



One-shot learning for the complex dynamical behaviors of weakly nonlinear forced oscillators

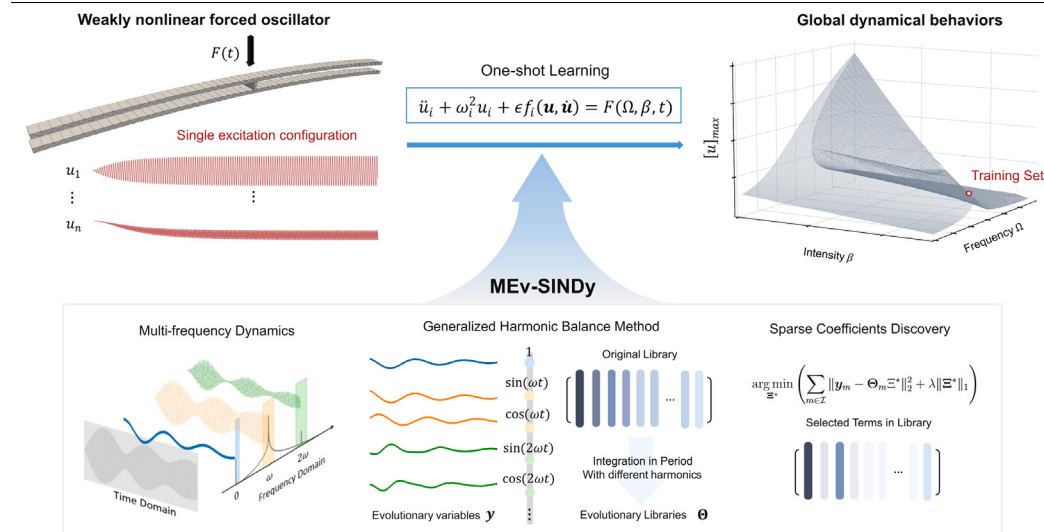
Teng Ma ^{a,b}, Luca Rosafalco ^b, Wei Cui ^a, Lin Zhao ^{a,c}, Attilio Frangi ^b

^a State Key Lab of Disaster Reduction in Civil Engineering, Tongji University, Shanghai, 200092, PR China

^b Department of Civil and Environmental Engineering, Politecnico di Milano, Milan, 20133, Lombardia, Italy

^c Guangxi Laboratory of Whole Life Safety for Land-Sea Corridor Engineering, Guangxi University, Nanning, 530004, Guangxi, PR China

GRAPHICAL ABSTRACT



ARTICLE INFO

Dataset link: <https://github.com/TengMa25/MEv-SINDy.git>

Keywords:

System identification
Nonlinear dynamics
Weak nonlinearity
Generalized harmonic balance

ABSTRACT

Extrapolative prediction of complex nonlinear dynamics remains a central challenge in engineering. This study proposes a one-shot learning method to identify global frequency-response curves from a single excitation time history by learning governing equations. We introduce MEv-SINDy (Multi-frequency Evolutionary Sparse Identification of Nonlinear Dynamics) to infer the governing equations of non-autonomous and multi-frequency systems. The methodology leverages the Generalized Harmonic Balance (GHB) method to decompose complex forced

* Corresponding author.

E-mail address: luca.rosafalco@polimi.it (L. Rosafalco).

<https://doi.org/10.1016/j.cma.2026.119332>

Received 13 April 2026; Received in revised form 15 June 2026; Accepted 17 August 2026

Available online 4 September 2026

0045-7825/© 2026 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

One-shot learning

responses into a set of slow-varying evolution equations. We validated the capabilities of MEv-SINDy on two critical Micro-Electro-Mechanical Systems (MEMS). These applications include a nonlinear beam resonator and a MEMS micromirror. Our results show that the model trained on a single point accurately predicts softening/hardening effects and jump phenomena across a wide range of excitation levels. This approach significantly reduces the data acquisition burden for the characterization and design of nonlinear microsystems.

1. Introduction

The rapid advancement of computational resources has enabled the high-fidelity simulation of complex physical phenomena, even within intricate multi-physics and multi-scale frameworks. However, characterizing the complex dynamical behaviors of physical systems remains a formidable challenge. In practical engineering applications, such as MEMS resonators and sensors, understanding the nonlinear response — including frequency shifts, hardening/softening effects, and jump phenomena — requires extensive frequency and amplitude sweeps. When relying on full-order models (FOMs), such as the finite element method (FEM), each individual simulation in the time domain or parametric space can be computationally very demanding. Consequently, exploring a broad range of initial conditions and parameter combinations becomes a computationally prohibitive task or even unfeasible for real-time monitoring and design optimization. Unlike simple output estimation, capturing the evolution of a high-dimensional nonlinear field is intrinsically difficult due to its time-dependent nature and the sensitivity of nonlinear regimes. These challenges drive the urgent need for efficient yet accurate reduced-order models (ROMs) that can bypass the heavy burden of FOMs while preserving the essential physics of the dynamical system.

Efficient ROM construction begins with dimensionality reduction [1]. This process projects high-dimensional FOM states onto a low-dimensional manifold from data previously collected from full-order simulations or experiments [2]. Historically, the reduced basis method has been the standard for building these reduced spaces [3–5]. It often utilizes Proper Orthogonal Decomposition (POD) [6–8] to approximate solutions. However, POD-based techniques are typically intrusive. They require direct access to the internal operators of the FOM. This requirement makes it difficult to use them with proprietary or commercial simulation software. Machine learning has emerged as a powerful non-intrusive alternative for ROM construction. These methods reduce dimensionality directly from data streams. They do not require access to the FOM operators. Dynamic Mode Decomposition (DMD) [9,10] is a prominent linear framework in this category. It extracts dominant frequencies and spatial modes from snapshot data. However, DMD is limited by its linear assumption. Autoencoder (AE) networks [11–13] provide a nonlinear alternative. They are particularly effective at uncovering latent features in complex systems. These networks offer better nonlinear compression than linear methods like POD or DMD. Advanced frameworks like deep learning-based ROMs (DL-ROMs) [14,15] and their enhanced version through POD (POD-DL-ROMs) [16] have further improved this process. They perform dimensionality reduction and learn the parameter-to-solution map at the same time.

Data-driven ROMs face significant challenges when predicting scenarios outside the training set. These methods are essentially interpolation tools. They provide high accuracy only within the observed parameter ranges. When these models encounter unseen parameter values, their performance often degrades. This lack of generalization limits their use in exploring unknown physical regimes. Furthermore, constructing reliable non-intrusive ROMs requires a large volume of training data. Generating these datasets through FOM simulations is computationally expensive. This high cost reduces the overall efficiency of the modeling process. There is a clear need for a framework that offers better extrapolation. Such a framework should also require less training data to maintain physical consistency.

Equation discovery offers a powerful alternative to black-box models. This approach infers governing equations directly from data. There are two main methodological paradigms in this field. The first paradigm is symbolic discovery [17,18]. It seeks to recover closed-form representations of physical laws. The second paradigm is sparse regression, SINDy [19,20]. This method identifies parsimonious models from a predefined library of candidate functions. These techniques were first developed for ordinary differential equations. Researchers later extended them to partial differential equations [21,22]. Recent advances in machine learning have improved these methods. They are now robust against sparse or noisy data [23–26]. This makes them viable for real-world scientific discovery. Equation discovery has catalyzed the field of scientific machine learning. It helps uncover interpretable laws in cyber-physical systems [27], fluid dynamics [28,29], biological and chemical networks [30,31] and complex networks [32,33]. The main appeal of this approach is the generation of an explicit ROM. The resulting model takes the form of a system of ordinary differential equations. Sparsity constraints ensure that the latent dynamics remain interpretable. Users can then analyze the identified system using standard numerical tools for time integration.

Most physical and engineering systems exhibit weakly nonlinear oscillatory behaviors with weak linear dissipation. These include applications in structural engineering and micro-electro-mechanical systems (MEMS). Weak nonlinearities often arise from geometric effects and can lead to pronounced impacts on dynamical performance. However, existing methods often overlook these challenges, as they focus on extracting only the most dominant terms in governing equations. This process often discards those contributions that seem insignificant but are physically important. Recent research has introduced evolutionary-based learning to address this issue. A prominent example is EvLOWN [34], which stands for Encoding Cumulation to Learn Perturbative Nonlinear Oscillatory Dynamics. This method reformulates optimization objectives using the method of averaging. It bridges the gap between weak nonlinearities and slow-varying evolutionary variables by selecting only the terms that are most informative about amplitude and

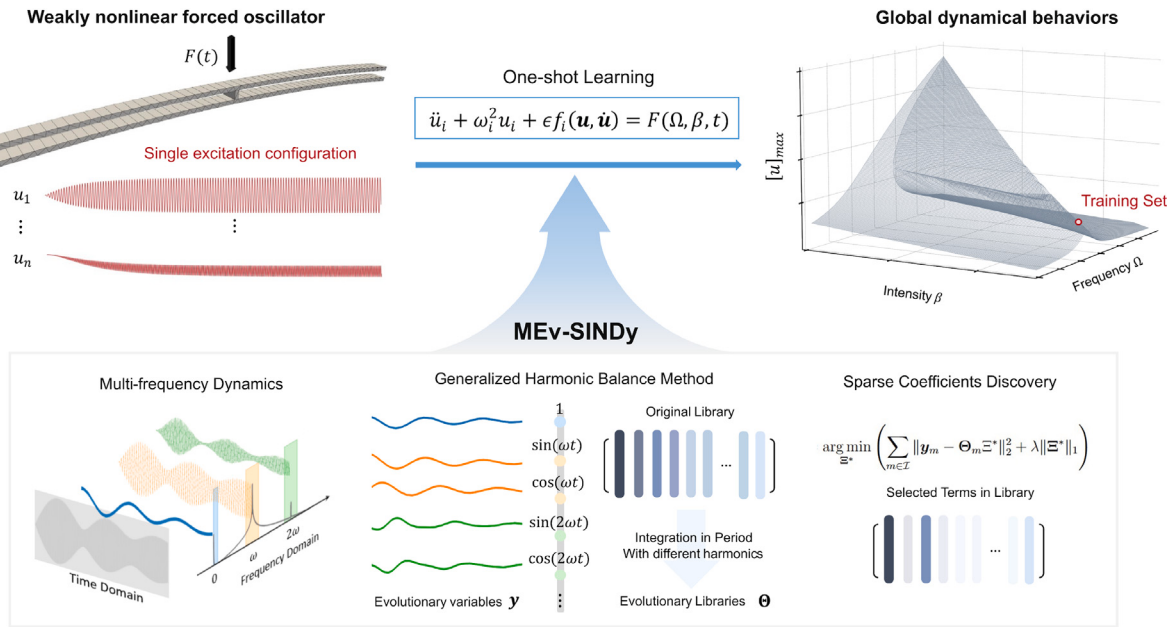


Fig. 1. One-Shot Learning framework of weakly nonlinear forced oscillator: The starting point of our framework is the dynamical response at a single excitation configuration, which is also the training data of MEv-SINDy. During the training stage, the multi-frequency dynamics are separated into distinct harmonic components. We use the generalized harmonic balance method to reformulate the sparse regression problem into an evolution space. A two-stage sparse regression process then discovers the final coefficient vector. The primary output of MEv-SINDy is the explicit governing equation of the weakly nonlinear forced oscillator. Finally, we apply a continuation method to the inferred equations to compute the global frequency-response curves.

phase evolutions. However, EvLOWN has specific limitations as it can only identify autonomous single-frequency responses but cannot handle multi-frequency oscillations or non-autonomous dynamics. In this work, we propose a novel data-driven approach, called Multi-frequency Evolutionary Sparse Identification of Nonlinear Dynamical systems (MEv-SINDy), which belongs to the vast family of SINDy approaches. This framework is designed to learn the governing equations of multi-frequency forced weakly nonlinear oscillators. It extends the advantages of evolutionary learning to more complex forced and multi-frequency regimes.

