

Bayesian Parameter Estimation of Vibrating Systems via Physics-Based Biases from Neural Networks

Marcus HAYWOOD-ALEXANDER ¹, Antonios KAMARIOTIS ¹, Giacomo ARCIERI ¹, Eleni CHATZI ¹

¹ ETH Zürich, Zürich, Switzerland, mhaywood@ethz.ch

Abstract. Often, Structural Health Monitoring (SHM) campaigns draw damage-identifiers from vibration-based monitoring data. One of the downstream tasks of vibration-based SHM is that of system identification, i.e., inference of a model which is able to describe the system. For vibration-based monitoring, such models often rely on partial-differential equation formulations, where the identification task may be tied to inference of the parameters of such models.

The probabilistic nature of SHM incentivises the extraction of stochastic estimates of these parameters, allowing for propagation of uncertainties associated to measurement noise or model imprecision. Simple approaches to generating such stochastic estimates require an assumption on the shape and independence of the posterior distributions. To circumvent these assumptions, a popular approach is the Hamiltonian Monte-Carlo (HMC) sampling approach, which simulates the posterior distribution, allowing for a comparatively larger degree of freedom in the posterior. However, this requires intensive computational effort, especially when the governing equations present increased complexity, such as inclusion of nonlinearities.

Physics-informed neural networks (PINNs) utilise physics-based objective functions in the form of known PDEs, exploiting the automatic differentiation of NN architectures. This implies that additional terms in the PDE, such as nonlinearities, or boundary conditions, can be accounted for. When the domain of observations encapsulates the collocation domain of the PINN, they can be exploited for parameter identification, by optimising over these parameters in addition to the network weights. However, classic objective functions in the form of mean-squared-norms underperform on noisy data, because of their deterministic nature.

In this work, PINN frameworks are combined with a HMC sampler to estimate the posterior distribution of the parameters of dynamic systems with noisy measurements. Noisy measurements from a nonlinear multi-degree-of-freedom oscillator are used as inputs to the framework. The results are assessed against the known values, and the correlated posterior distributions are discussed.

Keywords: physics-informed, Bayesian inference, parameter estimation, neural network



Introduction

One of the foundations of structural health monitoring (SHM) is the identification of features which are indicative of its health state [1, 2]. These identification/estimation schemes are also incentivised to be stochastic in their representation, as a result of the probabilistic nature of SHM [3]. This representation allows the propagation of uncertainties associated with measurement noise or model imprecision. Often, such campaigns draw from vibration-based monitoring data for damage-sensitive features [4]. In vibration-based SHM, one of the downstream tasks is that of system identification, where the identification task may be tied to that of parameter estimation [5]. Identification schemes which operate in the time-domain are often performed by solving the proposed solution of the vibration given specific conditions [6]. For complex dynamic systems, such as those involving nonlinearities, the governing equations cannot be solved analytically, and so the proposed solutions for a large time window are computationally intense, often requiring numerical or FE-based solvers [7]. Vibrating systems are often expressed in the form of ordinary- or partial-differential equations, which describe the system via balancing of the forces acting on, and as a result of the motion of, the system in the time domain [8].

As stated above, there exists a strong incentivisation for generating a probabilistic estimate of damage-sensitive features [9]. Vibration-signal measurements often are corrupted with errors, such as noise. Simple approaches to generating a stochastic estimation of the system parameters require an assumption on both the posterior shape and independence of parameters. This can ignore important information such as codependence [10, 11]. To circumvent these assumptions, a popular approach is the Hamiltonian Monte-Carlo [12], a type of Markov Chain Monte-Carlo [13] algorithm; which is a sampling procedure which allows one to emulate exploration over the posterior distributions. This allows for comparatively greater freedom in the posterior, and can capture dependence structures in the posterior, even when the priors are assumed dependent. However, as these often require an extensive number of samples, which increases along with the number of parameters, efficient computation of the likelihood is imperative [14]. When the estimated means of the likelihood are determined from the solutions of ODE/PDEs with additional complexities, such as nonlinearities, this can become computationally expensive.

