Chapter 3
Vibration-Based System Identification of Structural Dynamics using a Stochastic Gradient SINDy

Celso T. do Cabo, Raymond Joshua, and Zhu Mao

Abstract System identification is a challenging field of research. Oftentimes, it is possible to obtain certain information from a system based on measurements and models. In vibrational systems, for instance, natural frequencies and mode shapes can be obtained based on measurements. However, estimating the partial differential equations (PDEs) of a system is a complex challenge. Many approaches to model the system's behavior rely on physical information, such as in FEA models. When the physics information of the system is missing, data-driven approaches are a feasible solution to create a digital twin of the real system. Among the techniques for system identification available, the sparse identification of nonlinear dynamics (SINDy) is a library-based model that performs a linear regression using the measurement and its derivatives to estimate possible parameters existing in the system. However, structural vibrations consist of second-order PDEs with mass and stiffness dependences between the degrees of freedom. It is proposed that a modification of the current SINDy is employed by developing a gradient descent with penalty terms to keep the physical meaning of the PDEs. Results show that the model can predict the matrices of the system, with high signal reconstruction and mode shape estimation. The values obtained are compared with FEA models to verify their accuracy.

Keywords Machine learning · System identification · Sparse identification · Nonlinear dynamics · Structural dynamics

Introduction

Monitoring the dynamic behavior of structures is of great importance in evaluating the system's integrity. In many applications, it is possible to identify damage or to predict the remaining useful life of a system based on its vibration [1]. However, the system's dynamic often has complex behavior, and it is difficult to model. In many cases, a baseline is adopted as the initial system or nondamaged condition of the structure, and future changes are assumed to be damaged. In other cases, where some knowledge is available of the structure, it is possible to generate a mathematical or simulation model of the system, and deviations from the simulation are classified as damage.

In cases where physical knowledge is not available, data-driven techniques can be employed either to understand the complex behavior of the measurement collected or to simplify the system's behavior [2]. However, such approaches require large amounts of data collected and sometimes are not feasible to perform. In the case of nonlinear dynamics, it is even more difficult to obtain the governing equations without previous knowledge of the system [3]. Another approach is to identify some information about the system, based on the measurements obtained. Among such techniques, fuzzy logic, or

Celso T. do Cabo · Zhu Mao
Department of Mechanical and Materials Engineering, Worcester Polytechnic Institute, Worcester, MA 01609
e-mail: ctdocabo@wpi.edu; zmao2@wpi.edu

Raymond Joshua
Honeywell Federal Manufacturing & Technologies, LLC operates the Kansas City, National Security Campus for the United States Department of Energy / National Nuclear, Security Administration under Contract Number DE-NA0002839
e-mail: rjoshua@kcnsc.doe.gov

*Footnote for Authorization for:
NSC-614-5856
DE-NA0002893

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

Kristin B. Zimmerman, PhD (eds.), *Topics in Modal Analysis & Parameter Identification in Structural Dynamics, 2026, Vol. 8*,
Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/rp-9788743814191

physics-informed neural networks (PINNs) can be employed [4–6]. Some techniques rely on linear regression, such as the sparse identification of nonlinear dynamics (SINDy) [7]. This technique proposes performing measurements of a dynamical system, then calculating its derivatives or integral to create a library of parameters to estimate the governing equations. It proved to be successful in diverse applications, such as single degree of freedom systems (SDOF) or rotating machines [8–10]. SINDy was also employed in nonlinear problems for system identification and implicit equations [11, 12]. Previously, different modifications in its structure were proposed, including optimization methods to obtain the governing equation using SINDy, including PDEs estimation [13–15]. This work proposes the usage of the SINDy model to identify the PDEs of a continuous system's vibration. For that, SINDy is modified to employ a stochastic gradient descent with penalty terms to obtain the most realistic governing equations possible. The schematics of the proposed approach are presented in Figure 1.

