



# Physics-augmented neural networks for predicting nonlinear vibration energy harvesters

Catarina L. M. I. Barros<sup>1</sup> · Arthur Barbosa<sup>3</sup> · Rafael de O. Teloli<sup>1</sup> · Najib Kacem<sup>3</sup> · Nouredine Bouhaddi<sup>3</sup> · Paulo S. Varoto<sup>2</sup>

Received: 29 August 2025 / Accepted: 11 February 2026  
© The Author(s) 2026

## Abstract

This study investigates the use of physics-augmented neural networks to predict the frequency-response behavior of a nonlinear piezoelectric vibration energy harvester (PVEH). An analytical model is first developed under the assumption of geometrical linearity, incorporating a Duffing-type nonlinearity arising from the magnetic interaction, referred to as the baseline model. The physical parameters are calibrated using Particle Swarm Optimization (PSO) and the PVEH's behavior is experimentally validated, based on frequency response curves collected from low to high acceleration levels. While the baseline model predicts the PVEH behavior accurately at low accelerations, it fails at higher accelerations where nonlinear effects become significant. To extend the model's range of applicability, a physics-augmented neural network is introduced, resulting in an augmented model that integrates physical knowledge through loss and activation functions. Unlike existing data-driven approaches that typically focus on isolated response features, the proposed framework enables prediction of the frequency response in the presence of nonlinearities not accounted for in the analytical development. Performance evaluation within the -3 dB bandwidth region shows that the mean absolute percentage error (MAPE) of the baseline model is 81.1%, whereas the non-optimized and optimized augmented models achieve 7.8% and 2.2%, respectively. Furthermore, a purely data-driven neural network was implemented to evaluate the predictive capability of the augmented model. With a MAPE of 40.3%, the results underscore the importance of the custom loss terms used in the augmented model training. Given its superior performance, the augmented model presents a promising approach for enhancing predictive capabilities and optimizing vibration energy harvester design, extending its operational range and providing a new modeling tool for understanding the behavior of nonlinear PVEHs.

**Keywords** Piezoelectricity · Physics-augmented neural networks · Nonlinear oscillator · Vibration energy harvesting

Technical Editor: Alberto Di Matteo.

✉ Catarina L. M. I. Barros  
catarina.barros@femto-st.fr

Arthur Barbosa  
arthur.barbosa@femto-st.fr

Rafael de O. Teloli  
rafael.teloli@femto-st.fr

Najib Kacem  
najib.kacem@femto-st.fr

Nouredine Bouhaddi  
nouredine.bouhaddi@femto-st.fr

Paulo S. Varoto  
varoto@sc.usp.br

<sup>1</sup> Department of Applied Mechanics, Université Marie et Louis Pasteur, SUPMICROTECH, CNRS, FEMTO-ST Institute, 26 Rue de l'Épitaphe, 25000 Besançon, France

<sup>2</sup> Mechanical Engineering Department, São Carlos School of Engineering, University of São Paulo, Avenue Trabalhador São-Carlense, 400, 13566-590 São Carlos, São Paulo, Brazil

<sup>3</sup> Department of Applied Mechanics, Université Marie et Louis Pasteur, CNRS, FEMTO-ST Institute, 26 Rue de l'Épitaphe, 25000 Besançon, France

## 1 Introduction

In recent years, the search for sustainable energy generation has gained attention, driven by the pressing challenge of pollution associated with traditional energy production [1]. Concurrently, there is an increasing focus on developing miniaturized, flexible, and low-power devices that do not depend on conventional batteries, which might present limited lifespan [2]. Within this context, energy harvesting solutions have been emerging as viable alternatives for addressing both energy sustainability and device fatigue life efficiency [3].

Energy harvesting enables the conversion of energy from natural or human-generated phenomena into electrical energy. Various physical effects, such as motion, vibration, and pressure, can be utilized to capture energy from the environment and convert it into usable electrical power [4]. The primary mechanisms for transducing mechanical energy into electrical energy include electromagnetism [5–8], electrostatics [9, 10], triboelectricity [11–13], piezoelectricity [14–16] and hybrid strategies [17]. Among these, piezoelectric energy harvesting is particularly noteworthy due to its high power density, energy conversion efficiency, and simplicity to fabricate and implement [18, 19]. Consequently, there is a strong emphasis on developing efficient piezoelectric vibration energy harvesters (PVEHs) with enhanced power generation capabilities [20].

PVEHs are usually studied in systems that consist of beams with attached piezoelectric (PZT) patches. The beam structure enhances mechanical coupling by introducing additional mechanical stress to the piezoelectric material when the beam is in bending mode [21–23]. To increase bandwidth and therefore improve power generation efficiency in these systems, various design strategies have been employed, including multi-resonance [24, 25] and coupling effects [26, 27]. Another common strategy for bandwidth enhancement is the introduction of nonlinearities [28–31], with notable approaches including geometric optimization of the beam [32] and the use of magnetic interactions [33].

In order to accurately predict the behavior of a nonlinear vibration energy harvester, which might be useful for designing purposes, a validated physical model is essential [34]. Over the years, mathematical frameworks for modeling PVEHs in nonlinear regime of motion have been proposed. For instance, Sebald et al. [35] investigated a theoretical nonlinear Duffing oscillator for vibration energy harvesting, examining how electromechanical coupling influences the nonlinearities. Similarly, Mahmoudi et al. [36] developed a hybrid piezoelectric-electromagnetic energy harvester model that combines various sources of nonlinearity into a cubic stiffness term. Garg and Dwivedy [37] conducted both theoretical and experimental investigations of a cantilever

beam-based PVEH, taking into account nonlinear damping, as well as geometric and inertial nonlinearities. Additionally, Li et al. [38] developed a geometrically nonlinear model for cantilevered PVEHs subjected to base excitation, incorporating multiple sources of geometric nonlinearities to account for high displacement amplitudes. Recently, Alazemi and Kurdila [39] developed a model for hysteretically nonlinear piezoelectric beams using functional differential equations coupled with a Krasnosel'skii–Pokrovskii hysteresis operator, enabling the prediction of memory-dependent dynamics relevant to energy harvesting and vibration control. However, the simplifying assumptions adopted in the aforementioned models might constrain their generalization, often resulting in discrepancies between simulated and experimental results when the PVEH operates outside the initial operating conditions considered in the model. Such issue reflects a broader challenge in engineering modeling, where mathematical representations, although valuable for capturing general trends, might suffer from missing physics and simplifying assumptions, leading to discrepancies between predicted and observed behaviors.

In this context, data-driven techniques might offer a promising response to this challenge [40]. However, purely data-driven models often lack physical interpretation on how they achieve convergence, causing the identification of these black-box models to rely heavily on hyperparameter tuning and initialization choices, which can lead to physically inconsistent predictions, and also affect the generalization capacity outside the training domain [41]. To mitigate this issue, several strategies have been proposed to integrate data-driven algorithms with prior physical knowledge, incorporating physical laws and constitutive models into the algorithm's architecture [42]. Within this context, the goal is to combine strengths of both approaches, data and physics-based modeling [43]. Some commonly used strategies are: physics-informed models, where the system's equations are embedded in the loss function [44, 45]; hybrid modeling, where data-driven terms are used to identify missing physics [46, 47]; and physics-augmented models, where the data-driven model is constrained to satisfy physical and mechanical conditions [48]. Artificial neural networks (ANNs) are frequently employed in this context due to their capacity as universal approximators [49, 50].

Within the domain of physics-augmented neural networks, several studies demonstrate the advantages of integrating physical requirements into neural network-based models compared to their pure black-box counterparts, particularly for nonlinear dynamical systems and frequency-domain predictions [51]. For instance, Groote et al. [52] developed a physics-augmented neural network model for a slider-crank mechanism with partially unknown dynamics. This approach incorporates trainable physical layers that are

closely connected to the neural network to compensate for the missing dynamics. Benady et al. [53] trained an unsupervised neural network and validated it through two test cases: a non-quadratic potential in the small strain regime using synthetic optical fiber measurements, and a Mooney–Rivlin model in the hyperelastic case with synthetic digital image correlation observations. In their study, physical information is embedded through the loss function, the chosen architecture, and physics-guided initialization.

Other studies have proposed frameworks to ensure thermodynamic consistency or to overcome limitations related to partial observability. For example, Klein et al. [54] employed physics-augmented networks to address finite electro-elasticity problems by formulating the internal energy density as a convex neural network constrained by polyconvexity. Najera et al. [55] proposed a Hamiltonian-based framework using invertible neural networks and time-delay embeddings, enforcing physics consistency through the Euler–Lagrange formulation. Additionally, research has been devoted to improving existing physics-augmented neural network architectures, indicating the growing applicability of this strategy [56, 57]. In a recent study, Yinusa et al. [58] investigated nonlinear nanofluidic flow-induced vibrations using Euler–Bernoulli beam theory and Eringen’s nonlocal elasticity, combining analytical and numerical approaches with machine learning to optimize thermo-magnetic performance and identify key parameters for enhancing stability and reducing noise. The study highlighted promising applications in energy harvesting using similar mechanisms.

In the realm of PVEHs, machine learning techniques are increasingly being explored. For example, Cai and Harne [59] employed neural networks as surrogate models for a classification problem involving a nonlinear harvester with magnetic interaction. They generated a dataset based on an analytical model to train the network for classifying the system’s behavior. Similarly, Zhang et al. [60] created a dataset from experimental data for a wake-galloping PVEH, comparing the accuracy of various machine learning techniques in predicting the maximum displacement and RMS voltage of the system. Additionally, Peralta et al. [61] utilized a Kriging metamodel to optimize the shape of non-rectangular harvesters. Predicting full frequency response curves of nonlinear systems has attracted significant attention in the nonlinear dynamics community, where analytical, numerical, and reduced-order approaches have been explored to capture multivalued responses, nonlinear resonances, and backbone curves [62]. However, comparatively limited attention has been devoted to the use of physics-augmented data-driven techniques for predicting the full frequency response of nonlinear PVEHs, as existing studies tend to focus on specific response features rather than the complete

frequency-dependent behavior. It is worth noting that some features, although important for design purposes, such as the optimal electrical load, could be addressed in parallel with the system’s nonlinearity [63]. Therefore, analyzing frequency response curves is essential, as it provides comprehensive information that reflects the transition from linear to nonlinear behavior, capturing variations that single-point features alone cannot reveal. Predicting the full frequency response of nonlinear systems presents additional challenges, as it requires a comprehensive analysis of physical constraints, strategies to address the multiple solutions inherent in nonlinear systems, and the integration of physical information to ensure accurate predictions across a wide operational range.

Given this context, the present study focuses on a weakly nonlinear PVEH, where nonlinear effects are present but do not lead to jumps or bifurcations. As mentioned above, high-fidelity models for such PVEHs are available in the literature and can accurately capture their behavior; however, their complexity may not be justified in applications requiring fast evaluations, simplified control-oriented representations, or early-stage design studies. In these cases, low-fidelity models remain attractive due to their simplicity and reduced computational cost, despite their limited accuracy when nonlinear effects become more pronounced. In this work, an analytical model with a third-order stiffness term, referred to here as the baseline model, is developed and calibrated against experimental responses. While it matches experimental results at low acceleration amplitudes, its accuracy deteriorates at higher amplitudes, where the assumed nonlinear representation is insufficient. To address this limitation, we propose a general framework for augmenting the baseline physics-based model with data-driven components, specifically physics-augmented neural networks, to extend their validity over a wider operating range.

In the proposed framework, the physics-augmented neural network receives the electrical voltage response curve predicted by the baseline model as input and outputs an approximation of the experimental response, referred to as the augmented model. The loss function is specifically designed to emphasize accuracy within the -3 dB bandwidth region, which is of particular importance for energy harvesting applications. In addition, the formulation incorporates physical constraints, and the baseline model enforces physical principles. By integrating the physical insights from the baseline model with experimental data, this physics-augmented framework improves prediction accuracy, particularly in conditions where the former does not adequately capture the PVEH’s behavior. Accordingly, the augmented model is compared against both experimental results and the analytical model. The physics-augmented neural network demonstrates improved accuracy, making it a valuable

tool for predicting system behavior across a wider range of acceleration and frequency. This capability is particularly useful for the optimized design of energy harvesting mechanisms through enhanced bandwidth prediction, especially in contexts where not only specific features are required but also the full frequency response curve. The scope of this work is centered on pre-bifurcation regimes, where the uniqueness of the neural network solution is guaranteed for a given excitation frequency. Studies reporting the use of neural networks to address dynamics with strong nonlinearities suggest promising directions for extending the methodology presented herein [64–66]. This lies beyond the scope of the present study but is an important consideration for future work.

Based on the above, the paper is structured as follows. Section 2 presents the baseline analytical model of the nonlinear piezoelectric vibration energy harvester, including the simplifying assumptions, the electromechanical formulation, the derivation of a reduced-order model via modal expansion, and the analysis of its steady-state response. Section 3 outlines the experimental setup and the data acquisition procedure. Section 4 describes the parameter calibration using Particle Swarm Optimization and evaluates the predictive performance of the baseline model against experimental data under different acceleration levels. Section 5 introduces the physics-augmented neural network, describing the construction of the loss function, the training procedure, and the comparison between the augmented model and experimental results. Different neural network architectures are considered, allowing a comparison between gray-box model predictions and those obtained from the analytical formulation. Finally, Sect. 6 summarizes the main findings and outlines perspectives for future work.

