Chapter 7
Evidential Thresholding in Sparse Bayesian Learning for Nonlinear Dynamical System Identification

Jiawei Jian, Jerome Antoni, and Konstantinos Gryllias

Abstract Sparse regression provides a data-driven approach for the identification of nonlinear dynamical systems in cases where the prior knowledge of the nonlinear characteristics is unavailable. Specifically, a parsimonious governing equation can be distilled from a large dictionary of potential candidate components by imposing sparse regularization or thresholding techniques to the coefficients in a linear regression model. Owing to the capability of uncertainty quantification, sparse Bayesian learning (SBL) has been extensively studied for the identification of nonlinear dynamical systems. However, due to the interference of noise, existing SBL methods usually suffer from overfitting induced by redundant components. Traditional thresholding techniques may fail to eliminate these irrelevant components due to their inherent limitation, namely, evaluating the significance of components based solely on the magnitudes of their coefficients. As a result, redundant components with relatively large coefficients are difficult to eliminate, while the correct components with small coefficients (such as damping terms) are likely to be falsely discarded. To address this problem, this paper proposes a thresholding method, referred to as evidential thresholding, within a Bayesian framework. Instead of removing components with small coefficients, the proposed evidential thresholding approach evaluates the decoupled contribution of each candidate component to the evidence, and then eliminates those with negligible evidential contributions. Additionally, unlike traditional thresholding techniques relying on an unbounded threshold value, the threshold value in the proposed approach is naturally bounded within [0, 1], making it easier to define. From another perspective, the proposed approach can be considered as a component-wise extension of Bayesian model selection, where the components with relatively large evidential plausibility are selected. The proposed evidential thresholding method is subsequently integrated into a Gibbs sampler-based SBL (GB-SBL) framework for the identification of nonlinear dynamical systems. Numerical and experimental examples are investigated to demonstrate the effectiveness of the proposed method. The results show that it outperforms state-of-the-art approaches, namely, SINDy, relevance vector machine (RVM), and GB-SBL without evidential thresholding.

Keywords Sparse bayesian learning · System identification · Evidential thresholding · Gibbs sampler · Model selection

Introduction

Linear approximation is commonly employed to model the dynamical characteristics of engineering structures due to its simplicity and well-developed theoretical foundation. However, in practice, structures inevitably exhibit complicated nonlinear behavior, such as jumps, bifurcations, and limit cycles, which may greatly deviate the approximate linear models. Ignoring these nonlinear characteristics can lead to misinterpretations of structures. Consequently, precise modelling of nonlinear dynamics is important. To this end, nonlinear system identification has been widely studied over the past decades [1, 2].

Jiawei Jian · Konstantinos Gryllias
Department of Mechanical Engineering, KU Leuven, Celestijnenlaan 300, Leuven, 3001, Belgium
e-mail: jiawei.jian@kuleuven.be; konstantinos.gryllias@kuleuven.be

Jiawei Jian · Jerome Antoni
INSA Lyon, LVA, UR677, Villeurbanne, 69621, France
e-mail: jiawei.jian@kuleuven.be; jerome.antoni@insa-lyon.fr

Jiawei Jian · Konstantinos Gryllias
Flanders Make@KU Leuven, Belgium
e-mail: jiawei.jian@kuleuven.be; konstantinos.gryllias@kuleuven.be

© The Author(s), under exclusive license to River Publishers 2026

Kristin B. Zimmerman, PhD (eds.), *Nonlinear Structures & Systems, Vol. 1*, Conference Proceedings
of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814122

Numerous nonlinear system identification approaches have been developed. Most of them mainly focus on parameter estimation under a given mathematical equation model. However, nonlinearity characterization, namely determination of the equation form, is also a challenging and important task in practice. To address this problem, Brunton et al. proposed the sparse identification of nonlinear dynamics (SINDy), which learns parsimonious dynamical equations from a large dictionary of candidate components [3]. Due to its explainability and computational efficiency, SINDy and its derivatives have been widely used to simultaneously characterize nonlinearity and estimate parameters. However, SINDy is fundamentally a deterministic approach, which cannot quantify the uncertainty caused by noise. Moreover, its performance heavily relies on the predefined threshold value.

In recent years, probabilistic methods have gained increasing attention for uncertainty quantification in system identification. Notably, sparse Bayesian learning (SBL), which was originally proposed by Tipping [4], has been developed for nonlinear system identification [5, 6]. SBL-based approaches inherit the concept of dictionary learning, identifying the parsimonious model from numerous candidate components by imposing sparse priors on their coefficients. However, due to the interference of noise, existing SBL methods easily result in overfitted models induced by redundant components. To alleviate this issue, some efforts attempt to incorporate thresholding technique into SBL to produce sparser results [7, 8]. Nevertheless, in this case, the correct components with small coefficients, such as damping, are likely to be eliminated by thresholding, while some irrelevant components with relatively large coefficients may still retain in the model.