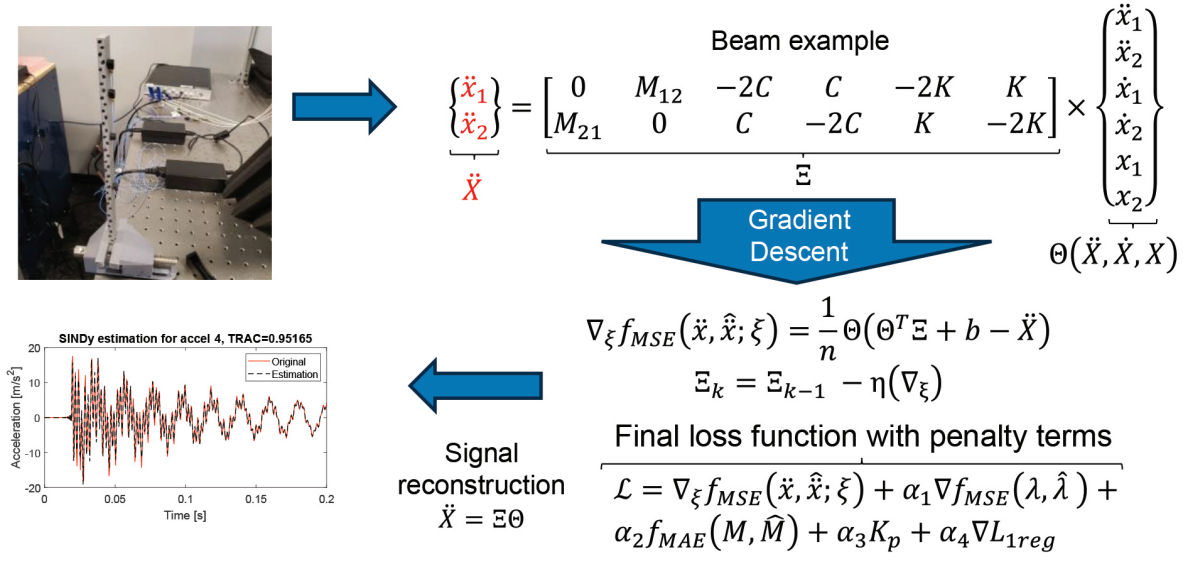


Fig. 1 Flowchart of proposed SGD SINDy model

Theory background

The Sparse Identification of Nonlinear Dynamics (SINDy) model was initially introduced by Brunton et al. [7] and consists of a linear regression method to estimate the parameters of a governing equation. For the case of vibrational systems, where the governing equation of the system consists of a second-order PDE. The measurement is given by a matrix with different measurements of acceleration $\left[\ddot{X}\right]$. This signal will be reconstructed based on a combination of parameters using the derivatives or integrals of the signal. For instance, if the acceleration is the measurement for this case, the library of parameters $\Theta\left(\dot{X}, X\right)$ will consist of velocities and displacements obtained from integrating the measurements. In addition, in case of nonlinear systems, other terms can be included, such as a cubic nonlinear spring. Finally, the parameters Ξ are estimated according to Equation (1).

$$\Xi = \left(\Theta^T\Theta\right)^{-1}\Theta^T\ddot{X} \tag{1}$$

The linear regression model has a high performance when estimating the parameters, obtaining the ideal solution with the smallest error. Although the model works well for second-order PDEs, continuous systems have dependency between the degrees of freedom (DOFs) in both mass and stiffness components. Thus, even the solution with the smallest error will have a low accuracy. A new model based on stochastic gradient descent (SGD) is proposed to counter such issues. For that, the gradient of the mean squared error (MSE) is calculated as shown in Equation (2).

$$\nabla_\xi f_{MSE}\left(\ddot{x}, \hat{\ddot{x}}, \xi\right) = \frac{1}{n}\Theta\left(\Theta^T\Xi + b - \ddot{X}\right) \tag{2}$$

After calculating the gradient of the MSE, the parameters are updated based on Equation (3).

$$\Xi_k = \Xi_{k-1} - \eta\left(\nabla_\xi\right) \tag{3}$$

Where η is the learning rate and hyperparameter of the model. In addition, when modifying the approach to an SGD, restrictions of the algorithm can be performed. For instance, the library of parameters can also include the mass terms on it, changing the existing one to a $\Theta\left(\ddot{X}, \dot{X}, X\right)$. Then, to avoid a case where $\ddot{X}_{ii} = \ddot{X}_{ii}$ the diagonal of the parameters is forced to be always zero. This will lead to a new solution where the masses are also dependent on different DOFs. However, inverse problems often are ill-posed problems, with different possible combinations of parameters to obtain the same signal reconstruction. Therefore, different penalty terms were included in the SINDy model, such that, with limited information of the system, it is possible to obtain a more realistic estimation of the governing equations. Some of those terms were previously introduced by the authors and are described in the following equations.

