLU stands for leaky rectified linear unitand introduces non-linearity to the network with reduced risk of deactivating neurons compared to simple ReLU, it is defined according to eq. (2), where $alpha$ is set to 0.01. HardTanh stands for hard hyperbolic tangent, is faster to compute than Tanh since it does not have exponentionas, and is defined according to eq. (3).

$$LeakyReLU(x) = \begin{cases} x, & \text{if } x > 0 \\ \alpha x, & \text{if } x \leq 0 \end{cases} \tag{2}$$

$$HardTanh(x) = \begin{cases} 1, & \text{if } x > 1 \\ x, & \text{if } -1 \leq x \leq 1 \\ -1, & \text{if } x < -1 \end{cases} \tag{3}$$

The hidden layers have 16 nodes each. The inputs to the network are the positions of the different joints, α_{1b}, α_{2b}, and l_{ext} (see fig. 1), and the limited velocity reference, $v_{tip,ref}$. The inputs are normalized.

The velocity controller is displayed in fig. 3. It takes in the current state of the crane, and together with the velocity reference and the extension speed command it calculates the 1b and 2b positions required to reach the reference tip position with the selected strategy. The difference between the current and the next state then sets the speed commands for the 1b and 2b cylinders, u_{1b} and u_{2b}. If no valid combination of cylinder positions can be found, the cylinder speed commands of all three cylinders are set to zero. The neural network does thus not control the crane directly. When deployed on a real crane, the velocity controller could also function as a safety layer and reduce the speed or even stop the crane in certain situations.

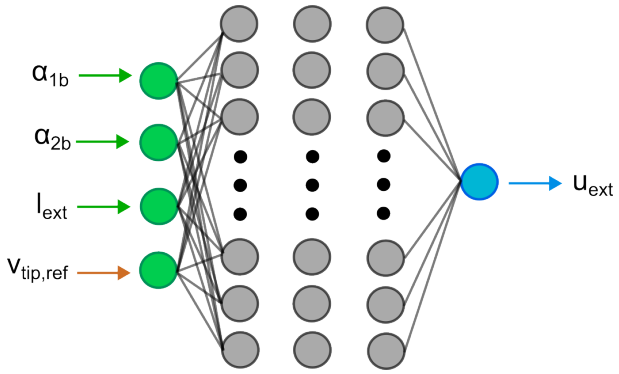


Figure 2: Neural network layout for the CTC strategy controller

2.2 Training Environment

The neural network is trained using the Pytorch framework in a Python environment where a model of the crane is implemented. The model has a geometry part for simulating the movements of the crane, and an energy consumption part for estimating the energy consumption. The geometry model does not include any dynamics and all movements are ideal, i.e. the commanded cylinder speeds always result in the desired movement. Once deployed on a real crane, an other feedback

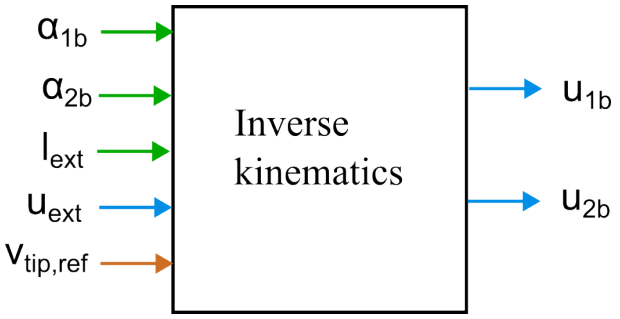


Figure 3: Layout of the CTC velocity controller

loop can handle the modeling errors. The energy consumption model has an ideal flow model, i.e. the pump flow directly corresponds the movement of the cylinders, according to eq. (4), where A_{cyl} denotes the active side cylinder areas, Δ_{clen} the cylinder position increment and T_s the step time. The total flow demand thus differs both with how far the cylinders need to travel and which cylinders that are used, since the areas of the cylinders are different.

The estimation of the pump pressure in the energy consumption model is based on models of the cylinder pressure, fitted to measurement data, together with flow dependent pressure drop models between the valve and the cylinders fitted to measurement data, according to eq. (5), where $p_{func,cyl}$ are the active side cylinder pressures, Δp_{func} are the pressure drops between the valve and the cylinders and Δp_{pump} is the pump pressure margin. The cylinder pressure level required to move each cylinder depends on the torque of that joint but also from counter pressure and friction in the system.

$$q_{pump} = \frac{\Delta 1 b_{clen}}{T_s} * A_{1b,cyl} + \frac{\Delta 2 b_{clen}}{T_s} * A_{2b,cyl} + \frac{\Delta ext_{clen}}{T_s} * A_{ext,cyl} \tag{4}$$

$$p_{pump} = \max[p_{1b,cyl} + \Delta p_{1b}, p_{2b,cyl} + \Delta p_{2b}, p_{ext,cyl} + \Delta p_{ext}] + \\ + \Delta p_{pump} \tag{5}$$

During training, a batch of cylinder positions that corresponds to a specified crane tip position is sampled. A simulation with a fixed timestep on a fixed horizon is rolled out on the batch for a number of episodes. The accumulated cost for each trajectory in the batch is calculated and forms the cost function according to eq. (6), where P_{scaled} is the power required during the current step, scaled with the speed of the current step according to eq. (7), $\Delta\dot{x}$ is the error with regards to the reference speed, and $q_{limited}$ is a boolean which is true if more than the maximum flow the pump can deliver is required. The cost is summarized over a fixed horizon, H, in order to promote long term gains rather than instant.

The policy of the strategy controller is deterministic so the only randomness comes from the sampling of starting positions in each batch. In future studies it would be of interest to see if a probabilistic policy can learn faster.

$$\sum_{i=1}^{H} C_i = \sum_{i=1}^{H} (P_{scaled,i} * w_e + (\Delta\dot{x}_i)^2 * w_x + q_{limited} * w_{qlim}) \tag{6}$$

$$P_{scaled,i} = \frac{q_{pump} * p_{pump}}{|\dot{x}_i|} \tag{7}$$

Even though the velocity reference following is handled by the velocity controller, the cost function includes the velocity error so that the model learns to avoid singularities and to stop when the outer boundaries of the working area are reached.

After each episode, the mean of the accumulated cost of all trajectories is calculated and set as the loss variable the optimization is set to minimize. The gradients of the loss with respect to the network parameters are computed during a backwards pass and the parameters are then updated by the optimizer in order to minimize the loss. The built in optimizer AdamW of Pytorch is used.

2.3 Test Cases

For this initial proof of concept study, two different networks for the strategy controller are trained on two different trajectories, with fixed starting crane tip positions. In each episode, the roll out is simulated for a fixed horizon of 50 samples, that with a sample time of 0.5 s corresponds to 25 s. The value is chosen so as to represent a rather long movement of the crane in a single direction, so that long trajectories can be compared to dynamic programming solutions and benchmarked. In real operation the direction will often change after a shorter time than that, and optimizing over longer trajectories would in many cases be sub-optimal. The horizon might have to be reduced when optimizing a network for the complete working envelop of the crane in future studies. However, a too small horizon might cause the optimizer to only being able to find solutions that are good in a very small, local area. Combining many such locally optimized parts of the controller might not result in a globally optimal solution for longer trajectories.

A fixed payload of 1000 kg and a fixed reference speed of 100 mm/s are used. A reference speed of 100 mm/s does not need to be limited in many areas which makes energy comparisons easier. If a higher reference speed that has to be limited in certain areas is set, a problem is that the optimizer can minimize the energy consumption be lowering the reference speed by finding cylinder configurations that have low maximum speed.

The networks are trained until convergence, which takes about 100 episodes with a learning rate of 0.0005.

The strategy controllers are evaluated together with the velocity controller on four different starting cylinder configurations each, on the same crane tip trajectory as they are trained. Table 1 shows the definition of the four test cases, TC1 to TC4, for the first trajectory, and Table 2 show the same for the second trajectory.

3 Results

The four different test cases for each trajectory are also given as input to the dynamic programming optimization in order to

Table 1: Test Cases for trajectory 1

	Tip start pos [x,y] [m]	Cyl start state $[\alpha_{1b}, \alpha_{2b}, l_{ext}]$ [deg, deg, m]
TC 1	[1, 0.8]	[54.03, -151.2, 0.8795]
TC 2	[1, 0.8]	[28.99, -144.5, 0.1855]
TC 3	[1, 0.8]	[43.76, -148.2, 0.6110]
TC 4	[1, 0.8]	[35.17, -145.9, 0.3686]

Table 2: Test Cases for trajectory 2

	Tip start pos [x,y] [m]	Cyl start state $[\alpha_{1b}, \alpha_{2b}, l_{ext}]$ [deg, deg, m]
TC 1	[3, 2]	[54.03, -151.2, 0.8795]
TC 2	[3, 2]	[28.99, -144.5, 0.1855]
TC 3	[3, 2]	[43.76, -148.2, 0.6110]
TC 4	[3, 2]	[35.17, -145.9, 0.3686]

give an indication of how well, in terms of energy efficiency, the different controllers perform. The dynamic programming algorithm uses the same energy model as the CTC strategy networks are trained on.

Table 3 shows the average power consumption for DP and NN controller respectively, on the four test cases from the trajectory where the first network is trained.

Table 3: Comparison of average power demand for DP and NN controllers for the four test cases of trajectory 1.

Variable	$P_{avg,DP}$ [kW]	$P_{avg,NN}$ [kW]	Diff
TC 1	0.7561	0.7633	+1.0%
TC 2	1.0522	1.0685	+1.6%
TC 3	0.8349	0.8509	+1.9%
TC 4	0.9501	0.9705	+2.2%

The average power consumption is very similar for the two controllers. By looking closer into one of the test cases, illustrated in Figure 4, it is clear that they perform very similar motions of the cylinders. The cylinder commands for the two controllers are displayed in Figure 5. A noticeable difference is that the command to the extension cylinder is not exactly zero for the NN controller since outputting exactly zero is very improbable for a neural network. In situations where the extension is the pump pressure setting function this could be an issue, but during this test case (as well the other three) the first boom is setting the pump pressure for the majority of the trajectory. This gives very similar pump pressures for the two controllers, see Figure 6.

If the second set of controllers, trained and optimized for another trajectory, is considered, another interesting difference regarding the speed command to the extension is highlighted. The average power consumption for the NN controller is actually lower than for the DP controller for all four test cases, see

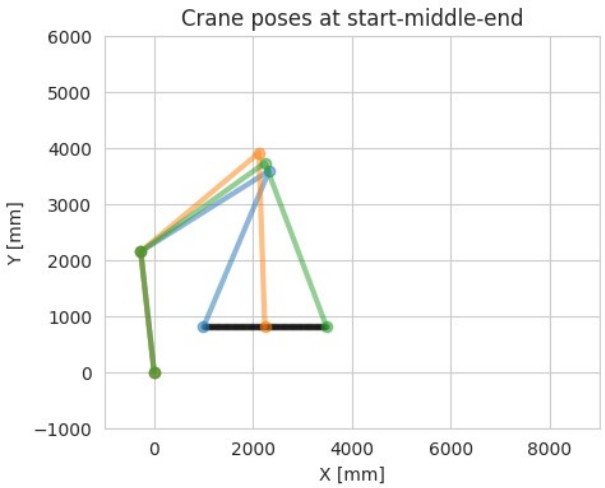


Figure 4: Crane poses at start, middle and end of test case 2 for the NN controller of trajectory 1.

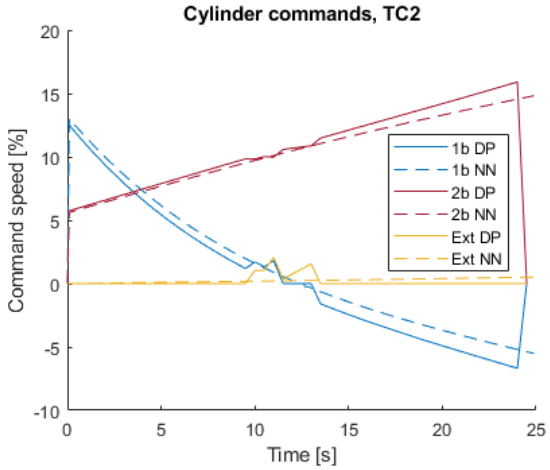


Figure 5: Cylinder commands for the two different controllers, test case 2, trajectory 1. The commands are very similar, but note that the extension command for the NN controller is not exactly zero.

Table 4. The reason is that the network finds solutions where the extension is driven at a very low speed, see the cylinder commands of test case 3 (illustrated in Figure 7) in Figure 8. This results in a lower pump flow during the whole trajectory, thanks to that the first and second boom can be operated more slowly. The pump pressure is not affected by operating the extension, the second pump is setting the pump pressure. The solution the network has found is, however, not valid. The speed command to the extension is lower than the lowest non-zero command that can be sent to the valve on the real system. The dynamic programming algorithm is discretized with a grid on the extension speed that has approximately the same resolution as the valve and can thus not find the illegal solution found by the network. The same issue appears for all four test cases.

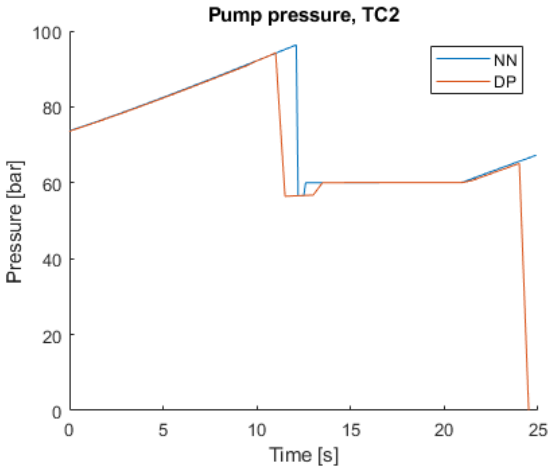


Figure 6: Pump pressure for the two different controllers, test case 2, trajectory 1.

Table 4: Comparison of average power demand for DP and NN controllers for the four test cases of trajectory 2.

Variable	$P_{avg,DP}$ [kW]	$P_{avg,NN}$ [kW]	Diff
TC 1	1.9493	1.8336	-5.9%
TC 2	2.2427	2.0894	-6.8%
TC 3	2.0567	1.8830	-8.4%
TC 4	2.1417	1.9327	-9.8%

4 Discussion

The results of the small CTC strategy networks trained in this work are promising, but only one step on the way towards a complete energy efficient CTC controller. In [2] four different networks, positive and negative x direction and positive and negative y direction, were used to form a complete CTC controller. For any other direction than horizontal or vertical, linear combinations of two of networks are used. Linear combinations of different network keeps the reference velocity but the energy consumption will not be optimal. Given that losses from simultaneous operation is a significant loss contributor, the energy consumption will be far from optimal since the linear combination rarely will result in only two functions being active at the same time. With the network and loss function setup of this paper it is however possible, at least in theory, to train a single network for covering all possible directions.

The network should also cover all possible payloads, all possible speeds and all possible starting positions. Given that the training time for one of the small networks in this study was about 40 minutes on a desktop computer with NVIDIA RTX 4070 GPU, training of such a network will most likely be very time consuming. It must also be considered that the inference time of a, possibly, much larger network must still fit within the sample time of the embedded system on the crane.

When comparing cylinder commands between the DP and the NN controllers, a general behavior is that the NN controllers are more smooth and unable to capture abrupt changes, that in some cases could allow the DP to find more optimal solutions.

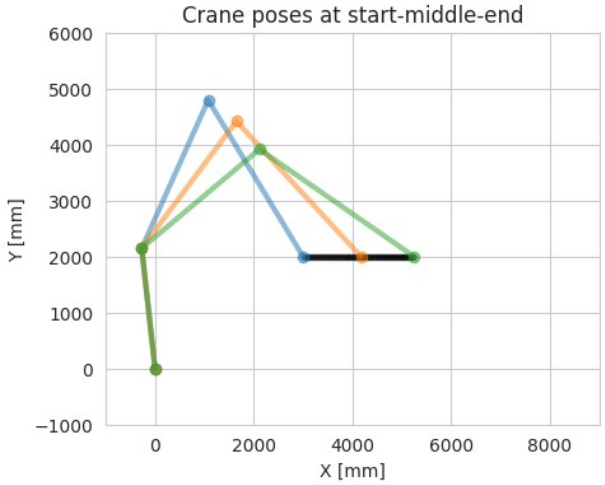


Figure 7: Crane poses at start, middle and end of test case 3 for the NN controller of trajectory 2.

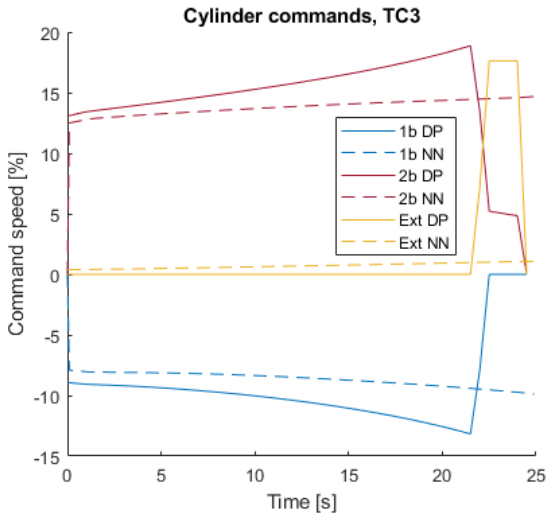


Figure 8: Cylinder commands for the two different controllers, test case 3, trajectory 2. The extension command for the NN controller is smaller than what the DP controller can find.

However, it remains to be evaluated whether this more on-off like control of the cylinders is as good when deployed on a real crane. Fast switching of cylinders can lead to more delays and dynamic forces on the crane, causing worse reference following and energy performance.

As seen in the results of the second trajectory, the more abrupt control of the the DP solutions could also be the cause of the discretization of the DP algorithm. The discretization matches the solutions to the resolution of the valve commands on the real system, and forces it to find solutions that are more energy consuming than if a continuous control signal had been possible, which it what the NN controller assumes. There is no straight forward way of introducing such a constraint on the output from the network during training, as it would be a non-differentiable function which would prevent the backwards propagation of gradients to flow properly. If a NN controller is deployed on the real system, the extension command

could be rounded to a legal value, and the commands to the first and second boom re-calculated if necessary to keep the crane tip within the tolerances of the planned trajectory. The energy consumption would be slightly less optimal, but the main task of CTC, driving the crane tip according to a set speed and direction, would still be carried out properly.

5 Conclusion

With sufficient training, a neural network for deciding CTC strategies can be taught that, together with a feed forward velocity controller, performs close to the optimal in terms of energy efficiency, on the small part of the working area that it is trained for. On the evaluated test cases, the energy consumption of the CTC controllers developed in this study consumes up to 2.2 % more energy than the near-optimal solutions given by a dynamic programming algorithm. The question remains how well the results will hold for a less generalized network that covers the whole working area of the crane.

Acknowledgment

References

[1] Amy Rankka, Marcus Rösth, and Alessandro Dell'amico. Evaluating the performance of semi-autonomous kinematically redundant loader crane operation. GFPS PhD Symposium, accepted for publication, 2024.

[2] Abdolreza Taheri, Amy Rankka, Pelle Gustafsson, Joni Pajarinen, and Reza Ghabcheloo. End-effector cartesian velocity control for redundant loader cranes using reinforcement learning. *IEEE Transactions on Robotics*, 2024.

[3] Hiab Official Webpage. Ctc – crane tip control taking ease of use to a new level, 7 2020.

[4] Rickard Nilsson. Inverse kinematics, 2009.

[5] Jarmo Nurmi and Jouni Mattila. Global energy-optimal redundancy resolution of hydraulic manipulators: Experimental results for a forestry manipulator. *Energies*, 10, 2017.

[6] Alessandro Tringali and Silvio Cocuzza. Finite-horizon kinetic energy optimization of a redundant space manipulator. *Applied Sciences (Switzerland)*, 11:1–15, 3 2021.

[7] Jennifer Andersson, Kenneth Bodin, Daniel Lindmark, Martin Servin, and Erik Wallin. Reinforcement learning control of a forestry crane manipulator. 2021.

Accelerating Product Development: The Effort of Asset Administration Shells for aggregated Systems

Denis Ritz, Malte Becker*, Jürgen Weber and Katharina Schmitz*

Technische Universität Dresden, Germany
*RWTH Aachen University, Germany
E-mail: denis.ritz@tu-dresden.de, malte.becker@ifas.rwth-aachen.de, mailbox.fluidronik@tu-dresden.de, katharina.schmitz@ifas.rwth-aachen.de

Abstract

When developing and revising products, a wide variety of engineers usually work together. Such heterogeneous development teams often work on the same product but with different models. Typically, each team works on its own partial version of a specific product.

At the same time, finding the right components for each subsystem is another challenge. The technical data sheets are often only available as PDF files. This makes it difficult to access the data they contain by machine. Furthermore, there is no standardized designation of product properties across the various manufacturers.

The joint research project "Fluid 4.0" is working on standardized semantics for fluid power technology. Among other things, existing standards such as "ECLASS" are utilized. This means that a certain basic stock of properties is already available. At the same time, the list of properties is analyzed initially, and missing features are detected.

This publication will show how these extended semantics can be used to create administration shells for aggregated systems. To this end, selected aspects of the specific extension of the semantics are explained. During the project, it is planned to submit this extension as a proposal to the standardization committees in order to provide a coordinated, machine-readable database for development processes.

An important point here is the linking of the individual components to form an aggregated system. For this purpose, a proposal for the creation of asset administration shells for aggregated systems is explained. The versioning and access control mechanisms of the administration shells make them ideal as a single source of truth for the collaborative development of complex, mechatronic systems. All of this is discussed using simplified examples.

Keywords Asset Administration Shell (AAS), System Development, Semantic

1 Introduction

In the product development process, the layout of the hydraulic powertrain is usually selected at the beginning. Choosing the layout is not a trivial part of this process, but it is usually much more straightforward than choosing the right components. During the development process, it starts with a few requirements at an abstract level of analysis, such as a specific hydraulic power. As the development process continues, decisions are made, such as setting the system pressure to a certain level. This then leads to more and more specific requirements for each component. This is where the tricky part begins: searching for components that fit the specified requirements. Among others, this is the widely varying quantity and quality of product data, which is generally difficult to compare across different manufacturers.

If a semantic search, which is possible across manufacturers, existed for this purpose, this step would be significantly simplified. This could be made possible with asset administration shells and a correspondingly specified submodel (SM) for the components.

In view of the wide range of specialist areas addressed, this paper attempts to discuss the various topics in such a way that they are not only comprehensible to the individual experts, but

also to people from outside the field. To conclude, the hydraulic engineer must be able to engage in meaningful discussions with the computer scientist and vice versa.

This paper presents the potential of the Asset Administration Shell to accelerate product development. Therefore, section 2 references the related work. After that, section 3 introduces the Asset Administration Shell, followed by a description of relevant terminology and standards in section 4. Section 5 integrates the AAS into the context of the product life cycle, after which the contribution is summarized in section 6, and potential benefits are highlighted.

2 Related Work

AAS has already been used in scientific work in the context of fluid power. In the publications [1–3], the AAS is studied in the context of the simulation-based development process. Use cases such as the requirements-based search for AAS in repositories and the parametrization of simulation models are initially considered. The dissertation [4] investigates the use case of commissioning and investigates the extent to which fluid power systems can be commissioned automatically.

In the research projects [5], [6], aspects for the development of AAS and an OPC UA specification for pumps [7] were shown. These two research projects, as well as project [8], are being worked on at the same chair. The latter project dealt with the OPC UA specification for compressed air stations [9]. It can be assumed that significant preliminary work for the two OPC UA specifications [7], [9] was carried out as part of these projects. Unfortunately, the authors only have insufficient sources for the content created within the aforementioned projects.

The papers [10–12] deal with the comparison of different semantics, the mapping of a source semantics into a target semantics, as well as simplified interoperability. In [13], a search algorithm for mapping proprietary features to ECLASS properties is described.

In addition to its application in fluid power, the AAS is already being used in various branches of industry. In electrical engineering, for example, the AAS was used in a control cabinet demonstrator to enable the automatic calculation of the product carbon footprint (PCF), whereby components from various manufacturers were included [14]. In addition, the AAS is also used in research activities in other areas, e.g., in the manufacture of electric motors [15] and water supply [16].

3 Asset Administration Shell

The Asset Administration Shell (AAS) is a virtual representation of an asset in the context of Industry 4.0 (I4.0), see Figure 1. In general, assets are physical, digital, or intangible objects that represent values for one or more stakeholders. An example of an asset in fluid power is a pump or a valve that is represented virtually with an AAS. In the context of the AAS, assets are described by SMs and submodel elements (SME) which represent properties. A SM is a set of

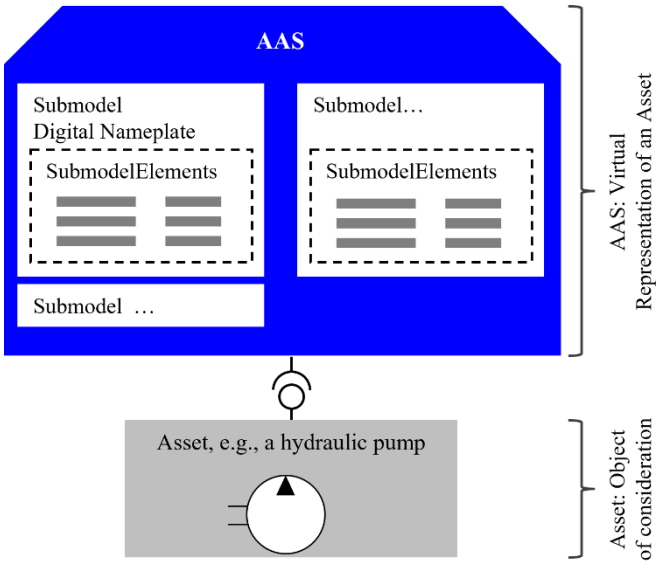


Figure 1: Exemplary AAS according to [20]

properties, collections, methods, etc. for a specific use case. SMs can then be standardized as templates to consider a range of relevant properties required for specific use cases. The SM template “IDTA 02006-3-0: Digital Nameplate for Industrial Equipment” [17] for an AAS represents a digital data sheet of the asset, while the SM template “IDTA 02005-1-0: “Provision of Simulation Models“ [18] can be used to exchange simulation-based properties and models of the asset. Another example for a SM template is “IDTA 02011-1-0: Hierarchical Structures enabling Bills of Material,” [19] which can be used to describe hierarchical structures of fluid power systems. Both SM and SMEs can be annotated with semanticIds, which will be discussed further in section 4.2. [20] Assets and thus also the corresponding AAS are distinguished in type and instance [21]. A type asset describes a product or component's general, abstract definition and is usually created during the development phase. An instance asset is a concrete, physical realization of a type asset. In production, instance assets are created based on the specifications of the type asset and linked to product-specific data (e.g., production, logistics, and test data). As a rule, there is a relationship between instance assets and a type asset. However, every instance asset does not need a corresponding type asset. [22]

In the context of fluid power, the type refers to a pump variant from the manufacturer's range, while an instance is a specific, individually available pump that has been installed in a specific application, for example. This work focuses on type AAS, which is decisive for the development process.

An AAS can represent rudimentary components, such as a “simple” non-return valve, the drive of a partial function of a production machine, but also the hydraulic ram of a deep-drawing press, a complex machine, as well as a hydraulic excavator, or even an entire production plant, such as a rolling mill with a connected press line for the production of semifinished products for the automotive industry.

4 Relevant Terminologies and Standards

According to the authors of [23], the AAS could have a huge impact on businesses and business models in general. The AAS can act as the interface between “our physical” world and the “digital” world. So, if a semantic is created for a specific SM, this is the vocabulary which has to be used in every domain. Therefore, naming the individual elements should follow a meaningful convention, from the point of view of every domain involved. Therefore, in the following existing standards, dictionaries, and semantics are reviewed. In the context of semantic descriptions, terms such as data, information, and knowledge are often used misleadingly [24]. To create a uniform understanding, this paper uses the definitions made by North in [25], which is explained in the following paragraph.

Data is formed from characters using syntax rules. It consists of arbitrary and uninterpreted character strings, such as any sequence of numbers and becomes information via semantics. Information is, therefore, data that has a context of meaning. It is created by linking information with current information or information stored in the past. Knowledge is characterized by the fact that it is shaped by individual experience, context-specific and personal. This hierarchy of terms is illustrated in Figure 2. [25]

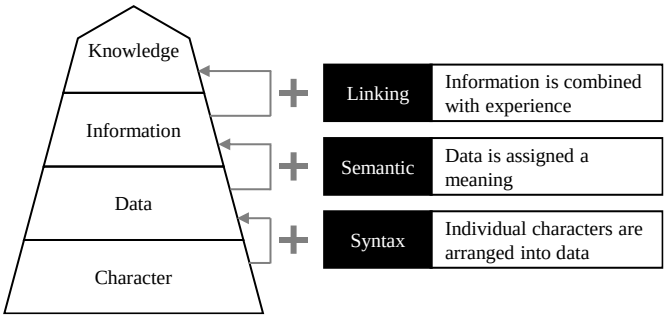


Figure 2: Conceptual hierarchy of data, information and knowledge according to [25], [26]

4.1 ISO 1219

The ISO 1219 standard [27] or the corresponding national implementations of the standard, like DIN ISO 1219 [28], represents the graphic symbols of all fluid technology components. Every hydraulic engineer and every service technician understands these symbols and can use them to identify the function of a component and a system. At the same time, the standard provides a structural classification of all standard components. Therefore, the contents of this standard should have a strong influence when defining the semantics, product groups of hydraulics and the description for the topology of systems.

Figure 3 shows the relevant component classifications as defined in ISO 1219-1[27]. It is striking that the standard explicitly divides directional, pressure, and flow control valves into “normal” and proportional valves. The vertical lines in the illustration indicate the related product groups.

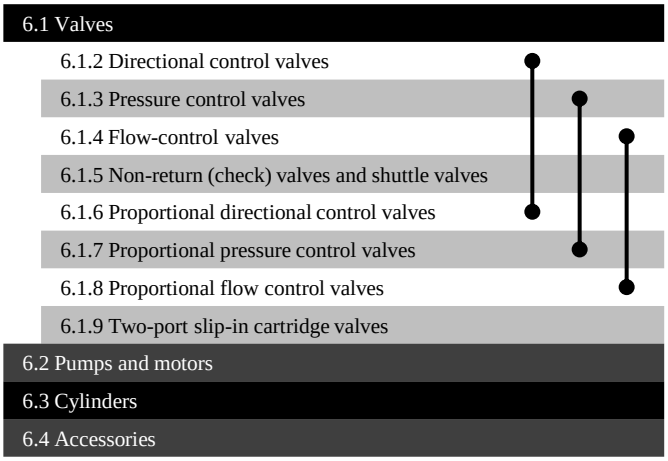


Figure 3: Component classification from the table of contents of ISO 1219-1 [27]

From the authors' point of view, this duplication between "normal" and proportional raises whether proportionality is not more than a property of the valves and, therefore, does not require a separate subdivision.

Another point that stands out is the category "Cartridge valves," which is the only place in the standard where the design of a valve is explicitly mentioned. These valves can also be described purely in terms of their function. All valves in this group can be classified according to their function, either as 2/2-way valves or as pressure-reducing or limiting valves.

ISO 1219-2 [29] explains the rules to be followed when presenting a hydraulic diagram. This is explicitly required in section 5.2 of this standard, which is necessary in order to be able to clearly identify the specific connections on a valve. Figure 4 shows this for a simple system.

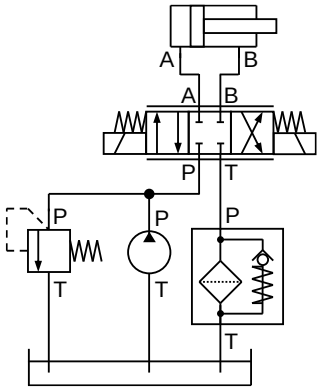


Figure 4: Example of the hydraulic scheme for a small aggregated hydraulic system

The last part of the standard, ISO 1219-3 [30], explains so-called symbol modules. These modules generally describe complex products that are offered by manufacturers of hydraulic components as ready-to-sell products. This often applies to the valve sections of mobile hydraulic systems. Figure 5 shows the module symbol for a LUDV valve disk from Bosch Rexroth [31].

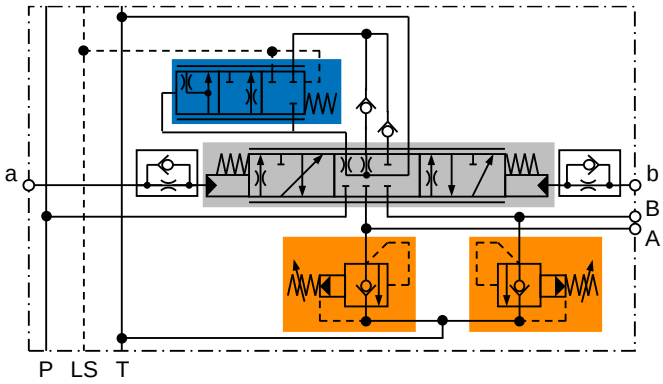


Figure 5: Module Symbol for the valve section of the LUDV system by Bosch Rexroth, based on [31]

This and similar valve systems are frequently used in construction machinery. Figure 5 shows that the LUDV valve itself is already a subsystem consisting of several components. The main valve, which controls the direction and speed of the connected actuator, is shown in gray. The pressure compensator is shown in blue. This valve system consists of at least these two components. The valves shown in orange are optional additional functions for secondary pressure limitation and post-suction, which often can be found in mobile machines.

For module symbols, only the connections need to be marked on the outside, but a frame must be used to delimit the module. This example also makes it clear that semantics must be flexible enough to describe valves that are not directly represented in the standard, such as the gray marked 7/3 directional valve.

4.2 Semantic Dictionaries

A unique semantic understanding of data is a prerequisite for implementing I4.0 use cases. According to Wang [32], this is referred to as “semantic interoperability” and is made possible using a unique semantic reference model.

However, data is often available in proprietary formats, e.g., based on manufacturer-specific data descriptions. An example illustrates this problem: One manufacturer's data sheet uses pressure-related properties like “Max continuous pressure”, “Peak pressure”, “Minimum Inlet Pressure”, “Maximum Inlet Pressure” [33] to describe a hydraulic pump, while another manufacturer's data sheet describes the hydraulic pump with pressure-related properties like “Nominal pressure,” “Minimal pressure,” and “Maximum pressure” [34]. This example alone demonstrates the potential added value of the algorithm presented in [13].

In addition to the challenge of the overarching comparability of components from different manufacturers, imposed by the need-to-know which property is related to the other, these different descriptions stand in the way of semantic interoperability. To counteract this challenge, there are semantic dictionaries such as ECLASS [35] or IEC 61360 CDD [36]. Semantic interoperability is given, under the assumption that corresponding stakeholders are committed to a semantic dictionary. In general, different semantic dictionaries are not interoperable with each other [11]. ECLASS itself is a non-profit association, and the steering committee is largely made up of companies from the pharmaceutical industry, automation technology and fluid power [37].

ECLASS provides defined properties for various domains, including fluid power. Within a domain, ECLASS is hierarchically structured. For example, an axial piston pump with variable displacement and bent design (product variant) belongs to the product type of axial piston pump. This product type, in turn, falls under the broader product family of pumps within the fluid power domain, which is shown in Figure 6. ECLASS properties or classifications are assigned an IRDI (International Registration Data Identifier), enabling unique identification. For instance, the shown product variant is uniquely identified by the IRDI 0173-1#01-AGW107#006.

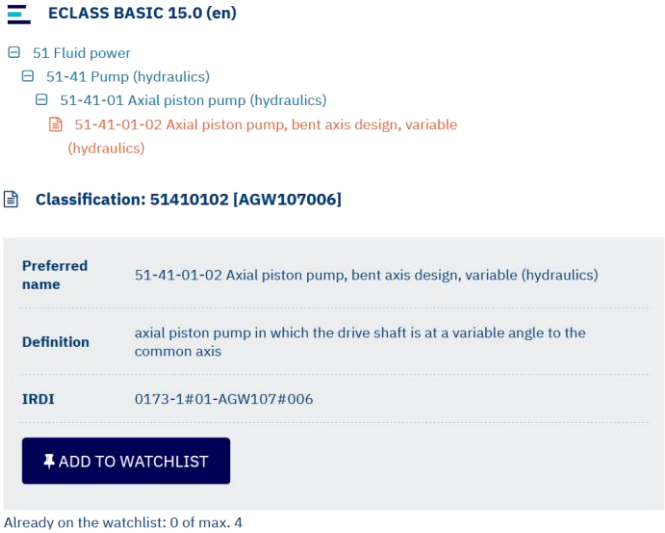


Figure 6: Exemplary Product Classification in ECLASS [35]

Furthermore, properties are assigned to the product variant. For example, the property “nominal pressure” with the IRDI 0173-1#02-ABL782#001 is assigned to the product variant. In the context of I4.0, the IRDI can be used as a semanticId [20]. In addition to the example above, ECLASS has numerous other product families, such as motors and valves with respective product types and product variants from fluid power.

ECLASS version 15 already has numerous general properties for fluid power. Missing properties are continuously being added, for example, the property displacement was not yet defined as a central feature for describing pumps in the previous version. In addition, duplications can still be found in ECLASS, such as the property “set pressure”, which appears under both IRDIs 0173-1#02-ABC506#003 and 0173-1#02-AAZ819#003, even though they represent the same concept. Discrepancies are to be expected if manufacturers refer to different semantics for this property.

Besides ECLASS, there is also IEC 61360 CDD [36], which serves as a semantic dictionary with a strong focus on electrical engineering. The properties defined in IEC are also assigned unique identifiers, such as the "actuating force" with the identifier AAE932-005. This characteristic can be expanded into a complete IRDI by specifying the standard source: 0112/2///61360_4#AAE932-005 [38]. In contrast to ECLASS, however, IEC does not provide specific semantic definitions for fluid power components.

4.3 Semantic Modelling

The publication [39] provides an overview of several different languages for semantic modeling. The authors of [39] also refer to ISO 15926 in their publication. The ISO 15926 standard describes the semantic modeling of data and integration during the product life cycle in detail. Part 2 of this standard [40] covers data modeling, which is particularly relevant for the development of semantic models, from the point of view of the authors of this paper. In [41], it is shown, among other things, how an AAS must be structured to comply with the rules of ISO 15926.

The Eclipse Semantic Modeling Framework (ESMF) [42] offers a suitable option for creating semantic models for administration shells. The Eclipse Foundation itself has been active since 2001 and has set itself the goal of providing open-source software for many areas of technical development. The members of the foundation cover many different industries, with many well-known, large I4.0 companies. [43].

The ESMF's basis is the Semantic Aspect Meta Model (SAMM), which forms the basis for the semantic description of SMs in so-called aspect models. These models express domain-specific semantics by linking structural elements, relations, namespaces, and verified concepts.

Figure 7 shows an example of modeling technical data according to the suggested structures of the SAAM [44]. In addition to integrating the ECLASS IRDI, the property's minimum and maximum value range and the properties' underlying unit are also considered. This comprehensive structure enables companies to prepare their proprietary data uniquely and semantically consistently. This significantly improves interoperability between different systems and data utilization in different contexts.

4.4 schema.org

The schema.org initiative provides a vocabulary for structuring data. The main drivers of this initiative are large digital companies such as Google, Microsoft, and Yahoo, to name just a few. The vocabulary can be used to represent entities of a wide variety of objects and link them together. According to schema.org, its vocabulary is already used by more than 45 million websites. [45]

The vocabulary of schema.org offers to categorize products hierarchically using the types "ProductGroup", "Product" and "ProductModel". Figure 8 shows a possible structure for

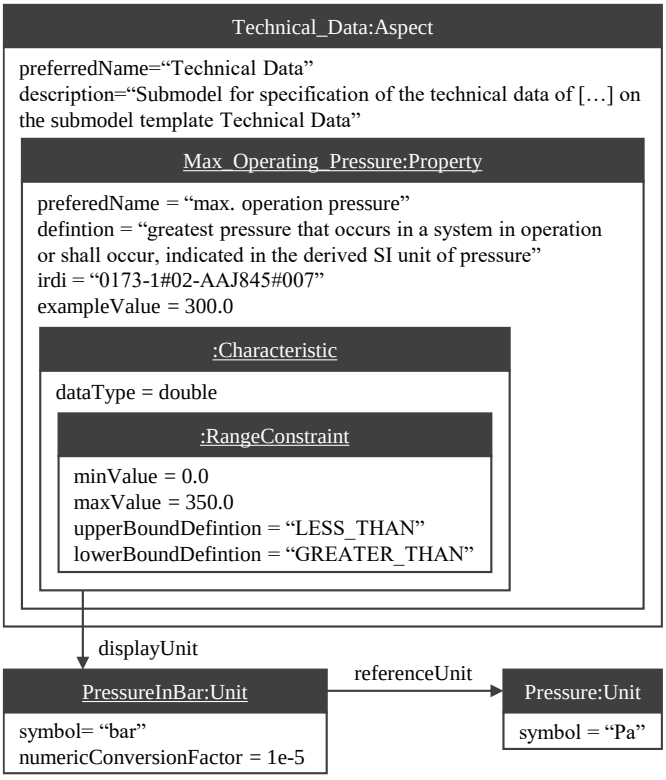


Figure 7: UML-like Object Diagram for a Semantic Modeling using the terms of the SAAM [44]

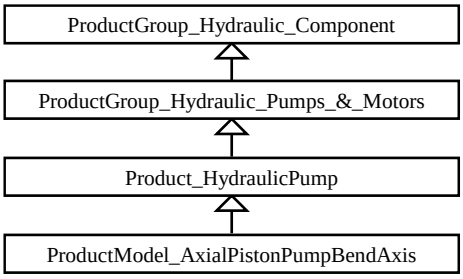


Figure 8: UML-like Class Diagram for structuring an axial piston pump in bent axis design

classifying an axial piston pump with inclined axis, which also corresponds to the example in Figure 6.

The hydraulic components correspond to the entire content of ISO 1219-1 [27]; the pumps and motors group is equivalent to section 6.2 of the standard. The symbol for a variable displacement pump can be found in subsection 6.2.1 of the standard. At this point, the standard does not differentiate between the possible designs of a pump, as this has no influence on the defined function of the component. In the schema.org description, a product model can also be derived from a product and thus represents the fact that every axial piston pump in bent axis design is a pump, but not every pump is an axial piston pump in bent axis design.

Compared with the classification according to ECLASS in Figure 6, some small differences are recognizable. For example, ECLASS already differentiates between pumps and motors at a high level and then displays the pump type in a

detailed manner over several steps. Likewise, adjustability is not used as a property of the pump as discussed in section 4.1, but as a distinction between several types.

When creating schema.org-compliant semantics, the question also arises, whether it makes sense to explicitly differentiate between pumps and motors, as the majority of all pumps can also be used as motors and vice versa. Whereby a “real” component is usually optimized for its primary purpose, which also must be readable in data of the component. This allows the number of product groups to be reduced, as long as the various products have comparable interfaces. From the authors' point of view, it makes sense to combine all valves in a “ProductGroup_Valves” and to provide this product group with subgroups for each configuration of the number of connections. This means that such a structure must be suitable for all components and flexible and expandable for components that do not fit into the standard or are not yet included. An example of this is the so-called “hydraulic transformer” from INNAS BV, which is described in [46].

As a final note in this section, it should be mentioned that the schema.org vocabulary, with “IndividualProduct”, contains an equivalent for the instance of an AAS.

4.5 Structured Information in Communication protocols

Another important aspect of I4.0 is the connectivity of systems. In [47], the possibilities for communication between a physical asset and its AAS, among other things, are summarized. The EDC (Eclipse Dataspace Connector) project [48] attempts to create a simple way of combining the various possible communication channels into an AAS with this software project. The authors will therefore discuss what they consider to be the most relevant communication standards, which are OPC UA and MQTT.

The origins of the OPC UA communication protocol date back to the 1990s. The main drivers at the time came primarily from automation technology. Even back then, there was a strong focus on interoperability. The first version of the OPC UA standard was released in 2006. [49]

The OPC Foundation has now published over 150 specifications. Within these specifications, basic properties are defined that build on and reference each other. In the “Data access” specification [50], for example, a proposal is made on how physical units should be represented. The advantage of this is the unambiguous assignment of a specific value based on the SI base units. This is a clear advantage over purely semantic definitions, which only describe the physical unit of a value within a text. One other core specification is “Devices” [51], it already contains a definition for “VendorNameplate”, which is similar to the IDTA SM Template "Digital Nameplate for industrial equipment" [17]. The recently published specification for “Powertrain” [52] contains more component definitions for the elements of electro-mechanical drives. Here, explicit attention is paid to object orientation for structuring; each derived class in turn contains objects which, if the link is followed, ultimately lead to formally described properties. Whenever possible, these properties should be linked with a semanticId, for example ECLASS.

Another important specification in this context is “Pumps and Vacuum Pumps” [7]. This specification defines many important properties and statuses of a pump. The specification explicitly references the respective physical unit of values. However, no distinctions are made regarding the design, but the different hydraulic connections are modeled. Based on this document, a proposal for the design of hydraulic systems has already been created in [53] using the example of a mobile excavator. The concrete modeling of the hydraulic connections and their linking is also used there. Another advantage of OPC UA is its capability to provide an automation interface. This was realized during the joint research project “Bauen4.0” [54]. As an extension of this system, a 3D safety system for a mobile excavator was implemented, which is described in [55].

The MQTT protocol was originally developed to reliably send messages in networks with low bandwidth and high latency. The “Technical Committee” of MQTT is mainly made up of software companies. [56]

The possibility of dividing the individual messages into different “topics” allows the messages to be structured and filtered for the corresponding recipients [57]. Although the protocol offers a good possibility for structuring, it has no explicit rules for structuring, for example which names are assigned and how. There is no recommendation from the “Technical Committee” as to how “topics” should be further subdivided. Even if this offers an exciting opportunity for distributed communication in an I4.0 network, few specific points can be derived for structuring. However, the division of messages into “topics” represents an interesting possibility for setting up field communication. This could be interesting when setting up field device communication with structure based on the semantic description.

From the authors' point of view, however, OPC UA is much better suited for this, as this protocol can already map a clear, unambiguous, semantically describable structure through the specifications.

4.6 Modeling and Model-Integration

In the context of this topic, two standards should be mentioned. The first one is the FMI/FMU standard, which allows integrating of nested simulation models in different tool. The other one is the Modelica language, which is suited for every domain, as long as the model can be described with ordinary differential equations. Both allowing to combine different model domains in one simulation environment. FMI/FMU is a way of exchanging simulation models between different tools. The respective simulation part model is packaged as a container, so to speak. [58]

These containers are described via their interfaces. It is important that the type of modeling is taken into account: what are model inputs, what are model outputs. The containers conceal the internal structure. There are now 235 different programs that support at least one version of the FMI standard, 69 of which support the current version 3.0 [59]. The consortium behind FMI consists largely of simulation software

manufacturers and large companies that use simulation in the automotive, automation and aviation sectors [60].

In summary, the FMI standard enables the simple cross-domain use of simulation models. However, the internal structure of the respective model is hidden, which is often in the interest of the providing companies. This means that the interfaces must be precisely described in order to be able to work with these models.

The object-oriented modeling language Modelica makes it possible to model complex technical systems using equations, regardless of the physical domain. In contrast to other simulation tools, which are typically geared towards a specific problem domain, e.g., for multi-body systems, electrical circuits or fluid dynamics. The Modelica language itself is maintained and further developed by the Modelica Association. In addition to simulation software manufacturers that distribute Modelica-compliant programs, the association also brings together several companies from the fields of automation technology and automotive engineering. The association is also involved in the development of the FMI/FMU standard. [61]

Two another relevant features of the Modelica language are that it explicitly supports the use of physical units based on SI-units and conversion to other units. Models of a system are based on components, which are linked via connectors. For hydraulic components those connectors could match exactly to the fluid ports and the electric pins. Modelica itself is also mentioned in [18] as an option for providing a simulation model from an AAS.

In principle, the Modelica standard libraries already provide many basic building blocks for creating complex simulation models. However, the standard library contains just few basic elements that can be used to build hydraulic systems directly. For more hydraulic component-oriented models, these must either be implemented by the user or by using third-party solutions such as the freely available “OpenHydraulics” package [62]. The package itself uses object-oriented, abstracted approaches for the efficient implementation of different but similar components. For example, there is a common base class for different 4/3 directional valves, which contains all the connections for such valves. The valves derived from this then differ, for example, in the linking of the connections in the center position. For “more complex” elements, such as a valve section for an LS system, this element is aggregated from existing models of individual components. The collection of examples in this library already includes some simulation models for hydraulic systems, which are also made up of individual components.

4.7 Conclusion to the reviewed Terminologies and Standards

Compared to Figure 2, each of the topics studied brings aspects of syntax, semantics or links to the hydraulic components and systems. With SAMM (section 4.3), it is possible to introduce syntax and semantics for parameters of components to fully describe those with standardized physical units. Ideally, such semantics are standardized with a semantic dictionary, such as ECLASS (section 4.2). For structuring and comparing components, it would be a good option to structure products in a hierarchical relationship to each other in order to make them easier to compare. Such a product structure can be realized with schema.org (section 4.4) and should be based on the product groups of ISO 1219-1 [27] (section 4.1) for hydraulics. Many of these functionalities are implemented in Modelica (section 4.6). At the same time, Modelica also allows components to be linked in sub-models like “module symbols” in accordance with ISO 1219-3 [30] or a complete hydraulic system to be assembled from individual components in accordance with ISO 1219-2 [29] (both section 4.1). In principle, compatibility and transferability to other standards, such as the Companion Specifications of OPC UA, should be sought wherever possible (section 4.5).

5 Integrating AAS into the Product Data Backbone and Product Life Cycle

The use cases described here address the product lifecycle phases of product planning and development. This is usually also the area in which PDM (product data management) is traditionally used [63]. This means that integration into the development processes also requires a link to the PDM system, and ideally also to the company's CAD program. Each reference in the AAS of the aggregated system should therefore correspond to a (purchased) part or an assembly in the PDM and CAD system. This has a deep impact on the product data backbone of a company. Thanks to the possibility of versioning and revising an AAS, this can also be used as a data basis for system engineering, which can thus be integrated into the typical working method of a PDM system. Figure 9

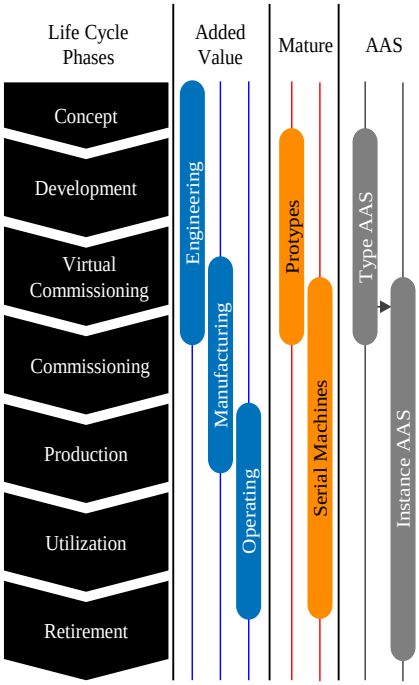


Figure 9: Use of AAS for a specific product during the product life cycle, the maturity of the product and the phases of value creation based on [21], [63–66]

shows the possible contexts of AAS along the product life cycle. For the product life cycle, a comparison was made using the documents [21], [63–66], as each of these documents has a slightly different focus.

Next to the administration shells is the maturity of the product and the value creation that takes place. It shows that the AAS type is relevant to large parts of engineering and can also be used for production planning. A type AAS of an aggregated system could be used to create virtual prototypes based on their linkage. At the same time, the type AAS should continue to mature with physical prototypes to serve as a blueprint for instance AAS, which are automatically derived from the type AAS during manufacturing. Ideally, the data from the manufacturer's ERP (Enterprise Resource Planning) system is linked during derivation. Figure 10 shows a proposal of which types of AAS should be used within a company's data and software architecture.

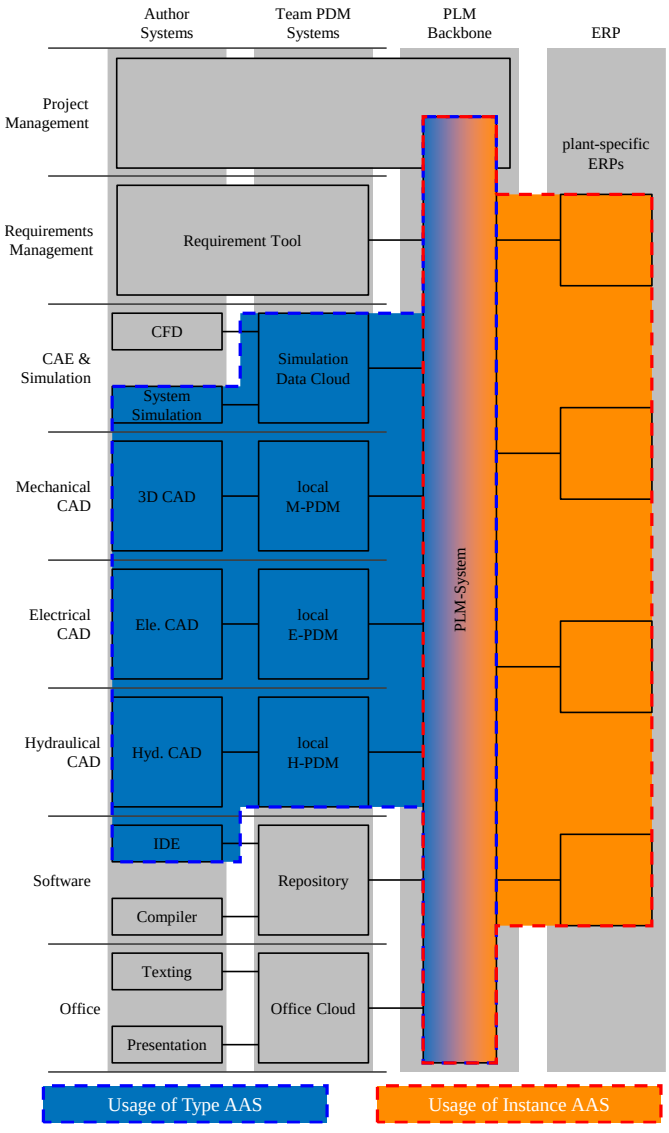


Figure 10: Use of AAS types in the various program systems in an exemplary PLM architecture based on [63]

As shown in the figure, all authoring systems that can map individual components should have a reference to the type shells of the corresponding components. Whether it is the system simulation with lumped parameters, which either loads concrete models of the components from the SM "Provision of Simulation Model" [18] or loads them from the SM "Technical Data for Fluid Components", or the "classic" 3D CAD program which loads the 3D model via the SM "Provision of 3D Models" [67] and automatically creates or updates the SM "Hierarchical Structures enabling Bills of Material" [19] when assemblies are created. This list could be extended further, but essential is that consistent identification of the same component must be ensured in all authoring systems. This could be done with the type AAS of the product as the "single source of truth" during the development process.

As already indicated, an instance AAS, that could be derived from the type AAS, should also be composed by the PLM system with all available data. This is also shown in Figure 10. In this case, the instance shell is created as a "copy" of the type shell. However, this copy receives some additional data, such as its own serial number. It is also possible to link contained parts and assemblies with their specific instance shell, which could be the hydraulic pumps, valves or even the combustion engine of a mobile machine. With this it all relations form parts and subassemblies could be described.

6 Conclusions and potential Benefits

In this paper, a wide variety of documents and templates for the development of semantics for specific products, the classification of products and the topology of systems were reviewed. Most of these documents are largely driven by the automation technology sectors and large digital companies. These documents can also be transferred to fluid power systems. In the opinion of the authors, the following statements should apply to the terminology of AAS for fluid power components and systems:

A hierarchical classification of the individual components, as already intended in ISO 1219-1 and the product classification according to ECLASS. However, the design solution does not play a role in the classification, as is envisioned e.g., in the OPC UA specification for "Pumps and Vacuum Pumps". The hierarchy is compatible with the conventions of schema.org. The principles of object orientation are used for inheritance from an abstract to a concrete description.

The characteristics of each component are uniquely described via an identifier that refers to an SI-compliant physical unit. As in Modelica, for example, this unit is specified to be converted into any unit describing the same physical quantity. The unit's assignment is not purely based on a text entry in the description of the identifier; this can lead to errors if a physical unit is written in different ways.

A component's connections (hydraulic ports, electrical pins, etc.) are part of the elements. By linking the connections, a more complex element can be composed of several

sub-elements, so aggregated systems can also be represented. This should also fulfill the requirements of ISO 1219-2 and ISO 1219-3.

From the authors' point of view, this would result in the following benefits.

With this semantic definition, components from different manufacturers can be directly compared based on their characteristic values. This enables a cross-manufacturer semantic search.

These aggregated systems also provide knowledge about the system's internal structure, which can also be used to utilize this structure for further applications. This can be used to derive technical documents such as the hydraulic circuit diagram or the BOM, or to create and parameterize simulation models as part of the development process, or to load the relevant parameters into the machine program code. The permissibility of the connection can also be checked by linking the connections. Errors can thus be detected much earlier in the development process. For instance, this can be the comparison of the burst pressure of a component with the set value of the pressure relief valve or the comparison of the signal type of the sensor and the input of an electrical controller.

When a product is manufactured, automatic parameterization for the machine can be carried out based on the instance AAS of the subcomponents by linking them to the individual instance AAS of the subcomponents. Thus, time-consuming calibration during the commissioning of a series product can be significantly reduced.

These aggregated systems also provide knowledge about the internal structure of the system, which makes the utilization of this structure for further applications possible. This can be employed for example, to derive technical documents such as the hydraulic circuit diagram, or the parts list, or to create and parameterize simulation models as part of the development process, or to load the relevant parameters into the machine program code.

7 Outlook

As part of the Fluid4.0 research project, the authors will attempt to implement the conclusions reached in this work and develop a proposal for a partial model for the technical data of hydraulic components and systems. During the project, this proposal will be suggested to IDTA as the basis for an official partial model.

In addition, the advantages also explained in the context of this work are to be demonstrated using an exemplary software implementation. Figure 11 shows a mock-up of the user interface of this application. Although there are already commercial solutions (e.g. Eplan [68]) for creating system structures, the planned solution primarily offers advantages in analyzing basic systems and component dependencies in the product development process independent of commercial

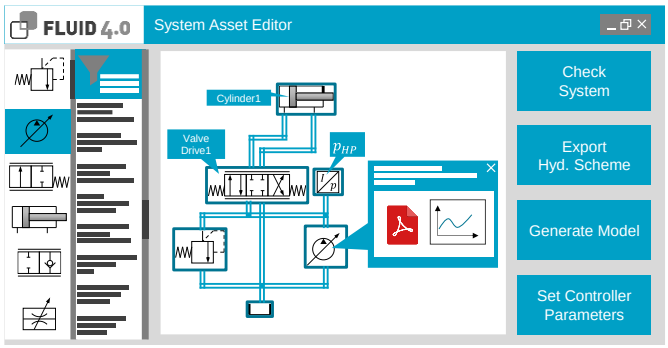


Figure 11: Mock-Up for the prototypical application, which is planned during the Fluid4.0 joint research project

solutions. At the same time, the tool will especially provide smaller and medium-sized companies with an opportunity for creating system descriptions based on the AAS, independent of license costs.

Another critical point is the standardization of the vocabularies and definitions and the definition of uniform translations. From the authors' point of view, this is currently one of the main challenges for the far-reaching implementation of I4.0.

8 Acknowledgement

This work has been funded by the European Union (EU) and supported by the German Federal Ministry for Economic Affairs and Climate Action (BMWK) in the project “Fluid4.0 – Revolutionizing Fluid Engineering: Implementation of digitalization for fluid power 4.0 in the cross-industry and cross-manufacturer data space using asset administration shells, submodels and demonstrators” (project number 13IK039L). The views and opinions expressed are solely those of the authors and do not necessarily reflect the views of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

Funded by the European Union
NextGenerationEU

Supported by:
Federal Ministry for Economic Affairs and Climate Action
on the basis of a decision by the German Bundestag

References

[1] M. Becker, S. Heppner, R. Alt, T. Kleinert, and K. Schmitz, “Improving Fluid Power System Simulation Through an AAS-Based Simulation Framework,” presented at the ASME/BATH 2023 Symposium on Fluid Power and Motion Control, American Society of Mechanical Engineers Digital Collection, Nov. 2023. doi: 10.1115/FPMC2023-109807.

[2] S. Heppner, T. Miny, T. Kleinert, M. Becker, K. Schmitz, and R. Alt, “Asset Administration Shells as Data Layer for Enabling Automated Simulation-based Engineering,” in *2023 IEEE 28th International Conference on Emerging Technologies and Factory Automation (ETFA)*, Sep. 2023, pp. 1–7. doi: 10.1109/ETFA54631.2023.10275474.

[3] R. Alt *et al.*, “Seamless integration of device and field data into the system simulation of a hydraulic servo-press using aas,” in *14th International Fluid Power Conference*, Dresden, Germany: River Publishers, 2024, pp. 548–565. doi: 10.13052/rp-9788770042222C43.

[4] R. M. Alt, “Service-oriented architecture for automated commissioning of fluid power systems,” Shaker Verlag, 2023. Accessed: Feb. 23, 2025. [Online]. Available: https://publications.rwth-aachen.de/record/958758

[5] “Industrie 4.0-Verwaltungsschale für Pumpen.” Accessed: Mar. 01, 2025. [Online]. Available: https://www.th-koeln.de/anlagen-energie-und-maschinensysteme/verwaltungsschale-pumpen-2_115652.php

[6] “Companion Specification for Pumps.” Accessed: Mar. 01, 2025. [Online]. Available: https://www.th-koeln.de/anlagen-energie-und-maschinensysteme/companion-specification-for-pumps_115547.php

[7] OPC 40223, *OPC UA for Pumps and Vacuum Pumps*, Apr. 19, 2021. Accessed: Feb. 06, 2025. [Online]. Available: https://opcfoundation.org/developer-tools/documents/view/244

[8] “Companion Specification for CAS.” Accessed: Mar. 01, 2025. [Online]. Available: https://www.th-koeln.de/anlagen-energie-und-maschinensysteme/companion-specification-for-cas_115546.php

[9] OPC 40250-1, *Compressed Air Systems - Main Control System*, Jul. 03, 2021. Accessed: Mar. 01, 2025. [Online]. Available: https://opcfoundation.org/developer-tools/documents/view/245

[10] M. Both, J. Müller, and C. Diedrich, “Automatisierte Abbildung semantisch heterogener I4.0-Verwaltungsschalen durch Methoden des Natural Language Processing,” - *Autom.*, vol. 69, no. 11, pp. 940–951, Nov. 2021, doi: 10.1515/auto-2021-0050.

[11] S. Heppner, M. A. Both, T. Miny, J. Müller, C. Diedrich, and T. Kleinert, “Semantic Matching Interface für semantisch heterogene Industrie 4.0 Verwaltungsschalen,” - *Autom.*, vol. 72, no. 10, pp. 919–927, Oct. 2024, doi: 10.1515/auto-2023-0225.

[12] A. Cartus, M. Both, N. Maisch, J. Müller, and C. Diedrich, “Interoperability of semantically heterogeneous digital twins through Natural Language Processing methods,” in *CLIMA 2022 conference*, May 2022. doi: 10.34641/clima.2022.143.

[13] J. Beermann, M. Both, C. Block, J. Nagel, and J. Müller, “Erstellung eines semantischen Suchalgorithmus zur Abbildung proprietärer Merkmale auf den ECLASS-Standard,” Jul. 2024.

[14] “ZVEI-Show-Case PCF@ControlCabinet,” ZVEI. Accessed: Apr. 27, 2025. [Online]. Available: https://www.zvei.org/en/subjects/zvei-show-case-pcfcontrolcabinet

[15] N. Klein, J. Gerner, A. Wang, M. Herm, F. Küßler, and J. Fleischer, “Implementation of an Information Model for Electric Motor Production Using the Asset Administration Shell,” in *2024 14th International Electric Drives Production Conference (EDPC)*, Nov. 2024, pp. 1–7. doi: 10.1109/EDPC63771.2024.10932846.

[16] S. Cavalieri and S. Gambadoro, “Digital Twin of a Water Supply System Using the Asset Administration Shell,” *Sensors*, vol. 24, no. 5, Art. no. 5, Jan. 2024, doi: 10.3390/s24051360.

[17] IDTA 02006-3-0, *Digital Nameplate for Industrial Equipment*, Frankfurt Am Main, Germany., Nov. 2024. [Online]. Available: https://industrialdigitaltwin.org/en/wp-content/uploads/sites/2/2024/11/IDTA-02006-3-0_Submodel_Digital-Nameplate.pdf

[18] IDTA 02005-1-0, *Provision of Simulation Model*, Frankfurt Am Main, Germany., Dec. 2022. [Online]. Available: https://industrialdigitaltwin.org/en/wp-content/uploads/sites/2/2023/01/IDTA-02005-1-0_Submodel_ProvisionOfSimulationModels.pdf

[19] IDTA 02011-1-1, *Hierarchical Structures enabling Bills of Material*, 06/24. Accessed: Feb. 23, 2025. [Online]. Available: https://industrialdigitaltwin.org/wp-content/uploads/2024/06/IDTA-02011-1-1_Submodel_HierarchicalStructuresEnablingBoM.pdf

[20] IDTA 01001-3-0-1, *Specification of the Asset Administration Shell Part 1: Metamodel*, IDTA 01001-3-0-1, Frankfurt Am Main, Germany., Jun. 01, 2024. [Online]. Available: https://industrialdigitaltwin.org/wp-content/uploads/2024/06/IDTA-01001-3-0-1_SpecificationAssetAdministrationShell_Part1_Metamodel.pdf

[21] DIN SPEC 91345, *Reference Architecture Model Industrie 4.0 (RAMI4.0) — English translation of DIN SPEC 91345:2016-04*, Apr. 2016.

[22] Industrial Digital Twin Association, “General :: Asset Administration Shell Documentation.” Accessed: Feb. 28, 2025. [Online]. Available: https://admin-shell-io.github.io/aas-specs-antora/IDTA-01001/v3.1/annex/general.html#types-and-instances

[23] E. Tantik and R. Anderl, “Potentials of the Asset Administration Shell of Industrie 4.0 for Service-Oriented Business Models,” *Procedia CIRP*, vol. 64, pp. 363–368, Jan. 2017, doi: 10.1016/j.procir.2017.03.009.

[24] R. H. L. Chiang, T. M. Barron, and V. C. Storey, “Data, Knowledge & Information in Database and Knowledge-Based Systems,” *J. Database Manag. JDM*, vol. 3, no. 3, pp. 12–20, Jul. 1992, doi: 10.4018/jdm.1992070102.

[25] K. North, *Wissensorientierte Unternehmensführung: Wissensmanagement im digitalen Wandel*. Wiesbaden: Springer Fachmedien, 2021. doi: 10.1007/978-3-658-32771-2.

[26] F. Bodendorf, *Daten- und Wissensmanagement*, 2., Aktualisierte u. erw. Auf. Berlin: Springer Berlin, 2005.

[27] ISO 1219-1, *Fluid power systems and components — Graphical symbols and circuit diagrams — Part 1: Graphical symbols for conventional use and data-processing applications*, Jun. 2012.

[28] DIN ISO 1219-1, *Fluidtechnik – Graphische Symbole und Schaltpläne Teil 1: Graphische Symbole für konventionelle und datentechnische Anwendungen*, Jan. 2019.

[29] ISO 1219-2, *Fluid power systems and components — Graphical symbols and circuit diagrams — Part 2: Circuit diagrams*, Sep. 2012.

[30] ISO 1219-3, *Fluid power systems and components — Graphical symbols and circuit diagrams — Part 3: Symbol modules and connected symbols in circuit diagrams*, Sep. 2016.

[31] Bosch Rexroth AG, *LUDV control block in mono block/sandwich plate design M7-20*. 2013.

[32] W. Wang, A. Tolk, and W. Wang, “The Levels of Conceptual Interoperability Model: Applying Systems Engineering Principles to M&S,” presented at the Spring Simulation Multiconference (SpringSim’09), San Diego, USA, 2009, p. 9. doi: 10.1145/1639809.1655398.

[33] Parker Hannifin Corporation, *P2/P3 Series Piston Pumps Variable Displacement*. Marysville, Ohio USA, 2009.

[34] Bosch Rexroth AG, *Axial piston variable pump A10VG series 60*. Elchingen, Germany, 2024.

[35] ECLASS, “ECLASS Datenstandard für Waren & Dienstleistungen,” ECLASS. Accessed: Feb. 04, 2025. [Online]. Available: https://eclass.eu/

[36] “IEC - Common Data Dictionary (CDD).” Accessed: Feb. 04, 2025. [Online]. Available: https://cdd.iec.ch/cdd/iec61360/iec61360.nsf/TreeFrameset?OpenFrameSet

[37] ECLASS, “ECLASS e.V. members,” ECLASS. Accessed: Feb. 10, 2025. [Online]. Available: https://eclass.eu/en/about-us/members/eclass-ev-members

[38] IDTA 01001-3-a-0-2, *Specification of the Asset Administration Shell Part 3a: Data Specification – IEC 61360*, Frankfurt Am Main, Germany., Jul. 01, 2024. [Online]. Available: https://industrialdigitaltwin.org/en/wp-content/uploads/sites/2/2024/07/IDTA-01003-a-3-0-2_SpecificationAssetAdministrationShell_Part3a_DataSpecification_IEC613601.pdf

[39] A. L. Opdahl, “A Platform for Interoperable Domain-Specific Enterprise Modelling Based on ISO 15926,” in *2010 14th IEEE International Enterprise Distributed Object Computing Conference Workshops*, Oct. 2010, pp. 301–310. doi: 10.1109/EDOCW.2010.29.