Physics-informed neural networks (PINNs) are an approach to estimate the solutions to systems which are defined by equations in the form of ODE/PDEs, by embedding these equations as additional loss terms to be optimised towards [15]. PINNs exploit the automatic differentiation capabilities of modern machine learning frameworks, and have gained significant attention in recent years as a result of the ease of implementation of additional complexities such as nonlinearities or boundary conditions [16]. By optimising over the system parameters along with the network weights, PINNs can also be exploited for system identification [17]. However, these can underperform on noisy data with classic mean-squared-error loss functions, as they provide a deterministic estimate. It is important to note that in this context, PINNs are limited in performing ‘instance’ modelling, in that they do not offer a model that can be reused for new data, and given new initial and forcing conditions, the minimisation procedure must be performed again. Pre-trained neural networks have been exploited for rapid, or in-time, output estimates of the solutions of complex physical equations [18]. By combining this motivation with the physics-equation solution that is embedded within PINNs, we here propose a framework for estimating the posterior distribution of vibrating systems from time-domain signals using HMC. In this framework, the neural network is posited to estimate the state-signal, and then the solutions of the physics equation can be estimated using the PINN-based method.

To demonstrate the framework, noisy data for a MDOF Duffing oscillators are generated from numerical simulations, and the posterior distributions of the parameters which govern the ODE equations are simulated. The paper begins with a definition of the simulated model, then discusses the combined NN-based HMC sampler, and finishes with the results of the simulated data having been passed through the framework.

1. Multi-Degree of Freedom Model

In this work, a generic multi-degree of freedom oscillator is employed as a working example, defined as a sequentially-connected n -DOF Duffing oscillator with a fixed connection at the first DOF, as shown in Fig. 1. The state space model of the system is,

$$\dot{\mathbf{z}} = \mathbf{A}\mathbf{z} + \mathbf{A}_n\mathbf{z}_n + \mathbf{H}\mathbf{f} \quad (1)$$

where $\mathbf{z} = \{\mathbf{u}, \dot{\mathbf{u}}\}^T \in \mathbb{R}^{2n}$ is the system state, and $\mathbf{f} \in \mathbb{R}^n$ is the force vector. An additional ‘nonlinear-state function is applied to account for the cubic stiffness;

$$\mathbf{z}_n = g_n(\mathbf{z}) = \{u_1^3, (u_2 - u_1)^3, \dots, (u_n - u_{n-1})^3\}^T \quad (2)$$

Moreover, it is assumed for simplicity that $m_n = m$ and the system matrices are defined as:

$$\mathbf{A} = \begin{bmatrix} 0 & \mathbf{I} \\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{C} \end{bmatrix}, \quad \mathbf{A}_n = \begin{bmatrix} 0 \\ -\mathbf{M}^{-1}\mathbf{K}_n \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} 0 \\ \mathbf{M}^{-1}\mathbf{I} \end{bmatrix} \quad (3)$$

and the linear and nonlinear stiffness and damping matrices, $\mathbf{K}$, $\mathbf{K}_n$, $\mathbf{C}$, are easily derived but are omitted here for brevity.