To address the problems, this paper proposes a probabilistic sparsity-promoting technique, referred to as evidential thresholding. Unlike conventional thresholding techniques, which remove candidate components solely according to the magnitudes of their coefficients, evidential thresholding evaluates the contribution of each component and eliminates those with small contributions. It is demonstrated that even in cases where the correct components possess small coefficients, their corresponding evidential contributions remain relatively large. Additionally, the proposed evidential thresholding approach can be considered as a component-wise extension of Bayesian model selection. In this paper, evidential thresholding is incorporated into a Gibbs sampler-based SBL (GB-SBL) method to illustrate its performance. However, it is noteworthy that it can be flexibly combined with general SBL approaches to promote sparsity.

The remainder of this paper is organized as follows: Section 2 firstly introduces the mathematical formulation of the sparse identification model for nonlinear dynamical systems, and then proposes the evidential thresholding method incorporated with the GB-SBL. A numerical example and an experimental benchmark are investigated in Section 3 to demonstrate the effectiveness of the proposed method. Finally, conclusions are drawn in Section 4.

Methodology

Sparse identification of nonlinear dynamical systems

The mathematical formulation of the sparse regression for nonlinear dynamical system identification is briefly introduced in this section. Consider an *m*-DOF nonlinear dynamical system as follows:

$$\mathbf{M}\ddot{\mathbf{y}}\left(t\right)+\mathbf{C}\dot{\mathbf{y}}\left(t\right)+\mathbf{K}\mathbf{y}\left(t\right)+\mathbf{P_N}\left(\mathbf{y}\left(t\right),\dot{\mathbf{y}}\left(t\right)\right)=\mathbf{P}\left(t\right)\#\left(1\right) \tag{1}$$

where $\mathbf{M},\ \mathbf{C},\ \mathbf{K}\in\mathbb{R}^{m\times m}$ are the mass, damping and stiffness matrices; $\ddot{\mathbf{y}}\left(t\right),\dot{\mathbf{y}}\left(t\right),\ \mathbf{y}\left(t\right)\in\mathbb{R}^{m}$ are the acceleration, velocity and displacement vectors; and $\mathbf{P_N}\left(\mathbf{y}\left(t\right),\dot{\mathbf{y}}\left(t\right)\right)$ and $\mathbf{P}\left(t\right)$ represent the nonlinear restoring force and external load respectively. By grouping the responses and considering their discrete forms, one can obtain

$$\mathbf{Z}=\left[\begin{array}{cccc}\mathbf{y}\left(t_1\right) & \mathbf{y}\left(t_2\right) & \cdots & \mathbf{y}\left(t_N\right)\\ \dot{\mathbf{y}}\left(t_1\right) & \dot{\mathbf{y}}\left(t_2\right) & \cdots & \dot{\mathbf{y}}\left(t_N\right)\end{array}\right]^{\mathrm{T}},\ \dot{\mathbf{Z}}=\left[\begin{array}{cccc}\dot{\mathbf{y}}\left(t_1\right) & \dot{\mathbf{y}}\left(t_2\right) & \cdots & \dot{\mathbf{y}}\left(t_N\right)\\ \ddot{\mathbf{y}}\left(t_1\right) & \ddot{\mathbf{y}}\left(t_2\right) & \cdots & \ddot{\mathbf{y}}\left(t_N\right)\end{array}\right]^{\mathrm{T}}\#\left(2\right) \tag{2}$$

where $\mathbf{Z},\dot{\mathbf{Z}}\in\mathbb{R}^{N\times 2m}$ are the assembled responses.

Subsequently, a dictionary including the potential candidate components of the nonlinear system is constructed as follows:

$$\mathbf{\Phi}\left(\mathbf{Z}\right)=\left[\begin{array}{ccccccccc}\mathbf{1} & \mathbf{Z} & \mathbf{Z_2} & \mathbf{Z_3} & \cdots & \mathrm{sign}\left(\mathbf{Z}\right) & \left|\mathbf{Z}\right| & \cdots & \mathbf{P}\end{array}\right]\#\left(3\right) \tag{3}$$

where the dictionary $\mathbf{\Phi}\left(\mathbf{Z}\right)\in\mathbb{R}^{N\times n}$ encompasses constants, polynomials, sign function terms, absolute values, and inputs etc. Once the dictionary is established, the system identification can be formulated through a linear regression model:

$$\dot{\mathbf{Z}}=\mathbf{\Phi}\left(\mathbf{Z}\right)\mathbf{W}\#\left(4\right) \tag{4}$$

where $\mathbf{W}\in\mathbb{R}^{n\times 2m}$ represents the coefficients of candidate components. Ideally, Eq. (4) can be solved by linear least squares regression in the absence of noise interference. However, in practical applications, it usually suffers from overfitting

due to noisy measurements, in which irrelevant components may be assigned non-zero coefficients. This can significantly compromise the accurate understanding of the target nonlinear system.

To mitigate overfitting, sparse regularization techniques are usually used to constrain the model coefficients. In principle, L_0-norm regularization is the optimal choice, as it directly counts the component number in the identified system. However, L_0-norm regularization is a non-convex optimization problem, which is difficult to solve. As a practical alternative, L_1-norm regularization, also known as LASSO, is widely applied to provide a sub-optimal solution. Alternatively, in SBL, sparsity is promoted by imposing the automatic relevance determination (ARD) prior to the coefficients. It will be illustrated in what follows.