[40] ISO 15926-2, *Industrial automation systems and integration — Integration of life-cycle data for process plants including oil and gas production facilities — Part 2: Data model*, Dec. 15, 2003.

[41] B. Kim *et al.*, “Use of Asset Administration Shell Coupled with ISO 15926 to Facilitate the Exchange of Equipment Condition and Health Status Data of a Process Plant,” *Processes*, vol. 10, p. 2155, Oct. 2022, doi: 10.3390/pr10102155.

[42] M. T. Delgado, “Eclipse Semantic Modeling Framework (ESMF) | projects.eclipse.org.” Accessed: Feb. 05, 2025. [Online]. Available: https://projects.eclipse.org/projects/dt.esmf

[43] M. Milinkovich, “About the Eclipse Foundation | The Eclipse Foundation.” Accessed: Feb. 10, 2025. [Online]. Available: https://www.eclipse.org/org/

[44] “Semantic Aspect Meta Model (SAMM) : Bosch Semantic Stack Documentation.” Accessed: Feb. 09, 2025. [Online]. Available: https://docs.bosch-semantic-stack.com/oss/samm-specification.html

[45] “schema.org.” Accessed: Feb. 05, 2025. [Online]. Available: https://schema.org/

[46] P. A. J. Achten, Z. Fu, and G. E. M. Vael, “Transforming future hydraulics : a new design of a hydraulic transformer,” *Fifth Scand. Int. Conf. Fluid Power SICFP 97 Swed. Linköping*, p. 24, 1997.

[47] S. Bader *et al.*, “What is the AAS.” Accessed: Feb. 09, 2025. [Online]. Available: https://www.plattform-i40.de/IP/Redaktion/EN/Downloads/Publikation/2021_What-is-the-AAS.pdf?__blob=publicationFile&v=1

[48] C. Schmidt and J. Müller, “GitHub - FraunhoferIOSB/EDC-Extension-for-AAS: EDC Extension supporting usage of Asset Administration Shells.” Accessed: Feb. 09, 2025. [Online]. Available: https://github.com/FraunhoferIOSB/EDC-Extension-for-AAS?tab=readme-ov-file

[49] “History,” OPC Foundation. Accessed: Feb. 06, 2025. [Online]. Available: https://opcfoundation.org/about/opc-foundation/history/

[50] OPC 10000-8, *OPC Unified Architecture - Part 8: Data Access*, Oct. 05, 2024. Accessed: Feb. 06, 2025. [Online]. Available: https://opcfoundation.org/developer-tools/documents/view/165

[51] OPC 10000-100, *OPC Unified Architecture - Part 100: Devices*. Accessed: Feb. 06, 2025. [Online]. Available: https://opcfoundation.org/developer-tools/documents/view/197

[52] OPC 40400-1, *OPC UA for the Powertrain - Part 1: Asset ManagementPower*, Nov. 01, 2024. Accessed: Feb. 06, 2025. [Online]. Available: https://opcfoundation.org/developer-tools/documents/view/240

[53] S. Köhler and J. Weber, "Integration of fault detection and diagnosis methods into OPC UA for mobile machinery," 2022, pp. 457–472.

[54] B. Beck *et al.*, "Bauen 4.0 - Effizientere und produktivere Bauprozesse durch Vernetzung und Kommunikation mobiler Maschinen : Abschlussbericht : Bauen 4.0 Bauprozesse - Vernetzung - Maschinen : Laufzeit des Vorhabens: 01.07.2019 bis 31.12.2022," Dresden, 2023. doi: 10.2314/KXP:1881651258.

[55] S. Köhler *et al.*, "Digital assisted collision avoidance for mobile machinery," in *14th International Fluid Power Conference*, Dresden, Germany: River Publishers, 2024, pp. 923–943. doi: 10.13052/rp-9788770042222C73.

[56] "FAQ." Accessed: Feb. 10, 2025. [Online]. Available: https://mqtt.org/faq/

[57] OASIS Open, "MQTT Version 5.0." Mar. 07, 2019. Accessed: Feb. 10, 2025. [Online]. Available: https://docs.oasis-open.org/mqtt/mqtt/v5.0/mqtt-v5.0.pdf

[58] "Functional Mock-up Interface." Accessed: Apr. 27, 2025. [Online]. Available: https://fmi-standard.org/

[59] "Tools - Functional Mock-up Interface." Accessed: Apr. 27, 2025. [Online]. Available: https://fmi-standard.org/tools/

[60] "About - Functional Mock-up Interface." Accessed: Apr. 27, 2025. [Online]. Available: https://fmi-standard.org/about/

[61] "Modelica." Accessed: Feb. 07, 2025. [Online]. Available: https://modelica.org/association/#organizational-members-of-the-modelica-association-listed-alphabetically

[62] D. Winkler, P. Harman, M. Tiller, K. Davies, and M. Sjölund, "GitHub - modelica-3rdparty/OpenHydraulics: A free Modelica package providing components describing 1-dimensional fluid flow in hydraulic circuits." Accessed: Feb. 07, 2025. [Online]. Available: https://github.com/modelica-3rdparty/OpenHydraulics

[63] M. Eigner and R. Stelzer, *Product Lifecycle Management – Ein Leitfaden für Product Development und Life Cycle Management*. 2008.

[64] ISO/IEC/IEEE 24748-2, *Systems and software engineering — Life cycle management — Part 2: Guidelines for the application of ISO/IEC/IEEE 15288 (system life cycle processes)*, Mar. 2024.

[65] ISO 15226, *Technical product documentation — Life cycle model and allocation of documents*, Apr. 01, 1999.

[66] R. Heidel, M. Hoffmeister, M. Hankel, and U. Döbrich, *Industrie 4.0 – Basiswissen RAMI4.0*. 2017.

[67] IDTA 02026-1-0, *Provision of 3D Models*, Jun. 2024. [Online]. Available: https://industrialdigitaltwin.org/wp-content/uploads/2024/06/IDTA_02026-1-0_Submodel_ProvisionOf3DModels.pdf

[68] EPLAN GmbH & Co. KG, "EPLAN – efficient engineering." Accessed: Feb. 12, 2025. [Online]. Available: https://www.eplan.de/

Data-Driven Serial Testing of Axial Piston Units: Integrating Signal Processing and Machine Learning for Roller Bearing Fault Detection

Maximilian Romeser[1,2] and Reinhold von Schwerin[1]

[1]Faculty for Informatics, Ulm University of Applied Sciences, Ulm, Germany
[2]Development of Axial Piston Units, Bosch Rexroth AG, Elchingen, Germany
E-Mail: maximilian.romeser@boschrexroth.de, reinhold.vonschwerin@thu.de

Abstract

Data-driven fault detection has emerged as an effective method for early anomaly detection and failure prevention, particularly through vibroacoustic analysis and the use of Machine Learning (ML) or Artificial Intelligence (AI). Given the central role of hydrostatic drives and axial piston units (APUs) in both conventional and electrified powertrains, their reliability is pivotal to prevent costly maintenance and machine downtime. Integral to the functionality of APUs are roller bearings, whose reliability and performance significantly influence the overall efficiency and longevity of the machinery. Given this context, the detection of faults in roller bearings within APUs is crucial to ensure operational safety and extend the lifespan of the systems. This is particularly important in the realms of condition monitoring and predictive maintenance, as well as serial testing, which is essential for comprehensive quality control, especially when considering alternative bearings. Therefore, advanced methods for detecting faulty roller bearings are needed, but challenges include hydraulically induced noise and manufacturing-related serial dispersion in serial testing scenarios. This paper presents an experimental investigation into roller bearing fault detection through vibroacoustic analysis. The study emphasizes the integration of advanced signal processing techniques and semi-supervised anomaly detection methods, utilizing various machine learning algorithms. Additionally, the research explores optimal sensor selection and operating conditions. The findings demonstrate that, under appropriate operating conditions and with the support of empirical knowledge, robust fault diagnosis can be achieved in the presented serial testing scenario.

Keywords: Axial Piston Units, Fault Detection, Anomaly Detection, Vibroacoustic Analysis, Machine Learning, Condition Monitoring, Serial Testing, Roller Bearings

1 Introduction

The robustness and high power density of hydrostatic drivetrains have long established them as a preferred choice in off-road applications. These systems are particularly valued for their precise control accuracy, high dynamics, and the ability to provide stepless transmission. Additionally, hydrostatic drives enable linear movements for various work functions, making them indispensable in demanding environments [1]. Given these advantages, hydrostatic drives, and APUs in particular, play a central role in the powertrain, whether it is a conventional drivetrain with an internal combustion engine or an electrified drivetrain, as shown in [2]. Therefore, the reliability of APUs is vital to prevent costly maintenance and machine downtime. As a result, methods for fault detection and diagnosis are the focus of various research efforts, primarily to enable condition monitoring (CM) and predictive maintenance, which are already well-established for gearboxes and turbines [3]. Additionally, robust and automated fault detection can serve as a valuable quality control measure during serial testing, particularly when products are in the ramp-up phase of production. Monitoring changes in suppliers is also crucial in this context.

Regarding the reliability of rotating machinery in general, and especially APUs [4], roller bearings are essential components [5]. Consequently, the detection of faulty bearings constitutes a significant area of research. In this context, the analysis of vibroacoustic behavior has emerged as an effective tool, enabling the detection of fault-related geometric deviations through anomalous excitations during operation [5]. This allows the detection of most failures in rotating machinery at an earlier stage than by monitoring temperatures or other

signals [3, 6] However, in the case of axial piston units, this is particularly challenging compared to other rotating machinery. The abrupt impulses induced by the reciprocating pistons are extremely dominant, making it difficult to detect fault-induced anomalies in the emitted structure-borne or airborne noise [4]. Furthermore, when considering scenarios for fleet-based CM [7] or serial testing processes, it is important to account for serial dispersion. Conventional CM approaches typically assume a nominal reference state for the machine under consideration, where the signal behavior is initially stable, and faults manifest as anomalies relative to this reference. However, in the aforementioned scenarios, such an approach is not feasible due to the short testing duration, varying operating conditions and the potential presence of initial faults. Instead, anomalies in the data of new test units must be evaluated in relation to the historical data of previous, healthy units. Due to serial dispersion, the vibroacoustic behavior also varies, and fault-induced anomalies must be sufficiently pronounced to stand out against this variability [8].

Given the complexity of the vibroacoustic and structural dynamic behavior of APUs, this study aims to develop a data-driven approach based on measurements from an ensemble of units. The approach follows the concept of semi-supervised anomaly detection and falls within the field of ML and AI. An ensemble of tested, healthy units will be used to characterize production variability and train the model, while various faulty units will be evaluated during model testing. Since it is not straightforward to determine how fault patterns manifest in vibroacoustic behavior, different signal processing and feature generation methods will be investigated. The goal is to extract fault-relevant information from the data to ensure robust, data-based fault detection. Since vibroacoustic behavior can be highly dependent on operating conditions and sensor application, these factors will be comprehensively evaluated to provide a holistic assessment of fault detection and to determine the optimal conditions for serial testing processes. To the best of the author's knowledge, no study has yet explored the data-driven detection of faulty roller bearings in an ensemble of APUs. Therefore, the contributions of this study are as follows:

- Proposal and evaluation of a semi-supervised approach for detecting APUs with faulty roller bearings in a serial testing scenario using ML
- Evaluation of the combination of various approaches for signal processing and feature generation under various operation conditions
- Comparison of different sensor applications for analysis of vibroacoustic behavior of APUs

2 State of the Art

Fault detection in axial piston units has become a prominent field of research in recent years, driven by advancements in ML and signal processing, as shown in [9]. Typically, this research is conducted within the framework of CM or in a more general context. Only Gaugel et al. [10] consider a serial testing scenario of APUs, without delving into the specific faults and signals considered and instead focusing on time series segmentation. In the following, the authors of this paper limit themselves to approaches that address vibroacoustic analysis. In particular, data-driven approaches including signal processing and ML techniques are frequently discussed, as model-based approaches are very challenging due to the system's complex vibroacoustic behavior [11]. A comprehensive study on the use of conventional ML in terms of supervised classification for CM scenarios is provided by Torrika [12], who particularly highlights the steps of feature extraction and reduction. An analogous approach to feature generation in a supervised classification setting can be found in [13]. Further intensive research on the condition monitoring (CM) of axial piston pumps using supervised methods such as Support Vector Machines (SVMs) and gradient boosting algorithms, utilizing a feature engineering approach based on tsfresh, was conducted by Horn et al. [14]. In contrast, Gnepper et al. [15] and Siyuan et al. [16] explore a semi-supervised anomaly detection scenario and employ principal component analysis to characterize the nominal state of the APU within the context of CM.

To better highlight fault information, signal processing techniques such as Fast Fourier Transformation (FFT) [15, 17, 18], Wavelet Packet Transformation (WPT) [9, 16, 17], Empirical Mode Decomposition (EMD) [18], Cepstrum Analysis [11], or Time Synchronous Averaging [13] are frequently employed. Specifically, in the detection of faulty bearings in APUs, various techniques are commonly employed to generate sensitive filter bands and enhanced signal envelopes [4,19-22]. For the fault detection of worn slippers, the principle of cyclostationarity plays a particularly important role here, as investigated by Casoli et al. [23]. By considering the results in [17], it can be observed that commonly studied fault patterns, such as wear on slippers, port plates or cylinders, can have immediate effects on the dominant harmonics of the piston frequency. However, this is not the case for roller bearings, which imposes different requirements on signal processing. Advanced signal processing techniques also support the application of Deep Learning, which has been extensively studied in the context of fault detection for APUs in recent years [24-31]. As such, model training typically involves feature extraction based on scalograms or spectrograms, rather than directly utilizing the raw time series data as in [31]. However, being supervised classification settings the foundation of these approaches relies on the identification of known, specific fault states resulting from wear and damages, as well as on sufficient data volumes without considering serial dispersion due to manufacturing tolerances. As Keller et al. [32] emphasize, this must be ensured for adequate validation. Nevertheless, in the case of CM, these factors are frequently irrelevant, since the goal is to identify faults relative to the normal condition of the same unit, as mentioned before. This contrasts with the present use-case.

Although the developed methods are often tested at various operating conditions, only Horn et al. [14] and Gnepper et al. [15] investigate a wide operating range and determines the quality of fault detection with respect to slipper clearance or also slipper wear and valve plate cavitation [14] depending on this range. Additionally, while the experimental setups of some of the mentioned studies include multiple structure-borne or airborne sound sensors, such as in [4] or [9], this

aspect is not always analyzed in detail. Casoli et al. [13, 17, 23] and Du et al. [11], for instance, only consider structure-borne sound in axial and radial directions, whereas Gnepper et al. [15] analyze different sensor concepts and positions. Furthermore, the investigations in [32] indicate that positioning the sensor close to the bearing is effective, even when the valve plate is damaged. Other studies address the fusion of multi-sensor data, including acoustic signals [31] or pressure signals [24].

3 Experimental Design

To replicate a scenario of serial testing with an ensemble of both healthy and faulty units, an experimental series is conducted on a laboratory test bench. The experimental design, including the materials and the test sequence, as well as the data acquisition setup is described below.

3.1 Material and Fault Types

In this study, APUs in bent axis design were tested using a laboratory hydraulic test bench, as shown in Figure 1. Each APU was initially tested in its original delivery condition. Subsequently, to characterize the range of vibroacoustic variation attributable to component interchangeability, the drive shafts were systematically interchanged among the housings. For example, in the first round of swaps, the drive shaft from the first APU was placed into the second housing, the second unit's drive shaft into the third housing, and so on. While a full exploration of all 49 possible unique pairwise combinations was desirable, time constraints limited this cross-assembly testing, constituting the first phase of the experimental series, to 30 unique drive shaft-housing combinations (achieved through 23 swaps). Throughout this phase, the drive assemblies remained equipped with their original, undamaged bearings, which were not removed. This approach enabled assessment of vibroacoustic variability while isolating the effects of drive shaft and housing variations.

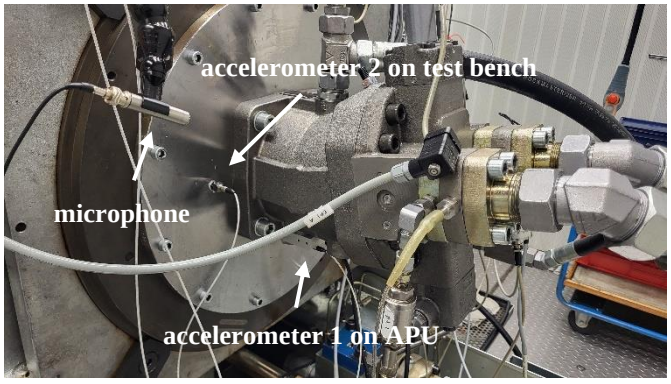


Figure 1: Experimental setup at the laboratory test bench

In the second phase of the experimental series, 14 drive assemblies with faulty roller bearings are installed. In 12 cases, this relates to the large roller bearing and in two cases the small roller bearing. Faults in other components of the drive assembly are neglected for the purposes of this investigation, meaning that these components remain in a healthy condition. The types of faults in the tapered roller bearings are

Figure 2: Manipulated roller elements (left) and outer ring (right) with abrasively generated spalls

intentionally chosen to be heterogeneous to generate a broad portfolio of potential anomalies in the data and to ensure comprehensive validity. The bearings are manipulated, with different sections of the raceways or end faces of the rolling elements abrasively processed, as exemplarily depicted in Figure 2. The defects vary in depth and width, as well as in number and location of the defects. Regarding further investigation, it is crucial to generate both easily detectable faults (e.g., signals with significantly higher signal energy) and subtle faults, identifiable only by the presence of small peaks in the frequency spectrum, to evaluate the robustness of the fault detection method. Due to modulation effects resulting from tolerances in the APU, various effects of impact excitations on the structure-borne sound emissions are observed. It should be noted that due to the cage, subsequent processing of the inner ring is not possible. Each faulty drive assembly is mounted once in one of the housings, and the assembled unit is then tested on the test bench. The assignment of faulty drive assemblies to housings is done randomly.

3.2 Test Sequence

Each unit or drive assembly-housing combination, regardless of the presence of a fault, undergoes the same automated test sequence on the test bench for high reproducibility. Initially, a 15-minute run-in period is conducted to achieve an oil temperature level of 40 °C. Since the regulation is not perfect, fluctuations can occur during operation, especially when load points are varied. In the end, the overall oil temperature range, across all following load points, was 42 °C to 56 °C. For individual load points, the experimentally determined temperature variation did not exceed 8 °C. It is plausible that these fluctuations, particularly in relation to oil viscosity, may influence the vibroacoustic emissions. However, this can also occur in real serial testing or fleet-based CM scenarios, and thus, it is treated as regular process variation.

The test procedure involves the traversal of a characteristic map within a three-dimensional operating range. An operating point is defined by the shaft speed, the pressure in the high-pressure line, and the swivel angle. The load in the closed circuit is adjusted in four stages between 100 bar and 400 bar using a load valve. The speed is varied in five equally sized stages between 400 and 2000 rpm. The swivel angle is set to either its minimum or maximum position. The swivel angle as well as the high-pressure influence the load on the bearing, which is why these parameters are referenced accordingly in the Table 1 for clarity. Consequently, the resulting section of the test sequence includes 40 load points. Between the load

points, the operating conditions are adjusted in a ramp-like manner. Each load point is then held for ten seconds.

Table 1: Indexing of the different bearing loads (0: Min. swivel angle, 1: Max. swivel angle)

Bearing Load	1	2	3	4	5	6	7	8
Pressure [bar]	100	200	300	400	100	200	300	400
Swivel Angle	0	0	0	0	1	1	1	1

3.3 Sensors and Data Acquisition

To capture the vibroacoustic behavior, three sensors were applied: Firstly, a free-field microphone is permanently installed to measure airborne noise, as shown in Figure 1, at approximately 30 cm above the APU. Secondly, a triaxial piezo-electric accelerometer is permanently attached to the test bench using an adhesive adapter to capture structure-borne noise emissions (accelerometer 2). Since the sensor application is performed just once before the test series, high reproducibility of the results is ensured. This permanent application is particularly advantageous for serial testing, as it eliminates the need for manual sensor placement by the operator or automatic positioning systems. Additionally, it is beneficial to measure as close as possible to the relevant source of structure-borne noise, i.e., the roller bearing, to achieve a high signal-to-noise ratio (SNR). For this purpose, another accelerometer of the same type is permanently mounted using an anodized screw adapter, as shown in Figure 3 (accelerometer 1). The adapter is attached to the bolting of the bearing cross-flushing and is therefore reattached with each new unit mounted on the test bench, i. e. for each test. Consequently, this has a negative impact on the dispersion between the conducted tests.

The sampling rate is consistently set to 50 kHz. To avoid aliasing effects, a Butterworth low-pass filter with a cutoff frequency of 25 kHz is employed. Data acquisition is performed in intervals and controlled by an automated test sequence. After a stabilization period of three seconds, interval measurements of five seconds duration are conducted at each of the 40 load points. For further investigations, the signal channel and the time series from the microphone are referenced as Lp, the channels from accelerometer 1 as a1x, a1y and a1z, and those from accelerometer 2 as a2x, a2y and a2z. Consequently, the relevant dataset consists of 40 x 7 time series, e. g. data instances, for each of the 44 tested units, with each time series comprising 250000 samples. Consequently, there are 40 x 7 data subsets including 44 instances per each.

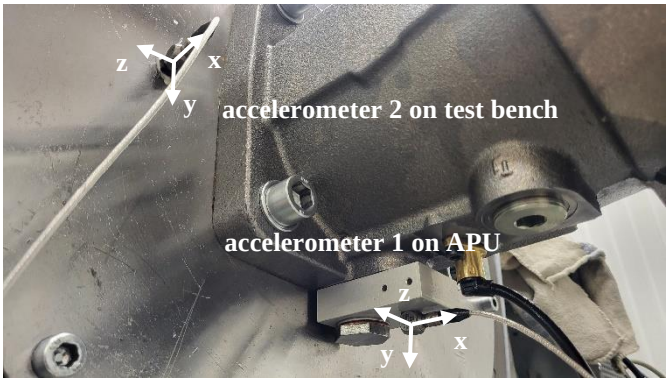


Figure 3: Detailed view of the attached accelerometers

4 Data-Driven Methodology

The methodology employed in this paper for generating a data-driven model for fault detection consists of three main steps: signal processing, feature generation, and anomaly detection, with the latter encompassing the actual ML. As illustrated in Figure 4, various methods for signal processing and different feature sets are utilized. The features thus serve as the input for the semi-supervised anomaly detection model, enabling it to learn the characteristics of vibroacoustic data from healthy units. While the features can be directly applied to the raw time series data $x(t)$, it is often more effective to precede the feature generation process with signal processing methods, as demonstrated by Bienefeld et al. [33] in a comprehensive study on this subject. To compare these methods while considering the dependency on the operating conditions and sensors, an appropriate metric is used to evaluate the anomaly detection and thus the quality of fault detection.

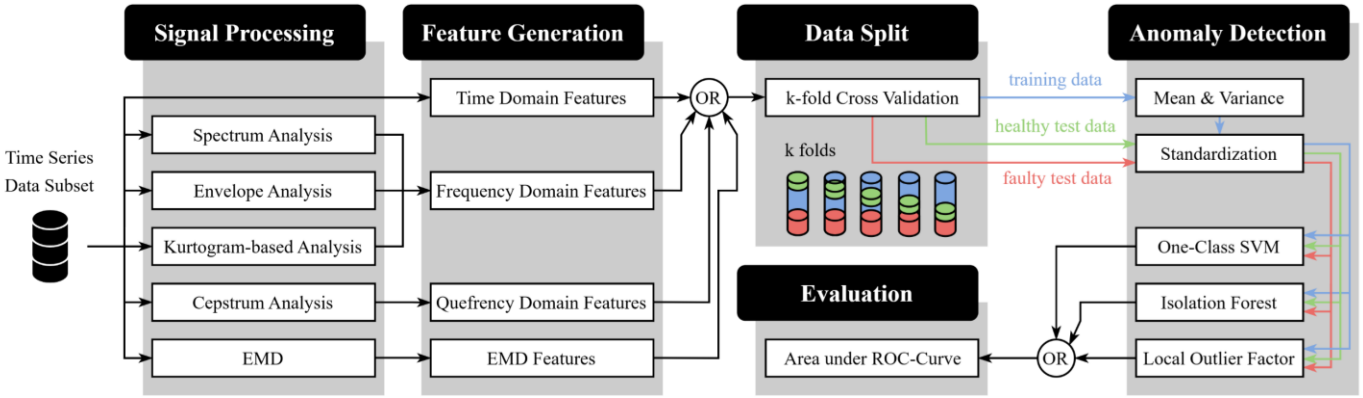


Figure 4: Integration of the data-driven methodology. The data pipeline enables the variation of signal processing techniques and their corresponding features as well as the variation of the anomaly detection model (OR gates). Note, that the pipeline is set up for each individual data subset regarding different operating conditions and sensor signals.

4.1 Signal Processing

In the scope of this work, five signal processing methods are investigated: spectrum analyses in frequency domain (FD) which includes the analysis of the raw spectrum, the envelope spectrum and the kurtogram-based envelope spectrum, cepstrum analysis and EMD. Additionally, the direct generation of features from the time series data is considered, too. A deeper understanding of the methods is provided by Randall et al. [5]. A low-pass filter with a cutoff frequency of 3 kHz was selected based on visual inspection of the raw time series data and consideration of the dominant frequencies observed in preliminary testing. Above 3 kHz, the spectra were dominated by broad-band elevations indicative of test bench and sensor coupling resonances, particularly pronounced for accelerometer 1. These resonances varied significantly between individual units, potentially masking the APU-specific vibroacoustic characteristics of interest. The 0-3 kHz range captured the most prominent features related to APU operation, as indicated by Figure 6.

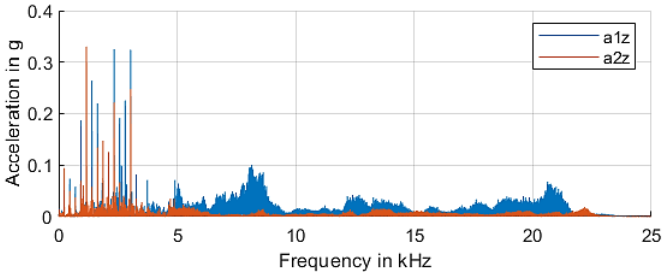


Figure 6: Entire spectra of a healthy unit in z-direction before low-pass filtering

4.1.1 Spectrum Analyses

Bearing faults often generate characteristic vibrations at specific frequencies depending on the type of fault, operating parameters, and bearing dimensions. By applying FFT to the vibroacoustic data, these frequency components can be identified in the frequency domain, i. e. in the raw spectrum $X(f) = \mathcal{F}\{x(t)\}$.

As the raw signal spectrum often provides limited diagnostic information about bearing faults, envelope analysis has become an established method for bearing diagnostics [5]. The envelope-based method exploits the fact that a high-frequency carrier signal, which reflects the structural resonances and is influenced by other excitations, is amplitude modulated by the recurring impulses caused by bearing faults. The relevant information in the raw signal $x(t)$ can thus be identified through demodulation, where the envelope $e(t)$ can be generated using the Hilbert transform $\mathcal{H}\{x(t)\}$ as follows:

$$e(t) = \sqrt{x(t)^2 + \mathcal{H}\{x(t)\}^2} \tag{1}$$

and the FFT of this envelope reveals the envelope spectrum $\mathcal{F}\{e(t)\}$, as exemplarily shown in Figure 5. It should be noted, however, that the envelope can be significantly affected by piston frequencies, as demonstrated by Xiao et al. [4]. As mentioned by Randall et al. [5], it thus makes sense to identify frequency bands that are primarily dominated by bearing resonances before demodulation. In the context of this work, these were determined empirically, and a bandpass filter with a passband of 2 to 3 kHz was chosen for the envelope-based analysis. This contrasts with the purely FFT-based approach, which examines the entire frequency band up to 3 kHz without performing envelope analysis.

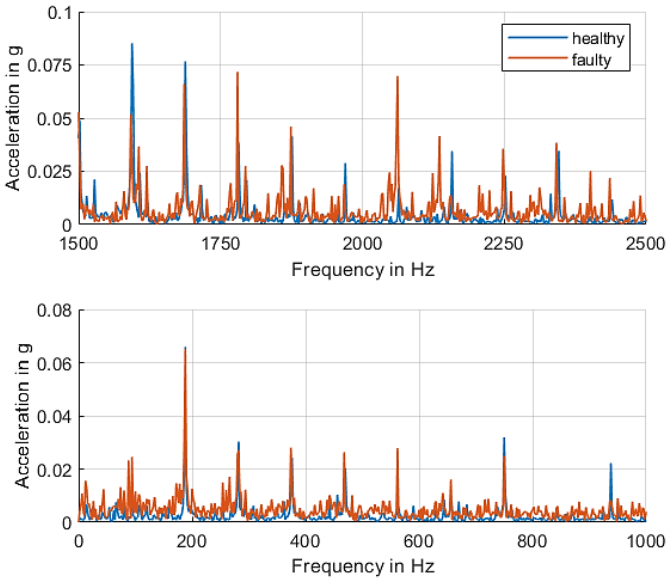


Figure 5: Raw spectrum (top) and envelope spectrum (bottom) of a healthy and faulty unit considering a2y

Nevertheless, the identification of a suitable filter is generally non-trivial. However, spectral kurtosis provides a tool that relates the impulsiveness of the signal to its frequency components. The underlying hypothesis is that frequency bands containing information relevant to bearing fault detection exhibit higher spectral kurtosis. For simplicity, a detailed explanation of the method is omitted here, and instead, explicit reference is made to [5]. However, it is important to understand that a window is slid over the time series and a FFT is performed for each window, resulting in the Short-Time Fourier Transformation (STFT) and resulting in the spectrogram $S(t, f)$. This process transforms the signal into the time-frequency domain. The spectral kurtosis $K(f)$ is calculated as

$$K(f) = \langle \frac{|S(t,f)|^4}{|S(t,f)|^2} \rangle - 2 \tag{2}$$

with $\langle \cdot \rangle$ characterizing the time-averaging operator. Frequency ranges with high spectral kurtosis are then suitable filter bands. Determining these ranges depends on the ratio of the window length to the unknown spacing between the periodically occurring impulses. Therefore, the kurtogram is constructed, which further represents $K(f)$ as a function of the window length. Due to the time-frequency uncertainty principle, the frequency resolution increases with longer window lengths.

4.1.2 Cepstrum Analysis

Another method to reduce the influence of disturbances in relation to fault-induced anomalies is cepstrum analysis. In the cepstrum, families of periodically spaced spectral components appear as prominent peaks, known as rahmonics, which include both harmonics and sidebands in the spectrum. These

phenomena are caused by impulsive forces due to roller element bearing faults combined with amplitude modulations resulting from phase shifts between the load and the fault positions [34]. According to Du et al. [11], this approach has been successfully applied in the context of bearing fault detection in APUs.

The cepstrum $C(T)$ is generated by first transforming the signal into the frequency domain using the FFT. After logarithmizing the spectral amplitudes, the signal is then transformed into the quefrency domain (QD) with $T = 1/f$, which can be expressed as

$$C(T) = |\mathcal{F}^{-1}\{\log|\mathcal{F}\{x(t)\}|^2\}|^2. \quad (3)$$

In simple terms, the cepstrum represents the inverse transformation of the logarithmized power spectrum $\log|\mathcal{F}\{x(t)\}|^2$, thereby presenting the cepstrum in relation to the quefrency components in time domain (TD) once again. The reciprocal quefrency T provides insights into dominant families of frequency components.

4.1.3 Empirical Mode Decomposition

A purely data-driven and heuristic approach to signal processing is EMD, which enables the decomposition of $x(t)$ into N so-called Intrinsic Mode Functions (IMFs) and a residual $r_N(t)$:

$$x(t) = \sum_{i=1}^{N} IMF_i(t) + r_N(t) \quad (4)$$

The IMFs characterize different frequency bands, which is particularly advantageous for diagnosing faulty roller bearings when there is no prior knowledge of relevant bands or the signal structure. To determine the individual IMFs, the algorithm iteratively decomposes $x(t)$ in TD through a process called sifting, which proceeds as follows:

1. Find the local minima and maxima of $x(t)$
2. Construct lower and upper envelopes $v_1^-(t)$ and $v_1^+(t)$ by using the local extrema
3. Calculate the mean of envelopes $m_1(t)$
4. Subtract $m_1(t)$ from $x(t)$ to obtain residual $r_1(t)$

This process is repeated N times with $x(t) := IMF_i(t) = r_i(t)$ till a specific condition is met: In this study, the process is completed after $N = 5$ repetitions, resulting in four IMFs and the final residual $r_5(t)$.

4.2 Feature Generation

Given the diverse nature of fault cases, it is challenging to generate fault-related features in advance, as these may not be suitable for extracting the corresponding anomalies from the signal in every fault scenario. It is worth mentioning that, even for roller bearings alone, a multitude of fault cases is possible, which cannot all be considered within the scope of this study. Therefore, it is even more important to develop appropriate features. If the features are to be derived directly from the time series, e. g. one data instance, statistical metrics are used for this purpose. This includes the first four statistical moments: mean $\mu_x = F_{TD_1}$, variance $\sigma_x^2 = F_{TD_2}$, skewness F_{TD_3}, and kurtosis F_{TD_4}, with the latter two defined as

$$F_{TD_i} = \mathrm{E}\left[\left(\frac{x-\mu}{\sigma}\right)^i\right] \; \forall\, i \in \{3,4\}. \quad (5)$$

Depending on the amplitude, outliers receive a higher weight. Additionally, the Root-Mean-Square (RMS) and the median are used as features F_{TD_5} and F_{TD_6}. Analogous to [32], it makes sense to relate features in FD (due to preprocessing with the methods from sections 4.1.1 to 4.1.4) to specific frequency bands. This approach allows, for example, the isolation of bands where bearing faults excite the corresponding bearing resonances [35] from bands where piston frequencies are predominantly dominant. Therefore, the spectrum is divided into five frequency bands, according to Table 2, and the mean and kurtosis are calculated for each band. The same approach for feature generation can also be applied to the cepstrum in QD. In the case of the kurtogram-based method, it is also useful to use the boundaries of the selected frequency band as an additional feature. Since EMD is performed in TD, the RMS is calculated instead of the mean to characterize signal power. An overview of the features is given in Table 3. Consequently, the dataset is divided into 6 x 40 x 7 subsets in the feature domain. Each subset is standardized to ensure that all features have equal influence in the ML process.

Table 2: Frequency bands for feature generation in frequency domain. The corresponding quefrency bands have reciprocal start and end values.

Band No.	1	2	3	4	5
Start Freq. [Hz]	0	120	360	840	1800
End Freq. [Hz]	120	360	840	1800	3200

Table 3: Overview of different feature sets

Set	Size	Feature Description
TD	6	Mean, variance, skewness, kurtosis, RMS, median
FD	10 (+ 2)	Mean and kurtosis per band (+ passband frequencies defined by the kurtogram)
QD	10	Mean and kurtosis per band
EMD	10	RMS and kurtosis per IMF or residuum

4.3 Anomaly Detection

In this study, it is assumed that faulty units generate anomalous data. Given that anomalous data are typically rare, it is often practical for anomaly detection to use normal data to construct a model that can reflect the characteristics of normal data (semi-supervised anomaly detection or novelty detection). These models are then capable of identifying anomalous data based on their deviating characteristics and thus a higher anomaly score. In this study, three ML techniques are employed for anomaly detection (see Figure 4): the One-Class Support Vector Machine (OC-SVM), the Isolation Forest (IF),

and the Local Outlier Factor (LOF). The reason for using multiple techniques lies in the fact that it is not trivial to determine which method will yield the best results for a given dataset. By applying these diverse approaches, we aim to comprehensively evaluate their performance and robustness in detecting anomalies.

To ensure comparable influence of features during model training, all feature sets are standardized to a similar scale. As test data must remain unseen, standardization is performed exclusively based on the training data, as illustrated in Figure 4. Specifically, the mean and variance are computed from the training set and subsequently applied to both training and test data (z-standardization).

4.3.1 One-Class Support Vector Machine

OC-SVM is a variant of the SVM designed for anomaly detection [35]. Unlike a binary SVM, which separates two classes with a hyperplane and is trained on normal and anormal data, OC-SVM is trained only on normal data. Hence, the objective is to find a hypersphere that maximizes the margin from the origin, where normal data lies, ensuring that the anomalous data lies in the positive half-space. A data instance p can be expressed as a vector with n_F Features. When slack variables are introduced to allow for some misclassifications, the optimization objective can be expressed as:

$$\min \frac{1}{2}\left|\left|w\right|\right|^2 + \frac{1}{\nu n_{Train}} \sum_{i}^{n_{Train}} \xi_i - b$$
$$\forall \langle w, \Phi(p_i)\rangle \geq b - \xi_i \ \wedge \ \xi i \geq 0 \tag{6}$$

where ξ is a vector of slack variables, Φ can be a nonlinear mapping function to the kernel space, b is the bias term, ν is a regularizing parameter for ensuring most training instances are within the hypersphere while minimizing the weights w, and n_{Train} is the number of data instances for training. Once the optimization is solved, an anomalous data instances o can be detected if the following condition is fulfilled:

$$\text{sgn}(\langle \text{w}, \Phi(\text{o})\rangle - b = 1 \tag{7}$$

In this work, the radial basis function is used as the kernel function. The parameter ν, which is a hyperparameter not adjusted during training, is pre-set to a value of 0.5. The closer ν is to 1, the more flexible and curvier the decision boundary becomes.

4.3.2 Isolation Forest

Instead of profiling normal data, the IF algorithm [37] isolates anomalies by recursively partitioning the data. Starting with n instances, the algorithm randomly selects an attribute a, e. g. a feature, and a split value a_{split} to divide the data. An internal node is created with the condition $a < a_{split}$, splitting the data into two subsets, where $a < a_{split}$ and where $a > a_{split}$. This process is recursively applied until a termination condition is met: The tree reaches a height limit, a subset contains only one instance, or all data points in a subset have identical values. The resulting so-called iTree is a binary tree and assuming that all instances are distinct, each instance is located as an external node when the iTree is fully grown. This tree structure ensures that anomalies, being fewer and more distinct, are located closer to the root, resulting in shorter path lengths. The path length $h(p)$ of an instance p is defined as the number of edges traversed from the root to an external node in an iTree. Anomalies tend to have shorter path lengths. The anomaly score $s(p)$ is then calculated as

$$s(p) = 2^{-\frac{E[h(p)]}{c(n)}}. \tag{8}$$

where $\text{E}[h(p)]$ is the average path length from several iTrees, in our case 100, and $c(n)$ is the average path length for unsuccessful searches in a binary search tree, where explicit reference is made to [37].

4.3.3 Local Outlier Factor

The LOF algorithm, first introduced in [38] is a density-based method where the process of calculating the LOF includes defining the degree of outlying. By introducing the concept of locality, LOF performs well in scenarios with a dataset with fluctuating density [39]. The LOF of an instance p can be calculated through

$$LOF_k(p) = \frac{\sum_{q \in N_k(p)} \frac{Lrd_k(q)}{Lrd_k(p)}}{|N_k(p)|}. \tag{9}$$

which is the average of the local reachability density (Lrd) of the instance and the k-nearest neighbors, where $N_k(p)$ is the number of neighbors of p. Note that $N_k(p) > k$ is possible for multiple instances q with the same distance. Lrd is defined as the inversion of the average reachability distance of the k-nearest neighbors of p:

$$Lrd_k(p) = 1/\frac{\sum_{q \in N_k(p)} Rd_k(p,q)}{|N_k(p)|} \tag{10}$$

The reachability distance $Rd_k(p,q)$ is defined as the distance $d(p,q)$ in the feature space but is at least the k-distance of q whereas the k-distance of q, denoted as $kd(q)$, is the distance of q to the k-nearest neighbor. This can be expressed as follows:

$$Rd_k(p,q) = \max\{d(p,q), kd(q)\} \tag{11}$$

In the present use-case, Euclidean distance is used and k is the number of training samples. In general, higher values for LOF indicate anomalous data instances whereas values around or below 1 indicate normal data instances.

4.4 Data Split and Evaluation

For the evaluation of anomaly detection quality, the concept of fold cross-validation is employed, as the dataset is relatively small and the significance of the results with randomly held-out test data (hold-out validation) would otherwise be low. As shown in Figure 4, five train-test splits of the dataset are generated, and a model is trained with the five resulting different training datasets, ensuring that the same instance does not appear in multiple splits. The exception to this rule is the faulty data, which is completely included in each split, so only the healthy data instances vary. The determination of model performance or the quality of anomaly detection is then based on the arithmetic mean of the evaluation metrics considered

across the different splits. The metric considered is the AUC, i.e., the area under the receiver-operating-characteristic-curve (ROC curve), which varies the threshold with respect to the anomaly scores of the aforementioned anomaly detection techniques. Therefore, the ROC curve plots the true positive rate (TPR) over the false positive rate (FPR).

In this study, data from faulty units are labeled as positive, anomalous instances and data from healthy units as negative, normal instances. According to this terminology, the model can incorrectly predict positive instances as negative or vice versa. There is an inherent trade-off, as a model that very confidently predicts positive instances may also frequently predict negative instances as positive, resulting in ambivalent effects regarding TPR and FPR. Therefore, a model is considered robust only if it provides correct predictions overall. Accordingly, a high AUC (close to 1) indicates robust fault detection, whereas a value close to 0.5 indicates random fault detection. In addition to the AUC, the TPR for a FPR of 0, as well as the FPR for a TPR of 0, are also analyzed to evaluate the quality of fault detection under specific requirements. In the former case, the TPR ideally takes a value of 1, while in the latter case, the FPR should be 0.

5 Results

To keep the evaluation of the experimental results clear and comprehensible, the analysis of the signal processing methods and signal channels is initially conducted using the averaged AUC values. Therefore, the arithmetic means of the AUC values using the OC-SVM, the IF and the LOF is calculated. To characterize fault detection performance, we consider results from numerous model training runs. Each training process is performed independently for each load point (as explained in Figure 4) and for each combination of measurement channel and signal processing method. The results are treated as individual data points within a distribution. These distributions are visualized as box plots in Figure 7. According to the box plots in Figure 7, it can be observed that an accelerometer is preferable to a microphone based on fault detection quality, provided no signal processing methods are applied. However, when choosing the envelope-based method, results can be achieved that are approximately equivalent to those obtained with the application of accelerometers. Here too, considering the median and the load point-specific dispersion in the representation, it is evident that the envelope-based method yields better results. The highest median AUC, around 0.954, and the smallest dispersion ranges are achievable using the envelope-based method and the accelerometer fixed to the test stand in the planar directions. However, methods based on Kurtogram, Cepstrum, and EMD generally tend to deliver poorer results, even when compared to the FFT-based method. For a more detailed examination of the load points and a deeper understanding of the sensors, only the envelope-based and FFT-based methods will be considered henceforth, as the latter does not require domain knowledge in the form of a suitable filter band in contrast to the envelope-based method.

Figure 9 clearly confirms that the envelope-based method, when considering accelerometer 2, delivers higher AUC values across the entire operating range. It further shows that the rotational speed can have a significant impact on the quality of fault detection. While with this combination, the AUC only drops to a worst-case value of 0.843 at 400 rpm, values of 0.630 can be achieved using the envelope-based method applied to accelerometer 1 in the y-direction, and values of 0.571 can be achieved using the microphone and the FFT-based method. When using accelerometer 1 and the FFT, there are partially opposing effects, as higher rotational speeds result in lower AUC values, especially at large swivel angles (bearing loads 5 to 8). However, the effect of bearing load on fault detection is not as apparent and necessitates a more detailed examination using the FPR assuming a TPR of 0, and the TPR assuming a FPR of 0. Figure 8 shows that the envelope-based method tends to yield higher TPR values and lower FPR values at smaller swivel angles, although there are outliers, such as at 1600 rpm and bearing load 1. In this context, it is important to emphasize that the FPR can increase rapidly when the TPR is 0, especially if a single faulty unit does not exhibit truly anomalous behavior. If the goal is to avoid falsely detecting good units (see TPR), higher rotational speeds remain essential for robust fault detection. In particular, in these scenarios, the superiority of the envelope-based method becomes evident. A summary of the maximum AUC values and TPR, as well as the minimum FPR values for the

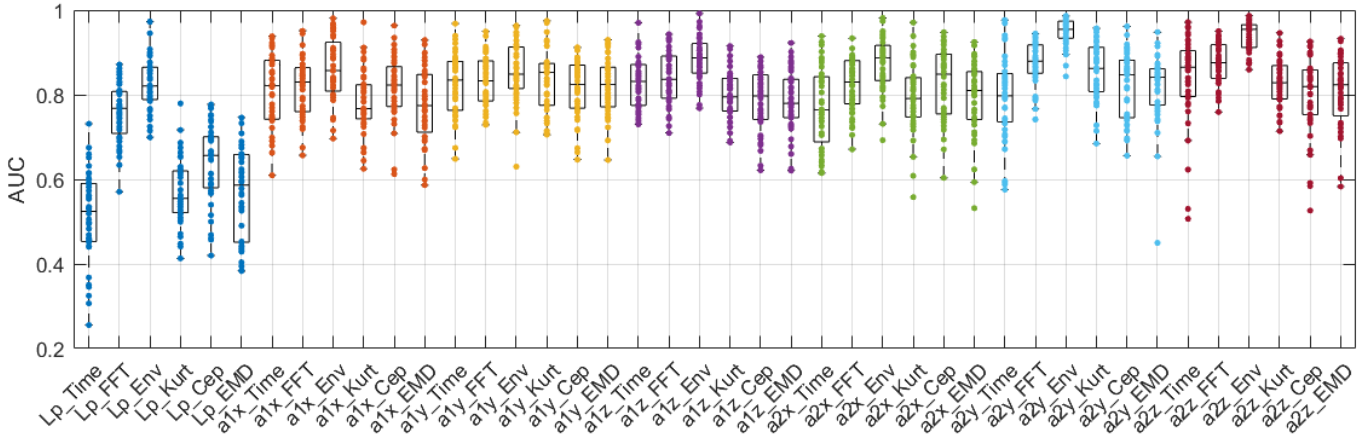


Figure 7: Distribution of AUC values obtained by averaging the models across the operating conditions. Each box represents the AUCs for a specific signal and signal processing method (Env: Envelope, Kurt: Kurtogram, Cep: Cepstrum).

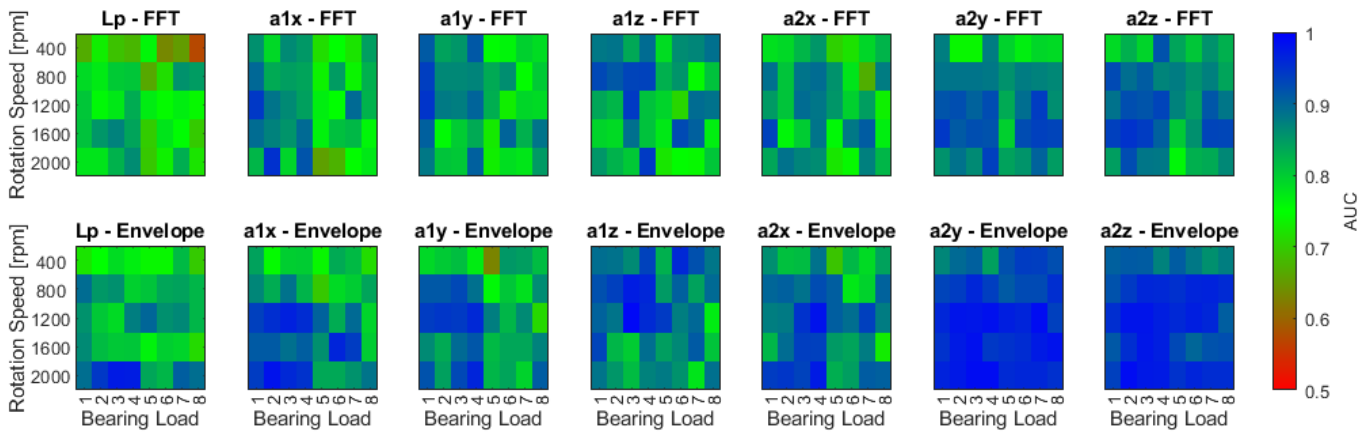


Figure 9: AUC values obtained by averaging the models depending on the operating condition

different sensors, is provided in Table 4 with corresponding signals, signal processing methods and operating conditions. The envelope method consistently delivers the best values for the metrics, and the z-direction is usually the most suitable.

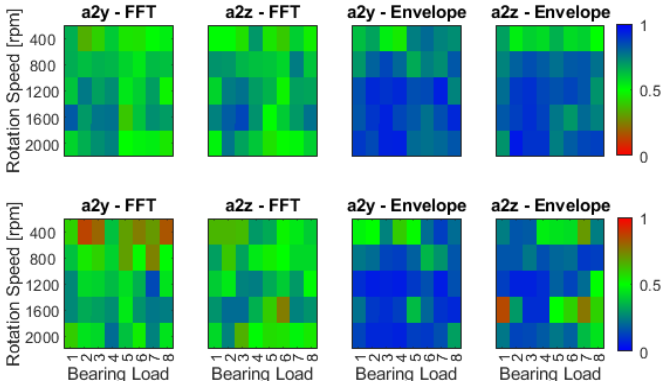


Figure 8: TPR (top) and FPR (bottom) values obtained by averaging the models depending on the operating condition

Table 4: Best values of the metrics for different sensors obtained by averaging the models (Acc: Accelerometer, Mic: Microphone)

Sensor	Metric	Value	Operating Condition	Signal	Signal Processing
Mic.	AUC	0.973	2000 rpm, BL 3	Lp	Env.
	TPR	0.906	2000 rpm, BL 3	Lp	Env.
	FPR	0.133	2000 rpm, BL 4	Lp	Env.
Acc. 1	AUC	0.992	1200rpm, BL3	a1z	Env.
	TPR	0.952	1200rpm, BL3	a1z	Env.
	FPR	0.056	1200rpm, BL3	a1z	Env.
Acc. 2	AUC	0.987	2000 rpm, BL 2	a2z	Env.
	TPR	0.952	2000 rpm, BL 2	a2z	Env.
	FPR	0.056	1200rpm, BL 2	a2y	Env.

Particularly when considering accelerometer 1, medium rotational speeds are well-suited, whereas for the microphone and accelerometer 2, the maximum rotational speed typically yields the highest results at low swivel angles. The FPR values, due to the averaging of the values for the three models, should be interpreted as follows: one of the models for anomaly detection has falsely detected a good unit as defective when it is required that all defective units must be correctly identified.

To gain a deeper insight into the performance of individual models for anomaly detection, it is crucial to compare their metrics rather than relying on the mean values as previously considered. The distributions, which arise from different operating conditions, reveal that no definitive general statement regarding performance can be made when varying scenarios for data collection and processing are taken into account, as illustrated in Figure 10. Using the IF (red), the median values for the metrics obtained from applying the FFT-based method to the signals from accelerometer 2 are inferior compared to those obtained using the OC-SVM (blue) and the LOF (yellow). However, an analysis of the metrics derived from the envelope-based method presents a different picture, with the OC-SVM yielding poorer median values and exhibiting significantly more variability depending on the operating conditions. Particularly in the lower plot of Figure 10, it becomes evident that the variability for signal a2y and the envelope-based analysis is considerably smaller, although the FPR can generally vary widely with respect to variable operating conditions. Given that the different models incorporate distinct concepts for anomaly detection, it is prudent to examine the correlations between the individual models in greater detail. Figure 11 depicts the AUC values of the models for the previously analyzed signals and signal processing methods. Compared to a perfect correlation (dashed line), the discrepancies between the IF and the other two models are apparent. A depiction with a linear regression (red line) is inadequate here, resulting in a lower Pearson Correlation of 0.52 and 0.65, respectively. In contrast, a value of 0.91 when comparing the OC-SVM and the LOF indicates a significantly more similar categorization of units and detection of faults.

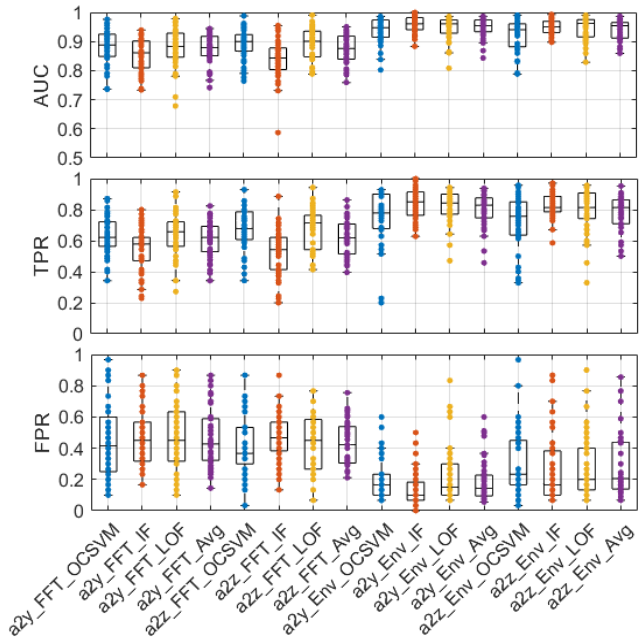


Figure 10: Distribution of AUC, TPR and FPR values for every single model across the operating conditions

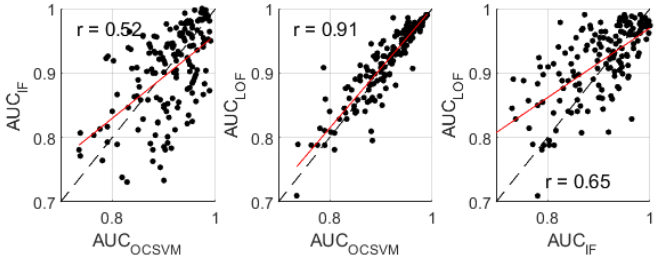


Figure 11: Correlation between the AUC values of the different models considering a2y and a2z (dashed line: perfect correlation, red line: linear regression)

6 Discussion and Future Work

The results of the experiments indicate that detecting faulty roller bearings in a serial testing scenario is a non-trivial task, as different methods of data acquisition and processing, as well as operating conditions, can have a significant impact. The choice of sensors is critical, especially in the absence of domain knowledge and when only limited signal processing methods, such as FFT, are available. In the time domain, robust fault detection using a microphone is not feasible, as the generated features are subject to significant fluctuations. On one hand, the noise emissions from the test bench infrastructure interfere with the measurements. On the other hand, the transmission path through the air attenuates the noise emissions of the test object compared to direct structure-borne sound measurements. However, the superimposed noise emissions of the test bench predominantly affect lower frequency bands, as good results can still be achieved using a filter band within the envelope analysis. Nonetheless, it must be considered that additional acoustic disturbances, such as music, conversations, or other nearby machinery, could induce interference, which was not addressed in this study. It is plausible that the impact on structure-borne sound measurements is lower due to the attenuation over the air transmission path, but this needs to be validated for the application, as disturbances can also be introduced through the structure, directly affecting structure-borne sound measurements. This also applies to other sources of interference, such as fluctuating oil temperatures or system changes due to maintenance work. Thus, further research is crucial to determine the influence of these fluctuations, especially oil temperature, on model performance. This research should investigate whether the impact of temperature variation is negligible compared to other variations (e.g., manufacturing tolerances), allowing models to leverage inherent fluctuations for robustness. If not, the need for temperature-specific training, feature integration, or test bench modifications should be evaluated.

In contrast to systematic disturbances, the repeatability of the measurements with the available data is well verified. It is confirmed that the repeatability of accelerometer 2 on the test bench is higher, as no reapplication is necessary when testing a new unit. The sensor coupling and thus the transmission path of the structure-borne sound does not vary, unlike for accelerometer 1 on the unit, where the variation in sensor application is compounded by the general serial dispersion. Particularly in the planar direction of the test bench flange, robust fault detection is possible across a wider operational range, as hydraulic pulsations have a lower impact here. Overall, it is evident that the advantage of the sensor being consistently applied outweighs the disadvantage of being further from the relevant noise source. This can be attributed to the rigid connection of the unit to the test bench. The dependency of fault detection on operating conditions can also be attributed to hydraulic pulsations: Lower swivel angles result in reduced noise and hydraulic pulsations due to the lower inertia of the displaced oil volume, leading to fewer noise disturbances that superimpose the structure-borne or airborne noise generated by the rolling bearing fault in the whole spectrum. The effects of rotational speed become particularly evident when high reproducibility, as seen with accelerometer 2, is achieved and noise disturbances are filtered out across a wide frequency range, as in the case of envelope analysis. Lower rotational speeds result in reduced excitation of the rolling bearing and the transmission path to the sensor. Consequently, there is a risk that minor faults may be masked by the background noise and the residual disturbances within this filter band. Conversely, high rotational speeds may have a more disruptive impact on accelerometer 1 if the adapter is induced to vibrate, which is application-dependent. Particularly in the z-direction, fault detection may be compromised at high rotational speeds.

The superiority of the envelope-based method, which is reflected in a high quality of fault detection with an AUC exceeding 0.9 across a wide range of operating conditions when considering accelerometer 2, and even with suboptimal sensors such as the microphone, is based on the empirically determined filter band. This filter band characterizes a range where hydraulic pulsations are attenuated by the structure, and bearing resonances are pronounced. However, it is critical to question whether this remains consistent in a serial testing

scenario where various variants and sizes of axial piston units are tested. The FFT-based method does not require such prior knowledge and can be generically applied to structure-borne or airborne sound signals. To ensure robust fault detection, it would be prudent to determine the filter band automatically, for example, through ML. A data-driven approach could involve determining anomaly detection models in individual frequency bands and isolating anomalies on a band-specific basis. Furthermore, within the scope of signal processing, there is the potential to create enhanced envelopes (see chapter 2) to calculate bands where hydraulic pulsations are minimized. The kurtogram is not a viable solution, as filter bands selected solely based on spectral kurtosis may be significantly influenced by the frequency peaks of hydraulic pulsations. Heuristic signal processing methods such as EMD offer no added value due to the dominance of hydraulic pulsations, and the cepstrum is also significantly influenced by the harmonics of the piston frequency and thus by hydraulic pulsations.

Outliers or faults in a2y and the envelope-based method, for example, are more effectively detected by the IF compared to the FFT method. For the latter, the location of faulty data instances in the feature space favors better detection by the OC-SVM and the LOF, with a notable strong correlation between the latter two. In this context, it makes sense to identify the faults that are or are not detected by individual models and to refer to the configuration in the feature space, which, however, is not within the scope of this paper. Nevertheless, it should thus be verified whether robust frameworks for fault detection can be constructed using an ensemble of models, a common approach in ML. Whether majority voting (hard voting) or soft voting is more suitable needs to be verified. The use of relatively simple models is justified due to the limited data available, ensuring the development of generalizable models. If more data can be generated, this opens the possibility of employing more complex models based on deep learning, such as autoencoders or convolutional neural networks for anomaly detection. Particularly, the application of the latter in conjunction with the transformation of time series data into image representations holds significant potential. The underlying scenario of semi-supervised anomaly detection facilitates the learning of the nominal behavior of relevant units. In serial testing processes with low variant diversity and high production volumes, model training can be effectively conducted using historical data. However, if the nominal behavior changes over time, this poses a challenge in a semi-supervised scenario, as the robustness of the solution concerning other variants must be reassessed. In such cases, unsupervised methods for outlier detection without model training (see [7]), as well as transfer or cross-domain learning approaches such as siamese neural networks, can provide viable solutions.

7 Conclusion

In this paper, we propose the integration of signal processing methods and ML in a semi-supervised anomaly detection scenario, which can be utilized for fault detection in serial testing processes. The proposal specifically focuses on the analysis of acoustic measurements but can be adapted for other signals as well. Additionally, the pipeline can be used for CM applications. The results indicate that detecting faulty bearings against the background of serial variation is not a trivial task. Operating conditions can have a significant impact on fault detection, with higher rotational speeds and lower swivel angles generally being preferable. When prior knowledge of relevant frequency bands exists, envelope analysis can yield comparatively favorable results even with microphones, demonstrating an improvement in the operating-point-specific median AUC of approximately 0.31 compared to time-based features. However, to obtain robust results in scenarios with a large variety of variants without such knowledge, it is crucial to generate repeatable results with a fixed installed accelerometer on the test bench. Using the envelope-based method, AUC values of at least 0.9 were achieved even at moderate speeds around 800 rpm, with a maximum observed value of 0.987. While the APU-mounted accelerometer achieved a best-case AUC of 0.992, its performance exhibited significantly higher variability depending on the operating point. Consequently, the potential exists to achieve more robust results with the investigated ML techniques through appropriate signal processing using a practicable test bench sensor. In our studies, it was possible to ensure nearly perfect fault detection in the best case considering the mentioned values for the AUC. This corresponds to a high TPR of 95,2% at a FPR of 0%. In Contrast to that, other signal processing methods commonly used in general bearing fault diagnosis like the cepstrum and kurtogram-based analysis as well as EMD are counterproductive for APUs.

References

[1] S. Haug, Ganzheitliche Optimierung einer Axialkolbenpumpe durch bedarfsangepasste Entlastung tribologischer Kontakte, Bd. 95, Karlsruhe: KIT Scientific Publishing, 2022.

[2] P. C. Palavicino, B. Sarlioglu, D. Bobba, W. Lee und T. Minav, „Electrification of hydraulic systems using high.efficiency permanent magnet motors,“ in *12th International Fluid Power Conference*, Dresden, 2020.

[3] M. Tiboni, C. Remino, R. Bussola und C. Amici, „A Review on Vibration-Based Condition Monitoring of Rotating Machinery,“ *applied sciences,* Bd. 12, Nr. 2, 2022.

[4] C. Xiao, H. Tang, Y. Ren, J. Xiang und A. Kumar, „A fault frequency bands location method based on improved fast spectral correlation to extract fault features in axial piston pump bearings,“ *Measurement,* Bd. 171, 2021.

[5] R. B. Randall und J. Antoni, „Rolling element bearing diagnostics - A tutorial,“ *Mechanical Systems and Signal Processing,* Bd. 25, pp. 485-520, 2011.

[6] M. Bengtsson und G. Lundström, „On the importance of combining “the new” with “the old” – One important prerequisite for maintenance in Industry 4.0,“ *Procedia Manufacturing,* Bd. 25, pp. 118-125, 2018.

[7] K. Hendrickx, W. Meert und Y. Mollet, „A general anomaly detection framework for fleet-based condition monitoring of machines," *Mechanical Systems and Signal Processing,* Bd. 139, 2020.

[8] J. Bonart und W.-G. Drossel, Entwicklung einer End-of-Line Körperschallprüfmethodik für hochintegrierte elektrifizierte Antriebstopologien, Wissenschaftliche Scripten, 2024.

[9] J. G. M. Lázaro und C. B. Pinilla, „A methodology for detection of wear in hydraulic axial piston pumps," *International Journal on Interactive Design and Manufacturing,* Bd. 14, pp. 1103-1119, 2020.

[10] S. Gaugel, B. Wu, A. Anand und B. Wu, „Supervised Time Series Segmentation as Enabler of Multi-Phased Time Series Classification: A Study on Hydraulic End-of-Line Testing," in *IEEE 21st International Conference on Industrial Informatics (INDIN)*, 2023.

[11] J. Du, S. Wang und H. Zhang, „Layered clustering multi-fault diagnosis for hydraulic piston pump," *Mechanical Systems and Signal Processing,* Bd. 36, pp. 487-504, 2013.

[12] T. Torrika, Bewertung von Analyseverfahren zur Zustandsüberwachung einer Axialkolbenpumpe, Aachen: Rheinisch-Westfälische Technische Hochschule Aachen, 2011.

[13] P. Casoli, M. Pastori, F. Scolari und M. Rundo, „A Vibration Signal-Based Method for Fault Identification and Classification in Hydraulic Axial Piston Pumps," *energies,* Bd. 12, Nr. 5, 2019.

[14] S. Horn, A. Shorbagy, M. Lenz, S. Knoll, M. Winter und J. Weber, „Causal Condition Monitoring for Axial Piston Pumps: Prediction of Different Faults under Variable Operating Conditions," in *2024 International Maha Fluid Power Conference*, West Lafayette, 2024.

[15] O. Gnepper, H. Hitzer und O. Enge-Rosenblatt, „Predictive Diagnosis in Axial Piston Pumps: A Study for High Frequency Condition Indicators Under Variable Operating Conditions," *International Journal of Prognostics and Health Management,* Bd. 14, Nr. 1, 2023.

[16] L. Siyuan, D. Linlin und J. Wanlu, „Study on Application of Principal Component Analysis to Fault Detection in Hydraulic Pump," in *Proceedings of 2011 International Conference on Fluid Power and Mechatronics*, Beijing, 2011.

[17] P. Casoli, M. Pastori und F. Scolari, „A multi-fault diagnostic method based on acceleration signal for a hydraulic axial piston pump," in *AIP Conference Proceedings* , Modena, 2019.

[18] S. Xia, Y. Xia und J. Xiang, „Piston Wear Detection and Feature Selection Based on Vibration Signals Using the Improved Spare Support Vector Machine for Axial Piston Pumps," *materials,* Bd. 15, Nr. 23, 2022.

[19] Y. Zhou, A. Kumar, C. Parkash, G. Vashishtha, H. Tang und J. Xiang, „A novel entropy-based sparsity measure for prognosis of bearing defects and development of a sparsogram to select sensitive filtering band of an axial piston pump," *Measurement,* Bd. 203, 2022.

[20] A. Kumar, C. P. Gandhi, H. Tang, G. Vashishtha, R. Kumar, Y. Zhou und J. Xiang, „Adaptive sensitive frequency band selection for VMD to identify defective components of an axial piston pump," *Chinese Journal of Aeronautics,* Bd. 35, Nr. 1, pp. 250-265, 2022.

[21] Y. Zhou, A. Kumar, C. Parkash, G. Vashishtha, H. Tang, A. Glowacz, A. Dong und J. Xiang, „Development of entropy measure for selecting highly sensitive WPT band to identify defective components of an axial piston pump," *Applied Acoustics,* Bd. 2023, 2023.

[22] A. Kumar, H. Tang, G. Vashishtha und J. Xiang, „Noise subtraction and marginal enhanced square envelope spectrum (MESES) for the identification of bearing defects in centrifugal and axial pump," *Mechanical Systems and Signal Processing* , Bd. 165, 2022.

[23] P. Casoli, A. Bedotti, F. Campanini und M. Pastori, „A Methodology Based on Cyclostationary Analysis for Fault Detection of Hydraulic Axial Piston Pumps," *energies,* Bd. 11, Nr. 7, 2018.

[24] Q. Chao, H. Gao, J. Tao, C. Liu, Y. Wang und J. Zhou, „Fault diagnosis of axial piston pumps with multi-sensor data and convolutional neural network," *Frontiers of Mechanical Engineering,* Bd. 17, Nr. 3, 2022.

[25] Y. Zhu, G. Li, S. Tang, R. Wang, H. Su und C. Wang, „Acoustic signal-based fault detection of hydraulic piston pump using a particle swarm optimization enhancement CNN," *Applied Acoustics,* Bd. 192, 2022.

[26] S. Tang, Y. Zhu und S. Yuan, „An improved convolutional neural network with an adaptable learning rate towards multi-signal fault diagnois of hydraulic piston pump," *Advanced Engineering Informatics,* Bd. 50, 2021.

[27] S. Tang, Y. Zhu und S. Yuan, „Intelligent fault diagnosis of hydraulic piston pump based on deep learning and Bayesian optimization," *ISA Transactions,* Bd. 129, Nr. Part A, pp. 555-563, 2022.

[28] Y. Zhu, L. Guangpeng, R. Wang, S. Tang, H. Su und K. Cao, „Intelligent Fault Diagnosis of Hydraulic Piston Pump Based on Wavelet Analysis and Improved AlexNet," *sensors,* Bd. 21, Nr. 2, 2021.

[29] S. Wang und J. Xiang, „A minimum entropy deconvolution-enhanced convolutional neural networks for fault diagnosis of axial piston pumps," *Soft Computing,* Bd. 24, pp. 2983-2997, 2020.

[30] S. Tang, Y. Zhu und S. Yuan, „A novel adaptive convolutional neural network for fault diagnosis of hydraulic piston pump with acoustic images,“ *Advanced Engineering Informatics,* Bd. 52, 2022.

[31] Y. He, H. Tang, Y. Ren und A. Kumar, „A deep multi-signal fusion adversarial model based transfer learning and residual network for axial piston pump fault diagnosis,“ *Measurement,* Bd. 192, 2022.

[32] N. Keller, A. Sciancalepore und A. Vacca, „Demonstrating a Condition Monitoring Process for Axial Piston Pumps with Damaged Valve Plates,“ *International Journal of Fluid Power,* Bd. 23, Nr. 2, 2022.

[33] C. Bienefeld, F. M. Becker-Dombrowsky, E. Shatri und E. Kirchner, „Investigation of Feature Engineering Methods for Domain-Knowledge-Assisted Bearing Fault Diagnosis,“ *entropy,* Bd. 25, Nr. 9, 2023.

[34] R. B. Randall, „A history of cepstrum analysis and its application to mechanical problems,“ *Mechanical Systems and Signal Processing,* Bd. 3, Nr. 19, 2017.

[35] C. Scheffer und P. Girdhar, Practical Machinery Vibration Analysis and Predictive Maintenance, Elsevier, 2004.

[36] P. Perera, P. Oza und V. M. Patel, „One-Class Classification: A Survey,“ arXiv, 2021.