In addition to leveraging multi-frequency data, the proposed approach supports a one-shot learning paradigm, whereby governing equations can be identified from a single training instance [35]. The concept originated in computer vision but remains rare in scientific machine learning (SciML) [36–38]. Most data-driven ROMs still require massive datasets to characterize physical systems [8,15,39]. However, generating such snapshots for complex nonlinear oscillators is computationally expensive. In our framework we extract knowledge of the system from a single excitation configuration. This analytical representation allows the model to predict dynamical behaviors under entirely different parameters. Consequently, MEv-SINDy can predict the complete frequency-response curve from one observation. This strategy eliminates the need for exhaustive training data. It also provides a robust tool for exploring unknown physical regimes with minimal cost. To our knowledge, this is one of the first applications of one-shot learning for multi-frequency forced oscillators.

A schematic diagram of this work is shown in Fig. 1. The paper is structured as follows. Section 2 contextualizes the weakly nonlinear systems inference problem and mathematically demonstrates why traditional SINDy and SINDy-like approaches fail. Section 3 details the methodology of the proposed MEv-SINDy framework. This section introduces the Generalized Harmonic Balance (GHB) method as an upgrade to the averaging method. The evolutionary variables are expanded from simple amplitude and phase to general harmonic coefficients. Based on GHB, we reformulate the sparse regression optimization targets. We also describe a joint sparse regression method designed for multiple equations. Section 4 presents numerical validations using a 1-DOF oscillator with quadratic and cubic nonlinearities. These tests validate the effectiveness of each step in the MEv-SINDy process. Section 5 considers two high-dimensional applications. These include a Beam MEMS resonator and a MEMS micromirror. We demonstrate that our approach achieves one-shot learning for the frequency-response curve. To further analyze the robustness of the method, we examine the influence of training point selection in Section 6. We describe a multi-point training strategy to improve predictive performance. Finally, Section 5 provides concluding remarks and summarizes the findings.

2. Problem statement and theoretical limitations

Data-driven equation discovery methods, spearheaded by the Sparse Identification of Nonlinear Dynamics (SINDy), have witnessed remarkable advancements in recent years. Extensive research efforts have been directed toward mitigating the effects

of measurement noise through robust optimization algorithms and refined derivative estimation schemes. However, the exact identification of weakly nonlinear systems remains a largely unaddressed frontier.

The fundamental difficulty in modeling weakly nonlinear vibrations lies in the massive scale disparity between physical terms. The governing equations simultaneously contain $O(1)$ dominant linear inertial terms and $O(\epsilon)$ microscopic nonlinear terms. Existing data-driven inference architectures exhibit structural limitations when attempting to capture these subtle, mixed-scale dynamical features. Because algorithms evaluate all terms based on instantaneous high-frequency data, the critical weak nonlinearities are easily overshadowed.

To clearly elucidate the theoretical contributions of the proposed MEv-SINDy framework, this section systematically reviews the inherent bottlenecks of existing methods and details the evolution of our mathematical strategy. The architecture of this section is as follows: Section 2.1 defines the mathematical setup of weakly nonlinear forced oscillators. Section 2.2 briefly reviews the standard SINDy framework. Section 2.3 dissects the ‘‘Pareto trap’’, explaining why hyperparameter tuning intrinsically fails to capture weak terms. Section 2.4 derives the mathematical identifiability limit, demonstrating how weak nonlinearities are buried by observation noise. Section 2.5 introduces EvLOWN, a preceding framework that utilizes classical averaging theory to identify autonomous systems. Finally, Section 2.6 synthesizes these discussions and outlines the theoretical advancements of the proposed MEv-SINDy methodology.

2.1. Problem setup

Consider the general form of weakly nonlinear oscillators with external forcing:

$$\ddot{x}_i(t) + \omega_i^2 x_i(t) + \epsilon f_i(\mathbf{x}(t), \dot{\mathbf{x}}(t)) = F_i(t), \quad (1)$$

for $i = 1, \dots, k$, where: $\mathbf{x}(t) = (x_1(t), \dots, x_k(t)) \in \mathbb{R}^k$ is the displacement vector at time t ; ω_i is the frequency of oscillator i ; $f_i(\mathbf{x}(t), \dot{\mathbf{x}}(t))$ is a smooth function of the velocity $\dot{\mathbf{x}}$ and the displacement $\mathbf{x}$ representing the weak nonlinearity and weak dissipation, which are from geometric, internal, material, contact and coupling nonlinearities; $F_i(t)$ is the external actuation that depends on time t ; ϵ is a small parameter quantifying the strength of nonlinearity compared to dominating linear spring effects $\omega_i^2 x_i$. Eq. (1) is well suited for modeling various classes of systems, whose dimensionality k represents the number of degrees of freedom arising, for example, from spatial discretization techniques (e.g. finite elements or finite volumes) applied to a system of partial differential equations (PDEs) governing the underlying physical problem.

Due to the weak nonlinearity f_i , the steady-state periodic response of the system governed by Eq. (1) under different harmonic actuation F_i exhibits distinct characteristics in terms of amplitude and phase, collectively referred to as the frequency-response curves (FRCs), which characterize the dynamic properties of the system. Typically, the actuation can be parameterized by two features, frequency and intensity, both of which strongly affect the steady-state periodic response. The goal of this study is to utilize data snapshots from a single time history and a given actuation to accurately predict responses under any other actuation conditions, thus reconstructing all the possible FRCs of the system. The proposed methodology employs only few set of time history of displacement $\mathbf{x}$ and its time-derivative $\dot{\mathbf{x}}$. In general, $\dot{\mathbf{x}}$ can be either computed directly or approximated numerically from $\mathbf{x}$.

2.2. Sparse identification of nonlinear dynamics: SINDy

To identify the system in Eq. (1), it is possible to rely on a linear combination of a set of predetermined functions collected in a library, as done in the SINDy method, [19]. To apply SINDy, snapshots of $\mathbf{x}$ at different time instants are collected in the following matrix:

$$\mathbf{X} = \begin{bmatrix} x_1(t_1) & x_2(t_1) & \cdots & x_n(t_1) \\ x_1(t_2) & x_2(t_2) & \cdots & x_n(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ x_1(t_T) & x_2(t_T) & \cdots & x_n(t_T) \end{bmatrix}. \quad (2)$$

where: $x_i(t_j)$ is the i th entry of $\mathbf{x}$ at the j th time instant; T is the number of time instants for each snapshot. Similarly, a matrix $\dot{\mathbf{X}}$ is constructed by collecting $\dot{\mathbf{x}}$.

A library $\Theta(\mathbf{x}) = [\theta_1(\mathbf{x}), \dots, \theta_p(\mathbf{x})] \in \mathbb{R}^p$ of p candidate functions is selected. In Section 3.4, it will be discussed how to retain only the most relevant functions among this set of candidates. The matrix $\Theta(\mathbf{X}) \in \mathbb{R}^{T \times p}$ is thus constructed by applying Θ to the rows of $\mathbf{X}$, for example, as in the following:

$$\Theta(\mathbf{X}) = \begin{bmatrix} | & | & | & | & & | & | & | \\ 1 & \mathbf{X} & \mathbf{X}^{p_2} & \mathbf{X}^{p_3} & \cdots & \sin(\mathbf{X}) & \cos(\mathbf{X}) & \cdots \\ | & | & | & | & & | & | & | \end{bmatrix}. \quad (3)$$

Any nonlinear function, such as polynomial and trigonometric, can be included in the library.

As mentioned, the model to be identified is a linear combination of the candidate functions, with combination coefficients stored in $\Xi = [\xi_1, \dots, \xi_k]$ with $\xi_k \in \mathbb{R}^p$. Combination coefficients are determined by solving a regression problem with a sparsity promoting term, as in the following:

$$\arg \min_{\Xi} \|\ddot{\mathbf{x}} - \Theta(\mathbf{x}, \dot{\mathbf{x}}, t)\Xi\|_2^2 + \lambda \|\Xi\|_1 \quad (4)$$

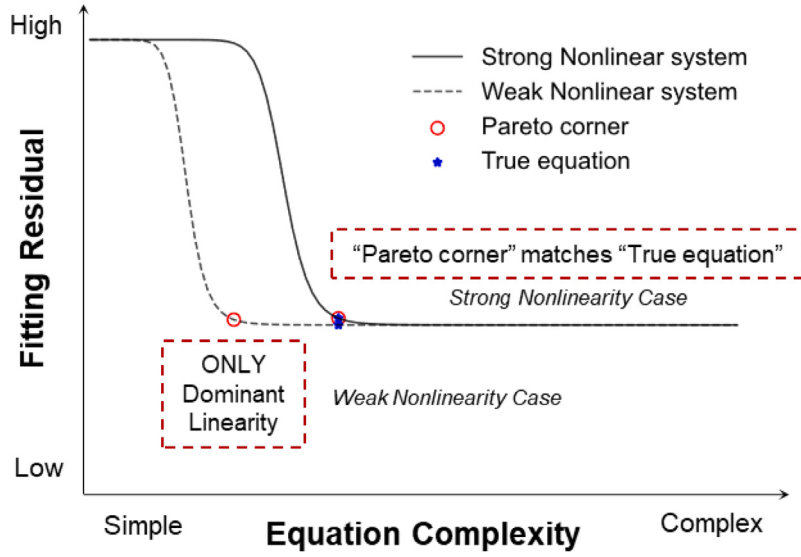


Fig. 2. Illustration of the Pareto front in sparse equation discovery for strongly versus weakly nonlinear systems. In strongly nonlinear systems (where the nonlinearity ϵf is comparable to the linear term $\omega^2 x$), pruning active nonlinear terms causes a sharp increase in the fitting residual, creating a distinct Pareto “corner” that accurately matches the true equation. Conversely, in weakly nonlinear systems ($\epsilon f \ll \omega^2 x$), the instantaneous signal variance is overwhelmingly dominated by the linear terms.

where $\|\cdot\|_2^2$ is the least squares term; $\|\cdot\|_1$ is the L_1 norm of the coefficient vector Ξ , which is the sum of absolute values of the coefficients; λ is the regularization parameter that controls the strength of the penalty. In Section 3.4, Eq. (4) will be modified for a better coupling with the proposed algorithm.