$$\hat{M} = \sum_{i=1}^{n} M_{ii} \tag{4}$$

In Equation (4), $\hat{M}$ is the total mass of the system, which can be estimated or measured depending on the application. To bound the system's mass matrix to a similar value to the total mass of the system, it is assumed that the sum of the diagonal terms of the mass matrix will be the same as the total mass of the system. The mass of the first DOF will be estimated based on the optimization of the main absolute error (MAE) as shown in Equation (5).

$$M_{11}^{k+1} = M_{11}^{k} + \eta_m \left| \hat{M} - \sum_{i=1}^{n} M_{ii} \right| \tag{5}$$

Then, the other terms of the diagonal can be calculated based on the following equation:

$$M_{ii} = \frac{\xi_{i-1,i}}{\xi_{i,i-1}} \times M_{i-1,i-1} \tag{6}$$

Where $\xi_{i,j}$ is one of the terms in the matrix Ξ. Another penalty term previously implemented was a term to encourage symmetry in the matrices obtained by the model. This is obtained by following Equation (7) below.

$$A_p = \frac{\left[A - A^T\right]}{\|A\|} \tag{7}$$

Here, the A matrix can be any of the matrices of the governing equations, such as the mass, stiffness, or damping matrices. The model, including the previously mentioned penalty terms, can estimate the governing equations of the system, with a time response assurance criterion (TRAC) of above 90%. However, the eigenvalues of the obtained matrices are not the same as the values measured in the experiment. In addition, in case of nonlinear systems, where multiple terms need to be implemented in the library, the model will predict all of the terms to achieve a higher accuracy, even if those terms are not present in the original system. Previously, an eigenvalue estimator was implemented in the framework using the derivatives of the eigenvalues. But this model only predicted correctly one of the natural frequencies. To improve the accuracy of the eigenvalue estimation, another approach was implemented using a perturbation analysis for eigenvalue estimation [16]. Starting with the generalized eigenvalue problem, described in Equation (8).

$$KU = \Lambda MU \tag{8}$$

where, K is the stiffness matrix of the system, M its mass matrix, U the unitary matrix with the mode shape vectors of the system and Λ the diagonal matrix with the eigenvalues of the system. It is possible to add a perturbation term in the equation, modifying it from Equation (8) to (9).

$$\left(K_0 + \delta K\right)\left(u_0 + \delta u\right) = \left(\lambda_0 + \delta\lambda\right)\left(M_0 + \delta M\right)\left(u_0 + \delta u\right) \tag{9}$$

The perturbation terms are represented with the δ, while the initial estimation for each of the terms is represented with the subscription 0. In addition, the desired eigenvalue can be estimated by approximating the eigenvalue as presented in Equation (10).

$$\lambda \approx \lambda_0 + \epsilon\delta\lambda \tag{10}$$

where, ϵ is the perturbation term. This term was included to add complexity to the solution. By simplifying Equation (9) and isolating the perturbation term for the eigenvalue $\delta\lambda$, the estimation of the eigenvalue will then be given by:

$$\lambda \approx \lambda_0 + \epsilon u^T \left(\delta K - \lambda_0 \delta M\right) u \tag{11}$$

This estimation of the eigenvalue can be included in the MSE between the natural frequencies measured from the experiment and the estimation from Equation (11). By deriving the MSE of the eigenvalue, the new term can be included in the optimization together with the gradient descent. The derivation of the MSE of the eigenvalue with respect to the stiffness and mass matrices are given by Equations (12) and (13), respectively.

$$\frac{\partial f_{MSE}\left(\lambda,\widehat{\lambda}\right)}{\partial K} = -2\epsilon u^T\left(\widehat{\lambda}-\lambda_0-\epsilon u^T\left(\delta K-\lambda_0\delta M\right)u\right)u \tag{12}$$

$$\frac{\partial f_{MSE}\left(\lambda,\widehat{\lambda}\right)}{\partial M} = -2\epsilon\lambda_0 u^T\left(\widehat{\lambda}-\lambda_0-\epsilon u^T\left(\delta K-\lambda_0\delta M\right)u\right)u \tag{13}$$