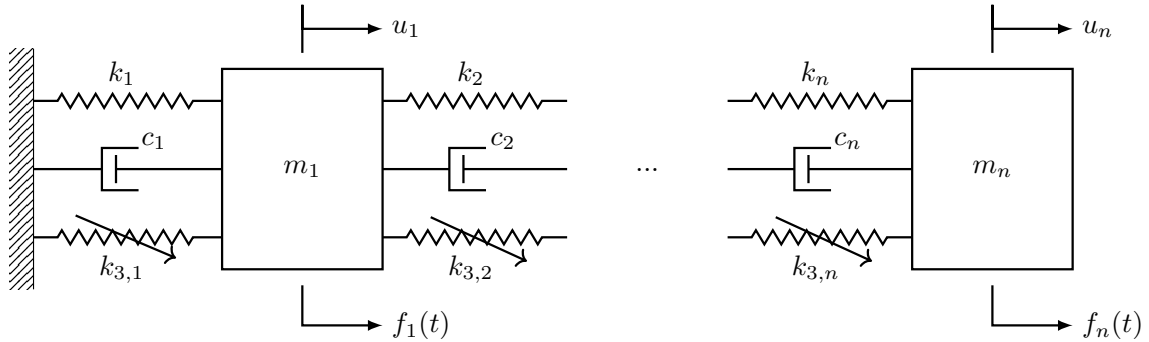


Fig. 1. n -DOF system

1.1 Simulated Model

For the work in this paper, a 2 degree-of-freedom (2DOF) system was simulated. The forcing $\mathbf{f}$ applied was white Gaussian noise, banded between frequencies of 0.1 and 0.5 Hz, with an amplitude of 1.0 N, applied to each degree of freedom. The values of the parameters are consistent at $m_i = 10\text{kg}$, $c_i = 1.0 \text{ N s m}^{-1}$, $k_i = 15 \text{ N m}^{-1}$, and $k_{n,i} = 100 \text{ N m}^{-3}$. The systems were then simulated with a fourth-order Runge-Kutte integration, over a 240s window with 4096 samples, and Gaussian noise added to the signal representing a signal-to-noise ratio of 20.0. Each simulation was repeated 5 times, with different instances of the added noise by changing the seed for the random-number-generator. These noisy data are then used as the observations throughout the paper.

2. Bayesian Identification of MDOF System

Consider the concept of an elementary system-identification procedure to estimate the set of system parameters $\boldsymbol{\theta}_s$ given noisy observations of the force, $\mathbf{f}^*$, a maximum-likelihood approach can be formulated to return the most likely parameters [19]. In this case, the model is assumed to be of the form,

$$\mathbf{f}^* = \mathcal{F}(\mathbf{z}, \boldsymbol{\theta}_s) + \varepsilon(0, \sigma_f) \quad (4)$$

where ε is white Gaussian noise, and $\mathcal{F}$ is the nonlinear differential operator which estimates the mean force, and is defined from Equation 1,

$$\mathcal{F}(\mathbf{z}, \boldsymbol{\theta}_s) = \hat{\mathbf{f}} = \hat{\mathbf{H}}^{-1}(\dot{\mathbf{z}} - \mathbf{A}\mathbf{z} - \mathbf{A}_n\mathbf{z}_n) \quad (5)$$

where $\hat{\mathbf{H}}^{-1} = [0, \mathbf{M}]^T$ is the quasi-inverted input matrix. Given a set of assumed-independent observations of state-force pairs, $\mathbf{y}^* = \{\mathbf{z}^*, \mathbf{f}^*\}$ the likelihood is then defined as,

$$L(\mathbf{y}^*|\boldsymbol{\theta}_s) = \prod_{i=1}^{n_s} \frac{1}{\sqrt{2\pi\sigma_f^2}} \exp\left(-\frac{\|\mathcal{F}(\mathbf{z}(t^{(i)}), \boldsymbol{\theta}_s) - \mathbf{f}^*(t^{(i)})\|^2}{2\sigma_f^2}\right) \quad (6)$$

where n_s is the number of observations. Simply maximising $L(\mathbf{y}^*|\boldsymbol{\theta}_s)$ provides a deterministic estimate of the most likely elastic constants; however, this is a limited level of information. Therefore, here we adopt an approach which makes it is also possible to retrieve information on the *distribution* of the parameters.

2.1 Sampling Over the Posterior Distribution

Given the set of variable parameters, which includes the system parameters and noise, $\boldsymbol{\Theta} = \{\boldsymbol{\theta}_s, \sigma_f\}$, Hamiltonian Monte-Carlo (HMC) sampling allows one to characterise the distribution of these parameters when the posterior distribution is intractable. With HMC sampling, the distribution is estimated by random sampling from the posterior distribution, where consecutive samples are chosen based on the posterior probability. A detailed description of HMC sampling is given by Barber [20], although a brief overview will be given here. During the iterative sampling procedure, the transition from on sample to the next is governed by the Metropolis-Hastings acceptance criterion. Given a target posterior distribution $p(\boldsymbol{\Theta}|\mathbf{y}^*)$, the acceptance ratio is defined as,

$$\hat{\alpha} = \frac{p(\mathbf{y}^*|\boldsymbol{\Theta}')p(\boldsymbol{\Theta}')}{p(\mathbf{y}^*|\boldsymbol{\Theta}_k)p(\boldsymbol{\Theta}_k)} \quad (7)$$