2.3. Pareto front and adaptive regularization

The fundamental premise of data-driven equation discovery involves reformulating the dynamical inference problem into an overdetermined linear regression system, typically expressed as $\dot{\mathbf{X}} = \Theta(\mathbf{X})\Xi$. Here, $\dot{\mathbf{X}}$ represents the derivative matrix, $\Theta(\mathbf{X})$ is the candidate library comprising nonlinear functions of the system states, and Ξ is the unknown sparse coefficient matrix. To prevent overfitting and promote physical interpretability, sparse regularization methods are extensively employed. Prominent techniques include the Least Absolute Shrinkage and Selection Operator (LASSO) using L_1 regularization [40], the Sequentially Thresholded Least-Squares (STLSQ) algorithm utilizing hard thresholding [19], and the unified Sparse Relaxed Regularized Regression (SR3) [41].

In these sparse optimization frameworks, identifying the true governing equation transforms into a multi-objective optimization problem: balancing model complexity (the number of active terms, $\|\Xi\|_0$) against data fitting accuracy (the residual error, $\|\dot{\mathbf{X}} - \Theta\Xi\|_2$). The optimal model is conventionally identified at the “elbow” or “corner” of the fitting residual versus complexity curve, known as the Pareto front [42,43]. However, to mitigate the noise susceptibility and inherent structural biases of standard sparse penalties in locating this optimal corner, advanced adaptive regularization schemes have been developed. Notably, Cortiella et al. [44] proposed an iterative reweighted L_1 -regularized least squares framework and coupled it with a priori denoising strategies [45].

While the strategy of locating the Pareto corner is highly effective for isolating strongly nonlinear dynamics, this paradigm inherently fails for weakly nonlinear systems. As conceptually illustrated in Fig. 2, even if the sparse optimization algorithm successfully converges to the geometric corner of the Pareto front, the resulting model will strictly retain only the dominant linear terms, entirely discarding the critical weakly nonlinear components. The reason is that the contribution of the weak nonlinearity ($O(\epsilon)$) to the reduction of the instantaneous fitting residual is so minuscule that its inclusion does not create a mathematically distinguishable “elbow” on the Pareto curve. While advanced adaptive methods like the reweighted L_1 scheme [44,45] help the optimization algorithm converge to a sharp Pareto corner, they paradoxically accelerate the elimination of weak nonlinearities. Because their iterative penalty is inversely proportional to the coefficient magnitude, the inherently small weakly nonlinear terms receive massive penalties and are mistakenly discarded as noise.

2.4. Mathematical identifiability limitations of weak nonlinearity under noise

While the previous section established that Pareto optimization naturally discards weakly nonlinear terms due to their negligible variance contribution, manual hyperparameter tuning could theoretically recover these subtle effects in a purely deterministic,

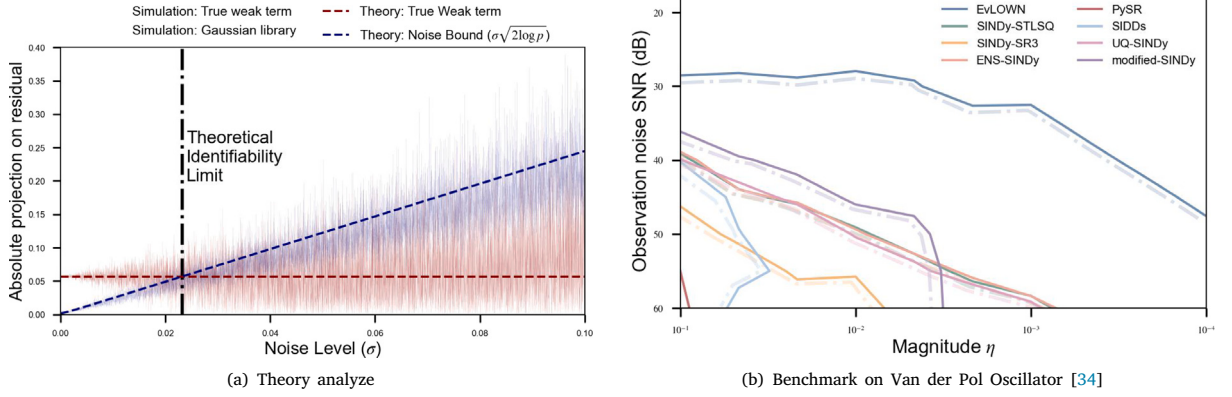


Fig. 3. Mathematical and empirical identifiability limits of weakly nonlinear dynamics under observation noise.

noise-free setting. However, the introduction of observation noise fundamentally alters this paradigm. When the intensity of the weak nonlinearity falls below the noise floor, it becomes practically unidentifiable. This section provides a mathematical derivation to establish the identifiability limits of weak nonlinear dynamics under noise.

Mathematically, unidentifiability arises because observation noise introduces spurious correlations within the candidate library. According to the Karush–Kuhn–Tucker (KKT) conditions for L_1 -regularized regression [46], a coefficient is strictly driven to zero if its projection onto the residual vector is less than the regularization penalty parameter λ . Consider an idealized extreme scenario where the dominant linear terms are perfectly known and subtracted, leaving exactly one true weak nonlinear term alongside a vast library of p candidate features composed entirely of Gaussian white noise. To successfully eliminate these p irrelevant noise terms, the L_1 penalty parameter λ must satisfy a strict theoretical lower bound, classically known as the universal threshold [47]:

$$\lambda \propto \sigma \sqrt{\log p} \quad (5)$$

where σ represents the standard deviation of the observational noise. This establishes a definitive information-theoretic boundary. If the magnitude of the weak nonlinearity satisfies the following condition:

$$\epsilon f_i(\mathbf{x}, \dot{\mathbf{x}}) < \sigma \sqrt{\log p} \quad (6)$$

the physical term is mathematically indistinguishable from ambient noise and will be unavoidably thresholded out (Fig. 3a). In realistic scenarios, candidate library terms are typically not orthogonal white noise but exhibit collinearity, which further exacerbates this masking effect and structurally enforces the elimination of the weak nonlinearity.

These theoretical boundaries are strongly supported by empirical evidence. In previous benchmark studies conducted by the authors utilizing the Van der Pol oscillator [34], the algorithmic success rates of eight existing inference methods were systematically evaluated across varying noise levels and weak nonlinearity intensities. Fig. 3b presents the identifiability boundaries under different noise levels and nonlinearity magnitudes. The boundary is defined as the region in which the success rate of equation identification exceeds 50% over 100 independent runs. The combination of solid and dashed lines indicates the boundary, and the dashed segment denotes the interior of the identifiable region. As Fig. 3b shown, the empirical results aligned perfectly with the derived mathematical limits, confirming that the identifiability of weak nonlinearities is strictly inversely proportional to the observation noise intensity.

2.5. Evolutionary learning oscillator with weak nonlinearity: EvLOWN

Because of the Pareto trap and identifiable limitations, most existing model identification frameworks like SINDy do not target the specific difficulties introduced by weak nonlinearities in practical systems. To address this gap, Ma et al. [34] proposed a data-driven approach called EvLOWN (Evolutionary Learning Oscillator with Weak Nonlinearity) for discovering weakly nonlinear oscillator. EvLOWN reformulates the original system of ODEs into an evolutionary-variable representation via averaging theory, naturally separating terms by their order of magnitude. The goal of EvLOWN is to discover the autonomous oscillatory dynamical system:

$$\ddot{x}_i(t) + \omega_i^2 x_i(t) + \epsilon f_i(\mathbf{x}(t), \dot{\mathbf{x}}(t)) = 0 \quad (7)$$

by learning the evolution of amplitude and phase, rather than directly learning the state variables $\mathbf{x}$.

Leveraging averaging theory [48], also known as the Krylov–Bogoliubov–Mitropolsky method [49], we approximate the response of the oscillator (7) when the nonlinear effect is weak (i.e., ϵ is small) as:

$$x_i(t) = A_i(t) \sin(\hat{\omega}_i t + \phi_i(t)) \quad (8)$$

where: $\hat{\omega}_i$ is the oscillating frequency; A_i and ϕ_i are the amplitude and phase, respectively.

In WNOs, the amplitude $A_i(t)$ and phase $\phi_i(t)$ evolve much more slowly over time compared to the oscillatory term, changing little over a single period $2\pi/\hat{\omega}_i$ [50]. This separation of timescales allows us to average the system dynamics over one oscillation cycle, yielding explicit relationships between the weak nonlinearities and the time evolution of the amplitude $A_i(t)$ and phase $\phi_i(t)$:

$$\begin{aligned} A_i'(t) &= \frac{A_i(t + \pi/\hat{\omega}_i) - A_i(t - \pi/\hat{\omega}_i)}{2\pi/\hat{\omega}_i} = \frac{\epsilon_i}{2\pi} \int_{t-\pi/\hat{\omega}_i}^{t+\pi/\hat{\omega}_i} f_i(\mathbf{x}(s), \dot{\mathbf{x}}(s)) \cos(\hat{\omega}_i s + \phi_i(s)) ds, \\ \phi_i'(t) &= \frac{\phi_i(t + \pi/\hat{\omega}_i) - \phi_i(t - \pi/\hat{\omega}_i)}{2\pi/\hat{\omega}_i} = -\frac{\epsilon_i}{2A_i(t)\pi} \int_{t-\pi/\hat{\omega}_i}^{t+\pi/\hat{\omega}_i} f_i(\mathbf{x}(s), \dot{\mathbf{x}}(s)) \sin(\hat{\omega}_i s + \phi_i(s)) ds, \end{aligned} \quad (9)$$

where $(\cdot)'$ denotes the rate of change over one period.

To identify the first-order ODEs governing the evolution of amplitude and phase, we integrate the basis terms in the original library Θ over a single period, which quantifies their contributions to the averaged dynamics. This transformation results in two new libraries: Θ_A for amplitude evolution and Θ_ϕ for phase evolution. The discovery task now becomes identifying a sparse subset of basis terms from these libraries that best approximate $A_i'(t)$ and $\phi_i'(t)$. Critically, this averaging process naturally separates terms by order of magnitude, allowing weak nonlinear contributions to be isolated from dominant dynamics.