Here, $\widehat{\lambda}$ represents the measured eigenvalue from the experiment, which is given by the natural frequency square. In addition, a regularization term was included in the model to encourage sparsity in the model obtained. In the case of nonlinear systems, it is difficult to estimate what type of nonlinearity is occurring. Therefore, when generating the matrix with possible parameters, multiple different types of nonlinearities can be included. In addition, in cases where it is unknown if the test has nonlinearities (e.g., boundary conditions are not ideal) it is encouraged to include such terms in the model. However, the model will predict the solution with the least error available, even if it has to estimate nonlinearities that are nonexistent in the system. Therefore, the L_1 norm or least absolute shrinkage and selection operator (LASSO) regression was implemented. This technique is known for encouraging sparsity in the model, increasing zero values for some of the parameters [17]. The model was implemented in the loss function as shown in Equation (14).

$$L_1 = |\xi| \tag{14}$$

Then, the gradient of the L_1 norm is calculated as presented in Equation (15).

$$\nabla L_1 = \begin{cases} 1 \text{ if } \xi \geq 0 \\ -1 \text{ if } \xi < 0 \end{cases} \tag{15}$$

The final loss function for the stochastic gradient SINDy model is presented in Equation (16).

$$\mathcal{L} = \nabla_\xi f_{MSE}\left(\ddot{x},\widehat{\ddot{x}},\xi\right) + \alpha_1 \nabla f_{MSE}\left(\lambda,\widehat{\lambda}\right) + \alpha_2 f_{MAE}\left(M,\hat{M}\right) + \alpha_3 A_p + \alpha_4 \nabla L_1 \tag{16}$$

The terms α_i in Equation (16) account for the weights for each of the penalty terms presented in the loss function. Normally, those weights are hyperparameters of the machine learning model and need to be manually specified. However, such tasks can be time-consuming and greatly impact the performance of the algorithm. Therefore, in addition to the penalty terms, a learning rate annealing technique, commonly employed in physics-informed neural networks (PINNs) was also implemented [18]. This technique automatically calculates the weights based on the maximum gradient of the reconstruction part, versus the average of the loss obtained by each of the specific terms as shown in Equation (17).

$$\widehat{\alpha}_i = \frac{\max_\xi\left\{\left|\nabla_\xi \mathcal{L}_\xi\left(\xi_n\right)\right|\right\}}{\left|\overline{\nabla_\xi \mathcal{L}_i\left(\xi_n\right)}\right|} \tag{17}$$

Equation (17) will estimate the term $\widehat{\alpha}_i$ for each epoch while training the model. Then, to increase the stability of the calculations, this term is averaged with previous iterations, leading to the final weight as described in Equation (18).

$$\alpha_i = (1-\beta)\,\alpha_i + \beta\widehat{\alpha}_i \tag{18}$$

Where β is the stabilization coefficient and select how aggressive the change in the weights is. Typically, $\beta = 0.1$. With this, the algorithm will learn each of the weights automatically and does not require extra hyperparameter tuning from the user.

Experimental Setup and Data Processing

The testing setup for the stochastic gradient descent SINDy, uses a simple cantilever beam with accelerometers attached to it. Since the cantilever beam has a simple behavior, its analytical solution can be obtained for its natural frequencies and mode

shapes. From Rao [19], the first two natural frequencies of the cantilever beam are at 82.9Hz and 519.9Hz. Since the model requires the integration of the measurement to obtain the velocity and displacement at each accelerometer, the sampling frequency (f_s) was set at 12.8kHz.

To detect nonlinearities in the system, a different setup was performed, with the same cantilever beam, however, with a soft boundary condition. In this case, another natural frequency appeared between the two mode shapes, and when evaluating the mode shapes via video-based sensing, it can be seen that the mode shape resonates with the boundary.

To process the acquired acceleration, initially, an integration method is performed. Currently, the integration technique employed is the trapezoid integration, that can be approximated following Equation (19) below.

$$\int_a^b f(t)\,dt \approx \sum_{n=1}^{N} \frac{(f(t_n) + f(t_{n+1}))}{2} \Delta t \tag{19}$$