Where $\boldsymbol{\Theta}'$ is the current estimate and $\boldsymbol{\Theta}_k$ is the previous estimate, $p(\boldsymbol{\Theta})$ is the prior distribution of $\boldsymbol{\Theta}$, and $p(\mathbf{y}^*|\boldsymbol{\Theta})$ is the likelihood. The proposed estimate is accepted if $\hat{\alpha} \leq 1$. In addition to the likelihood, the priors must be defined to embed any prior belief of the distribution of the parameters.

3. Modelling the Physics-bias Using Neural Networks

3.1 PINN Framework for MDOF system

In this framework, the neural network (NN) is posed to estimate the state of the MDOF system, given an individual time point of the signal,

$$\mathcal{N}_{\mathbf{z}}(t; \boldsymbol{\theta}_w) \quad (8)$$

where $\boldsymbol{\theta}_w$ is the network weights to be optimised over. The training of the network is then accomplished by minimising an objective function which guides the network output using observed values in the same domain. The most common approach of black-box NNs is to use the mean-squared error, which aims to reduce the distance between the estimated and observed values,

$$L_o = \langle \mathcal{N}_{\mathbf{z}} - \mathbf{z}^* \rangle, \quad \langle \bullet \rangle = \frac{1}{n_s} \sum_{i=1}^{n_s} \| \bullet \|^2 \quad (9)$$

Given a set of noisy observations of the state,

$$\mathbf{z}^*(t) = \mathbf{z}(t) + \varepsilon(0, \sigma_z) \quad (10)$$

one can adapt the objective function based on the likelihood form. For optimisation purposes the objective is formed as the negative log likelihood,

$$L_o = -\log(p(\mathbf{z}^* | \boldsymbol{\theta}_w)) = N_o \log(\sigma_z) + \frac{N_o}{2} \log(2\pi) + \frac{1}{2} \sum_{i=1}^{N_o} \frac{\| \mathbf{z}^* - \mathcal{N}_{\mathbf{z}}(t^{(i)}; \boldsymbol{\theta}_w) \|^2}{\sigma_z^2} \quad (11)$$

In PINN frameworks, if the physics of the system is known in the form of ordinary or partial differential equations, then this can be used to form additional objective terms. These are prescribed such that when they equal zero, the physics are satisfied. These methods exploit the capabilities of modern automatic differentiation to estimate the value of derivative terms in the PDE. These offer advantages over other methods such as numerical differentiation techniques, as they avoid numerical instabilities that often occur, particularly when estimating higher order derivative terms. For the MDOF system, the likelihood-based physics loss is formed using Eq. 6,

$$p(\mathbf{f}^* | \boldsymbol{\theta}_w, \boldsymbol{\theta}_s) = \prod_{i=1}^{N_c} \frac{1}{\sqrt{2\pi\sigma_f^2}} \exp \left(-\frac{1}{2} \frac{\| \mathbf{f}^* - \tilde{\mathbf{f}}(t^{(i)}, \boldsymbol{\theta}_s, \boldsymbol{\theta}_w) \|^2}{\sigma_f^2} \right) \quad (12)$$

$$L_p = -\log(p(\mathbf{f}^* | \boldsymbol{\theta}_w, \boldsymbol{\theta}_s)) = N_c \log(\sigma_f) + \frac{N_c}{2} \log(2\pi) + \frac{1}{2} \sum_{i=1}^{N_c} \frac{\| \mathbf{f}^* - \tilde{\mathbf{f}}(t^{(i)}, \boldsymbol{\theta}_s, \boldsymbol{\theta}_w) \|^2}{\sigma_f^2} \quad (13)$$