2.6. Contributions of MEv-SINDy

The core difficulty in discovering weakly nonlinear dynamics fundamentally stems from the coexistence of large dominant linear terms and small nonlinear terms within the same differential equation. As demonstrated in Sections 2.3 and 2.4, direct sparse regression on the instantaneous derivatives is structurally flawed for these systems. Because the weak terms contribute negligibly to the instantaneous variance, they do not form a distinguishable corner on the Pareto front and are easily completely masked by the mathematical lower bound of observation noise. Consequently, standard optimization paradigms structurally over-prune these vital physical terms.

To circumvent this bottleneck, the EvLOWN framework (Section 2.5) introduced a paradigm shift: rather than optimizing the differential equation directly, the regression target is reformulated into the slow-varying evolution of the system. By employing the method of averaging, the dominant linear oscillations are mathematically filtered out, projecting the dynamics onto a slow-evolution manifold where the weak nonlinearities become the primary driving forces of the amplitude and phase variations. This successfully rescues the $O(\epsilon)$ terms from the Pareto trap and noise masking. However, EvLOWN relies on a strict mathematical prior: it assumes the system exhibits a quasi-harmonic, single-frequency response. Therefore, it is exclusively applicable to autonomous oscillators. In practical engineering and real-world physics, most systems are non-autonomous; they are subjected to complex external forcing and exhibit multi-frequency responses containing higher-order harmonics, subharmonics, and strong modal coupling. In such scenarios, the classical method of averaging completely fails, rendering single-frequency evolutionary frameworks obsolete.

The primary contribution of the proposed MEv-SINDy (Multi-frequency Evolutionary Sparse Identification of Nonlinear Dynamics) framework is upgrading the slow-evolution identification philosophy from single-frequency autonomous vibrations to multi-frequency non-autonomous dynamics. To achieve this, MEv-SINDy abandons the classical method of averaging and instead embeds the Generalized Harmonic Balance (GHB) method directly into the data-driven optimization objective. By tracking the slow-varying evolution of multiple coupled harmonic coefficients simultaneously, MEv-SINDy successfully untangles complex forced responses. This theoretical breakthrough not only resolves the multi-scale sparsity conflict under noise but also enables the highly accurate one-shot learning of global frequency-response curves (FRCs) from a single transient excitation.

3. Methodology

3.1. Generalized harmonic balance method

As the magnitude of external forcing increases in the forced oscillator Eq. (1), the contribution of nonlinearities in the system, such as quadratic and cubic terms, becomes significant. Consequently, the assumption of a single-frequency solution, see Eq. (10) becomes inadequate for complex forced dynamics. To address this issue, we employ the generalized harmonic balance method to replace the averaging method, thus extending the EvLOWN approach to account for multiple-frequency responses. The solution of the weakly nonlinear forcing oscillator can be approximated as follows [51]:

$$x_i^*(t) = a_i^{(0)}(t) + \sum_{n=1}^N \{b_i^{(n)}(t) \cos(n\hat{\omega}_i t) + c_i^{(n)}(t) \sin(n\hat{\omega}_i t)\} \quad (10)$$

where the superscript $(\cdot)^{(n)}$ indicates the order of the harmonic frequency, corresponding to the n -multiple of the fundamental frequency $\hat{\omega}_i$. The terms $a_i^{(0)}(t)$, $b_i^{(n)}(t)$, and $c_i^{(n)}(t)$ are the slowly varying parameters associated with each harmonic component $n\hat{\omega}_i$.

The first- and second-order time derivatives of $\mathbf{x}^*(t)$ are:

$$\dot{x}_i^*(t) = \dot{a}_i^{(0)}(t) + \sum_{n=1}^N \{[\dot{b}_i^{(n)}(t) + n\hat{\omega}_i c_i^{(n)}(t)] \cos(n\hat{\omega}_i t) + [\dot{c}_i^{(n)}(t) - n\hat{\omega}_i b_i^{(n)}(t)] \sin(n\hat{\omega}_i t)\} \quad (11)$$

$$\begin{aligned}\ddot{x}_i^*(t) = & \ddot{a}_i^{(0)}(t) + \sum_{n=1}^N \{ [\ddot{b}_i^{(n)}(t) + 2n\hat{\omega}_i \dot{c}_i^{(n)}(t) - (n\hat{\omega}_i)^2 b_i^{(n)}(t)] \cos(n\hat{\omega}_i t) \\ & + [\dot{c}_i^{(n)}(t) - 2n\hat{\omega}_i \dot{b}_i^{(n)}(t) - (n\hat{\omega}_i)^2 c_i^{(n)}(t)] \sin(n\hat{\omega}_i t) \}\end{aligned}\quad (12)$$

Supposing that $a_i^{(0)}(t)$, $b_i^{(n)}(t)$, and $c_i^{(n)}(t)$ vary slowly with time, we substitute Eqs. (10)–(12) into Eq. (1), and then we average for each harmonic terms of $\cos(n\hat{\omega}_i t)$ and $\sin(n\hat{\omega}_i t)$ obtaining for $n = 1, 2, \dots, N$:

$$\begin{aligned}\ddot{a}_i^{(0)}(t) + \omega_i^2 a_i^{(0)}(t) = & -\frac{\hat{\omega}_i}{2\pi} \int_{t-\pi/\hat{\omega}_i}^{t+\pi/\hat{\omega}_i} [f_i^*(\mathbf{x}, \dot{\mathbf{x}}, s)] ds \\ \ddot{b}_i^{(n)}(t) + 2\hat{\omega}_i n \dot{c}_i^{(n)}(t) - \hat{\omega}_i^2 (n^2 - 1) b_i^{(n)}(t) = & -\frac{\hat{\omega}_i}{\pi} \int_{t-\pi/\hat{\omega}_i}^{t+\pi/\hat{\omega}_i} [f_i^*(\mathbf{x}, \dot{\mathbf{x}}, s)] \cos(n\hat{\omega}_i s) ds \\ \dot{c}_i^{(n)}(t) - 2\hat{\omega}_i n \dot{b}_i^{(n)}(t) - \hat{\omega}_i^2 (n^2 - 1) c_i^{(n)}(t) = & -\frac{\hat{\omega}_i}{\pi} \int_{t-\pi/\hat{\omega}_i}^{t+\pi/\hat{\omega}_i} [f_i^*(\mathbf{x}, \dot{\mathbf{x}}, s)] \sin(n\hat{\omega}_i s) ds,\end{aligned}\quad (13)$$

where:

$$f_i^*(\mathbf{x}, \dot{\mathbf{x}}, s) = \epsilon f_i(\mathbf{x}^*(s), \dot{\mathbf{x}}^*(s)) - F_i(\mathbf{x}^*(s), \dot{\mathbf{x}}^*(s), s) \quad (14)$$

It is worth stressing that the only assumption is related to the slow temporal variation of $a_i^{(0)}(t)$, $b_i^{(n)}(t)$, and $c_i^{(n)}(t)$, while no limitations on the type of response (steady state or transient) is instead required. In the next section, we will show how to exploit the evolutionary formulation in Eq. (13) to identify the system described, in its original formulation, by Eq. (1).

3.2. Multiple slow-varying evolutions evaluations

We now describe how to extract the slowly varying evolutionary variables $a_i^{(0)}(t)$, $b_i^{(n)}(t)$, and $c_i^{(n)}(t)$ ($n = 1, 2, \dots$) from the observed signal $x(t)$. These variables are subsequently used to perform sparse regression on the evolutionary formulation reported in Eq. (13) in order to identify the unknown components of the governing equations, namely ω_i , f_i^* .

Given an input signal $\mathbf{x}(t) = [x_1(t), \dots, x_d(t)]^T$, the oscillation frequency in the i th dimension can be directly computed from the Fourier transform $F(x_i(t)) : \mathbb{R} \rightarrow \mathbb{R}$, with:

$$\hat{\omega}_i = \arg \max_{\hat{\omega}_i} F(x_i(t)) = \arg \max_{\hat{\omega}_i} \int_{-\infty}^{+\infty} x_i(t) e^{-i\hat{\omega}_i t} dt \quad (15)$$

It is worth noting that the oscillation circular frequency $\hat{\omega}_i$ differs from the natural circular frequency ω_i . This deviation arises from the presence of weakly nonlinear terms and external forcing. To illustrate this, we consider a simple example of a linearly damped oscillator described by $\ddot{x} + 2\beta\dot{x} + \omega^2 x = F \cos(\Omega t)$. Under the assumption of light damping ($\beta \ll 1$), the solution of the linear damped oscillator is given by:

$$x(t) = x_0 e^{-\beta t} \cos(\sqrt{\omega^2 - \beta^2} t + \phi) + \frac{f}{\sqrt{(\omega^2 - \Omega^2)^2 + 4\beta^2 \omega^2}} \cos(\Omega t - \varphi)$$

Therefore, the oscillation circular frequency $\hat{\omega}$ obtained from Eq. (15) lies within the range $(\sqrt{\omega^2 - \beta^2}, \Omega)$, and the value of $\hat{\omega}$ depends on the selected portion of the data $x(t)$. If the transient response dominates, $\hat{\omega}$ will tend to $\sqrt{\omega^2 - \beta^2}$; otherwise, it will be close to Ω .

To obtain the multiple slow-varying evolutionary variables $a_i^{(0)}(t)$, $b_i^{(n)}(t)$, and $c_i^{(n)}(t)$ ($n = 1, 2, \dots$) from the given time history $x_i(t)$, we proposed a novel identification method involving the Hilbert transform. Unlike what was proposed in [34], the Hilbert transform cannot be directly applied to extract the slowly varying amplitude and phase, because the dynamic response of a weakly nonlinear forced oscillator contains multiple frequency components. To address this, we define N frequency bands to address the limitations of standard Fourier analysis. In classical spectral analysis, isolated peaks represent constant coefficients. However, our evolutionary coefficients depend on time. This time-dependency smears the frequency peaks in the spectrum. In particular, we determine the value of N depends on the number of distinct frequency peaks observed in the frequency spectrum. When, as in Fig. 4a, $x(t)$ consists of the zero, first and second harmonic components, we set $N = 3$. Each frequency band to be centered at $n\hat{\omega}$ and has a bandwidth of $\hat{\omega}/4$ (Fig. 4b). By passing the signal through the corresponding bandpass filter, each harmonic component is isolated, as shown in Fig. 4c. In practice, Finite Impulse Response (FIR) filters are designed using the window method. This procedure is implemented in readily available packages such as `scipy.signal`.