The error of this numerical integration technique was compared with other techniques, such as using the Fourier properties to integrate in the frequency domain or using the Simpson's rule. All the techniques had a similar performance. Thus, the trapezoid rule showed better computational efficiency and was selected for integration. In addition, when performing numerical integration, any offset component in the signal can lead to a drift in the integrated result. Then, a high-pass filter was employed to remove such drift. Different filters were tested to verify which has the best performance for this case. Starting with a fifth-order Butterworth filter was employed. To verify if the error obtained by this filter was the smallest possible error, a comparison study was conducted. A cantilever plate was excited with a sine sweep from 10Hz to 1kHz and a sampling frequency of 12.5kHz. For this experiment, both accelerometers and LDV were employed at the same time, such that the acceleration measurement could be integrated and compared with the velocity measured by the LDV. It was observed that without any filtering, the TRAC of the measurements would be 78%, while employing the Butterworth filter ended up around 98%. When estimating the governing equations of the system, a fast transition is desired. Therefore, an elliptic filter was employed instead. The performance of the elliptic filter for the validation test was similar to that of the Butterworth. However, when applying to the real test, an increase in the accuracy of the results was observed. Currently, all tests use an elliptic filter with 0.1 peak-to-peak passband ripple in dB and 30 stopband attenuation in dB scale.

System Identification Results and Discussion

In order to compare the results of the SGD SINDy model, the traditional technique was employed, where the linear regression described in Equation (1) is used to obtain the parameters. The results when using the linear regression approach are presented in Figure 2.

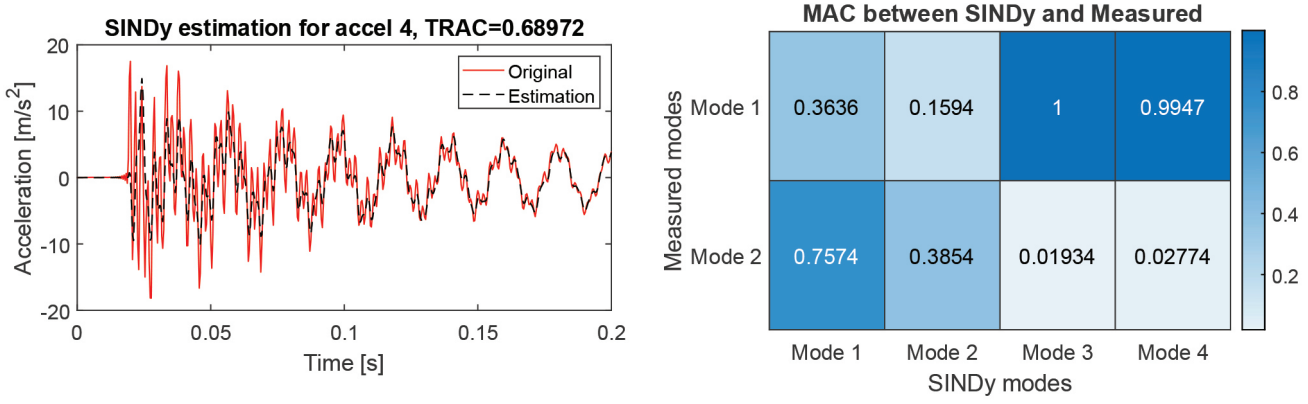


Fig. 2 Results from conventional SINDy; (a) Signal reconstruction of accelerometer 4 and (b) MAC between measurement and eigenvectors from SINDy

When employing the conventional linear regression model for SINDy, the assumption of independence between the terms in the mass matrix leads to low accuracy when reconstructing the measurements from the obtained model. As presented in Figure 2(a), the accelerometer closest to the base of the cantilever beam, which has most of the high-frequency content, will

have a TRAC of only around 68%. In addition, when comparing the mode shapes measured in the experiment with the ones obtained by the eigen decomposition of the mass and stiffness matrices, it can be seen in Figure 2(b) that although the first mode shape has a MAC of 100%, the second mode shape achieves only 75% of MAC. After implementing the SGD process with the penalty terms to obtain the PDEs of the system, the following hyperparameters were used for training:

Table 1 Hyperparameters for SGD SINDy training

Hyperparameter	# of Epochs	Batch Size	Learning Rate	Mass LR	$\alpha_{initial}$	ϵ
Value	10,000	128	1.5×10^{-4}	1.5×10^{-4}	0.1	0.01

When performing the SGD SINDy with the stiff boundary condition, it is possible to see that even the accelerometer with high-frequency content will have a high TRAC, with more than 90% reconstruction, as presented in Figure 3(a). In addition, the MAC for the mode shapes extracted in the experiment and the ones obtained by the eigen solution of the mass and stiffness matrices are above 90% for the first two mode shapes as presented in Figure 3(b).