$$\tilde{\mathbf{f}}(t) = \mathcal{F}(\mathcal{N}_{\mathbf{z}}(t; \boldsymbol{\theta}_n), \boldsymbol{\theta}_s) = \widehat{\mathbf{H}}^{-1}(\partial_t \mathcal{N}_{\mathbf{z}} - \mathbf{A} \mathcal{N}_{\mathbf{z}} - \mathbf{A}_n g_n(\mathcal{N}_{\mathbf{z}})) \quad (14)$$

where $\partial_t \mathcal{N}_{\mathbf{z}}$ is the estimated first order derivative of the state using automatic differentiation.

3.2 Embedding PINN framework for parameter estimation

In discussing this framework, we here use terms representing different forms of bias embedding [19, 20]. Without prior knowledge of the system parameters, the **observational bias** is the primary influence on the values of the network weights – and therefore estimated state values. When exploiting PINNs for parameter estimation, with no effort for domain extension or interpolation, the **inductive bias** introduced by the physics-knowledge, in the form of the differential operator $\mathcal{F}$, primarily drives the solution of the estimated system parameters. As the aim here is to estimate the posterior distribution of the system parameters, we first produce a deterministic estimate of the stochastic representation of the state in Equation 10. The physics-based inductive bias in Equation 6 is estimated using Equation 14 and is then used as the posterior likelihood for the HMC sampling procedure. This posterior sampling procedure then induces a **learning bias** in the form of assuming a stochastic representation of the system parameters. The neural network is built, and physics equation solution estimated, using PyTorch [23], and the HMC procedure is run using Pyro [24], employing the NUTS sampling algorithm [25].

4. Results

4.1 State and uncertainty estimation

The results of the estimated state and the uncertainty in the state are shown in Fig x, along with the uncertainty shown as the 2σ interval. It is clear to the results align well with the true unnoisy signal. This is not the main contribution of this work, as it is trivial, however it is shown here to ensure there is no effect of invalid state estimate on the parameter estimation stage.

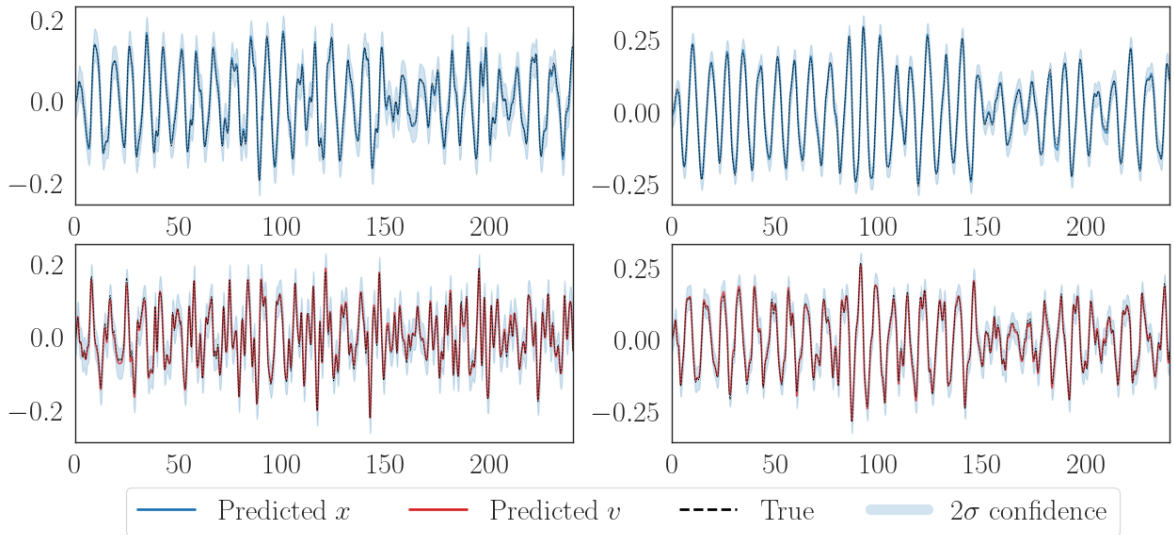


Fig. 2. Results of the state estimation as the output of the neural network, along with the estimated confidence.