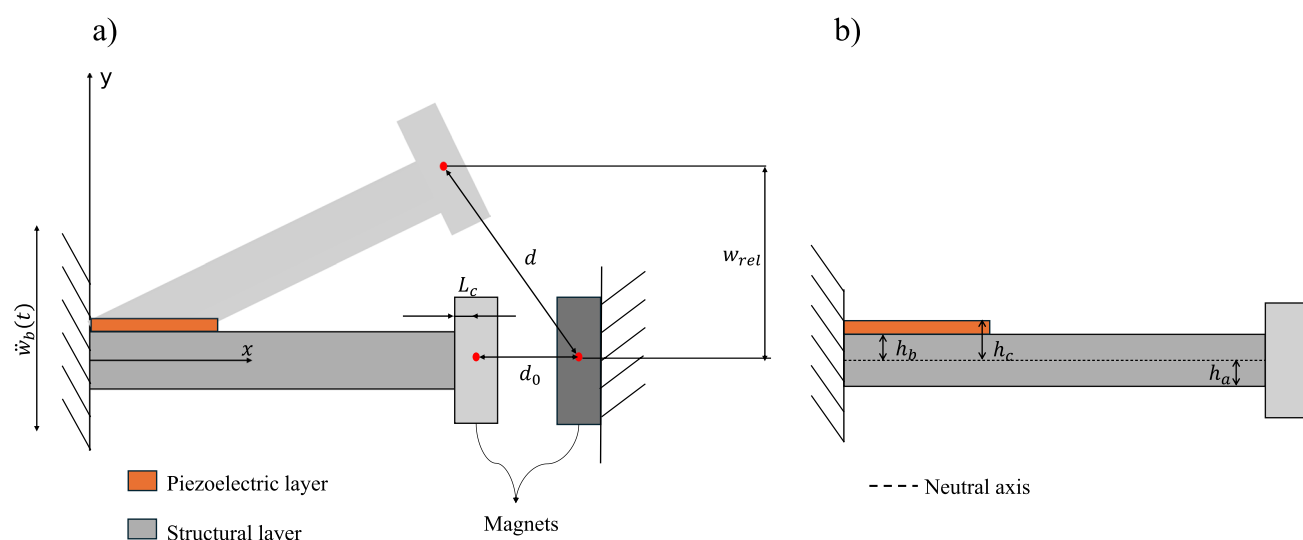
## 2 Physical system description and baseline model development

The development of a baseline model aims to provide an analytical framework for understanding the behavior of the studied PVEH, represented in Fig. 1. In the considered configuration, the cantilever beam is excited at the base by an imposed acceleration. To enhance energy generation, a piezoelectric layer is bonded near the clamped end. A tip mass, representing a magnet, is attached to the free end, with another magnet positioned nearby to induce magnetic interactions. The equations of motion and frequency response curves are derived following established formulations for PVEHs, resulting in a Duffing-type nonlinearity to represent the nonlinear stiffness effects. This model is intended to capture the system's dynamic response within a limited acceleration range, as determined by the simplifying assumptions adopted in its formulation. In what follows, 2.1 presents the hypotheses adopted in the model formulation. In 2.2, the Euler–Bernoulli beam model and the electromechanical coupling are introduced. Based on a modal expansion, 2.3 derives a reduced-order model of the PVEH. Finally, 2.4 presents the expressions associated with the steady-state frequency response of the PVEH.

### 2.1 Problem statement

The development of the baseline model relies on a set of simplifying assumptions aimed at reducing the complexity of the problem. The assumptions are as follows:

- the tip mass is modeled as a point mass;
- geometric nonlinearities are neglected;



**Fig. 1** **a** Schematic diagram of a cantilever beam under base excitation with a piezoelectric layer and magnetic interaction. **b** Distances with relation to the neutral axis

- the electromechanical coupling is assumed to be linear;
- the magnetic interaction is approximated by the force between two-point magnets;
- the Euler-Bernoulli beam theory assumptions are adopted;
- the boundary conditions are considered to be those of a beam clamped at one end and free at the other (cantilever beam);
- only the first vibration mode is considered; therefore, nonlinear modes are assumed to be uncoupled.

These assumptions constrain the model's applicability to scenarios involving low base acceleration levels and require specific design conditions, for instance, limiting the size of the tip mass. Additionally, they exclude the consideration of various forms of nonlinear electromechanical coupling, which may be relevant in more complex or high-energy systems.

Several studies have proposed more comprehensive models by relaxing the simplifying assumptions adopted in linear approaches, thereby extending their validity to a broader range of operating conditions. For instance, Jiajie Li et al. [38] studied how the inclusion of geometric nonlinearities allows the model to accurately predict system behavior at higher displacement amplitudes. Moreover, the tip mass is no longer treated as a point mass; instead, its distributed nature is taken into account by equivalent force and moment contributions at the beam tip, enabling improved accuracy when the point mass approximation is no longer valid. Similarly, in the work of Abdelkefi et al. [67], nonlinear electromechanical coupling is modeled and analyzed, showing that in some configurations, at least a third-order nonlinear representation is necessary to achieve accurate results.

In contrast, simpler models, such as the linear formulation presented by Erturk and Inman [68], can still provide accurate predictions, depending on the operational conditions and system configuration. Ultimately, the level of model complexity is a design choice that should balance analytical tractability with the need for generalization. The more a model seeks to generalize across operating conditions and configurations, the more it must incorporate nonlinear and distributed effects. For the purpose of this work, however, a geometrically linear model is deemed sufficient, as the physical parameters are calibrated experimentally, and the only source of nonlinearity is considered to arise from the magnetic force.

## 2.2 Euler-Bernoulli beam electromechanical model

Based on the Euler-Bernoulli formulation, the equation of motion of the PVEH represented in Fig. 1 is [69]:

$$m \frac{\partial^2 w_{rel}(x, t)}{\partial t^2} + c_a \frac{\partial w_{rel}(x, t)}{\partial t} + \frac{\partial^2 M(x, t)}{\partial x^2} = -m \frac{\partial^2 w_b(x, t)}{\partial t^2} - F_{mag} \delta(x - L), \quad (1)$$

where  $x$  is the spacial coordinate,  $t$  is the time coordinate,  $w_{rel}(x, t)$  is the transverse deflection of the beam relative to its base,  $w_b(x, t)$  is the excitation,  $M(x, t)$  is the internal bending moment,  $I$  is the equivalent area moment of inertia of the composite cross section,  $c_a$  is the viscous damping coefficient,  $F_{mag}$  is the magnetic force,  $m$  is the mass per unit length of the beam and  $\delta$  is the Dirac delta function. The value of  $m$  is a composition of the mass of the beam and the tip mass:

$$m = M_t \delta(x - L) + m_s, \quad (2)$$

where  $M_t$  is the tip mass and  $m_s$  is the mass per unit length of the substract.

The magnetic force,  $F_{mag}$ , between two-point magnets can be approximated by [70]:

$$F_{mag} = \frac{\alpha_{mag}}{d^2}, \quad (3)$$

where  $\alpha_{mag}$  is a constant related to the magnetic properties of the magnets and the surrounding medium, and  $d$  is the distance separating the magnets. This distance  $d$  is directly linked to the initial distance  $d_0$  between the magnets and the vertical displacement of the beam  $w$ , represented through the geometry of a right triangle, as depicted in Fig. 1 a). This relationship can be expressed as:

$$d = \sqrt{d_0^2 + w_{rel}^2}. \quad (4)$$

The magnetic force can be resolved into horizontal and vertical components, focusing on the vertical component in this model, since its the one parallel to the movement direction. Using the previously established relationships in Eqs. (3) and (4) along with the triangular geometry, the vertical component of the magnetic force is given by:

$$F_{mag,y} = \frac{\alpha_{mag} w_{rel}(x, t)}{(d_0^2 + w_{rel}^2(x, t))^{1.5}}. \quad (5)$$

This equation can be approximated by a third-order Taylor expansion:

$$F_{mag,y} = \frac{\alpha_{mag} w_{rel}(x, t)}{d_0^3} - \frac{1.5 \alpha_{mag} w_{rel}^3(x, t)}{d_0^5}. \quad (6)$$

With respect to the internal bending moment  $M(x, t)$ , it can be written as:

$$M(x, t) = EI \frac{\partial^2 w_{rel}(x, t)}{\partial x^2} + f_p(x) \theta_p V(t), \quad (7)$$

where  $\theta_p$  is the electromechanical coupling,  $f_p(x)$  the function defining the piezoelectric patch position,<sup>1</sup>  $V(t)$  is the electrical voltage and  $EI$  is the bending stiffness of the composite cross section, which is given by:

$$EI = b \left[ E_s (h_s^3 - h_a^3) + E_p (h_s^3 - h_c^3) \right] \frac{1}{3}, \quad (8)$$

where  $b$  is the width of the beam,  $h_a$  is the position of the bottom of the substructure layer from the neutral axis,  $h_b$  is the position of the bottom of the PZT layer (therefore, top of the substructure layer) from the neutral axis, and  $h_c$  is the position of the top of the PZT layer from the neutral axis, as represented in Fig. 1 b). Considering the experimental nature of the present work, the electromechanical coupling is simplified as follows:

$$f_p(x) \theta_p \rightarrow \theta_p, \quad (9)$$

thereby absorbing the effects of the positioning function and obtaining  $\theta_p$  directly from the experimental data. Therefore, substituting Eq. 7 and Eq. 6 into Eq. 1 the following equation is obtained:

$$m \frac{\partial^2 w_{rel}(x, t)}{\partial t^2} + c_a \frac{\partial w_{rel}(x, t)}{\partial t} + EI \frac{\partial^4 w_{rel}(x, t)}{\partial x^4} + V(t) \theta_p + \left( \frac{\alpha_{mag}}{d_0^3} w_{rel}(x, t) - \frac{1.5 \alpha_{mag}}{d_0^5} w_{rel}^3(x, t) \right) \delta(x - L) = -m \frac{\partial^2 w_b(x, t)}{\partial t^2}. \quad (10)$$

Regarding the electrical circuit equation, it can be derived from the voltage across the resistive load. A complete derivation can be found in [68], and the result is as follows:

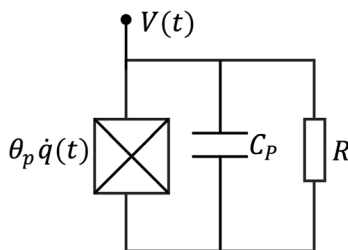


Fig. 2 Diagram of the electrical circuit

<sup>1</sup> It is important to note that the function  $f_p(x)$  is formulated according to the geometric design. It may be defined through a Heaviside function, or simply set to unity when the piezoelectric patch is bonded along the entire beam [71–73].

$$C_p \frac{dV(t)}{dt} + \frac{V(t)}{R} = -\theta_p \int_0^L \frac{\partial^3 w_{rel}(x, t)}{\partial x^2 \partial t} dx, \quad (11)$$

where  $C_p$  is the internal capacitance of the PZT layer and  $R$  is the load resistance. The circuit is represented in Fig. 2.

Since this work focuses on the first bending mode and the PVEH is considered weakly nonlinear, as consequence without coupling to higher modes, the next subsection introduces the reduced-order representation of the model. One should note that the central idea of the present article relies on a pre-existing model whose simplifying assumptions lose accuracy outside the limits for which it was originally derived. For instance the expression for the magnetic force, Eq. (3), follows the assumptions adopted in B.P. Mann and B.A. Owens (2010) [74], Bernard, Brian P. et al. (2013) [75], and Sharvari Dhote et al. (2018) [76], where the authors analyzed mechanisms in which the magnets oscillate along a single direction, thereby allowing the hypothesis of collinearity/parallelism of the magnetic field lines. In our specific case (Fig. 1), such assumption becomes inaccurate as the vibration amplitude increases, introducing additional nonlinear effects into the PVEH. A similar argument applies to geometric nonlinearity: the hypothesis of large displacements leads to nonlinear contributions not captured by the original model, as demonstrated by Farokhi et al. (2022) [77]. Possible extensions of the proposed framework to problems involving other forms of nonlinearity (such as bistable structures, Norenberg et al. (2023) [78]) must be built upon modeling approaches that more closely reflect the actual operating conditions.

## 2.3 Modal expansion

The modal superposition method is a solution often proposed to calculate the system's response directly from Eqs. (10) and (11). This approach assumes a modal basis for the displacement  $w$ , enabling a reduced-order solution to the problem. By applying this method, the response can be expressed as follows:

$$w(x, t) = \sum_{i=1}^{\infty} \phi_i(x) q_i(t), \quad (12)$$

in which  $\phi_i(s)$  and  $q_i(t)$  are the  $i^{\text{th}}$  modal shapes and modal coordinates, respectively, with the following boundary conditions for the studied clamped-free beam with a lumped mass at the tip [79]:

$$\begin{cases} \phi(0) = 0, \\ \phi'(0) = 0, \\ EI\phi''(L) = 0, \\ EI\phi'''(L) + \omega^2 M_t \phi(L) = 0, \end{cases} \quad (13)$$

The modal shapes can be expressed as [80]:

$$\phi_i(x) = A_i \left[ \cos \frac{\lambda_i}{L} x - \cosh \frac{\lambda_i}{L} x + \tau_i \left( \sin \frac{\lambda_i}{L} x - \sinh \frac{\lambda_i}{L} x \right) \right], \quad (14)$$

where  $\tau_i$  is given by:

$$\tau_i = \frac{M_t L (\cos \lambda_i - \cosh \lambda_i) \lambda_i + m L^2 (\sin \lambda_i - \sinh \lambda_i)}{-M_t L (\sin \lambda_i - \sinh \lambda_i) \lambda_i + m L^2 (\cos \lambda_i + \cosh \lambda_i)}, \quad (15)$$

and  $A_i$  is determined using the orthogonality conditions [81]:

$$\int_0^L \phi_q(x) m \phi_r(x) ds + \phi_q(L) M_t \phi_r(L) = \delta_{qr}, \quad (16)$$

where  $\delta_{qr}$  is the Kronecker delta function. Additionally,  $\omega_i$  is the  $i^{\text{th}}$  undamped natural angular frequency, written as:

$$\omega_i = \lambda_i^2 \sqrt{\frac{EI}{mL^4}}, \quad (17)$$

where  $\lambda_i$  is the solution of the following eigenvalue problem:

$$\frac{M_t}{mL} (\sinh \lambda_i \cos \lambda_i - \sin \lambda_i \cosh \lambda_i) \lambda_i + (1 + \cos \lambda_i \cosh \lambda_i) = 0. \quad (18)$$

Finally, considering the expansion around the first-mode only<sup>2</sup>, the reduced-order equations of motion are obtained as follows:

$$\ddot{q}(t) + 2\zeta\omega_0\dot{q}(t) + \omega_0^2 q(t) + k_{nl}q^3(t) + \theta_p V(t) = F_m A_0 \cos(\omega t), \quad (19)$$

$$C_p \dot{V}(t) + \frac{V(t)}{R} + \theta_p \dot{q} = 0, \quad (20)$$

in which:

- $\zeta$  is the mechanical damping ratio;
- $\omega_0^2 = \omega_1^2 + \frac{\alpha_{\text{mag}}}{d_0^3} \phi^2(L)$  is the frequency associated with the first bending mode of the beam, considering the magnetic effects;
- $\omega_1$  is the frequency associated with the first bending mode of the beam in the absence of magnetic effects;

- $k_{nl} = -\frac{1.5\alpha_{\text{mag}}}{d_0^5} \phi^4(L)$  is the nonlinear stiffness;
- $\int_0^L \phi(x) \theta_p dx \rightarrow \theta_p$  (for the sake of notation);
- $F_m = -\left( \int_0^L m_s \phi dx + M_t \phi(L) \right)$ .