After having obtained the slow-varying evolutionary variables $a_i^{(0)}(t)$, $b_i^{(n)}(t)$, and $c_i^{(n)}(t)$ through the Hilbert transform, we can elaborate Eq. (10) as:

$$x_i(t) = \sum_{i=0}^N x_i^{(n)}(t), \quad (16)$$

with:

$$x_i^{(n)}(t) = \begin{cases} a_i^{(0)}(t) & n = 0 \\ b_i^{(n)}(t) \cos(n\hat{\omega}_i t) + c_i^{(n)}(t) \sin(n\hat{\omega}_i t) & n = 1, 2, \dots \end{cases} \quad (17)$$

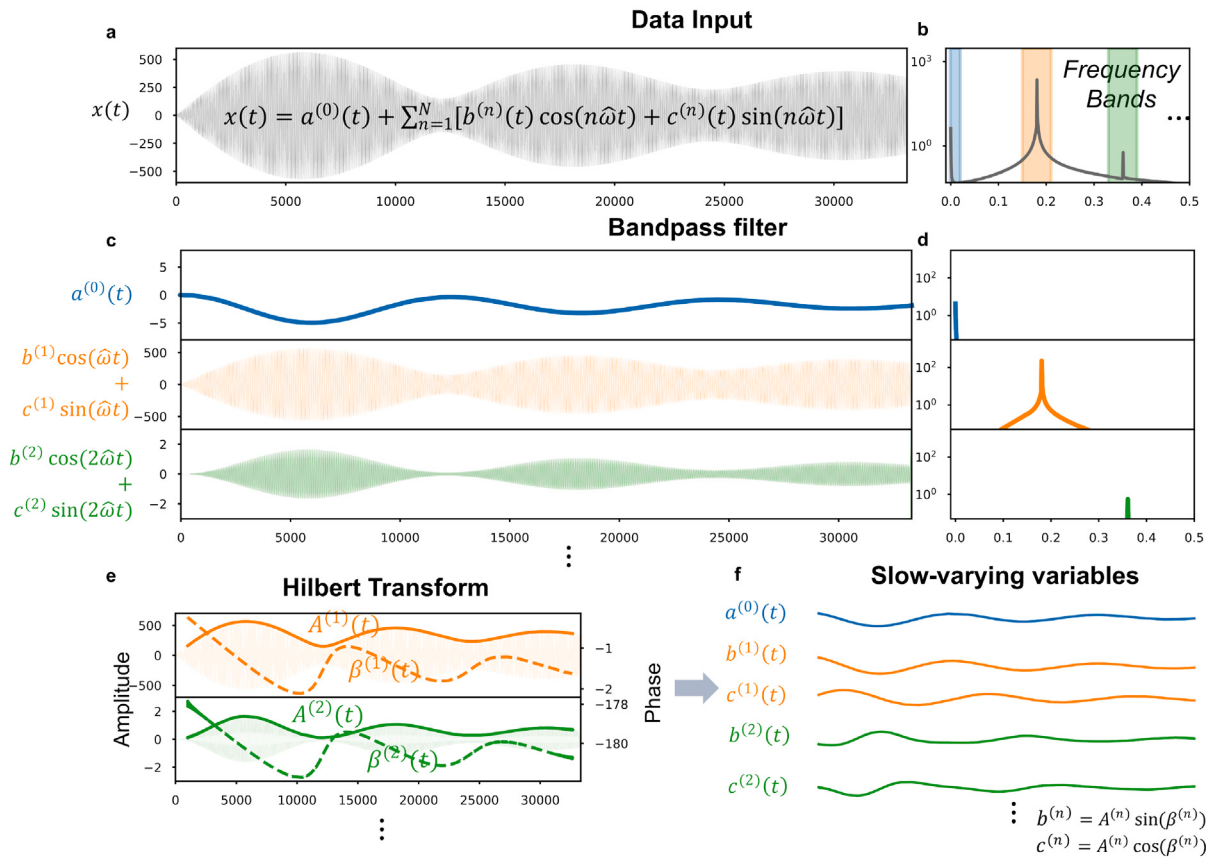


Fig. 4. Identification process of the slowly varying evolutionary variables $a^{(0)}(t)$, $b^{(n)}(t)$, and $c^{(n)}(t)$ by combining a bandpass filter and the Hilbert transform. (a) Input signal $x(t)$, representing the time-domain response of a weakly nonlinear forced oscillator; (b) $x(t)$ consists of multiple narrowband harmonic components centered at $n\hat{\omega}$, as shown in the frequency domain. Several frequency bands are defined to isolate each n th harmonic component; (c) Time-domain representation and (d) corresponding frequency-domain representation of the filtered narrowband signals; (e) Hilbert transform applied to each n th harmonic component to extract the instantaneous amplitude $A^{(n)}(t)$ and phase $\beta^{(n)}(t)$; (f) The slowly varying coefficients $b^{(n)}(t)$ and $c^{(n)}(t)$ are calculated from the extracted amplitude and phase, while the zero-order term directly corresponds to $a^{(0)}(t)$.

The zero-order harmonic component directly corresponds to $a^{(0)}(t)$. While for $n > 0$, the slow-varying evolutionary variables are calculated from amplitude and phase. The Hilbert transform $\mathcal{H}$ is used to compute the analytic signal $\hat{x}_i^{(n)}(t) = x_i^{(n)}(t) + i\mathcal{H}(x_i^{(n)}(t))$. The evolutionary variables $A_i^{(n)}(t)$ and $\beta_i^{(n)}(t)$ of time t can be considered as the averaging value of instantaneous evolutionary variables in one period:

$$A_i^{(n)}(t) = \frac{\int_{t-\pi/\hat{\omega}_i}^{t+\pi/\hat{\omega}_i} |\hat{x}_i^{(n)}(s)| ds}{2\pi/\hat{\omega}_i}$$

$$\beta_i^{(n)}(t) = \frac{\int_{t-\pi/\hat{\omega}_i}^{t+\pi/\hat{\omega}_i} \text{Arg}(\hat{x}_i^{(n)}(s)) ds}{2\pi/\hat{\omega}_i}$$
(18)

As Fig. 4f shows, the evolutionary variable $b_i^{(n)}(t)$ and $c_i^{(n)}(t)$ of higher harmonic components can be calculated as Eq. (19):

$$b_i^{(n)}(t) = A_i^{(n)}(t) \sin(\beta_i^{(n)}(t))$$

$$c_i^{(n)}(t) = A_i^{(n)}(t) \cos(\beta_i^{(n)}(t))$$
(19)

3.3. Reformulation of the regression problem in the evolutionary domain

In order to accurately identify the governing dynamics of weakly nonlinear forced oscillators, it is essential to reformulate the regression problem in a form that enhances the visibility of nonlinear effects.

Sharing the same starting point of EvLOWN, we transform the problem into an evolutionary domain. In the proposed MEv-SINDy approach, the system behavior is represented by slowly varying evolutionary variables based on generalized harmonic balance method. This reformulation allows the nonlinear interactions to be more effectively captured and facilitates the application of sparse regression for discovering the underlying equations. We thus exploit SINDy to approximate the evolutionary formulation reported in Eq. (13), instead of the original formulation (Eq. (1)). This is equivalent to perform regularized regression reported in Eq. (4), but here involving the evolutionary instead of the original variables to identify the weakly nonlinear coefficients.

According to Section 3.2, the slow-varying evolutionary variables $a_i^{(0)}(t)$, $b_i^{(n)}(t)$ and $c_i^{(n)}(t)$ ($n = 1, 2, \dots$), together with the oscillation frequency $\hat{\omega}_i$, can be computed from the given time history $x_i(t)$. Eq. (13) can then be reorganized such that known quantities are collected in the left hand side, while in the terms to be inferred are assembled in the right hand side.

$$\begin{aligned} \ddot{a}_i^{(0)}(t) + \hat{\omega}_i^2 a_0(t) &= -\frac{\hat{\omega}_i}{2\pi} \int_{t-\pi/\hat{\omega}_i}^{t+\pi/\hat{\omega}_i} \hat{f}_i^*(\mathbf{x}, \mathbf{x}; s) ds \\ \ddot{b}_i^{(n)}(t) + 2\hat{\omega}_i n \dot{c}_i^{(n)}(t) - \hat{\omega}_i^2 (n^2 - 1) b_i^{(n)}(t) &= -\frac{\hat{\omega}_i}{\pi} \int_{t-\pi/\hat{\omega}_i}^{t+\pi/\hat{\omega}_i} \hat{f}_i^*(\mathbf{x}, \mathbf{x}; s) \cos(n\hat{\omega}_i s) ds \\ \ddot{c}_i^{(n)}(t) - 2\hat{\omega}_i n \dot{b}_i^{(n)}(t) - \hat{\omega}_i^2 (n^2 - 1) c_i^{(n)}(t) &= -\frac{\hat{\omega}_i}{\pi} \int_{t-\pi/\hat{\omega}_i}^{t+\pi/\hat{\omega}_i} \hat{f}_i^*(\mathbf{x}, \mathbf{x}; s) \sin(n\hat{\omega}_i s) ds \\ \hat{f}_i^*(\mathbf{x}, \dot{\mathbf{x}}, s) &= (\omega_i^2 - \hat{\omega}_i^2) x_i^*(s) + \epsilon f_i(\mathbf{x}^*(s), \dot{\mathbf{x}}^*(s)) - F_i(\mathbf{x}^*(s), \dot{\mathbf{x}}^*(s), s) \\ n &= 1, 2, \dots, N \end{aligned} \quad (20)$$

where $\hat{f}_i^*(\mathbf{x}, \dot{\mathbf{x}}, t) = \boldsymbol{\Theta}^*(\mathbf{x}, \dot{\mathbf{x}}, t) \boldsymbol{\Xi}^*$ are unknown functions expressed as a sparse combination of a set of linear and nonlinear candidate basis function. Thanks to this rearrangement, the system is expressed in a standard regression form suitable for sparse identification. Unlike conventional sparse regression, our objective is to find a sparse coefficient vector $\boldsymbol{\Xi}^*$ that simultaneously satisfies all equations in Eq. (20), rather than a single one. This gives rise to the concept of joint sparse regression, where the goal is to estimate a shared sparse representation across multiple regression targets. Mathematically, given $M = 2N + 1$ equations reported in the following:

$$\mathbf{y}_i(t) = \boldsymbol{\Theta}_i(t) \boldsymbol{\Xi}^*, \quad (21)$$

$$\mathbf{y}_i(t) = \begin{bmatrix} y_i^{(0)}(t) \\ y_i^{(1,\cos)}(t) \\ y_i^{(1,\sin)}(t) \\ y_i^{(2,\cos)}(t) \\ y_i^{(2,\sin)}(t) \\ \vdots \end{bmatrix} = \begin{bmatrix} \ddot{a}_i^{(0)}(t) + \hat{\omega}_i^2 a_0(t) \\ \ddot{b}_i^{(1)}(t) + 2\hat{\omega}_i \dot{c}_i^{(1)}(t) \\ \ddot{c}_i^{(1)}(t) - 2\hat{\omega}_i \dot{b}_i^{(1)}(t) \\ \ddot{b}_i^{(2)}(t) + 4\hat{\omega}_i \dot{c}_i^{(2)}(t) - 3\hat{\omega}_i^2 b_i^{(2)}(t) \\ \ddot{c}_i^{(2)}(t) - 4\hat{\omega}_i \dot{b}_i^{(2)}(t) - 3\hat{\omega}_i^2 c_i^{(2)}(t) \\ \vdots \end{bmatrix}, \quad (22)$$