[37] F. T. Liu, K. M. Ting und Z.-H. Zhou, „Isolation Forest,“ in *2008 Eighth IEEE International Conference on Data Mining*, Pisa, Italy, 2009.

[38] M. M. Breunig, H.-P. Kriegel, T. N. Raymond und J. Sander, „LOF: Identifying Density-Based Local Outliers,“ *ACM SIMOD Record,* Bd. 29, Nr. 2, 2000.

[39] O. Alghushairy, R. Alsini, T. Soule und X. Ma, „A Review of Local Outlier Factor Algorithms for Outlier Detection in Big Data Streams,“ *big data and cognitive computing,* Bd. 5, Nr. 1, 2021.

First Steps Towards Model-Based Valve Design: Measuring Pressure Drop and Flow Forces of Industrial 2/2 Directional Poppet Valves

Timo Rothhardt, Katharina Schmitz

RWTH Aachen University – Institute for Fluid Power Drives and Systems (ifas),
Campus-Boulevard 30, 52074 Aachen, Germany
E-mail: timo.rothhardt@ifas.rwth-aachen.de, katharina.schmitz@ifas.rwth-aachen.de

Abstract

This paper introduces a method for the measurement of flow forces on industrial 2/2 directional poppet valves. The valves are both normally open (NO) and normally closed (NC), direct-acting oil hydraulic valves that allow fluid flow in both directions. Previous works have only examined abstracted and reduced geometries that were specially manufactured for measurement purposes. This work provides a detailed insight into the valve actuation and force measuring system for a wide range of geometries as well as elaborates on the conceived quasi-static measurement method with a moving poppet. A unit for the compensation of pressure forces during measurement was developed and both its advantages and disadvantages are discussed. The test bench setup and the assumptions for determining flow forces from two measurements with and without flow are explained. The results for the measurements of five NO valves and four NC valves are shown for one tank pressure and one pressure difference. The results indicate that for most NO valves, the flow direction alters the behavior of the flow forces, and in one valve, the flow force changed its effective direction. In contrast, the flow force behavior in NC valves was qualitatively the same in both flow directions across all four tested valves. These measurements provide a better understanding of the flow forces in industrial NO and NC valves and can be used to validate computational fluid dynamics (CFD) simulations in the future.

Keywords: model-based valve design, poppet valves, flow forces, measurements

1 Introduction

Oil hydraulic 2/2 directional poppet valves are used in a wide range of applications such as safety systems, locking mechanisms and systems with on/off functionalities. The development of valves for specific applications is currently time-consuming and cost intensive. This is due to the fact that valve development is an iterative process between tests and computational fluid dynamic (CFD) simulations, which require expert knowledge, and is usually based on the experience of the companies. In valve development, the geometries of the poppet and sleeve determine the pressure drop and the required actuator force. Both quantities should be as small as possible for 2/2 directional poppet valves. In order to simplify and accelerate the development of such valves in the future, a mathematical model is being developed that can predict the pressure drop and the flow force based on the valve geometry. The calculation of these two quantities could thus be carried out in a fraction of the time required for a CFD simulation, reducing the development times of new valves. In addition, the knowledge is conserved and thus also made accessible to inexperienced engineers.
The model for estimating the pressure drop and flow force based on the valve geometry is data-based. The data required to create this model is determined using CFD simulations. To ensure that the data from the CFD simulations is as accurate as possible, the simulations are validated in advance with measurement data from various valves.
Recent research concerning the measurement of flow forces elaborates measurement methods for spool valves [1], [2]. This paper focuses on the measurement of industrial 2/2 directional poppet valves that are directly actuated, close leak-free in both directions and permit flow in both directions. Previous works have only examined abstracted and reduced geometries that were specially manufactured for measurement. These studies are outlined in the state of the art in chapter 2. Two different units were developed to measure the poppet force. These are presented and their advantages and disadvantages discussed. The test rig developed for the valve measurement is then shown and the measuring cycle explained. The measurement results are presented and differences in the behavior of normally open (NO) and normally closed (NC) valves as well as the influence of the flow direction on NO and NC valves are worked out. Challenges in the measurements of the two valve types are also discussed. Finally, the knowledge gained about

the behavior of NO and NC valves is summarized and concluded with an outlook for future works.

2 State of the art

First, this chapter provides a definition of flow forces for a common understanding. Then, previous research about the influence of different poppet valve geometries on the flow force acting on the poppet is summarized. The literature is divided into research, only including valve measurements and research combining valve simulations with measurements. The focus rests on the test setup, the measurement procedure and the test results.

2.1 Definition of flow forces

Flow forces are forces that occur when a fluid flow changes direction and/or velocity. In order to keep the poppet at its position, further external forces are required to balance the flow forces [3]. Figure 1 shows all forces acting on the poppet of a NO valve in axial direction.

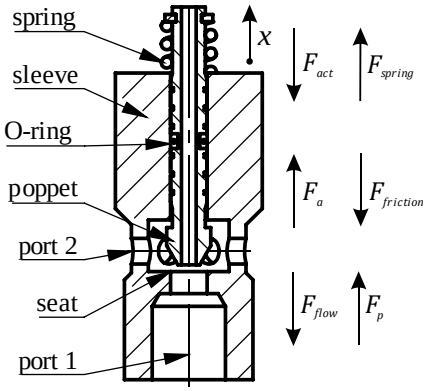


Figure 1: Axial forces on a normally open valve

The poppet is guided in the sleeve and sealed with an O-ring. In the guiding Coulomb friction forces, Newton friction forces and pressure-dependent friction forces caused by the O-ring appear. Hereinafter, all the named friction forces are combined in $F_{friction}$. In a NO valve, the spring opens the valve. The force of the spring F_{spring} is defined by Hooke's law with a pre-load and a position-dependent force. The valve can be closed by the force of an actuator F_{act} which is usually created by a magnet in commercially available valves. Inertia forces F_a are not considered in this work since the valves are not operated dynamically. The flow force F_{flow} results from a combination of static and dynamic forces due to the fluid's momentum, as well as Newton friction forces caused by the flow around the poppet. Static flow forces occur while the poppet is in a static position whereas the dynamic flow forces occur during poppet movement [4]. Since there is no dynamic operation considered in this paper, F_{flow} only contains static flow forces. Furthermore, steady-state and Newton friction forces are merely referred to as flow forces. The pressure forces in axial direction F_p act on every surface that is not parallel to the moving direction. A hole connects the bottom side to the top side of the poppet for pressure compensation. The literature commonly labels the flow direction 1-2 as diverged flow and the flow direction 2-1 as converged flow. This paper uses the numbered label and the convention that a positive flow force acts in closing direction.

2.2 Previous research: Measurements

Especially early research is characterized by measurements and analytical calculations. Stone [5] calculated discharge coefficients and flow forces for an ideal, sharp-edged poppet valve analytically. These calculations were compared to measurements of four different poppet configurations: 45° cone, 45° cone with a shoulder for flow force compensation, 45° sphere and 30° cone. All measurements were conducted at 43°C oil temperature and with an "infinite" volume behind the cone. Additionally, the 45° cone was measured with finite volumes of two different diameters behind the poppet and with the "infinite" volume at 60°C oil temperature. The calculated values were around 25 % below the measured values. It was found that a smaller chamber behind the cone leads to a higher flow force. Furthermore, three different methods for force compensation are discussed: a spring to preload the poppet, a piston with the same diameter as the poppet and a double poppet valve. The test bench used a piston with the same diameter as the poppet to compensate the force induced by the static pressure. The poppet position was controlled by regulating the pressure of the second piston chamber. The flow force was measured indirectly by multiplying the piston area with the pressure in the chamber. A dither was applied by shaking the assembly in order to reduce static frictional forces.

Scheffel [6] investigated the influence of different geometries on the static and dynamic behavior of a direct acting pressure relief valve (PRV) with a cone-shaped poppet. Eight different geometries with cone angles of 30°, 37.5° and 45° as well as different seat widths between 0.7 mm and 2.95 mm were considered. The geometries could be mounted in a variable test valve. The poppet was not pressure-compensated, enabling the attachment of a position and a force sensor directly to the poppet outside the fluid area. The study focusses on the pressure distribution at the poppet with the corresponding valve behavior and not the flow forces. The force sensor was only used to measure the preload of the spring. In the static measurements the Δp/Q-curve was recorded. The dynamic measurements investigated the pressure superelevation caused by the PRV due to a sudden rising system pressure. It was discovered that, with a constant flow rate, both a smaller seat and a smaller poppet angle lead to a larger Δp, while the valve opening was unknown. A mathematical model was developed, describing the correlation between the flow rate and the valve opening depending on the seat width, the poppet angle and the spring pre-load. With this model the valve opening was derived from the stationary measurements. At a constant flow rate, a wider seat and a smaller poppet angle lead to larger valve openings. The results can only give an indirect inference on the flow forces. A larger valve opening with the same external forces would entail lower flow forces.

Wobben [7] extended the investigations on the static and dynamic behavior of PRVs to pilot operated PRVs. The work focuses especially on the flow forces at the poppet of the main stage since these forces influence the dynamic behavior. Three different main stage geometries were tested. The poppets were not pressure-compensated and actuated with a rod which was

hydrostatically supported in the housing. The position of the poppet was set by a synchronous cylinder by controlling the chamber pressures. The flow forces on the poppet were measured indirectly by measuring the chamber and the valve pressures and calculating the resulting force with the piston and poppet areas. To address possible friction forces, a dither was applied to the servo valve controlling the chamber pressures. It was found that a sharp-edged poppet with a chamfered seat has lower flow forces than a chamfered poppet with a sharp-edged seat. Furthermore, a larger poppet angle leads to lower flow forces.

Oshima et al. [8], [9], [10] measured the flow forces of five different shaped poppet valves in both flow directions under flow conditions with cavitation using two different methods. The forces on a full-shaped model were measured directly with a load cell and the pressure distribution on the poppet was measured on a half-cut model which could be used to calculate the resulting force on the poppet. Both methods showed good agreement. Possible friction forces in the experimental setup were not taken into account. Valve "A" had a cone-shaped poppet and a chamfered seat with the poppet angle, like in the investigations of Scheffel [6]. Table 1 provides a brief overview of the influences of different parameters on the flow force found by Oshima et al. for poppet "A". Except for a valve with a sharp-edged seat in flow direction 2-1, all flow forces acted in closing direction. The arrows indicate the change of the flow force dependent on the change of the parameter, whereas the tilde represents approximately no change. A general correlation between poppet angle and flow force in flow direction 2-1 was not found, but it could be shown that the magnitude and the behavior of the flow force is strongly dependent on the occurrence of cavitation. The correlation between seat width and flow force is reversed compared to the interpreted findings of Scheffel [6]. This could be due to the fact that the valve opening was derived mathematically and cavitation was not taken into account by Scheffel [6].

Table 1: Influence of different parameters on the flow force

Parameter	Flow direction	Flow force
Valve opening ↑	1-2	F_{flow} ↑
	2-1	F_{flow} ↓
Seat width ↑	1-2	F_{flow} ↑
	2-1	F_{flow} ↑
Poppet angle ↑	1-2	F_{flow} ↑
	2-1	F_{flow} ~

The other four investigated valve geometries were based on standard geometries for PRVs at the time. The results of each valve are compared to the results of poppet "A".

Johnston et al. [11] measured flow forces on different shaped poppet and disc valves with a water-based fluid (95:5 water/oil) under non-cavitating, steady-state flow conditions. The poppets were mounted to a rod which was guided in the housing and sealed from its surroundings with a low-friction seal. The force sensor was mounted to the rod on the outside of the housing. Possible friction forces were not considered. Only flow direction 1-2 was investigated. For a sharp-edged seat, the results indicated that a smaller poppet angle leads to a larger flow force while the poppet angle has a bigger influence in large openings. During the tests, different flow patterns occurred while test conditions were maintained. The flow patterns influence the flow force and can also change the force direction. A flow-compensated poppet with an extra shoulder as in [5] was tested and confirmed the results of lower flow forces. The flow forces even assumed negative values. The measurements show discontinuities at the points of changing flow patterns. Investigations on poppets with a chamfered seat showed that the effect of the chamfered seat is neglectable when the opening is large compared to the chamfer.

2.3 Previous research: Simulations and measurements

The more recent research focusses mainly on 2D- or 3D- CFD simulations to investigate poppet valve flow characteristics. In some cases, also measurements were made to validate the CFD simulations.

One of the first people investigating the flow characteristics by measurements and CFD simulations was Dietze [12]. Dietze investigated the flow of simplified poppet valve geometries for static and dynamic operation. It was proven that CFD simulations give the right indication on flow characteristics by comparing a correct and known analytical solution with the numerical solution. Furthermore, a test bench was developed to visualize the flow in the valve. Two valve models scaled with a factor of 2.83 were used. A half cut 2D-model and a full 3D-model. Water was used as the fluid. In the 2D-model plastic particles were added whereas in the 3D-model a color additive was inserted to visualize the flow. The visualized flows match both each other and the simulated flows well. On a second test bench, the flow forces were measured on 4 different valve geometries for flow direction 1-2. The poppets were not pressure-compensated and the flow forces were measured directly with a force sensor. During the setup, the valve was closed first, then the opening position was set with a spindle and then the poppet was clamped at the set position. No seal was used between the poppet and the housing. The friction forces were less than 10 N but not further considered in the evaluation of the flow forces. For a sharp-edged piston with a chamfered seat, it was found, that a larger micro chamfer (range of 100 µm to 200 µm), a smaller opening (with the same flow) and a larger flow (with same opening) lead to a larger flow force. The dynamic investigations showed that the geometry, the opening and the system behind the valve influence flow forces on the poppet as well as the dynamic behavior.

At the same time, Latour [13] investigated four ways of flow force compensation in poppet valves and applied three of them in different valves. On the test bench, the force was measured directly with a rod connecting the poppet and the force sensor. To compensate the force difference caused by the rod guided outside the housing, a second rod was attached to the other side of the poppet and also guided outside the housing. With this setup, the poppet was pressure-compensated like in real valves. Seals were used between the rods and the housing. Possible friction of the seals was not considered in the force measurements. In the investigations, sharp-edged poppets with a chamfered seat were investigated in flow direction 1-2. It was found that a smaller seat angle (range of 15° to 75°) as well as, a larger radius (range of 55 µm to 350 µm) and a larger chamfer (range of 140 µm to 350 µm) at the poppet edge lead

to a larger flow force. A chamfer has higher flow forces than a radius. Both have higher flow forces than a sharp-edged poppet. Methods developed to compensate flow forces are 2D jet deflection, 3D jet deflection, pressure compensation with lowered static pressure and jet angle method with 3D jet deflection. All methods except the second one were applied successfully to different valves. The current state of the art in poppet valves is the 3D jet deflection for NO valves and the 2D jet deflection for NC valves.

Bergada et al. [14] derived analytic formulas for the flow and flow forces for laminar flow conditions in cone shaped poppet valves. The seat had the length of the cone, so that the opening was small compared to the length. It was shown that analytic formulas can be derived for laminar flow conditions.

Simic et al. [4] optimized the geometry of a poppet valve with the aim of lowering flow forces during static operation. The maximum opening was 0.1 mm. The valve was like a spool valve, with the difference that the sealing edge was designed like in a poppet valve. The 2D jet deflection was used to adapt the compensation of flow forces. The influence of five geometrical parameters on the flow forces was investigated via CFD simulations. For one parameter, additional experiments confirmed the accuracy of the CFD simulations with a maximum deviation of 7 %. The spool was not pressure-compensated, and the force sensor was attached directly to the spool with a rod. O-rings were used to seal between the spool and the housing. To take the friction of the seals into account, the force was first measured under static pressure and without flow at all positions. Afterwards, the measurements were repeated at all positions, with flow and a constant pressure difference of 25 bar. The difference in force between these two measurements is the flow force. Every parameter was optimized independently and then combined to an optimal solution.

In summary, previous research only varied single parameters in order to investigate their individual influence on the flow forces. The geometries used for these investigations were mostly simplified and do not occur in currently existing valves. Some experiments took the pressure compensation of the poppet or the friction of seals into consideration, but not at the same time. Mainly NO valves and the flow direction 1-2 were tested.

3 Test setup

This chapter shows the valve and force measuring system as well as the hydraulic test bench developed for measuring industrial poppet valves. Furthermore, the test procedure is explained in detail.

3.1 Valve actuation and force measuring system

One challenge in measuring the forces of the poppet in industrial valves is that the poppets are pressure-compensated. The hole in the poppet transmits the pressure from port 1 to the top side, where the armature pushes on the poppet. The armature is located in the pressurized area. To be able to measure the force on the poppet, the armature, the pole tube and the coil must be removed. In most valves, the pole tube is screwed into the sleeve and can be unscrewed to attach a measuring system. In few valves, the pole tube is welded or pressed to the sleeve and cannot be unmounted. For these valves the pole tube is machined to enable the attachment of a measuring system.

Figure 2 shows a step section view of the actuation and force measuring system developed for industrial poppet valves. The valve shown in the figure is a NC valve in closed position.

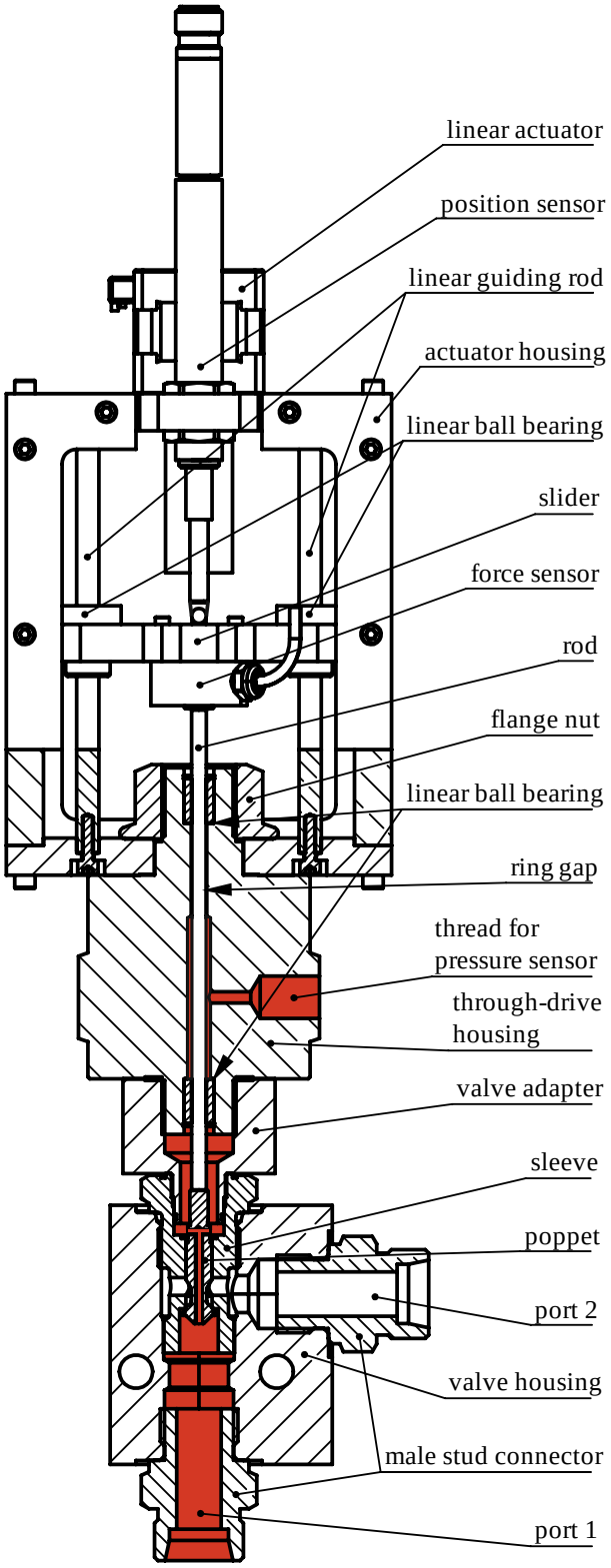


Figure 2: Valve actuation and force measuring system

The pressurized area is marked in red. For simplification, all seals and the spring of the poppet are not shown in the figure. During assembly, the poppet valve is screwed into the valve housing which is connected to the hydraulic circuit via the two ports. The matching valve adapter is screwed into the valve and the through-drive is screwed into the valve adapter. The actuator housing is fit to the through-drive and tightened with a flange nut. The through-drive housing contains linear ball bearings on both sides, guiding the rod that pushes on the poppet. To get a minimal friction force in the through-drive unit, a ring gap with a length of 25 mm is used to seal the pressure to the outside. The fit of the ring gap is 4 mm g6 for the rod and 4.05 mm H7 for the hole. The tolerance chain of the holes for the bearings and the bearings themselves is chosen such that the rod cannot touch the housing. A pressure sensor can be mounted to the unit and measure the compensation pressure of the poppet.
The actuator housing contains two guiding rods. A slider with two linear bearings is mounted on the guiding rods and is freely movable. The force sensor is screwed to the slider. The miniature force sensor HBM C9C, with a maximum force of 500 N and a linearity of 0.2 %, is used. The linear actuator is mounted to the top of the housing in line with the rod of the through-drive unit. It can only apply compressive and no tensile forces to the slider. In fig. 2 the position sensor covers the touching point of the actuator with the slider. The actuator is a Nanotec LGA35 with a maximum force of 242.4 N, a stroke length of 38.1 mm and a resolution of 3 µm per step. The actuator combines a stepper motor with a spindle in one unit. The position sensor is screwed into the housing above the actuator and measures the distance to the slider. The position sensor is a WayCon LVIT with a stroke length of 20 mm and a linearity of 0.5 %.

The maximum pressure all measured valves are rated for is 250 bar. With a rod diameter of 4 mm, the static pressure creates a force of 314 N. The spring forces of the spring opening or closing the poppet as well as the friction forces in the valve were unknown. Due to the size of the original valve coil, the flow forces were expected to be less than 50 N. In general, a force sensor should be selected that has a range equal to the force to be measured in order to obtain the highest possible resolution. In this case the range of the force sensor must cover the force induced by the static pressure, the spring forces and the flow forces, resulting in a force sensor with a capacity of at least 364 N. Consequentially, the measuring range of the force sensor must exceed the flow force that is to be measured by at least a factor of 7. This results in a lower resolution of the flow force measurement due to a lower signal-to-noise ratio. To achieve a more exact force measurement, at first a compensation unit was developed. The compensation unit allows the measurement of the poppet forces without forces induced by static pressure on the rod. It can be installed in the place of the through-drive. Figure 3 shows a section view of the developed compensation unit. The main part of the unit is the spool with three different sized shoulders. The two outer shoulders have a diameter of 8 mm, while the inner shoulder has a diameter of 11.3 mm. Each shoulder has a guided length of 10 mm and contains three pressure compensation grooves. Two of the shoulders are guided in the housing and the third shoulder is guided in the top cover. Not pictured in the section view are three M8 screws fixing the top cover at the housing. To reduce torque on the spool by possible lateral forces, the spool is as short as possible and two rods guided in linear ball bearings are located on both sides of the spool. The rods are rounded on the spool side. A pressure sensor can be screwed to the unit to measure the compensation pressure of the poppet. A leakage pipe can be connected to the unit to catch the leakage at ambient pressure. To explain the functionality, the pressurized area is colored red. The lower rod is pressurized from both sides so that no pressure forces arise. The spool seals the pressure to the outside and has three shoulders to compensate pressure forces. The working principle is that the pressure force acts on the lower shoulder in x direction. A force of the same magnitude occurs in the opposite direction, between the upper 8 mm shoulder and the 11.3 mm shoulder, since the differential area of the two shoulders equates the area of the small shoulder. This allows the measurement of the poppet forces with a more precise force sensor, characterized by a smaller range.

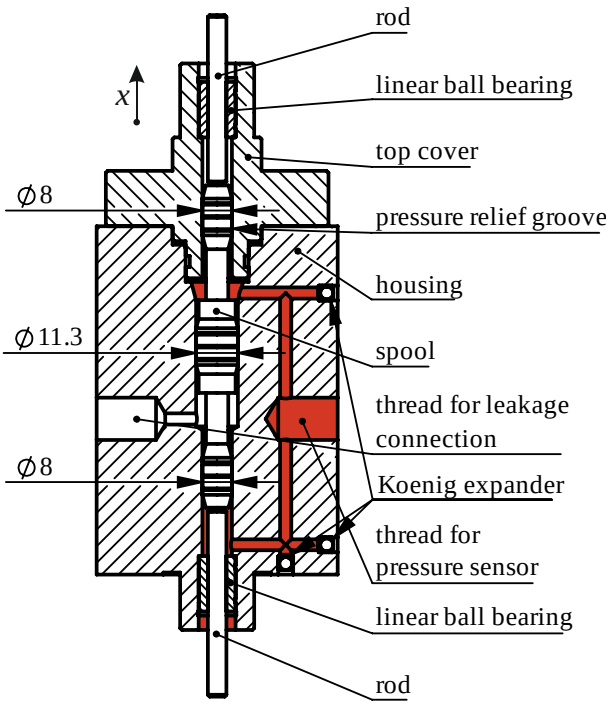


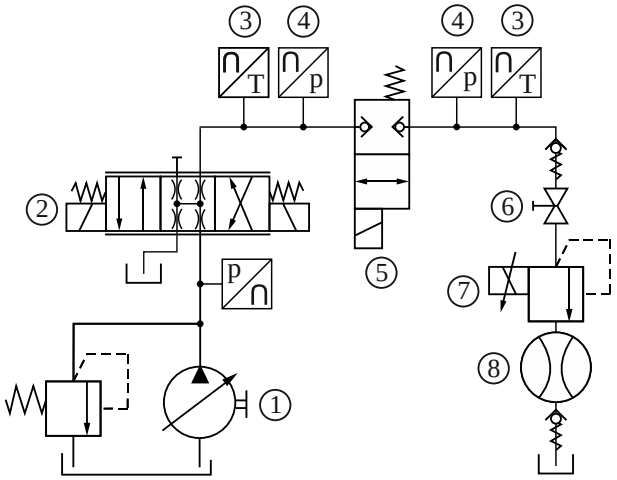
Figure 3: Compensation unit for static pressure forces

Tests showed that the unit works and pressure forces are compensated. A problem that occurred was that the spool jammed at certain positions. The jams could be seen in discontinuous force measurements. The peaks in the force measurements became larger with higher pressure in the compensation unit. Measurements also showed different pathways (force over poppet position) in two runs, when the pressure was changed between the runs. This means that measurements cannot be taken reliably and repeatedly under the same conditions and the compensation unit is unsuitable for measuring flow forces in this application. Possible reasons for the jamming are that the spool is not free from torque or that the housing deforms under pressure. However, the most likely reason is the double overfit of the spool. From a manufacturing perspective, it is generally not optimal to guide

more than one diameter on a spool since it leads to an overfit. This overfit exists between the spool and the housing. An additional overfit exists between the spool and the top cover. The top cover itself is fit to the housing. All dimensional, shape and position tolerances added lead to a non-optimal guiding and a jamming of the spool. Summarized, it can be said that the concept of pressure compensation works but for a reliable compensation unit an even more precise manufacturing or a different concept is required. For all measurements presented in this paper, the setup shown in fig. 2 is used.

3.2 Hydraulic test bench

A hydraulic test bench for measuring the poppet valves was developed and built. Figure 4 shows the hydraulic circuit diagram of the test setup.



① central pressure supply
② 4/3 directional proportional valve
③ temperature sensor
④ pressure sensor
⑤ test poppet valve
⑥ ball valve
⑦ proportional pressure relief valve
⑧ flow rate sensor

Figure 4: Hydraulic diagram of the test bench

The experimental setup is connected to the central pressure supply (1) of ifas. The pressure supply is set to a flow rate of 80 l/min with a temperature of 40°C. Both a temperature sensor (3) and a pressure sensor (4) are located in front of and behind the tested industrial valve. The temperature sensors are Danfoss MBT 3270 as PT1000 elements. As pressure sensors, Hydac HDA 4700 series sensors with a measuring range of 0 bar to 400 bar and an accuracy of 0.25 % were used. The distance from the pressure sensor to the inlet of the valve housing is 150 mm. From the outlet of the valve housing to the second pressure sensor, the distance is 230 mm. The distance between the temperature sensor and the pressure sensor is 60 mm. The 4/3 directional proportional valve (2) is a Bosch Rexroth 4WREE10-3X and is used to control the differential pressure over the test valve. The valve has a negative overlap. The proportional pressure relief valve (7) is an Argo-Hytos SR4P2-B2 and is used to control the outlet pressure of the test valve. The flow rate sensor (8) VSE VS2, with a measuring range of 0.1 l/min to 120 l/min and an accuracy of 0.3 %, measures the flow rate through the test valve. The ball valve (6) is used to stop the flow through the test valve to conduct measurements under static pressure. The hydraulic circuit diagram in fig. 4 is simplified. At the real test bench, a second and equal measurement string exists. It is connected to the closed port of the 4/3 directional proportional valve (2) and ends with a ball valve (6) in front of the proportional pressure relief valve (7). The second measurement string enables the user to change faster between valves for measurement. To change the measurement string, the 4/3 directional proportional valve is actuated in the opposite direction. The flow direction through the test valve can be changed by turning the housing of the test valve on the test bench. All signals are recorded with a frequency of 1000 Hz.

3.3 Test procedure

This chapter explains the general assumption that was made to measure the flow forces and describes the test procedure that was used for every valve in both flow directions. The idea is that the actuator force F_{act} can be measured with the previously described actuation and measuring unit. The force is equal to the sum of the four forces F_p, F_{spring}, $F_{friction}$ and F_{flow}. With just one measurement it is not possible to differentiate between the four forces. The assumption is that F_p, F_{spring} and $F_{friction}$ are the same with and without flow. With flow the flow force F_{flow} is added to the system. So, the difference between the measured force with flow and without flow must be the flow force. A restriction is that F_p and $F_{friction}$ depend on the pressure. This is addressed by taking measurements at different pressure levels. Furthermore, the friction coefficient is different when the poppet is at a static position or moves. Static friction would add uncertainty to the measurements. Therefore, the poppet should move at a very slow speed during the measurements to only have dynamic friction. So, all measurements are taken with a moving poppet in quasi-static operation.

Figure 5 shows all steps after mounting the valve until all measurements in one flow direction are finished. The first step is to measure the angle of the valve in the housing. During manufacturing, the start of the thread on the valve is random since it is a turned part. The start of the thread determines the positioning of the radial holes in the sleeve to port 2 in the valve housing. This position is later important for the CFD simulation because the valve angle should be included in the model. Thereafter the ball valve (6) is opened and the pressure supply is started in order to heat all components of the test bench. For this, the test valve is opened and a tank pressure p_t of 75 bar and a pressure difference Δp of 15 bar to 30 bar, depending on the test valve, is set. When the temperature in front of the test valve reaches 45°C, the pressure supply is stopped.

After this, the actuator is calibrated. During the calibration, the closed position of the valve is set in the test bench control. The actuator is moved slowly downwards to push on the poppet with a constant velocity. In the force over time plot, at first, a steep rising of the force can be seen until the spring pre-tension is reached. This is followed by a flatter zone with a linear increase defined by the spring stiffness. After this zone, a second peak can be seen. For NO valves this point is the closed

position. For NC valves, this point is the maximum technically possible opening position which, however, differs from the maximum opening position during operation. The maximum force is limited between 45 N and 95 N, depending on the valve. The closed position of NC valves is at the transition from the first steep rise to the flatter area.

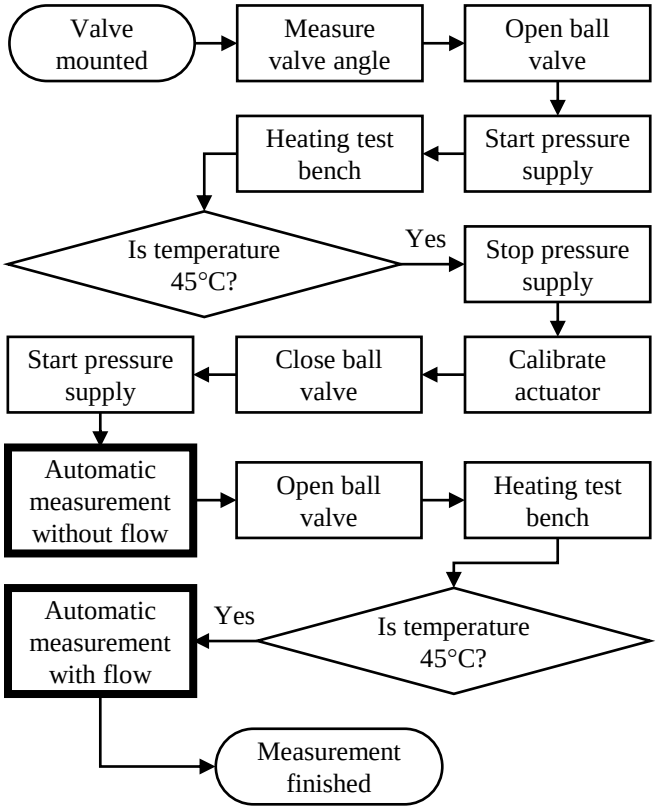


Figure 5: Measuring cycle for one valve in one flow direction

After the calibration, the ball valve (6) is closed and the pressure supply started again to begin with the measurements without flow. Figure 6 shows the measurement cycle without flow in detail. For the measuring cycles a maximum and minimum opening position is defined dependent on the valve. The minimum opening position lies between 2 % and 20 % and the maximum opening position between 60 % and 100 % of the overall stroke. The minimum positions are chosen so that the valve opening is at least 0.1 mm. The maximum opening position is 100 % for all NO valves, while the maximum opening position had to be limited for the NC valves to ensure that the valve can close by itself in all tested operating conditions. This phenomenon is discussed later in this paper. During the cycle, at first the poppet is moved to the maximum opening position and the pressure is set to 10 bar. When the pressure is reached, the poppet is moved to the minimum opening position and then again to the maximum opening position with a velocity of 0.05 mm/s. Before and after each movement, a waiting time of 5 s is included. After the moving cycle, the pressure is raised in 5 bar steps to 60 bar and then in 10 bar steps to 200 bar. Thereafter, the measuring cycle without flow is finished and the ball valve (6) is opened. Then the test bench is heated again to 45°C and the automatic measuring cycle with flow started. Figure 7 shows the cycle in detail.

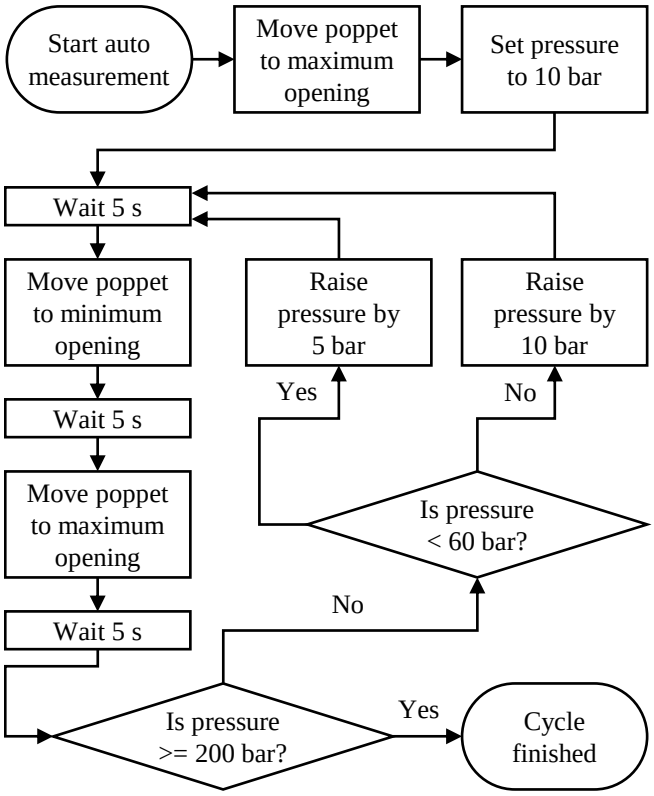


Figure 6: Cycle for automatic measurement without flow

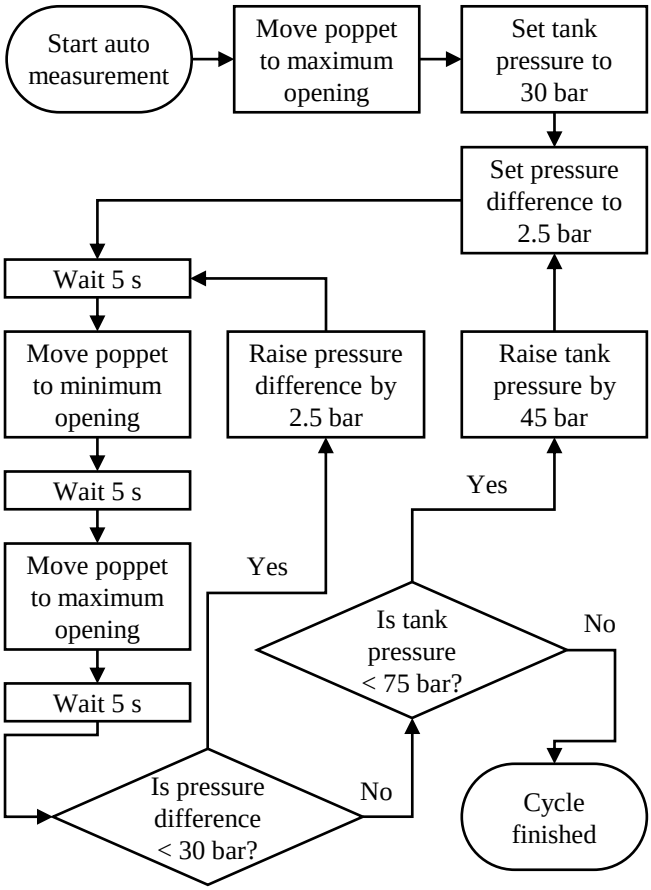


Figure 7: Cycle for automatic measurement with flow

At first, the poppet is moved to the maximum opening position. Next, the tank pressure is set to 30 bar and the pressure difference over the test valve is set to 2.5 bar. When both pressures are reached, the poppet is moved to the minimum opening position and then to the maximum opening position with a velocity of 0.05 mm/s. Before and after each movement, the position was held for 5 s. After each moving cycle, the pressure difference is raised by 2.5 bar until 30 bar is reached. The whole procedure is repeated with the tank pressure of 75 bar. Thereafter, all measurements for one valve in one flow direction are finished.

4 Results and discussion

The first step in the evaluation is to create interpolation (IP) models for the force, based on measurements without flow. For each valve and flow direction, one interpolation model for the opening and one IP model for the closing direction are created. The IP models can predict the actuator force without flow, dependent on the poppet position x and the compensation pressure p_{comp} measured in the through-drive. These IP models are used together with the measured force with flow to calculate the flow forces. For NO valves, the measured force with flow is subtracted from the force predicted by the IP model. For NC valves, the terms of the subtraction switch. This calculation is chosen to obtain the definition of positive flow forces acting in closing direction.
To create the IP model, only the areas of poppet movement from the measurements without flow are used. For the predictors x and p_{comp} continuously rising or falling signals are needed. Since the measured position signal has noise and does not fulfill the criteria, it cannot be used unfiltered. Therefore, a moving average with a range of 0.3 s was applied on x. For all other signals, moving averages with a range of 0.1 s were used. Figure 8 shows the interpolation models in opening and closing direction for a NO valve.

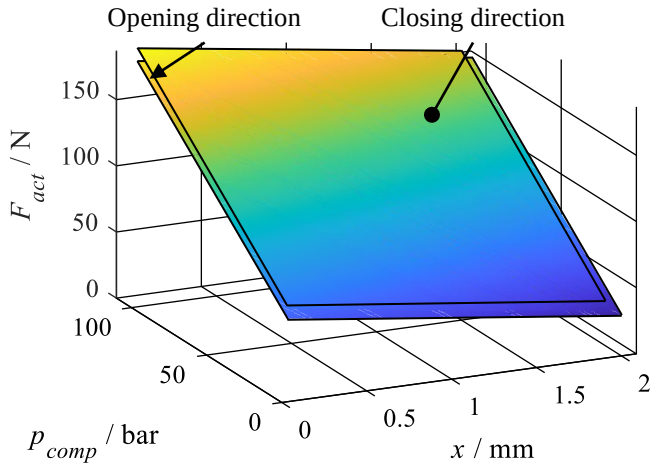


Figure 8: Two interpolation models for a NO valve in closing and opening direction

Both interpolation models are planes that are approximately parallel. The distance between the planes is hysteresis, induced by the friction of the poppet and the through-drive. In the figure, the plane from the closing direction nearly covers the plane from the opening direction completely. The hysteresis for this NO valve lies between 2.8 N and 9.9 N, depending on the compensation pressure and the poppet position. For NO valves at constant pressure, the force rises with a falling position due to the spring. At a constant position the force rises with the pressure due to the pressure acting on the rod of the through-drive. The same behavior under rising pressure can be seen at NC valves. For NC valves at constant pressure, the force rises with a rising position due to the spring. So, in x direction the gradient of the planes would be positive.

With the measurements made, different effects such as the influence of the tank pressure or the flow rate on the flow force can be evaluated for each valve. An influence of the tank pressure on the flow forces could not be seen in any measurements. A higher flow rate, respectively a higher pressure difference, lead to an amplification of the flow forces in general. This influence is not analyzed further in this paper. The aim is to gain a general understanding of flow forces in both flow directions in industrial poppet valves. Therefore, only the measuring cycle at a tank pressure of 30 bar and a pressure difference of 15 bar is shown and discussed in this paper. 15 bar is the maximum pressure difference at which all valves could be measured.
Figure 9 shows the measured flow force and flow rate over the poppet position for the five tested NO valves in both flow directions in opening (op.) and closing (cl.) direction. In flow direction 1-2, the maximum flow forces are between 3.5 N and 24.1 N, depending on the valve. The flow forces are proportional to the maximum flow rate. For valve 5, the maximum opening is not limited by the geometry, which means that this valve has the largest stroke. The behavior of the valves changes in flow direction 2-1. The valves 2, 3 and 5 have an almost constant flow force of 0 N to 4 N over the stroke. Valve 4 shows the same qualitative curve as in direction 1-2 but shifted downwards and flatter. The maximum flow force of this valve is 12 N. Valve 1 only has negative flow forces of up to -26.8 N. This means that the flow forces act in the opening direction. Valve 1 has a 20 % greater flow rate and valve 4 has an 80 % greater flow rate than in the direction 1-2.

Figure 10 shows the measured flow force and flow rate over the poppet position for the four tested NC valves in both flow directions in opening (op.) and closing (cl.) direction. In flow direction 1-2, the maximum flow forces are between 4.6 N and 16.7 N, depending on the valve. The flow force curves are qualitatively the same for all valves. In the first section, the force increases, reaches a maximum and then decreases again. For valves 1, 2 and 4, the maximum flow force is at the same position as the maximum flow rate. The fact that the flow force and the flow rate reach a maximum and then fall again when the valve is opened further is because the smallest cross-section is no longer at the sealing edge, but on the side of the radial holes in the direction of port 2. In this case, the flow force is partially negative and acts in the opening direction. When the opening or the pressure difference gets too large, the valves do not close on their own anymore. The poppet stays at an open position since the actuator cannot pull on the poppet. Then the pressure difference must be lowered. This is a state that does not occur in an original valve since the original magnet cannot open the valve that wide. Thus, preliminary investigations were made to find the maximum pressure

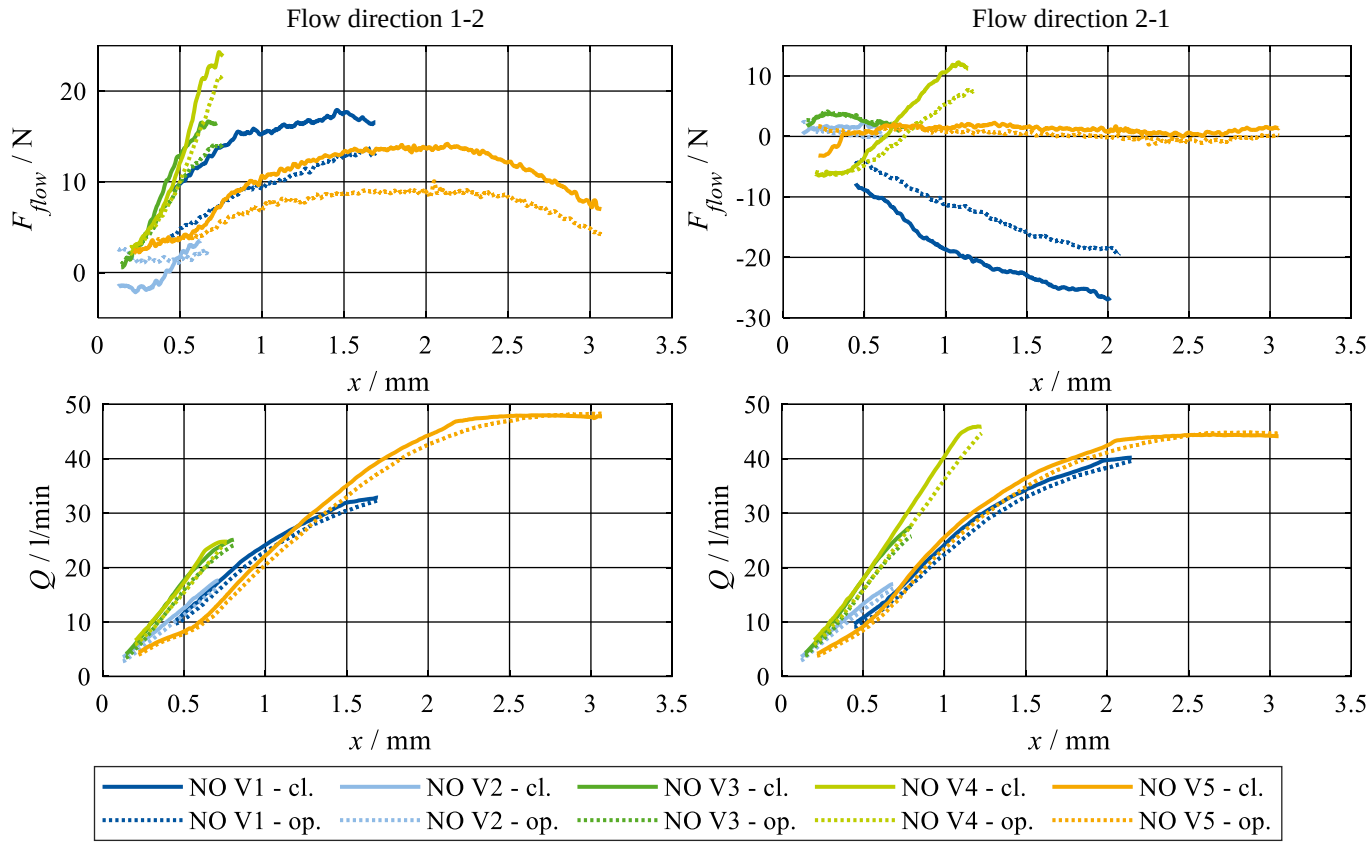


Figure 9: Flow force and flow rate over the poppet position for 5 NO valves at p_t = 30 bar and Δp = 15 bar

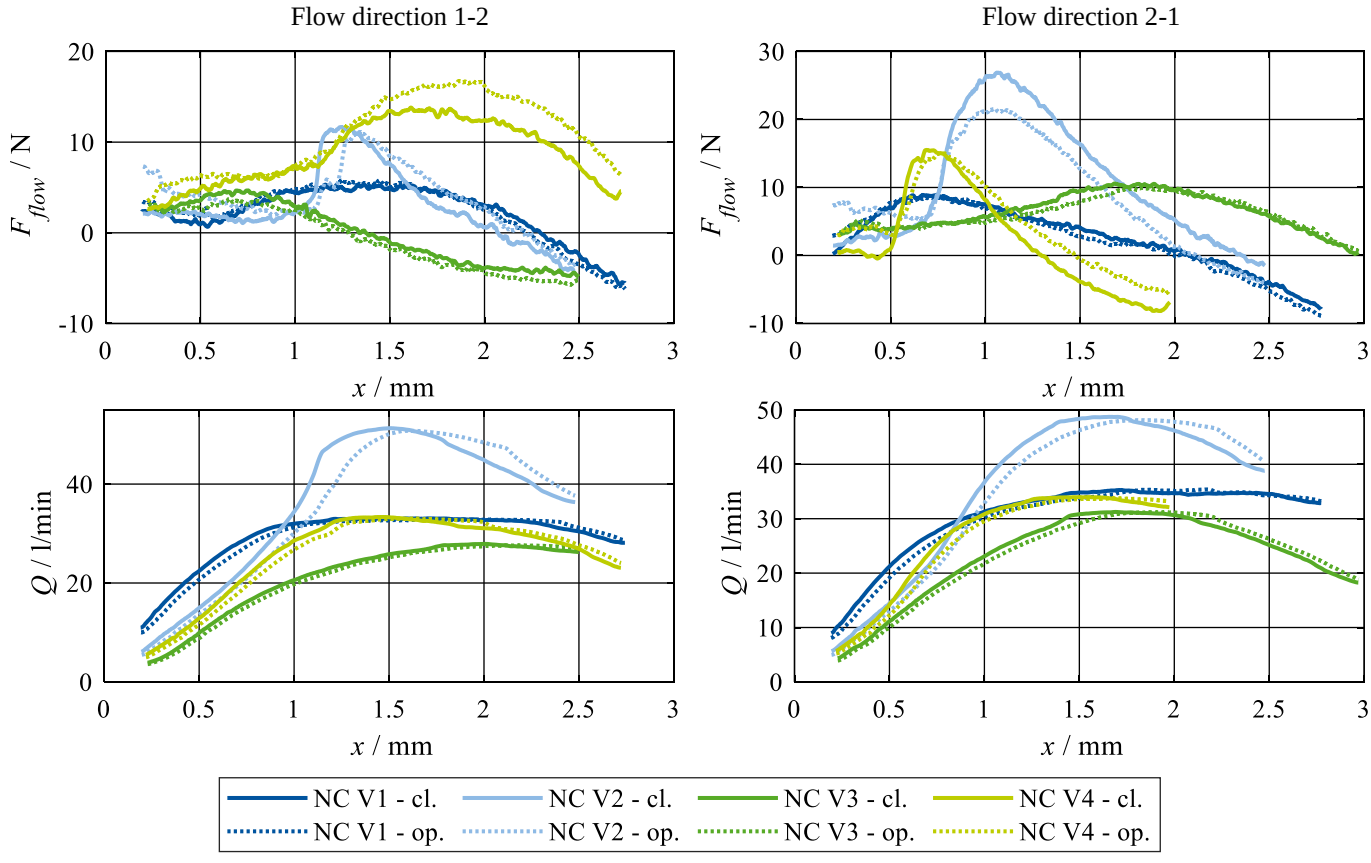


Figure 10: Flow force and flow rate over the poppet position for 4 NC valves at p_t = 30 bar and Δp = 15 bar

difference and opening width for the automatic measuring cycle. In flow direction 2-1, the curves of the flow force and the flow rate match qualitatively to the curves in flow direction 1-2. The peaks in flow force occur in two valves at lower positions and in one valve at a greater position. However, the maximum flow force of three valves is greater by a factor of 2 to 2.5 and reaches a maximum of 26.9 N. The flow rate also remains quantitatively unchanged.
Summarized, it can be said that in NO valves, the behavior of the flow forces changes with the flow direction. In one NO valve, the flow force changes the effective direction. The flow rate tends to be higher in flow direction 2-1. In NC valves, the flow direction does not influence the behavior of flow forces as much. The flow rate is qualitatively and quantitatively the same. The peak of the flow forces changes its position and tends to be higher in flow direction 2-1. Over all measurements made, the highest flow force measured was below 35 N.

Hereinafter, other characteristics that could be observed during the tests and the effect of not using the compensation unit will be discussed. One characteristic that can be seen in fig. 9 and fig. 10 is that the flow force in opening direction is different to the flow force in closing direction over a large range of the poppet position in almost all valves. The expectation is that the flow force is the same during opening and closing since hysteresis due to friction is considered in the IP model. Hence, during the measurements with flow, other forces must occur, or the flow pattern changes during the movement. One indication for the change of flow patterns is that the flow forces are equal in a specific range and then split up (e.g. NO valve 1). The change of flow patterns was investigated amongst others by Masuda et al. [15] and shows the same behavior. The most probable different force is a changing friction force. Due to the flow, an uneven pressure field can occur around the poppet. This leads to a slight tipping of the poppet in the sleeve and higher friction forces that are not considered in the IP model. This theory must be investigated by CFD simulations in the future.
All measurements were conducted as a basis for the validation of CFD simulations. The measurements were made with a moving poppet even though the valves are switching valves and the CFD simulations should be stationary. The aim is to preferably simulate the real operating point of the valve. Therefore, the poppet position during real valve operation must be found. For NO valves, the real operating point can be seen easily in measurements since the rod of the through-drive lifts from the poppet when the poppet is in the force equilibrium and consequently in the real point of operation. In NC valves, the force equilibrium forms together with the force of the magnetic actuator. This force is unknown thus the real point of operation is unknown in NC valves. So, assumptions based on measurements must be made to select one operating point for the validation of CFD simulations for NC valves.

All measurements were made using the through-drive unit and not the compensation unit. Therefore, a force sensor with a maximum force of 500 N had to be used. The resolution of the Beckhoff EL3351, used for data acquisition of the force sensor, is 16 bit. The typical noise measured is 0.05 % of the full-scale value. Using the 500 N sensor, this results in a noise of 0.25 N. The flow forces in fig. 9 and fig. 10 are in a wide range larger than 5 N. In these areas, the signal-to-noise ratio is larger than 20, which is accurate enough. The ability to use a 50 N sensor in combination with the compensation unit would lower the noise to 0.025 N, so that the signal-to-noise ratio is increased by a factor of 10. This would increase the accuracy of the flow force measurements below 5 N. Knowing that the real point of operation of the valves is in the areas of larger flow forces, the measurement method using the through-drive and the 500 N force sensor is valid and accurate enough to validate CFD simulations. Three NO valves in flow direction 2-1 show flow forces around 0 N. These points of operation need special attention when using the measurement data for validation.

5 Conclusion and outlook

The presented valve actuation and force measuring system with a through-drive can be used to measure flow forces. It was attempted to replace the through-drive with a unit that compensates the static pressure. This would have allowed to use a force sensor with a smaller measuring range and higher accuracy. The compensation unit worked, and the static pressure was compensated. Unfortunately, the spool in the compensation unit jammed at certain positions. The most likely reason for the jamming is the double overfit of the spool at the two shoulders in the housing and the one shoulder in the top cover. Since the accuracy of the force measuring system with the through-drive is high enough to validate CFD simulations, the compensation unit is not further improved and considered.
The flow forces are determined by two automatic measuring cycles. In the first cycle the poppet is moved at specific pressure levels up to 200 bar without flow. During the second cycle, the valves were flown through and the tank pressure as well as the pressure difference over the test valve controlled while the poppet was moved. The movement of the poppet was used to avoid static friction forces and the speed of 0.05 mm/s to achieve quasi-static measurements. The tested tank pressures were 30 bar and 75 bar and the maximum pressure difference was dependent on the valve between 15 bar and 30 bar. To determine the flow force it was assumed that, between the two cycles, all forces are identical except for the flow force. So, the difference in the measured force from both cycles must be the flow force. To calculate the flow force, interpolation models were created that can predict the force without flow based on the poppet position and the compensation pressure.
Five NO valves and four NC valves were measured. The results for the flow force and the flow rate over the poppet position were presented in both flow directions for the tank pressure of 30 bar and the pressure difference of 15 bar. The maximum flow force from all tested NO valves was 24.1 N. For NO valves it was found that the flow direction changes the behavior of the flow forces for most valves. In one valve, the flow force changed the effective direction to -26.8 N. For the NC valves the maximum flow force is 26.9 N. The flow force behavior is compared between the four NC valves qualitatively the same in both flow directions.
In future work, the measurements can be used to validate CFD simulations of the valves. With a validated CFD simulation,

the database for the prediction model of the flow forces and the pressure drop can be generated. The prediction model can be validated with measurements on further valve geometries using the valve actuation and force measuring system, the test bench and the test procedure presented in this paper.

Acknowledgements

The IGF research project 01IF22852N of the research association Forschungskuratorium Maschinenbau e. V. – FKM, Lyoner Straße 18, 60528 Frankfurt am Main was supported from the budget of the Federal Ministry of Economic Affairs and Climate Action through the DLR within the scope of a program to support industrial community research and development (IGF) based on a decision of the German Bundestag.

Nomenclature

Designation	Denotation	Unit
Δp	Pressure difference	bar
F_a	Acceleration Force	N
F_{act}	Actuator Force	N
F_{flow}	Flow force	N
$F_{friction}$	Friction force	N
F_p	Pressure force	N
F_{spring}	Spring force	N
p_t	Tank pressure	bar
Q	Flow Rate	l/min
x	Poppet position	mm

References

[1] P Bordovský. *Evaluation of Steady-State Flow Forces in Spool Valves.* Aachen, 2018. ISBN 978-3-8440-6699-9

[2] S Hucko, T Vonderbank, H Krampe, K Schmitz. Investigation of the Influence of the Fluid Temperature on Flow Forces in Spool Valves. *SICFP23 Proceedings*, (18):534-546, 2023.

[3] N Herakovič. *Die Untersuchung der Nutzung des Piezoeffektes zur Ansteuerung fluidtechnischer Ventile*. Mainz Verlag, Aachen, 1995. ISBN 3-89653-041-0.

[4] M Šimic, N Herakovič. Reduction of the flow forces in a small hydraulic seat valve as alternative approach to improve the valve characteristics. *Energy Conversion and Management*, 89:708-718, 2015.

[5] J A Stone. Discharge Coefficients and Steady-State Flow Forces for Hydraulic Poppet Valves. *Journal of Basic Engineering*, 82(1):144-154, 1960.

[6] G Scheffel. *Stationäres und dynamsiches Verhalten eines direktgesteuerten Kegelsitzventils unter dem Einfluss der Ventilgeometrie*. Aachen, 1977.

[7] G-D Wobben. *Statisches und dynamisches Verhalten vorgesteuerter DBV unter besonderer Berücksichtigung der Strömungskräfte*. Aachen, 1978.

[8] S Oshima, T Ichkawa. Cavitation Phenomena and Performance of Oil Hydraulic Poppet Valve (4th Report, Influence of Cavitation on the Thrust Force Characteristics). *Trans.JSME, B*, 51(467):2249-2257, 1985.

[9] S Oshima, T Ichkawa. Cavitation Phenomena and Performance of Oil Hydraulic Poppet Valve (5th Report, Influence of Dimensions of Valve on the Thrust Force Characteristics). *Bulletin of JSME*, 29(251):1427-1433, 1986.

[10] S Oshima. An Experimental Study on the Several Poppet Valves With Difference in Shape. *Proceedings of the JFPS International Symposium on Fluid Power*, (1):361-368, 1989.

[11] D N Johnston, K A Edge, N D Vaughan. Experimental investigation of flow and force characteristics of hydraulic poppet and disc valves. *Proceedings of the Institution of Mechanical Engineers*, (205):161-171, 1991.

[12] M Dietze. *Messung und Berechnung der Innenströmung in hydraulischen Sitzventilen*. Darmstadt, 1996.

[13] C Latour. *Strömungskraftkompensation in hydraulischen Sitzventilen*. Aachen, 1996.

[14] J M Bergada, J Watton. A Direct Solution for Flowrate and Force Along a Cone-Seated Poppet Valve for Laminar Flow Conditions. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, 218(3):197-210, 2004.

[15] S Masuda, F Shimizu, M Fuchiwaki, K Tanaka. Effect of Damping Force on Hysteresis Characteristics of Poppet Valves. *Proceedings of the BATH/ASME 2024 Symposium on Fluid Power and Motion Control*, 2024.

Model Predictive Control with Discrete-valued Input for Digital Valve Unit - Comparison of Differential Evolution and Genetic Algorithm -

Yuhao Song* and Wataru Kobayashi*

*Graduate School of Science and Engineering, Okayama University of Science, Okayama, Japan
*E-mail: r23smw4nx@ous.jp, w-kobayashi@ous.ac.jp

Abstract

This paper presents a novel Model Predictive Control (MPC) framework integrated with Differential Evolution (DE) for real-time discrete-valued control of Digital Valve Units (DVU). Addressing the combinatorial challenge of DVU's on/off states, our DE-based MPC achieves high-precision pressure tracking in artificial muscle systems while significantly reducing computational load. Simulation results demonstrate: (1) 83.7% faster computation (91ms vs. 558ms) compared to Genetic Algorithm (GA) approaches; (2) superior tracking performance with an average error of 6.3kPa (19kPa max) versus GA's 7.05kPa (24kPa max); (3) stable real-time operation within strict sampling constraints. The proposed method effectively handles the nonlinear dynamics of discrete flow systems through simulation, establishing DE-MPC as a computationally efficient solution for DVU control, outperforming conventional methods in both speed and accuracy.

Keywords: Digital Valve Unit, Model Predictive Control, Differential Evolution, Genetic Algorithm

1 Introduction

Fluid power systems, which use hydraulic and pneumatic actuators, have been widely employed in industrial applications due to their high power density and reliability. Traditional fluid power control relies heavily on proportional and servo valves, which enable precise control by continuously adjusting the valve opening. However, these systems have inherent drawbacks, such as high costs, energy losses due to throttling, and complex maintenance requirements. Digital Fluid Power (DFP) [1] has emerged as an alternative approach that aims to improve energy efficiency, cost-effectiveness, and system reliability. Unlike conventional proportional valves, DFP utilizes discrete on/off valves that switch between fully open and fully closed states. By combining multiple discrete valves in parallel, DFP systems can approximate continuous flow control while minimizing energy losses. This discrete approach reduces leakage, enhances robustness, and allows for modular and scalable system designs. The rapid advancements in digital control and computing power have further accelerated the development of DFP. With the integration of advanced sensors, real-time optimization algorithms, and embedded control systems, DFP can now achieve performance levels comparable to traditional fluid power systems while significantly reducing energy consumption. This has led to increased adoption of DFP in various fields, including industrial automation, robotics, and heavy machinery. Furthermore, the growing demand for environmentally friendly and energy-efficient systems has positioned DFP as a key technology for the future. Many industries are shifting towards electrification and digitalization, and DFP aligns well with these trends by offering precise, responsive, and low-energy solutions. Researchers and engineers continue to explore new methods to enhance the efficiency and flexibility of DFP, making it a promising area of study.

1.1 Digital Valve Unit

Among various DFP technologies, Digital Valve Unit (DVU) has attracted considerable attention as a practical implementation of discrete valve control [2]. A DVU consists of multiple on/off valves with different flow capacities, arranged in parallel to enable stepwise flow rate adjustments. Figure 1 shows the typical structure of DVU, by using valves with flow rates of $Q, 2Q$, and $4Q$, the system can achieve flow rates ranging from Q to $7Q$ by selectively opening and closing the valves. Table 1 shows the changing of Flow rates. This binary combination mechanism allows for efficient fluid control without requiring expensive proportional valves. Despite its advantages in energy efficiency and cost reduction, DVU presents several challenges. The primary difficulty lies in the discrete nature of its control inputs. Unlike proportional valves that offer continuous flow adjustment, DVU can only produce a finite number of flow rates, making it difficult to achieve smooth and precise control. Additionally, the combinatorial explosion of possible valve states increases the complexity of optimizing the control inputs in real-time applications. The adop-

tion of DVU has been driven by its potential to enhance system reliability and longevity. Since on/off valves have simpler mechanical structures than proportional valves, they tend to have longer operational lifetimes and lower maintenance costs. Additionally, DVU systems can be designed to operate under high-pressure conditions while maintaining robust performance, making them suitable for aerospace, automotive, and industrial automation applications.

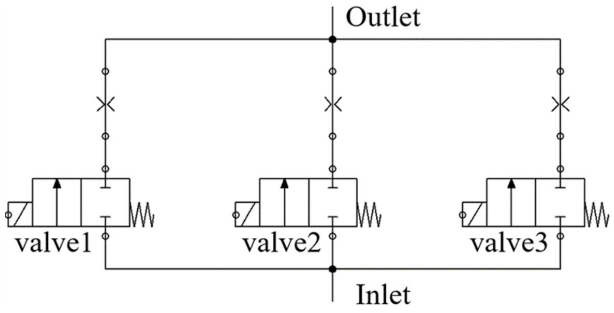


Figure 1: Digital valve unit (DVU)

Table 1: Flow rate change

0: closed valve		1: open valve	Outlet Flow ($\sum Q$)
$V1$ (Q)	$V2$ ($2Q$)	$V3$ ($4Q$)	
0	0	0	0
1	0	0	1
0	1	0	2
1	1	0	3
0	0	1	4
1	0	1	5
0	1	1	6
1	1	1	7

1.2 Problem statement and solution

The discrete nature of DVU poses significant challenges for conventional control methodologies [3]. Traditional controllers such as PID control and linear optimal control methods assume continuous control inputs, making them less effective for discrete-valued systems. While Model Predictive Control (MPC) has proven effective in handling system constraints and nonlinearity, its standard formulation is designed for continuous control variables, leading to computationally expensive optimization when applied to DVU [4]. Discrete-valued control requires specialized optimization techniques capable of efficiently searching large combinatorial spaces. Methods such as Genetic Algorithm (GA) have been explored but often suffer from prohibitive computational costs for real-time applications. Alternatively, heuristic and evolutionary approaches provide promising solutions by navigating the discrete search space more efficiently. To address these challenges, researchers have explored heuristic and evolutionary optimization techniques to improve control performance. Genetic Algorithm (GA) and Differential Evolution (DE) have been investigated for their ability to handle combinatorial optimization problems efficiently. DE, in particular, has demonstrated advantages in convergence speed and robustness for discrete-valued optimization, making it a promising candidate for DVU control. This study proposes a novel control framework that integrates MPC with DE to achieve real-time and high-precision control of DVU systems. By leveraging DE's capability to optimize discrete inputs efficiently, the proposed approach aims to reduce computational complexity while maintaining control accuracy. The effectiveness of the method is validated through simulation and experimental results, comparing DE-based optimization with traditional GA-based methods. The growing interest in DVU control methods highlights the need for more efficient algorithms capable of handling high-speed and high-precision operations. Future advancements in AI-based control and machine learning-driven optimization may further enhance the performance of DVU-based systems, leading to broader adoption across industries. The following sections provide a detailed discussion of the proposed control methodology, experimental validation, and performance evaluation.

2 Modeling of tap-water drive artificial muscle

To apply model predictive control, a model of objectives must be obtained. This study establishes a predictive model using system identification techniques for pressure control through an artificial muscle internal pressure response experiment [5]. This study selects pneumatic artificial muscles as the controlled object primarily due to their characteristic application features in discrete control. Previous research has demonstrated that effective control of artificial muscles can be achieved with a response time of approximately 0.2 seconds, making them an ideal testbed for validating discrete control algorithms. Specifically, the pressure response of artificial muscles is directly correlated with the switching states of discrete valve arrays. By adjusting the valve combinations in the DVU, rapid and precise control of internal pressure variations can be achieved. The 0.2-second response time not only meets the requirements for real-time control but also establishes a clear performance benchmark for algorithm optimization. This timescale effectively demonstrates the efficacy of control strategies while ensuring stable operation within hardware constraints. By selectively activating different valve combinations, the internal pressure response of the artificial muscle was measured. Based on the experimental data, a discrete-time model describing the dynamic behavior of the artificial muscle was obtained. Subsequently, the model was integrated into a MPC framework to optimize the valve control strategy and achieve real-time pressure tracking. Artificial muscles, as shown in Figure 2, are actuators that utilizes compressed air/water, with its structure composed of flexible materials such as rubber and fibers. Due to the characteristics, artificial muscles are lightweight while possessing a high power-to-weight ratio, enabling motion similar to that of biological muscles. Furthermore, unlike rigid electric or other hydraulic actuators, artificial Muscle can leverage their elasticity to achieve flexible movements, making them widely applicable in fields requiring human interaction, such as robotics and wearable devices. The most notable feature of the muscles is that their contraction and the force are determined by their in-

ternal pressure. Therefore, precise control of Artificial Muscle requires accurate regulation of the internal pressure.

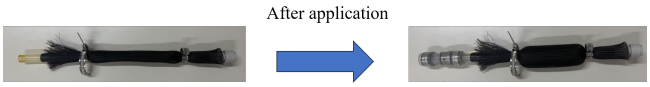


Figure 2: McKibben type artificial muscle

2.1 Experiment for system identification

In this study, the control of the internal pressure of the pneumatic artificial muscle is targeted, and a pressure control method using discrete flow adjustments with a Digital Valve Unit (DVU) is proposed. However, flow control with discrete inputs presents challenges, such as nonlinear characteristics and instability during transitions. Therefore, this research aims to achieve more accurate control by thoroughly analyzing the internal pressure response of the artificial muscle and constructing a mathematical model. As shown in Figure 3, the experiment setup consists a pressure sensor (PSE560-02, SMC), fourteen on/off valves (VDW31-6G-1-C4-XF, SMC), an artificial muscle with length of 150 [mm] and a microcomputer (ESP32, Espressif Systems Pte. Ltd.). In this experiment, fourteen on/off valves were individually or combinatorial opened and closed to measure the pressure variations under different flow conditions. Specifically, this experiment uses a Maximum Length Sequence (M-sequence) signal as an input signal to control the timing of water inflow and outflow. The M-sequence signal is a type of pseudo-random binary sequence with excellent autocorrelation properties, commonly used in system identification and dynamic response analysis. The M-sequence signal determines the on/off states of the valves, providing rich frequency content to excite the system. Water inlet and outlet flow are realized by switching designated valve combinations according to the signal.