Note that the variable  $q$  in Eqs. (19) and (20) represents the modal coordinate, which is not directly measurable. The measurable physical quantity is the displacement  $w$ , related to  $q$  through Eq. (12). As the function  $\phi(x)$  varies along the beam axis, measurements should be consistently taken at the same point  $p_0$  to ensure that  $\phi$  remains constant, allowing the derived constants to be meaningful. Consequently, the values of the modal force and nonlinear damping identified from the equations in terms of  $w$  rather than  $q$  will depend on the specific measurement point. To indicate this dependence, the parameters associated with the measurement point  $p_0$  are denoted with the superscript  $p_0$ .

## 2.4 PVEH frequency-domain description

For algebraic convenience, Eqs. (19) and (20) are normalized by introducing the following non-dimensional parameters [63]:

$$\begin{aligned} y &= \frac{q}{q_0}, & q_0 &= \frac{F_m A_0}{\omega_0^2}, & \Omega &= \frac{\omega}{\omega_0}, & \tau &= \omega_0 t, \\ \alpha &= \frac{k_{nl} q_0^2}{\omega_0^2}, & v &= \frac{C_p V}{\theta q_0}, & \beta &= \frac{1}{RC_p \omega_0}, & \text{and } \kappa^2 &= \frac{\theta_p^2}{\omega_0^2 C_p}, \end{aligned} \quad (21)$$

resulting in the non-dimensional form of the dynamic equations:

$$y'' + 2\zeta y' + y + \alpha y^3 + \kappa^2 v = \cos(\Omega\tau), \quad (22)$$

$$v' + \beta v = y'. \quad (23)$$

The frequency response curves are obtained by the harmonic balance method [84], considering only first-order harmonics. Assuming that the solutions of Eqs. (22) and (23) have the form:

$$y = Y \cos(\Omega\tau + \varphi_1), \quad (24)$$

$$v = V_{ap} \cos(\Omega\tau + \varphi_2), \quad (25)$$

in which  $Y$  and  $V_{ap}$  are the harmonic amplitudes of displacement and voltage, respectively, and  $\varphi_1$  and  $\varphi_2$  are the phase differences. The frequency-amplitude relation of the PVEH is obtained by substituting Eqs. (24) and (25) into Eqs. (22) and (23), neglecting the higher order harmonics and equating the coefficients of  $\cos(\Omega\tau)$  and  $\sin(\Omega\tau)$ , yielding to:

<sup>2</sup> The algebraic procedure associated with the modal expansion follows the classical treatment found in standard references on continuous system modeling [82, 83]. In the context of this work, the procedure unfolds as follows: the substitution of Eq. (12) into Eqs. (10) and (11), followed by multiplication of the resulting expression by the mode of interest  $\phi_1(x)$  and subsequent integration over the spatial domain, while exploiting the orthogonality properties of the mode shapes.

$$\frac{9\alpha^2}{16}Y^6 + \frac{3\alpha}{2}\left(1 - \Omega^2 + \frac{\kappa^2\Omega^2}{\Omega^2 + \beta^2}\right)Y^4 + \left[1 + \Omega^4 + (4\zeta^2 - 2)\Omega^2 + \frac{(2\kappa^2(1 - \Omega^2 + 2\zeta\beta) + \kappa^4)\Omega^2}{\Omega^2 + \beta^2}\right]Y^2 - 1 = 0, \quad (26)$$

$$V_{ap}^2 = \frac{\Omega^2 Y^2}{\Omega^2 + \beta^2}. \quad (27)$$

These nonlinear frequency response curves are typically solved numerically, for example using the Newton–Raphson method or other standard numerical approaches. Given the physical interpretation of Equation (26), only positive real solutions should be considered. In fact, Equations (26)–(27) define the baseline model considered hereinafter, whose parameter identification are described in Sect. 4. The next section presents the experimental setup of the study.

### 3 Experimental setup

The design considerations and experimental procedure are outlined to provide a comprehensive understanding of the approach used to obtain the PVEH's response data. The setup is designed to investigate the behavior of the PVEH under varying conditions of resistance and base acceleration, with a focus on the first bending mode of the beam, which is the most important for maximizing displacement and electrical voltage output. The experimental setup and physical mechanism are described in 3.1, followed by the dataset definition in 3.2.

#### 3.1 Design concept and experimental setup

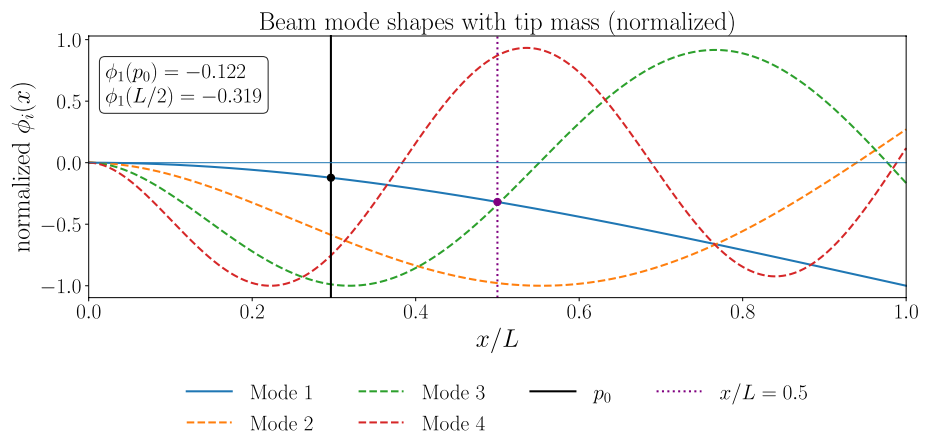
The beam is built from a magnetic material (steel) to enhance coupling with the tip magnet. To minimize unnecessary weight in the shaker base, the main structure, that hosts the beam, is made from aluminum. Additionally, the nature of the interaction between the magnets, whether attraction or repulsion, requires rigorous analysis. As indicated in Eq. (6), repulsion between magnets ( $\alpha_{\text{mag}} > 0$ )

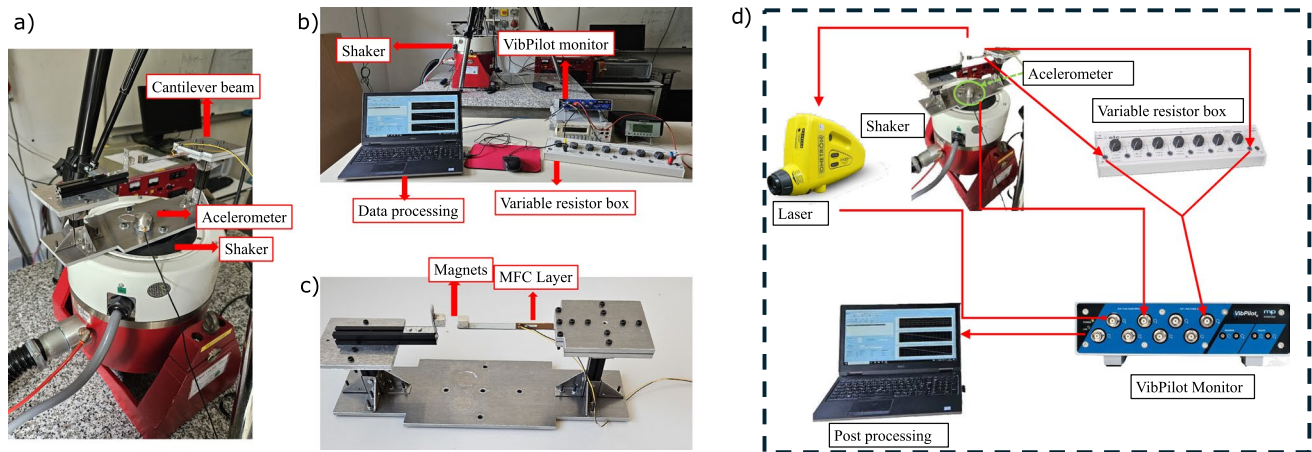
introduces a hardening effect, while attraction ( $\alpha_{\text{mag}} < 0$ ) results in a softening effect. In this study, the magnets are arranged in an attractive configuration to suppress bistability effects, and the base acceleration is limited to a low level to avoid inducing jump phenomena [85, 86].

Figure 3 provides a graphical perspective of the linear modal base, using the geometric and material parameters of the experimental beam. The experimental setup used for data acquisition is illustrated in Fig. 4. The cantilever beam, made of steel 1.0330 ( $135 \times 10 \times 0.8$  mm, tip mass value of  $M_t = 7.6$  g and  $d_0 = 13$  mm), is mounted on an aluminum base and excited by a Data Physics GW-V100 shaker (Fig. 4a). A DJB A/600 accelerometer is fixed to the beam to provide feedback for closed-loop control of the imposed base acceleration. The beam carries a microfiber composite (MFC) piezoelectric patch of type M-2807-P2 bonded near the clamped end, and a cube magnet (Supermagnete W-07-N) at the free end; a second magnet is positioned nearby to enable magnetic interaction (Fig. 4c).

The selection of the measurement point ( $p_0$ ) is also important. A location too close to the clamped end may exhibit low displacement amplitude, which can be affected by measurement noise. Conversely, a position too close to the free end may experience higher amplitude displacements, potentially causing perpendicularity issues due to increased horizontal motion. Therefore, it is essential to choose a constant measurement point that prioritizes vertical-axis displacement while minimizing these technical issues. The choice of an appropriate measurement point can be guided by analyzing  $\phi_i(x)$  around the mode of interest (Eq. (14), for  $i = 1$ , see Fig. 3). One should note that at  $x = L/2$ , the first mode reaches approximately 32% of its peak amplitude, which already mitigates the perpendicularity issue when compared with measurements taken at  $x = L$ . For experimental robustness, the PVEH output is therefore measured near the junction between the beam and the piezoelectric patch, at  $p_0 = 40$  mm, where the modal amplitude is approximately 13% in the vicinity of the first mode.

**Fig. 3** Normalized mode shapes  $\phi_i(x)$  of the cantilever beam with tip mass, computed using the geometric and material parameters of the experimental setup. The first bending mode ( $i = 1$ ) is highlighted, and the locations  $x = L/2$  and the measurement point  $p_0 = 40$  mm are indicated, illustrating the relative modal amplitudes used to guide the selection of the experimental measurement position





**Fig. 4** Experimental setup. **a** Cantilever beam mounted on the shaker, equipped with an accelerometer for vibration measurement. **b** Shaker generating controlled vibrations with the structure mounted on top,

Instrumentation (Fig. 4b) includes a VibPilot data acquisition system, a variable resistance box for adjusting the electrical load ( $1\ \Omega$ – $10\ \text{M}\Omega$ ), and a control computer for signal processing. A laser vibrometer (OMETRON VH300+) measures the beam's velocity. The schematic in Fig. 4d) shows the system connections: the shaker excites the base, accelerometer feedback maintains constant acceleration, and signals from the vibrometer and piezoelectric patch are measured through the VibPilot for acquisition. Data are then transferred to the control computer for post-processing and analysis.

### 3.2 Dataset description

A total of 21 experiments are conducted to investigate the PVEH's response under varying resistance and acceleration conditions. The tests focused on the frequency range from 40 to 44 Hz, targeting the first bending mode, providing greater electrical voltage output. A frequency sweep sine signal was applied at a sweep rate of 0.7 Hz/min, and the experiments are performed at three different acceleration levels: 0.02g, 0.04g, and 0.10g.

The dataset distribution is shown in Fig. 5, illustrating the different combinations of resistance and acceleration used to obtain the experimental curves. For an excitation amplitude of 0.10g, the maximum tested resistance is lower than for 0.02g and 0.04g. This is due to the 10 V voltage limit imposed by the acquisition system: as acceleration increases, the resulting displacement, and therefore the generated electrical voltage, also increases. At higher excitation levels, the voltage threshold is reached at lower resistance values, reducing the range of resistances that can be tested. The upper acceleration amplitude of 0.10g was chosen because it is close to the onset of jump phenomena in the frequency response curves, making it the maximum value

monitored via laptop with data acquisition software and signal conditioning. **c** Beam integrated with magnets and the MFC. **d** Connections between the main elements of the experimental setup

considered in this study. The other excitation amplitudes represent low (0.02g) and medium (0.04g) excitation levels, associated with the activation of nonlinear effects, whereas at lower acceleration levels the PVEH is expected to behave linearly.