4.2 Parameter estimation

The results of 5000 iterations of the sampling procedure applied to the 2DOF system are shown in Fig. 3. The diagonal elements of the gridded array show the univariate sample distributions for the i^{th} parameter, the upper triangle show the multivariate sample distributions comparing the i^{th} and j^{th} parameter, and the lower triangle shows the bivariate kernel density estimate of the cross-correlation between parameters, where the lighter colours represent a lower value of the probability density. Table 1 also shows the calculated mean,

along with the relative error with respect to the true value, and the calculated standard deviation, of the posterior samples of each parameter.

Table 1. Estimated parameters taken from kernel density estimates of the sampled posterior distributions.			
Parameter	Mean		Calculated Standard Deviation
	Calculated	Relative Error	
c_1	1.018	1.80%	0.045
c_2	0.913	8.7%	0.030
k_1	14.76	1.60%	0.090
k_2	14.74	1.77%	0.080
$k_{3,1}$	94.16	5.84%	5.78
$k_{3,2}$	78.66	21.34%	6.96
σ_f	0.0689	-	0.0005

From Fig. 3 one can see the strong correlation between the linear and nonlinear stiffness at the same connection points, (i.e. k_1 with $k_{3,1}$ and k_2 with $k_{3,2}$) which is to be expected as they are both dependent on the relative displacement of that connection. Other than this, there