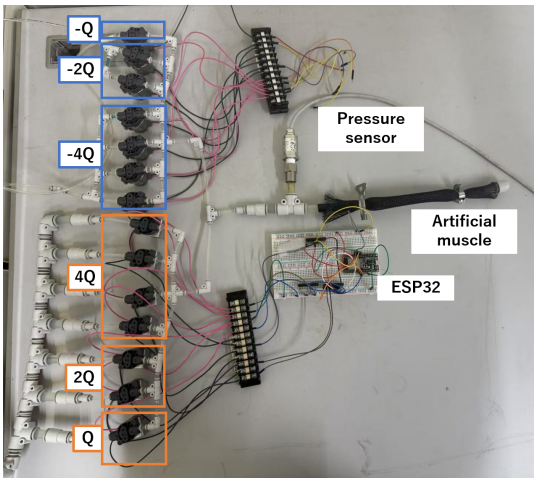


Figure 3: Experimental setup

As shown in Figures 4 and 5, the results clarify the characteristics of discrete flow control using a DVU and serve as fundamental data for modeling the pressure response of the artificial muscle.

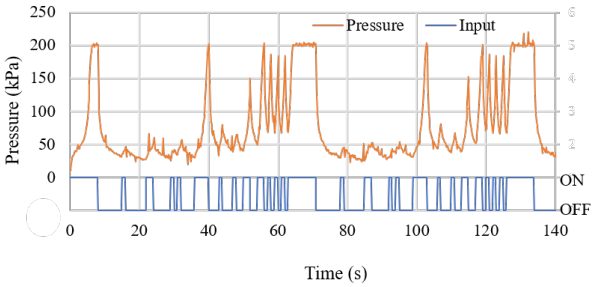


Figure 4: Pressure response of artificial muscle(1 valve)

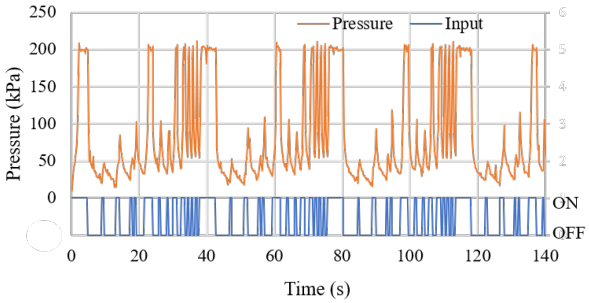


Figure 5: Pressure response of artificial muscle(7 valves)

2.2 Modeling

Based on the results of the previous step response experiment, a discrete-time first-order state-space model was constructed. The pressure response $y(k)$ of the artificial muscle is represented by the first-order delay system difference equation in Eq. (1) [6].

$$y(k) = \alpha y(k-1) + \beta u(k-1). \tag{1}$$

Here, $y(k)$ is the pressure of the artificial muscle at time step k, $u(k-1)$ is the input (applied voltage) in the previous time step, and α and β are constants obtained from the system step response. As a result of the experiment, seven different models were obtained, each corresponding to the state of the seven valves of the DVU. The discrete-time model can be expressed as Eqs. (2) to (8):

$$y(k) = 0.8966\, y(k-1) + 19.02\, u(k-1), \tag{2}$$
$$y(k) = 0.8229\, y(k-1) + 41.00\, u(k-1), \tag{3}$$
$$y(k) = 0.8026\, y(k-1) + 40.35\, u(k-1), \tag{4}$$
$$y(k) = 0.7474\, y(k-1) + 53.98\, u(k-1), \tag{5}$$
$$y(k) = 0.7318\, y(k-1) + 53.95\, u(k-1), \tag{6}$$
$$y(k) = 0.7132\, y(k-1) + 58.99\, u(k-1), \tag{7}$$
$$y(k) = 0.7381\, y(k-1) + 49.99\, u(k-1). \tag{8}$$

By using these models, the internal pressure of the artificial muscle can be predicted and utilized as foundational data for performing appropriate valve control. Note that the discretization method is zero-order hold and the model was discretized by using MATLAB. To verify the usefulness of the model, we performed one-step-ahead prediction using the generated

model. The results are shown in Figure 6, and the model has a precision of approximately 70%.

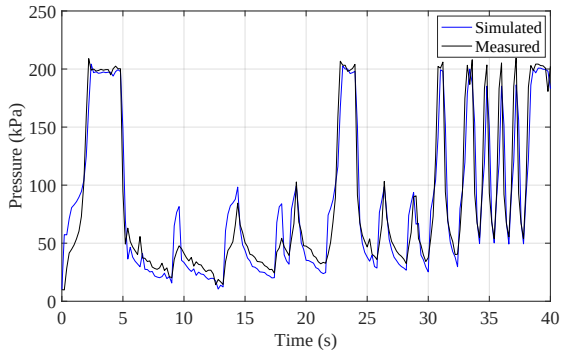


Figure 6: One step ahead prediction

3 Proposed control strategy using MPC with DE algorithm

This study proposes a novel control framework that integrates Model Predictive Control (MPC) with Differential Evolution (DE) to achieve real-time and high-precision control of Digital Valve Unit (DVU) systems, the block diagram of proposed method is shown in Figure 7. The discrete nature of DVU poses significant challenges for conventional control methodologies, which typically assume continuous control inputs. To address these challenges, the proposed approach leverages DE's capability to efficiently optimize discrete inputs, aiming to reduce computational complexity while maintaining control accuracy.

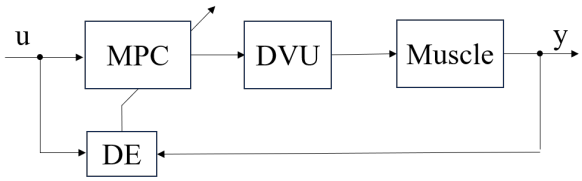


Figure 7: Block diagram of proposed method

3.1 Model predictive control

Model Predictive Control (MPC) is a control method that determines the optimal control inputs while predicting the future behavior of a system [7]. In MPC, the system's mathematical model is first used to calculate all the outputs within a specified prediction horizon. Next, a combination of inputs that minimizes an evaluation function (cost function) is sought, and this combination is applied. This calculation is repeated at each step, allowing the system to approach the target value while responding to its [8]. A major advantage of MPC is that, unlike simple feedback control, it takes into account the future state of the system when making control decisions. This enables high-precision control, even for systems with sudden fluctuations or nonlinear characteristics.As shown in Figure 8, constraints can be incorporated, improving the safety and energy efficiency of the system. For example, in fluid control systems, the number of valve openings can be restricted to reduce mechanical stress while achieving proper flow control. However, there are several challenges in applying MPC. To determine the optimal control input, an optimization problem must be solved in each control calculation, which can result in very high computational load [9]. This is particularly problematic for systems with discrete control inputs, such as Digital Valve Unit (DVU), where the number of combinations can increase exponentially. For example, in this study, seven on/off valves are used, and there are 15 possible combinations of flow rates ranging from $-7Q$ to $7Q$. Additionally, with a prediction horizon of 5 steps, considering all combinations results in 759,375 potential input candidates. Processing this enormous computational load within the sampling period is challenging, making real-time control difficult. Therefore, in this study, the Differential Evolution (DE) algorithm, a type of evolutionary computational algorithm, is applied to the optimization problem of MPC. DE is a method well-suited for combinatorial optimization problems, offering high search efficiency and the potential to significantly reduce computational costs. The next section will provide a detailed overview of the DE algorithm and explore the potential for improved control performance through its integration with MPC.

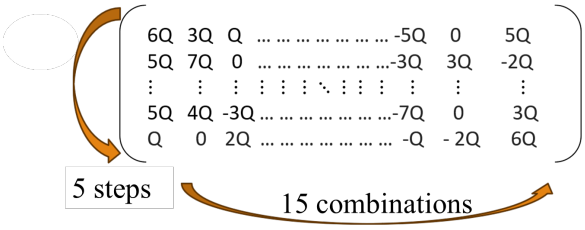


Figure 8: Prediction horizon

3.2 Differential evolution algorithm

Differential Evolution (DE) is a population-based heuristic optimization algorithm that belongs to the family of Evolutionary Computation algorithms [10]. It is primarily used to solve optimization problems that are high-dimensional, nonlinear, or multi-modal. The algorithm is inspired by the process of natural evolution, including mechanisms such as mutation, crossover, and selection, to optimize a given objective function. As shown in Figure 9, a distinguishing feature of DE is that it introduces differences between individuals in the population through mutation, thereby generating new candidate solutions [11]. This helps explore the search space effectively. Unlike traditional optimization methods, such as gradient descent, DE does not rely on derivative information of the objective function, making it well-suited for complex optimization problems that are nonlinear, discontinuous, or noisy. The core idea of DE is to simulate the natural evolutionary process, where individuals (candidate solutions) in a population are used to search for the optimal solution. In each generation (iteration), individuals undergo mutation, crossover, and selection processes, evolving progressively to better solutions. Unlike gradient-based algorithms, DE performs a global search, helping avoid local optima.

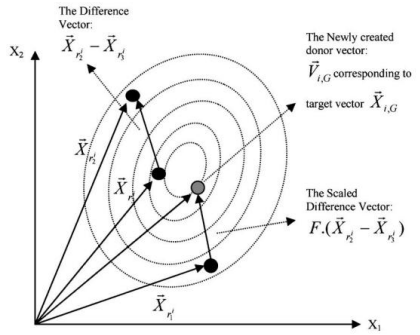


Figure 9: Illustrating a simple DE mutation scheme in 2-D parametric space [11].

Next, we will explain the DE algorithm process,the flow chart of DE as shown in Figure 10. Initialization is the first step in the DE algorithm, where a set of individuals (candidate solutions) is randomly generated to form an initial population. Each individual represents a potential solution to the problem at hand. Typically, these individuals are distributed randomly within the search space to ensure broad coverage.

In the initialization step, the population size and individual representation are crucial. A larger population size is usually chosen to ensure that the algorithm has enough diversity and global search capability. The population size is a user-defined parameter, such as 50, 100, or more individuals, depending on the problem's complexity. For example, in my research on controlling the DVU, each individual could represent a specific discrete valve combination that governs the flow rate.

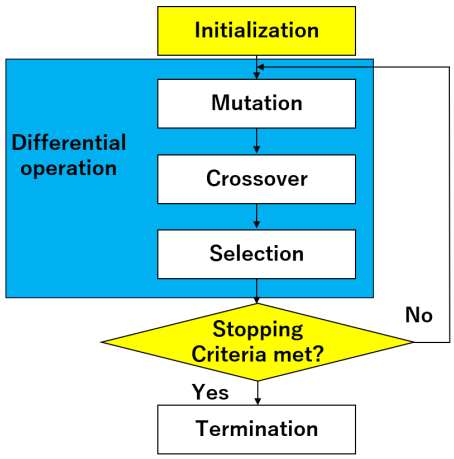


Figure 10: Flow chart of DE

Mutation is the core operation in DE, where new candidate solutions are generated by introducing differences between individuals. DE creates "mutant individuals" by selecting three distinct individuals from the population, computing the differences between them, and adjusting one of the individuals based on those differences. This is done using a scaling factor F that amplifies the difference. The mutation operation is typically performed using the following Eq.(9):

$$\mathbf{v_i} = \mathbf{x_{r1}} + F \cdot (\mathbf{x_{r2}} - \mathbf{x_{r3}}) \tag{9}$$

where $\mathbf{v_i}$ is the mutant vector for the target individual $\mathbf{x_i}$, $\mathbf{x_{r1}}, \mathbf{x_{r2}}, \mathbf{x_{r3}}$ are randomly selected individuals from the population, F is the mutation scaling factor, typically set to a value between 0 and 1. The crossover operation is used to generate "trial individuals" by exchanging information between the target individual and the mutant individual. The goal of the crossover is to combine the beneficial traits of the mutant and target individuals to create a new candidate solution. Crossover is typically controlled by a crossover probability CR, which determines the proportion of information exchanged between the two individuals. A common method is uniform crossover, where each gene (decision variable) is exchanged based on the probability CR. The mathematical formula for the crossover operation is represented by Eq. (10):

$$u_{i,j} = \begin{cases} v_{i,j}, & \text{if rand}_j \leq CR \text{ or } j = j_{\text{rand}} \\ x_{i,j}, & \text{otherwise} \end{cases} \tag{10}$$

where, $u_{i,j}$ is the j-th gene of the trial individual $\mathbf{u_i}$, $v_{i,j}$ is the j-th gene of the mutant vector $\mathbf{v_i}$, $x_{i,j}$ is the j-th gene of the target vector $\mathbf{x_i}$, rand_j is a random number between 0 and 1, CR is the crossover probability (typically between 0 and 1), j_{rand} is a randomly chosen index to ensure that at least one gene comes from the mutant vector. Crossover allows for the combination of different solutions, enhancing the algorithm's exploration capabilities and ensuring that the new individual inherits characteristics from both the mutant and target solutions. Selection is the process where the trial individual and the target individual are compared based on their fitness (objective function values). The better individual, with the lower fitness (objective function value), is chosen to survive to the next generation. The selection step guarantees the convergence of the algorithm since only individuals with better fitness are retained, and individuals with worse fitness are eliminated. This process ensures that the population evolves towards better solutions over time.

3.3 Simulation of proposed control with DE

The simulation was conducted to verify the effectiveness of the proposed control method. The goal was to optimize the valve operations to match the target pressure waveform as closely as possible using the Differential Evolution (DE) algorithm. In the control system using Arduino, pressure was measured in real time, and by applying the evolutionary algorithm, the optimal valve operations were determined. This process was simulated to demonstrate the control strategy [12] [13]. In this simulation, the target pressure P_{target} is defined as a sinusoidal wave, represented by the following Eq.(11):

$$P_{\text{target}}(k) = 75 \sin\left(\frac{\pi k}{25} - \frac{\pi}{2}\right) + 125 \tag{11}$$

where amplitude is 75 [KPa], indicating that the oscillation ranges from 50 to 200. The model uses the previously derived seven models, each corresponding to the number of on/off switches activated. When the flow rate combination is between 1 and 7, u (applied voltage) is equal to 1. When the flow rate combination is between -1 and -7, u is equal to 0. Eq.(12)

shows an fitness function in DE defined to minimize the evaluation function between the target pressure $P_{\text{target}}(k)$ and the actual pressure obtained from the simulation:

$$J(k) = \sum_{i=1}^{T} \left(P_{\text{target}}(k+i) - P_{\text{actual}}(k+i) \right)^2 \qquad (12)$$

where $P_{\text{actual}}(k)$ represents the pressure values obtained from the simulation. The objective is to find the sequence of flow rate combinations that minimizes this squared error. The parameters of DE is shown in Table 2.

Table 2: Parameters of DE

Parameter	Value
Population Size (*POP_SIZE*)	15
Maximum Generations (*MAX_GEN*)	70
Mutation Factor (F)	0.5
Crossover Probability (CR)	0.7

To verify the control performance of the proposed control method, it needs to be checked with simulation. The proposed control method is demonstrated to effectively manage the system, as shown in Figure 11. Within these 20 seconds, the average error between reference and output is 6.3[kPa], Maximum error is 19 [kPa].

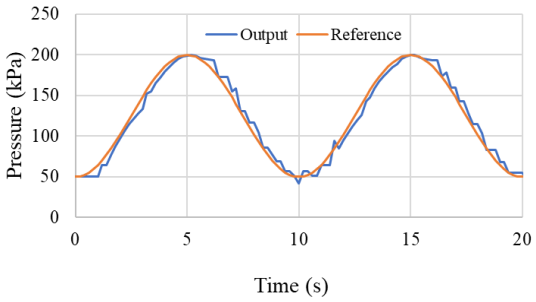


Figure 11: The result of simulation of proposed control

To maintain real-time control, accurately track the target pressure, and achieve optimal valve control while considering hardware performance, the computation time per step must be kept within 0.2[s]. The maximum computation time is 0.091 [s], and the minimum time is 0.07 [s].

4 Proposed control strategy using MPC with genetic algorithm

Since the DVU is based on discrete on/off control, the number of possible control input combinations becomes enormous, making optimization challenging. Genetic Algorithm (GA) has been widely used for this type of optimization; however, it has high computational costs and slow convergence speed. To demonstrate the effectiveness of the proposed method (MPC-DE), we compare it with the method combining MPC and GA (MPC-GA). The computational efficiency, control accuracy, and real-time applicability of both methods are evaluated and verified. The specific method is shown in the Figure 12.

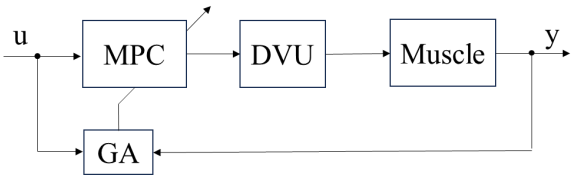


Figure 12: Block diagram of MPC-GA

4.1 Genetic algorithm

Genetic Algorithm (GA) is an optimization method that simulates the natural selection process [14]. By mimicking biological evolution, it continuously improves solutions to find the optimal solution to a problem. The basic steps of GA include population initialization, fitness evaluation, selection, crossover, mutation, and replacement operations [15]. The basic steps of GA is shown in Figure 13.

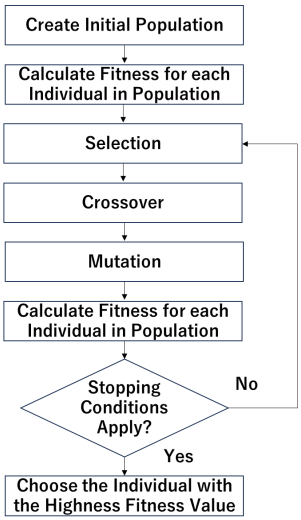


Figure 13: Flow chart of GA

First, the population is initialized. In this phase, the initial population is randomly generated, and the individual solutions are usually represented in some encoding form, such as binary encoding or real-valued encoding. The population size is predefined, and the number of individuals N affects the computational cost of the algorithm. Next, fitness evaluation is performed to assess the quality of each individual solution based on the objective function. The fitness function value reflects the quality of the individual solution. Typically, in minimization problems, the fitness function is the negative value of the objective function, while in maximization problems, the fitness function is the value of the objective function itself.

The Next, the selection operation is performed. The selection operation simulates the process of natural selection, with the goal of selecting individuals with higher fitness as parents. Common selection methods include roulette wheel selection and tournament selection. After selecting the parent individuals, the crossover operation is performed. The crossover operation generates new individuals by exchanging parts of the

parents' genes, simulating genetic recombination in nature. The most common form of crossover is single-point crossover, where a crossover point is randomly selected, and the genetic information of the parents is exchanged to generate two new offspring individuals. Assume the parent individuals are shown in Eq. (13):

$$P_1 = [x_1, x_2, \ldots, x_n] \quad \text{and} \quad P_2 = [y_1, y_2, \ldots, y_n]. \tag{13}$$

Let the crossover point be c, then the offspring individuals C_1 and C_2 are shown in Eqs. (14) and (15):

$$C_1 = [x_1, x_2, \ldots, x_c, y_{c+1}, \ldots, y_n] \tag{14}$$

$$C_2 = [y_1, y_2, \ldots, y_c, x_{c+1}, \ldots, x_n]. \tag{15}$$

Next, the mutation operation is performed. Mutation involves randomly changing a part of the individual's genes, which introduces new genetic combinations and increases the diversity of the population, helping to avoid the algorithm getting stuck in a local optimum. Mutation is usually done by flipping binary bits or randomly changing the value of a gene in a real-valued encoding. For example, if the individual is shown in Eq. (16):

$$x_i = [x_1, x_2, \ldots, x_n], \tag{16}$$

the mutation operation may randomly select a gene x_k to modify, resulting in a new individual x_i': After performing crossover and mutation, a replacement operation is performed. The replacement operation determines which individuals will enter the next generation. A common replacement method is the elitism strategy, which ensures that the best individuals from each generation directly enter the next generation. Finally, the termination condition is used to determine whether the algorithm should stop. Common termination conditions include reaching the maximum number of generations or the fitness function reaching a predefined threshold.

4.2 Simulation of proposed control with GA

A simulation was conducted to compare the computation time of MPC-DE and MPC-GA, evaluating the efficiency of the proposed method. The simulation conditions were unified for both approaches, and the difference in computation time for the same flow control task was assessed. As shown in Figure 14, the computation time results indicate that MPC-GA completed the optimization in a maximum of 558 [ms] and a minimum of 325 [ms], whereas MPC-DE completed it in a maximum of 91 [ms] and a minimum of 70 [ms]. This demonstrates that MPC-DE achieved a reduction in maximum computation time by approximately 83.7% and a reduction in minimum computation time by approximately 78.5% compared to MPC-GA. Additionally, in terms of control accuracy, MPC-DE achieved an average error of 6.3 [kPa] and a maximum error of 19 [kPa], while MPC-GA had an average error of 7.05 [kPa] and a maximum error of 24 [kPa]. As shown in Figure 15, these results clearly indicate that MPC-DE can significantly reduce computation time while maintaining a similar level of control accuracy as MPC-GA. This highlights the advantage of MPC-DE, especially in control systems where real-time performance is critical.

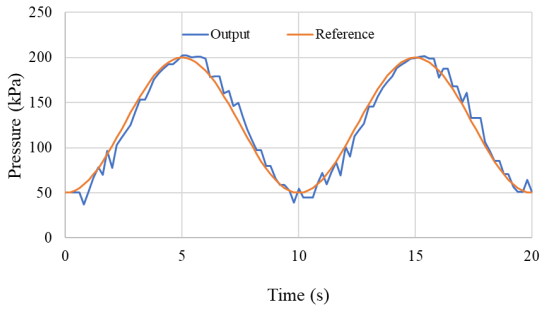


Figure 14: The result of simulation of MPC-GA

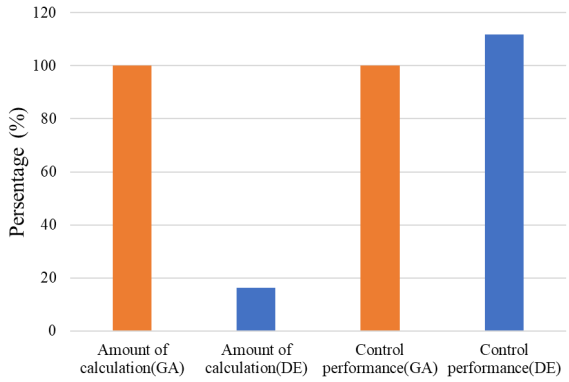


Figure 15: Comparison of DE and GA

5 Conclusion

This study presents an innovative control strategy that integrates Model Predictive Control (MPC) with Differential Evolution (DE) algorithm to address the real-time optimization challenges in Digital Valve Unit (DVU) systems with discrete inputs. Through systematic research involving theoretical analysis, simulation verification, and experimental testing, the following significant achievements have been made: In terms of algorithm design, this research creatively combines DE's global search capability with MPC's predictive optimization features to construct a hybrid optimization framework suitable for discrete control problems. Specifically, for DVU system characteristics, we developed a system identification method based on M-sequence and a multi-model switching strategy, effectively solving the modeling challenges of nonlinear systems. Simulation results demonstrate that in 120-second tracking control, this method not only reduced maximum computation time from 558[ms] (MPC-GA) to 91[ms] - an 83.7% performance improvement - but also maintained average tracking error within 0.735[kPa], showcasing excellent real-time performance and control accuracy. Regarding engineering implementation, the research team developed an embedded control system based on ESP32 microcontroller. Through optimized algorithm implementation and code architecture, the theoretical method was successfully transformed into practical application. Experimental data indicate that the system exhibits

excellent dynamic response characteristics in pressure tracking control, with maximum overshoot below 5% and steady-state error less than 1%, fully meeting industrial application requirements. Notably, this method requires relatively low hardware resources, laying the foundation for widespread adoption on low-cost controllers. Compared to existing technologies, this research features three main innovations: 1) A novel DE-based discrete optimization method that effectively solves the applicability issues of traditional continuous optimization algorithms in discrete control; 2) Development of a multi-model predictive control strategy that significantly enhances system adaptability to nonlinear characteristics; 3) Achievement of efficient embedded implementation, providing new ideas for intelligent upgrades of digital fluid control systems. The research findings have broad application prospects in industrial automation, robotic control, and related fields. Particularly in pneumatic servo systems requiring high precision and fast response, this method demonstrates remarkable technical advantages. Future research will focus on: 1) Incorporating deep learning to further enhance algorithm adaptability; 2) Expanding applications in multi-actuator coordinated control; 3) Developing smart diagnostics and maintenance functions for Industry 4.0. These efforts will advance digital fluid control technology toward smarter and more efficient solutions, providing new technical support for manufacturing digital transformation.

Nomenclature

Designation	Denotation	Unit
$y(k)$	System output (artificial muscle pressure) at time step k	kPa
$u(k)$	Control input (valve command) at time step k	-
α, β	Model coefficients (system dynamics parameters)	-
$P_{\text{target}}(k)$	Target pressure trajectory k	kPa
$P_{\text{actual}}(k)$	Actual measured pressure k	kPa
Q	Base flow rate of a single valve	L/min
F	DE mutation scaling factor	-
CR	DE crossover probability	-
POP_SIZE	Population size	-
MAX_GEN	Maximum generations	-
k	Discrete time step index	-
T	Prediction horizon length	steps
$V1, V2, V3$	Valve states (1: open, 0: closed)	-

References

[1] Matti Linjama. Digital fluid power â state of the art. *The Twelfth Scandinavian International Conference on Fluid Power*, 2011. Tampere University of Technology, Department of Intelligent Hydraulics and Automation, May 18–20.

[2] M. Paloniitty, M. Linjama, and K. Huhtala. High-performance digital hydraulic valve system. *International Journal of Fluid Power*, 16(2):77–86, 2015.

[3] T. O. Andersen and M. R. Hansen. Discrete-valued control of digital hydraulic systems using binary valve patterns. *IEEE Transactions on Industrial Electronics*, 65(9):7202–7211, 2018.

[4] Anders H. Hansen, Magnus F. Asmussen, and Michael M. Bech. Energy optimal tracking control with discrete fluid power systems using model predictive control. *The Ninth Workshop on Digital Fluid Power*, 2017. Aalborg University, September 7–8.

[5] B. Tondu and P. Lopez. Modeling and control of McKibben artificial muscle robot actuators. *IEEE Control Systems Magazine*, 20(2):15–38, 2000.

[6] C. P. Chou and B. Hannaford. Measurement and modeling of mckibben pneumatic artificial muscles. *IEEE Transactions on Robotics and Automation*, 12(1):90–102, 1996.

[7] J. M. Maciejowski. Predictive control with constraints, 2002.

[8] Mayuresh V. Kothare, Venkataraman Balakrishnan, and Manfred Morari. Robust constrained model predictive control using linear matrix inequalities. *Automatica*, 32(10):1361–1379, 1996.

[9] A. Bemporad and M. Morari. Control of systems integrating logic, dynamics, and constraints. *Automatica*, 35(3):407–427, 1999.

[10] R. Storn and K. Price. Differential evolution â a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11(4):341–359, 1997.

[11] S. Das and P. N. Suganthan. Differential evolution: A survey of the state-of-the-art. *IEEE Transactions on Evolutionary Computation*, 15(1):4–31, 2011.

[12] Dan Henriksson, Anton Cervin, and Karl-Erik Årzén. TRUETIME: Real-time control system simulation with MATLAB/Simulink. 2003.

[13] D. Verscheure et al. Time-optimal mpc for mechatronic applications. *IEEE Transactions on Control Systems Technology*, 17(5):1185–1197, 2009.

[14] J. H. Holland. *Adaptation in Natural and Artificial Systems*. University of Michigan Press, 1975.

[15] D. E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, 1989.

The Multi-Pump System Combinatorial Problem: A Filtering Approach Using Genetic Algorithms

Artur Tozzi de Cantuaria Gama , Samuel Kärnell , and Liselott Ericson

Fluid and Mechatronic Systems, Linköing University, Linköping, Sweden
E-mail: artur.tozzi@liu.se

Abstract

Hybrid and fully electric heavy machinery introduce new possibilities for hydraulic system design. Due to their lower energy density compared to conventional combustion engine systems, improving hydraulic efficiency is crucial. Electro-hydraulic actuators can achieve this but often require high installed power, as each actuator must be sized for maximum demand. The multi-pump system (MPS) in this paper addresses this by allowing all hydraulic machines to serve any actuator via a network of on/off valves, reducing losses and installed power. However, its multiple degrees of freedom make optimal operation non-trivial. This paper proposes a filtering strategy using a genetic algorithm to identify efficient operating points for the MPS. Although applicable for larger systems, the results here focus on an MPS with two pumps and one actuator as an example. A quasi-static system model is introduced, which the GA uses to determine steady-state control signals that minimise power consumption. The results highlight ideal operating conditions, significantly narrowing the range of viable valve combinations and pump/motor speeds. Finally, the paper discusses the limitations of the approach and its potential extension to more complex multi-pump systems for the development of dynamic control strategies.
Keywords: Multi-pump system, optimisation, genetic algorithm, hydraulic system modelling

1 Introduction

Electrification is nowadays one of the main trends in developing more efficient mobile machinery. Fully electric or hybrid vehicles aim to achieve a lower carbon footprint by removing or downsizing the conventional combustion engine, respectively. An electrified vehicle can use the conventional hydraulic system with variable displacement pumps and throttling valves to operate its functions, but these often maintain the high losses, leading to average efficiencies of around 30% [1] for the hydraulic system. Considering that battery-powered vehicles are less energy-dense, the low efficiency becomes a limitation and alternative solutions becomes more relevant.

One approach is to implement non-centralized architectures. For example, [2] compares baseline load-sensing system with displacement control, where a single prime mover drives multiple variable displacement pumps, and an alternative using electro-hydraulic actuators. A hybrid approach that combines displacement control with a smaller load sensing system is also presented in [3]. While all solutions offer higher efficiencies, they often require either increased pump capacity, installed power, or both.

An alternative solution is provided in [4] and [5], called variable-speed drive network system. This approach utilizes fixed displacement pump/motors driven by variable-speed electric machines. By eliminating valves and short-circuiting chambers or interconnecting them through the hydraulic machines, the system reduces installed motor power, but may require specific tuning for each vehicle.

Expanding on these approaches, [6] presented a new modular design, termed the multi-pump system (MPS). This architecture employs multiple fixed displacement hydraulic machines operated by variable speed electric machines, which for simplicity will often be called e-pumps in this paper, and a set of discrete valves to connected any pump/motor port to any actuator chamber. This shared configuration aims to avoid increasing the total installed power.

1.1 Combinatorial complexity in the MPS

One challenge with this system is the extremely large number of control signal combinations. A system with two main e-pumps, two actuators, and an auxiliary pump/motor would require a total of twenty valves, leading to $2^{20} = 1,048,576$ valve combinations. A relatively coarse discretization of the e-pump speed into 20 steps further increases this number to $2^{20} \times 20^2 = 419,430,400$.

A similar challenge is addressed in [7], where the authors propose a graphical approach to identify viable flow paths and systematically reduce the combinatorial complexity. Their

study on two individual metering systems demonstrates that this strategy simplifies optimization while improving energy efficiency. However, it focuses on the actuator side and does not consider the role of pump/motor control in the optimization process.

A previous study by the authors of this paper conducted a backward calculation analysis to optimize the e-pumps control to minimize power consumption for the operation of one actuator [8]. The architectures considered different number of e-pumps, operating in two quadrants, for each chamber. Additionally, the paper described a sizing strategy for the hydraulic machines using scaling laws described in [9]. The system analysis showed that:

- Using three pumps for the larger chamber and two for the smaller one was the most efficient architecture, as the system can better distribute the flow to avoid low-efficiency operating points;
- Some operating regions provide opportunities to unevenly divide the flow between hydraulic machines to increase overall efficiency;
- The number of e-pumps speed combinations required large computational storage using matrices calculations, indicating that expanding the system to include more actuators and four-quadrant operation would be prohibitive.

To expand on the previous points, this paper presents a method applicable to various MPS architectures to determine the most relevant operating modes for the e-pumps and valves. Assuming four-quadrant operation, analytically determining the viable flow paths can become cumbersome. To address this, a genetic algorithm (GA) is employed to explore the design space of control inputs — namely, valve signals and pump/motor speeds — with the objective of minimizing power consumption or maximizing energy recuperation.

The approach uses a quasi-static model and does not consider a drive cycle or the transition costs between different operating points. Instead, as a result of the explorative aspect of a GA, the method provides a list of recommended operating modes for each pair of actuator force and speed. This significantly reduces the number of control options by indicating, for example, when to activate the second e-pump or when to regenerate energy. The primary aim of this paper is to demonstrate this approach for an intuitively understandable system, with the broader goal of extending it to more complex configurations.

Section 2 describes the combinatorial problem in more detail and establishes the main goals of the quasi-static model and optimisation algorithm. The modelling strategy and the GA method are detailed in sections 3 and 4, respectively. The simulation results for an MPS with two pump/motors and one actuator is presented in section 5 to show the capabilities of this approach. Section 6 then discusses how this method can be expanded to more complex systems and how the results can support the development of dynamic control strategies for the MPS.

2 System overview and problem formulation

Figure 1 shows a multi-pump system with two fixed displacement e-pumps (P_1 and P_2) powered by electric machines with variable speed drives ($VSDs$) connected to an actuator (A_1) and an auxiliary pressure source (PS_{aux}) through a set of on/off valves.

This system minimizes throttling losses by eliminating proportional valves and instead relies on flow control through variable-speed machines. While proportional valves can assist with control, their inclusion would greatly increase the problem complexity and introduce losses that the architecture aims to avoid. Furthermore, this study focuses on steady-state conditions, rendering partially open valves irrelevant for the current analysis.

The valves are named according to the components they are connecting, e.g. V_{1A1P1A} allows flow from pump port P_{1A} to actuator chamber A_{1A}. This simplified system is used as a reference for the discussions in this paper, but the method applies, in theory, to a system with any number of components.

This design allows energy regeneration and recuperation and can be expanded. Including another actuator means adding another row of valve connections to the pump/motors, adding e-pumps requires another column of valves. A constant pressure source is considered for the auxiliary line to simplify the analysis.

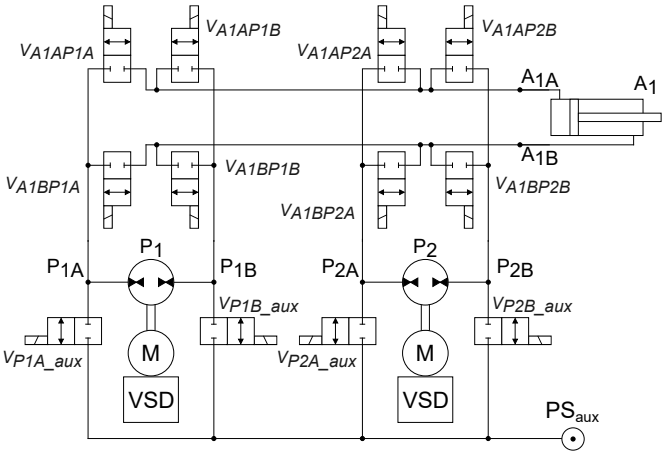


Figure 1: A diagram of a multi-pump system with one actuator controlled by two pumps with an auxiliary constant pressure source.

For the system in fig. 1, identifying valid valve combinations manually is feasible, especially because operating the hydraulic machines in four quadrants is redundant with a single actuator. Figure 2 shows three ways to achieve the same speed and force in the actuator, with each case detailed below. a) and b) show the conventional operating case with opposite pump speeds, while c) shows a regenerative case with the chambers interconnected. A similar configuration applies to motor operation.

In an operating scenario where the actuator extends against a resistive force and assuming that P_2 is off, P_1 can operate as a pump with a positive or negative speed by changing the set of

open valves from V_{A1AP1A}, V_{A1BP1B}, and V_{P1B_aux} to V_{A1AP1B}, V_{A1BP1A}, and V_{P1A_aux}, respectively.

P_1 can also operate with V_{A1AP1A}, V_{A1BP1A}, and V_{P1B_aux} open, as illustrated in fig. 2 c), depending on the external force and actuator area ratio. Adding the second machine enables alternative operating modes with two pumps or one pump and a motor, for example. Since they can be controlled individually, the system can achieve the same behaviour with different combinations of angular speeds.

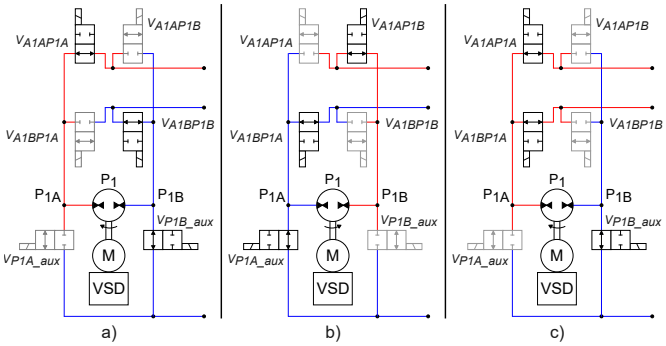


Figure 2: Operating examples for the multi-pump system in fig. 1 with a single pump. a) Pump mode with positive speed. b) Pump mode with negative speed. c) Pump mode with regeneration - chambers interconnected.

A complete system analysis would have to consider all the possible flow paths and directions to determine the optimal solution. As the number of components increases, the problem quickly becomes too complex for manual evaluation. The optimisation approach using a GA is one way to circumvent this problem for both simple and more complex systems.

2.1 Model and optimization requirements

This paper presents an approach to the aforementioned problem in two parts: a system model and an optimisation strategy to identify optimal operating conditions. The main objectives for the new model were defined as:

- Develop simplified component models that can be extended, balancing computational cost with system solvability;
- Ensure flexibility in defining the number of pumps and actuators, as the optimal architecture may vary across different applications;
- Enable both forward and backward simulation by allowing the inputs and outputs to be freely assigned, facilitating testing and results verification;
- Allow expansion of the loss models and facilitate post-processing calculation to estimate additional losses;
- Ensure smooth four-quadrant operation, eliminating the need for explicit checks on pressure difference to determine flow direction.

It was also determined that a filtering approach was needed to reduce the scope of the problem. Considering the large number of combinations, a strategy to identify the possible solutions to the system was sought after, thus this optimisation approach would need to:

- Handle a model that is non-linear and discontinuous and that can vary in size and complexity;
- Provide a reduced scope of control options to the system, thus facilitating future analysis and dynamic simulations.

The solution in this paper addresses these aspects by introducing, in sec. 3, a quasi-static model of the system based on a combination of symbolic expressions supporting multi-quadrant operation with a root-finding solver to compute the systems states. This allows the definition of different inputs and outputs for various types of simulations.

A genetic algorithm, described in sec. 4, explores the design space to find locally optimal solutions to the hydraulic machines' angular speeds and the corresponding valve signals to operate the system at steady-state conditions. The results greatly reduce the number of control options and provide insights into when the system transitions between different operating modes.

3 Quasi-static model of the hydraulic system

Based on the proposal for backward and forward simulation, the following structure is used to model the system:

- Each component may contain one or more power ports: hydraulic and mechanical. Electrical ports are omitted in this analysis;
- Ports hold the effort ($\mathcal{E}$) and flow ($\mathcal{F}$) values from one the the component's domains and share this information with other components;
- Each component includes one or more equations, including loss expressions, that define its flow and effort values;
- The communication between components is done through mechanical or hydraulic nodes, which transmit common effort values and calculate the flow balance;
- Each port is defined as a provider or receiver, thus setting the expected variable sign for the first quadrant operation. This approach handles the challenge of defining the variables signs but must be properly considered when defining the equations based on the other components.

Figure 3 illustrates a pump/motor component from fig. 1 along with its equations, connection ports, and nodes. The nodes connect to generic ports from other components, demonstrating the transmission of information. These hydraulic nodes act as abstract components that determine the flow balance between ports, mimicking the behaviour of a hydraulic volume but without accounting for dynamics.

The hydraulic flow (q) arrows show the expected positive sign for this model. The system models, for example, that the nodes receive flow from the pump/motor ports and from the ports on the bottom, and provides flows to the top ports. If the solver determines a negative value, then the flow direction is reversed. A similar consideration is performed for the valves and actuator. A more detailed explanation of the variables and equations is available in sections 3.1 and 3.4.

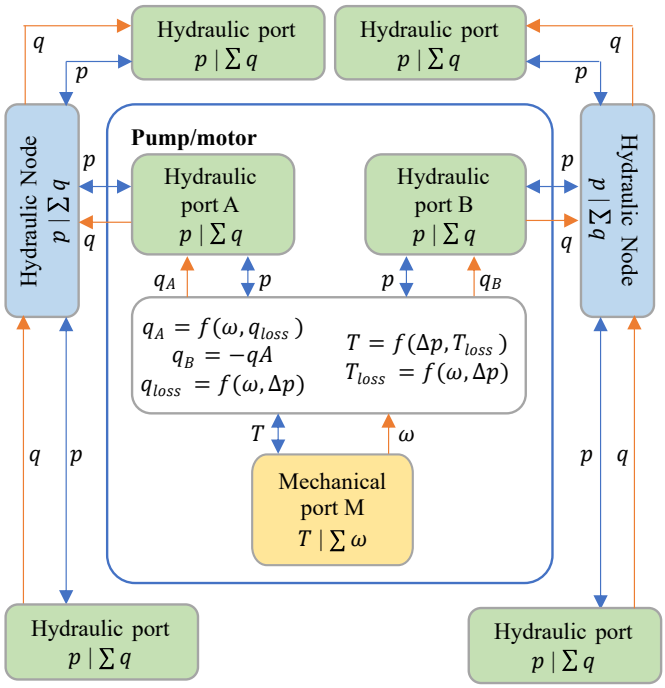


Figure 3: A diagram of a pump motor component with connection ports and nodes and the equations used to model the component.

The model assumes a positive input power at the pump/motor shaft is transmitted through the system and becomes a positive output power at the actuator rod. This means that the expected flow direction from the component ports must be taken into account in the equations, as presented in the next sections. The pressure source is an ideal component and simplifies the analysis by maintaining a constant pressure at the auxiliary line to the pump/motors.

The following subsections present the quasi-static models of pump/motors and their loss models, the linear actuator, valves, and nodes. The final subsection explains how these models are used to formulate the set of equations describing system behaviour and how variable values are provided to the solver to compute the solution.

3.1 Pump/motor model

The pump/motor model in fig. 3 considers two hydraulic ports and one mechanical port. Mechanical power going in the component, so pump operation, is assumed positive. The pump/motor is then expected to transmit positive power to the other components, so both hydraulic ports are assumed to have positive power leaving the component. The equations then explicitly define that a positive flow in one port would lead to a negative flow in the other. A summary of this logic and operating quadrants can be seen in table 1.

Table 1: Pump/motor quadrants, variable signs, and operating modes.

	T	ω	q_A	Δp	q_B	Mode
1st quad	+	+	+	+	-	Pump
2nd quad	-	+	+	-	-	Motor
3rd quad	-	-	-	-	+	Pump
4th quad	+	-	-	+	+	Motor

The equations that describe the pump/motor component are shown below, where $q_{p,A}$ is the flow at port A, D_p is the machine volumetric displacement, ω_p is the shaft angular speed, $q_{p,loss}$ is an expression for the internal leakage, T_p is the torque, Δp_p is the pressure difference between the hydraulic ports, and $T_{p,loss}$ is the torque loss.

$$q_{p,A} = D_p \omega_p - q_{p,loss}, \tag{1}$$

$$q_{p,B} = -q_{p,A}, \tag{2}$$

$$T_p = D_p \Delta p_p - T_{p,loss}. \tag{3}$$

The e-pumps' displacement vary with the number used in the system, and the sizing process follows a similar strategy to [8], with the machines sized to provide between 10% to 20% more flow than the maximum required actuator flow. The difference derives from how close the flow values are, since the new machines are sized with integer volumetric displacement. It also brings more flexibility to system operation at higher speeds.

3.1.1 4-quadrant loss models

The loss models are based on experimental data for a 32.4 cm^3/rev pump. For different machines' sizes, the losses and maximum angular speed are scaled according to the laws described in [9]. The scaling laws are shown below, where λ is the scaling factor and ω_{nom} is the nominal angular speed. The subscript ref means it is the reference pump value.

$$\lambda = \sqrt[3]{\frac{D_p}{D_{p,ref}}} \quad \Rightarrow \quad \begin{cases} \omega_{nom} = \frac{\omega_{nom,ref}}{\lambda}, \\ q_{p,loss} = \lambda^2 q_{p,loss,ref}, \\ T_{p,loss} = \lambda^3 T_{p,loss,ref}. \end{cases} \tag{4}$$

This provides loss values that cover the first quadrant of operation for the machine but to follow the model proposal, the $q_{p,loss}$ and $T_{p,loss}$ expressions should handle four quadrant operation. To cover this, the loss maps are extrapolated for the other scenarios (by inverting and mirroring the data as needed). To create the expressions, the Matlab Curve Fitting Toolbox was employed [10], generating non-linear polynomial equations to approximate the loss values as a function of the angular speed and pressure difference.

3.2 Linear actuator model

The actuator model assumes no losses and can therefore be described by the following equations, where $q_{c,A}$ and $A_{c,A}$ are

the flow and area of the piston side, respectively, $q_{c,B}$ and $A_{c,B}$ are the flow and area of the rod side, v_c is the actuator speed, F_c the actuator force, $p_{c,A}$ is the pressure in the piston side and $p_{c,B}$ in the rod side:

$$q_{c,A} = A_{c,A} v_c, \tag{5}$$

$$q_{c,B} = -A_{c,B} v_c, \tag{6}$$

$$F_c = p_{c,A} A_{c,A} - p_{c,B} A_{c,B}. \tag{7}$$

3.3 Valve model

All the valves in a single architecture have the same size. They are set to have a 10 bar pressure drop at the maximum flow, determined based on the maximum actuator flow ($q_{c,max}$). This is to cover cases of energy regeneration, since all the flow of a chamber can be redirected to the other chamber of the actuator, while the pump/motor(s) compensate for the area difference.

The model assumes no leakage; therefore, the flows $q_{v,1}$ and $q_{v,2}$, representing the flows in the pressure port and the actuator port, respectively, are determined by a linear flow coefficient C_v and the pressure difference Δp_v. The variable S_v denotes the control signal, which can take binary values of 0 or 1. The pressure drop is the same in both flow directions.

A linear model simplifies the calculations without significantly affecting the results. The primary impact is a slight preference for lower flow through the valves, which may lead to solutions that favour using two pumps more frequently.

$$q_{v,1} = S_v C_v \Delta p_v, \tag{8}$$

$$q_{v,2} = -q_{v,1}, \tag{9}$$

$$C_v = \frac{q_{c,max}}{10e5}. \tag{10}$$

3.4 Node model

The node is a component responsible for combining the information from the ports of other components. In this analysis, they directly transmit the effort $\mathcal{E}$ variables (pressure, torque, force) to the connected ports.

The flow variables $\mathcal{F}$ (hydraulic flow, linear speed, angular speed) have their signs determined directly by the components, while the node performs the summation. Thus, defining $\mathcal{F}_N$ as the the total node flow and $\mathcal{F}_\#$ as a generic representation of the other flow variables, the node equation becomes

$$\mathcal{F}_N = \sum \mathcal{F}_\#. \tag{11}$$

3.5 Symbolic representation of the system

A *Python* program utilises these definitions to construct systems with varying numbers of pumps and actuators, generating the required valve components. The *SymPy* library [11] facilitates the construction of symbolic expressions based on the components equations. For the system in fig. 1, for example, the program generates 38 equations with 54 variables. These include all node equations and component loss expressions, therefore the value of 16 variables must be defined to solve the system.

Since some models are non-linear, the equations are converted into numerical expressions. The *SciPy* library [12] enables the use of various solvers to compute a numerical solution when the necessary parameters are provided.

This model works for backward and forward simulations based on which symbols are chosen as variables. For instance, with a fixed pressure for PS_{aux}, an external force value for F_C, and the signals values (S_v) for the valves, 14 variables are set. For forward simulation, one can define each pump angular speed $\omega_{p,\#}$ and solve to find the the actuator speed v_c. Alternatively, if the actuator drive cycle is given, it is possible to calculate both $\omega_{p,\#}$ by defining the sharing ratio between e-pumps. Figure 4 summarizes this process and show the integration with the genetic algorithm, detailed in section 4.

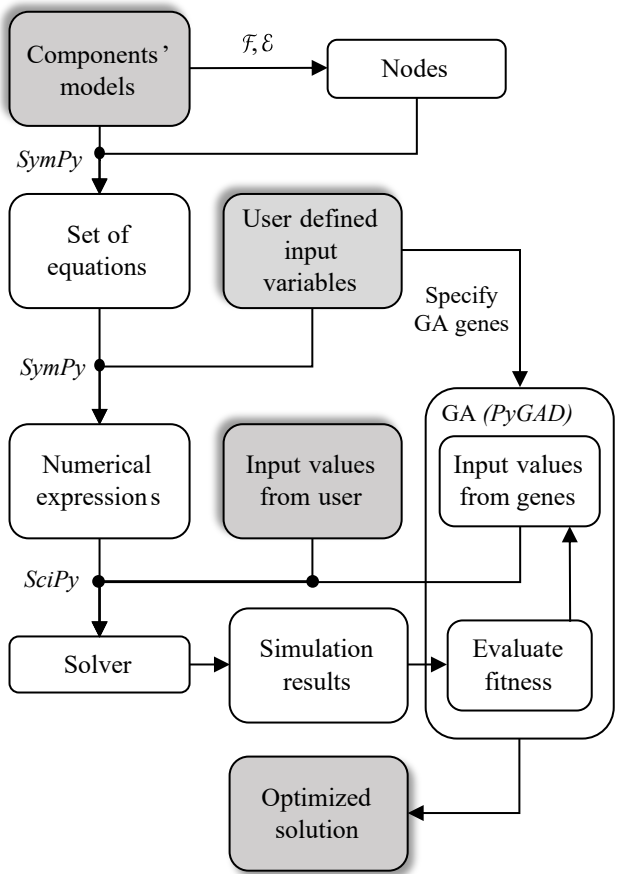


Figure 4: A diagram showing the model preparation process. Defining different input variables allow for forward and backward simulations.

4 Genetic algorithm (GA) method

As mentioned before, not all valve combinations are useful to the system, e.g. having a single valve open. In fact, many valve combinations are most likely infeasible or not useful but since they are coupled to the pump/motor speeds, it is not simple to define the available options.

Looking at the system as a control optimization problem is interesting to filter out the operating modes for the pump/motors and valves combinations. The genetic algorithm fits well into this approach because it can manage non-linear discontinuous problems and can explore a large area of the design space, thus reducing the likelihood of getting stuck in local minima.

GAs are inspired by natural evolution, where the fittest individuals are more likely to survive and pass on beneficial genes to their offspring. A chromosome is a set of genes, which, in this paper, represent the control signals to valves and pump/motors. In a simple example, two solutions (two members of the population), can provide genes to form a new chromosome to a new individual that can better fit the system's operation. To increase the variation, some genes may undergo mutation, slightly modifying the inherited information [13].

Figure 5 illustrates these steps and expands on the GA block shown in fig. 4. For each chromosome the GA creates, the genes values are sent to the solver to calculate the remaining variables. The complete results are then sent back to the GA and used to compute the fitness for each member of the population.

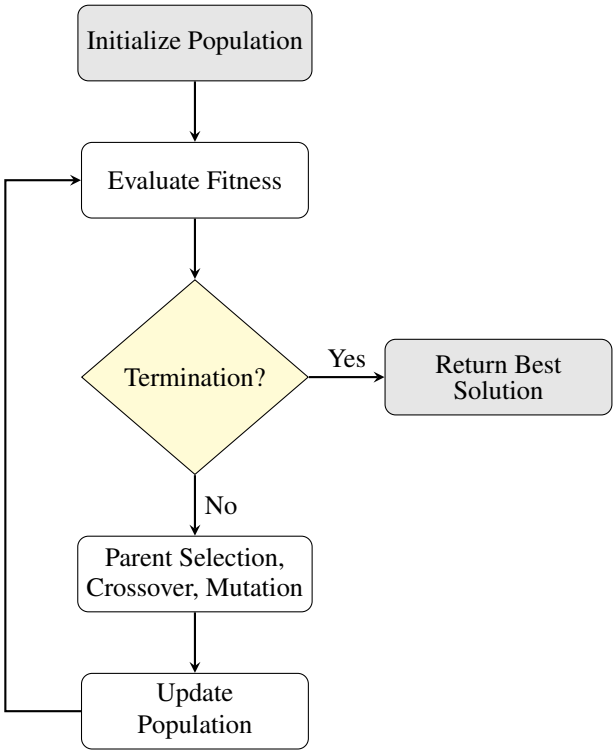


Figure 5: A simple genetic algorithm workflow

The algorithm requires a few other parameters, but the most relevant for this explanation are the population size, which determines how many individuals exist; a fitness (or cost) function that defines how good a certain individual is; and a termination criterion, which can be a fixed number of generations or another condition, such as convergence to a solution or minimal improvement over successive generations.

The reasons for choosing this method can be summarized to:

- The probabilistic approach to creating the population and offspring increases the chance of exploring distant solutions in the design space, which can converge in parallel to different local minima until the best solution is found;
- Many optimization strategies work well for continuous or discrete variables, but this system operates with binary signals for valves and continuous values for pump/motors. The GA can assign different constraints to each gene to overcome this, and the crossover step can provide large variation in values, which might lead to better exploration;

The *PyGAD* library [14] provides an easy-to-use framework for applying the algorithm in different problems. It includes standard options for gene constraints, crossover, and mutation steps, streamlining the implementation. The user is responsible for defining key parameters, such as population size, number of generations, and the fitness function.

Based on the system of equations described in the previous section, the GA can be set to create a population with any number of variables. This flexibility allows setting the valve signals and solving for the pump/motor angular speeds, or vice versa. In the more general case, both sets of values represent the genes of each chromosome. The number, types, and constraints of the genes are set from the components' parameters.

4.1 Fitness function definition

The fitness function provides a value that informs the GA on the quality of the individual. The *PyGAD* library tries to maximize the fitness value, which can combine different calculations and checks. To determine a good solution, the primary objective is to calculate the total power consumed or regenerated. The sign convention is set as: negative power represents consumption, positive power indicates regeneration.

Some constraints perform preliminary checks before calculating the fitness values. Solutions where only one valve is open can be given a value of $-inf$ to indicate infeasibility. Similar checks also exist to prevent cases where a pump/motor port is blocked if no viable path for the flow exists, if the pressure levels in the system are outside a certain range, or if a pump is mainly recirculating fluid with low Δp_p.

The total power calculation is given below, where ϕ_P is the power cost and is equal to P_{tot}, which is the total power from the pump/motors and the pressure source. n is the number of pumps, $T_{p,i}$ is the torque and $\omega_{p,i}$ is the angular speed of pump/motor i, and q_{ps} and p_{ps} are the flow rate and pressure at the pressure source, respectively. The pressure source is assumed lossless.

$$\phi_P = P_{tot} = \sum_{i=1}^{n} T_{p,i}\omega_{p,i} + q_{ps}p_{ps} \tag{12}$$

The external force applied to the system is an input that defines the pressure levels. The actuator reference speed is the target value, with the valve signals and pump/motor angular speeds as the control inputs. A cost function applies a penalty to the

fitness value based on the error between the reference and actual speed, as defined below.

Here, ϕ_{v_c} represents the total penalty cost due to actuator speed errors, λ_{v_c} is a weighting factor, m is the number of actuators, and $v_{c,j,ref}$ and $v_{c,j}$ are the reference and actual speeds of actuator j, respectively.

$$\phi_{v_c} = \lambda_{v_c} \sum_{j=1}^{m} (v_{c,j,ref} - v_{c,j})^2 \tag{13}$$

The weighting factor λ_{v_c} is introduced to scale the speed penalty value to match the magnitude of the power calculation and is determined heuristically. The speed error is squared to penalize larger deviations more severely.

Other penalty functions are also implemented but are not fully described here for the sake of simplicity. They primarily perform infeasibility checks or eliminate undesirable operating cases. The final fitness value is defined as the negative sum of the preceding equations and is expressed as Φ:

$$\Phi = -(\phi_P + \phi_{v_c}) \tag{14}$$

5 Case study for a system with one actuator and two pumps

This section presents the simulation results for the system with two pumps and one actuator in fig. 1. There are relatively few solutions to the problem, so it is simple to evaluate the simulations outputs. Simulations with more pumps were also performed, while cases with multiple actuators require different visualization approaches and are not covered in this paper.

This paper uses a population size of 200 with 100 generations — values determined heuristically to balance exploration, convergence, and simulation time. The optimisation process is repeated 12 times for each actuator speed/force pair, as this repetition tends to yield more diverse solutions and reduces the likelihood of converging to poor local minima more effectively than greatly increasing the population size. Multiple solutions are also interesting when designing a control strategy as transitioning solely between the optimal operating modes determined by the GA may incur significant energy costs.

5.1 Best operating mode between the 12 iterations

Figure 6 shows which operating modes lead to the best fitness for each actuator speed and force level over the 12 iterations. "P" and "M" are pump and motor operation, respectively, and "-" means the machine is off. It is important to reinforce that this is not necessarily the global optimum, but rather the best solution found by the GA.

This graph shows the transition points between using one or more e-pumps. The machines are sized to provide, combined, more that the total flow required by the actuator, thus, in theory, a second one would only be required when the actuator speed is more than 0.125 m/s. Instead, the system starts the flow sharing earlier than that, around -0.1 and 0.1 m/s with positive external force. This indicates that the GA encounters

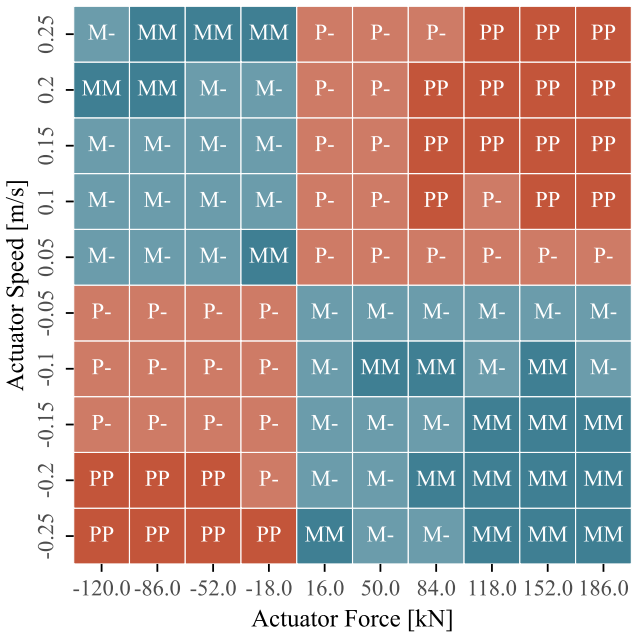


Figure 6: The best operating modes for the hydraulic machines for each actuator speed/force pair.

a region where it can minimize the system losses by using two machines at lower angular speeds.

5.2 Most frequent operating mode after 12 iterations

Defining a single operating option for the actuator's speed/force pair might not be ideal. This approach does not account for switching losses and having only one option could overly constraint the system operation. Figure 7 shows instead which operating mode most frequently happens between the 12 iterations.

The best solution available in 6 may not always be the one that most often happens. In fact, for some points the GA can determine two or more solutions that provide quite similar results just as often and with a small fitness difference, for instance around low positive forces with negative speed.

At higher speeds, both machines are typically required, as the flow demand exceeds the capacity of a single e-pump. A more interesting region is, for example, in the second quadrant, low positive speed and average negative force — table 1. From fig. 6, the best solution would be to use a single machine as a motor. However, the GA often converged to a pump-motor (PM) strategy in fig. 7. In a dynamic control context, this solution could be useful if both e-pumps were running before transitioning to this operating point, thereby avoiding the need to immediately shut down one of the machines.

5.3 Identification of regenerative cases for the best solution

Another interesting result is to analyse the pressure value in the chambers. When regeneration occurs, both actuator' chambers are connected to the same pump port, rising the pressure values. In normal operation, the low-pressure chamber

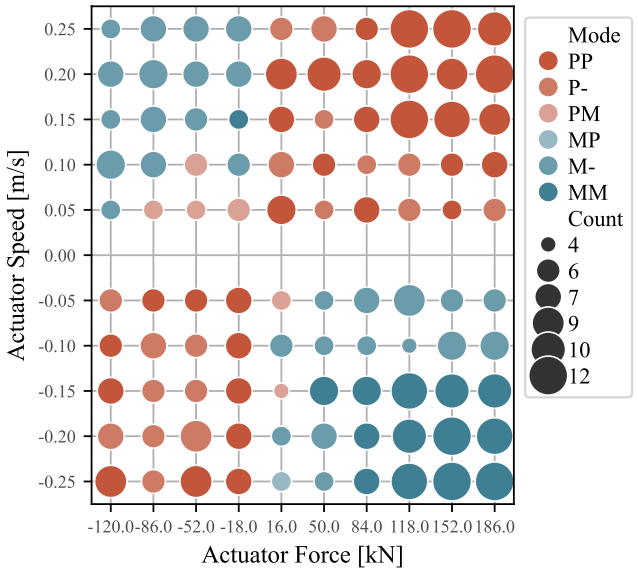


Figure 7: Most frequent operating mode for each pair speed/force.

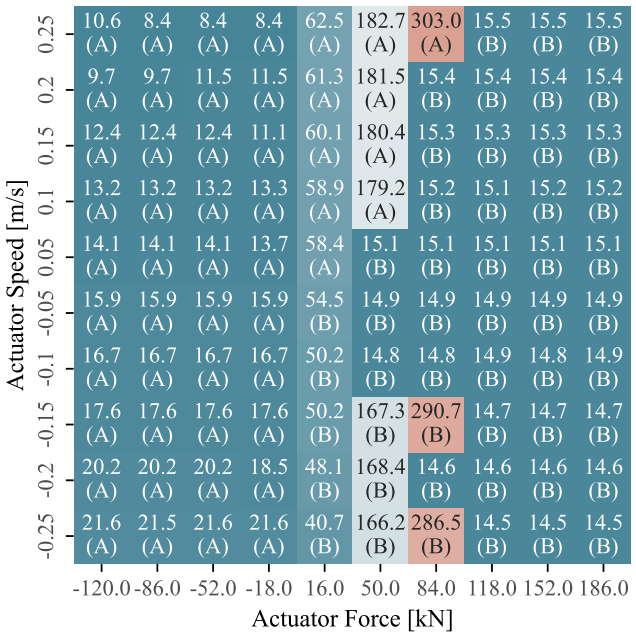


Figure 8: Low-pressure chamber values for the actuator. The higher values indicate conditions when the system is regenerating by interconnecting the chambers.

connects to PS_{aux}, thus the value varies between 10 to 20 bar, depending on flow direction and pressure drop in valves. Figure 8 shows the low pressure value for each operating point and indicates which chamber is the low-pressure side.

In the regenerative regions the average actuator pressure raises, but the e-pumps can operate at lower speeds, reducing the power cost. This seems to be the ideal case at low positive forces, and when the force increases, it looses relevance for lower speeds but stays as a reasonable option at higher absolute velocities.

The previous graphs are some examples of the information that can be extracted from the simulation results. It is also possible to identify how the flow is distributed between the e-pumps or approximate the system losses and average system efficiency at steady-state, for example.

6 Discussion and future work

Section 5 presented some results from the GA approach for a MPS with a single actuator and two pumps. The same study can be performed for a system with more e-pumps, and different constraints can be set, like limiting operation to two quadrants or focusing on a certain operating region for the actuator.

The main goal of this method was to reduce the complexity in designing a controller for a dynamic analysis. From the results, the scope of options can be greatly reduced and it provides better insights about which rules to implement. Thus, the proposal here is to use the quasi-static model to estimate the best operating modes for the system with different load conditions, and use the different options found by the GA as a guideline for the development of a dynamic control.

For example, an intuitive control strategy could start by using a single pump and adding the second machine when the required flow is too high. It is not simple to determine when to actually add this second machine, but the outcome of the GA provides a suggestion to determine the transition region between operating modes based on the actuator speed and force.

One of the challenges of using genetic algorithms arises from their probabilistic nature. Since a relatively large population is required and convergence can be slow depending on the problem, the system must be solved multiple times. A full execution of the algorithm may take days or weeks, depending on the number of components and the number of operating points set for the actuator, even with simplified models. Therefore, improvements or modifications may be necessary to apply this approach to extended architectures.

The GA could be adapted to evaluate a dynamic model and account for transition aspects during a drive cycle, but this would greatly increase the scope of the problem and constrain the results to the chosen case. Finding the ideal operating modes between two speed/force points does not guarantee optimality for the whole cycle, so the algorithm would have to account for all possible paths and compare the total energy consumption of the complete drive cycle. The probabilistic nature of the problem would most likely make it computationally too expensive.

Another common topic concerns global optimality. Dynamic programming (DP) is a well-established method often applied to analyse specific drive cycles. However, as described before, this system contains a vast number of operating points, rendering a pure DP approach computationally infeasible. The results from the GA can be used to constrain the DP algorithm by limiting the number of control options available. It is important to note, however, that this approach does not guarantee global optimality.

Nevertheless, neither strategy is feasible for implementation in an online control loop directly. The DP requires knowledge of the whole drive cycle, while the GA needs long computational time. Their main goals is to serve as a reference approach for the development and/or comparison with other control strategies which aim to mimic the results.

6.1 Future work

The results achieved here are consistent with those in [8], suggesting that the model is applicable for more complex problems. This approach is expected to be more relevant for a system with multiple actuators, as this should bring more options for system operation that are challenging to identify heuristically. Therefore, a key objective is to expand this analysis to include multiple actuators and pumps.

One constraint is that increasing the system complexity directly affects the time it takes to solve the set of equations. While memory limitation is often not a problem for GA, compared to other methods such as backward calculation or DP, computational time is a concern. Although including more actuators is feasible, adapting the algorithm to operate with parallelization could accelerate the process.

Finally, a dynamic analysis based on the results of this study can assess their practical applicability. A key question is the impact of valve switching between operating modes and the system's ability to follow the optimum operating point during different drive cycles. Additionally, GA results can support the development of real-time control strategies to support these studies.

7 Conclusions

This paper presented a filtering strategy to solve the combinatorial problem of the Multi-pump system. The large number of valves and the continuously variable speed pump/motors mean that the system can be controlled in different manners and it is not immediately clear what the viable options are. A modelling approach combined with optimisation is proposed to determine the ideal operating conditions for the system and to reduce the scope of options.

An algorithm was developed to generate symbolic quasi-static equations of the MPS architecture with varying number of pump/motors and actuators. By defining the values of certain parameters, a numerical solver is used to calculate the other variables, allowing for backward or forward simulation. A Genetic algorithm is then employed to generate possible solutions for the system and evaluate what control signals lead to minimized power consumption for different combinations of actuator speed and force.

A case study is presented for a MPS with two pumps and a single actuator as a proof of concept. The results bring insights regarding when the system should ideally transition between different operating modes for the hydraulic machines, and how to distribute the flow between them and the actuator, including regenerative and recuperative cases. This information can serve as a starting point for the development of a dynamic control to such a system.

Finally, the main interesting in the study of the MPS is to combine the operation with multiple actuators. The method proposed can be applied to extended systems, but care should be taken regarding their complexity and the time required to solve it. Nevertheless, the GA approach seems capable of handling the combinatorial problem and provides insightful information about the system behaviour that can be used for implementing dynamic control strategies.

Nomenclature

Designation	Denotation	Unit
$\mathcal{E}$	Effort variable	-
$\mathcal{F}$	Flow variable	-
q	Flow	$m/^3/s$
D_p	Volumetric displacement	$m/^3/rad$
ω	Angular speed	rad/s
T	Torque	Nm
Δp	Pressure difference	Pa
λ	Scaling factor	-
A	Area	$m/^2$
v	Linear speed	m/s^2
p	Pressure	Pa
C_v	Linear flow coefficient	$m^3/s/\sqrt{Pa}$
S	On/off signal	-
ϕ	Cost function	-
P	Power	W
λ_{v_c}	Weight factor	-
Φ	Total fitness value	-
Subscripts		
p	Pump	
A, B, 1, 2	Port or side A, B, 1 or 2	
loss	Loss value	
nom	Nominal value	
c	Cylinder (actuator)	
v	Valve	
N	Node	
P	Power	
tot	Total	
ps	Pressure source	

Acknowledgements

This research was funded by the Strategic Vehicle Research and Innovation (FFI – Fordonsstrategisk forskning och innovation) program within the Swedish Energy Agency (Energimyndigheten) under grant number P2023-00594.

References

[1] Lauren Lynch and Bradley T. Zigler. Estimating Energy Consumption of Mobile Fluid Power in the United States. Technical report, November 2017. Report Number: NREL/TP–5400-70240, 1408087.

[2] Timir Patel, Leonardo Dos Santos Franquilino, Andrea Vacca, and Charlie Young. Comparison Study of

Fully Individualized System Architectures for Electrified Mini-Excavators: Displacement Control (DC) vs. Electro-Hydraulic Actuation (EHA). In *14th International Fluid Power Conference*, pages 22–36, Dresden, Germany, 2024. River Publishers.

[3] David Fassbender, Christine Brach, and Tatiana Minav. Energy-Efficiency Comparison of Different Implement Powertrain Concepts to Each Other and Between Different Heavy-Duty Mobile Machines. In *International Journal of Fluid Power*, pages 127–144, July 2024.

[4] Lasse Schmidt, Mikkel van Binsbergen-Galán, Reiner Knöll, Moritz Riedmann, Bruno Schneider, and Edwin Heemskerk. Energy Efficient Excavator Implement by Electro-Hydraulic/Mechanical Drive Network. *International Journal of Fluid Power*, pages 413–438, December 2024.