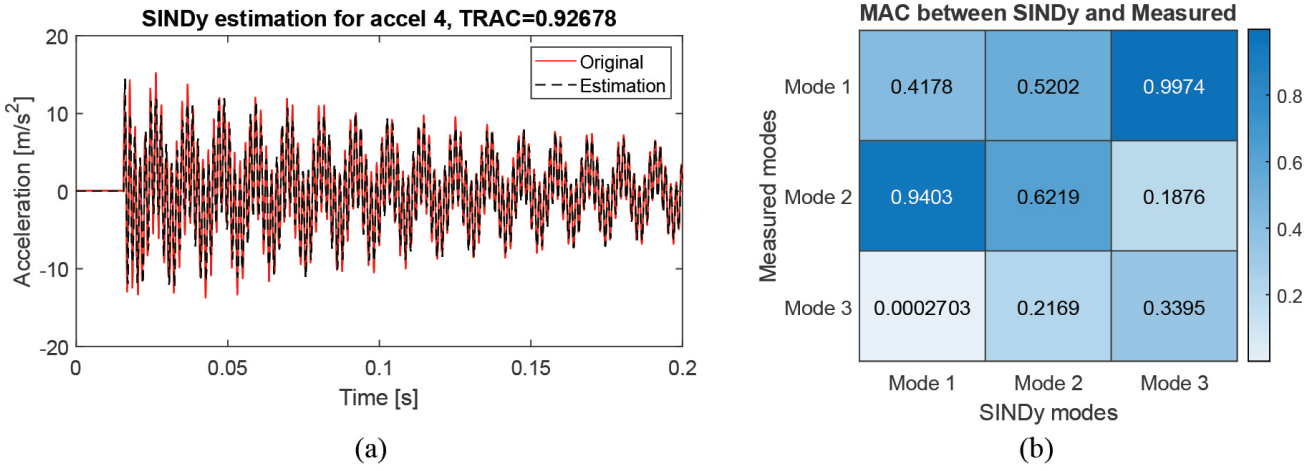


Fig. 3 Results from SGD SINDy; (a) Signal reconstruction of accelerometer 4 and (b) MAC between measurement and eigenvectors from SINDy

In addition, this approach had the gradient annealing implemented in the code, which led to the best selection of each of the weights of the penalty terms. The change in weight versus the number of epochs is given in Figure 4.

In Figure 4, the weights for the three penalty terms are presented. In the top left corner, the weight for the regularization factor is presented as $\nabla_\theta L_1$ is shown. The initial estimation for the weight is 1, which then will change based on each epoch until it converges to a value close to 0.1. On the top right, the values for the symmetry term are presented as $\nabla_\theta Symmetry$. This term has a different behavior than expected, while previously a small value was set as the symmetry penalty, when using learning rate annealing, the weight converges to a much higher value, on the 10,000 orders. Finally, the bottom part of the figure presents the weight value for the perturbation model. Since the symmetry term had a high value for the weights, the resulting matrices all had symmetric values, as shown in Equation (20).

$$M_{SINDy} = \begin{bmatrix} 0.038 & -0.0376 & 0.0182 \\ -0.0376 & 0.0388 & -0.0223 \\ 0.0182 & -0.0223 & 0.0232 \end{bmatrix} \text{ and } K_{SINDy} = \begin{bmatrix} 6.745 & 0 & -0.992 \\ 0 & -1.036 & -0.337 \\ -0.992 & -0.337 & 4.055 \end{bmatrix} 10^3 \tag{20}$$

In addition, the perturbation method used estimates the eigenvalue of the system, allows a good estimation of the natural frequencies, based on the eigen solution of the mass and stiffness matrices obtained in Equation (20). The comparison between the natural frequencies obtained via SINDy and the measured natural frequencies are presented in Table 2.

As presented in the table, the first natural frequency has an almost perfect estimation with errors in the decimals. While the second natural frequency, which has a much smaller amplitude and is more difficult to obtain, has an error smaller than 5%. Finally, the SINDy model was compared between linear and nonlinear problems. As presented previously, the SINDy model was employed in the linear example, which led to the results presented. In addition, a second experiment was generated, where the boundary condition is soft and has a new mode shape between the two mode shapes obtained. When generating the library of parameters, a cubic spring was implemented for both cases, while one of them has a linear behavior