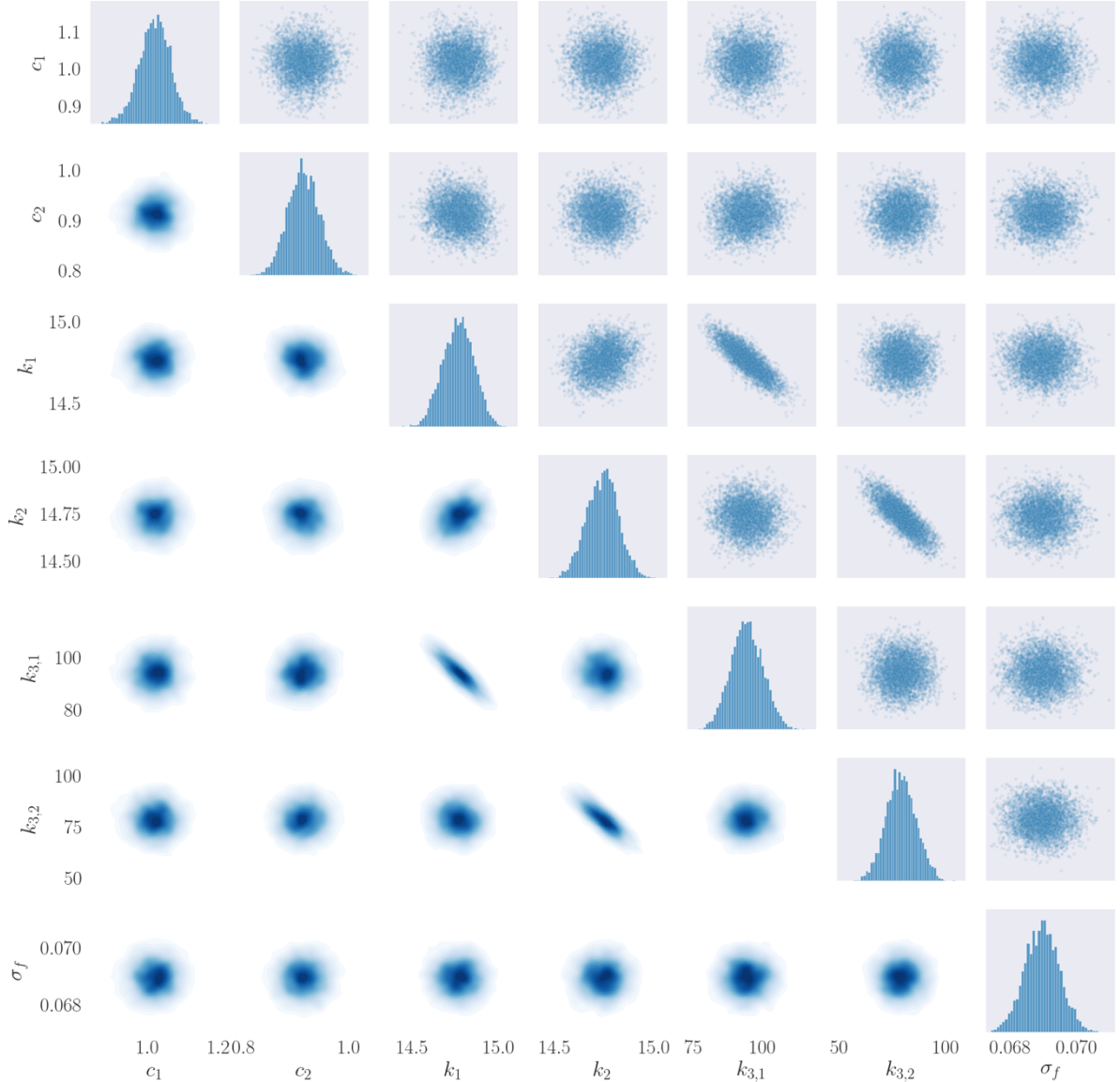


Fig. 3. Results of the parameter identification procedure applied to the 2DOF Duffing oscillator.

appears no correlation between parameters. This matches well with prior knowledge on the correlated distribution of parameters for a nonlinear MDOF oscillator [9, 26].

5. Conclusion

In conclusion, this paper presents a novel approach combining Physics-Informed Neural Networks (PINNs) with Hamiltonian Monte Carlo (HMC) sampling for estimating the posterior distribution of parameters in dynamic systems, particularly focusing on Structural Health Monitoring (SHM) applications. By leveraging the flexibility of PINNs to incorporate physics-based objective functions and the stochastic nature of HMC sampling, this method offers a promising solution for parameter identification in the presence of noisy measurements. Through application to a nonlinear multi-degree-of-freedom oscillator, the effectiveness of the proposed framework is demonstrated, with results showing correlated posterior distributions that align well with prior knowledge. This integrated approach not only bypasses the limitations of traditional deterministic objective functions but also easily accommodates complex system dynamics, making it a valuable tool for robust SHM campaigns and system identification tasks. Further exploration and refinement of this methodology hold great potential for advancing the field of vibration-based monitoring and uncertainty quantification in dynamic systems.

Acknowledgements

The authors gratefully acknowledge the funding from the Swiss National Science Foundation (SNSF) under the Horizon Europe funding guarantee, for the project ‘ReCharged - Climate-aware Resilience for Sustainable Critical and interdependent Infrastructure Systems enhanced by emerging Digital Technologies’ (grant agreement No: 101086413).

References

- [1] C. R. Farrar and K. Worden, ‘An introduction to structural health monitoring’, *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.*, vol. 365, no. 1851, pp. 303–315, 2007.
- [2] D. Balageas, C.-P. Fritzen, and A. Güemes, *Structural health monitoring*, vol. 90. John Wiley & Sons, 2010.
- [3] C. R. Farrar and K. Worden, *Structural health monitoring: a machine learning perspective*. John Wiley & Sons, 2012.
- [4] D. Goyal and B. Pabla, ‘The vibration monitoring methods and signal processing techniques for structural health monitoring: a review’, *Arch. Comput. Methods Eng.*, vol. 23, pp. 585–594, 2016.
- [5] J. P. Amezcua-Sanchez and H. Adeli, ‘Signal processing techniques for vibration-based health monitoring of smart structures’, *Arch. Comput. Methods Eng.*, vol. 23, pp. 1–15, 2016.
- [6] K. F. Tee, ‘Time series analysis for vibration-based structural health monitoring: A review’, *Struct. Durab. Health Monit.*, vol. 12, no. 3, pp. 129–147, 2018.
- [7] R. Astroza, *Vibration-based health monitoring and mechanics-based nonlinear finite element model updating of civil structures*. University of California, San Diego, 2015.
- [8] H. Chen, M. Kurt, Y. S. Lee, D. M. McFarland, L. A. Bergman, and A. F. Vakakis, ‘Experimental system identification of the dynamics of a vibro-impact beam with a view towards structural health monitoring and damage detection’, *Mech. Syst. Signal Process.*, vol. 46, no. 1, pp. 91–113, 2014.