[5] Mikkel van Binsbergen-Galán, Barbara Videbæk, Kenneth Vorbøl Hansen, and Lasse Schmidt. Experimental Investigation of Hydraulic Power Sharing Potential in a Dual Cylinder Electro-Hydraulic Variable-Speed Drive Network. American Society of Mechanical Engineers Digital Collection, October 2024.

[6] Artur Tozzi de Cantuaria Gama, Kim Heybroek, and Liselott Ericson. A Novel Multi-pump System for Hydraulic Actuation in Electric Mobile Machinery. pages 1–7. Linköping University Electronic Press, 2023.

[7] Grégory Tardy, Eric Bideaux, Christophe Gostomski, and Armando Fonseca. Optimization Based Energy Efficient Power Transmission Design Methodology Applied to a Compact Excavator. In *GFPS PhD Symposium 2024*, Hudiksvall, Sweden, June 2024.

[8] Artur Tozzi C. Gama, Kim Heybroek, and Liselott Ericson. An Analysis of a Multi-Pump System for Actuator Operation in Electric Mobile Machinery. In *Proceedings of the ASME/BATH 2023 Symposium on Fluid Power and Motion Control*. American Society of Mechanical Engineers Digital Collection, November 2023.

[9] Damiano Padovani, Søren Ketelsen, Daniel Hagen, and Lasse Schmidt. A Self-Contained Electro-Hydraulic Cylinder with Passive Load-Holding Capability. *Energies*, 12(2):292, January 2019.

[10] The MathWorks Inc. Curve Fitting toolbox version: 24.1 (r2024a), 2024.

[11] Aaron Meurer, Christopher P. Smith, Mateusz Paprocki, Ondřej Čertík, Sergey B. Kirpichev, Matthew Rocklin, AMiT Kumar, Sergiu Ivanov, Jason K. Moore, Sartaj Singh, Thilina Rathnayake, Sean Vig, Brian E. Granger, Richard P. Muller, Francesco Bonazzi, Harsh Gupta, Shivam Vats, Fredrik Johansson, Fabian Pedregosa, Matthew J. Curry, Andy R. Terrel, Štěpán Roučka, Ashutosh Saboo, Isuru Fernando, Sumith Kulal, Robert Cimrman, and Anthony Scopatz. SymPy: symbolic computing in Python. *PeerJ Computer Science*, 3:e103, January 2017.

[12] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. Van Der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul Van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3):261–272, March 2020.

[13] Frederick S. Hillier and Gerald J. Lieberman. Metaheuristics. In *Introduction to operations research*, pages 635–644. McGraw-Hill, Boston, Mass., 9. ed., internat. ed edition, 2010.

[14] Ahmed Fawzy Gad. PyGAD: an intuitive genetic algorithm Python library. *Multimedia Tools and Applications*, 83(20):58029–58042, June 2024.

Sliding contacts of 3D-printed direction control valves - tribological performance of suitable material pairs

Ana Trajkovski*, Jan Bartolj, Žan Spačal and Franc Majdič

Laboratory for Fluid Power and Controls, Faculty of Mechanical Engineering, University of Ljubljana, Aškrečeva cesta 6, 1000 Ljubljana, Slovenia
ana.trajkovski@fs.uni-lj.si, jan.bartolj@fs.uni-lj.si, zs1743@student.uni-lj.si, franc.majdic@fs.uni-lj.si

Abstract

Water hydraulics traditionally faces limitations, particularly concerning corrosion, lubrication challenges, and wear, as water lacks the natural lubricity of oil-based hydraulic fluids. However, with the precision and design flexibility provided by metal 3D printing, key components such as valves, or their parts, can be produced with complex internal geometries that both reduce component mass and improves flow. This enhancement has a direct impact on the efficiency and responsiveness of the hydraulic system, which is essential for proportional control applications.

In our study we have analysed materials suitable for spools in sliding contacts of Maraging steel 3D printed direction control valve housing. Tribological tests focused on identifying material pairs that exhibit low friction and wear in a water-based environment, which is crucial for extending the service life and reliability of water-hydraulic systems. For comparison reason, material pairs were also tested in hydraulic oil ISO VG46. Tested materials were nitride stainless steel, anodised aluminium and bronze. The durability tests on three valves with selected material pairs were also performed.

Based on tribological testing conducted in water, the lowest coefficient of friction was observed when a bronze pin slide over a 3D-printed steel plate. The minimal specific wear volume was achieved with a nitride steel pin. Additionally, anodisation was found to significantly enhance the wear resistance of aluminium. Durability tests of the valves indicated that the most suitable material combination involved a bronze spool. This study emphasizes the potential of water hydraulics in specific applications where sustainability and environmental considerations are critical.

Keywords: hydraulic valve, friction, wear, water, tribological tests, durability tests

1 Introduction

Hydraulic fluid power systems are vital in various engineering fields, including transportation, heavy industry and robotics. However, hydraulic oil leaks can cause significant environmental damage. Hose defects and leaking couplings are the most common causes of such accidents, accounting for 50 % of cases. On construction sites, 74 % of fluid leaks involve hydraulic oil, with 30 % of these spills exceeding 20 litres [1]. Contaminated soils or waters must be cleaned at considerable cost to restore them to their natural condition. Additionally, limited amount of mineral oils in the nature, underscores the need for sustainable alternatives and efficient resource management. Increased environmental awareness has led to an ecological transition, which aims to use both novel materials and environmentally acceptable lubricants to replace the traditional oil-based hydraulics.

1.1 Lubricants

Current EU and national emission regulations, based on United Nations memoranda, have already imposed restrictions on the use of certain key lubricants. Green lubricants are becoming mandatory in maritime and hydropower units, water-based hydraulics can be found in nuclear engineering, coal mines, steel foundries, desalination plants, firefighting, food/beverage production and plastics moulding. Green lubricants are also increasingly desirable in farming, forestry and mobile machinery, food processing and pharmaceutics [1], with high potential and concern on contamination or fire due to oil leakages. Growing concern about environmental issues has led to renewed interest in water hydraulics.

Water is non-toxic, environmentally friendly and readily available. Water-hydraulic systems are usually made of very high-quality materials and fabricated with precision machining, which for consequence has between 30 % and 400 % higher costs of water-based compared to the oil-based system [2]. Using water as a pressurized medium demands high corrosion resistance and presents potential issues such as cavitation, requiring the use of materials like stainless steel, bronze, or ceramics. Furthermore, the low viscosity of water necessitates the use of narrow clearances to prevent excessive

leakage, components must be fabricated with high precision finishes. Water also does not provide adequate lubrication for commonly used stainless steel components. Consequently, the number of manufacturers and components suitable for water hydraulics remains limited [3–7]. The higher costs of water-hydraulic systems are a trade-off for their environmental, safety, and operational advantages. While upfront costs are steep, the benefits can outweigh these costs in specific applications where environmental sustainability and safety are paramount. Additional improvements can be achieved by improving valve construction and manifold design, potentially by using the advantages of 3D printing method.

1.2 Materials

Additive manufacturing (AM) has recently enabled the production of hydraulic valves and manifolds with intricate geometries that were often impossible using conventional manufacturing techniques such as CNC machining or casting. AM, a layer-by-layer assembly process based on 3D computer-aided design (CAD) data, offers significant design flexibility and allows for the realization of complex internal structures. The inherent advantages of AM include the elimination of design constraints associated with traditional methods which can contribute to optimized fluid flows and pressure drops. With this approach we can produce fully functional metal parts in a single process, without the need for extensive material removal. This results in products that are up to 70 % lighter than metal products made with traditional manufacturing methods. This process opens up new possibilities for design flexibility, production efficiency and waste reduction [8]. With AM capabilities, hydraulic systems can be designed with smoother, more efficient flow paths, resulting in improved performance, higher operational precision, increased pressure resistance, and enhanced reliability. In our recent work, we have shown that designing the component such as hydraulic proportional spool valve is complicated work but has many positive effects. We successfully reduced the mass of the final part by 63 % and increased the flow characteristic compared to conventionally manufactured counterpart [9].

Companies such as Aidro are already incorporating AM in production of hydraulic manifolds and valves production for diverse applications. John Deere has incorporated AM in manufacturing stainless steel valves for tractor fuel systems, achieving enhanced efficiency, smaller size, reduced material usage, and approximately 50 % cost savings compared to traditional manufacturing methods [10–14]. Also, USA based hydraulic components producer MOOG, has shown the 3D printed version of All-in-one hydraulic system with reservoir, pump, control components and actuator, showing another strong point of AM, which is consolidation of components into single unit. Beyond improvements in mass and flow characteristics, AM enables the integration of advanced functionalities such as embedded sensors or special designs such as a number of connected small on/off valves for digital flow control unit.

Incorporating new materials and manufacturing procedures consequently require the knowledge of mechanical and tribological properties of these materials. For hydraulic components, such as control valves, the tribological properties of incorporated materials pairs in combination with used lubricant are of special interest. By combining the advantages of the new surfaces, materials, and lubrication technologies for friction reduction and wear protection in various equipment worldwide, energy losses caused by friction and wear could potentially be reduced by 18 % in the short term (8 years), or as much as 40 % in the long term [15]. Although 3D metal printing seams very promising, it also brings challenges that require special knowledge to ensure the desired results of 3D printed metal. Achieving high surface quality, dimensional accuracy and appropriate mechanical properties are areas where much research has still to be done. In 3D metal printing, the porosity of the printed material itself is very important. It is known that the faster the printing, the greater the porosity of the material, which usually means poorer mechanical properties. However, high porosity when oil is present in contact can even be desired, since the oil retained in the porous holes can contribute to better tribological properties and lower friction [16].

In order to reduce the component mass, aluminium and as well as high-performance polymers are often used [17,18]. Due to aluminium low specific density, low weight, high thermal and electrical conductivity, and availability, aluminium and its alloys are the most widely used metals for reducing the total mass of the system and the energy consumed for their manufacture and operation [19]. However, due to their low hardness and higher possibility of successive abrasive wear, the use of these materials is often limited to applications with low friction and lower loads [20,21]. To increase the surface hardness of aluminium parts, various surface protection processes have been developed, such as conventional and hard anodising [21], micro-ring oxidation (plasma electrolytic oxidation) [20] or heat hardening with oxygen [22]. Conventional anodising typically produces aluminium oxides up to 25 μm thick layer, while hard anodising produces thicknesses between 25 and 150 μm thick layer [21] similar to HVOF [22], with surface hardness increased by a factor of 5 to 8 [19]. The anodising process and parameters can significantly affect the corrosion resistance of aluminium alloys [23]. Despite the widespread use of aluminium and its alloys in various industrial applications, tribological properties are most often investigated under dry conditions [19,21] or in hydraulic oil [24]. Another material known for excellent corrosion resistance and good sliding properties (even in aqueous environment) is bronze [25]. This alloy consisting mainly of cooper also has good strength and hardness even at elevated temperatures.

In this research, we have tested a 3D printed directional control valve body in combination with tree different materials for control spools: aluminium, bronze and for comparison reason standard stainless steel. These spools, and sliding contacts are essential for the operation of hydraulic systems. Basic tribological tests allowed us to study different material pairs and measure the coefficient of friction and wear, thereby determining the most suitable materials pairs for use in water hydraulics. Tribological tests were done in water and for comparison reason in standard hydraulic oil. In addition, we developed a measuring set-up for testing of hydraulic valves, for durability and reliability test under real operating

conditions. Our research opens up new possibilities for the improvement of water hydraulic systems.

2 Methods

2.1 Tribological Tests

2.1.1 Materials – samples and lubricants

For the tribological tests (fig. 1.a) we prepared 3D printed discs (25 x 25 x 6 mm, fig. 1.c), made of maraging steel (MS1/1.2709) which represented valve body. The discs samples were polished in several steps using a RotoForce-3 automatic sample polishing and preparation device, to a final roughness of 0.1 µm. The pins samples (Φ 10 x 20 mm, with 12,5 mm radius, fig. 1.b), were prepared of three different materials: bronze (CuAl10Fe3Mn2/2.0936, fig. 1.b-1), a commercially available, standard hardened stainless steel (AISI 440-C/1.4125, fig. 1.b-2) and aluminium alloy (EN AW 6082/3.2315, fig. 1.b-3). The pin samples were also polished with the same procedure as MS1 discs, and after that aluminium samples were anodised by standard procedure (EN AW 6082 an, fig. 1.b-4). Pin samples represented valve spools. Materials microhardness was measured with standard Vickers hardness tester and an indentation force of 0.5 N. The main mechanical properties of used materials are represented in tab.1. Before each test, the samples, clamps and pins were cleaned in ethanol and dried in air.

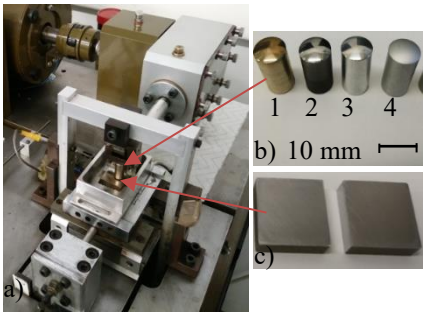


Figure 1. a) Tribological tests set-up; b) pins: 1- Bronze; 2- AISI 440-C; 3-Al 6082; 4-Al 6082 an; c) disc - 3D printed MS1.

First lubricant was commonly used ISO VG46 hydraulic oil. The second lubricant was demineralized water (W).

Table1. Material properties of prepared samples

Material	Density [kg/m^3]	Young modulus [GPa]	Ultimate tensile strength [MPa]	Hardness [HV]
pin				
Bronze	7,6	120	560-690	~261
AISI440-C	7,6	200	760-1960	~695
Al 6082	2,7	70	260	~132
Al 6082 an				~200
disc				
MS1	8-8,1	180	2080	~460

2.1.2 Experimental procedure

Tribological tests were performed on a Camern-Plint TE 77 high-frequency tribometer (fig.1a). During the test, the pin slides in a reciprocal mode on the 3D printed disk-like plate. The average sliding velocity was set to 0.05 m/s (frequency 10 Hz), with stroke length 2.4 mm, which represents the valve spool stroke in a CETOP 3 size 4/3-way slide valve. The tested 3D printed valves used in this research fall into this size class. The normal load was set to 65 N. Before each test, the tested samples were completely immersed in the selected lubricant. During the tests, the current value of the friction coefficient was automatically recorded. The tests were performed for 60 minutes based on our preliminary tests, which showed a stable value of the coefficient of friction and measurable wear. Each test was repeated three times in water and hydraulic oil, and the average values of the steady-state friction coefficient are used for the comparison analysis.

2.1.3 Wear evaluation

The wear volume of the pins was calculated from the dimensions of the wear tracks, which were measured using a 3D digital microscope with an additional nano-point confocal profilometer Hirox HRX-01 & NPS. We determined a 3D profile for each calotte and the dimensions of the wear calotte with Mountain Map software [17]. The digital microscope also made it possible to determine the shape and dimensions of the wear tracs and surface films on the steel discs.

2.2 Hydraulic durability tests

For the control spools, we used the same materials as for the pins during the durability test: aluminium alloy Al 6082, bronze and stainless still AISI 440-C (fig.2). Due to the problems with tolerances after the anodisation, we have only tested valves with standard aluminium alloy spools (Al 6082).

Figure 2. Prepared control valve spools: a) aluminium alloy Al 6082, b) bronze and c) hardened stainless steel AISI 440-C respectively.

The control valve body was produced with the EOS M290 metal 3D printer with the SLM (selective laser melting) method (fig.3). The Maraging tool steel X3NiCoMoTi (MS1) was used, with the properties of stainless steel due to its high nickel content. The printing of the valve with a final volume of 49 cm^3 took about 5 hours. The valve body was modelled and topologically optimised in few successive steps, the modelling process is explained in detail, in our previous studies [9].

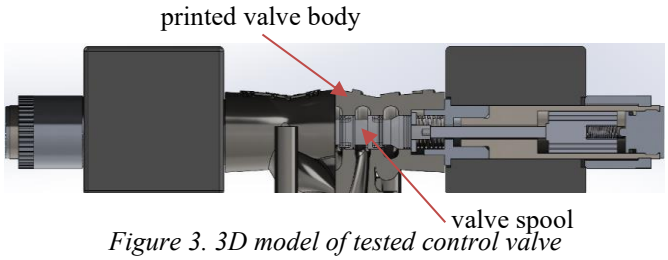


Figure 3. 3D model of tested control valve

2.2.1 Hydraulic test-bench

Figure 4 represents simplified hydraulic scheme of the test rig used for the durability test. An electric motor drives a high-pressure water pump in the main water tank. From the high-pressure pump water flows through a check valve, a flow meter and an adjustable flow valve with pressure compensation. After the pressure filter, the water continues its path into the hydraulic manifold (pos. 8), to which three 3D printed proportional hydraulic valves were attached (pos. 9.1, 9.2 and 9.3). From the P-line, water flows forward the A- or B-line of the individual valve (depending on the valve switching position). The water then continues its path to the orifices (pos. 10.1-10.6), which cause a load in the hydraulic system, simulating real working conditions. From here, the water travels forward into the common return T-line, and through a return filter back into the water tank.

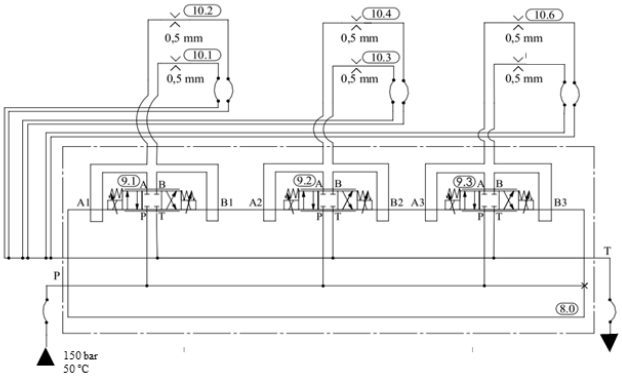


Figure 4. Simplified scheme of hydraulic test-bench

2.2.2 Experimental procedure

All connections to the hydraulic manifold and connections to the hydraulic pipes as well as rigid pipes were made of stainless steel.

The durability tests were performed for about 4 million switching cycles, with the pressure relief valve set at 150 bar. The water thermostat in the tank was set to 50 °C. The test bench operated in such a way that all three valves switch simultaneously.

For the measurement of internal leakage of the valves we installed new pipes on all A and B outputs from the hydraulic manifold, through which water flowed into the measuring cylinders. Test was performed for 90 seconds, while the amount of water that flowed into a specific measuring cylinder was measured. The internal leakage measurements were repeated 3 times for each valve. These tests were performed before the durability tests, and once a week for three following weeks, afterwards. The pressure of the relief valve was set to 50 bar during the internal leakage measurement. To be able to measure for a longer period of time. The water temperature in tank was set at 50 °C, with the cooling system operated when needed.

3 Results and discussion

3.1 Tribological testes

3.1.1 Coefficient of friction evolution

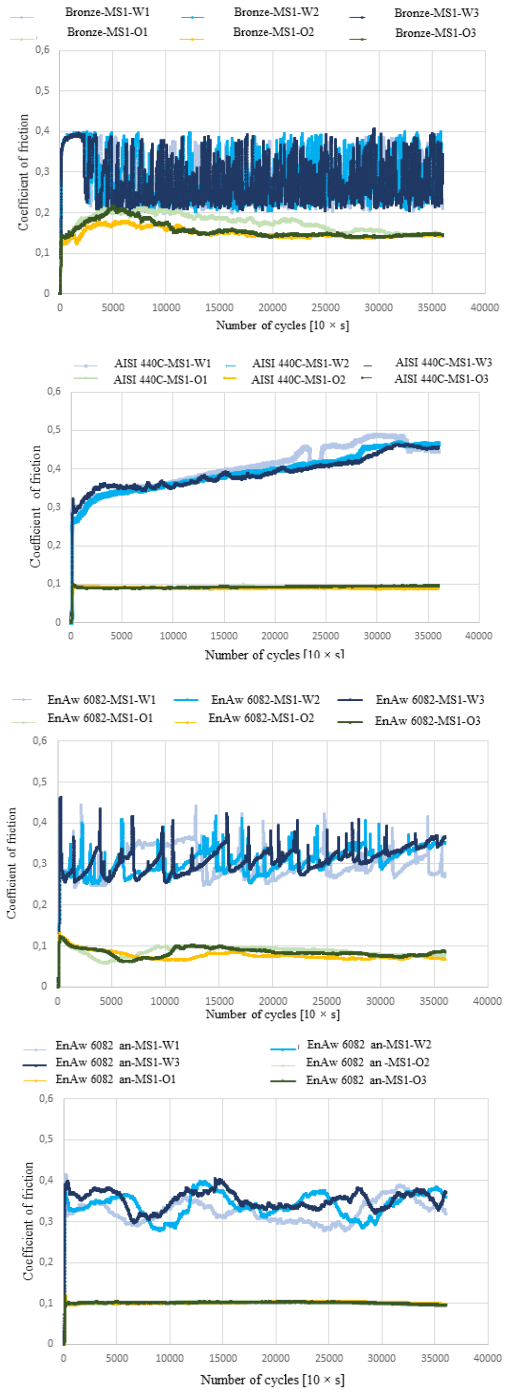


Figure 5. Coefficient of friction evolution during the tribological test of: a) bronze; b) AISI 440-C; a) EN AW Al 6068 d) EN AW Al 6068 an - Pin against MS1 in water and hydraulic oil

Figure 3 shows the evolution of measured coefficient of friction during individual tests of bronze, AISI 440-C, aluminium alloy Al 6082 and anodised aluminium alloy pin samples, respectively. When lubricated with ISO VG46 hydraulic oil, coefficient of friction reaches a stable value almost immediately for AISI 440-C and anodised aluminium, and relatively quickly, with in the first 5000 – 10000 cycles (500 - 100 s) for bronze and basic aluminium alloy respectively.

However, the coefficient of friction measured in the water doesn't show typical running in, it rather has unstable value around the relatively higher coefficient of friction compared to oil lubricated condition for all tested materials. Based on the results of measurements, the fluctuation observed in case of water lubricated conditions indicate, that water incorporates some instabilities into the sliding contact. Although, fluctuations do not disappear, there is a trend of normalization after about 200 s of running in time in case of aluminium alloy in water, also these instabilities are not as significant for anodised aluminium. In case of bronze the fluctuations are constant and persistent. In case of standard AISI 440-C these instabilities are not evident, although coefficient of friction is linearly increasing during the observed period of time (almost 50 %).

3.1.2 Coefficient of friction

Figure 6 represents average values of measured steady coefficient of friction for all tested samples in both water and hydraulic oil. Lower values of coefficient of friction were observed in hydraulic oil, for all tested samples compared to water lubricated conditions. However, the coefficient of friction increased in water, from about 1.75 times for bronze up to 4.5 times for AISI 440-C and aluminium alloys.

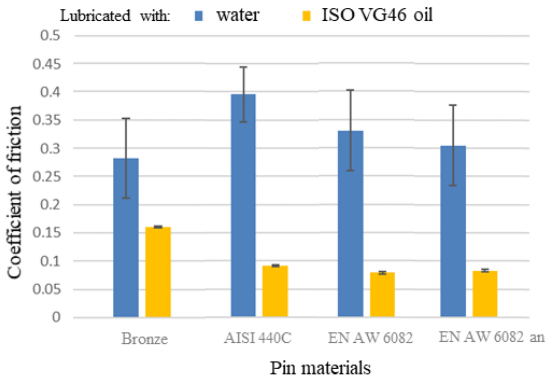


Figure 6. Coefficient of friction (average value ± sd) for all tested Pin-s against MS1 in water and hydraulic oil

In oil, EN AW 6082 showed lowest measured coefficient of fiction with comparable values to measured values for anodised aluminium and AISI 440-C, all around ~ 0.09. Under same lubricated conditions, bronze showed about 50 % higher values of coefficient of friction (~ 0.16).

When water was used as lubricant, the lowest value of coefficient of friction was measured for bronze, around ~ 0.28, followed by anodised aluminium (~ 0.33), standard aluminium EN AW 6082 (~ 0.33) and considerably higher values were measured for AISI 440-C (~ 0.39). In water both bronze and aluminium alloy give good response compared to standard stainless steel, and from that point of view could be potentially good materials for water hydraulics. When we account for fluctuations in measured values seen in for bronze, aluminium alloy is expected to enable smoother working conditions even in water. Additionally, process of anodising aluminium alloy, enabled considerable decrease of coefficient of friction.

3.1.2 Specific wear

Based on repeated measurements in water and hydraulic oil, fig. 7 represents average values of calculated specific wear from analyzed pictures of wear tracs, of all tested pin materials.

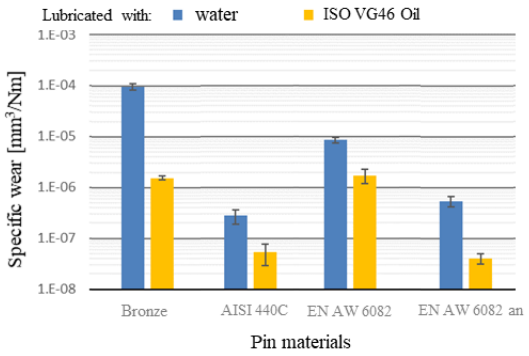


Figure 7. Specific wear (average value ± sd), for all tested Pin-s against MS1 in water and hydraulic oil

In hydraulic oil, we can see that both AISI 440-C and anodised aluminium showed lower specific wear (~ 5×10^{-8} mm^3/Nm), compared to EN AW 6082 and bronze with comparable and almost two orders higher specific wear (~ 1.5×10^{-6} mm^3/Nm).

As expected, specific wear was higher in water and again lowest measured was for AISI 440-C (~ 2.5×10^{-7} mm^3/Nm), fairly followed by anodised aluminium (~ 5×10^{-7} mm^3/Nm). The highest specific wear was measured for bronze (~ 1×10^{-4} mm^3/Nm). However, the increase of specific wear in water, for EN AW 6082 was not as significant, and stayed at the same order as calculated in oil (~ 8×10^{-6} mm^3/Nm).

Based on the results, we can conclude that the best results were accomplished by AISI 440-C and anodised aluminium. However, since water had the lowest impact on increase of specific wear of EN AW 6082 it can be considered as an interesting alternative, which can potentially be used for very different and demanding working conditions.

3.1.3 Surface analyses

After the tests, wear tracks of both pin and disc in sliding contact, were analyzed with digital microscope (fig.8). From the pictures it is clear, that the most significant wear and material removal area is evident in bronze, when lubricated with water (W). Also, for the bronze pin the stretches in the

sliding direction are the most intensive, thick and prominent, suggesting abrasion as main wear mechanism.

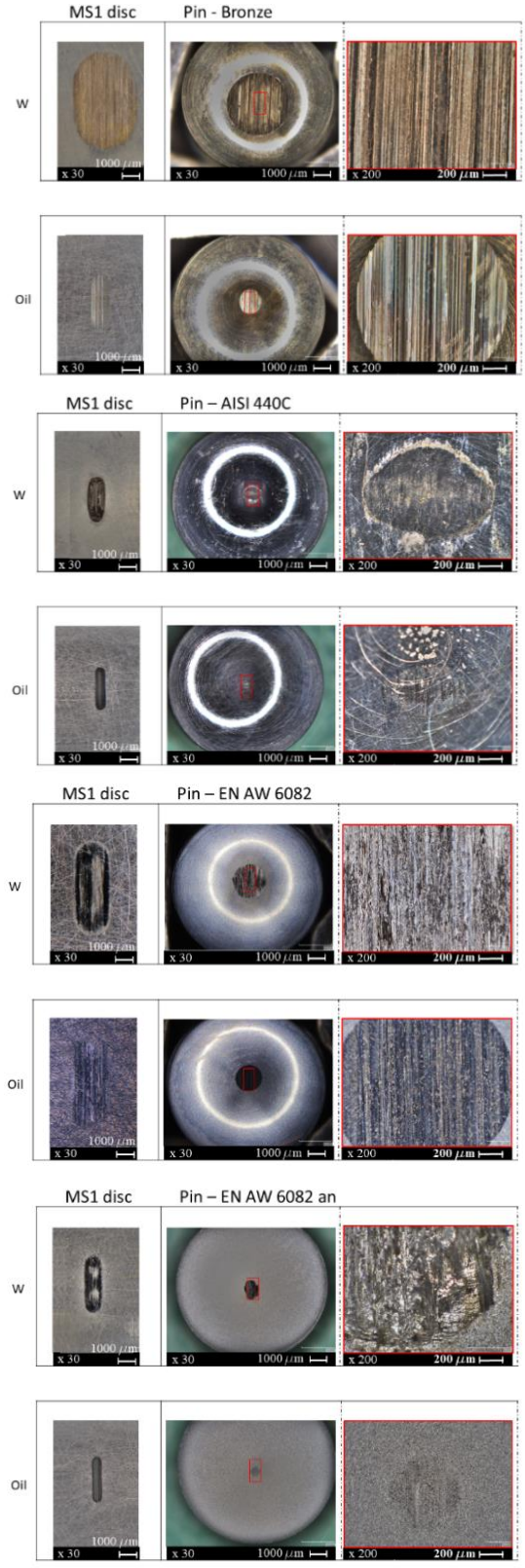


Figure 8. Digital microscope images of wear tracs on MS1 discs and all tested Pin-s after sliding in water(W) and ISO VG46 hydraulic oil (Oil)

For other tested materials, the difference between wear track on MS1 disc in water in oil is not as significant. But there is a trend in more prominent pin wear and material removal in water as compared to oil lubricated condition for AISI 440-C and anodised aluminium (EN AW 6082 an). Stretches in the sliding direction are also present in AISI 440-C and aluminium, but not as intensive. Anodised aluminium results suggest relatively small wear area, with area comparable to AISI 440-C, but different wear mechanism present. There were irregularities present, that probably present peeled anodised protective layer in water, thus suggesting adhesive wear.

3.2 Hydraulic durability tests

3.2.1 Internal leakage

The internal leakage was measured for each valve on lines A and B and averaged afterwards (Q_L). Figure 9 shows the internal leakage of all three valves, measured before the durability test, and after about 1, 2.7 and 4.2 million of cycles. Switching cycle represented the control spool movement from the zero position to the first working position, followed by switching from the zero to the second working position and returning the control spool back to the zero (middle) position.

As can be seen in fig. 9, the internal leakage increased by far the most on the valve with the control spool made of aluminium. If we compare the spools, we can see that in a valve with an aluminium spool, internal leakage increased 32 times more than in a steel one and about 220 times more compared to bronze spool. The leakage increased the least on the control spool made of bronze. After the end of the test, it increased by only 11 ml/min (~ 50 %). On the steel spool, the internal leakage increased by 172 ml/min (~ 8.5 times).

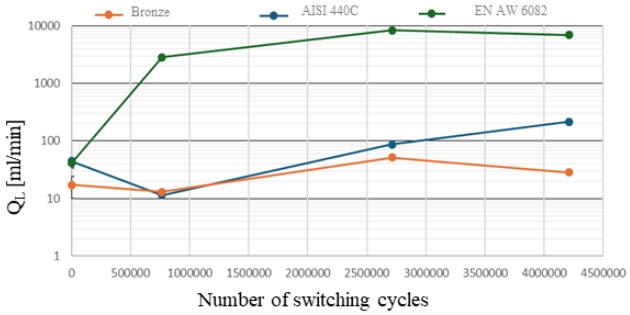


Figure 9. Measured internal leakage for all three tested control spool materials

During the durability test there was a problem probably due to the intense MS1 3D printed housings rusting. The rust caused the control spools to stuck in the valve. As a result, we had to regularly disassemble and clean the valves. Additionally, spools could potentially stuck due to different thermal expansions between the housing material and the control spool materials. This could also reduce the gap required for the spool to slide smoothly in the housing. This raises the question of how many switches some of the control spools actually made,

because some got stuck much before others, and we could not precisely control the process.

3.2.2 Control valve spool surfaces

When the durability tests were stopped, for the leakage measurements, all control valve spools were inspected visually (fig.10).

a) Control valve spools after ~ 800.000 cycles

b) Control valve spools after ~ 2.7 million cycles

c) Control valve spools after ~ 4.2 million cycles

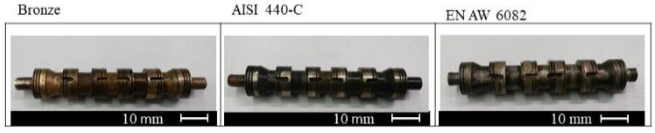


Figure 10. Control valve spools after durability tests in water

After about 800.000 cycles both bronze and AISI 440-C spools were stuck, with to the evident rust on valve housing. There were no evident damages on the spools. The spools and valve bodies were cleaned to remove rust particles and other impurities before measuring internal leakage. After about 2.5 million cycles, there was no severe damage evident on both bronze and AISI 440-C spools edges. But the aluminium spool after a number of cycles started deforming and even the washer on the left side of the spool could not be removed without using considerable force. After about 4.2 million switching cycles, we ended the durability test. The valves were stuck often during the test and needed to be cleaned regularly. In addition, the aluminium control spool was no longer switching because it had become so worn. After the last leakage measurement, all the control spools were analyzed with digital microscope. Special emphasizes was given to control edges to find the areas where the greatest wear (fig.11).

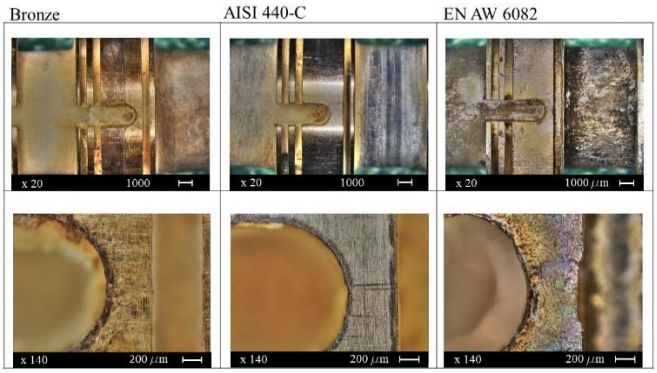


Figure 11: Wear of the control spool grooves on the P side of after durability test, top view

As seen also on the fig.11 the rust is present on all the spools. However, there is also evidence of damage of spool control edges, especially for the aluminium spool. This is why we further analyzed P and T port of the aluminium spool (fig.12)

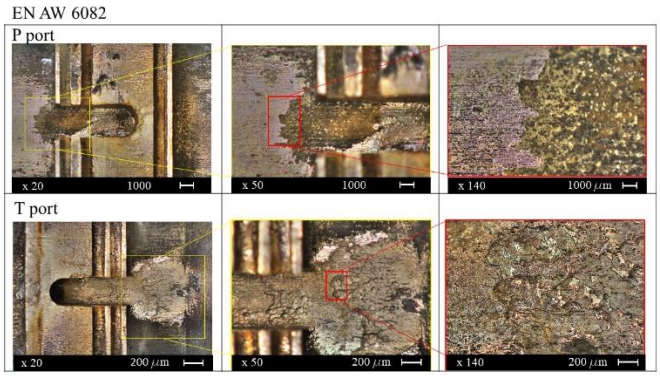


Figure 12. Aluminium control spool edges wear (P and T port)

Further deeper analyses are needed to accurately find the wear mechanism. However, we can see that erosion and cavitation were partially present here (fig,12), because aluminium peeling is visible on the spool surface, which is characteristic of material fatigue. Fatigue probably occurred due to continuous pressure shocks caused by water.

3.2.3 Control valve housing surface

Since dust was evident on MS1 3d printed housing of the valve, at the end of durability tests, the housing was cut in half, and analyzed with digital microspore. There was the most prominent evidence of intense rust on slide surfaces (fig.13a); edge intensive wear and erosion (fig.13b) and material abrasion and intense scratches, probably from rust impurities or worn particles (fig.13c).

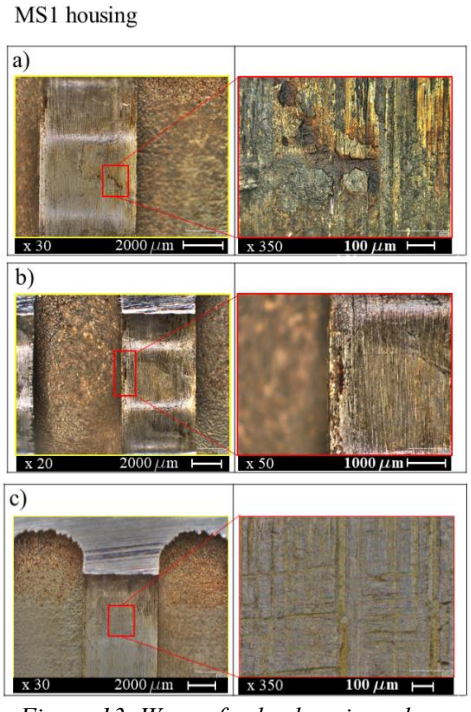


Figure 13. Wear of valve housing edges

4 Conclusions

As part of the research on effective components for water-hydraulics, in this study we have analyzed the potential of using 3D printed maraging steel (MS1) in combination with some materials typically used for traditional oil hydraulics. Based on tribological measurements in both hydraulic oil and water, we can conclude that water presents challenging lubrication conditions for all tested materials, with both the coefficient of friction and specific wear being therefore higher. The materials tested ranged from standard stainless steel, over corrosion resistant bronze down to lightweight and low-cost aluminium.

The highest specific wear was observed in bronze, the intense abrasive wear was observed especially in water and was also related to the unstable coefficient of friction during the experiments. However, these deep grooves in the material, which were probably filled with lubricant, allowed for smoother movement of the control valve spool, and a response closer to the valve with standard stainless-steel spool.

Aluminium was the material with the lowest microhardness compared to the other materials tested. However, the coefficient of friction was stable, and the specific wear was higher, but comparable to that observed in stainless-steel under the low-load conditions tested. The anodising process improved both the coefficient of friction and specific wear, bringing aluminium closer to the behaviour of stainless-steel under the conditions tested. The wear mechanism also included adhesions of material. In the traditionally used hydraulic oil, anodised aluminium also followed the stainless steel quite closely in tribological pair with the 3D printed steel.

Further tests and microscopic analyses would enable as to precisely analyse the effect of anodising on improving the tribological properties of aluminium alloys towards water-hydraulic components. In addition, the manufacturing process of such components should be precisely controlled, as we had problems with testing the anodised control spool with already prepared printed manifolds.

Leakage in spool-type valves is highly sensitive to clearance due to its cubic relationship to small gap, causing significant variation (up to 5-10 times) even within standard tolerances. Although this study did not specifically evaluate the tolerance limits of 3D printed parts, consistent post-processing—sandblasting, micro hardening, CNC machining, and fine grinding—was used to minimize housing variation. Given the tight tolerance requirements, especially for sliding surfaces, final machining plays a crucial role in performance. There are some experimental data on the influence of surface finishing of 3D printed parts, on friction and wear, which present a valuable direction for future studies of particular material and lubricant combination, as well as loading parameters.

Although, there are studies that confirm good corrosion resistance of miraging steel, within our research there were problems with intense rust of valve manifold. The results suggesting that special attention must be paid to the manufacturing and final thermal processing of miraging steel, before it is considered for use in water-lubricated conditions.

Acknowledgment: This research work was funded by the Slovenian Research and Innovation Agency (ARIS) as part of the interdisciplinary project of the University of Ljubljana: Green Urban Communities of the Future (Phase I: Preparatory Stage).

References

[1] Deuster S, Schmitz K, Widomski K, Barnat-Hunek D, Musz-Pomorska A. Bio-Based Hydraulic Fluids and the Influence of Hydraulic Oil Viscosity on the Efficiency of Mobile Machinery. Sustainability 2021, Vol 13, Page 7570 2021;13:7570. https://doi.org/10.3390/SU13147570.

[2] Krawczyk J, Sobczyk A, Styczek J, Walczak P. Tests of New Methods of Manufacturing Elements for Water Hydraulics. Materials Research Proceedings 5, 2018, p. 200–5. https://doi.org/10.21741/9781945291814-35.

[3] Tapio L. Miniature Digital Hydraulic Valve System. PhD. Aalto University, 2018.

[4] Paloniitty M. Novel Water Hydraulic On/Off Valves and Tracking Control Method for Equal Coded Valve System. PhD. Tampere University of Technology, 2018.

[5] Janus 2/2 On/Off Valve - The Water Hydraulics Company n.d. https://www.waterhydraulics.co.uk/janus-2-2-on-off-valve/ (accessed January 28, 2025).

[6] - Amiad Water Systems Eliptix R-30 & R-31 Hydraulic Control Valve n.d. https://au.amiad.com/products/eliptix-r-30-r31-hydraulic-control-valve-irr/ (accessed January 28, 2025).

[7] 2/2 dir. contr. propotional valve n.d. https://tiefenbach-wasserhydraulik.eu/en/products/proportional-valves/2/2-dir-contr-propotional-valve/ (accessed January 28, 2025).

[8] Ngo TD, Kashani A, Imbalzano G, Nguyen KTQ, Hui D. Additive manufacturing (3D printing): A review of materials, methods, applications and challenges. Compos B Eng 2018;143:172–96. https://doi.org/10.1016/j.compositesb.2018.02.012.

[9] Bartolj J, Čelik A, Majdič F. Development of metallic 3d-printed water hydraulic proportional directional control valve. International conference Fluid Power 2021: Conference Proceedings, University of Maribor Press; 2021, p. 41–52. https://doi.org/10.18690/978-961-286-513-9.4.

[10] Klimek A, Kluczyński J, Łuszczek J. Performance Analysis of Additively Manufactured Hydraulic Check Valves with Different Postprocessing.

Materials 2023;16:7302. https://doi.org/10.3390/ma16237302.

[11] Peng T, Wang Y, Zhu Y, Yang Y, Yang Y, Tang R. Life cycle assessment of selective-laser-melting-produced hydraulic valve body with integrated design and manufacturing optimization: A cradle-to-gate study. Addit Manuf 2020;36:101530. https://doi.org/10.1016/j.addma.2020.101530.

[12] Li D, Dai N, Wang H, Zhang F. Mathematical Modeling Study of Pressure Loss in the Flow Channels of Additive Manufacturing Aviation Hydraulic Valves. Energies (Basel) 2023;16:1788. https://doi.org/10.3390/en16041788.

[13] Arena M, Ambrogiani P, Raiola V, Bocchetto F, Tirelli T, Castaldo M. Design and Qualification of an Additively Manufactured Manifold for Aircraft Landing Gears Applications. Aerospace 2023;10:69. https://doi.org/10.3390/aerospace10010069.

[14] Hofmann U, Fankhauser M, Willen S, Inniger D, Klahn C, Löffel K, et al. Design of an additively manufactured hydraulic directional spool valve: an industrial case study. Virtual Phys Prototyp 2023;18. https://doi.org/10.1080/17452759.2022.2129699.

[15] Friedrich K. Polymer composites for tribological applications. Advanced Industrial and Engineering Polymer Research 2018;1:3–39. https://doi.org/10.1016/J.AIEPR.2018.05.001.

[16] Zhu Y, Lin G, Khonsari MM, Zhang J, Yang H. Material characterization and lubricating behaviors of porous stainless steel fabricated by selective laser melting. J Mater Process Technol 2018;262:41–52. https://doi.org/10.1016/j.jmatprotec.2018.06.027.

[17] Trajkovski A, Novak N, Pustavrh J, Kalin M, Majdič F. Performance of Polymer Composites Lubricated with Glycerol and Water as Green Lubricants. Applied Sciences 2023;13:7413. https://doi.org/10.3390/APP13137413.

[18] Trajkovski A, Novak N, Kalin M, Majdič F. On polyoxymethylene composite for sustainable hydraulic valves. 14th International Fluid Power Conference, Dresden: 2024, p. 298–307.

[19] Musza A, Windisch M, Török M, Molnár T, Kovács S, Sátán-Papp S, et al. Tribological Examination of Anodized Al-356 for Automotive Use. Coatings 2023;13:1642. https://doi.org/10.3390/coatings13091642.

[20] Rodriguez L, Vieu A, Balsarin M, Combes P, Alexis J, Esvan J, et al. Physico-chemical characterisation and tribological behaviour of ground micro-arc oxidation coating on aluminium alloy – Comparison with hard anodised oxidation. Wear 2023;516–517:204591. https://doi.org/10.1016/j.wear.2022.204591.

[21] Soffritti C, Fortini A, Nastruzzi A, Sola R, Merlin M, Garagnani GL. Dry Sliding Behavior of an Aluminum Alloy after Innovative Hard Anodizing Treatments. Materials 2021;14:3281. https://doi.org/10.3390/ma14123281.

[22] Bolelli G, Lusvarghi L, Barletta M. HVOF-sprayed WC–CoCr coatings on Al alloy: Effect of the coating thickness on the tribological properties. Wear 2009;267:944–53. https://doi.org/10.1016/j.wear.2008.12.066.

[23] Miramontes JC, Gaona Tiburcio C, García Mata E, Esneider Alcála MÁ, Maldonado-Bandala E, Lara-Banda M, et al. Corrosion Resistance of Aluminum Alloy AA2024 with Hard Anodizing in Sulfuric Acid-Free Solution. Materials 2022;15:6401. https://doi.org/10.3390/ma15186401.

[24] Nakase T, Sakuragi K. Surface Treatment of Aluminium Alloy for Tribological Applications. KYB Technical Review 2017;55:18–24.

[25] Jia JH, Chen JM, Zhou H Di, Wang JB, Zhou H. Friction and wear properties of bronze–graphite composite under water lubrication. Tribol Int 2004;37:423–9. https://doi.org/10.1016/J.TRIBOINT.2003.12.013.

An Approach for Evaluating Structure-Borne Noise Emission of Axial Piston Units Using Blocked Forces

Matthias Vogt[1,2] and Marcus Geimer[1]

[1]Institute of Mobile Machines (Mobima), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
[2]Development of Axial Piston Units, Bosch Rexroth AG, Horb am Neckar, Germany
E-mail: matthias.vogt@partner.kit.edu, marcus.geimer@kit.edu

Abstract

In the context of electrification and the development of quieter internal combustion engines, the noise generated by axial piston units is becoming more prominent. While established methods and standards exist for evaluating airborne and fluid-borne noise of hydraulic pumps, the structure-borne noise emitted by the pump and transferred to mobile machinery has been rarely investigated. A method that has gained popularity in recent applications for characterizing structure-borne noise sources is the in situ blocked force method. Blocked forces are receiver-system-independent interface quantities, allowing for a component-specific assessment of structure-borne noise. This publication introduces the fundamental approaches for determining blocked forces and presents the results of a metrological assessment of blocked forces for an axial piston pump. Besides evaluating the hydraulic unit using blocked forces, these forces can also be applied to any machine structure as they are receiver-system-independent. This enables the prediction of how structure-borne noise emitted by the axial piston pump propagates in a mobile working machine. The findings demonstrate promising outcomes according to common validation methods for blocked forces and provide a simple single-value metric for evaluating the emitted structure-borne noise of an axial piston pump.

Keywords: Structure-Borne Noise, Axial Piston Pumps, Source Characterization, Blocked Forces

1 Introduction

The development of increasingly quieter internal combustion engines, along with the growing adoption of considerably quieter electric motors, has heightened the demand for low-noise components in mobile machinery [1]. This trend presents a significant challenge for component manufacturers, but it also offers substantial potential to reduce overall noise levels of mobile machines. A key aspect in the development of low-noise mobile machinery is the ability to accurately assess and evaluate the contribution of individual noise sources. One of the primary sources of noise in mobile machinery are axial piston units. These units are widely used due to their high power density, the adjustable displacement volume, and the achievable high pressure level [2]. The challenge in assessing the contribution of axial piston pumps to the overall noise lies in the fact that these pumps emit all three types of sound: airborne noise (ABN), fluid-borne noise (FBN), and structure-borne noise (SBN). To enable a comprehensive assessment of the acoustic behavior of a pump, all three types of noise must be evaluated separately.

Figure 1 presents a schematic illustration of the noise transfer paths within a mobile machine, assuming the axial piston pump is the sole noise source and the SBN transmitted through the tubing can be neglected. SBN vibrations that directly affect the driver, such as those transmitted through the seat, are also not considered here. The primary excitation mechanisms within an axial piston pump can be divided in a SBN excitation and a FBN excitation, often referred to as flow ripple or fluid pulsation. These excitation mechanisms lead, on the one hand, to an ABN directly emitted from the pump housing ① ④ (see Figure 1) and, on the other hand, to FBN and SBN transmitted via the pump discharge port ③ and the pump mounting ② to the vehicle. The following section briefly describes the most important excitation mechanisms within an axial piston pump.

The FBN excitation or fluid pulsation in an axial piston pump can be divided into kinematic pulsations, due to the finite number of displacement elements, and pulsations resulting from compressibility of the hydraulic oil. It occurs during the transition of a piston from the suction port to the discharge port, when the pressure in the piston chamber differs from that in the discharge port. [2–4]

Since fluid pulsation can lead to not only noise issues but also other problems within the hydraulic system, various studies

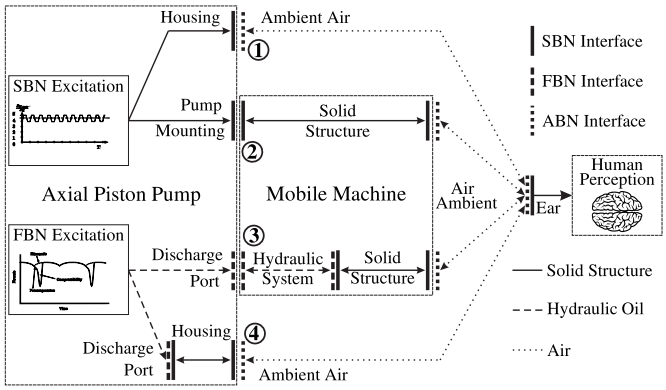


Figure 1: Noise transfer paths of a mobile machine with axial piston unit as the only source

have been conducted to characterize the FBN excitation of hydraulic units. The most common methods are the "High Impedance Pipe" Method [5], the "Reflectionless Termination (RaLA)" Method [3,6], the "Secondary Source" Method [7,8], and the "Two Pressures Two Systems" Method [9, 10]. Some of these methods have also been standardized in ISO 10767, enabling a comparable source characterization of hydraulic pumps with respect to FBN. Concerning the ABN caused by high-pressure pulsations in the discharge port of the pump ④, Haarhaus [11] demonstrated that this plays a minor role in the noise emission of axial piston pumps.

Regarding SBN excitation mechanism, previous studies [2, 11–14] have indicated that the primary excitation is due to the axial force resulting from pressure fluctuations in the cylinder chambers of the pump's power unit. These forces lead to two primary effects: firstly, they induce vibrations in the pump housing, causing the pump to directly radiate ABN ①; secondly, the SBN is transmitted at the pump's mounting points to the mobile working machine, thereby exciting the entire machine structure ②.

For the characterization of ABN radiated from the housing of hydraulic pumps and motors, a test procedure is described in the ISO 4412-1/2 standard. This standard defines necessary properties for the test setup to ensure comparable results that are independent of external boundary conditions.

In contrast to FBN and ABN, there is currently no standard or established method for the source characterization of hydraulic pumps with respect to SBN. Hence, this publication presents an approach that enables the evaluation of SBN based on so-called blocked forces, which provide a receiver-independent description of the source at the interface to the receiving structure. In addition to evaluating the pump, source characterization using blocked forces also enables the prediction of vehicle noise resulting from SBN excitation.

The publication is structured as follows: First, various methods of transfer path analysis are introduced, focusing on component-based transfer path analysis. Next, the application of the in situ blocked force method to an axial piston pump is presented, and a metric for evaluating SBN is derived. Subsequently, the results of the in situ blocked force measurement and the evaluation metric are discussed. Finally, the findings are summarized, and an outlook is provided.

2 Methods

The method used to investigate the transmission paths through which noise is transferred from a source to a receiver is referred to as Transfer Path Analysis (TPA). The description of a source using blocked forces is one method of source characterization within the framework of TPA. Therefore, an introduction to TPA will be provided in this chapter.

According to [15] a typical TPA workflow aiming to identify critical noise sources or transmission paths can be divided into four steps:

1. Operational measurement of the active component
2. Determination of the structural dynamic behavior of the passive receiver
3. Calculation of interface quantities between the source and the receiver
4. Assessment of the contribution of individual sources and paths to the overall noise

Depending on the TPA method used, only some of these steps may be performed, or they may be carried out in different order. Since this publication focuses on evaluating the SBN transmitted from an axial piston pump to a mobile working machine, the emphasis is on the third step, the calculation of interface quantites. An overview of the different TPA methods can be found in [15]. The notation and theoretical framework introduced in the following are also based on this work. In the next sections, the TPA methods most relevant to this publication will be introduced.

Figure 2 illustrates the general transfer path problem applied to hydraulic pump **A** connected to a receiver system **B**, such as the engine flange in a mobile working machine. In addition to the mounting points, the pump is also connected to the receiver via hydraulic hoses or pipes. These interfaces are not considered here. The system consists of three nodes: the excit-

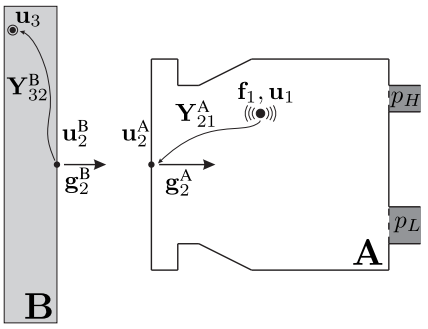


Figure 2: General transfer path problem applied to a hydraulic pump

ation point within the source **1**, the interface **2**, and an arbitrary point on the receiver **3**. The excitation at node **1** is described by the force $\mathbf{f}_1$ and the displacement quantity $\mathbf{u}_1$. As displacement quantities, accelerations, velocities and displacements can be considered, as these can be directly converted into one another. In the following, it is assumed that **u** represents a velocity. The response at the receiver point to the excitation at **1** is denoted by $\mathbf{u}_3$. At the interface, the displacement quantities on the

source and receiver sides are described by $\mathbf{u}_2^{\mathrm{A}}$ and $\mathbf{u}_2^{\mathrm{B}}$, respectively, and the interface forces by $\mathbf{g}_2^{\mathrm{A}}$ and $\mathbf{g}_2^{\mathrm{B}}$. At this point, it is important to note that the superscript indicates the component to which the quantity refers. Accordingly, $(\cdot)^{\mathrm{A}}$ represents a source quantity, while $(\cdot)^{\mathrm{B}}$ denotes a receiver quantity. The superscript $(\cdot)^{\mathrm{AB}}$ indicates that the quantity is associated with the assembled system. The structural dynamic behavior of a node, as well as the transfer behavior between two nodes, is described using admittances $\mathbf{Y}$. Admittance, also known as mobility, is a measure of how a structural dynamic system responds to external forces. It is defined as the ratio of the system's velocity response to the applied force in the frequency domain. Since the systems under investigation are typically multi-input multi-output (MIMO) systems, forces and velocities at the nodes are described by column vectors, and the transfer behavior is represented using admittance matrices.
The described transfer path problem can be represented by the following system of equations, where the subsystem admittance matrices $\mathbf{Y}^{\mathrm{A}}$ and $\mathbf{Y}^{\mathrm{B}}$ are arranged in a block-diagonal format. The admittances matrices with identical indices describe the point admittances at the nodes, while those with different indices represent the transfer admittances between the nodes.

$$\underbrace{\begin{bmatrix}\mathbf{u}_1\\ \mathbf{u}_2^{\mathrm{A}}\\ \mathbf{u}_2^{\mathrm{B}}\\ \mathbf{u}_3\end{bmatrix}}_{\mathbf{u}} = \underbrace{\begin{bmatrix}\mathbf{Y}_{11}^{\mathrm{A}} & \mathbf{Y}_{12}^{\mathrm{A}} & \mathbf{0} & \mathbf{0}\\ \mathbf{Y}_{21}^{\mathrm{A}} & \mathbf{Y}_{22}^{\mathrm{A}} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{Y}_{22}^{\mathrm{B}} & \mathbf{Y}_{23}^{\mathrm{B}}\\ \mathbf{0} & \mathbf{0} & \mathbf{Y}_{32}^{\mathrm{B}} & \mathbf{Y}_{33}^{\mathrm{B}}\end{bmatrix}}_{\mathbf{Y}} \underbrace{\left(\begin{bmatrix}\mathbf{f}_1\\ \mathbf{0}\\ \mathbf{0}\\ \mathbf{0}\end{bmatrix} + \begin{bmatrix}\mathbf{0}\\ \mathbf{g}_2^{\mathrm{A}}\\ \mathbf{g}_2^{\mathrm{B}}\\ \mathbf{0}\end{bmatrix}\right)}_{(\mathbf{f}+\mathbf{g})} \tag{1}$$

For an assembled system, assuming a rigid behavior of the interface, the compatibility condition and the equilibrium condition at the interface must also be satisfied. The compatibility condition states that the velocities on both sides of the interface must be identical, i.e., $\mathbf{u}_2^{\mathrm{B}} = \mathbf{u}_2^{\mathrm{A}}$. The equilibrium condition, which is fundamentally based on Newton's Third Law, asserts that the interface forces on the receiver side must be equal in magnitude and opposite in direction to the forces on the source side, i.e., $\mathbf{g}_2^{\mathrm{B}} = -\mathbf{g}_2^{\mathrm{A}}$. By employing equation 1 in conjunction with the compatibility and equilibrium conditions, it is possible to forecast the response at the receiver $\mathbf{u}_3$ to an excitation $\mathbf{1}$, provided that the excitation and the admittances of the subsystems are known.
Although predicting the response at the receiver using the internal excitation, as just described, is theoretically possible, it presents certain practical issues. In particular, the internal forces $\mathbf{f}_1$ are typically unknown with regard to their position, orientation, and magnitude. Consequently, it is not possible to define the transfer behavior, i.e., the admittances $\mathbf{Y}_{11}^{\mathrm{A}}$ and $\mathbf{Y}_{21}^{\mathrm{A}}$.
A solution to this problem is to find an alternative description of the source activity that: a) is capable of physically representing the internal forces (i.e., inducing the same vibration in the structure), b) has known positions, and c) has magnitudes that can be practically determined [16]. Such a description of the source activity can be achieved using the interface quantities at interface $\mathbf{2}$, which makes these interface quantities also suitable for evaluating the source with respect to SBN.

The TPA approaches presented in the following differ primarily in the interface quantities used to describe the source and how the response at the receiver $\mathbf{u}_3$ can be predicted based on these interface quantities.

2.1 Classical TPA

In the first investigations in the field of TPA, the objective was to describe the source activity using the forces acting at the interface $\mathbf{g}_2^{\mathrm{A}}$ and $\mathbf{g}_2^{\mathrm{B}}$, whereby $\mathbf{g}_2^{\mathrm{A}}$ and $\mathbf{g}_2^{\mathrm{B}}$ differ only in sign due to the equilibrium condition. From the fourth row of Equation 1, we obtain:

$$\mathbf{u}_3 = \mathbf{Y}_{32}^{\mathrm{B}}\mathbf{g}_2^{\mathrm{B}} \tag{2}$$

From this, it can be directly seen that the response at the receiver $\mathbf{u}_3$ is determined by the interface force $\mathbf{g}_2^{\mathrm{B}}$ and the transfer behavior from the interface to the receiver point $\mathbf{Y}_{32}^{\mathrm{B}}$. This type of TPA is referred to as "Classical TPA".
In practice, the transfer behavior is often determined using impact hammer or shaker measurements. When measuring $\mathbf{Y}_{32}^{\mathrm{B}}$, the source must be decoupled, as only the transfer behavior of the receiver itself (denoted by the superscript $(\cdot)^{\mathrm{B}}$) is required. The interface forces $\mathbf{g}_2^{\mathrm{B}}$ can be determined using various methods, namely the "Direct Force" method or the "Mount Stiffness" method. While the interface forces in the "Direct Force" method are directly measured using force transducers placed between the source and the receiver, the "Mount Stiffness" method determines these forces by measuring the displacement, assuming the stiffness of the mounting is known. Another well-studied method is the "Matrix Inverse" method, which shares many similarities with the later introduced "In situ Blocked Force" method. In this method, it is assumed that the transfer admittance $\mathbf{Y}_{42}^{\mathrm{B}}$ from the interface to additional receiver points $\mathbf{u}_4$ positioned on the receiver can be determined in an initial step, for example, through an impact hammer measurement. In the second step, an operational measurement of the assembled system is conducted to determine $\mathbf{u}_4$ when the source is running. Using the relationship from Equation 2 and replacing $\mathbf{u}_3$ with $\mathbf{u}_4$, it follows that

$$\mathbf{g}_2^{\mathrm{B}} = \left(\mathbf{Y}_{42}^{\mathrm{B}}\right)^{-1}\mathbf{u}_4 \tag{3}$$

which allows the interface force $\mathbf{g}_2^{\mathrm{B}}$ to be computed. Since the receiver points $\mathbf{u}_4$ are not used as target quantities for purposes such as predictions, but rather for determining the interface forces, they are often referred to as indicator points.
Although classical TPA is a useful tool for many applications and still used in practice, it has a fundamental drawback: the interface forces $\mathbf{g}_2^{\mathrm{B}}$ used in this approach are influenced not only by the internal excitation but also by the structural dynamic behavior of the receiver. As a result, these forces do not represent a pure source quantity, making them not transferable to other receiver systems and unsuitable for evaluating the source itself.

2.2 Component-based TPA

The limitations of classical TPA are overcome in component-based TPA (cTPA). In this approach, the source is described using receiver-independent interface quantities, which offers

several advantages. Firstly, a receiver-independent source description allows the prediction of the vibration response $\mathbf{u}_3$ for different receivers, assuming the transfer behaviour of the receiver can be determined. Secondly, and particularly relevant for this publication, the receiver-independent interface quantities enable the evaluation of the source itself in terms of SBN emission. Figure 3 shows an overview of the most important source characterisation methods, which are introduced in the following.

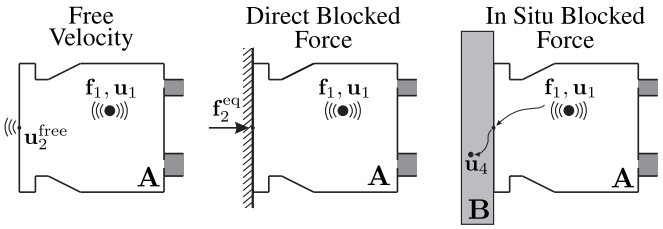


Figure 3: Source characterisation methods in cTPA

2.2.1 Free velocity

The most obvious method to characterize a source without receiver system influence is to determine an interface quantity when the source is freely vibrating, i.e., not connected to any receiver system. In this case, the force at the interface vanishes, i.e., $\mathbf{g}_2^{\text{A}} = 0$. The velocity at the interface is referred to as the free velocity, $\mathbf{u}_2^{\text{free}}$. Under this condition, the second row of Equation 1 becomes

$$\mathbf{u}_2^{\text{free}} = \mathbf{Y}_{21}^{\text{A}} \mathbf{f}_1 \tag{4}$$

from which it is evident that the free velocity depends solely on source quantities, thereby allowing for a description of the source excitation at the interface. Although the characterization of SBN sources using free velocities is theoretically possible, this method is not applicable to every kind of source. This is primarily because the source must be operated under free conditions. For hydraulic pumps in particular, this is not practical, as it would eliminate any means of driving the pump.

2.2.2 Direct blocked force

Another method for eliminating environmental influences is to theoretically attach the source to an infinitely stiff receiver. In this scenario, the velocity at the interface vanishes, i.e., $\mathbf{u}_2^{\text{A}} = \mathbf{u}_2^{\text{B}} = 0$. The force at the interface is then referred to as the blocked force, $\mathbf{f}_2^{\text{eq}}$ (The superscript $(\cdot)^{\text{eq}}$ denotes "equivalent," as the blocked forces offer an equivalent representation of the source excitation). Under these conditions, the second row of Equation 2 becomes:

$$\begin{aligned} 0 &= \mathbf{Y}_{21}^{\text{A}} \mathbf{f}_1 + \mathbf{Y}_{22}^{\text{A}} \mathbf{g}_2^{\text{A}} \quad \text{for} \quad \mathbf{f}_2^{\text{eq}} = \mathbf{g}_2^{\text{bl}} = -\mathbf{g}_2^{\text{A}} \\ \Rightarrow \mathbf{f}_2^{\text{eq}} &= \left(\mathbf{Y}_{22}^{\text{A}}\right)^{-1} \mathbf{Y}_{21}^{\text{A}} \mathbf{f}_1 \end{aligned} \tag{5}$$

From this, it can be seen that the blocked forces depend solely on the source itself, making them an appropriate interface quantity as well. Furthermore, by substituting the term $\mathbf{Y}_{21}^{\text{A}} \mathbf{f}_1$ in Equation 4 with the relationship from Equation 5, it can be seen that the blocked forces and the free velocities are directly related through the equation

$$\mathbf{u}_2^{\text{free}} = \mathbf{Y}_{22}^{\text{A}} \mathbf{f}_2^{\text{eq}}. \tag{6}$$

As with the interface forces in the classic TPA, there are several methods for determining the blocked forces. The most obvious method is direct measurement using force transducers between the source and a receiver with sufficient stiffness. Nevertheless, previous studies [17, 18] have demonstrated that the condition of sufficiently high receiver stiffness can only be met within a limited frequency range. Furthermore, mounting force transducers between the source and the receiver presents several of challenges. Firstly, the stiffness of the force transducers must be sufficiently high to meet the stiffness requirements. Secondly, the sensors must not alter the interface itself. While this is less of an issue for point contacts, it becomes a significant challenge for surface-like interfaces, as is typically the case for hydraulic pumps. Moreover, directly measuring torques and in-plane forces is difficult to achieve due to the complex sensor alignment. Consequently, direct measurement of blocked forces is impractical for many applications, including hydraulic pumps.

2.2.3 In situ blocked force

Another method, first introduced by Moorhouse et al. [19], is the so-called "In Situ Blocked Force" method. This approach is comparable to the "Matrix Inversion" method of classical TPA but allows for the determination of blocked forces "in situ," meaning in the typical installation condition of the source. This method has gained popularity in recent years, particularly in the automotive industry, leading to numerous scientific publications in this field [20, 21]. Additionally, the procedure for conducting an in situ blocked force measurement has been defined in the ISO 20270 standard since 2019. In [19], Moorhouse et al. demonstrated that the blocked force can be determined using the following relationship (for a detailed derivation, see [19] or [15]):

$$\mathbf{f}_2^{\text{eq}} = \left(\mathbf{Y}_{42}^{\text{AB}}\right)^{-1} \mathbf{u}_4 \tag{7}$$

Analogous to the matrix inversion method, $\mathbf{u}_4$ denotes additional indicator points on the receiver structure. The right-hand side of of Equation 7 differs from that of Equation 3, which describes the "Matrix Inversion" method of classical TPA, in that it requires the admittance of the assembled system, $\mathbf{Y}_{42}^{\text{AB}}$, instead of the admittance $\mathbf{Y}_{42}^{\text{B}}$ of the receiver only. From Equation 7, it follows that the measurement process for determining the blocked forces consists of two steps. In the first step, the admittance matrix $\mathbf{Y}_{42}^{\text{AB}}$ from the interface to the indicator points for the assembled system is determined. For this purpose, the structure is excited close to the interface using an impulse hammer or shaker, and the response at the indicator points is measured using accelerometers. In the second step, the response at the indicator points $\mathbf{u}_4$ is measured while the source is operating, with the accelerometers remaining at the same locations as

during the impulse hammer measurement. The blocked forces are then determined via Equation 7 at the positions where the excitation was applied in the first step using the impulse hammer. By following these steps, the source and receiver do not need to be separated at any stage of the measurement process. This allows the blocked forces to be determined in the coupled state (in situ).

Figure 4, adapted from [22], visually highlights the crucial difference between describing the interface excitation using actual interface forces $\mathbf{g}_2^{\mathrm{A}}$ (classical TPA) and blocked forces $\mathbf{f}_2^{\mathrm{eq}}$ (component-based TPA). When the source is fully characterized by the corresponding interface quantity, both methods ultimately lead to the original vibration response at the receiver. In classical TPA, the interface forces are combined with the admittance $\mathbf{Y}_{32}^{\mathrm{B}}$ of the receiver only, whereas in component-based TPA, the blocked forces are applied to the coupled system $\mathbf{Y}_{32}^{\mathrm{AB}}$.

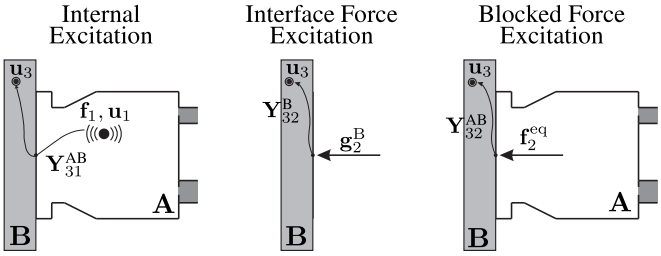


Figure 4: Description of the source excitation with Interface Forces and Blocked Forces

Although "in situ" can be translated to "in its original place," the blocked forces, due to their independence from the receiver system, do not necessarily need to be determined in the target application. Instead, they can also be obtained on a specially designed test bench. The main advantage here, besides the measurement under controlled laboratory conditions, is primarily the accessibility of the interface. The receiver can be designed in such a way that the interface, or the area around the interface, can be easily excited with an impact hammer or a shaker. This is often not possible in a mobile machine with limited space. Once a test bench for the in situ method is developed, the blocked forces for any operating point of the pump can be determined under controlled test conditions.