Gibbs sampler-based SBL

A SBL framework based on Gibbs sampling is proposed in this section to solve the aforementioned sparse regression model. The proposed approach follows the concept of the relevance vector machine (RVM) [4]. However, instead of maximum a posteriori (MAP) estimation, it utilizes a Gibbs sampler to obtain the full posterior distribution.

To start with, an error term is introduced into the regression model:

$$\mathbf{z} = \mathbf{\Phi w} + \boldsymbol{\varepsilon}_{\mathbf{z}},\ \boldsymbol{\varepsilon}_{\mathbf{z}} \sim \mathcal{N}\left(0, \sigma^2 \mathbf{I}\right) \# (5) \tag{5}$$

where $\mathbf{z} = [z_1, z_2, \ldots, z_N]^{\mathrm{T}}$ and $\mathbf{w} = [w_1, w_2, \ldots, w_n]^{\mathrm{T}}$ denote one column of matrices $\dot{\mathbf{Z}}$ and $\mathbf{W}$; $\boldsymbol{\varepsilon}_{\mathbf{z}}$ denotes the error vector, which follows an independent multivariate Gaussian distribution with zero mean and variance σ^2. Consequently, the likelihood can be obtained as

$$p\left(|\mathbf{z} \mid \mathbf{w}, \sigma^2\right) = \mathcal{N}\left(\mathbf{\Phi w}, \sigma^2 \mathbf{I}\right) \# (6) \tag{6}$$

Subsequently, the ARD prior is imposed to the coefficients $\mathbf{w}$. Additionally, owing to the conjugate property, Gamma distributions are selected as the hyperpriors to the precision parameters of coefficients and data. Specifically, the prior and hyperpriors are specified as follows:

$$\begin{aligned} p(\mathbf{w} \mid \boldsymbol{\alpha}) &= p_{\mathrm{ARD}}(\mathbf{w} \mid \boldsymbol{\varepsilon}) = \prod_{i=1} \mathcal{N}\left(w_i \mid 0, \varepsilon_i\right) \\ p(\boldsymbol{\alpha} \mid \mathbf{a}, \mathbf{b}) &= \prod_{i=1}^{n} \mathrm{Gamma}\left(\alpha_i \mid a_i, b_i\right) \\ p\left(\sigma^2 \mid c, d\right) &= \mathrm{Gamma}\left(\frac{1}{\sigma^2} \middle| c, d\right) \end{aligned} \tag{7}$$

where $\boldsymbol{\varepsilon} = [\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n]^{\mathrm{T}}$ is the coefficient variances; $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_n]^{\mathrm{T}} = \left[\frac{1}{\varepsilon_1}, \frac{1}{\varepsilon_2}, \ldots, \frac{1}{\varepsilon_n}\right]^{\mathrm{T}}$ is the corresponding precision parameters; $\mathbf{a}$, $\mathbf{b}$, c, and d are the shape and rate parameters of Gamma distributions.

By applying Bayes theorem, a hierarchical Bayesian model is established:

$$p\left(|\mathbf{w}, \boldsymbol{\alpha}, \sigma^2 \mid \mathbf{z}, \boldsymbol{\theta}\right) = \frac{p\left(|\mathbf{z}|\mathbf{w}, \sigma^2\right) p(|\mathbf{w}|\boldsymbol{\alpha}) p\left(|\boldsymbol{\alpha}|\mathbf{a}, \mathbf{b}\right) p\left(|\sigma^2|c, d\right)}{p(|\mathbf{z}|\boldsymbol{\theta})} \tag{8}$$

where $\boldsymbol{\theta} = \{\mathbf{a}, \mathbf{b}, c, d\}$ denotes the hyperparameters, $p\left(|\mathbf{w}, \boldsymbol{\alpha}, \sigma^2 \mid \mathbf{z}, \boldsymbol{\theta}\right)$ represents the joint posterior, and $p\left(|\mathbf{z} \mid \boldsymbol{\theta}\right)$ is the evidence (or marginal likelihood). Theoretically, the target variables $\left\{\mathbf{w}, \boldsymbol{\alpha}, \sigma^2\right\}$ should be estimated simultaneously by maximizing the joint posterior. However, it is a computational nontrivial task.

To render the problem tractable, a Gibbs sampler is applied to infer the conditional posterior distributions of individual variables. Since the specified distributions follow conjugacy, the conditional posteriors of all variables have the same distribution forms as their priors. Specifically, the derived conditional posteriors for the target variables are provided as follows:

$$\begin{aligned} p\left(|\mathbf{w} \mid \mathbf{z}, \boldsymbol{\alpha}, \sigma^2\right) &\propto \mathcal{N}\ \mathbf{w} \left| \left(\mathbf{\Phi}^{\mathrm{T}} \mathbf{\Phi} + \sigma^2 \mathbf{A}\right)^{-1} \mathbf{\Phi}^{\mathrm{T}} \mathbf{z}, \left(\sigma^{-2} \mathbf{\Phi}^{\mathrm{T}} \mathbf{\Phi} + \mathbf{A}\right)^{-1} \right. \\ p\left(|\boldsymbol{\alpha} \mid \mathbf{z}, \mathbf{w}, \mathbf{a}, \mathbf{b}\right) &\propto \prod_{i=1}^{n} \mathrm{Gamma}\left(|\alpha_i \middle| a_i + \frac{1}{2}, b_i + \frac{1}{2} w_i^2\right) \\ p\left(|\sigma^2 \mid \mathbf{z}, \mathbf{w}, c, d\right) &\propto \mathrm{Gamma}\left(|\frac{1}{\sigma^2} \middle| c + \frac{N}{2}, d + \frac{1}{2} \|\mathbf{z} - \mathbf{\Phi w}\|_2^2\right) \end{aligned} \tag{9}$$

where $\mathbf{A} = \mathrm{diag}(\boldsymbol{\alpha})$ is the diagonal precision matrix. The Gibbs sampler draws samples of each variable from the conditional posterior, with other variables fixed as their latest samples. The initial samples generated by the Gibbs sampler are typically unstable and deviated from the true values. Thus, a burn-in period $N_{burn} = 500$ is predefined, while the maximum sampling number is set to 2500.

As mentioned before, SBL may fail to remove irrelevant components, resulting in overfitted models. In the next section, the evidential thresholding approach is proposed to solve this issue.

Evidential thresholding

Evidence serves as an important indicator for assessing the plausibility of candidate models in the Bayesian model selection [9]. Inspired by this principle, the decoupled evidential contributions corresponding to the candidate components in the linear regression model are derived and employed as an indicator to evaluate the plausibility of each component. Thus, it can also be interpreted as component-wise Bayesian model selection. A probabilistic thresholding method, termed evidential thresholding, is then proposed to eliminate components with relatively small evidential contributions. The method is integrated into the SBL framework to enhance sparsity.

The nonlinear system to be identified is only associated with the coefficients $\mathbf{w}$. Consequently, the corresponding evidence is calculated by marginalizing out the coefficients from the likelihood and the prior:

$$p\left(|\mathbf{z} \mid \boldsymbol{\alpha}, \sigma^2\right) = \int_{\mathbf{w}} p\left(|\mathbf{z} \mid \mathbf{w}, \sigma^2\right) p\left(|\mathbf{w} \mid \boldsymbol{\alpha}\right) d\mathbf{w}. \tag{10}$$

After applying the Sherman-Morrison formula to separate the *i*-th component from the others, one can derive the logarithmic evidence as follows:

$$\begin{aligned} \ln p\left(|\mathbf{z} \mid \boldsymbol{\alpha}, \sigma^2\right) &= \ln p\left(|\mathbf{z} \mid \boldsymbol{\alpha}_{-i}, \sigma^2\right) + \mathcal{L}\left(\alpha_i\right), \\ \mathcal{L}\left(\alpha_i\right) &= \frac{1}{2}\left[-\ln\left(1+\frac{s_i}{\alpha_i}\right) + \frac{q_i^2}{\alpha_i + s_i}\right], \end{aligned} \tag{11}$$

where $s_i = {\boldsymbol{\Phi}_i}^{\mathrm{T}}\left(\mathbf{C}_{-i}\right)^{-1}\boldsymbol{\Phi}_i$ and $q_i = {\boldsymbol{\Phi}_i}^{\mathrm{T}}\left(\mathbf{C}_{-i}\right)^{-1}\mathbf{z}$, with $\mathbf{C}_{-i} = \sigma^2\mathbf{I} + \sum_{j\neq i}\varepsilon_j\boldsymbol{\Phi}_j{\boldsymbol{\Phi}_j}^{\mathrm{T}}$. $\ln p\left(|\mathbf{z} \mid \boldsymbol{\alpha}_{-i}, \sigma^2\right)$ represents the logarithmic evidence with the *i*-th component removed, while $\mathcal{L}\left(\alpha_i\right)$ denotes the decoupled contribution of the *i*-th component to the logarithmic evidence. The evidence can then be expressed as:

$$p\left(|\mathbf{z} \mid \boldsymbol{\alpha}, \sigma^2\right) = \exp\left[\mathcal{L}\left(\alpha_i\right)\right]\ p\left(|\mathbf{z} \mid \boldsymbol{\alpha}_{-i}, \sigma^2\right). \# \tag{12}$$

Eq. (12) shows that if the evidential contribution $\mathcal{L}\left(\alpha_i\right)$ is positive, including the *i*-th component will increase the evidence; conversely, if it is negative, the *i*-th component will decrease the evidence. This provides a criterion to select the candidate components, where the components with large positive evidential contributions should be included in the identified model. Additionally, it is found that the irrelevant components due to noise interference usually have relatively small evidential contributions, even though their coefficients may be large. Therefore, to avoid overfitting, the components with relatively small evidential contributions should also be eliminated. To better evaluate the relative contributions, the evidential contribution ratios are calculated as:

$$\mathcal{R}\left(\alpha_i\right) = \frac{\mathcal{L}\left(\alpha_i\right)}{\sum_{i=1}^{n}\mathcal{L}\left(\alpha_i\right)}. \tag{13}$$

The ratio is bounded as $\mathcal{R}\left(\alpha_i\right) \in [0, 1]$ and $\sum_{i=1}^{n}\mathcal{R}\left(\alpha_i\right) = 1$.