$$\boldsymbol{\Theta}_i(t) = \begin{bmatrix} \boldsymbol{\Theta}_i^{(0)}(t) \\ \boldsymbol{\Theta}_i^{(1,\cos)}(t) \\ \boldsymbol{\Theta}_i^{(1,\sin)}(t) \\ \boldsymbol{\Theta}_i^{(2,\cos)}(t) \\ \boldsymbol{\Theta}_i^{(2,\sin)}(t) \\ \vdots \end{bmatrix} = \begin{bmatrix} -\frac{\hat{\omega}_i}{2\pi} \int_{t-\pi/\hat{\omega}_i}^{t+\pi/\hat{\omega}_i} \boldsymbol{\Theta}^*(s) ds \\ -\frac{\hat{\omega}_i}{\pi} \int_{t-\pi/\hat{\omega}_i}^{t+\pi/\hat{\omega}_i} \boldsymbol{\Theta}^*(s) \cos(\hat{\omega}_i s) ds \\ -\frac{\hat{\omega}_i}{\pi} \int_{t-\pi/\hat{\omega}_i}^{t+\pi/\hat{\omega}_i} \boldsymbol{\Theta}^*(s) \sin(\hat{\omega}_i s) ds \\ -\frac{\hat{\omega}_i}{\pi} \int_{t-\pi/\hat{\omega}_i}^{t+\pi/\hat{\omega}_i} \boldsymbol{\Theta}^*(s) \cos(2\hat{\omega}_i s) ds \\ -\frac{\hat{\omega}_i}{\pi} \int_{t-\pi/\hat{\omega}_i}^{t+\pi/\hat{\omega}_i} \boldsymbol{\Theta}^*(s) \sin(2\hat{\omega}_i s) ds \\ \vdots \end{bmatrix}, \quad (23)$$

the objective is to determine a common coefficient vector $\boldsymbol{\Xi}^*$ that minimizes the overall residual while maintaining sparsity. Only when this condition is met, the reconstructed system can accurately recover the original formulation expressed in Eq. (1). In Eqs. (22) and (37), the superscript $(\cdot)^{(n,t)}$ identifies the harmonic term in Eq. (20). Specifically, $n \in \mathbb{R}^+$ represents the order of the harmonic frequency, corresponding to the n -fold multiple of the fundamental frequency $\hat{\omega}_i$, and $\{\cos, \sin\}$ indicates the cosine and sine components.

3.4. Joint sparse regression across multiple equations

To infer the sparse coefficient vector $\boldsymbol{\Xi}^*$ that best satisfies all evolutionary equations, we propose a joint sparse regression method across multiple equations all evolutionary equations Eq. (21), rather than fitting each equation independently. The problem is formulated as minimizing the total residual across the $2N + 1$ evolutionary equations derived from the weakly nonlinear oscillator, each corresponding to a harmonic component over $N + 1$ frequencies, with an L_1 regularization term to ensure sparsity.

A key algorithmic choice is setting which harmonics are considered in the evolutionary representation among $\mathbf{y}_{full}(t) = [y^{(0)}(t), y^{(1,\cos)}(t), y^{(1,\sin)}(t), y^{(2,\cos)}(t), y^{(2,\sin)}(t), \dots]$. Once set, goal of the optimization is to identify a sparse coefficient vector $\boldsymbol{\Xi}^*$ that simultaneously satisfies all evolutionary equations derived from weakly nonlinear forcing oscillator as shown in the following:

$$\boldsymbol{\Xi}^* = \arg \min_{\boldsymbol{\Xi}^*} \left(\sum_{m \in I_S} \|\mathbf{y}^{(m)} - \boldsymbol{\Theta}^{(m)} \boldsymbol{\Xi}^*\|_2^2 + \lambda \|\boldsymbol{\Xi}^*\|_1 \right) \quad (24)$$

where: $\mathcal{I}_S$ collects all the index of the effective harmonic components (e.g., 0, (1, cos), (1, sin), ...); $\mathbf{y}^{(m)}$ and $\boldsymbol{\Theta}^{(m)}$ denote the response and library matrices for the m th equation; $\lambda > 0$ controls the sparsity level of the solution. We introduce a two-phase sparse regression method to solve Eq. (29). The first phase utilizes a convex optimization algorithm to select an initial subset of the library. In this work, we employ the CVXOPT solver [52] to handle the L_1 regularization term. This step provides a preliminary sparse structure for the coefficient vector. The second phase involves a sparse fine-tuning process that removes terms with negligible contributions, thus ensuring a concise and generalized model. We apply an approach inspired by orthogonal matching pursuit to narrow down the model space. We select only the most relevant basis functions from the initial subset. The contribution c_{ij} of each basis function j for the i th dimension is calculated using Eq. (25):

$$c_{ij} = \int_0^T \left(\sum_{m \in \mathcal{I}_S} \frac{(\hat{\Theta}_{mij}(t) \Xi_{ij}^*)^2}{(y_{im}(t))^2} \right) dt \quad (25)$$

These values are then normalized to a range between 0 and 1. Finally, we fine-tune the model by applying a cut-off threshold. Any basis function with a contribution value lower than this threshold is removed from the library. This two-phase strategy ensures the identification of a parsimonious dynamical system. It maintains high predictive accuracy while preventing overfitting to numerical noise.

3.5. Parameter continuation method

In nonlinear dynamical systems, periodic solutions often depend on one or more system parameters, such as the excitation frequency and intensity and can be computed using several classical numerical techniques, such as the harmonic balance method [53], the shooting method [54], and collocation-based approaches [55]. These methods transform the governing differential equations into nonlinear algebraic systems by enforcing periodicity conditions, and the steady-state responses are then obtained through iterative solvers such as the Newton–Raphson method [56]. However, these traditional approaches are often computationally expensive because they require repeated integrations or nonlinear solves for each excitation parameter. In addition, they usually depend on initial guesses that are close to stable solutions, which makes them incapable of accurately tracing unstable branches or capturing turning points in the frequency-response curves. As a result, their applicability is limited when a complete description of the system dynamics, including both stable and unstable periodic responses, is required.

The continuation method provides a systematic procedure to trace the evolution of these solutions as parameters vary, allowing for the identification of critical phenomena such as bifurcations or stability transitions. Consider a general system of nonlinear algebraic equations derived from the steady or periodic conditions of a dynamical system:

$$\mathbf{G}(\mathbf{u}, \zeta) = 0 \quad (26)$$

where $\mathbf{u}$ denotes the unknown state vector (e.g., displacement and velocity), and ζ represents the control parameter to be varied (e.g., frequency and amplitude of external forcing).

A straightforward approach would be to increment ζ in small steps and solve for $\mathbf{u}$ at each step using the previous solution as the initial guess $\zeta_{k+1} = \zeta_k + \Delta\zeta$. However, this simple “parameter sweep” method fails e.g. near turning points or folds in the solution curve, where the slope $\partial\mathbf{u}/\partial\zeta$ becomes infinite and the solution path cannot be followed uniquely.

To overcome this limitation, arc-length continuation (also known as pseudo-arc-length continuation) is commonly employed [57]. In this method, the solution path is parameterized by an artificial variables, representing the arc length along the curve in the $(\mathbf{u}, \zeta)$ space. The continuation step is then constrained by:

$$(\Delta\mathbf{u})^T (\Delta\mathbf{u}) + (\Delta\zeta)^2 = (\Delta s)^2 \quad (27)$$

where Δs is the prescribed step size along the solution trajectory.

By solving the extended system composed of Eq. (26) together with the arc-length constraint, the continuation algorithm can robustly track both stable and unstable branches of the solution family, even across fold and bifurcation points. This technique provides a powerful and general framework for the numerical exploration of nonlinear response characteristics. These methods are implemented in many ready-to-use packages like Auto07p [58], that implements collocation methods in FORTRAN to perform numerical continuation and bifurcation analysis; Manlab, a Matlab tool that uses HB methods and Asymptotic Numerical Method [59]; Nlvib, that also exploits HB methods [60], and many others among which we mention COCO [61]. The numerical examples addressed in this work will be solved using the Matlab MATCONT package [62], that exploits collocation methods to perform the continuation of periodic orbits.

3.6. Mean chamfer distance of frequency response curves (MCDRC)

Frequency response curves (FRCs) are the primary goals for one-shot learning and prediction, describing complex dynamic behaviors of weakly nonlinear oscillators. FRCs characterize a steady-state output quantity of interest, such as maximum deflection, across various load multipliers β for all excitation frequency values Ω within a specified range. To quantitatively assess the predictive ability, an error metric named Mean Chamfer Distance of frequency response curves (MCDRC) between two FRCs is defined herein.

It is worth mentioning that FRCs are generally not single-valued functions of the excitation frequency. For a given frequency, multiple periodic solutions may exist due to nonlinear effects such as jump phenomena, hysteresis, and multi-stability. Furthermore,

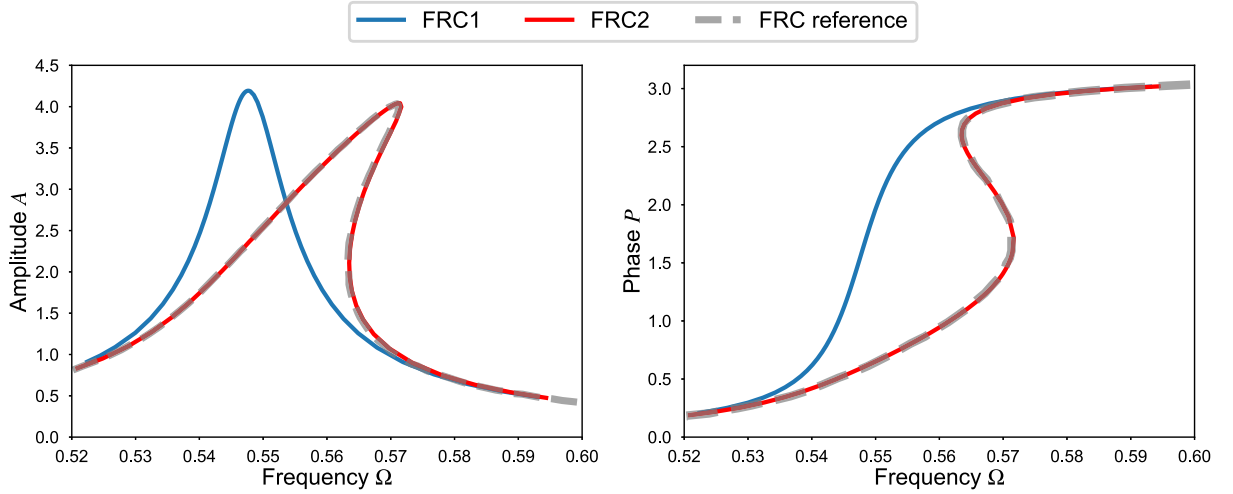


Fig. 5. Two examples of mean chamfer distance of frequency response functions (MCDRC): FRC1 (blue solid line), whose MCDRC is 0.24, exhibits a noticeable deviation from the reference (gray dashed line); FRC2 (red solid line) has a MCDRC of 0.01 shows a much better correlation with the reference. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

continuation-based computations generally produce differently spaced samples along solution branches, and distinct response branches may be specified using varying frequency grids. These properties violate the assumptions required by most pointwise or function-based error measures.