Additionally the in situ blocked force relationship from Equation 7 not only allows for the determination of blocked forces but also opens up possibilities for Virtual Acoustic Prototyping (VAP). Since the blocked forces $\mathbf{f}_2^{\mathrm{eq}}$ are independent of the receiver system, they can be combined with any arbitrary receiver. More specifically, once a source is characterized by blocked forces, it can be virtually combined with various receiver systems to simulate the resulting noise for those receivers. This opens up possibilities such as the virtual investigation of how noise mitigation measures, tested on the test bench, affect the noise in a mobile working machine. The transfer behavior $\mathbf{Y}_{32}^{\mathrm{AB}}$ required for this purpose can be determined either through an initial measurement in the vehicle or by employing dynamic substructuring approaches.

2.3 Quality criteria in situ blocked force method

The quality of the results when determining the blocked forces using the in situ method can be evaluated using various criteria. A fundamental prerequisite for achieving satisfactory results according to these quality criteria is that the transfer admittance $\mathbf{Y}_{42}^{\mathrm{AB}}$ can be determined with sufficient accuracy. There are specific quality criteria for this, which will not be discussed here, as they are general criteria for the experimental determination of transfer functions, as elaborated upon in [16]. In the following, the criteria of On-Board Validation and Transferability, which are specifically suitable for the in situ blocked force method, are introduced.

2.3.1 On-Board Validation

The most straightforward and particularly easy-to-apply method for in situ determination of blocked forces is the On-Board Validation (OBV). In this approach, the blocked forces, determined using indicator points in conjunction with Equation 7, are used to make a prediction for the receiver points $\mathbf{u}_3$:

$$\mathbf{u}_3 = \mathbf{Y}_{32}^{\mathrm{AB}} \mathbf{f}_2^{\mathrm{eq}} \tag{8}$$

The required transfer admittance $\mathbf{Y}_{32}^{\mathrm{AB}}$ and vibration response $\mathbf{u}_3$ are typically obtained simultaneously with the determination of $\mathbf{Y}_{42}^{\mathrm{AB}}$ and $\mathbf{u}_4$, making it an easy-to-apply method. To evaluate the source description, the predicted vibrations $\mathbf{u}_3$ at the receiver points are compared with the directly measured ones. The OBV method can be particularly useful for assessing whether the number of interface degrees of freedom (DOFs), and thus the number of blocked forces, is suitable to accurately describe the source. Additionally, OBV can help identify errors or issues in the inversion of the admittance matrix $\mathbf{Y}_{42}^{\mathrm{AB}}$.

In addition to OBV during the operation of the source, an OBV for so-called artificial excitations is typically conducted. In this case, the source itself is not in operation, and artificial excitations are applied to the source using an impact hammer or shaker. This approach offers the advantage that the artificial excitations provide a broad-band excitation and do not contain complex features like the operational response, which can complicate the inspection of agreement [16]. Consequently, the description of the interface can be evaluated more easily.

2.3.2 Transferability

As mentioned, OBV is a straightforward method for evaluating blocked forces. However, a good result in OBV is a necessary but not sufficient condition. While OBV can assess the description of the interface and the matrix inversion, it does not verify whether the blocked forces are independent of the receiver system, which is a fundamental requirement for a reasonable source description in cTPA. To evaluate this, the concept of transferability is used, which examines whether the determined blocked forces can be transferred to another receiving system. For this purpose, the blocked forces are determined for an identical source excitation in two different assemblies **AB** and **AC**. If the condition of receiver independence is satisfied, the response at the receiver points $\mathbf{u}_3$

of assembly **AC** can be predicted using the blocked forces determined in assembly **AB**. Or mathematically expressed:

$$\mathbf{u}_3^{\mathrm{AC}} = \mathbf{Y}_{32}^{\mathrm{AC}} \mathbf{f}_2^{\mathrm{eq,AB}} \tag{9}$$

It should be noted that deviations can arise not only due to nonlinearities, insufficient interface description, or numerical errors but also from variations in the excitation mechanism. Because of this, it is crucial to ensure that the operating conditions are as similar as possible in both operational tests. For hydraulic pumps, this particularly includes pressure and rotational speed. Additionally, the passive properties on the source side, such as the hosing, should be as identical as possible for both systems.

2.4 In situ blocked force determination for an axial piston pump

As with many SBN sources, the in situ method is considered the most suitable source characterization method for hydraulic pumps. The determination of free velocities and the direct measurement of blocked forces are less suitable due to the aforementioned practical reasons.
In the case of the axial piston pump examined in this section, the blocked forces were not determined in a typical application, such as a mobile machine; instead, a test bench was specifically developed for this purpose, which offers the aforementioned advantages. Additionally, for a comprehensive characterization of the source excitation at the interface, it is crucial to excite various interface Degrees of Freedom. This means that excitation with the impulse hammer or shaker should be possible in different directions and at various positions around the interface. In practical applications, especially excitation in the plane direction is often not feasible. These challenges can be mitigated by conducting measurements on the test bench, thereby enhancing the quality of the results.

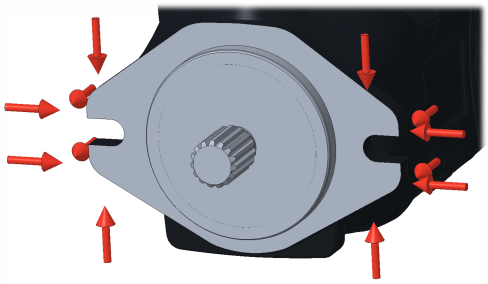


Figure 5: Possible Impact Positions for an Axial Piston Pump (Receiver not shown)

Figure 5 illustrates a potential positioning of excitation points for an impulse hammer or shaker in proximity to the mounting of an axial piston pump, which features a two-bolt mounting flange according to DIN ISO 3019-1 and is typically used in 6-ton mini excavators. It should be noted that the coupled receiver is not shown in Figure 5, as the design of the receiver can theoretically be arbitrary. However, this only applies when the internal excitation mechanism of the source is not influenced by the the receiver. Since in an axial piston pump the main SBN excitation is caused by pressure fluctuation within the power unit and this is not directly influenced by the external movement of the pump, it is assumed that this condition is met.
To measure $\mathbf{Y}_{32}^{\mathrm{AB}}$ and $\mathbf{Y}_{42}^{\mathrm{AB}}$, the pump and receiver must be coupled, making direct excitation at the interface infeasible. Therefore, excitation points must be positioned around the interface. In the illustrated case, six impacts are arranged around each of the two mounting points, based on the assumption that the excitation from the source is primarily transmitted to the receiver through the mounting points. In order to excite all interface DoF, impacts are applied in the normal direction and the in-plane directions.
The accelerometers used to determine the admittance matrices can theoretically be positioned anywhere downstream of the interface on the receiver. However, in practice, they should be placed in close proximity to the interface to achieve a good signal-to-noise ratio (SNR) and to minimize sources of error such as nonlinearities along the transmission path. To solve Equation 7, it is necessary to measure at least as many response signals (indicator DoFs) as there are excitation points (interface DoFs). However, according to ISO 20270-1-2019, it is recommended to overdetermine the system by using more indicator DoFs than interface DoFs. This results in an overdetermined system, and the inverse matrix $\left(\mathbf{Y}_{42}^{\mathrm{AB}}\right)^{-1}$ in Equation 7 is replaced by the pseudoinverse $\left(\mathbf{Y}_{42}^{\mathrm{AB}}\right)^{+}$. Consequently, the matrix inversion becomes a least-squares optimization, which reduces the influence of measurement errors and noise (for a detailed explanation, see [16]). For the measurement setup of the axial piston pump, seven triaxial indicator sensors ($\mathbf{u}_4$) were used, resulting in an approximately twofold overdetermined system of equations. Additionally, two triaxial receiver sensors ($\mathbf{u}_3$) were employed for evaluation using OBV and transferability.

2.5 Evaluation of structure borne noise emission of axial piston Pumps

In the previous section, the procedure for determining the blocked forces of an axial piston pump was explained. The following introduces a metric for evaluating an axial piston pump's SBN emission based on the blocked forces.
The characterization of structure-borne sound sources has been a subject of investigation for many years. The Technical Committee on Acoustics TC43 of the International Organization for Standardization (ISO) has defined four requirements in [16, 23] that should be met by a characterization method:

1. Comparison of structure-borne sound emissions of similar machines
2. Examination of compliance with limit values
3. Determination of input data for acoustic planning (e.g. prediction of sound levels when installed)
4. Collection of necessary data for the development of quiet machines

In addition to the listed requirements, the Committee also

proposed various characterization methods. These methods, along with many subsequently developed approaches such as the Source Descriptor method [24] and the Characteristic Power method [25], aim to fulfil the aforementioned requirements using one or two frequency-dependent single-value parameters. In these approaches objective is to facilitate a straightforward characterization of the source, analogous to the use of sound power in the characterization of ABN.

However, it has been demonstrated that, unlike ABN, a complete characterization of SBN sources using one or two frequency-dependent single-value parameters is hardly feasible. This is due to the significantly different nature of SBN (for a detailed explanation, see [16]). To specifically address the third requirement, which is the prediction of noise levels when the source is installed, the mentioned single-value-based methods are rather inadequate, particularly when an accurate prediction is required. Consequently, new approaches such as the in situ blocked force method have been developed in recent years [16]. These methods provide a more comprehensive description of the source, thereby enabling advanced concepts like cTPA and VAP, which are particularly relevant in later stages of development when specific applications for the source are known or for root cause analysis in a vehicle exhibiting poor noise performance.

However, in an early stage of hydraulic pump development, these concepts are of less interest. This is primarily due to the wide range of applications for a hydraulic pump, which vary significantly in their dimensions and acoustic properties. At an early stage of pump development, it is particularly important to investigate the effectiveness of measures, compare different pump designs, and identify critical operating points (1st and 4th requirement for a source characterization method). For this purpose, a straightforward evaluation method is needed, which may not provide a complete source description but offers a simple assessment that is independent of a receiving system.

To provide both a comprehensive source description for application-specific investigations and a straightforward evaluation method for early-stage development, the evaluation metric should be derived from a complete source description. In the case of the axial piston pump described above, the evaluation quantity should be derived from the twelve determined blocked forces.

There are various methods [25–27] to combine interface forces such as blocked forces. In the evaluation method introduced here, the forces are aggregated in a straightforward manner by calculating the vector norm of the resultant force vector of the blocked forces $\mathbf{m}_2^{\mathrm{eq}} = \begin{bmatrix} m_{2,X}^{eq} & m_{2,Y}^{eq} & m_{2,Z}^{eq} \end{bmatrix}^T$. In the first step, all blocked forces are combined into three resultant blocked forces $m_{2,X}^{eq}, m_{2,Y}^{eq}, m_{2,Z}^{eq}$, one for each coordinate direction. The combined blocked forces are denoted by the variable m in accordance with the Virtual Point Transformation, which will be briefly described in the subsequent paragraph. Assuming that the interface behaves like a rigid body, no information is lost during this combination step; however, this condition is only met within a limited frequency range. In the second step, the vector norm of the resultant blocked forces vector is calculated as follows:

$$\|\mathbf{m}_2^{\mathrm{eq}}\| = \sqrt{\left(m_{2,X}^{eq}\right)^2 + \left(m_{2,Y}^{eq}\right)^2 + \left(m_{2,Z}^{eq}\right)^2} \tag{10}$$

yielding a single-valued, frequency-dependent force quantity. This combination using the vector norm results in the loss of phase information, rendering the combined force is unsuitable for cTPA and VAP.

It should be noted that the aggregation of the blocked forces into three resultant forces in the first step corresponds to the so-called Virtual Point Transformation (VPT), without considering the torques. A detailed discussion of the VPT is beyond the scope of this paper and can be found in [17, 27]. However, it should be mentioned that neglecting the torques also represents a significant simplification and, as previous studies have shown, can lead to substantial errors. Nevertheless, they are not considered here for the sake of simplicity as a complete VPT would be required, which need additional information. Moreover, since the resultant forces and torques of a comprehensive VPT are physically different quantities, they cannot be directly combined into a single evaluation metric. To achieve this, an advanced approach, such as a power-based method, would be required.

3 Results

In this chapter, the first part presents the results of the in situ blocked force measurement for the axial piston pump described above and evaluates them using the introduced quality criteria. The second part demonstrates the behavior of the introduced evaluation metric for two characteristic operating points of the pump.

3.1 Results in situ blocked force determination

Before presenting the results in the following section, it should be noted that, although velocities and admittances were used to describe the physical relationships in the previous chapter, the subsequent analysis will compare predicted and measured acceleration signals. This is because accelerations were directly measured and contain the same information. Thus, the measured transfer functions, also known as Frequency Response Functions (FRFs), describe the relationship between the resulting acceleration and the applied force excitation.

Figure 6 shows the result of the OBV for an artificial excitation using an impulse hammer impact on the pump housing. Specifically, the vector norm of the acceleration signal from all three measurement channels of a triaxial accelerometer at a receiver point in the frequency domain is depicted. The y-axis represents the acceleration in decibels (dB), with a reference value of $10^{-6}\,\mathrm{m/s^2}$. This reference value is uniformly used in all subsequent figures where the acceleration in dB is shown. It can be observed that for the frequency range $150\ \mathrm{Hz} < f < 2500\ \mathrm{Hz}$, the prediction and measurement of the signal at the receiver point match very well. Beyond 2500 Hz, the signal becomes significantly noisier, indicating a poorer SNR. This can be attributed to the decreasing frequency response of the impulse hammer with a plastic tip at higher

frequencies. This outcome indicates that, in further analysis, good results can only be expected for the frequency range of 150 Hz $< f <$ 2500 Hz. Therefore, this frequency range will be examined in more detail in the following.

Figure 7 shows the OBV for a typical operating point with $n = 1500$ rpm, $p = 200$ bar, and $V_d = V_{d,max}$ for a frequency range up to 2500 Hz. As can be seen, the prediction and measurement align very well across the entire frequency range.

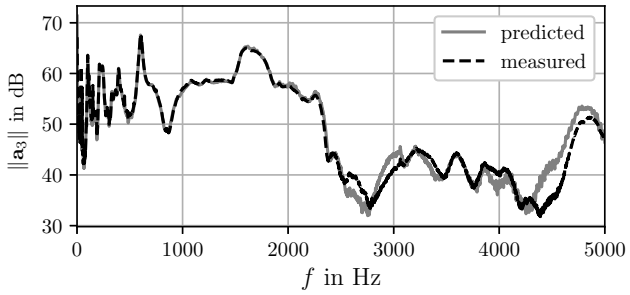


Figure 6: Result OBV for an artificial excitation

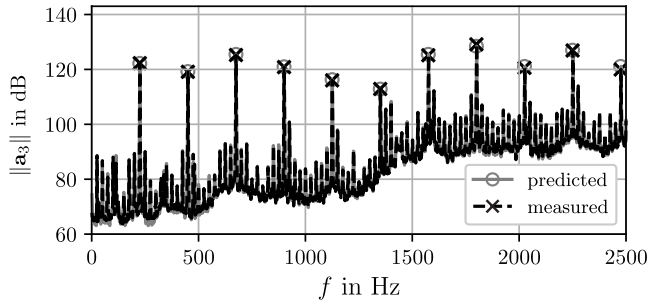


Figure 7: Result OBV for the stationary operating point: $n =$ 1500 rpm, $p = 200$ bar, $V_d = V_{d,max}$

As explained in the section on OBV, the result demonstrates that a set of twelve blocked forces is suitable for describing the source in this frequency range. However, what the OBV cannot yet demonstrate is whether the determined blocked forces are independent of the receiver system. To address this, the next step involves investigating the transferability of the blocked forces. For this purpose, the blocked forces determined in system **AB** are applied to another system **AC**. To ensure a valid transferability check, systems **AB** and **AC** must differ in their structural dynamic behavior. For the axial piston pump under investigation, the behavior of the receiver was influenced by varying the number of mounting screws on the receiver flange. The extent to which the admittance matrices $\mathbf{Y}_{42}^{\mathrm{AB}}$ and $\mathbf{Y}_{42}^{\mathrm{AC}}$ differ can be estimated by the mean magnitude difference of the indicator FRFs, as shown in Figure 8. The plot illustrates the percentage difference between the mean magnitude of all FRFs in system **AB** and those in system **AC**, relative to the values from system **AB**, as a function of frequency. From the plot, it can be observed that the deviation is relatively small for the frequency range 150 Hz $< f <$ 450 Hz. It is only at frequencies above 450 Hz that the dynamic behaviour of the structure can be expected to be significantly affected, making the investigation of transferability only meaningful beyond this frequency.

Figure 9 presents the transferability check for the same operating point as in Figure 7. The insights from Figure 8 can also be observed when comparing Figure 7 and 9. While the peak values at the harmonics for frequencies up to 1000 Hz are nearly identical in Figures 7 and 9, they diverge more significantly at higher frequencies.

Examining the transferability check, it is evident that the measured and predicted accelerations align well for $f <$ 1300 Hz, with greater deviations occurring beyond this point. These discrepancies can be attributed to various factors, such as measurement errors during the impulse hammer test, numerical errors, or nonlinearities. It is also important to note that, unlike the OBV, the transferability check pertains to the same operating points but not exactly the same operational measurement. Consequently, differences in operating conditions, such as variations in oil temperature or air content in the oil, can also influence the results.

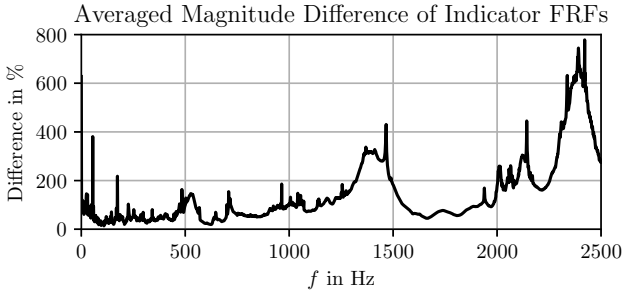


*Figure 8: Averaged magnitude difference of the indicator FRFs of System **AB** and **AC***

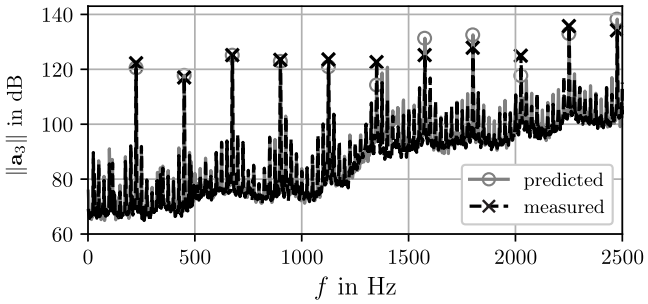


Figure 9: Result Transferability for the stationary operating point: $n = 1500$ rpm, $p = 200$ bar, $V_d = V_{d,max}$

3.2 Results evaluation metric

As the results of the in situ measurements demonstrate, the structure-borne noise excitation by an axial piston pump can be described for a using twelve blocked forces. Based on the blocked forces, the evaluation metric introduced in the previous chapter can be computed. Figures 10 and 11 illustrate the behavior of the evaluation metric for maximum and minimum displacement volumes ($V_{d,min}$: minimum displacement to maintain the target pressure), with values obtained from both system **AB** and system **AC**. Since the evaluation metric is an aggregation of the blocked forces, it has the unit Newton. For the determination of the decibel value, a reference value of 10^{-6} N is used.

When analyzing Figure 10, it is noticeable that the qualitative behavior of the evaluation metric significantly differs from that of the acceleration at the receiver points in Figures 7 and 9. This difference is primarily due to the rigid test setup, which causes the receiver sensors to experience stronger vibrations only at higher frequencies. Instead, Figure 10 clearly shows that the structure-borne noise excitation at this operating point is mainly dominated by the first harmonic, with the influence of higher harmonics generally decreasing. Comparable values of the evaluation metric are obtained for both systems **AB** and **AC**, confirming the independence from the receiver system.
Figure 11 shows the evaluation metric for the operating point with the same rotational speed and pressure but with minimal displacement volume. As can be seen, the spectrum does not fundamentally change due to the same rotational speed; however, the quantitative values of some harmonics in Figure 10 and Figure 11 differ significantly. While the evaluation metric for maximum displacement is clearly dominated by the first harmonic, the peaks at minimum displacement are at a similar level over a broader frequency range. The evaluation metric for both systems also aligns well for this operating point.

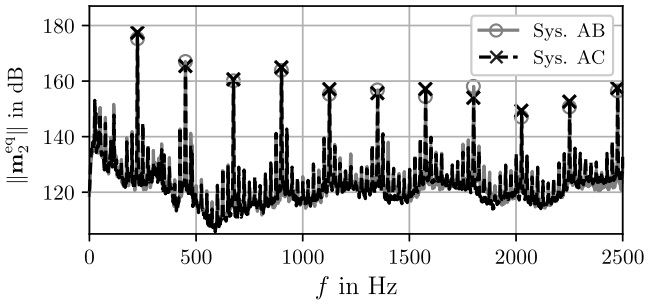


Figure 10: Evaluation quantity $\|\mathbf{m}_2^{\mathrm{eq}}\|$ for the operating point: $n = 1500$ rpm, $p = 200$ bar, $V_d = V_{d,max}$

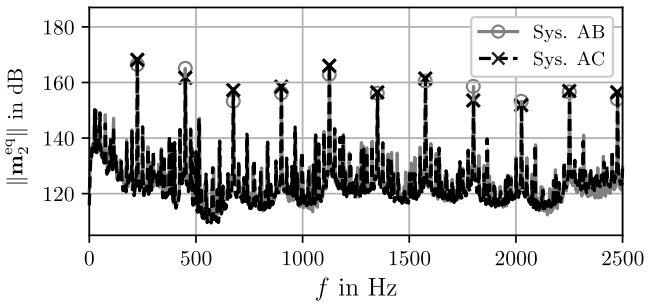


Figure 11: Evaluation quantity $\|\mathbf{m}_2^{\mathrm{eq}}\|$ for the operating point: $n = 1500$ rpm, $p = 200$ bar, $V_d = V_{d,min}$

4 Discussion

The presented results demonstrate that the in situ method is suitable for determining the blocked forces of an axial piston pump. According to the OBV results, the pump can be characterized using 12 blocked forces within a frequency range of $150\,\mathrm{Hz} < f < 2500\,\mathrm{Hz}$. However, the analysis of transferability indicates that these forces can only be transferred to another receiver system up to a frequency of approximately $f < 1300\,\mathrm{Hz}$. This limitation could be attributed not only to insufficient source characterization but also to measurement errors or variations in the operational measurement.
Given that the frequency range up to approximately 5000 Hz is particularly relevant for the acoustics of axial piston pumps [28], the procedure for determining the blocked forces must be further optimized to enable reliable source characterization within the relevant frequency range. To achieve this, sufficient energy must be introduced into the system across the entire frequency range of interest during the FRF measurement. This can be accomplished by using different impact hammer tips or a shaker. Additionally, an investigation of the required interface degrees of freedom could help analyze whether the twelve blocked forces used are adequate for characterizing the axial piston pump.
The introduced single-valued evaluation method significantly simplifies the source description using twelve blocked forces. However, it facilitates a straightforward comparison of different operating points or potential pump measures.The suitability of this evaluation method for assessing the SBN emission of axial piston pumps remains to be ascertained, as it necessitates validation through additional measurements. Nevertheless, it represents an initial approach to evaluating the emitted SBN independently of a receiver system and provides a foundation for further investigations.

5 Conclusion and outlook

To meet increasing noise requirements for mobile machinery, it is crucial to evaluate the acoustic behavior of axial piston pumps. Unlike ABN and FBN, no established method exists for assessing SBN emissions from hydraulic pumps. Therefore, a method has been introduced that derives a single-valued evaluation metric based on a comprehensive source description using blocked forces, enabling a straightforward assessment of the pump's SBN characteristics.
As the investigation shows, the in situ blocked force method is a suitable approach for the comprehensive source characterization of axial piston pumps with respect to SBN. However, to accurately describe the pump across the entire relevant frequency range, the procedure for determining the blocked forces must be further optimized. This particularly requires improvements in the measurement procedure.
The single-valued evaluation metric derived from the source description using blocked forces represents a significant simplification of the source characterization. However, it enables straightforward comparisons of operating points and measures, which is essential in early stage pump development.
The validity of the evaluation method cannot be conclusively assessed at this stage and must be examined in further studies. Additionally, the metric should be compared with other approaches for summarizing interface quantities like power-based methods.

6 Acknowledgements

Acknowledgements are given to Bosch Rexroth AG for the financial support and the provision of the acoustic test bench, as well as to the Bosch Sound Propagation expert team for their technical and professional support.

References

[1] T. Pietrzyk, M. Georgi, S. Schlittmeier, and K. Schmitz. Psychoacoustic Evaluation of Hydraulic Pumps. *Sustainability*, 13(13):7320, June 2021.

[2] T. Nafz. *Aktive Ventilumsteuerung von Axialkolbenpumpen zur Geräuschreduktion hydraulischer Anlagen*. PhD thesis, RWTH Aachen, Aachen, 2011. ISBN 978-3-8440-0511-0.

[3] H. Theissen. Volumenstrompulsation von Kolbenpumpen. *O+P Ölhydraulik und Pneumatik*, 24, 1980.

[4] U. Bittner. *Strukturakustische Optimierung von Axialkolbeneinheiten: Modellbildung, Validierung und Topologieoptimierung*. PhD thesis, Karlsruher Institut für Technologie, Karlsruhe, 2013. ISBN 978-3-86644-938-1.

[5] D. E. Bowns, K. A. Edge, and D. G. Tilley. The Assessment of Pump Fluid Borne Noise. In *Proceedings of I. Mech. E. Conf.-Quiet Oil Hydraulic Systems*, London, England, 1977.

[6] P. Larsson, J. Palmberg, and K. Weddfelt. Analysis and measurement of pressure ripple of fluid power pumps. In *Proceedings of Eighth IASTED International Symposium on Measurement, Signal Processing and control*, Taormina, Italy, 1986.

[7] K. A. Edge and D. N. Johnston. The 'Secondary Source' Method for the Measurement of Pump Pressure Ripple Characteristics Part 1: Description of Method. *Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy*, 204(1):33–40, February 1990.

[8] K. A. Edge. The 'Secondary Source' Method for the Measurement of Pump Pressure Ripple Characteristics Part 2: Experimental Results. *Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy*, 204(1):41–46, February 1990.

[9] K. Weddfelt. Measurement of pump source characteristics by the two-microphone method. In *Proceedings of The Second Tampere International Conference on Fluid Power*, Tampere, Finland, 1991.

[10] E. Kojima. A new method for the experimental determination of pump fluid-borne noise characteristics. In *Proceedings of 5th Bath International Fluid Power Workshop, Circuit, Component and System Design*, pages 111–135, Bath, England, 1992.

[11] M. Haarhaus. *Geräuschentstehung und Geräuschminderung bei Axialkolbenpumpen in Schrägscheibenbauweise*. PhD thesis, RWTH Aachen, Aachen, 1984.

[12] M. Bergemann. *Systematische Untersuchung des Geräuschverhaltens von Kolbenpumpen mit ungerader und mit gerader Kolbenanzahl*. PhD thesis, RWTH Aachen, Aachen, 1994. ISBN 978-3-930085-23-1.

[13] M. Pettersson. *Design of fluid power piston pumps: with special reference to noise reduction*. PhD thesis, Linköping University, Linköping, 1995. ISBN 978-91-7871-592-3.

[14] B. Müller. *Einsatz der Simulation zur Pulsations- und Geräuschminderung hydraulischer Anlagen*. PhD thesis, RWTH Aachen, Aachen, 2002. ISBN 978-3-8322-0793-9.

[15] M. V. van der Seijs, D. de Klerk, and D. J. Rixen. General framework for transfer path analysis: History, theory and classification of techniques. *Mechanical Systems and Signal Processing*, 68-69:217–244, February 2016.

[16] J. W. R. Meggitt. An introduction to the in-situ blocked force method for vibration source characterisation. Technical report, Acoustics Research Centre - University of Salford, 2022.

[17] M. V. van der Seijs. *Experimental Dynamic Substructuring: Analysis and Design Strategies for Vehicle Development*. PhD thesis, Delft University of Technology, 2016.

[18] M. V. van der Seijs, E. Pasma, D. de Klerk, and D. J. Rixen. A robust Transfer Path Analysis method for steering gear vibrations on a test bench. In *Proceedings of ISMA 2014*, Leuven, Belgium, 2014.

[19] A. T. Moorhouse, A. S. Elliott, and T. A. Evans. In situ measurement of the blocked force of structure-borne sound sources. *Journal of Sound and Vibration*, 325(4-5):679–685, September 2009.

[20] C. L. Balogh, T. Kimpián, and T. Ungvári. An Examination of Component-Based TPA Performance and Feasibility for Steering Motors. *Journal of Physics: Conference Series*, 2677(1):012011, December 2023.

[21] M. Burkhardt, E. Hensel, and W. Drossel. Messdatenbasierte Charakterisierung von Körperschallquellen in sechs Freiheitsgraden. In *DAGA 2020 - 46. Jahrestagung für Akustik*, Hannover, Germany, 2020.

[22] A. T. Moorhouse, A. S. Elliott, and J. W. R. Meggitt. Virtual Acoustic Prototyping - A story of four decades. In *Proceedings of Inter-Noise 2022*, Glasgow, 2022.

[23] T. ten Wolde and G. Gadefelt. Development of Standard Measurement Methods for Structureborne Sound Emission. *Noise Control Engineering Journal*, 28(1):5, 1987.

[24] J. M. Mondot and B. Petersson. Characterization of structure-borne sound sources: The source descriptor and the coupling function. *Journal of Sound and Vibration*, 114(3):507–518, May 1987.

[25] A. T. Moorhouse. On the characteristic power of structure-borne sound sources. *Journal of Sound and Vibration*, 248(3):441–459, November 2001.

[26] M. M. Späh and B. M. Gibbs. Reception plate method for characterisation of structure-borne sound sources in buildings: Assumptions and application. *Applied Acoustics*, 70(2):361–368, February 2009.

[27] M. van der Seijs, D. van den Bosch, D. Rixen, and D. de Klerk. An improved methodology for the virtual point transformation of measured frequency response functions in dynamic substructuring. In *Proceedings of COMPDYN 2013*, pages 4334–4347, Kos Island, Greece, 2014.

[28] H. Münch. *Acoustically extended virtualization of industrial productsAkustisch erweiterte Virtualisierung von Industrieprodukten*. PhD thesis, FAU Erlangen-Nürnberg, Erlangen, 2021. DOI 10.25593/978-3-96147-387-8.

A Class-incremental Learning Method based on Knowledge Distillation for Multi-fault Diagnosis of Axial Piston Pump

Dandan Wang, Shihao Liu, Junhui Zhang*, Fei Lyu, Weidi Huang, and Bing Xu

State Key Laboratory of Fluid Power and Mechatronic System, Zhejiang University, Hangzhou, China
E-mail: ddwang@zju.edu.cn, shihaoliu@zju.edu.cn, benzjh@zju.edu.cn, feilv@zju.edu.cn, wdhuang@zju.edu.cn, bxu@zju.edu.cn

Abstract

Fault diagnosis of the axial piston pump is critical for maintaining the safety and reliability of modern hydraulic transmission and control systems. Existing deep-learning fault diagnosis models demonstrate effective performance but fail to generalize if new fault patterns occur. Simply fine-tuning these models only with newly collected data leads to the catastrophic forgetting problem, whereas retraining a new fault diagnosis model with the entire historical data is both resource-intensive and time-consuming. To address these issues, a novel class-incremental learning method based on knowledge distillation is proposed for multi-fault diagnosis of the axial piston pump, which can continually learn new fault patterns and preserve fault diagnosis ability on old fault patterns with a limited amount of historical data. To mitigate catastrophic forgetting in subsequent stages, the discrepancy between output-logits of the previous model and that of the current one is introduced in the incremental learning process. Additionally, intermediate feature relationships with different important weights are also employed to further enhance fault diagnosis performance on old fault patterns. Both the comparison experiment and the ablation experiment demonstrate the effectiveness of the proposed method.

Keywords: Axial piston pump, class-incremental learning, fault diagnosis, knowledge distillation

1 Introduction

The axial piston pump serves as a pivotal power component in advanced hydraulic applications, such as aerospace actuators, submarine vessels, and mobile machinery, owing to its superior power density and energy efficiency[1-3]. Ensuring its reliable operation is paramount to maintaining the stability and performance of the whole hydraulic transmission and control system[4]. Nevertheless, uncertain manufacturing and assembling errors, complex structure, and extreme working conditions make the axial piston pump prone to various fault patterns, posing great risks to its reliability and safety in operation[5]. Therefore, multi-fault diagnosis of the axial piston pump is significant for both the reliability-oriented design guidance and the service safety enhancement.

Fault diagnosis techniques for the axial piston pump have witnessed tremendous progress since deep learning enabled data-driven fault pattern recognition from an end-to-end perspective[6-8]. These intelligent fault diagnosis methods typically rely on building a historical dataset and training a class-fixed fault diagnosis model for deployment. However, in practical engineering applications, it is infeasible to gather a comprehensive dataset that contains all possible fault patterns in advance due to the complex structure of the axial piston pump[9]. Consequently, such a class-fixed fault diagnosis model fails to generalize if new fault patterns arise over time. While combining the newly collected data with the whole historical dataset and retraining the fault diagnosis model is theoretically viable, it is time-consuming, resource-demanding, and potentially constrained by data privacy. Conversely, fine-tuning the fault diagnosis model solely with new data leads to

catastrophic forgetting, dramatically degrading the fault diagnosis performance on the previously learned fault patterns[10]. Therefore, it is necessary to develop an effective multi-fault diagnosis method that preserves knowledge of old fault patterns while continuously integrating new fault patterns.

To this end, class-incremental learning, a framework ensuring compatibility with both the old and the new fault patterns, presents a promising solution. Generally, existing incremental learning algorithms can be categorized into architecture-based[11, 12], replay-based[13, 14], parameter regularization-based[15], and knowledge distillation-based methods[16, 17]. Among them, the knowledge distillation paradigm treats the previous fault diagnosis model as the teacher model and initializes a new one as the student model, mitigating catastrophic forgetting by transferring knowledge from the teacher to the student[18]. With minimal or even zero historical data, knowledge distillation-based methods outperform the parameter regularization-based methods on the previously learned fault patterns[19]. Learning without Forgetting (LwF)[16] records the teacher model's output logits on new data as the augmented labels, and minimizes the discrepancy between the student model's output logits and the augmented labels without the engagement of historical data. Incremental Classifier and Representation Learning (iCaRL)[17] also follows a similar idea but selects a fixed amount of prioritized examples from the historical dataset to further retain previous knowledge. Based on iCaRL, a Forward-Back Compatible Representation (FBCR)[20] reserves sufficient embedding space for new patterns to avoid invading previous knowledge. Though these knowledge distillation-based methods alleviate the catastrophic forgetting to some extent through output-layer logits alignment, they overlook the preservation of the intermediate-layer representations that encode discriminative fault features, thereby resulting in suboptimal fault diagnosis performance on old fault patterns.

Therefore, a novel class-incremental learning method based on knowledge distillation is proposed for the multi-fault diagnosis of the axial piston pump. Firstly, a limited number of exemplars are randomly sampled from the historical dataset, and new data are progressively added to form the incremental dataset. Secondly, to tackle the problem of insufficient preservation of old fault patterns, a knowledge distillation loss considering both the output-layer logits and the intermediate-layer representations is specifically designed. Finally, by minimizing the designed loss between the teacher model and the student model, the proposed method achieves good fault diagnosis performance on both the old and new fault patterns.

2 Methodology

The axial piston pump, mainly composed of a shaft, a cylinder block, a valve plate, a retainer, slipper-piston components, etc., is susceptible to multiple fault patterns. In real-world scenarios, the initial training dataset typically covers only a small number of fault patterns. With the operation time going on, new fault patterns that are not presented in the initial training dataset emerge. To learn these new fault patterns, the fault diagnosis model is trained incrementally. The overall training process is split into T+1 stages. Specifically, *Stage* 0 is the initial training for a base model, and *Stage* 1 ~ *Stage T* is the incremental learning to integrate new fault patterns that occur at each stage, as shown in Figure. 1.

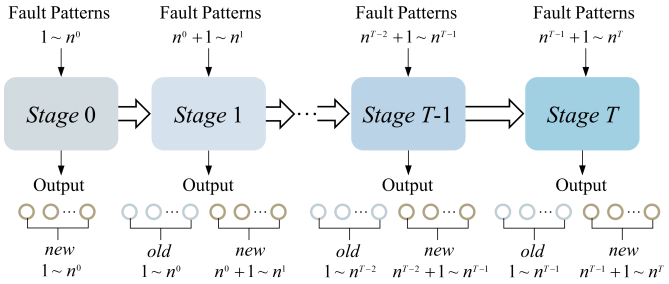


Figure 1: Overall training process for multi-fault diagnosis

2.1 Dataset construction

Ideally, the training dataset $\mathcal{D}_{\text{ideal}}^t$ for the incremental learning *Stage t* (t = 1, 2, …, T) is

$$\mathcal{D}_{\text{ideal}}^t = \{\underbrace{\mathcal{X}^0, \mathcal{X}^1, ..., \mathcal{X}^{t-1}}_{\mathcal{H}_{\text{ideal}}^t}, \mathcal{X}^t\} \tag{1}$$

$$\mathcal{X}^j = \left\{\left(x_i^j, y_i^j\right)\right\}_{i=1}^{N^j}, y_i^j \in \mathcal{Y}^j, j = 0,1,...,T \tag{2}$$

where $\mathcal{X}^0$ is the training dataset for the initial training, $\mathcal{X}^1 \sim \mathcal{X}^T$ is the newly collected dataset from the incremental *Stage* 1 to *Stage T*, respectively, x_i^j and y_i^j are training samples and true labels, respectively, $\mathcal{Y}^j$ is the label space, if $j \neq j'$, then $\mathcal{Y}^j \cap \mathcal{Y}^{j'} = \varnothing$, meaning that new fault patterns introduced to each stage are distinct and non-repetitive, N^j is the number of training samples contained in $\mathcal{X}^j$. N^j is enough large so that the fault diagnosis model has adequate historical and new training samples to learn both the old and the new fault patterns. For the incremental learning *Stage t*, $\mathcal{X}^0 \sim \mathcal{X}^{t-1}$ forms the ideal historical dataset $\mathcal{H}_{\text{ideal}}^t$, and $\mathcal{X}^t$ is the newly collected dataset only containing new fault patterns emerging at *Stage t*.

However, due to limited storage resources and data privacy, not all the historical data can be recorded in practice. Therefore, the real-world training dataset $\mathcal{D}_{\text{real}}^t$ at the *Stage t* is revised as

$$\mathcal{D}_{\text{real}}^t = \mathcal{H}_{\text{real}}^t \cup \mathcal{X}^t \tag{3}$$

$$\mathcal{H}_{\text{real}}^t = \mathcal{H}_{\text{real}}^{t-1} \cup \left(\mathcal{X}_{\text{sub}}^{t-1} \sim \mathcal{P}\left(\mathcal{X}^{t-1}, \alpha\right)\right), \mathbb{E}_{\mathcal{P}}\left(\left|\mathcal{X}_{\text{sub}}^{t-1}\right|\right) = \alpha N^{t-1} \tag{4}$$

where $\mathcal{H}_{\text{real}}^0 = \varnothing$, $\alpha \in [0,1)$ is a storage ratio on the historical dataset. Actually, α is a negligible value close to or equal to 0, and $\alpha = 0$ indicates that the fault diagnosis model at the *Stage t* is trained solely on newly collected data. According to (3) and (4), a small subset of exemplars $\mathcal{X}_{\text{sub}}^{t-1}$ is randomly selected from the historical dataset in a storage ratio of α while newly collected data is continuously incorporated to construct the real-world training dataset for the *Stage t*.

2.2 Model architecture

Denote $\mathcal{M}^t$ as the fault diagnosis model consisting of a feature extractor $\mathcal{F}^t$ with L intermediate layers $\{f_i^t; i = 1,2,...,L\}$ and a classifier $\mathcal{C}^t$ at the incremental *Stage t*[21]

$$\boldsymbol{h}_i^t = f_i^t \circ \cdots \circ f_2^t \circ f_1^t(x) \tag{5}$$

$$\hat{\boldsymbol{y}}^t = \mathcal{M}^t(x) = \mathcal{C}^t\left(\mathcal{F}^t(x)\right) = \mathcal{C}^t\left(h_L^t\right) \tag{6}$$

where $\circ$ denotes function composition, $\boldsymbol{h}_i^t \in \mathbb{R}^{d_i}$ is the i-th intermediate-layer representations, $\hat{\boldsymbol{y}}^t \in \mathbb{R}^{n^t}$ is the output-layer logits, n^t is the number of the observed fault patterns until the *Stage t*.

In this article, a one-dimensional convolution neural network (1D-CNN), mainly composed of convolution modules (Conv), batch normalization, activation function, and a two-layer fully connected network (FCN), is employed as the fault diagnosis model. As shown in Figure. 2, all the convolution modules form the feature extractor $\mathcal{F}^t$, and the two-layer FCN forms the classifier $\mathcal{C}^t$. Notably, the detailed architecture of the feature extractor, i.e., the number and the size of the convolution kernels, the convolution strides, and the paddings, remains the same during the incremental learning process. The number of nodes in the input layer of the classifier is unchanged as well, while nodes for new fault patterns are progressively added to the output layer.

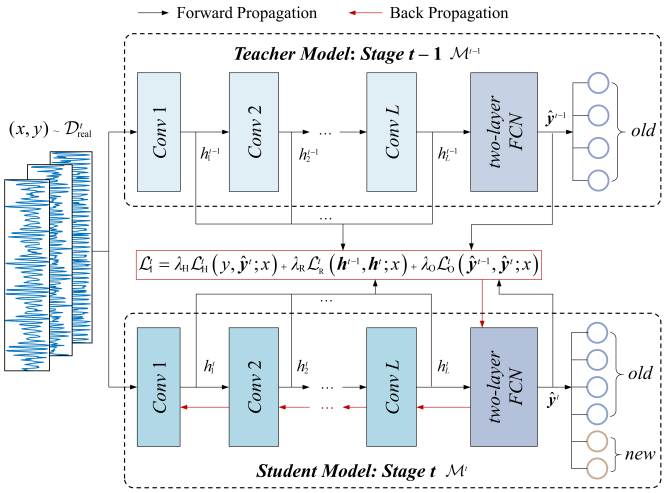


Figure 2: Class-incremental learning based on knowledge distillation for multi-fault diagnosis

2.3 Knowledge distillation loss design

At the incremental learning *Stage t*, the fault diagnosis model $\mathcal{M}^t$ is expected to learn new fault patterns and maintain the capability to diagnose fault patterns learned from *Stage* 0 to *Stage* t-1. Generally, given the training dataset $\mathcal{D}_{\text{real}}^t$, the common loss is designed as

$$\mathcal{L}_H^t(y,\hat{\boldsymbol{y}}^t;x) = -\sum_{i=1}^{n^t} \mathcal{I}(y)_i \log \frac{\exp(\hat{y}_i^t)}{\sum_j \exp(\hat{y}_j^t)} \tag{7}$$

$$\mathcal{I}(y) = \left[\mathcal{I}(y)_1, \mathcal{I}(y)_2, ..., \mathcal{I}(y)_i, ..., \mathcal{I}(y)_{n^t}\right] \tag{8}$$

$$\mathcal{I}(y)_i = \begin{cases} 1, i = y \\ 0, i \neq y \end{cases}, i = 1,2,..,n^t \tag{9}$$

where $\mathcal{I}(\cdot)$ maps the true label y to a one-hot vector, $\mathcal{I}(y)_i$ is the i-th item in the one-hot vector, and $\hat{y}_i^t$ is the i-th logit in $\hat{\boldsymbol{y}}^t$.

Simply optimizing (7) enables $\mathcal{M}^t$ to achieve good fault diagnosis performance on new fault patterns but cannot guarantee comparable performance on old fault patterns, probably leading to catastrophic forgetting. Following the teacher-student paradigm of knowledge distillation proposed by Hinton[18], the fault diagnosis model $\mathcal{M}^{t-1}$ which has been well-trained at *Stage t*-1, serves as the teacher model, and $\mathcal{M}^t$ established by adding new nodes at the output layer of the teacher model, serves as the student model. As shown in Figure. 2, the knowledge of the teacher model, i.e., the diagnostic capability for old fault patterns, is transferred to the student model by aligning the output-layer logits and the intermediate-layer representations accordingly.

The loss based on output-layer logits is designed as

$$\mathcal{L}_O^t(\hat{\boldsymbol{y}}^{t-1},\hat{\boldsymbol{y}}^t;x) = \sum_{i=1}^{n^{t-1}} \tilde{y}_i^{t-1} \log \frac{\tilde{y}_i^{t-1}}{\tilde{y}_i^t} \tag{10}$$

$$\tilde{y}_i^{t-1} = \frac{\exp(\hat{y}_i^{t-1}/\varepsilon)}{\sum_j \exp(\hat{y}_j^{t-1}/\varepsilon)}, \tilde{y}_i^t = \frac{\exp(\hat{y}_i^t/\varepsilon)}{\sum_j \exp(\hat{y}_j^t/\varepsilon)} \tag{11}$$

where ε is a temperature factor to adjust the softness of the output logits distribution.

To further preserve the fault diagnosis performance on previously learned fault patterns, the loss based on intermediate-layer representations that encode discriminative fault features is further introduced

$$\begin{aligned}\mathcal{L}_R^t(\boldsymbol{h}^{t-1},\boldsymbol{h}^t;x) &= \sum_{i=1}^{L} \omega_i \left(1 - \cos\left(\boldsymbol{h}_i^{t-1},\boldsymbol{h}_i^t\right)\right) \\ &= \sum_{i=1}^{L} \omega_i \left(1 - \frac{\boldsymbol{h}_i^{t-1}\cdot\boldsymbol{h}_i^t}{\left\|\boldsymbol{h}_i^{t-1}\right\|_2 \left\|\boldsymbol{h}_i^{t-1}\right\|_2}\right)\end{aligned} \tag{12}$$

where cos(·) denotes cosine similarity, ω_i is a weight to adjust the importance of the i-th intermediate-layer representations, and $\omega_i \propto i$, i.e., representations encoding higher-level features in the deeper intermediate layer are attached more importance.

Therefore, the total knowledge distillation loss is designed as

$$\mathcal{L}_1^t(y,\hat{\boldsymbol{y}}^{t-1},\boldsymbol{h}^{t-1},\hat{\boldsymbol{y}}^t,\boldsymbol{h}^t;x) = \lambda_H \mathcal{L}_H^t + \lambda_R \mathcal{L}_R^t + \lambda_O \mathcal{L}_O^t, (x,y) \sim \mathcal{D}_{\text{real}}^t \tag{13}$$

where λ_H, λ_R, and λ_O are hyper-parameters that balance the contributions of the three types of loss. By optimizing (13) with appropriate hyper-parameters, the fault diagnosis model can effectively retain diagnostic performance on old fault patterns while learning new fault patterns.

2.4 Class-incremental learning framework

The overall class-incremental learning framework includes initial training at *Stage* 0 and class-incremental learning from *Stage* 1 to *Stage T*, detailed in Algorithm 1.

At *Stage* 0, the fault diagnosis model $\mathcal{M}^0$ is optimized according to (14) and (15).

$$\mathcal{L}_H^0(y,\hat{\boldsymbol{y}}^0;x) = -\sum_{i=1}^{n^0} \mathcal{I}(y)_i \log \frac{\exp(\hat{y}_i^0)}{\sum_j \exp(\hat{y}_j^0)}, (x,y) \sim \mathcal{X}^0 \tag{14}$$

$$\arg\min_{\mathcal{M}^0} \mathcal{L}_{\mathrm{H}}^0\left(y, \hat{\boldsymbol{y}}^0; x\right) \tag{15}$$

From *Stage* 1 to *Stage* T, the fault diagnosis model $\mathcal{M}^t$ (t = 1, 2, …, T) is optimized according to (7) ~ (13) and (16).

$$\arg\min_{\mathcal{M}^t} \mathcal{L}_{\mathrm{I}}^t\left(y, \hat{\boldsymbol{y}}^{t-1}, \boldsymbol{h}^{t-1}, \hat{\boldsymbol{y}}^t, \boldsymbol{h}^t; x\right), t = 1, 2, ..., T \tag{16}$$

Algorithm 1 Class-incremental learning method based on knowledge distillation

Input: Initial training dataset $\{x_i^0, y_i^0\}_{i=1}^{N^0} \sim \mathcal{X}^0$, incremental stage T, storage ratio α, the number of intermediate layers L, weights $\omega_i\left(i = 1, 2, ..., L\right)$, temperature factor ε, hyper-parameters λ_{H}, λ_{R}, and λ_{O}, initial training epoch E_0, incremental learning epoch E_1, batch size B

Initialize $\mathcal{M}^0$ according to (5) and (6)
Define $N = N^0$
For e = 0, 1, …, E_0
  For b = 0, 1, …, N/B-1
    Calculate $\mathcal{L}_{\mathrm{H}}^0$ according to (6) and (14)
    Optimize $\mathcal{M}^0$ according to (15)
  End
End
For t = 1, 2, …, T
  Initialize $\mathcal{M}^t$ according to (5) and (6)
  Collect $\mathcal{X}^t$ and form $\mathcal{D}_{\mathrm{real}}^t$ according to (3) and (4)
  Update $N = \left|\mathcal{D}_{\mathrm{real}}^t\right|$
  For e = 0, 1, …, E_1
    For b = 0, 1, …, N/B-1
      Calculate $\mathcal{L}_{\mathrm{I}}^t$ according to (7) to (13)
      Optimize $\mathcal{M}^t$ according to (16)
    End
  End
End

Output: well-train $\mathcal{M}^T$

3 Experiment and Result Analysis

3.1 Experimental setting and fault injection

The test rig is shown in Figure. 3. Specifically, the tested axial piston pump is HP5V45 with 9 slipper-piston components, operating under a rated displacement of 45 cm^3/r, a rated rotation speed of 2700 rpm, and a rated discharge port pressure of 32 MPa. To monitor periodic vibrations of the tested axial piston pump, a triaxial integrated electronic piezoelectric accelerometer is mounted at the center of its backend cover. The sampling frequency is 20 kHz and every 1024 consecutive sampling points are considered as a training sample.

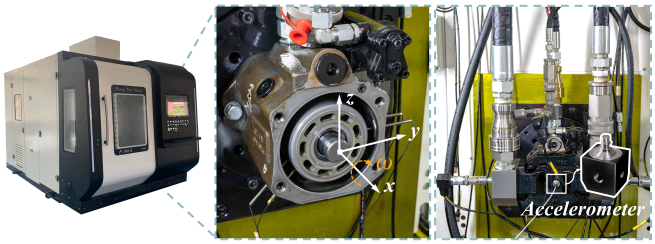


Figure 3: Test rig

In addition to the normal state, four types of fault patterns shown in Figure. 4, including main shaft spline spalling, retainer crack, valve plate scoring, and slipper-piston pair worn, are injected into the tested axial piston pump by replacing the normal components, respectively. Their time-domain vibration signals are depicted in Figure. 5. The fault injection sequence and dataset description for the proposed method are detailed in Table. 1 and Table. 2, respectively.

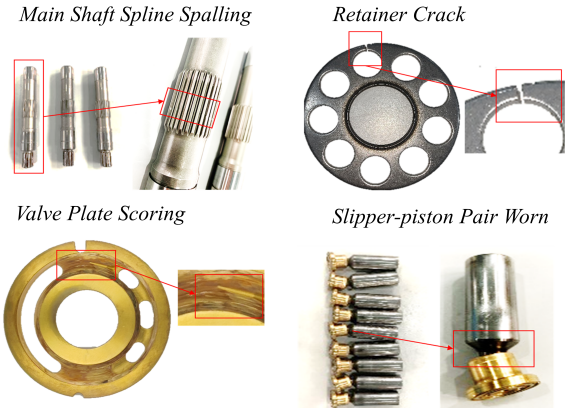


Figure 4: Faulty components

Table 1: Fault injection sequence

Stage	Injected Patterns	Labels
0	Normal State	1
	Main Shaft Spline Spalling	2
1	Retainer Crack	3
2	Valve Plate Scoring	4
3	Slipper-piston Pair Worn	5

Note: During the initial training and incremental learning process, labels are mapped into one-hot vectors according to (8) and (9).

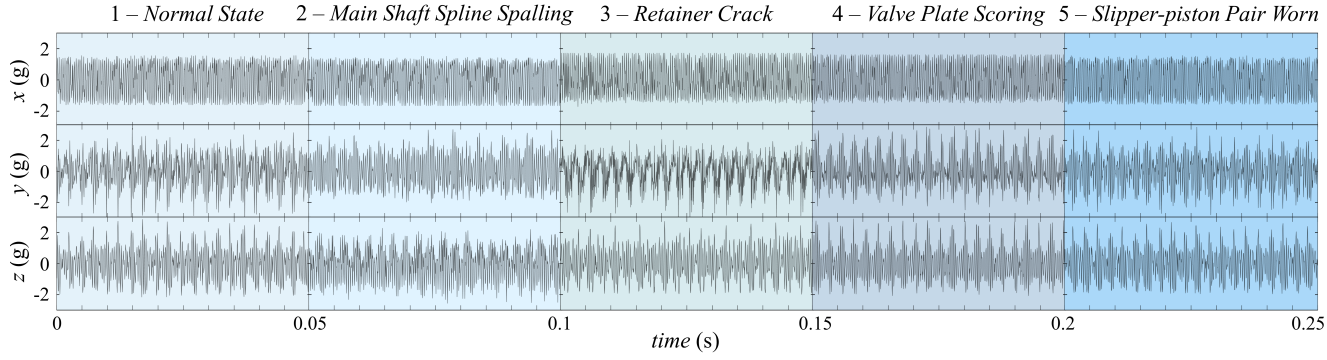


Figure 5: Time-domain vibration signals

Table 2: Dataset description for the proposed method

Stage	Number of Training Samples				Number of Testing Samples	
	α = 0.005		α = 0.01			
	New	Old	New	Old	New	Old
0	1728	0	1728	0	192	0
1	864	8	864	17	96	192
2	864	12	864	25	96	288
3	864	16	864	33	96	384

Table 3: Detailed architecture of base model

Layer	Kernel Size/Strides	Kernel Number	Input	Output	Padding
Conv 1	64×1/4×1	4	3×1024	4×256	31
Conv 2	4×1/2×1	4	4×256	4×128	1
Conv 3	4×1/2×1	4	4×128	4×64	1
Conv 4	4×1/2×1	4	4×64	4×32	1
FCN	/	/	128	2	/

3.2 Comparison experiment with other methods

The detailed architecture of the base model, i.e., the fault diagnosis model at *Stage* 0, is illustrated in Table. 3. Model parameters in this architecture follow the established 1D-CNN designed for time series input[22]. The input size of *Conv 1* is determined by the number of channels and the length of the input sample. The output size of *FCN* is determined by the number of fault patterns at the current stage. The activation function is LeakyReLU. As mentioned in Section 2.2, the architecture of *Conv* 1 ~ *Conv* 4 remains the same from *Stage* 0 to *Stage* 3, while the number of output nodes in *FCN* increases progressively according to Table. 1.

At *Stage* 0, the optimizer is SGD with a momentum of 0.9 and a weight decay of 10^{-4}. The learning rate is 10^{-3}, the max training epoch is 20, and the batch size is 48. From *Stage* 1 to *Stage* 3, the proposed method is compared with LwF[16] and fine-tuning. The optimizer of the proposed method is Adam, the batch size is 48, the learning rate is 10^{-2}, the max training epoch is 150, the temperature factor is 10, $[\omega_1, \omega_2, \omega_3, \omega_4] = [0.05, 0.05, 0.2, 0.7]$ and $[\lambda_H, \lambda_R, \lambda_O] = [0.5, 1, 0.5]$. For LwF, the optimizer is Adam, the batch size is 48, the warm-up rate is 10^{-6} and the warm-up epoch is 400, the learning rate is 10^{-4}, and the max training epoch is 200. For fine-tuning, the optimizer is Adam, the batch size is 48, the learning rate is 10^{-2}, and the max training epoch is 150. These training settings are manually configured to ensure that all the methods are trained to converge and reach a training accuracy of 100% for fair comparison.

To evaluate the effectiveness of each method, three performance metrics are defined as follows

$$\mathcal{A}_{\text{old}}^{i} = \frac{C_{\text{old}}^{i}}{N_{\text{old}}^{i}} \times 100\%, i = 1, 2, 3 \tag{17}$$

$$\mathcal{A}_{\text{new}}^{i} = \frac{C_{\text{new}}^{i}}{N_{\text{new}}^{i}} \times 100\%, i = 0, 1, 2, 3 \tag{18}$$

$$\mathcal{A}_{\text{avg}}^{i} = \frac{C_{\text{new}}^{i} + C_{\text{old}}^{i}}{N_{\text{new}}^{i} + N_{\text{old}}^{i}} \times 100\%, i = 0, 1, 2, 3 \tag{19}$$

where C_{old}^{i} and C_{new}^{i} are the number of correctly diagnosed testing samples of old fault patterns and new fault patterns at *Stage i*, respectively, N_{old}^{i} and N_{new}^{i} are the number of testing samples of old fault patterns and new fault patterns at *Stage i*, respectively, $\mathcal{A}_{\text{old}}^{i}$ indicates the fault diagnosis performance on old fault patterns at *Stage i*, $\mathcal{A}_{\text{new}}^{i}$ indicates the fault diagnosis performance on new fault patterns at *Stage i*, and $\mathcal{A}_{\text{avg}}^{i}$ indicates the comprehensive fault diagnosis performance on all fault patterns at *Stage i*. All the methods are trained to converge. The training process is repeated 5 times, and the average performance metrics are considered as the fault diagnosis performance. The comparison results are listed in Table. 4.

Table 4: Comparison experiment results (%)

Stage		Proposed Method (α = 0.005)	Proposed Method (α = 0.01)	LwF	Fine-tuning
0	$\mathcal{A}_{\text{new}}^{0}$	100	**100**	100	100
	$\mathcal{A}_{\text{avg}}^{0}$	100	**100**	100	100
1	$\mathcal{A}_{\text{old}}^{1}$	55.10±8.28	**90.52±11.57**	12.08±11.30	0
	$\mathcal{A}_{\text{new}}^{1}$	100	**100**	99.58±0.83	100
	$\mathcal{A}_{\text{avg}}^{1}$	70.01±5.52	**93.08±7.71**	41.25±7.58	33.33
2	$\mathcal{A}_{\text{old}}^{2}$	37.01±14.04	**69.44±11.02**	7.57±3.52	0
	$\mathcal{A}_{\text{new}}^{2}$	100	**100**	97.08±1.38	100
	$\mathcal{A}_{\text{avg}}^{2}$	52.76±10.53	**77.08±8.27**	29.95±2.67	25
3	$\mathcal{A}_{\text{old}}^{3}$	37.86±7.03	**69.06±4.25**	6.56±5.53	0
	$\mathcal{A}_{\text{new}}^{3}$	100	**100**	97.08±1.21	100
	$\mathcal{A}_{\text{avg}}^{3}$	50.29±5.62	**75.25±3.40**	24.67±4.53	20

For new fault patterns, all the methods demonstrate superior fault diagnosis performance at each stage, where the accuracy of LwF is slightly lower than 100%, and that of the proposed method and fine-tuning is 100%.

In contrast, for old fault patterns, all the methods suffer from different degrees of decline in accuracy, as shown in the Figure. 6. Particularly, the accuracy of fine-tuning drops directly to 0% at each stage, indicating its poor ability in maintaining fault diagnosis performance on old fault patterns. LwF outperforms fine-tuning, with an accuracy of 6.56±5.53% on old fault patterns and 24.67±4.53% on all fault patterns at *Stage* 3. Given a small storage ratio, the proposed method surpasses LwF and fine-tuning. As the storage ratio increases from 0.005 to 0.01, the accuracy of the proposed method on old fault patterns and all fault patterns increases by 31.20% and 24.96%, respectively, at *Stage* 3. This highlights the significance of leveraging historical data in the incremental training process.

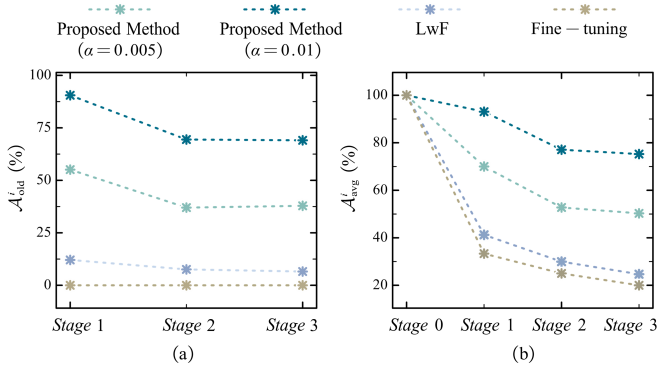


Figure 6: Comparison regarding $\mathcal{A}^i_{\text{old}}$ and $\mathcal{A}^i_{\text{avg}}$

Furthermore, the recalls are calculated at *Stage* 3 to analyze the degradation of the fault diagnosis performance for each fault pattern

$$\mathcal{R}^i = \frac{C^i}{N^i} \times 100\%, i = 1,2,3,4,5 \tag{20}$$

where N^1, N^2, N^3, N^4, and N^5 are the number of testing samples of the normal state, main shaft spline spalling, retainer crack, valve plate scoring, and slipper-piston pair worn at *Stage* 3, respectively, C^1, C^2, C^3, C^4, and C^5 are the correctly diagnosed testing samples of the normal state, main shaft spline spalling, retainer crack, valve plate scoring, and slipper-piston pair worn at *Stage* 3, respectively, $\mathcal{R}^1$, $\mathcal{R}^2$, $\mathcal{R}^3$, $\mathcal{R}^4$, and $\mathcal{R}^5$ are the recall of the normal state, main shaft spline spalling, retainer crack, valve plate scoring, and slipper-piston pair worn at *Stage* 3, respectively. $\mathcal{R}^i$ indicates fault diagnosis performance on each fault pattern at *Stage* 3. The average recalls at *Stage* 3 are listed in Table. 5 and visualized in Figure. 7. The results generally demonstrate a common trend among all the methods, i.e., as the incremental learning progresses, fault diagnosis performance degrades more severely on older fault patterns compared to new ones, suggesting the catastrophic forgetting phenomenon. However, two exceptions are observed in the proposed method, i.e., $\mathcal{R}^2$ exceeds $\mathcal{R}^3$ and $\mathcal{R}^4$ at a storage ratio of 0.005, and $\mathcal{R}^4$ falls below both $\mathcal{R}^2$ and $\mathcal{R}^3$ at a storage ratio of 0.01.

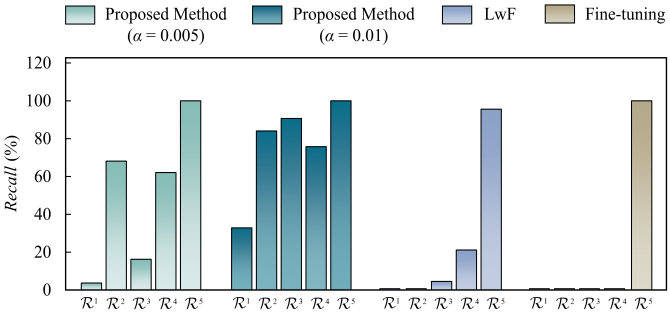


Figure 7: Recalls of injected patterns at Stage 3

3.3 Ablation experiment of the proposed method

To verify the contribution of each type of loss in (13), the ablation experiment is conducted. Specifically, Ablation 1 sets $[\lambda_{\text{H}}, \lambda_{\text{R}}, \lambda_{\text{O}}] = [1,0,0]$ and Ablation 2 sets $[\lambda_{\text{H}}, \lambda_{\text{R}}, \lambda_{\text{O}}] = [0.5,0,0.5]$. Other training settings are consistent with those described in Section 3.2. The ablation experiment results are listed in Table. 5, and results regarding $\mathcal{A}^i_{\text{old}}$ and $\mathcal{A}^i_{\text{avg}}$ are drawn in Figure. 8.

Similarly, all the methods achieve an accuracy of 100% on new fault patterns and generally degrade on old fault patterns during the incremental training process. At a storage ratio of 0.005, ablation 2 and the proposed method always outperform ablation 1, indicating that aligning the output-logits and intermediate-representations enhances the capability to preserve fault diagnosis performance on old fault patterns. According to Figure. 8(a) and 8(c), the proposed method performs worse than Ablation 2 on old fault patterns at *Stage* 2. However, from *Stage* 2 to *Stage* 3, the accuracy of Ablation 2 on old fault patterns decreases by 9.86% and 14.97%, at a storage ratio of 0.01 and 0.005, respectively. While that of the proposed method only decreases by 0.38% at a storage ratio of 0.01, and even increases by 0.85% at a storage ratio of 0.005. This indicates that, compared to relying solely on output-layer logits, joint alignment of both output-layer logits and the intermediate-layer representations further improves the ability to retain fault diagnosis performance on old fault patterns.

Table 5: Recalls of injected patterns at Stage 3 (%)

Recall	Proposed Method (α = 0.005)	Proposed Method (α = 0.01)	LwF	Fine-tuning
$\mathcal{R}^1$	3.62	32.79	0	0
$\mathcal{R}^2$	68.04	84	0	0
$\mathcal{R}^3$	16.17	90.67	4.46	0
$\mathcal{R}^4$	62.04	75.71	21.13	0
$\mathcal{R}^5$	100	100	95.50	100

Table 6: Ablation experiment results (%)

Stage		$\alpha = 0.005$			$\alpha = 0.01$		
		Ablation 1	**Ablation 2**	**Proposed Method**	**Ablation 1**	**Ablation 2**	**Proposed Method**
0	$\mathcal{A}^0_{new}$	100	100	**100**	100	100	**100**
	$\mathcal{A}^0_{avg}$	100	100	**100**	100	100	**100**
1	$\mathcal{A}^1_{old}$	35.42±21.49	46.98±10.62	**55.10±8.28**	71.25±16.25	81.87±13.78	**90.52±11.57**
	$\mathcal{A}^1_{new}$	100	100	100	100	100	**100**
	$\mathcal{A}^1_{avg}$	56.94±14.33	64.65±7.08	**70.01±5.52**	80.83±10.84	87.92±9.18	**93.08±7.71**
2	$\mathcal{A}^2_{old}$	28.40±11.48	**46.74±8.35**	37.01±14.04	69.72±8.51	**75.90±9.75**	69.44±11.02
	$\mathcal{A}^2_{new}$	100	100	**100**	100	100	**100**
	$\mathcal{A}^2_{avg}$	46.30±8.61	**60.05±6.26**	52.76±10.53	77.29±6.39	**81.93±7.31**	77.08±8.27
3	$\mathcal{A}^3_{old}$	22.76±8.97	31.77±4.29	**37.86±7.03**	65.10±3.40	66.04±6.78	**69.06±4.25**
	$\mathcal{A}^3_{new}$	100	100	**100**	100	100	**100**
	$\mathcal{A}^3_{avg}$	38.21±7.18	45.42±3.43	**50.29±5.62**	72.08±2.72	72.83±5.42	**75.25±3.40**

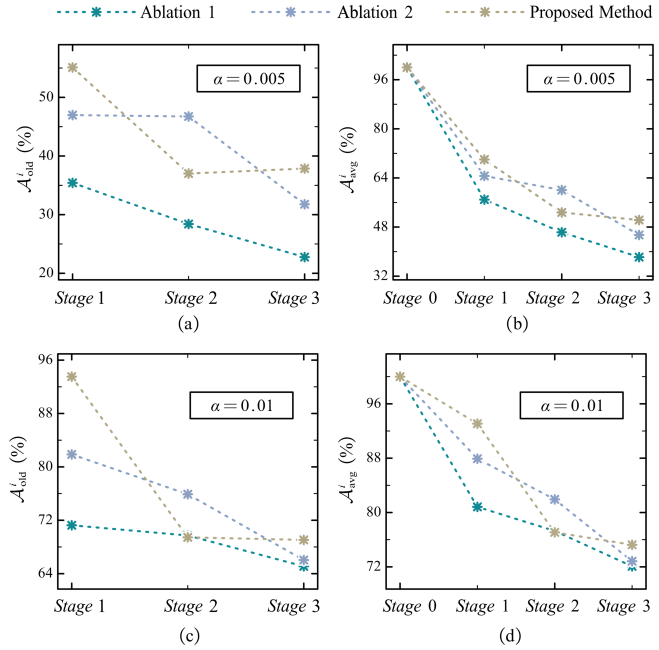


Figure 8: Ablation regarding $\mathcal{A}^i_{old}$ and $\mathcal{A}^i_{avg}$

4 Conclusions

For multi-fault diagnosis of the axial piston pump, a class-incremental learning method based on knowledge distillation is proposed to integrate newly emerging fault patterns as well as preserving fault diagnosis performance on old fault patterns. Comparison experiment with LwF and fine-tuning, and ablation experiment around each loss item are conducted to verify the effectiveness of the proposed method. Main conclusions are drawn as follows.

(1) Storing only a small portion of historical data can greatly improve the fault diagnosis performance on old fault patterns. At *Stage* 3, increasing the storage ratio from 0.005 to 0.01 boosts the accuracy of the proposed method on old fault patterns by 31.20%.

(2) Joint alignment of the output-layer logits and the intermediate-layer representations enhances knowledge preservation on old fault patterns. Adding the loss item regarding the output-layer logits boosts the accuracy of the proposed method on old fault patterns by 9.01% (α = 0.005) and 0.94% (α = 0.01), respectively. Further adding the loss item regarding the intermediate-layer representations boosts the accuracy of the proposed method on old fault patterns by 6.09% (α = 0.005) and 3.02% (α = 0.01), respectively.

There is still room for further discussions and enhancements.