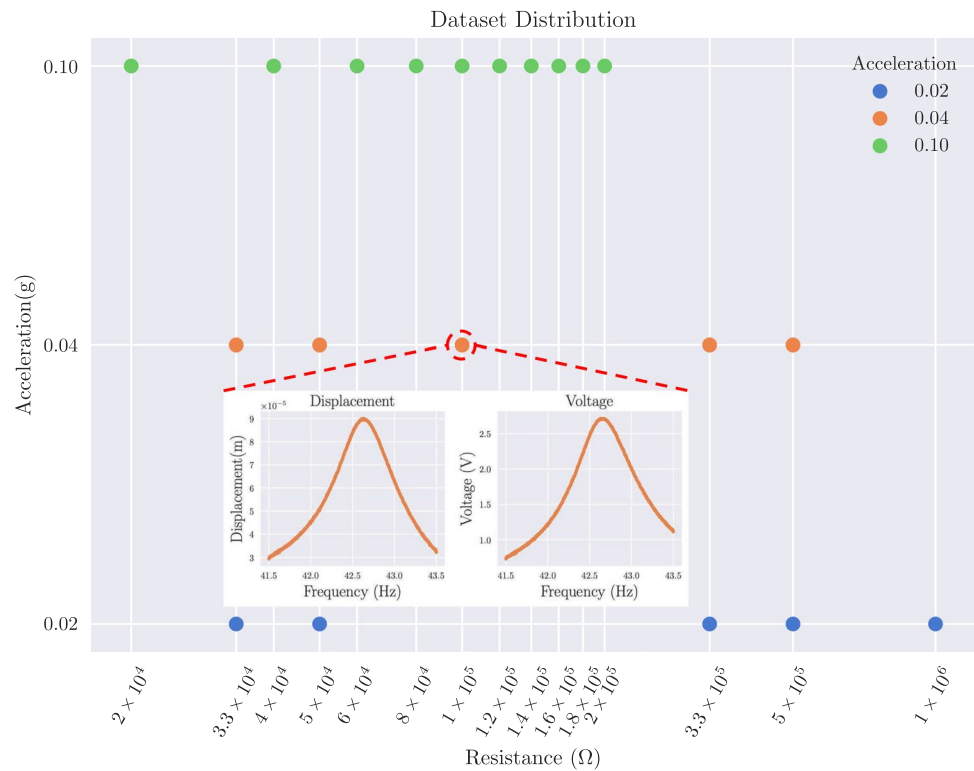
To illustrate the PVEH's nonlinear behavior for a given resistive load, Fig. 6 compares the frequency response curves for the three studied acceleration levels, all obtained under a fixed resistive load of  $10^5\ \Omega$ . As the excitation level increases, a shift in the resonance frequency is observed. Specifically, the peak frequency at 0.10g occurs at a lower value compared to those at 0.02g and 0.04g. This downward shift is indicative of a softening behavior, typically associated with nonlinear dynamic systems. The result suggests that nonlinear effects become increasingly significant as the input energy grows, providing experimental evidence of the presence and activation of nonlinearities in the studied PVEH, even if jump phenomena is not present.

The experimental setup and the resulting dataset play a key role in the subsequent analysis. They provide the necessary basis for calibrating the physical parameters of the baseline model, assessing its predictive accuracy, and serving as data for the physics-augmented neural network introduced in later sections.

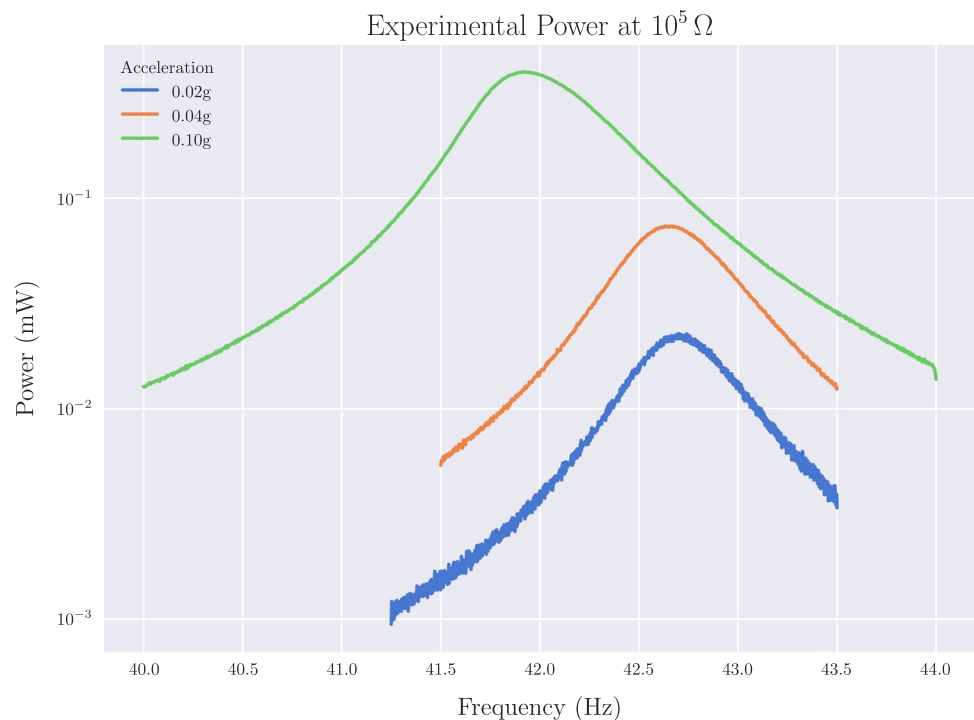
## 4 Analytical model parameter identification

Based on the experimental data obtained by the setup described in the previous section, the calibration of the physical parameters is carried out using the Particle Swarm Optimization (PSO) algorithm [87], a heuristic optimization technique. The use of PSO as the identification method for the experimental parameters is motivated by its straightforward implementation and tunability. This approach has

**Fig. 5** Dataset distribution showing the different combinations of resistance and acceleration selected for the experimental tests. For each point, frequency response curves for both displacement and voltage were measured. An example is highlighted for 0.04g and  $1 \times 10^5 \Omega$ . The acceleration range was chosen to avoid jumps and bistability effects, while the resistance range was selected to optimize electric power generation in piezoelectric devices by using high resistances, but without reaching values that could cause voltage saturation



**Fig. 6** Experimental frequency response of the power output for a resistive load of  $10^5 \Omega$ , showing the effect of increasing acceleration and the resulting nonlinear softening behavior



been adopted in similar works involving piezoelectric systems [88, 89], with particular emphasis on Alcântara et al. (2025) [23], where the physical parameters of a nonlinear PVEH were identified using PSO. The algorithm identifies the optimal set of parameters ( $\Theta = \zeta, \omega_0, F_m^{p0}, k_{nl}^{p0}, \kappa$ ) that minimizes the discrepancy between the analytical frequency

response curves of displacement and voltage (please, see Eqs. (26)–(27)), derived from the baseline model, and the experimental data. The value of  $C_p$ , which is required for the analytical model, is directly measured using a multimeter and is found to be 22nF.

The optimization process is conducted using the medium and high excitation level data, as the nonlinear behavior is present in PVEH output. The cost function is defined as the mean absolute error between the experimental data and the baseline model prediction, for both voltage and displacement frequency responses, considering all resistances in the selected acceleration. The optimization process uses 50 particles and runs for 200 iterations. The resulting values of the parameters are presented in Table 1. Considering that  $N$  is the number of experimental curves used, the optimization process is formulated as follows:

$$\Theta^* = \arg \min_{\Theta} J(\Theta). \quad (28)$$

3 where

$$\Theta = \zeta, \omega_0, F_m^{p_0}, k_{nl}^{p_0}, \kappa, \quad (29)$$

$$J(\Theta) = \sum_{n=1}^N \left( 3 \frac{|\hat{V}(\Omega, \Theta) - V_n(\Omega)|^2}{\sigma_{V_n}^2} + \frac{|\hat{Y}(\Omega, \Theta) - Y_n(\Omega)|^2}{\sigma_{Y_n}^2} \right),$$

subject to:

$$\begin{aligned} 40 \text{ Hz} &\leq \Omega \leq 44 \text{ Hz}, \\ 35 &\leq \omega_0 \leq 50, \\ 10^{-4} &\leq F_m^{p_0} \leq 10^{-1}, \\ 10^{-4} &\leq \zeta \leq 10^{-2}, \\ -10^9 &\leq k_{nl}^{p_0} \leq -10^7, \\ 10^{-3} &\leq \kappa \leq 1. \end{aligned}$$

where  $\sigma^2$  is the data variance,  $\hat{V}$  and  $\hat{Y}$  represent the predicted amplitudes of voltage and displacement, respectively, obtained from the baseline model, whereas  $V_n$  and  $Y_n$  denote the experimental amplitude values of voltage and displacement for the  $n^{\text{th}}$  curve. The constraints are determined through an heuristic approach, based on identifying the order of magnitude of the parameters that yield a physically consistent response when compared to the experimental data. A weight of 3 is applied to the voltage error to prioritize higher accuracy in the electrical voltage domain, which is more relevant for energy harvesting applications. Details regarding the convergence of the optimization process are presented in Appendix A.

Having calibrated the parameters of the baseline model through the procedure previously described, the PVEH is evaluated against experimental results for acceleration

levels not included in the calibration process. The comparison, in terms of the electrical voltage response amplitude for selected resistance–acceleration combinations, is shown in Fig. 7.

The results indicate that the model calibrated reproduces the experimental response with good accuracy at both low (0.02g) and medium (0.04g) excitation levels, capturing the PVEH behavior in these regimes. In contrast, at higher excitation amplitudes, noticeable deviations emerge, with the model underestimating or misrepresenting the response. This discrepancy arises from the increasing influence of nonlinear effects not captured by the calibrated model, which includes only a third-order stiffness nonlinearity. Consequently, the model does not generalize well to higher acceleration levels but remains effective for predicting the PVEH's response under lower excitation levels.

## 5 Physics-augmented neural networks for the PVEH response prediction

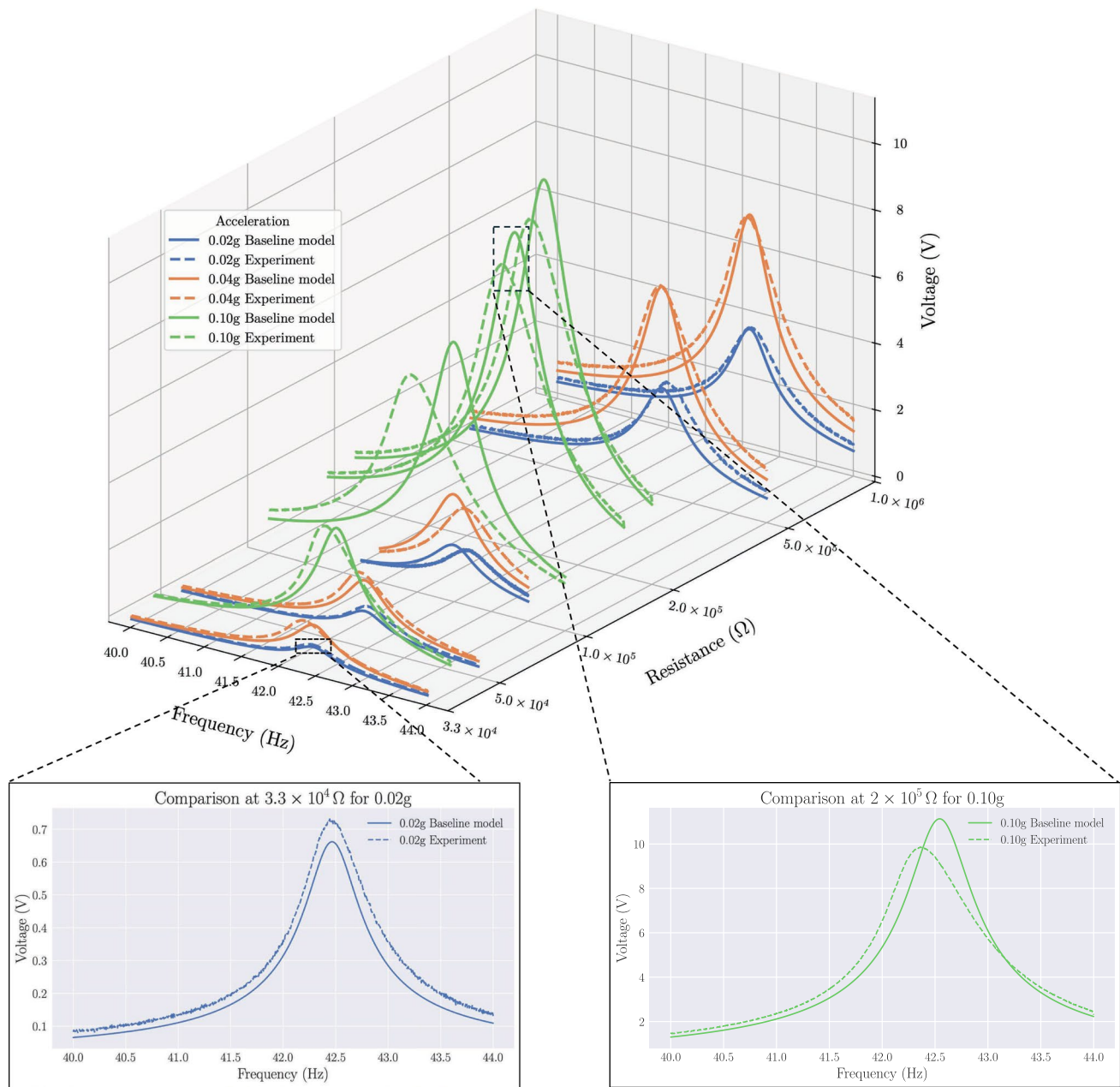
Aiming to improve the generalizability of the baseline model, particularly at higher acceleration levels, a physics-augmented neural network approach is employed, seeking to extend the model's range of applicability. Unlike purely data-driven black-box models, this type of neural network incorporates physical knowledge into its optimization procedure by means of adaptations performed on the loss function, providing greater interpretability and physical consistency in the results. In this work, such integration is achieved by using the amplitude of the baseline voltage frequency response curve as input, along with a custom loss function and the activation function in the network architecture. The output of the network, referred to as the augmented model, represents an approximation of the experimental results, with varying resistances. Since the baseline model already provides satisfactory accuracy at low and medium excitation levels, the benefits of the augmented approach become more relevant at higher accelerations, where larger output amplitudes, and consequently greater power generation for harvesting, are achieved. For this reason, only the higher-amplitude data are used for training the network. In 5.1, the formulation of the loss function is described, whereas 5.2 presents the network architecture and the training loop. In 5.3, the performance of the augmented model is evaluated, highlighting the advantages of the proposed strategy over the analytically derived model.