To address this limitation, we represent FRCs as point sets in the frequency-response space. This allows to naturally support multi-valued responses, discontinuities, and nonuniform sampling, hence offering a consistent and stable foundation for quantitative comparison of nonlinear frequency response curves.

We then exploit the Chamfer distance $d_{ch}(C \rightarrow C_{ref})$ [63] to establish a novel metric termed Mean Chamfer Distance of Frequency Response Curves (MCDRC). The chamfer distance $d_{ch}(C \rightarrow C_{ref})$ is defined as the average distance from each point on the predicted curve $C = \{(x_i, y_i)\}_{i=1}^N$ to its nearest neighbor on the reference curve $C_{ref} = \{(x_j, y_j)\}_{j=1}^M$. It can be computed as:

$$d_{ch}(C \rightarrow C_{ref}) = \frac{1}{N} \sum_{p \in C} \min_{q \in C_{ref}} \|p - q\| \quad (28)$$

By selecting the oscillation amplitude (A) and phase (P) as primary steady-state descriptors, the FRC is discretized into two point sets: $C^A = \{(\Omega_i, A_i)\}_{i=1}^N$ and $C^P = \{(\Omega_i, P_i)\}_{i=1}^N$. Here, A and P denote the amplitude and the phase lag relative to the external excitation, respectively. The MCDRC is defined as follows:

$$\text{MCDRC} = d_{ch}(C^A \rightarrow C_{ref}^A) + d_{ch}(C^P \rightarrow C_{ref}^P) \quad (29)$$

Prior to calculating the distance between points, a max–min scaling normalization pre-processing step is introduced to ensure that the scaling disparity between the axes does not spoil the representativeness of MCDRC. Fig. 5 presents an illustrative example of MCDRC comparing two predicted FRCs against a reference FRC. The MCDRC for FRC1 is 0.24, indicating poor predictive accuracy, whereas the value for FRC2 is 0.01, representing excellent agreement with the reference. This comparison provides an intuitive demonstration of how MCDRC quantifies the similarity between FRCs.

4. Numerical validations

To demonstrate the effectiveness of the proposed framework, we first consider a benchmark one-dimensional (1D) forced oscillator characterized by linear dissipation, weak quadratic and cubic nonlinearities. This model is widely used in nonlinear dynamics to represent various physical systems [64]. The objective is to evaluate the framework capability in formulating the governing equations and to numerically verify the equivalence between the original formulation (Eq. (1)) and the generalized evolutionary formulation (Eq. (13)). The equation of motion is expressed as:

$$\ddot{x} + \omega^2 x + c\dot{x} + \alpha_1 x^2 + \alpha_2 x^3 = \beta \cos(\Omega t) \quad (30)$$

where x , $\dot{x}$ and $\ddot{x}$ denote the displacement, velocity and acceleration of the oscillator, respectively; ω is the natural frequency, $\alpha_1, \alpha_2 \in \mathbb{R}$ are parameters characterizing the weak nonlinearities, $c \in \mathbb{R}$ is the linear damping coefficient, β and Ω are the amplitude and frequency of the external forcing. To generate the training trajectory, the oscillator described by Eq. (30) is numerically integrated using the following model parameters: $c = 1 \times 10^{-2}$, $\alpha_1 = 1 \times 10^{-2}$, $\alpha_2 = 1 \times 10^{-4}$, $\omega = 2$, $\Omega = 1.999$, $\beta = 0.5$, with initial states $x_0 = [0, 0]$. Indeed, the magnitudes of α_1, α_2 and c are orders of magnitude smaller than ω .

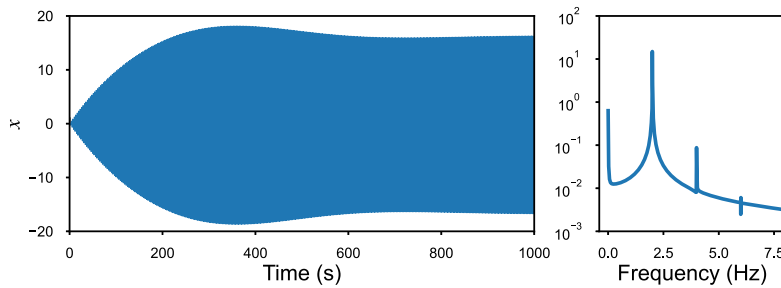


Fig. 6. Benchmark 1D forced oscillator. Training data of numerical quadratic and cubic nonlinear oscillator in time and frequency domain.

The resulting time-domain and frequency-domain responses are illustrated in Fig. 6. The frequency spectrum clearly indicates that the dominant dynamical response occurs near the fundamental frequency of approximately 2. Additionally, two secondary peaks are observed at the frequencies 0 and 4, which correspond to the DC component and second-order harmonic, respectively. These multi-frequency features underline the necessity of a higher-order analytical treatment. It is worth mentioning that the previously proposed EvLOWN approach [34] is unable to handle the presence of higher-order harmonics.

Through this numerical example, we aim to demonstrate the efficacy of the proposed method regarding three key aspects:

- First, to justify the necessity of incorporating multi-frequency responses in forced nonlinear oscillators;
- Second, to validate the evaluation framework for multiple slowly varying evolutionary variables and the generalized harmonic balance (GHB) approach;
- Third, to showcase the effectiveness of the method in both governing equation inference and prediction of complex nonlinear dynamical behaviors.

4.1. Multi-frequency response under external forcing

The emergence of multi-frequency response stems from the system nonlinearities, specifically where the nonlinear restoring force exhibits a nonlinear dependence on the displacement x . As the amplitude of displacement increases, the contributions of higher-order nonlinear terms may grow order of times faster (e.g. x^2 and x^3), leading to the emergence of pronounced harmonic components in the response spectrum. For forced system, the external excitation may drive the system into a regime where higher-order nonlinearities dominate, resulting in a rich multi-frequency spectrum. Conversely, in the absence of external forcing, a free oscillator typically exhibits diminished vibration amplitudes, effectively oscillating at a single frequency (i.e., its natural frequency), as other harmonic components are negligible.

To quantitatively illustrate this phenomenon, we perform a numerical test using the nonlinear oscillator ruled by Eq. (30). All parameters remain constant except for the amplitude of the external forcing β , which varies from 0 to 1. The additional harmonic components, specifically the DC offset (zero-order) and second-order harmonic responses, are evaluated under different values of β . The initial conditions are adjusted to $x_0 = [1, 0]$. As shown in Fig. 7, when β approaches zero, the magnitudes of these extra harmonic components vanish, indicating a nearly quasi-linear single-frequency response. However, as β increases, these higher-order harmonics exhibit a steady growth. This trend confirms that intensified external excitation amplifies nonlinear effects, thereby underscoring the necessity of incorporating multi-frequency responses when analyzing forced nonlinear oscillators.

4.2. Numerical validation of generalized harmonic balance

We evaluate the theoretical foundation of the proposed method, specifically focusing on the validity of slow-varying evolutionary variables and the generalized harmonic balance formulation. The fundamental assumption is that the response $x(t)$ of a weakly nonlinear forced oscillator can be accurately represented as a superposition of multiple harmonic components. Fig. 8(a) presents a comparison between the ground-truth response $x(t)$, obtained via numerical integration, and the approximated response $x^*(t)$ reconstructed using Eq. (10). Based on the spectral characteristics identified in Fig. 6, only the zero-order and first-order harmonic components are selected for the oscillator in Eq. (30). We neglect the second-order harmonic in this specific case. It is important to note that including insignificant harmonic components can degrade the inference results. The slowly varying evolutionary variables $a^{(0)}(t)$, $b^{(1)}(t)$, and $c^{(1)}(t)$ are extracted following the procedure described in Section 3.2. As illustrated in the zoomed-in view in Fig. 8(b), the reconstructed signal $x^*(t)$ exhibits excellent agreement with the reference response $x(t)$, thereby confirming the robustness of the proposed approximation.

Furthermore, the validation of the GHB-based reformulation (Eq. (20)) is presented in Fig. 8(c), comparing the terms derived from the evolutionary variables, with those computed by numerically integrating the weak nonlinearities $f(x, \dot{x})$ over a single period. The close match verifies the accuracy of the GHB method and establishes the equivalence between the original physical formulation (Eq. (1)) and the transformed evolutionary representation (Eq. (13)).

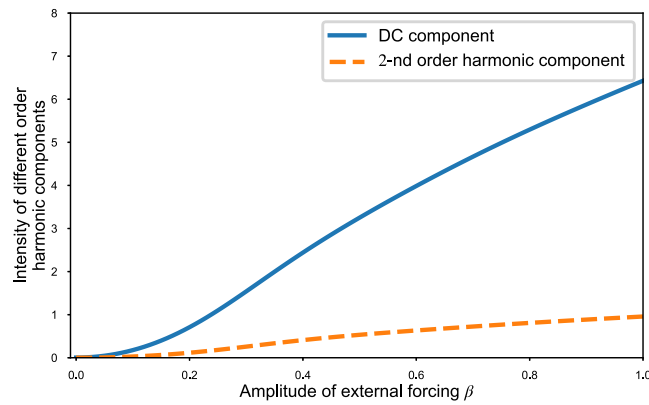


Fig. 7. Benchmark 1D forced oscillator. Influence of external forcing amplitude β on the intensities of the DC offset (zero-order) and second-order harmonic components.