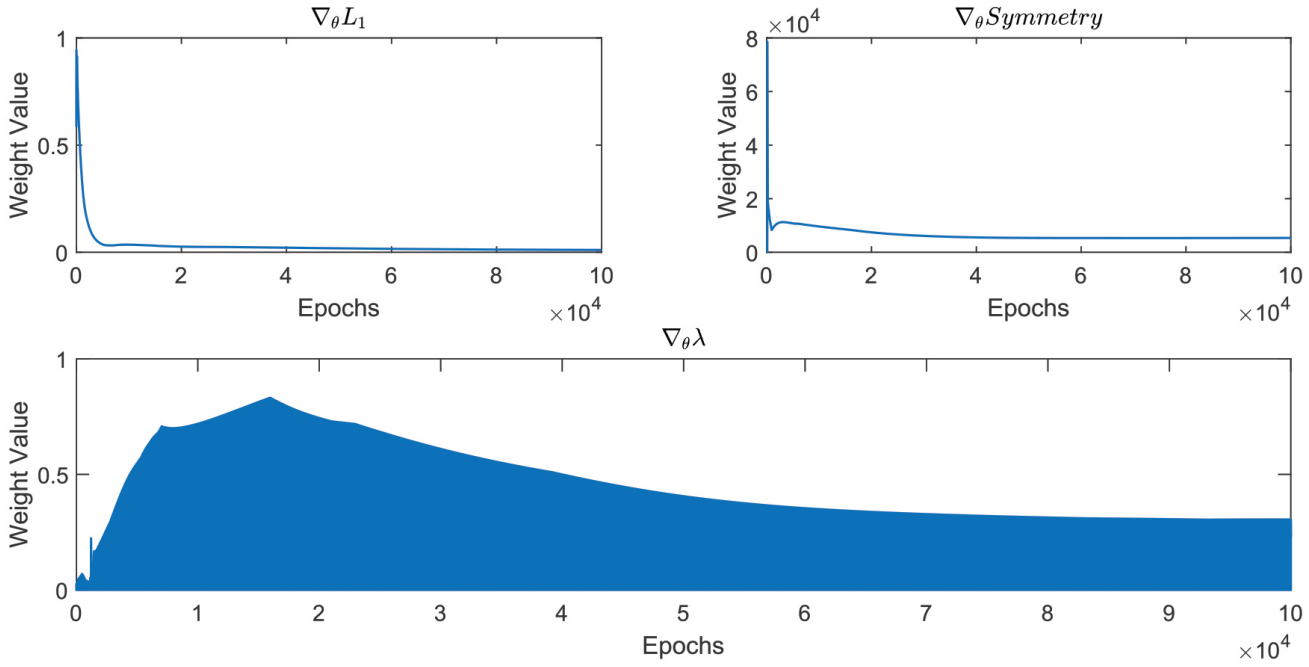


Fig. 4 Weights of learning rate annealing versus epochs while training SGD SINDy

Table 2 Natural frequency estimation using SGD SINDy

Measured Frequency [Hz]	SINDy Estimation [Hz]	% Error
89.6	89.7	0.1
576.3	591.2	2.6

and is expected to have zero values for the cubic spring. The second application has nonlinearities from the boundary conditions and is expected to have more terms to describe the behavior of the system accurately. In Equation (21) the comparison between the cubic spring matrices is given.

$$[K_3]_{linear} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3.12 \end{bmatrix} \times 10^{10} \text{ and } [K_3]_{nonlinear} = \begin{bmatrix} 1.93 & 0.07 & -0.15 \\ 0 & -0.49 & -0.07 \\ -0.16 & -0.08 & 0.07 \end{bmatrix} \times 10^{12} \tag{21}$$

The nonlinear condition has terms in almost every position of the stiffness matrix. Which is needed to maintain a good accuracy of the model. However, when the same nonlinear spring is implemented in the linear case, most of its values become zero, with the exception of one of the terms. However, when observing the amplitude of this value, it is two orders of magnitude smaller than the ones obtained by the nonlinear case. Thus, the value of the linear would be smaller than the smallest value obtained by the nonlinear case.

Conclusion

The proposed approach to modify the SINDy algorithm developed by Brunton et al. [7], is focused on obtaining second-order PDEs, more specifically, in continuous vibration, where there exists a dependence between the degrees of freedom in both mass and stiffness matrices. The modification of the algorithm to a stochastic gradient descent allowed the implementation of different penalty terms, such that the resulting PDE has a similar behavior to the one expected for a continuous vibration system. The reconstruction of the signal and the mode shapes estimated by the model have above 90% accuracy, and the eigenvalues estimated are also close to the natural frequencies measured. In addition, the model is capable of predicting nonlinear terms as long as they are included in the library of parameters.