### 5.1 On the composition of the loss function

A neural network is a mathematical model composed of a layered structure of nonlinear activation functions. In this

**Table 1** Identified values of the physical parameters

| Symbol         | Quantity                                | Value                                |
|----------------|-----------------------------------------|--------------------------------------|
| $\zeta$        | Damping ratio                           | $5.39 \times 10^{-3}$                |
| $\omega_0$     | Natural Frequency                       | 42.46 Hz                             |
| $F_m^{p_0}$    | Modal Force                             | $9.94 \times 10^{-2}$ N              |
| $k_{nl}^{p_0}$ | Nonlinear stiffness                     | $-5.94 \times 10^8$ N/m <sup>3</sup> |
| $\kappa$       | Adimensional Electromechanical coupling | $8.33 \times 10^{-2}$                |



**Fig. 7** Comparison between the baseline model and the experimental results for the electrical voltage amplitude in frequency response curves, considering different combinations of acceleration and load

resistance. Two cases are highlighted:  $3.3 \times 10^4 \Omega$  at 0.02g and  $2 \times 10^5 \Omega$  at 0.10g. The results show that increasing the excitation acceleration reduces the predictive accuracy of the baseline model

work, it is adopted a dense fully-connected architecture (also referred as multi-layer perceptron), which is commonly used and presents a flexible framework. Such configuration is made up of  $L$  hidden layers, each containing  $N_\ell$  neurons, where every neuron is connected to all neurons in the subsequent layer through a nonlinear activation function  $\sigma(\cdot)$ . Let  $\mathbf{x} \in \mathbb{R}^D$  denote the input and  $y \in \mathbb{R}$  the target output. The neural network aims to approximate the target function  $y(\mathbf{x})$  by a function  $\hat{y}(\mathbf{x})$ , which can be expressed as:

$$y(\mathbf{x}) \approx \hat{y}(\mathbf{x}) = \mathcal{H}^{(L)} \circ \mathcal{H}^{(L-1)} \circ \dots \circ \mathcal{H}^{(1)} \circ \mathcal{H}^{(0)}(\mathbf{x}), \quad \text{for } \mathbf{x} \in \mathbb{R}^D, \quad (30)$$

where the transformations  $\mathcal{H}$  in each layer are defined as:

$$\begin{aligned} \mathbf{x}_0 &:= \mathcal{H}^{(0)}(\mathbf{x}), \\ \mathbf{x}_\ell &:= \mathcal{H}^{(\ell)}(\mathbf{x}_{\ell-1}) = \sigma(\mathbf{W}_\ell \mathbf{x}_{\ell-1} + \mathbf{b}_\ell), \quad 0 < \ell < L, \\ \hat{y}(\mathbf{x}; \mathbf{W}, \mathbf{b}) &= \mathcal{H}^{(L)}(\mathbf{x}_{L-1}) := \mathbf{W}_L \mathbf{x}_{L-1} + \mathbf{b}_L. \end{aligned} \quad (31)$$

In these expressions,  $W_\ell \in \mathbb{R}^{N_\ell \times N_{\ell-1}}$  and  $b_\ell \in \mathbb{R}^{N_\ell}$  denote the weights and biases of the  $\ell$ -th layer, respectively, and  $\sigma$  is the element-wise activation function. The neural network output  $\hat{y}(x) = \hat{y}(x; W, b)$  is thus fully determined by the set of weights and biases, hereafter denoted by  $\theta := \{W, b\}$ .

In purely data-driven models, a widely adopted choice is the mean absolute error (MAE) between the predicted and the observed data, here denoted as  $\mathcal{R}_{\text{MAE}}$ , and defined by:

$$\mathcal{R}_{\text{MAE}}(\hat{y}(\cdot; \theta)) = \frac{1}{N} \sum_{n=1}^N |\hat{y}(x_n; \theta) - y(x_n)|, \quad (32)$$

where  $N$  denotes the number of observed data samples.

The -3 dB bandwidth is commonly used as a metric to define the operational range in energy harvesting systems, as it represents the frequency interval over which the output power remains close to its maximum value [90]. To incorporate this physical insight into the learning process, a loss term focused on the -3 dB region is introduced. This region is defined based on the experimental amplitude of the voltage frequency response curve, where the amplitude remains above  $y_{\text{peak}}/\sqrt{2}$ , with  $y_{\text{peak}}$  being the peak value of the experimental response. The corresponding loss term, denoted as  $\mathcal{R}_{\text{MAE-3dB}}$ , is given by:

$$\mathcal{R}_{\text{MAE-3dB}}(\hat{y}(\cdot; \theta)) = \frac{1}{N_{\text{3dB}}} \sum_{n=n_{\text{low}}}^{n_{\text{up}}} |\hat{y}(x_n; \theta) - y(x_n)|, \quad (33)$$

where the indices that delimit the -3 dB bandwidth,  $n_{\text{low}}$  and  $n_{\text{up}}$ , are defined as:

$$n_{\text{peak}} = \arg \max_n y(x_n), \quad (34)$$

$$n_{\text{low}} = \max \left\{ n < n_{\text{peak}} \mid y(x_n) < \frac{y(x_{n_{\text{peak}}})}{\sqrt{2}} \right\}, \quad (35)$$

$$n_{\text{up}} = \min \left\{ n > n_{\text{peak}} \mid y(x_n) < \frac{y(x_{n_{\text{peak}}})}{\sqrt{2}} \right\}, \quad (36)$$

and  $N_{\text{3dB}} = n_{\text{up}} - n_{\text{low}} + 1$  corresponds to the number of samples within the 3 dB region.

Another quantity of practical relevance in energy harvesting systems is the maximum voltage amplitude, as it defines the best operational point in terms of power extraction. To ensure the neural network accurately captures this peak behavior, a loss term denoted as  $\mathcal{R}_{\text{max-amp}}$  is introduced:

$$\mathcal{R}_{\text{max-amp}}(\hat{y}(\cdot; \theta)) = \left| \max_n \hat{y}(x_n; \theta) - \max_n y(x_n) \right|. \quad (37)$$

Although the information captured by  $\mathcal{R}_{\text{MAE-3dB}}$  and  $\mathcal{R}_{\text{max-amp}}$  is implicitly reflected in the global MAE metric, these additional terms provide localized supervision on key physical features of the PVEH. Their explicit inclusion in the loss function serves as a form of regularization and supports a multi-objective training strategy, encouraging the model to accurately capture both overall trends and critical performance regions.

Lastly, to avoid nonphysical oscillations and to promote smoothness in the predicted curve, a regularization term is included to penalize large variations between adjacent frequency points. This term enhances local continuity and reflects the smooth behavior observed in the experimental data. The smoothness loss is defined as:

$$\mathcal{R}_{\text{smooth}}(\hat{y}(\cdot; \theta)) = \frac{1}{N-1} \sum_{n=1}^{N-1} |\hat{y}(x_{n+1}; \theta) - \hat{y}(x_n; \theta)|, \quad (38)$$

where  $N$  is the number of frequency samples. Since this term depends solely on the predicted response  $\hat{y}$ , it acts as a prior to encourage physically plausible solutions. Given this terms, the total loss function ( $\mathcal{L}$ ) is defined as:

$$\mathcal{L} = \gamma_1 \mathcal{R}_{\text{MAE}} + \gamma_2 \mathcal{R}_{\text{MAE-3dB}} + \gamma_3 \mathcal{R}_{\text{max-amp}} + \gamma_4 \mathcal{R}_{\text{smooth}}, \quad (39)$$

where  $\gamma_n$  are weights used to calibrate the order of magnitude of the different loss terms. To the best of the authors' knowledge, this work is the first to propose such a loss function structure for addressing energy harvesting applications based on vibration signals.

## 5.2 On the network's architecture and the training loop

Two distinct neural network architectures are considered for comparison. For sake of simplicity, the models are referenced herein as #1 and #2. In the model #1 architecture,  $L = 4$  layers are considered, having each of the four fully connected layers 100 neurons. Number of layers and neurons are selected without any prior optimization criterion. Conversely, in the model #2 architecture, the number of neurons per layer and layers are determined through an optimization process using the Optuna library [91] available for Python, with the search space ranging from 10 to 100 neurons per layer, and the number of layers ranging from 2 to 6. In this model selection step, cross-validation (CV) using 4 folds is employed to minimize overparameterization risks. To train the network, the dataset composed of 10

curves is randomly divided into training, validation, and test sets, following a 70:15:15 split. The later is never used during the learning steps, and the validation set is used just to evaluate model overfitting. At the end of the architecture optimization step considering 25 trials, the best performing architecture for model #2 was found to have 6 layers, and 67 neurons per layer. Both models use the hyperbolic tangent as activation function, whereas the final layers employ the softplus activation function to ensure non-negative outputs, consistent with the physical meaning of the amplitude–frequency response. A fixed random seed is used to ensure the reproducibility of the results.

The training set is used to optimize the weights and biases of the network by minimizing the loss function  $\mathcal{L}$  described previously (please, see Eq. (39)). At each epoch, the loss on the validation set is monitored to prevent underfitting and overfitting. After training, the performance of the network is evaluated on the test set to assess its ability to generalize on unseen data. The optimization process is carried out using the Adam optimizer [92], which combines adaptive learning rates with momentum. To further improve convergence, the learning rate is dynamically adjusted during training using the OneCycle policy [93], following its standard configuration, with a maximum learning rate of 0.001, reached after 30% of the total training iterations.

To enhance robustness and reduce the risk of overfitting, data augmentation is applied to the inputs. For each original sample  $(x_n, y_n)$ , a set of  $M$  augmented inputs is created by adding zero-mean Gaussian noise:

$$\tilde{x}_n^{(m)} = x_n + \varepsilon_n^{(m)}, \quad \text{with } \varepsilon_n^{(m)} \sim \mathcal{N}(0, \sigma_m^2 \mathbf{I}), \quad m = 1, \dots, M, \quad (40)$$

where  $\sigma_m \in \{0.01, 0.02, 0.03, 0.04, 0.05\}$  are the noise levels applied to each augmentation. The augmented training set is then given by:

$$\mathcal{D}_{\text{aug}} = \left\{ \left( \tilde{x}_n^{(m)}, y_n \right) \mid n = 1, \dots, N; m = 1, \dots, M \right\}. \quad (41)$$

This strategy has been shown to be effective in mitigating overfitting [94], as it increases the diversity of training samples and encourages the network to learn features that are invariant to small input perturbations.

Additionally, early stopping is employed as a regularization strategy [95]. Training is terminated when the validation loss exceeds twice the training loss for 50 consecutive epochs. This criterion prevents the model from over-adapting to the training data and helps preserve its generalization capability.