A threshold value for the evidential contribution ratio can then be defined to remove components with ratios smaller than it. Similar to SINDy, the threshold value is a hyperparameter that can significantly affect the results. However, since it is bounded, the threshold value is much easier to be defined compared to conventional thresholding techniques. The threshold value is empirically set to 10^{-4} in this paper.

The proposed Bayesian framework for nonlinear system identification is then constructed. Firstly, the GB-SBL is conducted to obtain a preliminary estimate of the system equation. Subsequently, evidential thresholding is performed to further eliminate irrelevant components. The updated coefficients are then used as initial values for the next iteration of the GB-SBL. This process is repeated until all evidential contribution ratios exceed the predefined threshold value.

Demonstration

Numerical 3-DOF nonlinear system

Firstly, a numerical 3-DOF nonlinear dynamical system, including two cubic stiffness springs and a Coulomb friction damper, as shown in Fig. 1, is applied to demonstrate the effectiveness of the proposed method. The linear mass, damping and stiffness coefficients are defined as: $m_i = 2,\ c_i = 2,\ k_i = 200$, for $i = 1,\ 2,\ 3$, while the nonlinear parameters are set to $k_{N1} = 100,\ K_{N2} = 200,\ C_{N1} = 4$. The third mass is excited by a harmonic input $P(t) = 200\sin(1.8\pi t)$. The system responses are measured at a sampling rate of 200 Hz for 150 seconds. The data are contaminated by Gaussian white noise with 3% standard deviation.

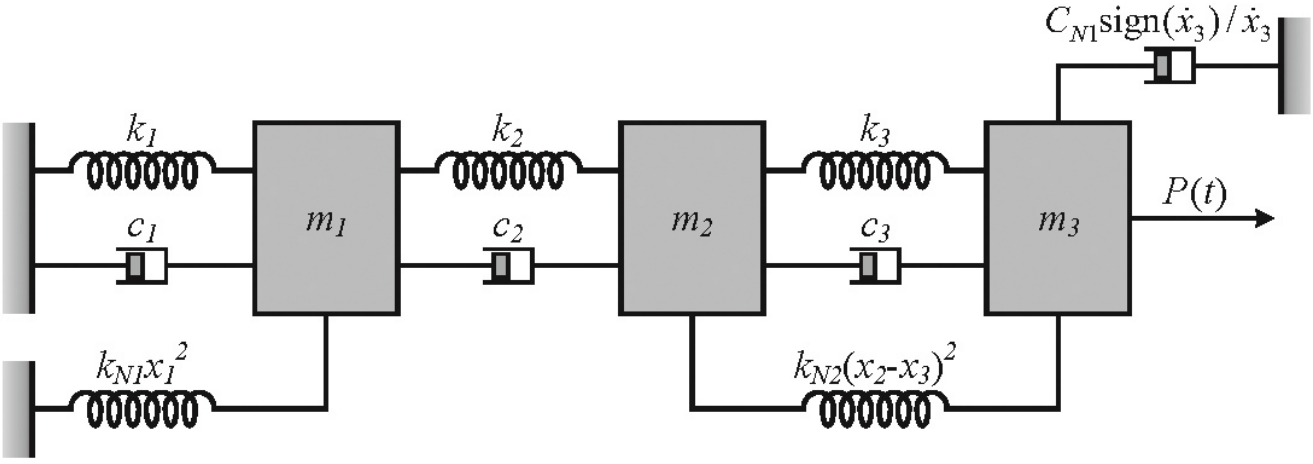

Fig. 1 Numerical 3-DOF nonlinear dynamical system

A dictionary including linear, second-order and third-order polynomials, sign function terms, absolute value terms, and inputs is constructed for the implementation of sparse regression. The time-domain responses of the identified system equation are presented in Fig. 2. It shows good agreement between the identified responses and the exact values.

The GB-SBL without evidential thresholding, RVM, and SINDy are also applied to identify the system for comparison. For brevity, only the results of the identified nonlinear terms are shown below, where the values in the brackets denote the standard deviations.

– Proposed method:

$$\mathbf{P_N}(\mathbf{t}) = \begin{bmatrix} -49.8\,(\pm 0.02)\,x_1^3 \\ -99.3\,(\pm 1.17)\,x_2^3 + 297.5\,(\pm 3.10)\,x_2^2x_3 - 296.7\,(\pm 2.75)\,x_2x_3^2 + 98.5\,(\pm 0.85)\,x_3^3 \\ 99.2\,(\pm 1.13)\,x_2^3 - 297.5\,(\pm 2.98)\,x_2^2x_3 + 297.1\,(\pm 2.53)\,x_2x_3^2 - 98.6\,(\pm 0.45)\,x_3^3 - 2.1\,(\pm 0.21)\,\text{sign}(\dot{x}_3) \end{bmatrix}$$