The future work of the project can be focused on further improving the technique for different types of forced vibration tests. The current model works for impact tests, which are similar to a free vibration test. In addition, the number of sensors and natural frequencies measured must be similar. Investigation into implementing SINDy with modal reduction techniques can be a solution for better estimating the system dynamics.

Acknowledgments This work is supported by the Department of Energy Kansas City National Security Campus Managed by Honeywell under contract # N000510678.

References

1. C. Zhang, A. A. Mousavi, S. F. Masri, G. Gholipour, K. Yan, and X. Li, "Vibration feature extraction using signal processing techniques for structural health monitoring: A review," Mechanical Systems and Signal Processing, vol. 177, p. 109175, 2022.
2. Y.-J. Cha, R. Ali, J. Lewis, and O. Büyük?ztürk, "Deep learning-based structural health monitoring," Automation in Construction, vol. 161, p. 105328, 2024.
3. O. Nelles and O. Nelles, Nonlinear dynamic system identification. Springer, 2020.
4. F. Abdullah and P. D. Christofides, "Data-based modeling and control of nonlinear process systems using sparse identification: An overview of recent results," Computers & Chemical Engineering, vol. 174, p. 108247, 2023.
5. T. Liu and H. Meidani, "Physics-informed neural networks for system identification of structural systems with a multiphysics damping model," Journal of Engineering Mechanics, vol. 149, no. 10, p. 04023079, 2023.
6. 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, vol. 378, pp. 686-707, 2019.
7. S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Discovering governing equations from data by sparse identification of nonlinear dynamical systems," Proceedings of the national academy of sciences, vol. 113, no. 15, pp. 3932-3937, 2016.
8. C. Lathourakis and A. Cicirello, "Physics-enhanced sparse identification of nonlinear oscillator with coulomb friction," in International Conference on Nonlinear Dynamics and Applications, 2023: Springer, pp. 507-517.
9. C. Lathourakis and A. Cicirello, "Physics enhanced sparse identification of dynamical systems with discontinuous nonlinearities," Nonlinear Dynamics, pp. 1-28, 2024.
10. S. Piramoon, M. Ayoubi, and S. Bashash, "Modeling and Vibration Suppression of Rotating Machines Using the Sparse Identification of Nonlinear Dynamics and Terminal Sliding Mode Control," IEEE Access, 2024.
11. S.-C. Huang, H.-W. Chen, and M.-H. Tien, "Predicting Nonlinear Modal Properties by Measuring Free Vibration Responses," Journal of Computational and Nonlinear Dynamics, vol. 18, no. 4, p. 041005, 2023.
12. K. Kaheman, J. N. Kutz, and S. L. Brunton, "SINDy-PI: a robust algorithm for parallel implicit sparse identification of nonlinear dynamics," Proceedings of the Royal Society A, vol. 476, no. 2242, p. 20200279, 2020.
13. P. Goyal and P. Benner, "Discovery of nonlinear dynamical systems using a Runge–Kutta inspired dictionary-based sparse regression approach," Proceedings of the Royal Society A, vol. 478, no. 2262, p. 20210883, 2022.
14. Y. Liu, S. Shi, Y. Fang, and A. Fu, "Robust data-driven wave excitation force estimation for wave energy converters with nonlinear probabilistic modelling," Ocean Engineering, vol. 310, p. 118726, 2024.
15. K. Champion, B. Lusch, J. N. Kutz, and S. L. Brunton, "Data-driven discovery of coordinates and governing equations," Proceedings of the National Academy of Sciences, vol. 116, no. 45, pp. 22445-22451, 2019.
16. A. Greenbaum, R.-c. Li, and M. L. Overton, "First-order perturbation theory for eigenvalues and eigenvectors," SIAM review, vol. 62, no. 2, pp. 463-482, 2020.
17. I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. MIT Press, 2016.
18. S. Wang, Y. Teng, and P. Perdikaris, "Understanding and mitigating gradient flow pathologies in physics-informed neural networks," SIAM Journal on Scientific Computing, vol. 43, no. 5, pp. A3055-A3081, 2021.
19. S. S. Rao, Vibration of continuous systems. John Wiley & Sons, 2019.