The weights  $\gamma_n$  in Eq. 39 are chosen to ensure that all loss terms have the same order of magnitude, and are set as follows:

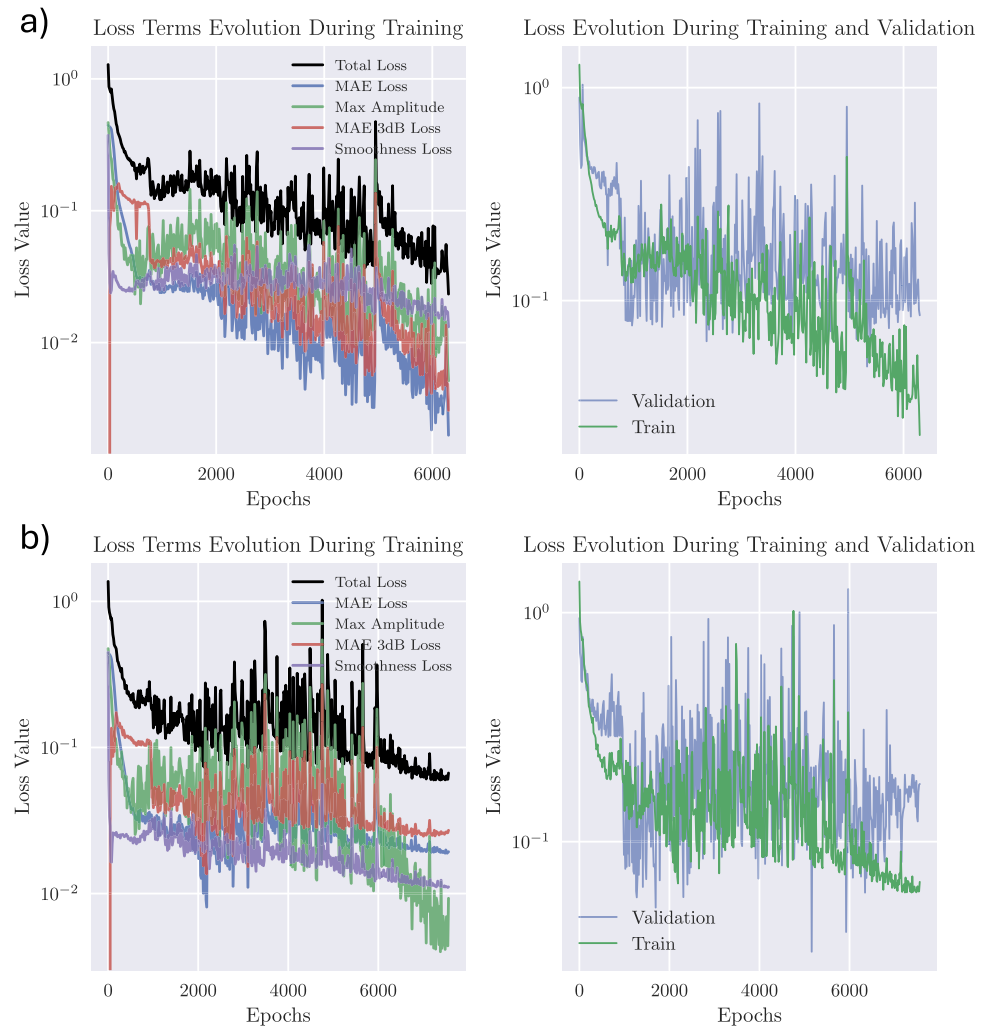
$$\gamma_1 = 1, \quad \gamma_2 = 1, \quad \gamma_3 = 2, \quad \gamma_4 = 10.$$

For the model architecture #1, training is terminated after 6781 epochs, according to the early stopping criterion. For the model architecture #2, training ends after 7571 epochs. Figure 8 illustrates, on the left, the evolution of each individual loss term along with the total loss in the training set. On the right, the figure shows the evolution of the total loss for both training and validation sets over the epochs. The absence of overfitting is a result of the regularization strategies previously described. The loss curves indicate stable convergence, with all individual loss components decreasing steadily and the total loss showing no abrupt fluctuations. The training and validation losses follow similar trajectories throughout, confirming the absence of overfitting and the effectiveness of the applied regularization strategies.

### 5.3 Evaluation of the augmented model's performance

To further assess the improvements of considering the regularization terms in the total loss function of the neural networks in their predictions capabilities, a purely data-driven neural network also is proposed, such that  $\gamma_2 = \gamma_3 = \gamma_4 = 0$  in Equation (39). The architecture considered for the data-driven network is the same as the model #2, and the evolution of its loss function evolution is in Appendix B. After training the neural network, the next step is to evaluate the performance of the augmented models in comparison to the experimental results, the purely data-driven neural network and to assess whether the proposed methodology offers an improvement over the baseline model. To achieve this hypothesis, the mean absolute percentage error (MAPE) is used as the performance metric. As shown in Fig. 9, the comparison between the augmented and baseline models demonstrates a significant improvement. The mean MAPE of the baseline model is 81.1%, whereas the non-optimized and optimized augmented models achieve 7.8% and 2.22%, respectively, highlighting the performance gains achieved through the proposed optimization. Considering the purely data-driven network, the mean MAPE is 40.3%. These results indicate the improvements obtained by including the regularization terms in the loss functions. Also, as shown in Fig. 9, the optimized model performs better on the test

**Fig. 8** Evolution of loss values during the training process. **a** First architecture: 100 neurons per layer and 4 layers, without optimization. **b** Second architecture: 67 neurons per layer and 6 layers, obtained via optimization

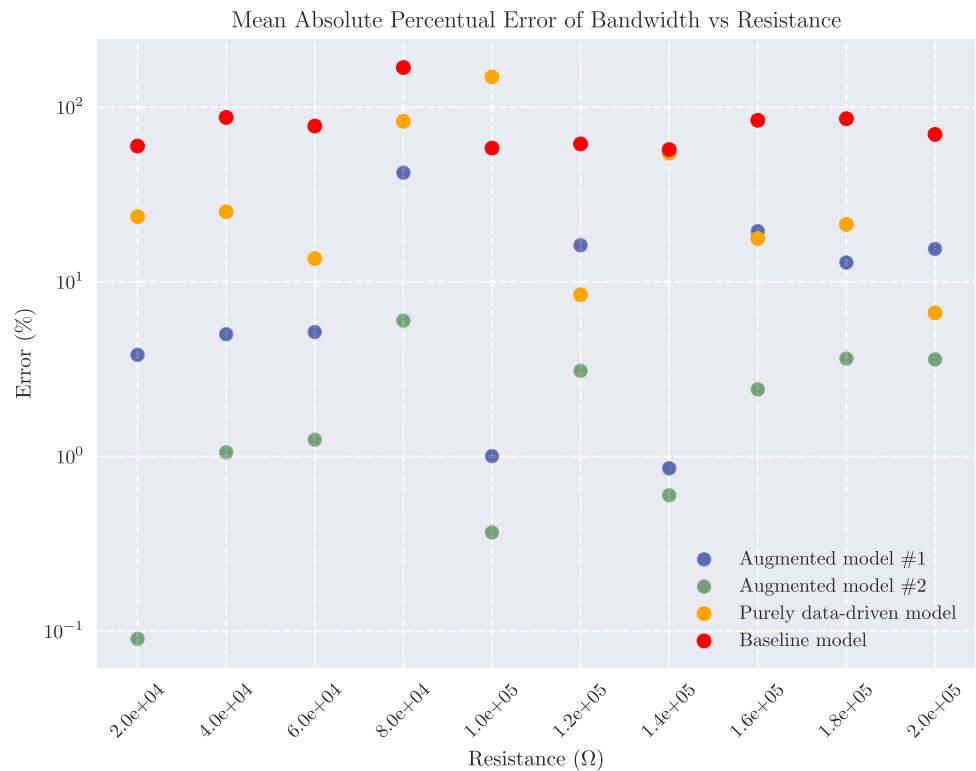


resistances, indicating a greater ability to generalize to unseen data.

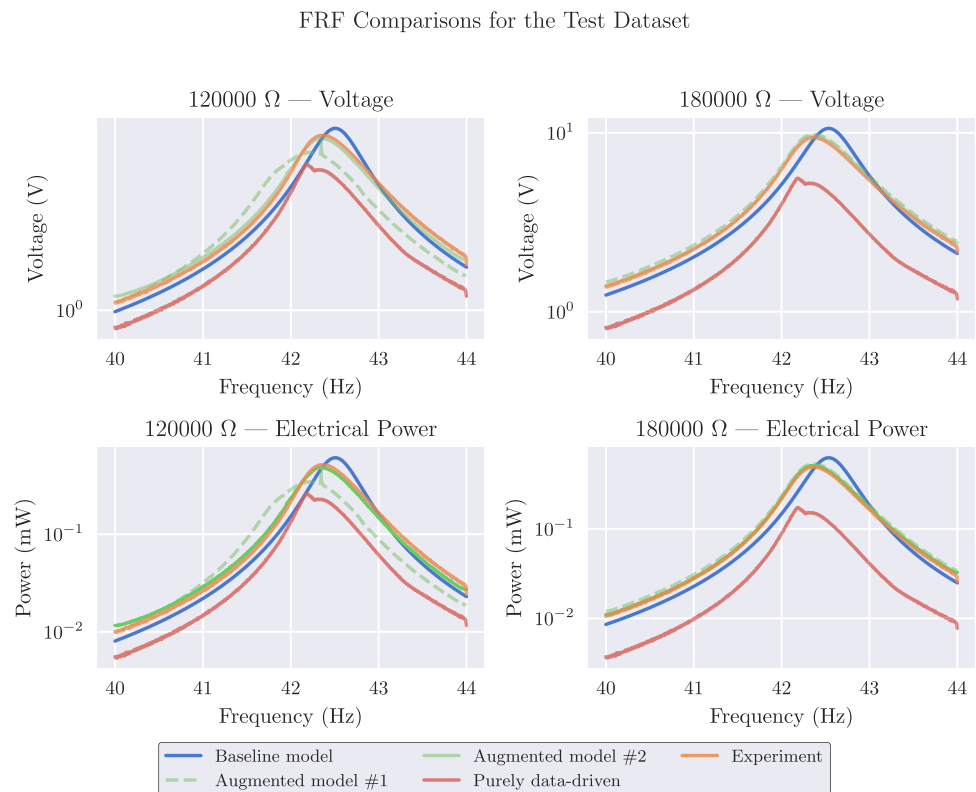
It is crucial to evaluate how the augmented model reconstructs the amplitude of the electrical voltage and power frequency response curves and how it compares to the baseline model, the experimental responses, and the purely data-driven network. This comparison is shown in Fig. 10 for the cases from the test dataset, which, unlike the training and validation sets, was never included in the training loop. From these graphs, it is evident that the augmented model provides a significantly better approximation of the voltage and electrical power response across the entire frequency range, not just within the -3 dB bandwidth. Additionally, the augmented model offers improved predictions for both the maximum

value and the peak frequency. In the test cases presented in the graph, the augmented model's predictions nearly coincide with the experimental amplitudes, whereas the baseline model displays a more considerable discrepancy.

While the baseline model and the purely data-driven network appear qualitatively reasonable, the augmented model achieves systematically higher accuracy across all evaluated conditions. This improvement is particularly relevant for predicting key design parameters such as peak amplitude, peak frequency, and bandwidth, which are critical for optimizing energy harvesting performance. Moreover, the proposed framework extends the baseline model's validity to regimes where its assumptions no longer hold, offering a transferable methodology for other weak nonlinear energy harvesting systems.

**Fig. 9** MAPE for each resistance

**Fig. 10** Frequency response curves comparisons for the test dataset. The top row shows the voltage response, while the bottom row presents the corresponding electrical power response. The left column corresponds to a resistance of 120 k $\Omega$ , and the right column to 180 k $\Omega$ . Overall, the augmented model shows closer agreement with the experimental observations than the baseline model and than the purely data-driven model



## 6 Conclusions

This study explores the behavior of a nonlinear piezoelectric vibration energy harvester consisting of a cantilever beam with a PZT patch and a tip mass, subjected to magnetic forces. The PVEH is initially modeled analytically, assuming geometric linearity with the only nonlinear term arising from the magnetic interaction. The PVEH's equations are derived based on existing literature, resulting in a Duffing-type nonlinearity, referred to as the baseline model. To investigate the PVEH's behavior, an experimental setup is designed, and tests are conducted at accelerations of 0.02g, 0.04g, and 0.10g, with varying load resistance values. The physical parameters of the equations are identified experimentally using Particle Swarm Optimization to fit the model.

After calibrating the parameters, the baseline model is compared to the experimental frequency response curves for 0.02g, 0.04g, and 0.10g. The results show that while the model accurately predicts the behavior at 0.02g and 0.04g, it fails to capture the dynamics at 0.10g. This mismatch highlights that nonlinear effects not accounted for in the analytical formulation, such as geometric nonlinearities and magnetic field reorientation, become increasingly significant at higher excitation levels, leading to distortions in both amplitude and resonance frequency. These deviations explain why the baseline model underestimates the softening behavior observed experimentally.

To extend the model's range of applicability, a physics-augmented neural network is introduced with two distinct architectures, resulting in two augmented models. Physics is integrated into the ANNs in three ways: the input is the voltage amplitude from the baseline model, the loss function is customized to prioritize accuracy within the -3 dB bandwidth and enforce physically consistent values, and the softplus activation function is used to ensure non-negative outputs, as required by the frequency response curve amplitude definition. A direct comparison with a purely data-driven network, obtained by removing the physical regularization terms from the loss function, shows a reduction in predictive accuracy and generalization capability, highlighting that these differences are associated with the presence of physical regularization terms from the physics-based constraints rather than from network complexity alone. The architecture of the optimized augmented model is determined through a hyperparameter search using the Optuna library, resulting in 6 layers and 67 neurons per layer. The main conclusion regarding the loss function design is that the augmented model remains consistent with the physical nature of the response while gaining sufficient

flexibility to capture unmodeled nonlinearities, thereby mitigating the “black-box” character typically associated with neural network approaches.

The results demonstrate a substantial improvement in predictive capability: the baseline model exhibits an average MAPE of 81.1%, whereas the non-optimized and optimized augmented models achieve 7.8% and 2.2%, respectively, and the purely data-driven model achieves 40.3%. These findings show not only a quantitative improvement but also a qualitative one, as the frequency-response curves predicted by the augmented model closely reproduce experimental peak positions, bandwidths, and curve shapes across the evaluated resistances. Furthermore, the optimized model generalizes to unseen test resistances, since the qualitative trends predicted by the analytical model are preserved by the augmented framework.

Overall, the findings confirm that integrating physics-based modeling with data-driven learning extends the validity of simplified analytical models into regimes where their underlying assumptions lose accuracy. In the context of vibration energy harvesting, such improved predictive capability is particularly relevant, as reliable estimation of nonlinear frequency-response behavior is essential for design objectives such as bandwidth maximization (here defined as the -3 dB range). In addition, the augmented model developed in this work expands the applicability of metamodels traditionally limited to the prediction of isolated features (such as the maximum voltage amplitude or the resonant frequency) by addressing the full frequency domain. Another conclusion of particular relevance to design concerns physical consistency: by retaining the analytical formulation in parallel with the correction of experimental discrepancies, the insights provided by the simplified model (such as the identification of optimal load resistances) are preserved, despite its restricted assumptions associated with the simplified hypothesis.