– GB-SBL without evidential thresholding:

$$\mathbf{P_N}(\mathbf{t}) = \begin{bmatrix} -49.8\,(\pm 0.02)\,x_1^3 + 1.2\,(\pm 0.15)\,x_2x_3 - 0.66\,(\pm 0.07)\,x_3^2 + 0.31\,(\pm 0.05)\,\text{abs}\,(x_2) + \ldots \\ -98.7\,(\pm 0.30)\,x_2^3 + 298.3\,(\pm 0.82)\,x_2^2x_3 - 297.3\,(\pm 0.78)\,x_2x_3^2 + 98.8\,(\pm 0.26)\,x_3^3 + 1.1\,(\pm 0.19)\,x_1^3 - 0.6\,(\pm 0.14)\,\text{abs}\,(x_3) + \ldots \\ 98.7\,(\pm 0.43)\,x_2^3 - 297.8\,(\pm 0.66)\,x_2^2x_3 + 296.8\,(\pm 1.42)\,x_2x_3^2 - 98.5\,(\pm 0.82)\,x_3^3 - 2.5\,(\pm 0.36)\,\text{sign}\,(\dot{x}_3) + 0.9\,(\pm 0.07)\,x_1^2 + \ldots \end{bmatrix}$$

– RVM:

$$\mathbf{P_N}(\mathbf{t}) = \begin{bmatrix} -49.8\,(\pm 0.02)\,x_1^3 + 2.7\,(\pm 0.19)\,x_2x_3 - 1.2\,(\pm 0.09)\,x_3^2 + 0.43\,(\pm 0.07)\,\text{abs}\,(x_2) + \ldots \\ -98.6\,(\pm 0.30)\,x_2^3 + 297.9\,(\pm 0.74)\,x_2^2x_3 - 298.6\,(\pm 0.70)\,x_2x_3^2 + 98.9\,(\pm 0.24)\,x_3^3 - 4.8\,(\pm 0.88)\,x_1^2 - 3.6\,(\pm 0.35)\,\text{x}_1x_2 + \ldots \\ 98.1\,(\pm 0.34)\,x_2^3 - 297.2\,(\pm 0.76)\,x_2^2x_3 + 298.1\,(\pm 0.65)\,x_2x_3^2 - 98.2\,(\pm 0.52)\,x_3^3 - 1.7\,(\pm 0.16)\,\text{sign}\,(\dot{x}_3) + 4.3\,(\pm 0.75)\,x_1^2 + \ldots \end{bmatrix}$$

– SINDy:

$$\mathbf{P_N}(\mathbf{t}) = \begin{bmatrix} -49.8x_1^3 \\ -97.4x_2^3 + 301.6x_2^2x_3 - 302.8x_2x_3^2 + 100.9x_3^3 + 22.9x_1^2x_2 + 18.5\text{x}_1x_2x_3 - 15.3x_1^2x_3 + \ldots \\ 98.2x_2^3 - 295.1x_2^2x_3 + 295.2x_2x_3^2 - 98.4x_3^3 - 2.21\text{sign}\,(\dot{x}_3) - 3.1x_1^2 + 3.8x_1x_2 - 1.8x_2x_3 + \ldots \end{bmatrix}$$

The results indicate that only the proposed approach successfully identified both the equation form and the parameters of the nonlinear dynamical system, while other methods failed due to the presence of redundant components in their identified models.