(1) Theoretically, the proposed method can simultaneously learn several fault patterns at every incremental learning stage, but the case of concurrent faults remains to be empirically explored due to the limited types of fault patterns introduced in this article. More fault patterns of the axial piston pump need to be considered and collected in future work.

(2) Hyper-parameters in this article are given manually and may probably lead to a suboptimal performance in real-world scenarios when operating condition changes frequently. Therefore, how to set appropriate hyper-parameters adaptive to different pumps and operating conditions needs to be analyzed to achieve a more generalized fault diagnosis performance.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (U24B2049).

References

[1] S Mukherjee, L Shang, and A Vacca. Numerical analysis and experimental validation of the coupled thermal effects in swashplate type axial piston machines. *Mechanical Systems and Signal Processing*, 220:111673, 2024.

[2] P Michael, K Stelson, D Williams, and H Malik. Dynamometer testing of hydraulic fluids in an axial piston

pump under simulated backhoe loader trenching conditions. In *Fluid Power Systems Technology* (American Society of Mechanical Engineers), 86335:V001T01A014, 2022.

[3] L V Larsson and P Krus. Modelling of the swash plate control actuator in an axial piston pump for a hardware-in-the-loop simulation test rig. In *Fluid Power Systems Technology* (American Society of Mechanical Engineers), 50473:V001T01A044, 2016.

[4] I Baus, R Rahmfeld, A Schumacher, and H C Pedersen. Systematic methodology for reliability analysis of components in axial piston units. In *Fluid Power Systems Technology* (American Society of Mechanical Engineers), 59339:V001T01A009, 2019.

[5] R Ivantysyn and J Weber. Advancing thermal monitoring in axial piston pumps: Simulation, measurement, and boundary condition analysis for efficiency enhancement. *International Journal of Fluid Power*, 547-590, 2024.

[6] S Wang, J Xiang, Y Zhong, and H Tang. A data indicator-based deep belief networks to detect multiple faults in axial piston pumps. *Mechanical Systems and Signal Processing*, 112:154-170, 2018.

[7] Y He, H Tang, Y Ren, and A Kumar. A semi-supervised fault diagnosis method for axial piston pump bearings based on DCGAN. *Measurement Science and Technology*, 32(12):125104, 2021.

[8] S Wang et al. Few-shot fault diagnosis of axial piston pump based on prior knowledge-embedded meta learning vision transformer under variable operating conditions. *Expert Systems with Applications*, 269:126452, 2025.

[9] S Liu, J Zhang, W Huang, F Lyu, D Wang, and B Xu. Temporal–spatial attention network: A novel axial piston pump coupled fault diagnosis method. *IEEE Transactions on Instrumentation and Measurement*, 73:1-15, 2024.

[10] E Belouadah, A Popescu, and I Kanellos. A comprehensive study of class incremental learning algorithms for visual tasks. *Neural Networks*, 135:38-54, 2021.

[11] A A Rusu et al. Progressive neural networks. *arXiv preprint* arXiv:1606.04671, 2016.

[12] Y-X Wang, D Ramanan, and M Hebert. Growing a brain: Fine-tuning by increasing model capacity. Proceedings of *the IEEE Conference on Computer Vision and Pattern Recognition*, 2471-2480, 2017.

[13] R Aljundi, M Lin, B Goujaud, and Y Bengio. Gradient based sample selection for online continual learning. *Advances in Neural Information Processing Systems*, 32, 2019.

[14] A Odena, C Olah, and J Shlens. Conditional image synthesis with auxiliary classifier gans. *International Conference on Machine Learning*, 2642-2651, 2017.

[15] J Kirkpatrick et al. Overcoming catastrophic forgetting in neural networks. Proceedings of *the National Academy of Sciences*, 114(13):3521-3526, 2017.

[16] Z Li and D Hoiem. Learning without forgetting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(12):2935-2947, 2017.

[17] S-A Rebuffi, A Kolesnikov, G Sperl, and C H Lampert. icarl: Incremental classifier and representation learning. Proceedings of *the IEEE Conference on Computer Vision and Pattern Recognition*, 2001-2010, 2017.

[18] G Hinton, O Vinyals, and J Dean. Distilling the knowledge in a neural network. *arXiv preprint* arXiv:1503.02531, 2015.

[19] G M Van de Ven, T Tuytelaars, and A S Tolias. Three types of incremental learning. *Nature Machine Intelligence*, 4(12):1185-1197, 2022.

[20] S Yan, H Shao, X Wang, and J Wang. Few-shot class-incremental learning for system-level fault diagnosis of wind turbine. *IEEE/ASME Transactions on Mechatronics*, 2024.

[21] M Kang, J Park, and B Han. Class-incremental learning by knowledge distillation with adaptive feature consolidation. Proceedings of *the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 16071-16080, 2022.

[22] M. Ji, et al. A neural network compression method based on knowledge-distillation and parameter quantization for the bearing fault diagnosis. *Applied Soft Computing*, 127:109331, 2022.

Sensorless Position Estimation of Electro-Hydraulic Proportional Valve Using Dual Solenoids

Jiasheng Wang, Qi Su*, Junhui Zhang, Xukun Zhang, Jinheng Liu, and Zhi Qiu

State Key Laboratory of Fluid Power and Mechatronic Systems, Zhejiang University, Hangzhou, China
E-mail: zjuwjs@zju.edu.cn, suqi@zju.edu.cn, benzjh@zju.edu.cn, zhangxk987@163.com, ljh1640914541@163.com, qiuzhi@zju.edu.cn

Abstract

Solenoid-actuated valves used in construction machinery are typically not equipped with position sensors to reduce costs. This hinders the acquisition of spool position information during operation and adversely affects degradation prediction and fault diagnosis. This study presents a soft measurement method for spool position estimation in solenoid-actuated valves without position sensors. The technique utilizes the position information fusion of the two solenoids in the valve. First, a spool displacement model is derived based on the inductance variations of the working solenoid, and the solenoid simulation model is established using data from the solenoid tests and the measurement of the solenoid structure. Meanwhile, a high-frequency excitation signal is applied to the idle solenoid, and feedback from the circuit is utilized to measure changes in inductance, allowing for the derivation of the spool position. Subsequently, an improved adaptive Kalman filter algorithm is developed to fuse these two position estimates, establishing an adaptive soft measurement model for dynamic spool position estimation. Finally, simulation validation is conducted on a Simulink model based on the proposed method. The results indicate that the proposed position estimation method can achieve high-precision spool position information over the full spool stroke.

Keywords: Proportional valve, Soft measurement, Position estimation, Solenoid inductance, Kalman filter

1 Introduction

Solenoid-actuated valves are widely used in construction equipment, agricultural machinery, and mining machinery [1]. Therefore, these valves often operate under conditions of significant vibration and harsh working environments. Moreover, due to cost constraints, it is challenging to equip precise displacement sensors, e.g., linear variable displacement transducer (LVDT), in solenoid-actuated valves, which hinders the acquisition of valve spool position information [2]. Currently, these machines predominantly depend on the operator's experiential knowledge for their operation. For example, adjusting the position of components, such as the terminal section of excavating machinery, relies heavily on the operator's visual perception. This reliance substantially hinders the advancement of automation in construction and production processes, as well as the development of effective fault diagnosis methods. Therefore, the research on position soft measurement for solenoid-actuated valves is of great significance.

In recent years, soft measurement technology, an indirect measurement technique where a mathematical model or algorithm is used to estimate a process variable based on other available sensor data rather than measuring it directly, has been gradually applied to hydraulic actuators [3-5]. By exploiting the inherent correlations between internal parameters of controlled systems, the estimation of critical states by mathematical models or data-driven approaches can effectively reduce the dependence on traditional sensors. In electrohydraulic valve applications, proportional solenoids have been widely adopted due to their cost-effectiveness, high output force, and favorable linearity [6]. There is a clear functional relationship between the air gap length of the solenoid and its inductance value, which provides both the theoretical basis and the technical path for position sensing. Existing electromagnetic position soft sensing methods can be categorized into three classes. The first category targets low-speed applications, including inductance-based position estimation via pulse injection, current waveform monitoring for position inference, and methods utilizing switching signals from power electronics hardware [7]. Nagai et al [8] developed a low-cost position and force sensing strategy for compact solenoid actuators by superimposing an alternating current (AC) component on a direct current (DC) power supply. This enables simultaneous real-time estimation of position and force. Their subsequent work [9] implemented inductance-displacement soft sensing combined with disturbance observer

control for dual solenoids. Considering the ambiguities caused by hysteresis and saturation effects, Kramer et al [10] proposed a self-sensing design of the proportional solenoids, which balances the force generation ability and the sensor properties. The second category focuses on medium to high speed scenarios and typically incorporates observer-based position estimation and neural network approaches that establish mappings between position, magnetic flux, and current [11]. Zhao et al [12] achieved sensorless linear control of spool position in electromagnetic switching valves by integrating sliding mode controllers with nonlinear sliding mode observers. The third category employs auxiliary components for position estimation, predominantly in motor applications, such as additional stator coils for current variation detection, bootstrap circuits for initial mover positioning, eddy current sensors that correlate signals with position, and dedicated sensing units with auxiliary stator/rotor laminations [13]. Jing et al [14] realized gap self-inductance sensing in maglev systems using an auxiliary coil with high-frequency excitation, demonstrating enhanced sensitivity in suspension gap measurement. Verma et al [15] proposed a self-sensing electromagnetic actuator for vibration suppression in flexible structures, where structural displacement was evaluated by back electromotive force measurements in the coil.

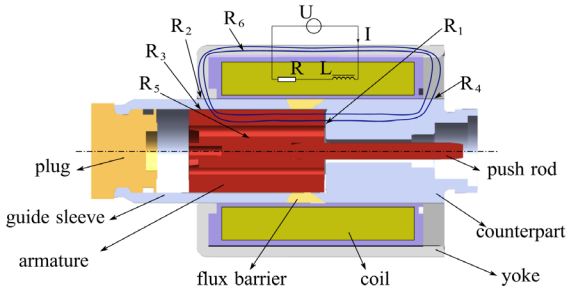


Figure 1: Structure of the electrohydraulic proportional solenoid

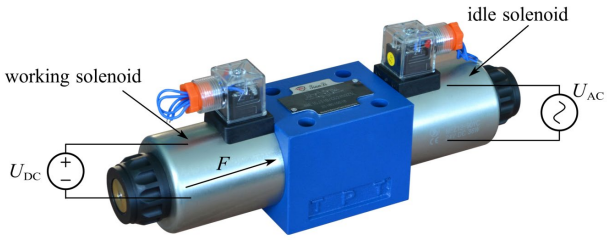


Figure 2: Working principle of dual solenoids in a valve

Based on the low speed and auxiliary component methodology, this study proposes a soft measurement method for spool position in electrohydraulic proportional valves with the collaborative operation of dual solenoids. Fig. 1 illustrates the structure of the proportional solenoid adopted in this paper. During directional valve operation, one working solenoid is typically energized to generate force while the other remains idle, as shown in fig. 2. Firstly, during high-current force generation, the working solenoid achieves self-sensing by exploiting the relationship between inherent inductance and air gap, enabling real-time position monitoring during actuation. Meanwhile, the idle solenoid will be used to serve as an auxiliary sensing component for position monitoring. Sinusoidal excitation signals are applied to the idle solenoid, and the air gap-inductance relationship is pre-calibrated. Displacement information is inversely derived from inductance values calculated under sinusoidal excitation. Finally, an adaptive Kalman filtering algorithm merges the two displacement estimates to obtain the final measurement.

The major innovations of this study are summarized as follows:

(1) Dual-solenoid position soft measurement method. A novel soft sensing strategy using coordinated dual-solenoid operation is proposed: The working solenoid simultaneously performs thrust generation and self-sensing functions, thus achieving integration of displacement measurement and actuation. The idle solenoid enables real-time spool displacement inversion through sinusoidal excitation signal injection and pre-calibrated inductance between air gap relationships. In particular, the initial position of the working solenoid is obtained directly through step excitation without the need for additional inductance and position calibration.

(2) An adaptive Kalman filter is designed to perform data fusion, which enhances measurement accuracy and noise immunity. The precision of the dual-solenoid method shows better results than the single-solenoid soft measurement methods. The position soft measurement method works with existing solenoid structures without additional components, ensuring full compatibility with standard electromagnetic valve operation.

The methodological framework of this study is organized as follows: Section 2 presents the proposed approach in detail. Section 3 describes the experimental data acquisition setup and the established simulation model for the validation simulations to be performed. Section 4 analyzes the results of the simulations and their practical implications. Section 5 summarizes the results of the study, discusses limitations, and suggests directions for future research. Finally, all symbols used in this paper are listed and defined in the Nomenclature section.

2 Principle of dual-solenoid position soft measurement

2.1 Position self-sensing method for working solenoid

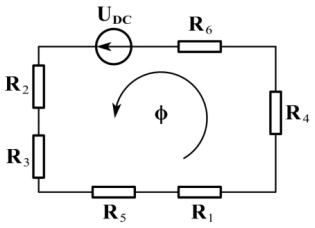


Figure 3: Magnetic circuit of the solenoid

The magnetic circuit characteristics of the working solenoid are shown in fig. 3, and the analysis of the reluctance positions is annotated in fig. 1. The relationship between magnetic reluctance and air gap dimensions is as follows:

$$R_1 = \frac{x_1}{\mu_0 A_s} \tag{1}$$

$$R_2 = \int_{r_{22}}^{r_{21}} \frac{dr}{2\pi r \mu_0 l_2} = \frac{\ln(r_{21}/r_{22})}{2\pi\mu_0 l_2} \tag{2}$$

$$R_3 = \int_{r_{32}}^{r_{31}} \frac{dr}{2\pi r \mu_0 l_3} = \frac{\ln(r_{31}/r_{32})}{2\pi\mu_0 l_3} \tag{3}$$

$$R_4 = \int_{r_{42}}^{r_{41}} \frac{dr}{2\pi r \mu_0 l_4} = \frac{\ln(r_{41}/r_{42})}{2\pi\mu_0 l_4} \tag{4}$$

Since the solenoid is a wet-type solenoid, the theoretical magnetic permeability u_0 is that of hydraulic oil. However, commonly used hydraulic oils are non-magnetic fluids, and their magnetic permeability is similar to that of air. In this study, $u_0 = 4\pi \times 10^{-7}$ H/m. Due to the significantly higher magnetic permeability of the metal components of the solenoid compared to air, the effects of the metal reluctances R_5 and R_6 can be neglected. The structure parameters of the solenoid satisfy $r_{21} = r_{41}$, $r_{22} = r_{42}$, and $l_2 = l_4$. Therefore, $R_2 = R_4$.

According to Kirchhoff's second law:

$$NI_w = \phi_1 (R_1 + R_2 + R_3 + R_4 + R_5 + R_6) \tag{5}$$

and the definition of self-inductance in electromagnetism:

$$\psi = N\phi_1 = L_w I_w \tag{6}$$

It follows:

$$\begin{aligned} L_w &= \frac{N\phi_1}{I_w} \\ &= \frac{N}{I_w} \times \frac{NI_w}{R_1 + R_2 + R_3 + R_4 + R_5 + R_6} \\ &\approx \frac{N^2}{R_1 + 2R_2 + R_3} \end{aligned} \tag{7}$$

During the operation of the solenoid, there are

$$U_{DC} = U_L + U_R \tag{8}$$

$$U_R = I_w R \tag{9}$$

$$U_L = U_{DC} - I_w R \tag{10}$$

$$\psi_0 = \int_0^{t_0} U_L(t)\,dt \tag{11}$$

Since the solenoid operates in the non-saturation region, combining eq. (6) and eq. (11), it follows:

$$L_0 = \frac{\psi_0}{i_0} \tag{12}$$

Combining eq. (7), it follows:

$$L_0 = \frac{N^2}{\frac{\delta_0}{\mu_0 A_s} + \frac{\ln(r_{21}/r_{22})}{\pi\mu_0 l_2} + \frac{\ln(r_{31}/r_{32})}{2\pi\mu_0 l_3}} = \frac{\psi_0}{i_0} \tag{13}$$

The initial position δ_0 can be calculated based on the given parameters.

$$\delta_0 = A_s \left(\frac{\mu_0 N^2 i_0}{\psi_0} - \frac{\ln(r_{21}/r_{22})}{\pi l_2} - \frac{\ln(r_{31}/r_{32})}{2\pi l_3} \right) \tag{14}$$

The magnetic flux $\Delta\psi$ of the solenoid is:

$$\Delta\psi = \int_{t_0}^{t_0+\Delta t} U_L(t)\,dt \tag{15}$$

From eq. (6), it follows:

$$d\psi = dLi = L\,di + i\,dL \tag{16}$$

Therefore, based on the iterative method for calculating the real-time inductance:

$$\begin{cases} L_1 = \dfrac{\Delta\psi - L_0 \Delta i}{i_0} + L_0 \\ \quad\quad ... \\ L_k = \dfrac{\Delta\psi - L_{k-1}\Delta i}{i_{k-1}} + L_{k-1} \end{cases} \tag{17}$$

and

$$L_k = \frac{N^2}{\frac{\delta_0 + \Delta x}{\mu_0 A_s} + \frac{\ln(r_{21}/r_{22})}{\pi\mu_0 l_2} + \frac{\ln(r_{31}/r_{32})}{2\pi\mu_0 l_3}} \tag{18}$$

By solving eq. (18), the iron core position x_1 is:

$$\begin{aligned} x_1 &= \delta_0 + \Delta x \\ &= A_s \left(\frac{\mu_0 N^2}{\frac{\Delta\psi - L_{k-1}\Delta i}{i_{k-1}} + L_{k-1}} - \frac{\ln(r_{21}/r_{22})}{\pi l_2} - \frac{\ln(r_{31}/r_{32})}{2\pi l_3} \right) \end{aligned} \tag{19}$$

2.2 Position estimation method for idle solenoid

When a sinusoidal voltage excitation is applied to the idle solenoid, the following relationship applies:

$$U_{AC} = \bar{U}_m \sin(\omega t) + U_k \tag{20}$$

The solenoid current will also have sinusoidal fluctuations:

$$I_{AC} = \bar{I}_m \sin(\omega t - \theta) + I_k \tag{21}$$

The solenoid voltage equation is:

$$\begin{aligned} U_{AC} &= I_{AC} R + \frac{d\phi_2}{dt} \\ &= I_{AC} R + I_{AC} \frac{dL_s}{dt} + L_s \frac{dI_{AC}}{dt} \end{aligned} \tag{22}$$

When the current is small, the inductance has a one-to-one correspondence with the position [10], it follows:

$$U_{AC} = I_{AC} R + L_s \frac{dI_{AC}}{dt} \tag{23}$$

By combining eq. (20), eq. (21), and eq. (23), it follows:

$$\begin{aligned} &\bar{U}_m \sin(\omega t) + U_k \\ &= \bar{I}_m \sqrt{R^2 + (\omega L_s)^2} \sin(\omega t - \theta + \theta') + I_k R \end{aligned} \tag{24}$$

where

$$\theta' \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right),\ \tan\theta' = \frac{\omega L_s}{R},\ U_k = I_k R$$

By equating the corresponding terms, it can be concluded that

$$\theta = \theta', \ \bar{U}_{\mathrm{m}} = \bar{I}_{\mathrm{m}} \sqrt{R^2 + (\omega L_{\mathrm{s}})^2}$$

Therefore, the L_{s} is as follows:

$$L_{\mathrm{s}} = \frac{1}{\omega} \sqrt{\left(\frac{\bar{U}_{\mathrm{m}}}{\bar{I}_{\mathrm{m}}}\right)^2 - R^2} \tag{25}$$

The relationship between core position and inductance is calibrated beforehand for the idle solenoid used in this paper, and a fitted relationship is established:

$$x_2 = aL_{\mathrm{s}}^4 + bL_{\mathrm{s}}^3 + cL_{\mathrm{s}}^2 + dL_{\mathrm{s}} + e \tag{26}$$

Based on this equation, the real-time position of the solenoid iron core can be obtained.

2.3 Adaptive Kalman filter algorithm

(1) The first step is to build the state model. Defining the state vector:

$$X_{\mathrm{k}} = \begin{bmatrix} x_{\mathrm{k}} \\ v_{\mathrm{k}} \end{bmatrix} \tag{27}$$

In a discrete-time system with sampling period T, the state transition model is:

$$x_{\mathrm{k}} = Ax_{\mathrm{k}-1} + w_{\mathrm{k}-1}, \ w_{\mathrm{k}-1} \sim N(0, Q_{\mathrm{k}}) \tag{28}$$

where the state transition matrix A is:

$$A = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \tag{29}$$

(2) The second step is to build the observation model. The system has two measurement signals: $z_{1,\mathrm{k}}$ (working solenoid position x_1 calculated from the solenoid model) and $z_{2,\mathrm{k}}$ (idle solenoid position x_2 derived from a sensor-like method). The observation matrix C is defined as:

$$C = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \tag{30}$$

The observation equation is:

$$z_{\mathrm{k}} = Cx_{\mathrm{k}} + r_{\mathrm{k}}, \ r_{\mathrm{k}} \sim N(0, R_{\mathrm{k}}) \tag{31}$$

The initial observation noise covariance is:

$$R_{\mathrm{base}} = \begin{bmatrix} r_{\mathrm{base}\,1} & 0 \\ 0 & r_{\mathrm{base}\,2} \end{bmatrix} \tag{32}$$

(3) Adaptive adjustment of noise covariance design. First, the observation noise covariance R_{k} is dynamically modified. Since the inductance L_{w} is calculated iteratively in the self-sensing model of the working solenoid, the position estimation $z_{1,\mathrm{k}}$ may become unreliable during abrupt position changes (e.g., step). To mitigate the impact of $z_{1,\mathrm{k}}$ inaccuracies, increase the confidence in $z_{2,\mathrm{k}}$. Adjust R_{k} based on the discrepancy between $z_{1,\mathrm{k}}$ and $z_{2,\mathrm{k}}$:

$$R_{\mathrm{k}} = \begin{cases} mR_{\mathrm{base}}, & \text{if } |z_{2,\mathrm{k}} - z_{1,\mathrm{k}}| > x_{\mathrm{Ru}}, \\ nR_{\mathrm{base}}, & \text{if } |z_{2,\mathrm{k}} - z_{1,\mathrm{k}}| < x_{\mathrm{Rl}}, \\ R_{\mathrm{base}}, & \text{otherwise.} \end{cases} \tag{33}$$

A dynamic adjustment for the process noise Q_{k} is then performed. Since the idle solenoid only plays an auxiliary role in position measurement during the process, it cannot be subjected to an excitation voltage that is too high to avoid generating a magnetic force that would affect the control of the solenoid valve. Therefore, the response current of the idle solenoid is generally small, within 30 mA. The current acquisition circuit has limited accuracy and is susceptible to external interference. It is necessary to process the noise of the idle solenoid position estimation. The noise is dynamically updated according to the deviation $\mathrm{d}x_2$, of x_2 from the current estimate:

$$Q_k = \begin{cases} \begin{bmatrix} q_{\mathrm{k}1} & 0 \\ 0 & q_{\mathrm{k}2} \end{bmatrix}, & \text{if } \mathrm{d}x_2 > x_{\mathrm{Qu}}, \\ \begin{bmatrix} q_{\mathrm{k}3} & 0 \\ 0 & q_{\mathrm{k}4} \end{bmatrix}, & \text{if } \mathrm{d}x_2 < x_{\mathrm{Ql}}, \\ \begin{bmatrix} q_{\mathrm{k}5} & 0 \\ 0 & q_{\mathrm{k}6} \end{bmatrix}, & \text{otherwise.} \end{cases} \tag{34}$$

When x_2 varies drastically, increasing Q_{k} allows the filter to track abrupt changes more quickly. When x_2 is smooth, decreasing Q_{k} improves smoothing.

(4) A fusion factor α is designed to fuse the data of x_1 and x_2 weightedly. Define the step error as

$$s_{\mathrm{err}} = |z_{1,\mathrm{k}} - \hat{x}_{\mathrm{k}-1}| \tag{35}$$

and the noise level as

$$n_{\mathrm{L}} = |z_{2,\mathrm{k}} - \hat{x}_{\mathrm{k}-1}| \tag{36}$$

Calculate the fusion factor there:

$$\alpha = 1 - \frac{n_{\mathrm{L}}}{n_{\mathrm{L}} + s_{\mathrm{err}} + \varepsilon} \tag{37}$$

If $z_{2,\mathrm{k}}$ has higher noise levels and $z_{1,\mathrm{k}}$ is more stable, α tends to 0, relying more on data from x_1. If $z_{2,\mathrm{k}}$ differs less from the current estimate (low noise) and $z_{1,\mathrm{k}}$ differs more from the current estimate, α tends to 1, relying mainly on measurements from x_2. A weighted fusion of the position components in the state:

$$\hat{x}_{\mathrm{k}} = \alpha \hat{x}_{\mathrm{k}} + (1 - \alpha) z_{1,\mathrm{k}} \tag{38}$$

(5) Synthesize the above design scheme to update the Kalman filter. Its prediction steps are:

$$\hat{x}_{\mathrm{k}|\mathrm{k}-1} = A\hat{x}_{\mathrm{k}-1|\mathrm{k}-1} \tag{39}$$

$$P_{\mathrm{k}|\mathrm{k}-1} = AP_{\mathrm{k}-1|\mathrm{k}-1}A^{\mathrm{T}} + Q_{\mathrm{k}} \tag{40}$$

The Kalman gain calculation in data updating is as follows:

$$K_{\mathrm{k}} = P_{\mathrm{k}|\mathrm{k}-1}C^{\mathrm{T}} \left(CP_{\mathrm{k}|\mathrm{k}-1}C^{\mathrm{T}} + R_{\mathrm{k}}\right)^{-1} \tag{41}$$

The status update will be as follows:

$$\hat{x}_{\mathrm{k}|\mathrm{k}} = \hat{x}_{\mathrm{k}|\mathrm{k}-1} + K_{\mathrm{k}} \left(z_{\mathrm{k}} - C\hat{x}_{\mathrm{k}|\mathrm{k}-1}\right) \tag{42}$$

and updating the error covariance as follows:

$$P_{\mathrm{k}|\mathrm{k}} = (I - K_{\mathrm{k}}C) P_{\mathrm{k}|\mathrm{k}-1} \tag{43}$$

3 Solenoid test and simulation model

3.1 Electromagnet Test Bench

Solenoid test calibration experiments use a self-made electromagnet test bench, mainly composed of the electric servo cylinder, servo control system, solenoid fixing block, solenoid controller, DC power supply, data acquisition device, and other components, shown in fig. 4. The solenoid fixed block can maintain the overall position of the solenoid unchanged. The servo control system controls the expansion and contraction of the electric servo cylinder, which in turn can control the position of the solenoid core. The DC power supply powers the proportional solenoid controller, which injects voltage signals into the solenoid, while a current sensor is integrated to obtain the current. The data is uniformly collected and saved by the data acquisition device. In this paper, after the calibration test of the proportional solenoid on the electrohydraulic proportional valve, the simulation model of the solenoid is established. Then, the effect of the dual-solenoid position soft measurement method is verified in the Simulink simulation of MATLAB software.

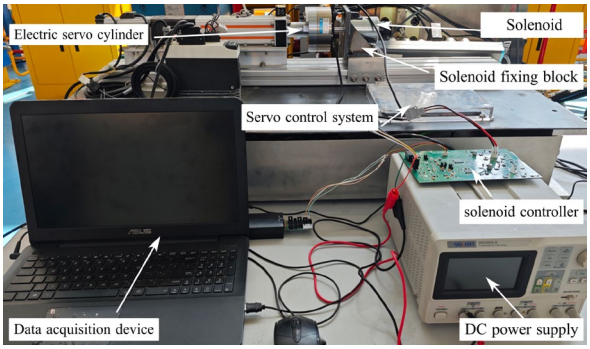


Figure 4: Electromagnet Test Bench

3.2 Solenoid characterization data acquisition and processing

3.2.1 Solenoid simulation model

In this paper, a solenoid simulation model is established, as shown in fig. 5, and the specific establishment method is described in detail in the previous research results [16-17]. The test temperature of the solenoid is room temperature. For the solenoid to apply voltage excitation to obtain data, it is quickly completed, so do not consider the solenoid operation process temperature changes on the resistance.

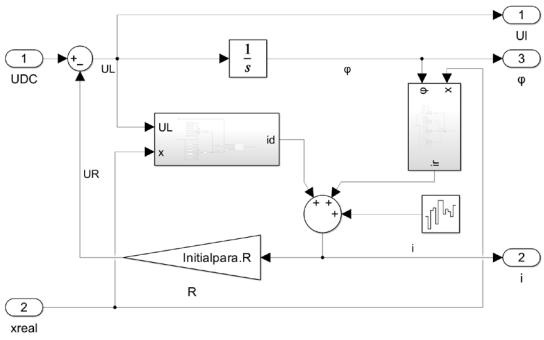


Figure 5: Solenoid simulation model

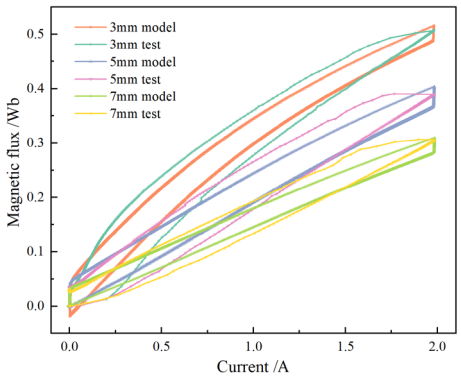


Figure 6: Comparison of solenoid test and simulation

Fig. 6 shows that when the 0-9 V step voltage command is applied to the solenoid model, the response current and magnetic flux change the relationship. Compared with the test data, the fitted solenoid simulation model can represent the performance of the actual solenoid and reflect the change rule of the current and magnetic flux. Therefore, the follow-up will be based on this solenoid model to carry out the dual-solenoid position soft measurement method effect verification.

3.2.2 Structure parameters of solenoid

In the self-sensing model of the working solenoid, the necessary parameters can be obtained from the valve manufacturer or by disassembling the solenoid for measurement. In a more general implementation, the solenoid parameters should be provided by the valve manufacturer during production. This integration eliminates the need for additional measurements of the solenoid's structural characteristics, thereby streamlining the design and manufacturing process. However, in this paper, the size parameters of the solenoid are obtained by disassembling it and are listed in tab. 1. After checking the samples, the rated stroke of the solenoid is 4 mm. The core is compressed inward by 3 mm from the outermost end to the compression of 7 mm, which is the working interval. Therefore, the subsequent position soft measurement simulation is also carried out in this interval.

Table 1: Structure parameters of solenoid

Parameters	Value	Parameters	Value
r_{21}, r_{41}	15.80	l_3	0.0285
r_{22}, r_{42}	15.70	N	780
l_2, l_4	0.006	A_s	443.35e-06
r_{31}	12.65	R	5.07
r_{32}	12.50		

3.2.3 Calibrating idle solenoid inductance and position

The calibration of the inductance and position relationship of the idle solenoid is carried out using the electromagnet test bench, which is necessary for the idle solenoid to measure the position. According to the test, under the existing control board hardware facilities, the solenoid response to excitation is best

when the input excitation voltage amplitude is 0.5 V, the bias is 0.5 V (i.e., the voltage fluctuates between 0 and 1 V), and the frequency is 100 Hz. The current amplitude shows the most pronounced monotonic change with increasing air gap length. As shown in fig. 7, a quartic curve was selected to fit the inductance and position relationship after comparing the effects of the quadratic, cubic, and quartic curve fits. The final fitting parameters are shown in tab. 2.

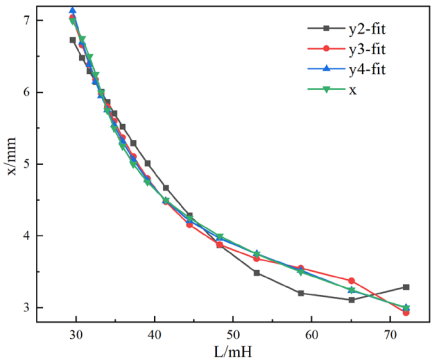


Figure 7: Inductance-position fitting

3.3 Dual-solenoid position soft measurement simulation model

According to the dual-solenoid position soft measurement method derived in Section 2, combined with the solenoid model already established in Section 3.2, the Simulink model of the dual-solenoid position soft measurement method simulation is built, as shown in fig. 8.

Fig. 8 shows the dual-solenoid position soft measurement method model. The purple part in the figure is the input signal module, which will generate sine and step position signal commands in the next section of the simulation. The green part in the figure is the working solenoid position self-sensing module, which inputs a 9 V voltage step signal at the beginning of position sensing (at 0.5 s) to generate the initial excitation and maintain the solenoid in the energized state. The equivalent inductance voltage, current, and flux signals of the solenoid are input into the position self-sensing model to output the position of the solenoid x_1. The yellow part in the figure is the idle solenoid position sensing module, to which a sinusoidal excitation signal is continuously applied with the voltage selection, as mentioned in Section 3.2.3. The current is acquired as the input to the subsequent idle solenoid position estimation module, and the final output is the position x_2. The

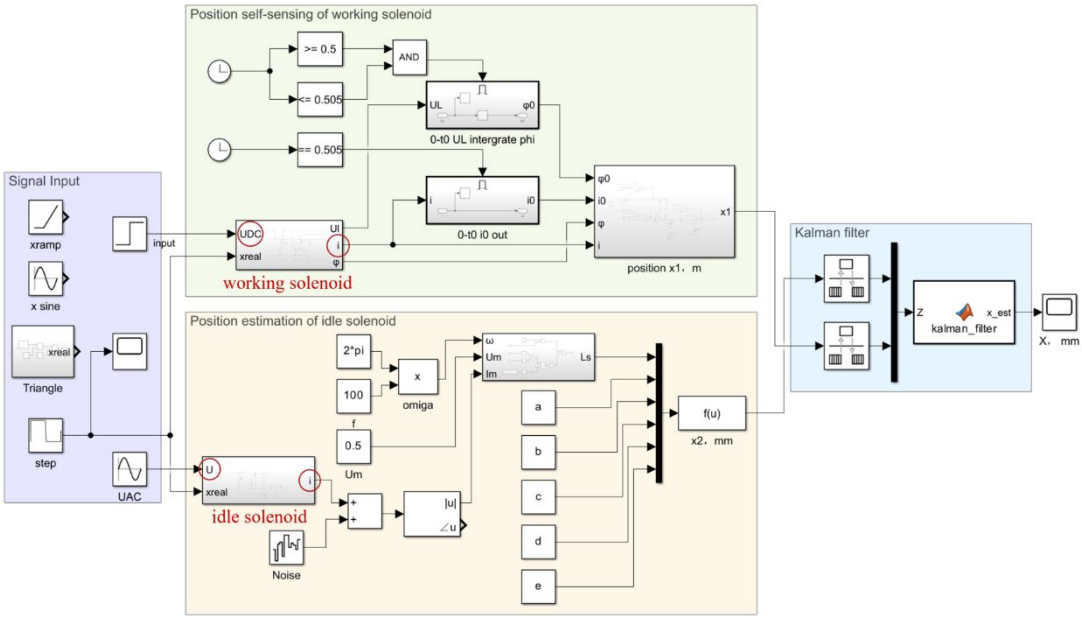


Figure 8: The simulation model of the dual-solenoid position soft measurement method

Table 2: Inductance-position fitting parameters of idle solenoid

Parameters	Value
a	3.4588e-06
b	-8.1919e-04
c	0.0724
d	-2.8758
e	47.3696

necessary parameters required for the dual solenoids soft measurement method are labeled in the figure. The system captures and archives the input voltage (DC for the working solenoid and AC for the idle solenoid) along with the corresponding response current for subsequent processing. Concurrently, other variables, such as magnetic flux and inductance voltage, are computed by integrating these acquired measurements with the solenoid pre-measured parameters, as described in Section 2. The blue part in the figure is the data fusion output module using the adaptive Kalman filtering method, where x_1 is used as the predicted value input and x_2 is used as the observed value input.

4 Simulations

4.1 Sinusoidal signal simulation

In this simulation, the input solenoid position is sinusoidally varied according to the interval of 3 mm - 7 mm with a variation frequency of 0.5 Hz. Fig. 9 shows the effect of position soft measurement when no noise is added. The position estimation results of the working solenoid and the idle solenoid can both accurately obtain the initial value and follow the actual position change. In more detail, the estimation results of the idle solenoid stabilize faster than those of the working solenoid. The fused results are closer to the actual value, and the accuracy is 1.91% after the measurement is stabilized.

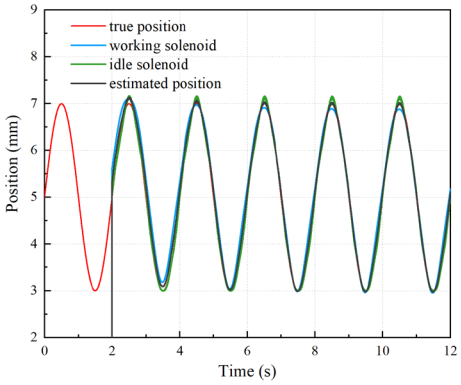
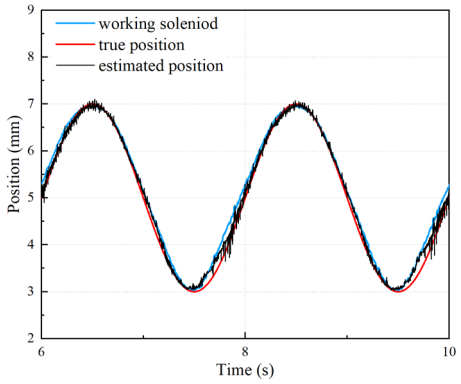
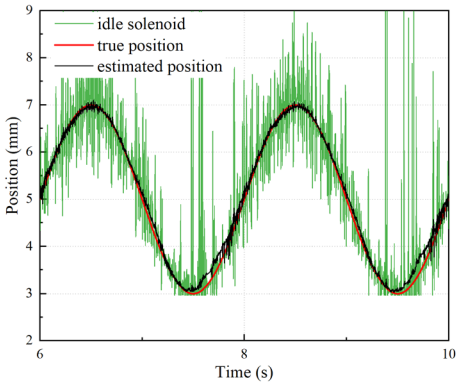


Figure 9: Sinusoidal signal simulation position estimation results (no noise)

Then, Gaussian white noise with a mean of 0 and a standard deviation of 0.005 A is added to the current sample of the solenoid model to simulate the measurement error of the current sensor due to inaccuracy in the actual application process and to verify the anti-interference ability of the model. Fig. 10(a) and 10(b) show that the estimation results of the working solenoid are more noise-resistant than those of the idle solenoid. After adaptive Kalman filter data fusion, the position soft measurements are well denoised and more closely match the actual displacement values than the individual estimates from either solenoid. The average accuracy of the position measurement remains within 6%, and the accuracy is slightly worse when the curve changes direction because the measured value of the working solenoid is somewhat different from the actual value, which affects the final result of the soft measurement to some extent.



(a)

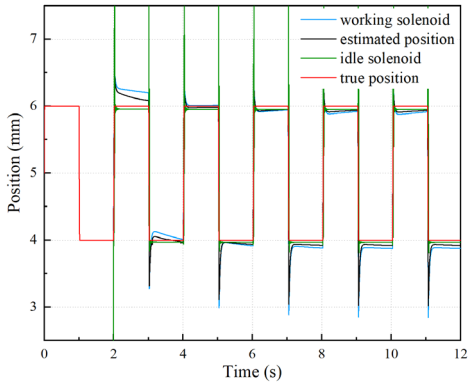


(b)

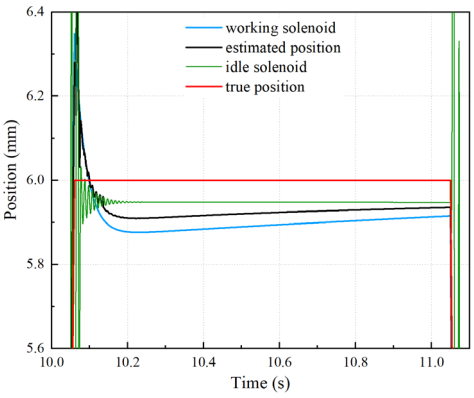
Figure 10: Sinusoidal signal simulation position estimation results (with noise)

4.2 Step signal simulation

In this simulation, the input solenoid position is made a step change between the position of 4 mm and 6 mm with a period of 2 s. Fig. 11(a) shows the effect of the position soft measurement method. The measured position follows the actual position change better, and the position soft measurement accuracy is within 2.33% of the whole stroke at the smooth time. Due to the inductance of the working solenoid being calculated iteratively, it is susceptible to a large error at the time of position switching. From fig. 11(b), it can be seen that this problem can be effectively mitigated by the dual-solenoid soft measurement method.



(a)

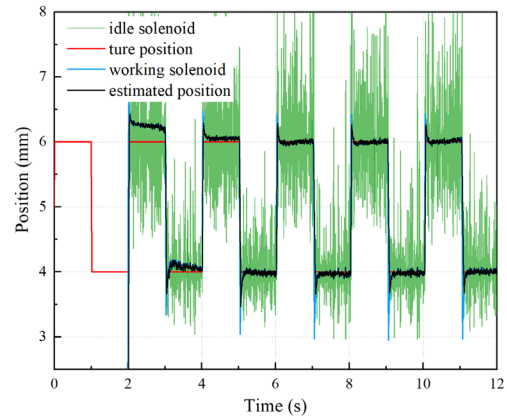


(b)

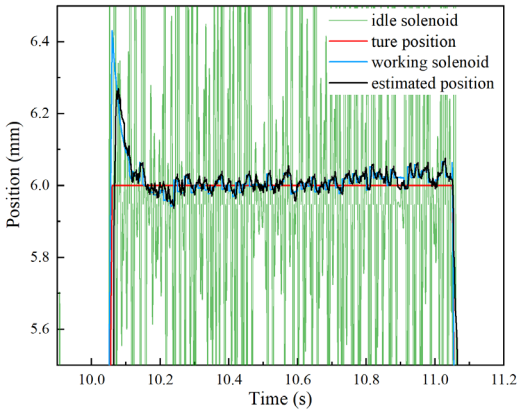
Figure 11: Step signal simulation position estimation results (no noise)

Similarly, Gaussian white noise with a mean value of 0 and a standard deviation of 0.005 A was later added to the current samples of the solenoid model to verify the noise immunity of the model. A comparison of fig. 12(a) and 12(b) show that the position soft measurements after data fusion by adaptive Kalman filtering have a good denoising effect. The accuracy is 2.06% after the measurement is stabilized. Similarly, the impact on the working solenoid due to the sudden change in position generation during the step is mitigated, as can be clearly seen in fig. 12(b).

In the step signal simulation, the results also indicate that the position estimation method based on the idle solenoid exhibits a faster dynamic response, whereas the method based on the working solenoid demonstrates superior robustness against noise. The dual-solenoid position soft measurement method proposed in this study effectively integrates the strengths of both methods, leading to improved estimation accuracy.



(a)



(b)

Figure 12: Step signal simulation position estimation results (with noise)

5 Conclusion

This paper proposes a position soft measurement method based on dual solenoid data fusion to realize an accurate estimation of spool position. The method solves the problems of inaccuracy, low adaptability, and susceptibility to noise interference brought by a single solenoid. Besides, the method utilizes only the original parts of the solenoid valve without adding additional auxiliary components. After prior calibration, it can realize the effect of accurately detecting the initial position and precisely following the position change. Simulation results demonstrate that the soft measurement method based on dual-solenoid position data fusion yields promising outcomes under various operating conditions. In noise-free scenarios, the measurement accuracy after stabilization remains within 2.5% of the full stroke. When subjected to Gaussian noise interference with a standard deviation of 5 mA, the method demonstrates strong noise immunity, ensuring accurate position identification across most of the stroke. Although a transient degradation in precision (exceeding 5%) is observed during spool directional reversals, the system rapidly recovers to yield correct measurements. Currently, proportional valves without displacement sensors but with mechanical hydraulic feedback achieve ±1% accuracy and under 3% hysteresis. The performance achieved by the method proposed in this paper nearly meets this criterion, indicating its potential to replace conventional valves for improved automation in production processes. Moreover, this advancement can facilitate condition monitoring and support fault diagnosis applications, such as detecting spool sticking.

The method proposed in this paper is specifically for dual-solenoid proportional valves, where the coordinated operation of two solenoids enables the final position estimation. This dual-solenoid configuration effectively compensates for the shortcomings inherent in single-solenoid systems. In the industrial applications where valves are equipped with a single solenoid, the working solenoid position estimation method may be more suitable for simultaneously achieving both force generation and position sensing. It is critical to ensure the accuracy of the solenoid's structural parameters, as the position calculation relies on data from the solenoid. Simulation results show that the working solenoid method exhibits greater resilience to noise. Conversely, although the idle solenoid method operates independently of the structural parameters, it is more sensitive to the noise and requires additional filtering to mitigate noise interference. Besides, if this approach were to be applied to a single-solenoid valve, it would necessitate a redesign of the valve structure to incorporate an auxiliary coil for sensing the changes in inductance.

As the use effect of this study has mainly been conducted using an established solenoid model, future experiments will be carried out on a valve equipped with a displacement sensor. The sensor data will be compared to validate the efficacy of the dual-solenoid position soft measurement method. Additionally, the prolonged operation of the solenoid may lead to an increase in temperature, which could theoretically affect both the resistance and the accuracy of the position estimation. Therefore, future work will need to address the impact of temperature variations on these parameters.

Nomenclature

Designation	Denotation	Unit
R_1, R_2, R_3, R_4	Magnetic reluctance of air gaps 1, 2, 3, and 4	1/H

x_1	Position of the iron core /Length of air gap 1	m
μ_0	Hydraulic oil magnetic permeability	H/m
A_s	Cross-sectional area of air gap 1	m^2
r_{21}, r_{31}, r_{41}	Outer radius of air gaps 2, 3, and 4	mm
r_{22}, r_{32}, r_{42}	Inner radius of air gaps 2, 3, and 4	mm
l_2, l_3, l_4	Length of air gaps 2, 3, and 4	m
R_5	Armature magnetic reluctance	1/H
R_6	Yoke magnetic reluctance	1/H
ψ	Working solenoid magnetic flux linkage	Wb
N	Coil turns	/
ϕ_1	Average magnetic flux	Wb
L_w	Working solenoid inductance	H
I_w	Working solenoid current	A
U_{DC}	DC supply voltage	V
U_L	Working solenoid inductive voltage	V
U_R	Working solenoid resistive voltage	V
R	Coil resistance	Ω
ψ_0	Working solenoid initial magnetic flux linkage	Wb
L_0	Working solenoid initial inductance	H
i_0	Working solenoid initial current	A
t_0	Initial voltage generation time	s
Δt	Time increment	s
δ_0	Initial armature position	m
$\Delta\psi$	Magnetic flux linkage variation	Wb
L_1	Inductance after the L_0	H
Δi	Current variation	A
i_{k-1}	Current at the previous time step	A
L_{k-1}	Inductance at the previous time step	H
L_k	Present inductance	H
Δx	Position variation	m
U_{AC}	AC supply voltage	V
$\bar{U}_m$	AC voltage amplitude	V
ω	Angular frequency	rad/s
U_k	DC offset of the AC voltage	V
I_{AC}	Response current	A
$\bar{I}_m$	Response current amplitude	A
θ	Phase lag between the current and voltage	degree
I_k	DC offset of the response current	A
ϕ_2	Idle solenoid magnetic flux	Wb
L_s	Idle solenoid inductance	H
x_2	Estimated position of the idle solenoid	m
a, b, c, d, e	Coefficients of the quartic fitting polynomial	/
X_k	State vector	/
x_k	Solenoid core displacement	m
v_k	Solenoid core velocity	m/s
T	Sampling period	s
A	State transition matrix	/
Q_k	Process noise covariance	/
w_{k-1}	Process noise vector	/
$z_{1,k}$	Working solenoid position x_1	m
$z_{2,k}$	Idle solenoid position x_2	m
C	Observation matrix	/
z_k	Observation equation	/
R_k	Observation noise covariance	/
R_{base}	Initial observation noise covariance	/
r_{base1}, r_{base2}	Initial observation noise covariance parameters	/
x_{Ru}, x_{Rl}	Upper and lower segmentation points of the difference between $z_{2,k}$ and $z_{1,k}$	m
q_{k1}, q_{k2}, … q_{k6}	Process noise covariance parameters	/
x_{Qu}, x_{Ql}	Upper and lower segmentation points of the difference between x_2 and current estimation	m
α	Fusion factor	/
s_{err}	Step error	m
n_L	Noise level	m
ε	Small constant to prevent the denominator from being zero	/
P_x	Error covariance	/
$\hat{x}_k$	Final estimated position	m

Acknowledge

This work was supported by the State Key Program of National Natural Science of China (509116-N12501ZJ).

References

[1] M Yang, K Lian, S Luo and et al. Simulation analysis and optimization of pilot type electro-hydraulic proportional directional valve based on multi field coupling. Flow Measurement and Instrumentation, 100:102726, 2024.

[2] J Ren, B Zhao, L Quan, et al. Research on proportional valve spool wear diagnostic method hybrid-driven by the valve port energy loss mechanism model and data. Mechanical Systems and Signal Processing, 219:111606, 2024.

[3] J Shen, H Zong, M Cheng and et al. Force soft measurement based on cylinder decoupling model and stepwise parameter identification for hydraulic manipulators. IEEE Transactions on Instrumentation and Measurement, 73:7504912, 2024.

[4] W Ma, L Tan, H Feng, et al. A data-driven LSTM soft sensor model based on Bayesian optimization for hydraulic pressure measurement of excavator. IEEE Sensors Journal, 23(21):25749-25759, 2023.

[5] D Xie, Y Yang, Y Zhang and et al. Precision positioning based on temperature dependence self-sensing magnetostrictive actuation mechanism. International Journal of Mechanical Sciences, 272:109174, 2024.

[6] H C Pedersen, T Bak-Jensen, R H Jessen, et al. Temperature-independent fault detection of solenoid-actuated proportional valve. IEEE/ASME Transactions on Mechatronics, 27(6):4497-4506, 2022.

[7] B A Reinholz and R J Seethaler. Sensor fusion of self-sensed measurements for position control of a constant air-gap solenoid. IEEE Sensors Journal, 22(24):23997-24005, 2022.

[8] S Nagai, T Nozaki and A Kawamura. Environmental robust position control for compact solenoid actuators by sensorless simultaneous estimation of position and force. IEEE Transactions on Industrial Electronics, 63(8):5078-5086, 2016.

[9] S Nagai and A Kawamura. Position sensorless position control for dual solenoid actuator. Proc. of 2018 International Power Electronics Conference (IPEC-Niigata 2018 -ECCE Asia), 1687-1691, 2018.

[10] T Kramer and J Weber. Self-sensing design of proportional solenoids. Proc. of Proceedings of the BATH/ASME 2020 Symposium on Fluid Power and Motion Control, 2811, 2020.

[11] T Braun, J Reuter and J Rudolph. Observer design for self-sensing of solenoid actuators with application to soft landing. IEEE Transactions on Control Systems Technology, 27(4):1720-1727, 2019.

[12] X Zhao, L Li, J Song, et al. Linear control of switching valve in vehicle hydraulic control unit based on sensorless solenoid position estimation. IEEE Transactions on Industrial Electronics, 63(7):4073-4085, 2016.

[13] G Cao， H Hu, S Huang et al. An assistant-mover-based position estimation method for planar switched reluctance motors. IEEE Transactions on Industrial Electronics, 67(7):6066-6077, 2020.

[14] Y Jing, S Wang, Q Liu et al. Gap self-sensing method for maglev system using partial electromagnet coil. IEEE Transactions on Industrial Electronics, 71(4):4273-4282, 2024.

[15] M Verma, V Lafarga, and C Collette. Perfect collocation using self-sensing electromagnetic actuator: Application to vibration control of flexible structures. Sensors and Actuators A: Physical, 313: 112210, 2020.

[16] J Fang, X Kong, X Zhu, et al. The modeling and experimental verification of a servo-proportional valve. Applied Mechanics and Materials, 220-223:1018-1022, 2012.

[17] B Xu, Z Lu, J Zhang, et al. Dynamic experiment and modeling of a proportional solenoid electromagnet. Chinese Hydraulics & Pneumatics. 9:1-5, 2015.

A Vision-Based Vibration Measurement Method for Large Hydraulic Manipulator

Shuwei Yang, Weidi Huang*, Junhui Zhang, Bing Xu, Yaoxing Wang, and Qi Wang

State Key Laboratory of Fluid Power and Mechatronic System, Zhejiang University, Hangzhou, China
E-mail: swyang@zju.edu.cn, wdhuang@zju.edu.cn, benzjh@zju.edu.cn, bxu@zju.edu.cn, zjuyxwang@zju.edu.cn, qiwang98@zju.edu.cn.

Abstract

Vibration frequently occurs at the Tool Center Point (TCP) of large hydraulic manipulators during operation, making the monitoring of TCP vibrations imperative for ensuring the safety and operational efficiency of the manipulator. Traditional methods for TCP vibration measurement generally rely on contact-based techniques, such as Inertial Measurement Units (IMU). However, these methods face limitations, including an inability to directly quantify vibration amplitudes and susceptibility to noise interference, which can negatively impact the manipulator's performance. To address these challenges, the present study proposes a novel, non-contact vibration measurement technique that utilizes vision-based technology. Specifically, a high-resolution camera is employed to capture images of the manipulator's TCP, with the Harris corner detection algorithm applied to identify key features. The changes in these features are then tracked and converted into positional shifts within the manipulator's coordinate system. This approach ensures the precision and stability of the measurement and prevents any direct interference with the equipment's operation. Experimental results across various operational scenarios reveal that, when compared to conventional methods, the proposed vision-based technique offers superior accuracy and stability in non-contact settings. Additionally, it demonstrates enhanced performance in vibration data extraction and processing efficiency. The method delineated in this study provides a viable solution for the vibration measurement and maintenance of large hydraulic manipulators, offering novel insights into enhancing operational safety and precision control.

Keywords: Vibration measurement, Image recognition, Manipulator kinematics, Signal processing.

1 Introduction

Large hydraulic manipulators are widely used in industries such as material handling, port loading and unloading, and municipal construction, often operating in high-load environments [1,2]. As a result, ensuring the stability and safety of these manipulators has become a critical concern. The nonlinear characteristics of the hydraulic system, combined with the coupling effects of the multi-link structure, lead to low-frequency vibrations (typically between 2-25Hz) during transient operations such as start-up, stop, and reversal. Although these vibrations generally have minimal impact on the manipulator's overall functionality, they can cause issues such as reduced operational accuracy, structural fatigue, and hydraulic pipeline leaks. Therefore, the development of high-precision vibration measurement methods, suitable for complex industrial environments, has become an urgent technical challenge that must be addressed.

Traditional vibration monitoring techniques commonly employ contact sensors, such as piezoelectric accelerometers and fiber optic IMUs, for vibration measurement [3,4]. While these methods offer certain advantages in low-frequency vibration monitoring, their inherent limitations restrict their widespread application in complex industrial environments. Contact sensors generally require precise mounting on the manipulator's end effector, a process that demands high accuracy during installation. In specific operational environments, such as those with elevated temperatures, oil contamination, and intense vibrations, these sensors are particularly susceptible to interference and signal drift. Additionally, the installation process may alter the mass distribution of the measured object, introducing vibrational disturbances that compromise measurement accuracy. Conventional contact methods, like IMUs, provide only acceleration data, necessitating secondary integration to estimate displacement amplitudes. This process is prone to cumulative errors and cannot directly yield the actual vibration amplitude. In large workspace environments, the installation and wiring process is often complex, requiring cables to be routed along the entire length of the manipulator. This process

is both time-consuming and labor-intensive, and it may disrupt the normal operation of the manipulator.

Non-contact devices, such as the Laser Doppler Vibrometer (LDV), offer micron-level measurement accuracy and are capable of effectively monitoring vibrations in complex environments [5]. However, the high cost of these devices (typically exceeding $100,000 per unit), intricate optical calibration procedures, and sensitivity to line-of-sight obstructions present significant challenges for their widespread use in practical construction scenarios. As a result, the cost-performance ratio of these devices is substantially lower than that of traditional contact sensors.

In recent years, with the rapid development of computer vision technology, vision-based vibration measurement methods have emerged as a significant research direction [6-8]. These methods extract and analyze the motion trajectories of target feature points in image sequences to directly obtain spatial displacement information. They offer notable advantages, such as non-contact measurement, full-field monitoring, and low cost. Vision-based vibration measurement methods can currently be categorized into two types: the marker-based template matching method, which relies on manually calibrated markers and is susceptible to occlusion in complex scenarios [9], and the markerless tracking method, which utilizes optical flow. The latter is independent of manual intervention and offers enhanced flexibility [10]. However, it is vulnerable to noise interference when employing the traditional Lucas-Kanade algorithm, leading to tracking failures in scenarios with significant displacements [11,12]. The operation of the hydraulic manipulator typically results in low-frequency, small-amplitude oscillations at the TCP, which imposes stringent requirements on the positioning accuracy and real-time performance of vision algorithms.

To overcome the limitations of existing methods, this paper proposes a vision-based vibration measurement method integrating Harris corner detection with optical flow tracking. The Harris corner detection algorithm is used to extract target feature points, and optical flow tracking enables precise tracking, allowing for efficient and accurate extraction of vibration displacement data. Unlike traditional contact sensors, the vision-based measurement method does not require direct contact with the equipment or environment, enabling stable operation in harsh conditions with strong robustness. To validate the accuracy of the method, a comparison with existing contact-based measurement methods is conducted. Experimental results demonstrate that the vision-based vibration measurement method ensures accuracy and stability while exhibiting excellent performance in vibration data extraction and processing efficiency. This approach offers an innovative solution for vibration measurement and maintenance of large hydraulic manipulators, providing effective technical support for enhancing the safety and precision control of equipment.

The structure of this paper is as follows: Chapter 2 presents the research object and technical approach, providing a detailed explanation of the proposed method. Chapter 3 outlines the specific implementation steps for vision-based vibration measurement. Chapter 4 describes the procedure for contact-based vibration measurement. Chapter 5 details experiments conducted under various operating conditions to validate the effectiveness of the proposed method. Chapter 6 concludes the paper and discusses future research directions.

2 Technical approach

The proposed methodology employs a vision-based, non-contact vibration measurement technique for high-precision monitoring of the TCP of large hydraulic manipulators (Figure 1). A camera is used to capture real-time images of the manipulator's end effector, and image processing techniques are applied to detect small vibration variations. The integration of Harris corner detection and optical flow-based feature tracking allows for the extraction of vibration displacement data from the TCP without requiring direct contact with the equipment, providing an efficient and stable solution for vibration measurement.

Figure 1: Concrete pumping equipment fitted with a large hydraulic manipulator.

Initially, markers with distinct geometric features are affixed to the TCP of the manipulator to serve as target points for subsequent image processing. A high-resolution camera is used during the image acquisition process to ensure an adequate frame rate and resolution, allowing for precise capture of small vibrations at the TCP. Each image frame is processed using the Harris corner detection algorithm to extract feature points, which correspond to the positions of the markers on the manipulator's TCP. This process forms the foundation for subsequent displacement calculations.

The Lucas-Kanade algorithm, based on optical flow, is then applied to track the extracted feature points and capture their motion trajectories across adjacent image frames [13,14]. Optical flow operates under the assumption of image brightness consistency and calculates pixel displacements between consecutive frames to derive relative position changes of the TCP. Given that the vibrations of hydraulic manipulators typically manifest as minute displacements, optical flow can accurately track these subtle movements, ensuring high precision in vibration measurement.

This study utilizes camera calibration to obtain the camera's intrinsic and extrinsic parameters, employing geometric

transformation methods to convert displacement information from the image coordinate system into the manipulator's work coordinate system. The camera's intrinsic parameters (e.g., focal length and principal point) and extrinsic parameters (describing the relationship between the camera and the world coordinate systems) are used to achieve precise mapping from the image plane to the actual working space through perspective transformation and other algorithms.

To validate the accuracy of the proposed method, a comparison is made between the vibration measurement results obtained using the vision-based method and the experimental results from a contact-based measurement method (IMU). As a traditional vibration measurement tool, the IMU provides dynamic information, such as acceleration. However, due to the error accumulation characteristics and the inherent limitations of the acceleration integration method, significant errors may be introduced over long measurement durations. To address this, signal processing techniques are applied to extract vibration components from the IMU data. By combining the manipulator's kinematic model with the angular data provided by the joint sensors, displacement data for the TCP is obtained, excluding vibration signals. The comparison of experimental results from both the vision-based and contact-based methods confirms the superiority of the former in terms of accuracy and effectiveness.

3 Vision-based measurement method

Vibration measurement of the TCP requires non-contact vibration signal detection using computer vision technology, which involves precise camera calibration and image processing techniques to extract displacement information. This section provides a detailed explanation of the camera calibration methods, including intrinsic and extrinsic parameters, Harris corner detection principles, and optical flow applications. Additionally, the relationship between the checkerboard markers and the camera matrix is utilized to convert pixel coordinates into the manipulator's working coordinate system, enabling precise measurement of TCP position variations.

3.1 Camera intrinsic and extrinsic parameters calibration

To accurately convert pixel information from the image coordinate system into physical coordinates, the intrinsic and extrinsic parameters of the camera must first be calibrated. The intrinsic parameters include the focal length, principal point, and distortion coefficients. The extrinsic parameters describe the relationship between the camera coordinate system and the world coordinate system. Calibration of both intrinsic and extrinsic parameters can be achieved using a calibration board with known dimensions, with the Zhang checkerboard calibration method [15] being a widely used approach.

During the calibration process, a calibration board with a known square size is used, and multiple images are captured to extract the positions of the chessboard corners. The 2D coordinates of the chessboard corners in the image are denoted as (u_i, v_i), and the corresponding 3D coordinates in the world coordinate system are (X_i, Y_i, Z_i). The intrinsic and extrinsic parameters of the camera satisfy the following projection relationship:

$$\begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \mathbf{K} \begin{bmatrix} \boldsymbol{R} & \boldsymbol{T} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{bmatrix} \tag{1}$$

Where $\boldsymbol{K}$ is the intrinsic parameter matrix of the camera, represented as:

$$\mathbf{K} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

Where f_x and f_y represent the focal lengths, and c_x and c_y represent the principal point coordinates. $\boldsymbol{R}$ and $\boldsymbol{t}$ are the rotation matrix and translation vector in the camera's extrinsic parameters, respectively, which describe the transformation from the camera coordinate system to the world coordinate system.

By minimizing the reprojection error, the intrinsic and extrinsic parameters of the camera are computed, enabling the conversion from image coordinates to world coordinates.

3.2 Harris corner detection and optical flow tracking

The Harris corner detection algorithm used in this paper is a widely applied feature point extraction method in image processing. Corners are detected by computing the local autocorrelation matrix of each pixel in the image. The image gradients in the x and y directions are represented by Ix and Iy, respectively. The autocorrelation matrix within the image window is expressed as:

$$M = \begin{bmatrix} \Sigma I_x^2 & \Sigma I_x I_y \\ \Sigma I_x I_y & \Sigma I_y^2 \end{bmatrix} \tag{3}$$

Where $\sum I_x^2$ is the sum of the squared gradients in the x-direction within the window, and $\sum I_x I_y$ is the sum of the product of the gradients in the x and y directions. By calculating the eigenvalues of the matrix M, it is possible to determine whether a point is a corner. Specifically, the Harris corner response function is defined as:

$$R = \det(M) - k(\text{trace}(M))^2 \tag{4}$$

Where k is an empirical constant, typically in the range of 0.04 to 0.06, det(M) is the determinant of the matrix M, and trace(M) is the trace of matrix M. In the event that the value of R is sufficiently large, the point is designated as a corner, thereby indicating the presence of significant features in the image. In the context of vibration measurement, the markers situated at the TCP of the manipulator can be utilized to extract multiple stable corner points through the implementation of Harris corner detection. These corner points subsequently serve as feature points for the purpose of subsequent optical flow

tracking, thereby providing the foundational data necessary for the calculation of TCP position changes.

After corner points are detected in each image frame, optical flow is applied to compute the pixel displacements between adjacent frames, thereby tracking the movement of objects in the image. The fundamental assumption of optical flow is that the intensity values between adjacent frames remain constant, meaning the object in the image undergoes translational motion within a short time. Based on this assumption, the intensity values of a feature point in the image at times t and t+1 are represented by $I(x, y, t)$ and $I(x+\Delta x, y+\Delta y, t+1)$, respectively. The fundamental assumption of optical flow is as follows:

$$I(x, y, t) = I(x + \Delta x, y + \Delta y, t + 1) \tag{5}$$

Performing a Taylor expansion and neglecting higher-order terms results in the optical flow equation:

$$I_x u + I_y v + I_t = 0 \tag{6}$$

Where u and v represent the velocities of the point in the x and y directions, respectively, and I_x and I_y are the image gradients in the x and y directions, while I_t is the time gradient. By using the pixel gradients within a local window, optical flow can be applied to calculate the displacement of each feature point in the image, thus providing the vibration displacement information of the TCP.

3.3 Coordinate transformation

The integration of Harris corner detection and optical flow tracking facilitates the acquisition of variations in the coordinates of the corner points in the image plane. These corner points represent changes in the position of the TCP of the manipulator, and by analyzing the trajectory of these points, the displacement of the TCP can be accurately tracked. In this paper, the intrinsic and extrinsic parameter matrices derived from the camera calibration process are used to convert pixel coordinates in the image plane to 3D physical coordinates in the world coordinate system. These parameter matrices describe the relationship between the camera coordinate system and the world coordinate system, enabling effective mapping of pixel points in the image coordinate system to the physical workspace. A detailed explanation of the coordinate transformation process is provided in Equations 1 and 2.

4 Contact Measurement method

When using the IMU for measurements, it must be securely mounted on the target to ensure it moves in conjunction with the object being measured [16]. The IMU is capable of recognizing acceleration signals that encompass the rigid motion of the arm, external disturbances, and vibration components. These signals are highly susceptible to low-frequency drift (e.g., static acceleration caused by gravity) and high-frequency noise. To mitigate the impact of these factors on the measurement results, the acquired acceleration signals are preprocessed, with only the vibration-related valid signals being extracted. The IMU is positioned near the end of the manipulator to capture its acceleration signals. Through data processing, the vibration of the Tool Center Point (TCP) is characterized. Angle encoders are installed at the joints of the manipulator to measure displacement signals that do not include vibration components. The rigid displacement of the arm is determined through a combination of joint angle information and the kinematic model. This dual approach is integrated to ensure the precision of the TCP displacement.

4.1 Bandpass filtering

Band-pass filtering serves as the initial phase in the processing of IMU acceleration signals. Since the vibrations of the manipulator typically occur within a specific frequency range, the band-pass filter is used to efficiently eliminate undesired low-frequency and high-frequency components. Low-frequency components may include gravitational acceleration or slower rigid body motion signals, while high-frequency components may arise from noise or other disturbances.

In this study, a Butterworth band-pass filter is applied due to its smooth and flat frequency response within the passband, which ensures that only the relevant vibration frequencies are retained while effectively suppressing unwanted components. This configuration helps eliminate low-frequency components associated with the manipulator's rigid motion and sensor noise, thereby preserving the frequency components related to vibration. The band-pass filtered acceleration signal primarily contains components that correspond to the vibration frequency of the manipulator, providing a clear signal foundation for subsequent displacement extraction.

4.2 Freuqency-domain processing and double integration

Given that acceleration is defined as the second derivative of displacement, direct integration of the acceleration signal in the time domain introduces drift errors, which can lead to inaccurate results. To address this issue, a frequency-domain integration method for displacement extraction is proposed in this paper. This approach involves transforming the band-pass filtered acceleration signal into the frequency domain using the Fourier transform, which converts the time-domain signal into a frequency-domain signal, allowing for more precise manipulation of each frequency component. In the frequency domain, the acceleration signal A(f) is represented as a summation of frequency components, where f denotes the frequency.

The primary advantage of frequency-domain analysis is that it avoids the error accumulation that occurs during time-domain integration. In the frequency domain, integration is equivalent to multiplying each frequency component, a process that is more stable and accurate. This method effectively eliminates the drift errors introduced during time-domain integration.

In the frequency domain, each frequency component of the acceleration signal is integrated. Since acceleration is the second derivative of displacement, two integrations are

required. Initially, the processed acceleration signal $A(f)$ is integrated to yield the velocity signal $V(f)$:

$$V(f) = \frac{A(f)}{i2\pi f} \tag{7}$$

The velocity signal $V(f)$ is then integrated again to obtain the displacement signal $D(f)$:

$$D(f) = \frac{V(f)}{i2\pi f} = \frac{A(f)}{(i2\pi f)^2} \tag{8}$$

Once the vibration displacement signal D(f) has been obtained in the frequency domain, the inverse Fourier transform is applied to convert the frequency-domain signal back to the time domain, yielding the final vibration displacement signal. The inverse Fourier transform recovers the individual frequency components in the frequency domain and reconstructs a continuous signal in the time domain. The resulting time-domain signal represents the pure vibration displacement, effectively eliminating errors caused by initial conditions and noise. The frequency-domain integration method facilitates the precise extraction of vibration components from the acceleration signal, bypassing the error accumulation that can occur with time-domain integration.

4.3 Rigid displacement calculation

During the movement of the manipulator, the positional changes of the TCP involve not only the minor variations caused by vibrations but also substantial changes resulting from rigid body motion. Therefore, after obtaining the pure vibration displacement signal from the IMU data, it is necessary to calculate the rigid motion changes of the manipulator based on the angular information between the joints (i.e., positional changes that exclude vibration components).