Future developments of this work will address two main directions. First, the framework will be extended to account for jump phenomena in the model output. Since these nonlinear effects are associated with the coexistence of multiple solutions near resonance, the neural network will need to be adapted to represent multivalued responses. In this context, physics-informed regularization based on energy or power balance principles, in the spirit of physics-informed neural networks, could be leveraged to guide the learning process toward physically admissible solution branches and improve the representation of jump-related nonlinearities. Although a few recent studies have begun to explore such capabilities, their application to energy harvesting remains

largely unexplored. Second, the proposed approach could be integrated into optimization procedures aimed at maximizing harvested power. In this context, the augmented model could serve as a fast and accurate surrogate for design exploration, enabling the identification of optimal PVEH configurations and operating conditions.

## Appendix

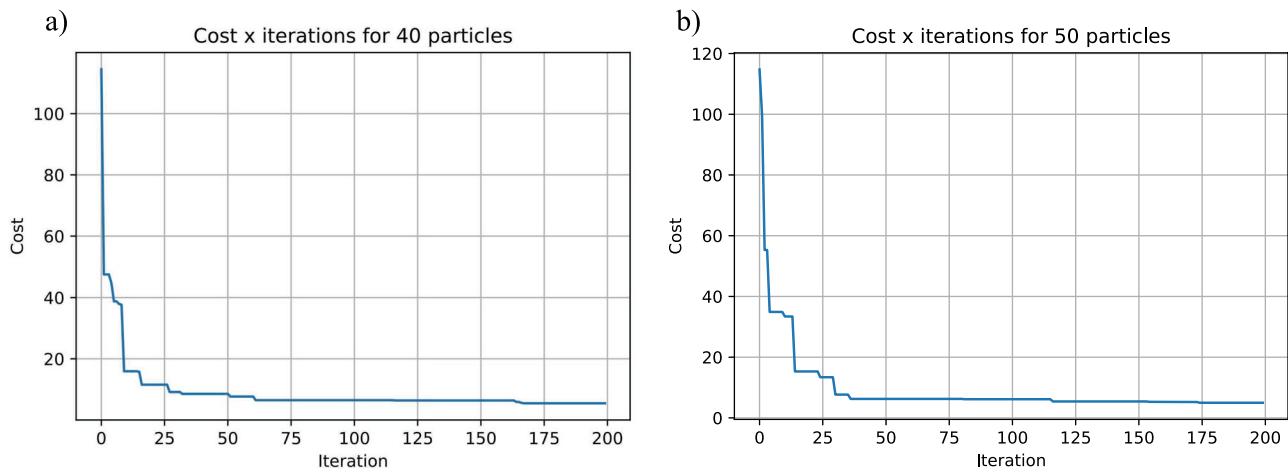
### Number of particles and maximum number of iterations selection

In the particle swarm optimization, each particle represents a candidate solution in the parameter space, defined by the vector of unknown parameters  $\Theta = (\zeta, \omega_0, F_m^{p_0}, k_{nl}^{p_0}, \kappa)$ , balancing exploration and computational cost. Tests with

smaller swarms (10–15 particles) showed weaker convergence and occasional trapping in local minima, while larger swarms ( $\gtrsim 70$ ) provided no substantial improvement in solution quality or parameter consistency. Figure 11 illustrates the optimization process in two scenarios: in (a), 40 particles are used, whereas in (b) the population size is set to 50. It is worth noting that, in both cases, increasing the number of iterations beyond 100 does not significantly improve the identification of the physical parameters. In order to ensure convergence of the search, a total of 50 particles and 200 iterations is adopted in the PSO, balancing computational cost and numerical accuracy.

### Loss evolution of purely data-driven model

In Fig. 12, it is shown the loss evolution corresponding to the purely data-driven model, where the loss corresponds only to the mean absolute error.



**Fig. 11** Convergence behavior of the particle swarm optimization. (a) Evolution of the cost function using 40 particles. (b) Evolution of the cost function using 50 particles. In both cases, the cost function sta-

bilizes after approximately 100 iterations, indicating that increasing the number of iterations beyond this threshold does not significantly improve the identification of the physical parameters

**Fig. 12** Evolution of loss values during the training process for the purely data-driven model



**Acknowledgements** This research work has been carried out within the framework of EIPHI Graduate School (Contract No. ANR-17-EURE-0002), supported by Bourgogne-Franche-Comté Region, and achieved in the frame of FEMTO-ST technological facilities AME-TISTE. Authors acknowledge the financial support granted by São Paulo Research Foundation (FAPESP) (grants #2022/09350-3 and #2023/15931-1).

**Funding** Open access funding provided by Université Marie et Louis Pasteur.

**Data Availability** The data that support the findings of this study are available from the corresponding author upon reasonable request.

## Declarations

**Conflict of 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.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Brown TW et al (2018) Response to 'burden of proof: A comprehensive review of the feasibility of 100% renewable-electricity systems'. *Renew Sustain Energy Rev* 92:834–847. <https://doi.org/10.1016/j.rser.2018.04.113>
- Sezer N, Koç M (2021) A comprehensive review on the state-of-the-art of piezoelectric energy harvesting. *Nano Energy* 80:105567. <https://doi.org/10.1016/j.nanoen.2020.105567>
- Wu Z, Wang Y, Hong Q, Sun H, Lin S (2025) High output and durable piezoelectric energy harvesting system utilizing non-contact collision-induced self-vibration mechanism. *J Braz Soc Mech Sci Eng*. <https://doi.org/10.1007/s40430-025-05546-5>
- Ku M-L, Li W, Chen Y, Liu KR (2015) Advances in energy harvesting communications: Past, present, and future challenges. *IEEE Commun Surv Tutor* 18:1384–1412. <https://doi.org/10.1109/COMST.2015.2497324>
- Lv H et al (2021) A flexible electromagnetic wave-electricity harvester. *Nat Commun* 12:834. <https://doi.org/10.1038/s41467-021-21103-9>
- Zhang L, Dai H, Yang Y, Wang L (2019) Design of high-efficiency electromagnetic energy harvester based on a rolling magnet. *Energy Convers Manage* 185:202–210. <https://doi.org/10.1016/j.enconman.2019.01.089>
- Wang S, Miao G, Zhou S, Yang Z, Yurchenko D (2022) A novel electromagnetic energy harvester based on the bending of the sole. *Appl Energy* 314:119000. <https://doi.org/10.1016/j.apenergy.2022.119000>
- Ebrahimi R, Ziaei-Rad S (2024) Design, modelling and experimental verification of a tunable electromagnetic generator for multi-directional vibration energy harvesting. *J Braz Soc Mech Sci Eng*. <https://doi.org/10.1007/s40430-024-04703-6>
- Zhang Y et al (2018) Micro electrostatic energy harvester with both broad bandwidth and high normalized power density. *Appl Energy* 212:362–371. <https://doi.org/10.1016/j.apenergy.2017.12.053>
- Vysotskiy B, Campos J-FA, Lefeuvre E, Brenes A (2024) Dynamic analysis of a novel two-sided nonlinear mems electrostatic energy harvester. *Mech Syst Signal Process* 206:110932. <https://doi.org/10.1016/j.ymssp.2023.110932>
- Kim W-G et al (2021) Triboelectric nanogenerator: Structure, mechanism, and applications. *ACS Nano* 15:258–287. <https://doi.org/10.1021/acsnano.0c09803>
- Zhao C, Yang Y, Upadrashta D, Zhao L (2021) Design, modeling and experimental validation of a low-frequency cantilever triboelectric energy harvester. *Energy* 214:118885. <https://doi.org/10.1016/j.energy.2020.118885>
- Wang C et al (2022) Double-spring-piece structured triboelectric sensor for broadband vibration monitoring and warning. *Mech Syst Signal Process* 166:108429. <https://doi.org/10.1016/j.ymssp.2021.108429>
- Safaei M, Sodano HA, Anton SR (2019) A review of energy harvesting using piezoelectric materials: state-of-the-art a decade later (2008–2018). *Smart Mater Struct* 28:113001. <https://doi.org/10.1088/1361-665X/ab36e4>
- Yang Z, Zhou S, Zu J, Inman D (2018) High-performance piezoelectric energy harvesters and their applications. *Joule* 2:642–697
- Fan K, Tan Q, Liu H, Zhang Y, Cai M (2019) Improved energy harvesting from low-frequency small vibrations through a monostable piezoelectric energy harvester. *Mech Syst Signal Process* 117:594–608. <https://doi.org/10.1016/j.ymssp.2018.08.001>
- Li H, Liu K, Deng J (2024) Modeling and evaluation of a multi-stable hybrid energy harvester. *Vibration* 7:662–686. <https://doi.org/10.3390/vibration7030035>
- Wang J, Geng L, Ding L, Zhu H, Yurchenko D (2020) The state-of-the-art review on energy harvesting from flow-induced vibrations. *Appl Energy* 267:114902. <https://doi.org/10.1016/j.apenergy.2020.114902>
- Zhang Z et al (2020) A pneumatic piezoelectric vibration energy harvester based on the compressed air-transducer-structure interaction. *Energy Convers Manage* 213:112861. <https://doi.org/10.1016/j.enconman.2020.112861>
- Schimidt CS, de Sousa VC, De Marqui JC (2023) Piezoelectric energy harvesting in graded elastic metastructures using continuous and segmented electrodes. *J Braz Soc Mech Sci Eng*. <https://doi.org/10.1007/s40430-023-04248-0>
- Ajitsaria J, Choe S-Y, Shen D, Kim D (2007) Modeling and analysis of a bimorph piezoelectric cantilever beam for voltage generation. *Smart Mater Struct* 16:447. <https://doi.org/10.1088/0964-1726/16/2/024>
- Ly R et al (2011) Modeling and characterization of piezoelectric cantilever bending sensor for energy harvesting. *Sens Actuators, A* 168:95–100. <https://doi.org/10.1016/j.sna.2011.04.020>
- Alcântara JDL et al (2025) A dual-function nonlinear electromechanical oscillator for thermal sensing and vibration energy harvesting. *Smart Mater Struct* 34:115010. <https://doi.org/10.1088/1361-665X/ae1916>
- Baù M, Alghisi D, Dalola S, Ferrari M, Ferrari V (2020) Multi-frequency array of nonlinear piezoelectric converters for vibration energy harvesting. *Smart Mater Struct* 29:085047. <https://doi.org/10.1088/1361-665X/ab9a8d>
- Kouritem SA, Al-Moghazy MA, Noori M, Altabay WA (2022) Mass tuning technique for a broadband piezoelectric energy harvester array. *Mech Syst Signal Process* 181:109500. <https://doi.org/10.1016/j.ymssp.2022.109500>