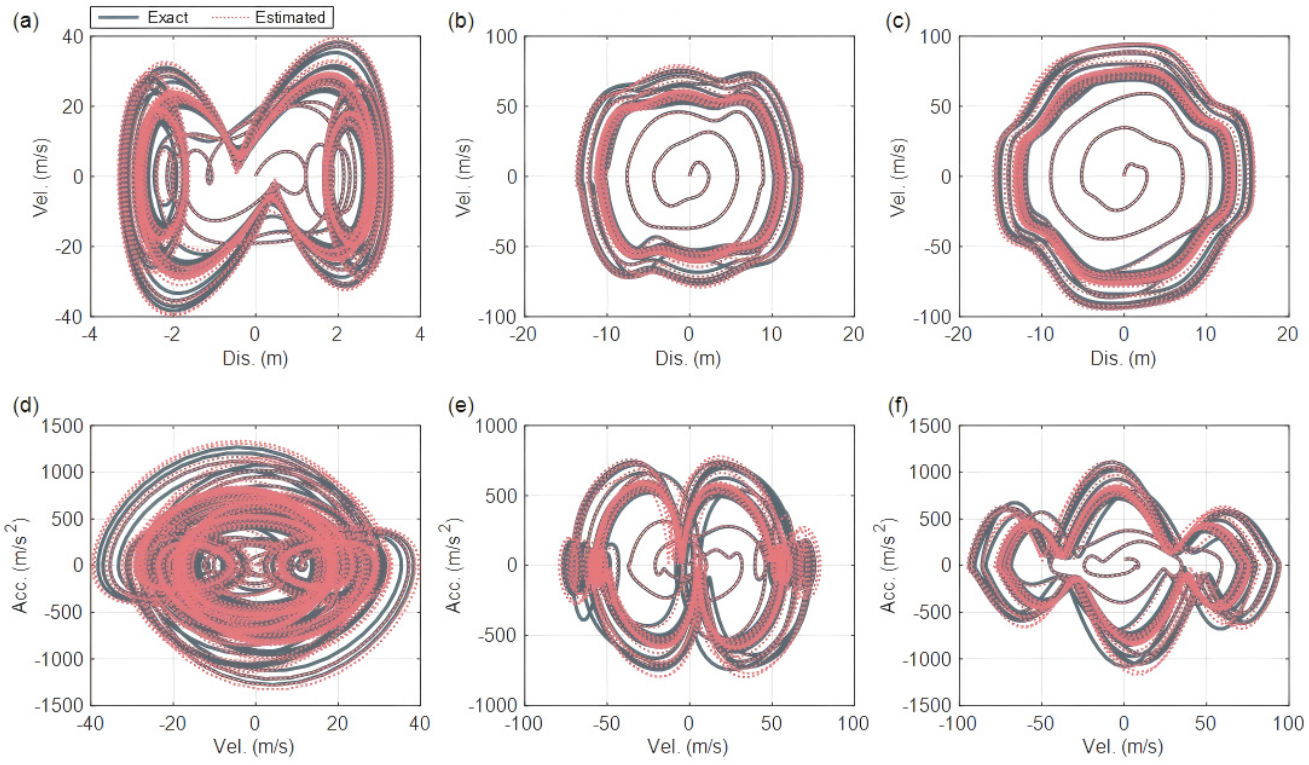

Fig. 2 The identified phase portraits of the numerical 3-DOF nonlinear dynamical system: (a) - (c) displacements and velocities of mass 1 - 3, (d) - (f) velocities and accelerations of mass 1 - 3

Experimental benchmark

An experimental benchmark of an electro-mechanical positioning system [10] is studied to further demonstrate the effectiveness of the proposed method in practical applications. The experimental setup is shown in Fig. 3, which represents a standard drive system of prismatic joints. The primary nonlinear component of the system results from friction. The detailed description of the system can be found in [10].



Fig. 3 Experimental setup of the electro-mechanical positioning system

The proposed method is applied to identify the governing equation of the system, in comparison with GB-SBL without evidential thresholding, RVM, and SINDy. The identified equations are listed in Table 1. The results indicate that the proposed method provided the most parsimonious equation of the system, while other methods identified more redundant components. However, compared to the reference model, it still includes a second-order polynomial term. It is not surprising, as the perfect Coulomb friction hardly exists in real applications [5]. The estimated accelerations are presented in Fig. 4, which correspond well to the measurements.

Table 1 The identified equation of the electro-mechanical positioning system

Reference	$\ddot{y}(t) = 0.033 - 2.140\dot{y}(t) - 0.214\text{sign}(\dot{y}(t)) + 0.370P(t)$
Proposed method	$\ddot{y}(t) = 0.014\left(\pm 2.2 \times 10^{-4}\right) - 2.146\left(\pm 3.6 \times 10^{-3}\right)\dot{y}(t) + 2.465\,(\pm 0.02)\,\dot{y}^2(t)$ $-0.213\left(\pm 3.2 \times 10^{-4}\right)\text{sign}(\dot{y}(t)) + 0.369(\pm 1.24 \times 10^{-4})P(t)$
GB-SBL	$\ddot{y}(t) = 0.020\left(\pm 3.4 \times 10^{-4}\right) - 0.077\,(\pm 0.01)\,y(t) - 2.166\left(\pm 5.5 \times 10^{-3}\right)\dot{y}(t) + 0.156\,(\pm 0.04)\,y^2(t)$ $+0.310(\pm 0.03)y(t)\dot{y}(t) + 2.454\,(\pm 0.04)\,\dot{y}^2(t) - 0.214\left(\pm 3.0 \times 10^{-4}\right)\text{sign}(\dot{y}(t))$ $+0.367(\pm 1.23 \times 10^{-4})P(t)$
RVM	$\ddot{y}(t) = 0.020\left(\pm 3.4 \times 10^{-4}\right) - 0.078\,(\pm 0.01)\,y(t) - 2.165\left(\pm 5.1 \times 10^{-3}\right)\dot{y}(t) + 0.165\,(\pm 0.04)\,y^2(t)$ $+0.299(\pm 0.03)y(t)\dot{y}(t) + 2.455\,(\pm 0.04)\,\dot{y}^2(t) - 0.214\left(\pm 3.3 \times 10^{-4}\right)\text{sign}(\dot{y}(t))$ $+0.368(\pm 1.26 \times 10^{-4})P(t)$
SINDy	$\ddot{y}(t) = 0.020 - 0.082y(t) - 2.164\dot{y}(t) + 0.180y^2(t) + 0.297y(t)\dot{y}(t) + 2.465\dot{y}^2(t) - 0.214\text{sign}(\dot{y}(t))$ $+0.368P(t)$