- [9] T. J. Rogers and T. Friis, ‘A latent restoring force approach to nonlinear system identification’, *Mech. Syst. Signal Process.*, vol. 180, p. 109426, 2022.
- [10] S.-K. Au and F.-L. Zhang, ‘On assessing the posterior mode shape uncertainty in ambient modal identification’, *Probabilistic Eng. Mech.*, vol. 26, no. 3, pp. 427–434, 2011.
- [11] M. Haywood-Alexander, N. Dervilis, K. Worden, R. S. Mills, P. Ladpli, and T. J. Rogers, ‘A Bayesian method for material identification of composite plates via dispersion curves’, *Sensors*, vol. 23, no. 1, p. 185, 2022.
- [12] M. Betancourt, ‘A conceptual introduction to Hamiltonian Monte Carlo’, *ArXiv Prepr. ArXiv170102434*, 2017.
- [13] D. Gamerman and H. F. Lopes, *Markov chain Monte Carlo: stochastic simulation for Bayesian inference*. Chapman and Hall/CRC, 2006.
- [14] M. Bicher, M. Wastian, D. Brunmeir, and N. Popper, ‘Review on Monte Carlo simulation stopping rules: how many samples are really enough?’, *Simul Notes Eur*, vol. 32, no. 1, pp. 1–8, 2022.
- [15] 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’, *J. Comput. Phys.*, vol. 378, pp. 686–707, 2019.
- [16] S. Moradi, B. Duran, S. Eftekhar Azam, and M. Mofid, ‘Novel Physics-Informed Artificial Neural Network Architectures for System and Input Identification of Structural Dynamics PDEs’, *Buildings*, vol. 13, no. 3, p. 650, 2023.
- [17] S. A. Faroughi *et al.*, ‘Physics-guided, physics-informed, and physics-encoded neural networks in scientific computing’, *ArXiv Prepr. ArXiv221107377*, 2022.
- [18] T. Simpson, N. Dervilis, and E. Chatzi, ‘Machine learning approach to model order reduction of nonlinear systems via autoencoder and LSTM networks’, *J. Eng. Mech.*, vol. 147, no. 10, p. 04021061, 2021.
- [19] L. Le Cam, ‘Maximum likelihood: an introduction’, *Int. Stat. Rev. Int. Stat.*, pp. 153–171, 1990.
- [20] D. Barber, *Bayesian reasoning and machine learning*. Cambridge University Press, 2012.
- [21] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, ‘Physics-informed machine learning’, *Nat. Rev. Phys.*, vol. 3, no. 6, pp. 422–440, 2021.
- [22] C. Lathourakis and A. Cicirello, ‘Physics Enhanced Sparse Identification of Dynamical Systems with Discontinuous Nonlinearities’, 2023.
- [23] A. Paszke *et al.*, ‘Automatic differentiation in PyTorch’, 2017.
- [24] E. Bingham *et al.*, ‘Pyro: Deep universal probabilistic programming’, *J. Mach. Learn. Res.*, vol. 20, no. 28, pp. 1–6, 2019.
- [25] M. D. Hoffman, A. Gelman, and others, ‘The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo.’, *J Mach Learn Res*, vol. 15, no. 1, pp. 1593–1623, 2014.
- [26] J. Brynjarsdóttir and A. O’Hagan, ‘Learning about physical parameters: The importance of model discrepancy’, *Inverse Probl.*, vol. 30, no. 11, p. 114007, 2014.