26. Gibus D et al (2020) Strongly coupled piezoelectric cantilevers for broadband vibration energy harvesting. *Appl Energy* 277:115518. <https://doi.org/10.1016/j.apenergy.2020.115518>
27. Na Y, Lee M-S, Lee JW, Jeong YH (2020) Wind energy harvesting from a magnetically coupled piezoelectric bimorph cantilever array based on a dynamic magneto-piezo-elastic structure. *Appl Energy* 264:114710. <https://doi.org/10.1016/j.apenergy.2020.114710>
28. Asanuma H (2021) Selecting nonlinear piezoelectricity for fully autonomous self-sensing synchronized switch damping on inductor technique. *Mech Syst Signal Process* 159:107846. <https://doi.org/10.1016/j.ymssp.2021.107846>
29. Asanuma H, Komatsuzaki T (2020) Nonlinear piezoelectricity and damping in partially-covered piezoelectric cantilever with self-sensing synchronized switch damping on inductor circuit. *Mech Syst Signal Process* 144:106867. <https://doi.org/10.1016/j.ymssp.2020.106867>
30. Burda I (2022) Nonlinear piezoelectric energy harvester: Experimental output power mapping. *Vibration* 5:483–496. <https://doi.org/10.3390/vibration5030027>
31. Barbosa A, Kacem N, Bouhaddi N (2024) A design methodology for nonlinear oscillator chains enabling energy localization tuning and soliton stability enhancement with optimal damping. *Mech Syst Signal Process* 213:111358. <https://doi.org/10.1016/j.ymssp.2024.111358>
32. Fan Y, Zhang Y, Niu M-Q, Chen L-Q (2024) An internal resonance piezoelectric energy harvester based on geometrical nonlinearities. *Mech Syst Signal Process* 211:111176. <https://doi.org/10.1016/j.ymssp.2024.111176>
33. Wang G et al (2019) Nonlinear magnetic force and dynamic characteristics of a tri-stable piezoelectric energy harvester. *Nonlinear Dyn* 97:2371–2397. <https://doi.org/10.1007/s11071-019-05133-z>
34. de Lima AMG, Delgado-Filho M, da Silva A, Bouhaddi N (2025) An efficient modeling methodology of piezoaeroviscoelastic systems for vibration-based energy harvesting and subsonic flutter suppression. *J Braz Soc Mech Sci Eng*. <https://doi.org/10.1007/s40430-025-05410-6>
35. Sebald G, Kuwano H, Guyomar D, Ducharme B (2011) Simulation of a duffing oscillator for broadband piezoelectric energy harvesting. *Smart Mater Struct* 20:075022. <https://doi.org/10.1088/0964-1726/20/7/075022>
36. Mahmoudi S, Kacem N, Bouhaddi N (2014) Enhancement of the performance of a hybrid nonlinear vibration energy harvester based on piezoelectric and electromagnetic transductions. *Smart Mater Struct* 23:075024. <https://doi.org/10.1088/0964-1726/23/7/075024>
37. Garg A, Dwivedy SK (2020) Piezoelectric energy harvester under parametric excitation: a theoretical and experimental investigation. *J Intell Mater Syst Struct* 31:612–631. <https://doi.org/10.1177/1045389X19891523>
38. Li J, He X, Yang X, Liu Y (2020) A consistent geometrically nonlinear model of cantilevered piezoelectric vibration energy harvesters. *J Sound Vib* 486:115614. <https://doi.org/10.1016/j.jsv.2020.115614>
39. Alazemi AH, Kurdila AJ (2025) Modeling hysteretically nonlinear piezoelectric composite beams. *Vibration* 8:37. <https://doi.org/10.3390/vibration8030037>
40. LeCun Y, Bengio Y, Hinton G (2015) Deep learning nature 521:436–444. <https://doi.org/10.1038/nature14539>
41. Mohajerin N, Waslander SL (2019) Multistep prediction of dynamic systems with recurrent neural networks. *IEEE Trans Neural Netw Learn Syst* 30:3370–3383. <https://doi.org/10.1109/TNNLS.2019.2891257>
42. Santos VMDS, Martins YA, dos Santos HE, Sales TDP, Rade DA (2025) Stochastic modeling of periodic beams under uncertain boundary conditions and environmental fluctuations. *Int J Mech Sci* 285:109779. <https://doi.org/10.1016/j.ijmecsci.2024.109779>
43. Rosenkranz M, Kalina KA, Brummund J, Sun W, Kästner M (2024) Viscoelasticity with physics-augmented neural networks: Model formulation and training methods without prescribed internal variables. *Comput Mech* 74:1279–1301. <https://doi.org/10.1007/s00466-024-02477-1>
44. Teloli RDO et al (2025) A physics-informed neural networks framework for model parameter identification of beam-like structures. *Mech Syst Signal Process* 224:112189. <https://doi.org/10.1016/j.ymssp.2024.112189>
45. Raissi M, Perdikaris P, Karniadakis GE (2019) Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J Comput Phys* 378:686–707. <https://doi.org/10.1016/j.jcp.2018.10.045>
46. Kashyap S, Rogers TJ, Nayek R (2024) A gaussian-process assisted model-form error estimation in multiple-degrees-of-freedom systems. *Mech Syst Signal Process* 216:111474. <https://doi.org/10.1016/j.ymssp.2024.111474>
47. Brunton SL, Proctor JL, Kutz JN (2016) Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proc Natl Acad Sci* 113:3932–3937. <https://doi.org/10.1073/pnas.1517384113>
48. Klein DK, Roth FJ, Valizadeh I, Weeger O (2023) Parametrized polyconvex hyperelasticity with physics-augmented neural networks. *Data-Centric Eng* 4:e25. <https://doi.org/10.1017/dce.2023.21>
49. Hornik K, Stinchcombe M, White H (1989) Multilayer feedforward networks are universal approximators. *Neural Netw* 2:359–366. [https://doi.org/10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8)
50. Barbosa A, Sena J, Kacem N, Bouhaddi N (2023) An artificial intelligence approach to design periodic nonlinear oscillator chains under external excitation with stable damped solitons. *Mech Syst Signal Process* 205:110879. <https://doi.org/10.1016/j.ymssp.2023.110879>
51. Huang D, Wang P, Li W, Li R, Liu L (2025) Modeling, response and bpn-pid control of symmetric multistable galloping energy harvesters. *Chaos Solitons Fract* 199:116736. <https://doi.org/10.1016/j.chaos.2025.116736>
52. De Groote W, Kikken E, Hostens E, Van Hoecke S, Crevecoeur G (2021) Neural network augmented physics models for systems with partially unknown dynamics: application to slider-crank mechanism. *IEEE/ASME Trans Mechatron* 27:103–114. <https://doi.org/10.1109/TMECH.2021.3058536>
53. Benady A, Baranger E, Chamoin L (2024) Nn-mcre: A modified constitutive relation error framework for unsupervised learning of nonlinear state laws with physics-augmented neural networks. *Int J Numer Meth Eng* 125:e7439. <https://doi.org/10.1002/nme.7439>
54. Klein DK, Ortigosa R, Martínez-Frutos J, Weeger O (2022) Finite electro-elasticity with physics-augmented neural networks. *Comput Methods Appl Mech Eng* 400:115501. <https://doi.org/10.1016/j.cma.2022.115501>
55. Najera-Flores DA, Todd MD (2024) State-space reconstruction from partial observables using an invertible neural network with structure-preserving properties for nonlinear structural dynamics. *Nonlinear Dyn* 112:18055–18077. <https://doi.org/10.1007/s11071-024-09642-4>
56. Fuhg JN, Jones RE, Bouklas N (2024) Extreme sparsification of physics-augmented neural networks for interpretable model discovery in mechanics. *Comput Methods Appl Mech Eng* 426:116973. <https://doi.org/10.1016/j.cma.2024.116973>
57. Franke M, Klein DK, Weeger O, Betsch P (2023) Advanced discretization techniques for hyperelastic physics-augmented neural networks. *Comput Methods Appl Mech Eng* 416:116333. <https://doi.org/10.1016/j.cma.2023.116333>
58. Yinusa A et al (2025) Machine learning approach to nonlinear fluid-induced vibration of pronged nanotubes in a thermal-magnetic environment. *Vibration* 8:35. <https://doi.org/10.3390/vibration8030035>
59. Cai W, Harn RL (2022) A machine learning approach for maximizing direct current power of nonlinear energy harvesting systems subjected to periodic impulse excitation. *Mech Syst Signal*

- Process 164:108262. <https://doi.org/10.1016/j.ymssp.2021.108262>
60. Zhang C et al (2021) Machine learning based prediction of piezoelectric energy harvesting from wake galloping. *Mech Syst Signal Process* 160:107876. <https://doi.org/10.1016/j.ymssp.2021.107876>
  61. Peralta P, Ruiz R, Natarajan S, Atroshchenko E (2020) Parametric study and shape optimization of piezoelectric energy harvesters by isogeometric analysis and kriging metamodeling. *J Sound Vib* 484:115521. <https://doi.org/10.1016/j.jsv.2020.115521>
  62. Huang D, Zhou S, Litak G (2019) Theoretical analysis of multi-stable energy harvesters with high-order stiffness terms. *Commun Nonlinear Sci Numer Simul* 69:270–286. <https://doi.org/10.1016/j.cnsns.2018.09.025>
  63. Xiao H, Qie H, Bowen CR (2021) Modelling of the circular edge-clamped interface of a hydraulic pressure energy harvester to determine power, efficiency and bandwidth. *Mech Syst Signal Process* 146:107013. <https://doi.org/10.1016/j.ymssp.2020.107013>
  64. Saeed N, Ainamon C, Çiçek S, Kingni ST, Wei Z (2021) Pitchfork and hopf bifurcations in quantum dot light emitting diode: Analysis and prediction by using artificial neural network. *Eur Phys J D* 75:168. <https://doi.org/10.1140/epjd/s10053-021-00188-3>
  65. Yuting D, Haoran R, You W, Chao Y, Yuntao X (2023) Stall flutter prediction based on multi-layer gru neural network. *Chin J Aeronaut* 36:75–90. <https://doi.org/10.1016/j.cja.2022.07.011>
  66. Terasaki Y, Nakajima K (2025) Thermodynamic limit in learning period three. *Phys Rev Res* 7:023119. <https://doi.org/10.1103/PhysRevResearch.7.023119>
  67. Abdelkefi A, Nayfeh AH, Hajj MR (2012) Effects of nonlinear piezoelectric coupling on energy harvesters under direct excitation. *Nonlinear Dyn* 67:1221–1232. <https://doi.org/10.1007/s11071-011-0064-9>
  68. Erturk A, Inman DJ (2008) A distributed parameter electromechanical model for cantilevered piezoelectric energy harvesters. *J Vib Acoust* 130:041002. <https://doi.org/10.1115/1.2890402>
  69. Erturk A, Inman DJ (2008) On mechanical modeling of cantilevered piezoelectric vibration energy harvesters. *J Intell Mater Syst Struct* 19:1311–1325. <https://doi.org/10.1177/1045389X07085639>
  70. Luttinger J, Tisza L (1946) Theory of dipole interaction in crystals. *Phys Rev* 70:954. <https://doi.org/10.1103/PhysRev.70.954>
  71. Barbosa AS, da Silva MM (2023) Macro fiber composite-actuated soft robotic fish: a gray box model-predictive motion planning strategy under limited actuation. *Soft Rob* 10:948–958. <https://doi.org/10.1089/soro.2022.0061>
  72. Barbosa AS, Tahara LZ, da Silva MM (2023) Motion planning of a fish-like piezoelectric actuated robot using model-based predictive control. *J Vib Control* 29:411–427. <https://doi.org/10.1177/1077546321104825>
  73. da Silva MM, Barbosa AS, Savio TL, de Oliveira LPR (2025) Mfc-based propulsion system for body/caudal fin robotic fish locomotion: Experimental and numerical results. *J Vib Eng Technol* 13:250. <https://doi.org/10.1007/s42417-025-01808-7>
  74. Mann B, Owens B (2010) Investigations of a nonlinear energy harvester with a bistable potential well. *J Sound Vib* 329:1215–1226. <https://doi.org/10.1016/j.jsv.2009.11.034>
  75. Bernard BP, Owens BAM, Mann BP (2013) Uncertainty propagation in the band gap structure of a 1d array of magnetically coupled oscillators. *J Vib Acoust* 135:041005. <https://doi.org/10.1115/1.4023821>
  76. Dhote S, Yang Z, Behdinan K, Zu J (2018) Enhanced broadband multi-mode compliant orthoplanar spring piezoelectric vibration energy harvester using magnetic force. *Int J Mech Sci* 135:63–71. <https://doi.org/10.1016/j.ijmecsci.2017.11.012>
  77. Farokhi H, Xia Y, Erturk A (2022) Experimentally validated geometrically exact model for extreme nonlinear motions of cantilevers. *Nonlinear Dyn* 107:457–475. <https://doi.org/10.1007/s11071-021-07023-9>
  78. Norenberg JP, Luo R, Lopes VG, Peterson JVL, Cunha A (2023) Nonlinear dynamics of asymmetric bistable energy harvesters. *Int J Mech Sci* 257:108542. <https://doi.org/10.1016/j.ijmecsci.2023.108542>
  79. Yang Y, Zhao L, Tang L (2013) Comparative study of tip cross-sections for efficient galloping energy harvesting. *Appl Phys Lett* doi 10(1063/1):4792737
  80. Bibo A, Abdelkefi A, Daqaq MF (2015) Modeling and characterization of a piezoelectric energy harvester under combined aerodynamic and base excitations. *J Vib Acoust* 137:031017. <https://doi.org/10.1115/1.4029611>
  81. Tan T, Yan Z, Lei H, Sun W (2017) Geometric nonlinear distributed parameter model for cantilever-beam piezoelectric energy harvesters and structural dimension analysis for galloping mode. *J Intell Mater Syst Struct* 28:3066–3078. <https://doi.org/10.1177/1045389X17704922>
  82. Balachandran B, Magrab EB (2018) *Vibrations*. Cambridge University Press
  83. Rao SS (2019) *Vibration of continuous systems*. John Wiley & Sons
  84. Peng Z, Lang Z, Billings S, Tomlinson G (2008) Comparisons between harmonic balance and nonlinear output frequency response function in nonlinear system analysis. *J Sound Vib* 311:56–73. <https://doi.org/10.1016/j.jsv.2007.08.035>
  85. Gao Y, Leng Y, Fan S, Lai Z (2014) Performance of bistable piezoelectric cantilever vibration energy harvesters with an elastic support external magnet. *Smart Mater Struct* 23:095003. <https://doi.org/10.1088/0964-1726/23/9/095003>
  86. Zou H-X et al (2021) A magnetically coupled bistable piezoelectric harvester for underwater energy harvesting. *Energy* 217:119429. <https://doi.org/10.1016/j.energy.2020.119429>
  87. Kennedy J, Eberhart R (1995) Particle swarm optimization. Vol. 4, 1942–1948 (IEEE, Perth, Australia)
  88. Rahman MA, Al Mamun A, Yao K (2015) Rate dependent direct inverse hysteresis compensation of piezoelectric micro-actuator used in dual-stage hard disk drive head positioning system. *Rev Sci Instrum* 8:86. <https://doi.org/10.1063/1.4928478>
  89. Yin R, Xue B, Brousseau E, Geng Y, Yan Y (2024) Characterizing the electric field-and rate-dependent hysteresis of piezoelectric ceramics shear motion with the bouc-wen model. *Sens Actuators, A* 367:115044. <https://doi.org/10.1016/j.sna.2024.115044>
  90. Cammarano A, Neild S, Burrow S, Inman D (2014) The bandwidth of optimized nonlinear vibration-based energy harvesters. *Smart Mater Struct* 23:055019. <https://doi.org/10.1088/0964-1726/23/5/055019>
  91. Akiba T, Sano S, Yanase T, Ohta T, Koyama M (2019) Optuna: A next-generation hyperparameter optimization framework. 2623–2631
  92. Kingma DP, Ba J (2014) Adam: A method for stochastic optimization. arXiv preprint [arXiv:1412.6980](https://arxiv.org/abs/1412.6980). <https://doi.org/10.48550/arXiv.1412.6980>
  93. Smith LN, Topin N (2019) Super-convergence: Very fast training of neural networks using large learning rates, Vol. 11006, 369–386 (SPIE, 2019)
  94. Karystinos GN, Pados DA (2000) On overfitting, generalization, and randomly expanded training sets. *IEEE Trans Neural Netw* 11:1050–1057. <https://doi.org/10.1109/72.870038>
  95. Caruana R, Lawrence S, Giles C, Leen T, Dietterich T, Tresp V (eds) Overfitting in neural nets: Backpropagation, conjugate gradient, and early stopping. (eds Leen, T., Dietterich, T. & Tresp, V.) *Advances in Neural Information Processing Systems*, Vol. 13 (MIT Press, 2000). [https://proceedings.neurips.cc/paper\\_files/paper/2000/file/059fdcd96baeb75112f09fa1dcc740cc-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2000/file/059fdcd96baeb75112f09fa1dcc740cc-Paper.pdf)