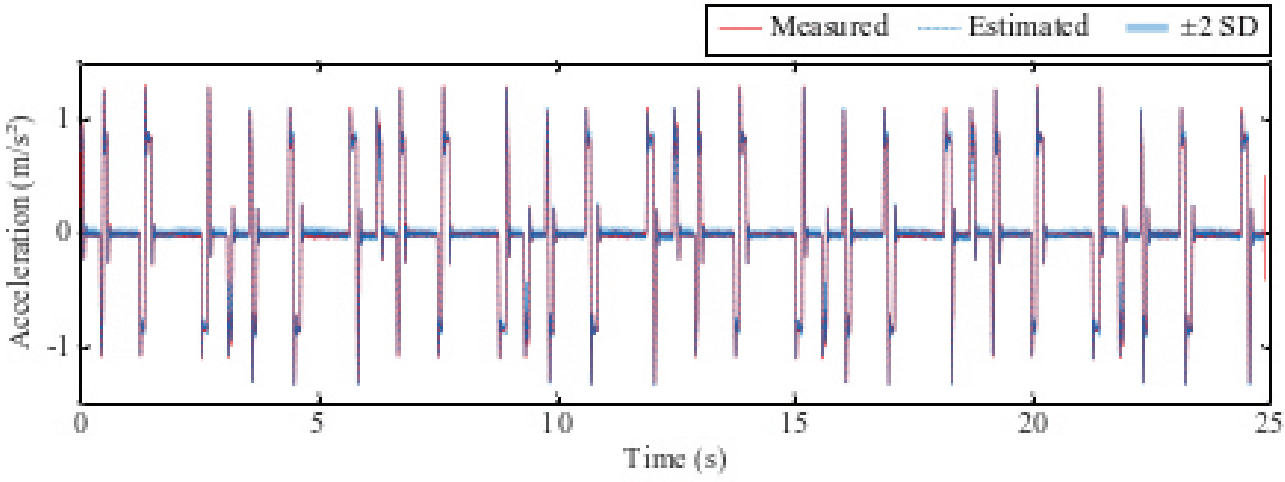

Fig. 4 The estimated accelerations with double standard deviation

Conclusion

Existing sparse identification methods for nonlinear dynamical systems commonly suffer from overfitting due to the presence of redundant components. To address this issue, this paper proposes a probabilistic thresholding technique, referred to as evidential thresholding, which is cooperated with GB-SBL to enhance the sparsity of the identified model. Firstly, the formulation of GB-SBL is briefly introduced. Subsequently, the theoretical foundation of the proposed evidential thresholding is derived. Instead of removing candidate components with small coefficients, evidential thresholding evaluates the decoupled evidential contribution of each component and eliminates those with negligible contributions. This approach can also be interpreted as a component-wise Bayesian model selection technique. It avoids circumstances where the correct components are falsely removed due to their small coefficients, while the irrelevant components with large coefficients are still retained in the model. Evidential thresholding is then integrated into GB-SBL to form the proposed Bayesian method for nonlinear dynamical system identification. A numerical example and an experimental benchmark have been studied to verify the effectiveness of the proposed method, which demonstrates its superiority over SINDy, RVM, and GB-SBL without evidential thresholding.

Acknowledgments Jiawei Jian gratefully acknowledges the support of the China Scholarship Council. The authors acknowledge the support of KU Leuven Internal Funds.

References

1. Kerschen G, Worden K, Vakakis A F, et al. Past, present and future of nonlinear system identification in structural dynamics. Mechanical systems and signal processing, 2006, 20(3): 505-592.
2. Noël J P, Kerschen G. Nonlinear system identification in structural dynamics: 10 more years of progress. Mechanical Systems and Signal Processing, 2017, 83: 2-35.
3. Brunton S L, Proctor J L, Kutz J N. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. Proceedings of the national academy of sciences, 2016, 113(15): 3932-3937.
4. Tipping M E. Sparse Bayesian learning and the relevance vector machine. Journal of machine learning research, 2001, 1(Jun): 211-244.
5. Fuentes R, Nayek R, Gardner P, et al. Equation discovery for nonlinear dynamical systems: A Bayesian viewpoint. Mechanical Systems and Signal Processing, 2021, 154: 107528.
6. Chatterjee T, Shaw A D, Friswell M I, et al. Sparse Bayesian machine learning for the interpretable identification of nonlinear structural dynamics: Towards the experimental data-driven discovery of a quasi zero stiffness device. Mechanical Systems and Signal Processing, 2023, 205: 110858.
7. Hirsh S M, Barajas-Solano D A, Kutz J N. Sparsifying priors for Bayesian uncertainty quantification in model discovery. Royal Society open science, 2022, 9(2): 211823.
8. Sun L, Huang D, Sun H, et al. Bayesian spline learning for equation discovery of nonlinear dynamics with quantified uncertainty. Advances in neural information processing systems, 2022, 35: 6927-6940.
9. Beck J L, Yuen K V. Model selection using response measurements: Bayesian probabilistic approach. Journal of Engineering Mechanics, 2004, 130(2): 192-203.
10. Janot A, Gautier M, Brunot M. Data set and reference models of EMPS. Nonlinear system identification benchmarks, 2019.