In this study, a kinematic model is developed for the mechanical structure of the large manipulator (Figure 2). The three sections of the manipulator's arms can rotate around the joints, and hydraulic actuators drive the amplitude-changing mechanism, which controls the lifting and lowering of each arm section. The joints are subject to angular constraints, and by utilizing the angular information between the joints, the TCP position can be computed by combining the arm lengths, as expressed by the following equation:

$$\begin{cases} x_{TCP} = c_0(a_3c_{123} + a_2c_{12} + a_1c_1 + a_0) \\ y_{TCP} = s_0(a_3c_{123} + a_2c_{12} + a_1c_1 + a_0) \\ z_{TCP} = a_3s_{123} + a_2s_{12} + a_1s_1 + d_0 \\ \xi = \varphi_1 + \varphi_2 + \varphi_3 \end{cases} \tag{9}$$

where φ_0, φ_1, φ_2, φ_3 denote angular vectors of joints 1-3, $\boldsymbol{P} = [x, y, z]^{\mathrm{T}} \in \mathbf{R}^3$ specifies the TCP position. $\zeta \in (-\pi/2,\ \pi/2)$ represents the inclination angle between the terminal arm section and horizontal plane. The specific corresponding parameters are shown in the table. φ_{lim} define rotational constraints per joint. All physical parameters are quantified in Table 1.

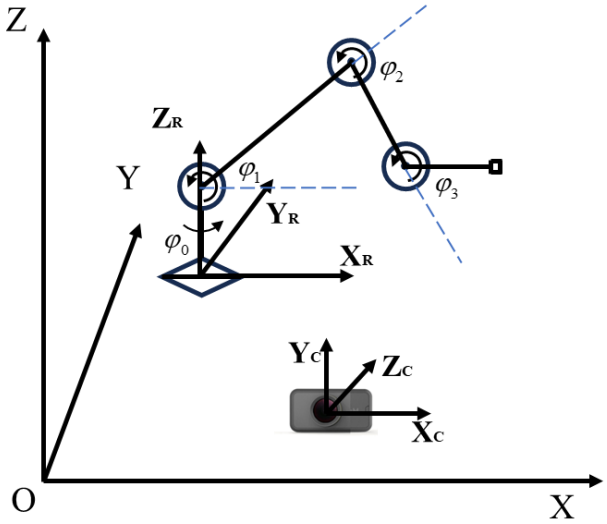


Figure 2: D-H coordinates of manipulator.

Table 1: D-H parameters.

Link	φ_i(°)	$d_i(mm)$	$a_i(mm)$	α_i(°)	φ_{lim}(°)
0	φ_0	2100	0	90	-90 ~ 90
1	φ_1	0	5500	0	0 ~ 88
2	φ_2	0	3750	0	-180 ~ 0
3	φ_3	0	3700	0	-60 ~ 180

The rigid displacement, obtained from the forward solution of the manipulator's kinematics, and the vibration displacement, obtained from the frequency-domain integration of the IMU data, are accumulated to calculate the TCP position change, including the vibration displacement.

5 Experiment

5.1 Experimental setup

The experimental apparatus in this study consists of a concrete pumping system equipped with a manipulator. The device features a three-degree-of-freedom rotational manipulator, with a total length of 13.5 meters when fully extended. Each section of the manipulator is elevated and lowered by hydraulic actuators, combined with an amplitude-changing mechanism. The rotational motion of the manipulator is driven by a rotary hydraulic motor that operates a gear system. The technical specifications of the manipulator are detailed in Table 2. To obtain angular information between the joints, angle encoders are installed, and an IMU sensor is positioned near the TCP of the manipulator. The experimental setup and

the installation of the corresponding sensors are shown in Figure 3.

The controller sampling frequency is 1000 Hz, the IMU sampling frequency is 500 Hz, and the camera sampling frequency is 60 Hz. The sampling frequency of the control multiway valves is 20 Hz.

Table 2: Experimental platform structure parameters.

Nomenclature	Value
Boom 1 length l_1	5.55 m
Boom 2 length l_2	3.75 m
Boom 3 length l_3	3.70 m
Teeth of Rotary Hydraulic Motor Gear z_1	15
Teeth of Slewing Platform Gear z_2	124
Rotary Mechanism Ratio	15/124

To validate the effectiveness of the vision-based vibration measurement method, two sets of experimental conditions are conducted. For each set, vibration measurements are performed using both vision-based and contact-based methods, and the results from both methods are compared and validated. The experimental conditions are as follows:

Experiment A: The camera is placed on the side of the manipulator, with a vertical distance between the manipulator and the camera ranging from 5m to 6m. This distance range simulates a typical scenario where the camera is closer to the manipulator, allowing it to capture higher-resolution images and clearer vibration information. This experiment is designed to verify that the vision-based vibration measurement method can achieve higher accuracy within the ideal distance range and provide an effective comparison with the contact method.

Experiment B: The camera is placed on the side of the manipulator, with a vertical distance between the manipulator and the camera exceeding 10m. This configuration simulates a scenario where vibration measurements are taken at longer distances. This setup allows for the evaluation of the vision-based approach at greater distances, along with the challenges it may face in real-world applications, such as image blur and degradation of vibration tracking accuracy.

Prior to the commencement of the experiments, the intrinsic and extrinsic parameters of the camera were calibrated. The intrinsic parameter matrix was obtained by capturing images of the checkerboard pattern in various poses and employing the Zhang calibration method. During the calibration process, the coordinates of multiple corner points on the checkerboard were measured, providing intrinsic parameters such as focal length and principal point position. The mean error in pixels is shown in Figure 4. The extrinsic parameter matrix was determined based on the relative position between the camera and the manipulator. This process ensured the accuracy and reliability of the camera measurements. To ensure precision, checkerboard markers were affixed near the IMU installation point at the TCP. During the experiment, the Harris corner detection algorithm was used to detect the four corner points of the black square at the center of the checkerboard, and optical flow was applied to track the motion trajectories of these corner points in real-time.

Considering the IMU sensor errors and noise, the IMU was initialized before its signals were used to calculate the TCP vibration displacement. Acceleration signals from the IMU were collected while the manipulator was stationary to eliminate errors caused by static gravity or other factors. During the experiment, acceleration signals were collected in real-time, and in the subsequent offline processing stage, the data were filtered and integrated. A band-pass filter was applied to remove low-frequency static acceleration and high-frequency noise, ensuring the signal's accuracy. Subsequently, the frequency-domain integration method was employed to convert the acceleration signal into displacement data, thereby obtaining the pure vibration displacement signal.

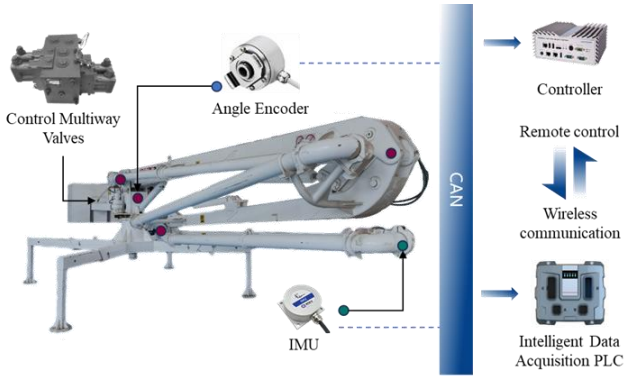
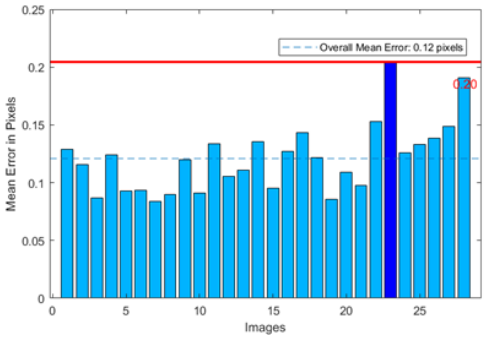
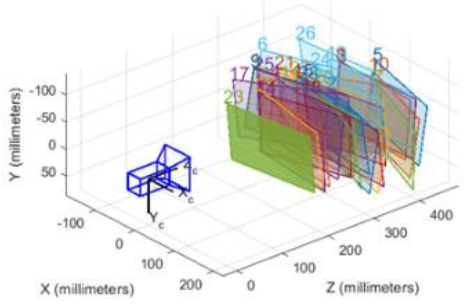


Figure 3: Experimental platform setup on the concrete pumping equipment.



(a) Camera inherent parameter calibration error



(b) Schematic diagram of camera parameter calibration

Figure 4: Camera inherent parameter calibration and parameter error.

5.2 Experiment A

During the motion of the manipulator, the vibration amplitude in the horizontal direction is found to be less significant than in the vertical direction. To provide a clear comparison of the effectiveness of the vision-based and contact-based measurement methods, the present experiment focuses on the extraction of the more prominent vertical vibrations. The results of Experiment A are shown in Figures 5-9.

In Experiment A, the motion duration of the manipulator is 50 seconds. During the start-up and stop phases, slight vibrations were observed, even when minimal TCP velocity was set, due to the impact of the hydraulic system. These vibrations are evident in the acceleration signal shown in Figure 6, predominantly occurring between 0-10 s and 40-50 s. Figure 7 illustrates the Z-direction vibration displacement changes obtained after processing the acceleration signal collected by the IMU. Figure 8 displays the angular data collected by the joint angle encoders of the manipulator. Figure 9 presents a comparison of the measurement results obtained from both the vision-based method and the contact-based methods (IMU and angle encoders).

The experimental results demonstrate that the proposed method exhibits nearly identical overall trends in the experimental curve when compared to the contact-based measurement methods. During the initial and final stages of the experiment, slight vibrations were observed in the TCP of the manipulator. The results indicate that both the vision-based and contact-based measurement methods accurately represented the vibration trends and amplitudes.

Errors in the vibration amplitudes measured by both methods were observed during certain time periods. This phenomenon is primarily attributed to the accumulation of errors in the IMU during extended measurement periods. As a result, the amplitude may be either overestimated or underestimated.

5.3 Experiment B

The specific experimental process for Experiment B is illustrated in Figure 10, with the manipulator's motion duration set to 60 seconds. The vibration of the manipulator's TCP is evident in the acceleration signal shown in Figure 11, predominantly occurring between 10-20 s and 50-60 s. Figure 12 presents the Z-direction vibration displacement changes derived from the processed acceleration signal collected by the IMU. The angular data collected by the joint angle encoders of the manipulator is shown in Figure 13. Figure 14 presents the measurement results obtained from both the vision-based and contact-based methods (IMU and angle encoders), along with a comparative analysis of the results.

Similar to Experiment A, slight vibrations were observed at the TCP of the manipulator during the initial and final stages of the experiment. The results indicate that both the vision-based and contact-based measurement methods accurately represented the vibration trends and amplitudes. However, some errors in the vibration amplitude were still observed during certain periods. This discrepancy is primarily attributed to the accumulation of errors within the IMU during prolonged measurement periods, resulting in the overestimation or underestimation of the amplitude at specific instances.

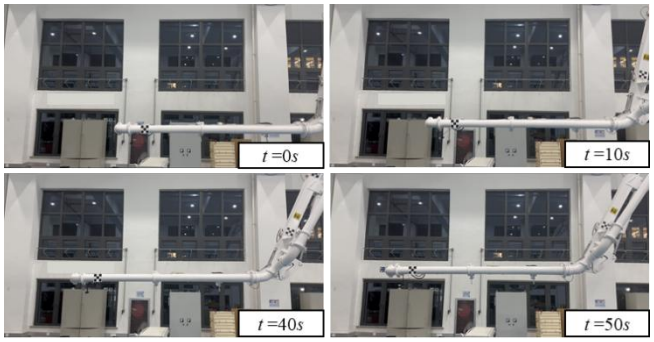


Figure 5: Manipulator motion process.

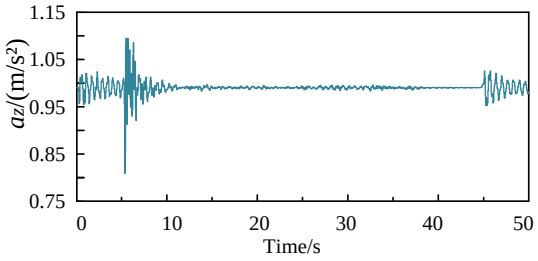


Figure 6: Z-direction acceleration variation.

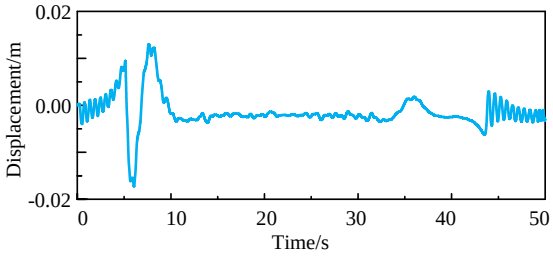


Figure 7: Variation of vibration displacement.

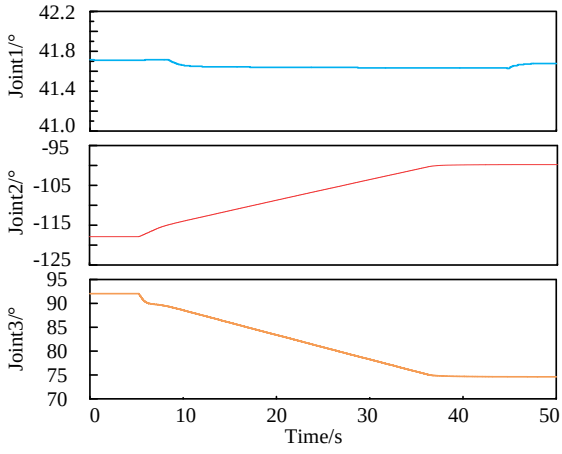


Figure 8: Manipulator Joint Angle Variation.

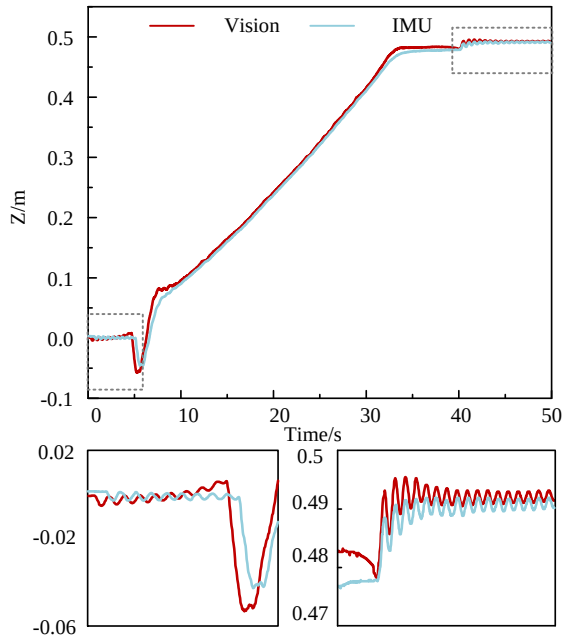


Figure 9: Position of the manipulator's TCP in the Z-direction.

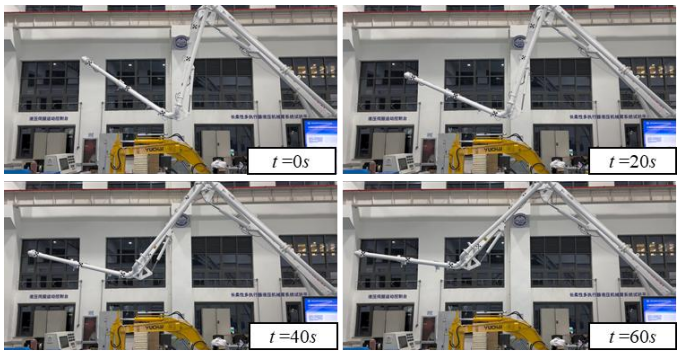


Figure 10: Manipulator motion process.

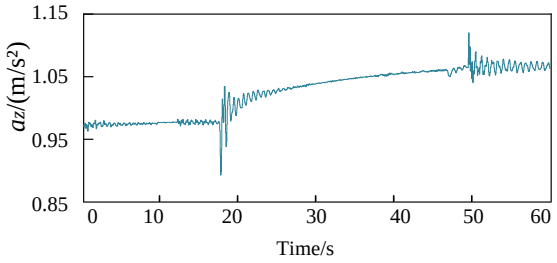


Figure 11: Z-direction acceleration variation.

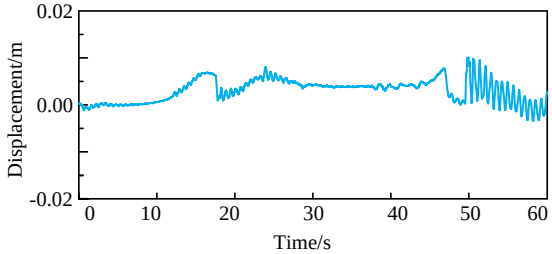


Figure 12: Variation of vibration displacement.

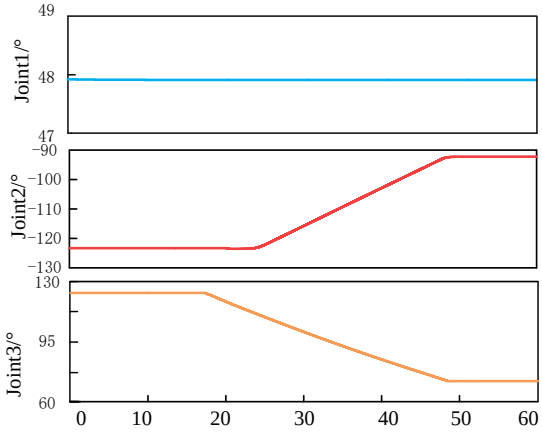


Figure 13: Manipulator Joint Angle Variation.

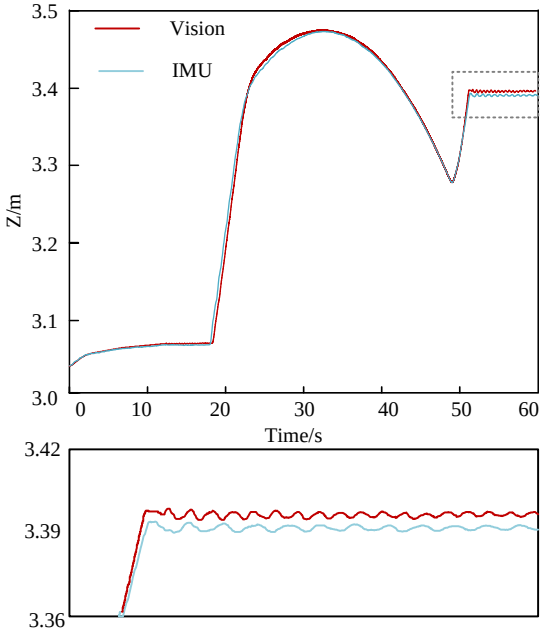


Figure 14: Position of the manipulator's TCP in the Z-direction.

5.4 Experiment

The experimental data demonstrate that both the vision-based and IMU methods can effectively characterize the vibration of the manipulator's end. Both the vibration amplitude and trend measured by the two methods show strong consistency. Although the vision-based method is more complex in terms of camera setup and calibration, it ultimately avoids the time-consuming process of sensor installation, which is prone to errors. In contrast, IMU methods require precise calibration and frequent maintenance during operation to minimize errors caused by drift effects.

Additionally, vision-based methods can incorporate advanced feature tracking and other techniques to enable real-time monitoring of displacement changes at the end of a large manipulator, offering high accuracy in measuring small vibrations. In contrast, the IMU method struggles to achieve

high-precision real-time displacement measurements, especially in the presence of small vibrations in complex environments where significant errors can arise.

In conclusion, the experimental results highlight the significant advantages of the vision-based vibration measurement method in terms of accuracy, stability, and ease of use. This method is particularly suitable for vibration monitoring and maintenance of large hydraulic manipulators, offering robust non-contact measurement capabilities. It shows broad application potential, especially in vibration monitoring scenarios that require high accuracy and stability.

6 Conclusions and outlook

This paper proposes a vision-based, non-contact vibration measurement method, which includes the calibration of the camera's intrinsic and extrinsic parameters, Harris corner detection, and feature point tracking based on optical flow. This study effectively overcomes the limitations of traditional contact-based sensors in complex environments, offering a stable and accurate non-contact vibration measurement solution.

The proposed method is compared with existing contact-based measurement methods (i.e., IMU and angle encoders) in a series of experiments conducted under two different operating conditions. The experimental results demonstrate the vision-based method's superiority in terms of accuracy and stability, accurately extracting displacement data related to manipulator vibrations.

This paper proposes an innovative solution for the vibration monitoring of large hydraulic manipulators. Future research can further optimise the corresponding visual detection algorithm to improve the adaptability and stability of the method in more complex environments. And the accuracy of the proposed method can be further verified by comparing it with other measurement means (laser total station tracking the end of the manipulator).

Acknowledgment

This work is supported by the National Natural Science Foundation of China (No. U21A20124).

References

[1] Geimer, M. Mobile Working Machines. SAE International, Warrendale, PA, 2020.

[2] Machado, T., Fassbender, D., Taheri, A., Eriksson, D., Gupta, H., Molaei, A., Forte, P., Rai, P. K., Ghabcheloo, R., Mäkinen, S., et al. Autonomous heavy-duty mobile machinery: A multidisciplinary collaborative challenge. In Proceedings of the 2021 IEEE International Conference on Technology and Entrepreneurship (ICTE), pp. 1–8, Kaunas, Lithuania, August 2021. IEEE.

[3] Try, P., and Gebhard, M. A vibration sensing device using a six-axis IMU and an optimized beam structure for activity monitoring. Sensors, 23(19):8045, 2023.

[4] Patrizi, G., Carratù, M., Ciani, L., Sommella, P., Catelani, M., and Pietrosanto, A. Analysis of inertial measurement units performances under dynamic conditions. IEEE Transactions on Instrumentation and Measurement, 72:3520713, 2023.

[5] Staszewski, W. J., Lee, B. C., Mallet, L., and Scarpa, F. Structural health monitoring using scanning laser vibrometry: I. Lamb wave sensing. Smart Materials and Structures, 13(2):251–260, April 2004.

[6] Baqersad, J., Poozesh, P., Niezrecki, C., and Avitabile, P. Photogrammetry and optical methods in structural dynamics—A review. Mechanical Systems and Signal Processing, 86:17–34, March 2017.

[7] Li, M., Yang, S., He, J., Gu, X., Xu, Y., Gu, F., and Ball, A. D. Full-field extraction of subtle displacement components via phase-projection wavelet denoising for vision-based vibration measurement. Mechanical Systems and Signal Processing, 224:112021, 2025.

[8] Deng, R., Lin, Y., Li, M., Gu, F., and Ball, A. Vision-based vibration mode shape extraction for identifying cracks in a cantilever beam. Journal of Physics: Conference Series, 2184(1):012037, May 2022.

[9] Kalpoe, D., Khoshelham, K., and Gorte, B. V. Vibration measurement of a model wind turbine using high speed photogrammetry. In Proceedings of SPIE, vol. 8085, p. 80850J, June 2011.

[10] Gautama, T., and Van Hulle, M. M. A phase-based approach to the estimation of the optical flow field using spatial filtering. IEEE Transactions on Neural Networks, 13(5):1127–1136, September 2002.

[11] Jepson, A. D., and Fleet, D. J. Phase singularities in scale-space. American Journal of Physics, 67(1):55–60, 1999.

[12] Fleet, D. J., and Jepson, A. D. Stability of phase information. In Proceedings of the IEEE Workshop on Visual Motion, vol. 4, no. 1, p. 1, October 1991.

[13] Lucas, B. D., and Kanade, T. An iterative image registration technique with an application to stereo vision. In Proceedings of the 7th International Joint Conference on Artificial Intelligence, pp. 674–679, 1981.

[14] Balazadeh Bahar, M. R., and Karimian, G. High performance implementation of the Horn and Schunck optical flow algorithm on FPGA. In Proceedings of the 20th Iranian Conference on Electrical Engineering, pp. 736–741, 2012.

[15] Zhang, Z. A flexible new technique for camera calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11):1330–1334, November 2000.

[16] Try P , Gebhard M .A Vibration Sensing Device Using a Six-Axis IMU and an Optimized Beam Structure for Activity Monitoring[J].Sensors, 23(19):24, 2023.

CFD Analysis of Hydraulic Oscillator for Flow Separation Control

Ryuto Yoshida*, Wataru Kobayashi*, and Seiichiro Tanibuchi*

*Graduate School of Science and Engineering, Okayama University of Science, Okayama, Japan
*E-mail: r23smt2mn@ous.jp, w-kobayashi@ous.ac.jp, r24smh1jt@ous.jp

Abstract

The objective of this study is to control the separation flow of ship. Previous studies on ship separation control has carried out passive methods and active methods. These studies have some issue for example energy and another conditions. There, this study focuses on fluidics as one of the passive methods, applied to oscillators using switching elements. However, previous study has been clarified that a bit error between design to actual of chamber and orifice make difference oscillation. Therefore, this study proposes integrated type oscillator that it integrates chamber, orifice and switching element. Also, this study analyzes the reduction of fluid resistance around the ship using this oscillator. Based on this research background, the study first analyzes the proposed oscillator's characteristics, focusing on parameters like orifice and chamber volume, which influence the oscillation period and frequency. By adjusting the output port diameter (1.4 to 2.8 mm), the oscillation period inside the chamber is halved. Also, increasing the chamber volume raises the average pressure and the maximum pressure difference during oscillation. However, the effectiveness of this approach in reducing fluid resistance around the ship remains unclear. To investigate further, this study created a ship model, and separation behavior is analyzed with the oscillator installed.

Keywords: Flow separation control, CFD, Hydraulic oscillator, Fluidics

1 Introduction

In Japan, maritime transportation plays a crucial role in international trade, with ships accounting for 99% of overseas cargo transport [1]. Therefore, improving the transport efficiency of ships is essential and various studies have been conducted to improve this efficiency. These include optimization of operation routes [2], enhancement of engine power of ships [3], and reduction of fluid resistance around ships. In optimization of operation routes, optimization of operation routes has been proposed to quantitatively estimate changes in route selection when the ship size increases. This estimation uses simulation models that reproduce the current container cargo volume. In studies for fluid resistance reduction, the focus has been on the environment around the ship and the reduction of fluid friction [4]. One focusing on the environment around the ship is the development of special ship bottom paint to reduce resistance caused by bio-fouling. Thus, various approaches aim to improve transport efficiency. Among these, this study focuses specifically on methods to reduce fluid resistance around ships.

1.1 Fluid separation and turbulence

This study focuses on methods to reduce fluid resistance. Therefore, attention is given to fluid separation, which is one of the causes of fluid resistance in ships. As shown in Figure 1, fluid separation refers to the phenomenon in which a fluid flowing along the surface of an airfoil gradually decreases in velocity due to viscous friction with the surface. As the fluid moves toward the rear of the object, it encounters increasing difficulty in maintaining its flow against the opposing pressure. Eventually, the fluid separates from the surface and flows away. Moreover, the flow with reduced velocity may even acquire a reverse velocity, changing direction due to friction with the separated flow.

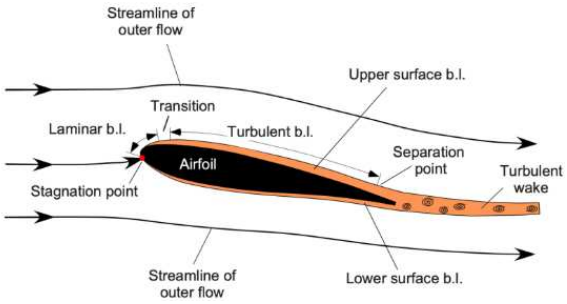


Figure 1: Flow patterns around airfoil [5]

This separation results in a pressure difference between the front and rear of the object, which generates resistance as the

object moves forward [5] [6]. This phenomenon-induced resistance is considered an important issue in maritime vessels. Also, the point where the fluid velocity becomes zero and the flow separates from the object's surface is called the separation point. To reduce fluid resistance, a method is to intentionally disturb the laminar flow near the separation point, which helps to move the separation point downstream [7]. For example, aircraft wings use flaps as shown in Figure 2, and propellers positioned at the rear of the aircraft create a suction effect to move the separation point [8]. In ships, one method to reduce frictional resistance involves introducing microbubbles along the ship surface as shown in Figure 3. These microbubbles create a lubricating layer between the ship and the surrounding water, effectively decreasing viscous drag and improving fuel efficiency. It has been demonstrated that the use of microbubbles can reduce wall shear stress by up to 30% [9]. However, to achieve a more efficient reduction in frictional resistance, ships need to specifically design built-in microbubble generators. Additionally, it is essential to assess the effects of microbubble generator in the propeller area to ensure optimal performance. Beyond passive techniques, active control methods utilize external power sources, such as electricity, to reduce fluid resistance and improve transport efficiency. However, The efficiency improves only by the amount of energy subtracted from the power due to the reduction of fluid resistance. In some cases, these methods may offer only marginal benefits or even lead to increased energy consumption.

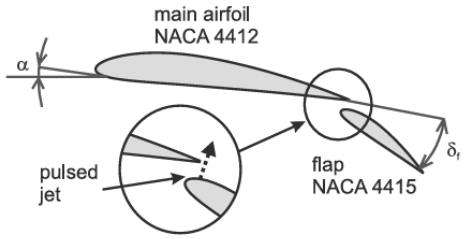


Figure 2: Schematic diagram of flap surface moving device [8]

Figure 3: Overview of ship and built-in micro bubble generator [9]

In contrast, passive methods, which reduce fluid resistance without using energy, have been gaining much more attention. Some of them include the dimples on golf balls as shown in Figure 4, and vortex generators on the surfaces of cars and airplanes as shown in Figure 5 [10]. The dimples on the ball create turbulence near the separation point, preventing flow separation during flight and enhancing the ball flight distance. In related study, it shows dimpled golf balls experience lower drag due to delayed flow separation at the rear end, allowing them to travel farther. The dimples cause the boundary layer to transition from laminar to turbulent flow, similar to vortex generators. This turbulent flow stays attached longer, delaying flow separation and reducing form drag. As a result, total drag, which includes form drag and skin friction, is minimized, improving the ballâs distance [11] .

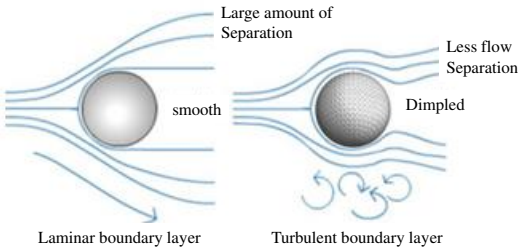


Figure 4: Simulation results of velocity vector distribution of golf ball (left: w/o dimple, right: w/ dimple) [11]

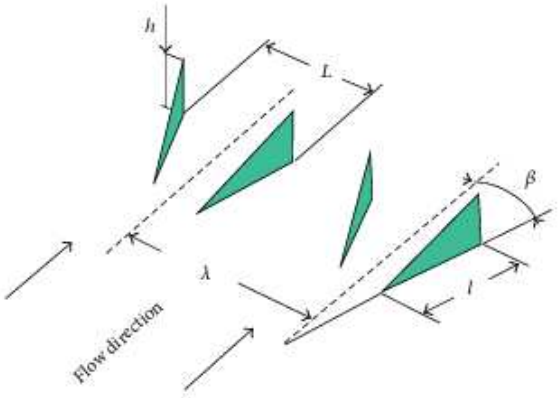


Figure 5: Vortex generator geometry [10]

1.2 Fluidics

One of the passive methods is an application of fluidics. Fluidics is a technology that controls fluids by utilizing phenomenon generated by the fluid itself, without mechanical moving parts. Since it controls the fluid solely through its own properties, it has been extensively studied for its potential use in environments where electronic devices are difficult to apply. Figure 6 shows an application of fluidics for medical fields, which called micro fluidics [12] [13].

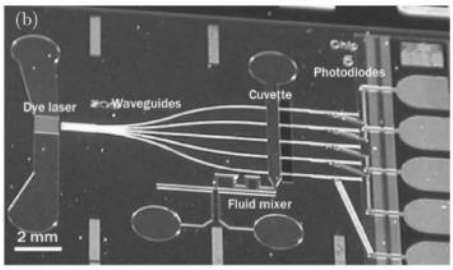


Figure 6: Micro fluidics [13]

Fluidics are highly reliable and durable because they do not have any mechanical moving parts. However, due to the constant flow, they have a large energy loss and lower response

time compared with typical electric systems [14]. Fluidic elements that can perform some functions such as OR, AND, flip-flops, and Switching have been developed [15]. By performing drug synthesis on microchannels carved into silicon or glass, the amount of samples and reagents can be significantly reduced, offering benefits such as time savings and reduced waste disposal. On the other hand, due to the phenomenon occurring in microscopic environments, challenges include the difficulty of measuring temperature and flow rate, as well as the need for technology to inject fluids into the microchannels [16]. Although the devices are manufactured by inserting channels, it has become possible to produce devices, including channels, using 3D printers with advancements in technology. Additionally, improvements in the precision of channel carving enable the creation of smaller channels and components. As a result, manufacturers can decrease costs, and they can also achieve the miniaturization of components. Thus, fluidics is expected to have various elements and applications, and useful for passive detachment suppression for ships as well.

1.3 Fluidic switching elements

In this study, the focus is on a fluid switching device that can change the flow path using only fluid. Also, we are considering a valveless switching system with the aim of realizing a purely aqua-driven system in the future. However, as demonstrated in this paper, this oscillator also has the potential to reduce fluid resistance. Therefore, to reduce the fluid resistance of ships, this study is developing an oscillator that is driven solely by fluid, as described below. First, this section presents the structure of the fluid switching element and describes its operating principle. Then, the effective design revealed by previous research is discussed [17]. A fluidic switching element has three types of ports: an input port, a control port, and an output port, as shown in Figure 7. The roles of these ports are as follows: the input port supplies pressure, the control port determines the direction of output, and the output port is responsible for the final output of pressure and flow velocity. After the control port determines the output direction of the main flow, the Coanda effect draws the main flow toward the side wall due to the negative pressure generated between the side wall and the main flow. This mechanism helps maintain the flow path. As shown in Figure 8, the operating principle

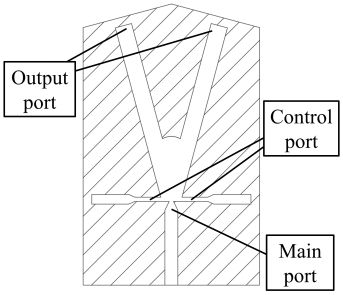


Figure 7: Structure of fluidic switching element

of the fluid switching device is explained as follows: 1) The main jet flow entering from the input port flows toward the left and right output ports. 2) By applying flow from the left and right control ports, the main jet is deflected toward the sidewalls. 3) The deflected main jet maintains its direction due to the Coanda effect. Through these operations, switching is made possible using only fluid.

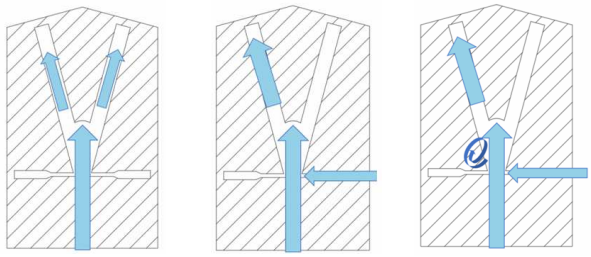


Figure 8: Operating principle of switching element

In the optimal shape for switching, however, various factors such as the diameter of the split, the distance to the split, the diameter of the input port, and the diameter of the output port all have an impact. Therefore, in previous studies, analyses were conducted on the switching and maintenance of flow paths, clarifying the effects of the design parameters of switching elements and confirming the feasibility of a switchable design. This analysis focuses on three parameters shown in Figure 9 (L_s: split distance, H: offset distance, and θ: sidewall inclination angle) which are considered to be important factors influencing the switching of flow channels and the maintenance of flow direction.

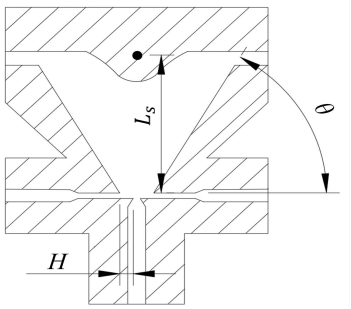


Figure 9: Structure of fluidic switching element

2 Fluidic oscillator using switching element

Fluidic switching elements, which is previously analyzed in our studies , have been applied in various fields. One of them is the fluidic oscillator, which is possible to oscillate by connecting a volumetric chamber and an orifice to a fluid switching element as shown in the figure 10. In a related study, a fluidic oscillator was applied to artificial hearts as a medical device. The study noted that periodic pressure fluctuations strongly influence the frequency and amplitude of pressure fluctuations, especially when the orifice diameter is reduced. On the other hand, the study pointed out a concern that, depending on the structure of the element, pressure may switch to the opposite direction before reaching the calculated pressure [18]. This study investigates the use of this fluidic oscillators to prevent fluid separation around the ship.To understand the characteristics of the fluidic oscillator, This study constructed an experimental setup. As shown in Figure 11, the experimental setup

consists of a fluidic switching element, a pressure chamber, and a pressure sensor (PSE560-02, SMC Corporation). The supply pressure is set to 280 kPa corresponding to the tap water pressure in our laboratory. Also, this study adjusted the orifice diameter in the center of the oscillator by controlling the switching element's control port. Figure 12 shows the experimental results, showing oscillation frequencies of up to 150 Hz depending on the pressure chamber volume and orifice diameter. However, both the frequency and amplitude of pressure fluctuation exhibited instability. This instability suggests that parameters such as the pressure chamber volume, tube volume, and orifice diameter have a significant impact on the oscillation characteristics. Small discrepancies in the design can strongly influence the oscillation behavior.

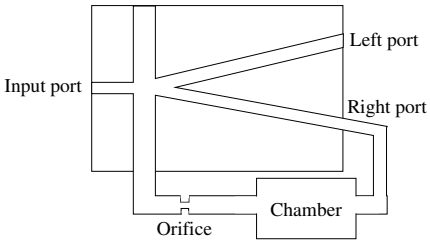


Figure 10: Schematic diagram of conventional oscillator

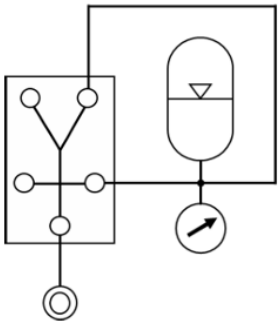


Figure 11: Experiment setup for conventional oscillator

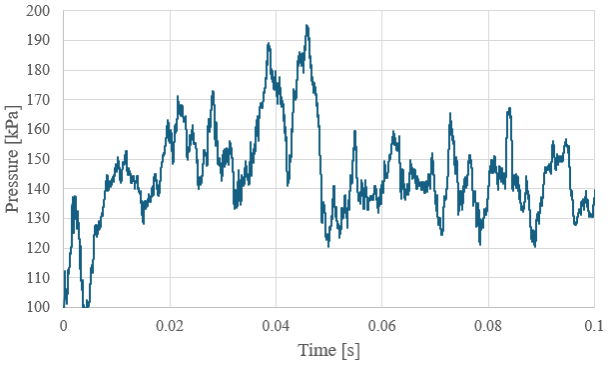


Figure 12: Measured pressure response in oscillator chamber

2.1 Integrated type fluidic oscillator

As mentioned in the previous section, even slight design discrepancies can make it difficult to operate the oscillator when manufactured according to theoretical specifications. To address this issue, this study proposes an integrated type fluidic oscillator that takes into account of the effects of the chamber, tubes, and orifices. Figure 13 shows the design of the proposed integrated type fluidic oscillator.

Building upon a conventional oscillator design, the configuration consists of a fluidic switching element, left and right pressure chambers, and an orifice connecting the right pressure chamber to the right control port. The proposed oscillator has following two key characteristics: 1) the right pressure chamber connects to the control port, and 2) the pressure chamber and orifice are integrally molded. By integrally molding the pressure chamber and orifice with the fluidic element, the impact of these components on oscillation can be minimized. Figure 14 shows the operating principle of the oscillator. The process is as follows: 1) Fluid is supplied from the input port and flows into both the left and right pressure chambers, 2) The internal pressure in the pressure chambers, 3) Flow occurs from the right control port, directing the main flow to the left pressure chamber, 4) Coanda effect occurs along the side wall of the oscillator, maintaining the flow direction, 5) The internal pressure of the left pressure chamber increases and the flow to the right pressure chamber decreases, 6) The internal pressure of the right pressure chamber decreases and also the flow from the right control port decreases, 7) The main flow switches from the left to the right pressure chamber. By repeating this process, passive oscillation can be achieved without any external power sources.

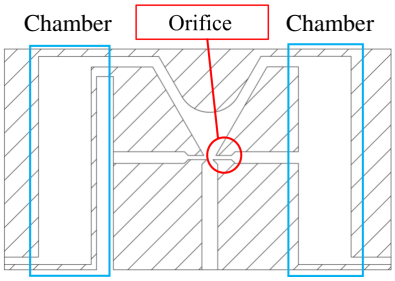


Figure 13: Structure of integrated type fluidic oscillator

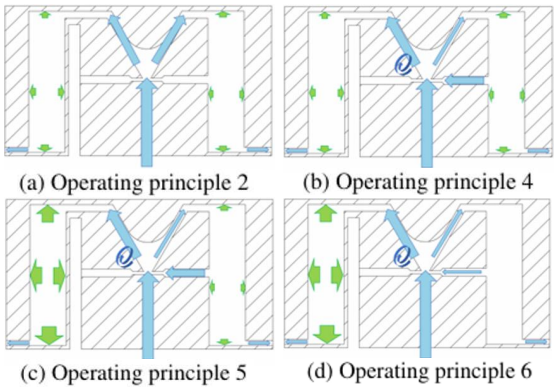


Figure 14: Operating principle of proposed oscillator

2.2 Analysis of oscillator: Initial conditions

This section evaluates the oscillation performance of the proposed oscillator and analyzes the effect of design parameters

of the oscillator for its oscillation behavior. The analyzed oscillator has dimensions of 77.5 mm in length and 43 mm in width. Table 1 lists the design parameters while Figure 15 visually illustrates these parameters. In these analyses, internal pressure of chambers p_1, p_2 and flow velocity of output ports u_1, u_2 are measured. The first analysis focuses on the output port diameters, T_R and T_L, of the oscillator, as these diameters influence the oscillation period. Another key parameter is the thickness D, which affects the amplitude of pressure fluctuation. Regarding the inlet flow conditions, both the output port and the left control port of the oscillator open to the atmosphere, and supply pressure for the oscillator is 300 kPa. Additionally as shown in Figure 16, the mesh number of the oscillator set to 640,000, with further refinement by a factor of 1/8 at the center of the oscillator. Also, this study needs to observe a detailed understanding of the internal flow within the oscillator, particularly phenomena such as the coanda effect. For this reason, this study adopted a CFD-based approach.

Table 1: Design parameters of proposed oscillator for analysis

Variable	Symbol	Value
Chamber width	W_L and W_R	15 mm
Output port diameter	T_L and T_R	1.4 mm
Oscillator thickness	D	10 mm

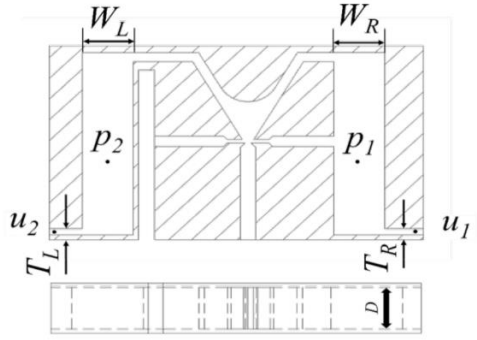


Figure 15: Design parameters of proposed oscillator

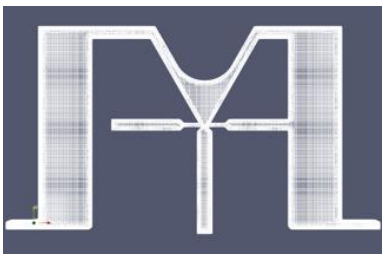


Figure 16: Analysis model of proposed oscillator

Figures 17 and 18 shows the analysis results for the design parameters listed in Table 1. Figure 17 describes the velocity distribution at 0.1 and 0.44 s, respectively. The flow velocity from the input port can switch from the left pressure chamber to the split section. Figure 18 shows the time response of p_1 and p_2. The average pressure in the pressure chamber is 90 kPa and the oscillation period is 0.1 s, and the maximum pressure difference is approximately 25 kPa. Additionally, from Figures 18, it was confirmed that oscillation is possible, as the internal pressure is fluctuating at the left output port.

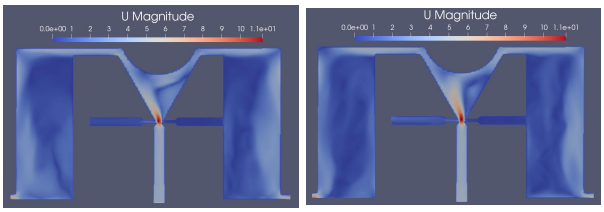


Figure 17: Velocity distribution of proposed oscillator (left: 0.1 s, right: 0.44 s)

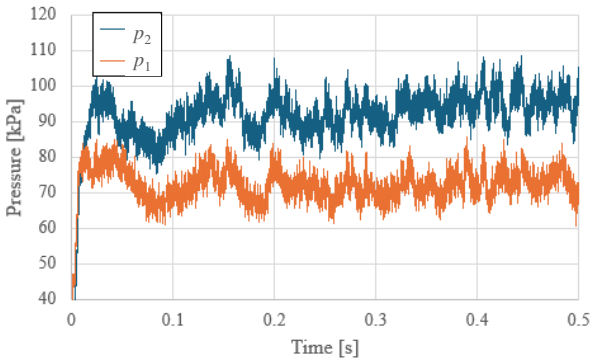


Figure 18: Pressure response of chambers in proposed oscillator

2.3 Analysis of oscillator: Output port diameter

First, the output port diameter of the oscillator, which affects the oscillation period, is analyzed. Figure 19 shows the results when W_L is changed from 1.4 to 2.8 mm. Top-left to bottom-right of Figure 19 show the velocity distribution at 0.1, 0.3, 0.4 and 0.5 s respectively. As shown in Figure 19, the flow velocity oscillates from the input port towards both the left and right pressure chambers.

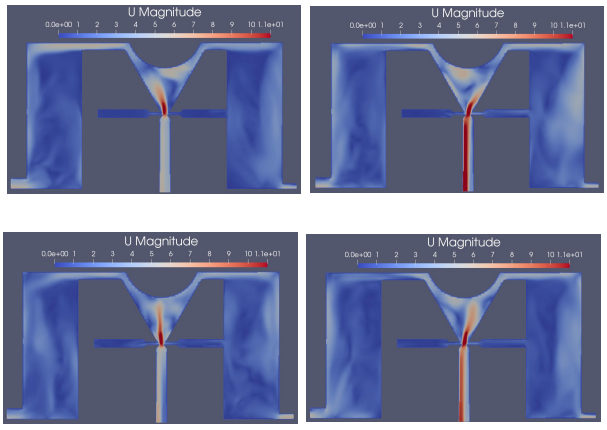


Figure 19: Velocity distribution of proposed oscillator (top-left: 0.1 s, top-right: 0.3 s, bottom-left: 0.4 s, bottom-right: 0.5 s)

Figure 20 describes the time response of internal pressure of the chambers at p_1 and p_2 when T_L is 1.4 and 2.8 mm. Fig-

ure 20 shows that the average pressure inside the left chamber is approximately 30 kPa at the case that the left output port diameter is 2.8 mm. Also, by performing analysis of frequency components using FFT (Fast Fourier Transform), the frequency component was confirmed to be 20 Hz, as shown in the figure 21. This means 60 kPa reduction in average pressure compared to the previous case with the diameter of 1.4 mm, and the oscillation period is halved. Oscillations between the left and right pressure chambers and the center split can be observed when the pressure difference between the left and right chambers reversed, and when the pressure difference decreased to 15 kPa. Therefore, it can be concluded that the output port diameter significantly influences the oscillation period and plays a crucial role in determining the pressure difference between the left and right pressure chambers.

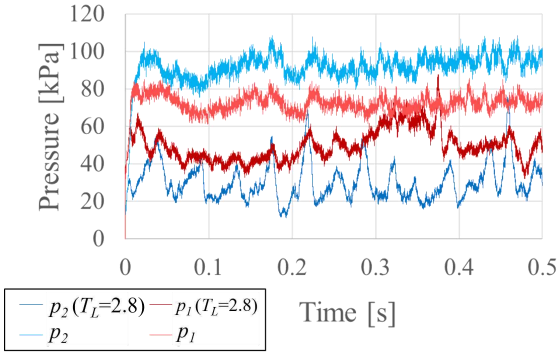


Figure 20: Comparison of internal pressure of chambers

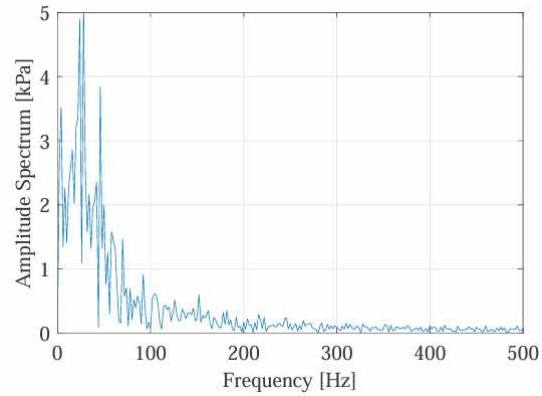


Figure 21: Frequency component analysis result

Figures 22 and 23 shows the results when T_R is changed from 1.4 to 2.8 mm. Figure 22 describes the velocity distribution at 0.1, 0.19, 0.31, and 0.42 s, while Figure 23 shows the time response of pressure at p_1 and p_2. As seen in Figure 22, the flow velocity from the input port primarily oscillates towards the left pressure chamber and the center split. Figure 23 indicates that p_1 increases significantly by 0.03 s, after which oscillatory fluctuations are observed. This is thought to be due to the decreased flow towards the left pressure chamber and the relatively increased flow towards the right pressure chamber. These results suggest that varying the diameter of the right output port T_R affects the oscillation.

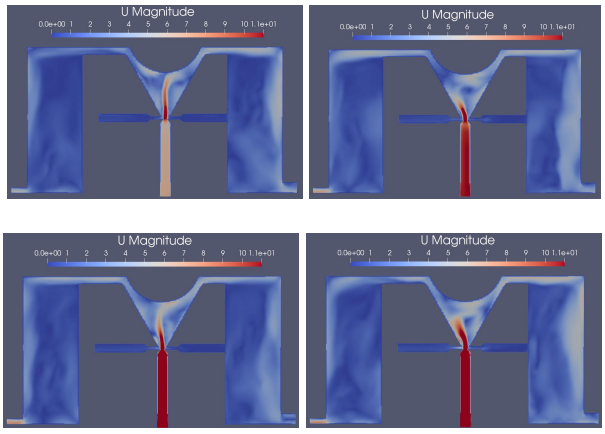


Figure 22: Velocity distribution of proposed oscillator with T_L: 2.8 mm (top-left: 0.1 s, top-right: 0.3 s, bottom-left: 0.4 s, bottom-right: 0.5 s)

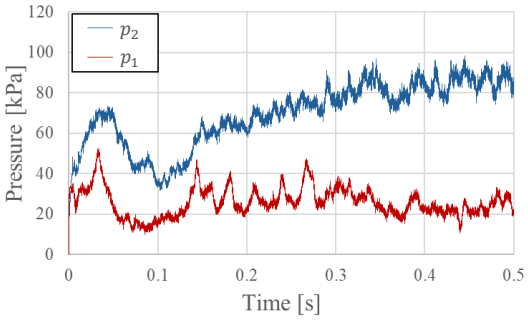


Figure 23: Pressure response of chambers in proposed oscillator ($T_R = 2.8$ mm)

2.4 Analysis of oscillator: Chamber volume

Next, this study focused on the chamber volume of the oscillator, which is thought to affect the amplitude of pressure fluctuation, and conducted analysis by varying several parameters. Figures 24 and 25 show the analysis results when the height D is changed from 10 to 15 mm. Figure 24 presents the flow velocity distribution inside the element at 0.1 and 0.3 s, respectively. Figure 25 shows the time response of pressure p_1 and p_2. As shown in Figure 24, the flow velocity from the input port moves towards the left chamber. On the other hand, no clear switching of flow direction to the right chamber is observed. In Figure 25, the average pressures in the left and right chambers are 75 and 60 kPa, respectively. The pressure difference between the maximum and minimum pressures is small and this indicates that the volume of the chambers significantly affects the amplitude of pressure fluctuation.

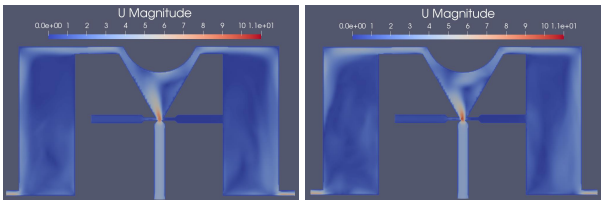


Figure 24: Velocity distribution of proposed oscillator with D: 15 mm (left: 0.1 s, right: 0.3 s)

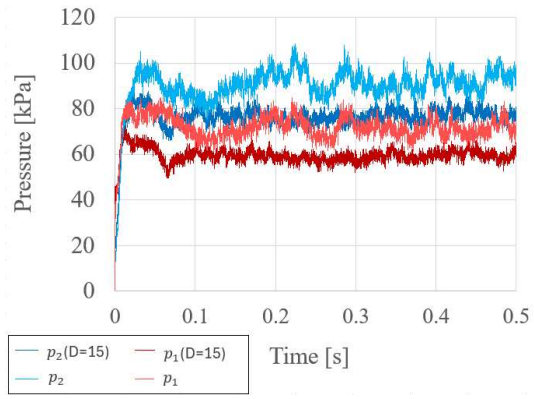


Figure 25: Pressure response of chambers in proposed oscillator (D =10, 15 mm)

Figures 26 and 27 show the analysis results when the pressure chamber height D is changed from 10 to 5 mm. Figure 26 presents the flow velocity distribution inside the element at 0.1 and 0.45 s, respectively. Figure 27 shows the time response of pressure p_1 and p_2. As shown in Figure 26, the flow continues towards the left chamber. The average pressure in the left chamber is 120 kPa, and the maximum differential pressure is 30 kPa. This clearly indicates that the volume of the chambers affects the amplitude of pressure fluctuation. Regarding the periodic oscillation, the period is faster compared to the period of previous oscillator with the height of 10 mm. This phenomenon is due to the reduced time required to accumulate pressure, as the volume of the chamber is decreased. From the results, it is evident that the diameter of the left output port in the oscillator affects the oscillation period.

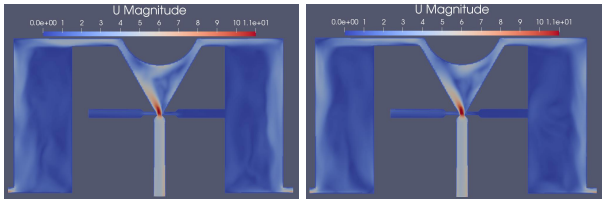


Figure 26: Velocity distribution of proposed oscillator with D: 5 mm (left: 0.1 s, right: 0.45 s)

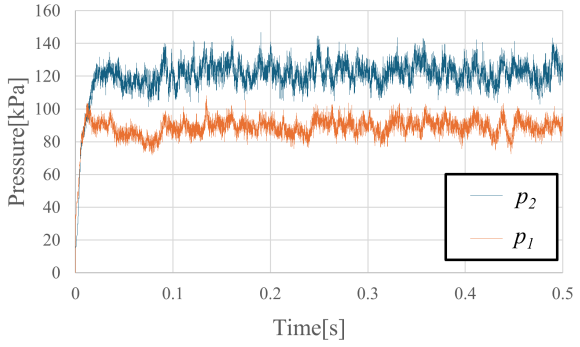


Figure 27: Pressure response of chambers in proposed oscillator (D = 5 mm)

3 Analysis of flow separation around ship

In this chapter, this study created a simple ship model and then CFD analysis shows its behavior when turbulent flow generated by the proposed oscillator is introduced around the ship model. This will help clarify the effective oscillation frequency for separation control of ships. The proposed oscillator is integrated at the separation point on the model shown in Figure 28. Figure 29b shows the created ship model, and its shape is taken from the A-A cross-section in Figure 29a. The model is 3.5 m in length, 1.6 m in width and 0.5 m in height.

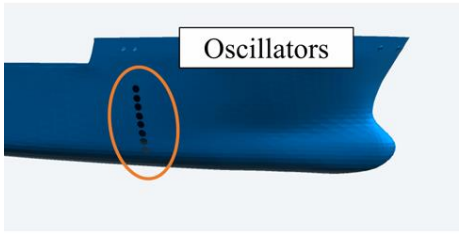
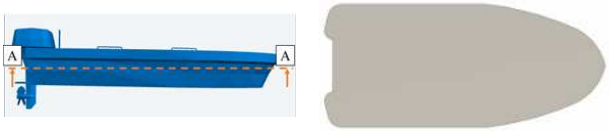


Figure 28: Image of oscillators arrangement at ship



(a) Side view of ship *(b) Cross section view*

Figure 29: Analysis model of ship

3.1 Analysis of separation control without oscillator

To clarify the optimal oscillation for separation control, the first analysis is conducted on separation control for the conventional shape without the integrated oscillator. In this analysis, the applied turbulence model and sampling period are the same as those in the previous section. The mesh resolution is 640,000, with an additional 1/8 division around the separation point and a further 1/64 division around the output port of the oscillator. While this may seem relatively small, the mesh resolution around the injection point was additionally increased, and the analysis was performed again ant then no significant differences were observed. Therefore, this study believes that this simulation has sufficient validity. The cruising speed of the ship is 20.5 m/s. Figure 30 shows the analysis results and illustrates the velocity distribution in the X-direction which means heading direction of the ship. It is confirmed that the flow velocity around the ship decreases and spreads from near the ship side toward the rear of the ship. Additionally, the result of supplying a steady flow from the oscillator's output port, meaning it generates a steady flow instead of a periodic turbulent flow is shown in Figure 31. It illustrates the flow velocity distribution in the X-direction, where a negative flow velocity indicates movement toward the stern. The flow rate from the oscillator is 20.5 m/s and the oscillator applies only to the starboard side of the ship (bottom side), with no oscillator on the port side (top side). The decrease in flow velocity is mitigated on the starboard side compared to conventional flow, as shown in Figure 31, but no significant reduction in

flow velocity occurs at the rear of the ship. Even with steady flow applied, separation fails to be sufficiently controlled.

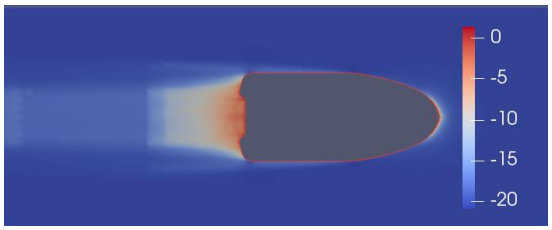


Figure 30: Velocity distribution (w/o oscillator)

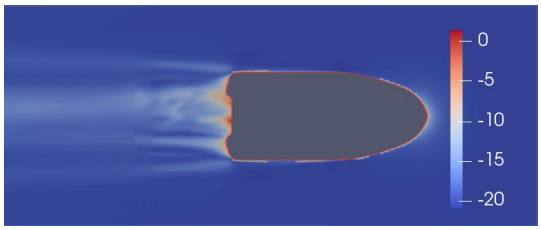


Figure 31: Velocity distribution (w/ steady flow)

3.2 Analysis of separation control with oscillator

In this section, to achieve flow separation control around the ship, the integrated oscillator is installed into the ship model, and analyses are conducted to evaluate its effect on flow separation. The optimal oscillation frequency for separation control is determined. This study draws on research related to aircraft wings, which encounter similar separation issues. In the studies, the concept of reduced frequency F^+, commonly used in aircraft wing separation control, is introduced. It is defined by following eq. (1)

$$F^+ = \frac{fl}{U} \tag{1}$$

where f is the oscillation frequency, l the characteristic length, and U the flow velocity. By adjusting this frequency, the effective oscillation frequency for separation control can be determined. However, the optimal reduced frequency varies depending on objectives and is often determined experimentally. Therefore, while referencing reduced frequency, this study investigates separation control around the ship using an integrated oscillator, examining the effects of oscillation frequencies of 1, 10, and 100 Hz. The ship model, analysis conditions, and flow boundary conditions are as same as those described above. First, an analysis is conducted with oscillation frequency of 1 Hz applied from the oscillator. Figure 32 illustrates the X-direction flow velocity distribution. The calculated F^+ at 1 Hz is 0.17, which is lower than the conventional optimal value for related studies. No significant difference in flow velocity is observed compared to the previous analysis shown in Figure 31. In addition, the maximum flow velocity around the port side of the ship is 16 m/s, while on the starboard side with oscillator, the maximum velocity remains approximately 16 m/s. These results suggest that a low F^+ is insufficient for effective flow separation control around the ship. Figure 33 illustrates the analysis results for oscillation frequency of 10 Hz along with the flow velocity distribution in the X-direction. The calculated F^+ at 10 Hz is 1.7, which corresponds to the conventional optimal value reported in related studies. Compared to the previous analysis shown in Figure 31, a reduction in flow velocity is observed and the maximum flow velocity on the starboard side of the ship is approximately 18 m/s. These results indicate that the given F^+ is effective for controlling flow separation around the ship. Figure 34 illustrates the analysis results for oscillation frequency of 100 Hz along with the flow velocity distribution in the X-direction. The calculated F^+ at 100 Hz is 17, which is larger than the conventional optimal value reported in related studies. As shown in Figure 34, however, the decrease in flow velocity is effectively mitigated with the maximum velocity of 18.5 m/s. The area where the flow velocity decreases near the rear of the ship seems to be smaller than the area of the previous analysis in Figure 31. Furthermore, a reduction in the area where the flow velocity significantly decreases in the wake of the ship can be observed. The analyses presented in this chapter demonstrate that the reduced frequency F^+ significantly influences flow separation control in ships. The observed discrepancy is likely attributable to the omission of fluid property considerations, such as viscosity differences between air and water, in the formulation of the reduced frequency. Further investigations that incorporate fluid-specific parameters are required to refine the flow separation control strategy for this application.

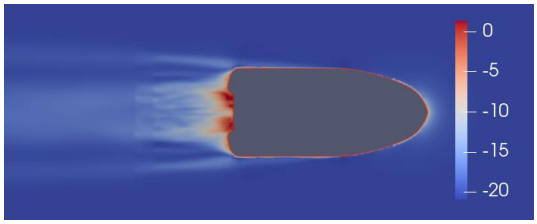


Figure 32: Simulation result of flow velocity of ship (1 Hz)

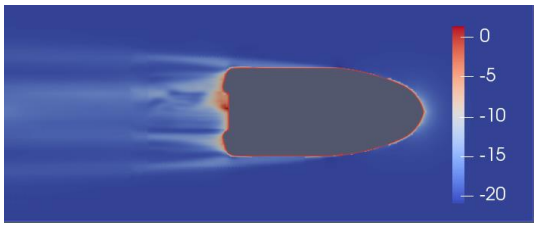


Figure 33: Simulation result of flow velocity of ship (10 Hz)

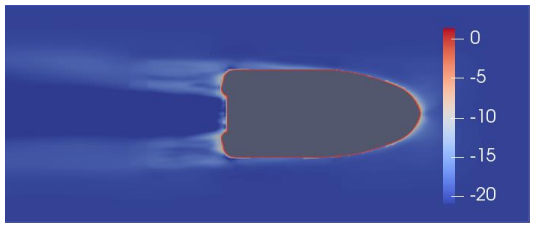


Figure 34: Simulation result of flow velocity of ship (100 Hz)

4 Conclusion

This study focuses on fluidic switching elements that can switch fluid paths using only fluid, with the aim of developing an oscillator that can be driven solely by fluid in order

to reduce fluid resistance in ships. In previous fluidic oscillators using switching elements, oscillation occurred without external input due to the influence of chambers, tubes, and orifices. However, achieving the desired oscillation period and frequency was difficult due to small design errors in these components. Therefore, this study proposes an integrated fluidic oscillator that combines the chamber and tube into a single unit. Also, to evaluate the proposed oscillator and clarify the effects of various design parameters on oscillation, an analysis was conducted. Results showed that changing the diameter of the output port of each oscillator modifies the oscillation period, and adjusting the chamber volume changes the amplitude of pressure fluctuation. This enables the design of an oscillation period and frequency effective in reducing the fluid resistance of ships. While there have been numerous studies on the reduction of resistance through fluidic separation, there are few examples of study focused on ships. Therefore, this study aims to clarify the oscillations that are effective in reducing the fluid resistance of the ship. First, a simplified ship model was created, and through analysis, it was confirmed that separation occurs. For this analysis, this study focused on separation at a specific height on the ship. Next, this study assumed the installation of the oscillator at a part of the ship model and applied oscillations under specific conditions to analyze the separation behavior. The results of the analysis revealed that an oscillation period in the range of 10-100 Hz most effectively control separation. As future work, it is necessary to investigate the separation behavior at different heights of the ship and compare it with the current results. Furthermore, based on these analysis results, this study needs to design an oscillator with an optimal oscillation period for separation control and conduct a further analysis of the oscillation when input conditions are altered. Additionally, since this study mainly focuses on analysis, it will be important to verify the consistency with experimental results.

Nomenclature

Designation	Denotation	Unit
D	Oscillator thickness	mm
f	Oscillation frequency	Hz
F^+	Reduced frequency	-
H	Offset distance	mm
l	Characteristic length	m
L_s	Split distance	mm
p_1	Right chamber pressure	kPa
p_2	Left chamber pressure	kPa
T_L	Left output port diameter	mm
T_R	Right output port diameter	mm
θ	Sidewall inclination angle	degree
U	Flow velocity	m/s
u_1	Right outport pressure	kPa
u_2	Left outport pressure	kPa
W_L	Left chamber width	mm
W_R	Right chamber width	mm

Acknowledgments

Part of this work was supported by JSPS KAKENHI Grant Number JP21K14133.

References

[1] Trade Ministry of Economy and Industry. White paper on international economy and trade [outline]. 2020.

[2] Ryuichi SHIBASAKI and Tomihiro WATANABE. Future forecast of trade amount and international cargo flow in the apec region: An application of trade-logistics forecasting model. *Asian Transport Studies*, 2(2):194–208, 2012.

[3] Simone Lion, Ioannis Vlaskos, and Rodolfo Taccani. A review of emissions reduction technologies for low and medium speed marine diesel engines and their potential for waste heat recovery. *Energy Conversion and Management*, 207:112553, 2020.

[4] Liang Li, Heting Hong, Jingyi Cao, and Yange Yang. Progress in marine antifouling coatings: Current status and prospects. *Coatings*, 13(11):1893, 2023.

[5] J. G. Leishman. Introduction to aerospace flight vehicles. 2022.

[6] William S. Vorus. A theory for flow separation. *Journal of Fluid Mechanics*, 132:163–183, 1983.

[7] Dennis Bushnell. Separation control: Review. *Journal of Fluids Engineering*, 113, 04 1991.

[8] Ralf Petz and Wolfgang Nitsche. Active separation control on the flap of a two-dimensional generic high-lift configuration. *Journal of Aircraft - J AIRCRAFT*, 44:865–874, 05 2007.

[9] Ali Dogrul, Yasemin Arıkan Özden, and Fahri Çelik. A numerical investigation of air lubrication effect on ship resistance. *SMOOTH 2010*, 01 2010.

[10] Shi-bo Tao, Aiping Tang, Da-bo Xin, Ke-tong Liu, and Hong-fu Zhang. Vortex-induced vibration suppression of a circular cylinder with vortex generators. *Shock and Vibration*, 2016:1–10, 01 2016.

[11] Shabir Grover, B B Arora, Vaibhav Khanna, Tushar Kaushik, and Akhilesh Arora. International journal of advanced production and industrial engineering innovative design computational modelling analysis on drag reduction of bluff body using dimples. *International Journal of Advanced Production and Industrial Engineering*, pages 4–11, 01 2017.

[12] Vigneswaran Narayanamurthy, Z. E. Jeroish, K. S. Bhuvaneshwari, Pouriya Bayat, R. Premkumar, Fahmi Samsuri, and Mashitah M. Yusoff. Advances in passively driven microfluidics and lab-on-chip devices: a comprehensive literature review and patent analysis. *RSC Adv.*, 10:11652–11680, 2020.

[13] Henrik Bruus. *Theoretical microfluidics*, volume 18. Oxford university press, 2007.

[14] Kin Fong Lei. Microfluidic systems for diagnostic applications: A review. *Journal of laboratory automation*, 17(5):330–347, 2012.

[15] Animesh Pratap Singh, Marcel Tintelott, Elmira Moussavi, Sven Ingebrandt, Rainer Leupers, Xuan-Thang Vu, Farhad Merchant, and Vivek Pachauri. Logic operations in fluidics as foundation for embedded biohybrid computation. *Device*, 1(6), 2023.

[16] Microfluidics for nano-drug delivery systems: From fundamentals to industrialization. *Acta Pharmaceutica Sinica B*, 13(8):3277–3299, 2023.

[17] Masashi YOKOTA Ryuto YOSHIDA, Wataru KOBAYASHI. Analysis of design parameters of aqua drive fluidics using openfoam. *The Japan Fluid Power System Society*, 2024 (In Japan).

[18] Barna Becsek, Leonardo Pietrasanta, and Dominik Obrist. Turbulent systolic flow downstream of a bioprosthetic aortic valve: velocity spectra, wall shear stresses, and turbulent dissipation rates. *Frontiers in physiology*, 11:577188, 2020.