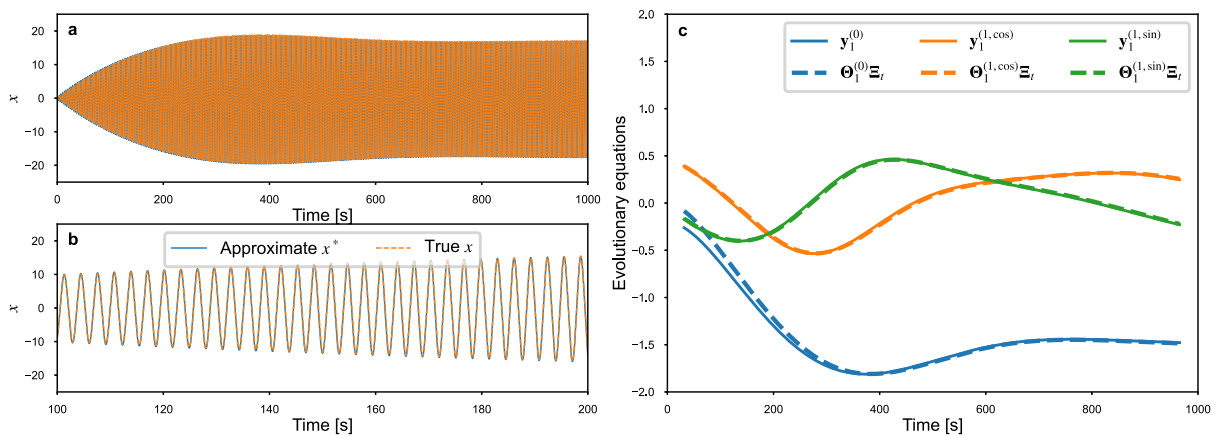


Fig. 8. Benchmark 1D forced oscillator. Numerical validation of the Generalized Harmonic Balance (GHB) method: (a) Comparison between the ground-truth displacement x (numerical integration) and the reconstructed x^* obtained via harmonic superposition, see Eq. (10); (b) zoomed-in view of the transient response highlighting the reconstruction accuracy; (c) validation of the evolutionary terms, comparing the analytical transformations with their numerical counterparts for various harmonic components.

4.3. Inferred equations

The proposed method is applied to the benchmark 1D oscillator to discover the governing equations and reconstruct its complete dynamic portrait, including frequency response curves (FRCs). To capture potential nonlinear behaviors, a candidate library Θ is constructed, consisting of ten pure polynomial terms up to the fifth order and trigonometric functions ($\cos(\Omega t)$ and $\sin(\Omega t)$). Notably, cross-coupling terms between displacement and velocity are omitted to maintain model parsimony. The algorithm is implemented with the following hyper-parameters: the effective harmonic orders are selected as $[0, 1]$, the residual tolerance is set to 1×10^{-1} , and the lowest contribution ratio is 5×10^{-2} . A significant advantage of the proposed approach is its data efficiency: the exact equation structure is successfully identified using data from only a single excitation case (indicated by the red point in Fig. 9). The estimated coefficients, reported in Table 1, exhibit high precision relative to the reference values.

The proposed method successfully identifies the correct equation structure from the library by only one excitation case (red Point in Fig. 9). Estimated coefficients are reported in Table 1. These results demonstrate that the proposed approach achieves accurate and parsimonious discovery of the weakly nonlinear dynamics.

After the governing equation is identified, the system frequency response function (FRF) can be computed either by time integration or by a parameter continuation method. The FRF here describes the stationary oscillation amplitude of the state x as a function of the excitation frequency Ω . As shown in Fig. 9, the predicted FRF closely matches the ground-truth results for various excitation frequencies and amplitudes. It is worth emphasizing, although the governing equation is inferred using only a single excitation condition, the entire frequency-response space under arbitrary forcing conditions can subsequently be predicted with high accuracy.

Table 1

Benchmark 1D forced oscillator. The “Reference” columns report the parameters used for the training; in “MEv-SINDy” columns, we list the parameter values identified by our method.

Coefficients	Relevant term	Reference	MEv-SINDy
ω^2	x	4.0000	4.0012
c	$\dot{x}$	1.0000×10^{-2}	1.0495×10^{-2}
α_1	$\dot{x}^2$	1.0000×10^{-2}	1.0071×10^{-2}
α_2	x^3	1.0000×10^{-4}	9.6691×10^{-5}
β	$\cos(\Omega t)$	5.0000×10^{-1}	5.0748×10^{-1}

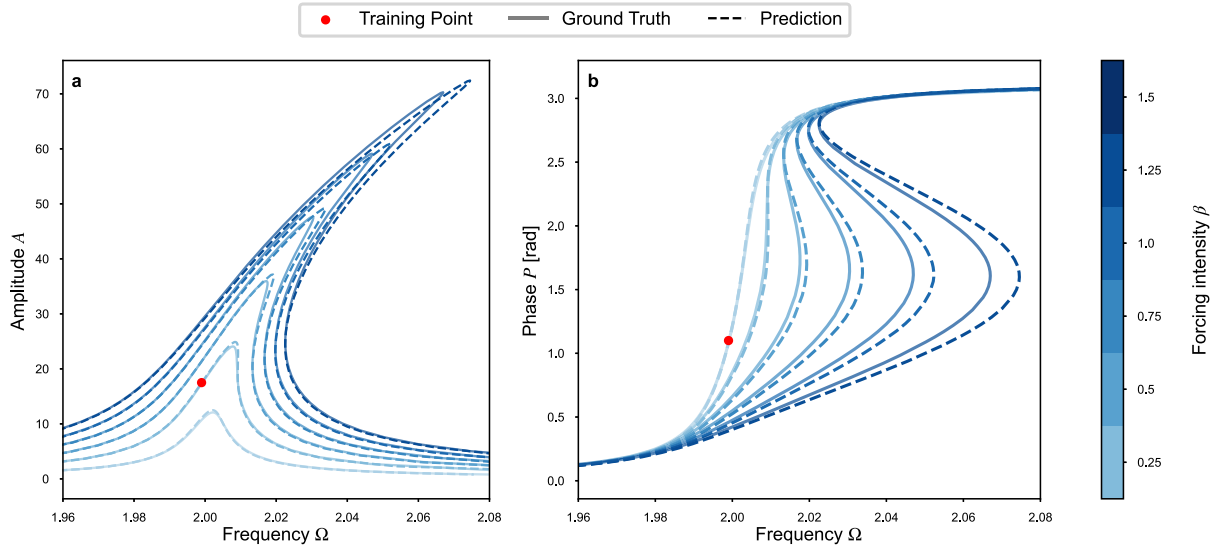


Fig. 9. Benchmark 1D forced oscillator. Comparison between predicted and reference frequency response functions under different excitation amplitudes. The proposed method accurately captures the hardening dynamical performance of oscillator.

5. Applications

To highlight the capabilities of the proposed approach, we investigate high-dimensional structural dynamical systems exhibiting large-amplitude vibrations and geometric nonlinearities. Micro-Electro-Mechanical Systems (MEMS) are typically actuated near their resonance frequencies and consequently undergo substantial deformations. These effects are further intensified by the fact that MEMS are monolithic devices often encapsulated in near-vacuum environments, where energy dissipation is minimal. As a result, the system dynamics are strongly nonlinear and characterized by complex coupled behaviors, showing highly nonlinear dynamical features that are rarely observed at the macro scale. Therefore, MEMS provide a representative and valuable platform for applying the proposed approach to evaluate its capability in predicting dynamical performance through the inferred governing equations.

In this section, we demonstrate the application of the proposed method to high-dimensional structural dynamical systems using synthetic data. The discussion begins with the formulation of the full-order model (FOM) based on the finite element discretization of the governing equations. Subsequently, the system is reduced through the Proper Orthogonal Decomposition (POD) technique to obtain a reduced-order representation that retains the dominant dynamical features. To illustrate the procedure and evaluate the performance of the method, two case studies are presented: a clamped–clamped beam, and a more complex MEMS micromirror device that demonstrates the scalability of the proposed approach.

5.1. Formulation of full-order model

The framework of structures subjected to large displacements and small strains is considered for full-order model simulations [65]. This is the operating range of most microsystems since they are often actuated at resonance and large aspect ratios allow reaching displacements within the linear elastic range of the material. In this framework, the Saint Venant–Kirchhoff constitutive model is the most appropriate choice, and is given by:

$$\mathbf{S} = \mathcal{A} : \mathbf{E}, \quad \mathbf{E} = \frac{1}{2}(\nabla \mathbf{d} + \nabla^T \mathbf{d} + \nabla^T \mathbf{d} \cdot \nabla \mathbf{d}) \quad (31)$$

where $\mathbf{S}$ is the second Piola–Kirchhoff strain tensor, $\mathcal{A}$ is the fourth-order elasticity tensor and $\mathbf{E}$ is the Green–Lagrangian Strain tensor. Here we denote by $\mathbf{d}$ the displacement field and by $\nabla(\cdot)$ the (material) gradient defined with respect to the reference

configuration. The weak form of the linear momentum conservation law is:

$$\int_{\Omega_0} \rho_0 \ddot{\mathbf{d}} \cdot \mathbf{w} d\Omega_0 + \int_{\Omega_0} \mathbf{P}[\mathbf{d}] : \nabla^T \mathbf{w} d\Omega_0 = \int_{\Omega_0} \rho_0 \mathbf{F} \cdot \mathbf{w} d\Omega_0 + \int_{S_{T0}} \mathbf{f} \cdot \mathbf{w} dS_0, \quad \forall \mathbf{w} \in H_0^1(\Omega_0) \quad (32)$$

where the integrals are expressed in the reference configuration Ω_0 . Here, ρ_0 denotes the material density, $\mathbf{P}[\mathbf{d}] = (1 + \nabla \mathbf{d}) \cdot \mathbf{S}$ is the first Piola–Kirchhoff stress tensor, $\mathbf{F}$ is the body force per unit mass, $\mathbf{f}$ is the surface tractions prescribed on the surface S_{T0} and $\mathbf{w}$ is the test velocity selected in $H_0^1(\Omega_0)$, i.e. the space of functions with finite energy that vanish on the portion $S_U \subset \partial\Omega_0$ where Dirichlet boundary conditions are prescribed.

By applying spatial discretization to Eq. (32), for instance through the finite element method, and incorporating a Rayleigh damping model, the system can be transformed into a set of coupled nonlinear ordinary differential equations of the following form:

$$\mathbf{M}\ddot{\mathbf{D}} + \mathbf{C}\dot{\mathbf{D}} + \mathbf{K}\mathbf{D} + \mathbf{G}(\mathbf{D}, \mathbf{D}) + \mathbf{H}(\mathbf{D}, \mathbf{D}, \mathbf{D}) = \mathbf{B}(\mathbf{D}, \beta, \Omega, t), \quad t \in (0, T) \quad (33)$$

where the vector $\mathbf{D} \in \mathbb{R}^k$ collects all the unknown displacement nodal values, $\mathbf{M} \in \mathbb{R}^{n \times n}$ is the mass matrix, $\mathbf{C} = (\omega_0/Q)\mathbf{M}$ is the Rayleigh model mass proportional damping matrix, defined with respect to a reference eigenfrequency ω_0 and a quality factor Q . $\mathbf{B} \in \mathbb{R}^n$ is the nodal force vector which depends on the actuation intensity parameters β , the angular frequency of actuation Ω and in general also on $\mathbf{D}$. The internal force vector has been exactly decomposed in linear, quadratic, and cubic power terms of the displacement: $\mathbf{K} \in \mathbb{R}^{n \times n}$ is the linear stiffness matrix, while $\mathbf{G} \in \mathbb{R}^n$ and $\mathbf{H} \in \mathbb{R}^n$ are vectors given by monomials of second and third order, respectively. The components of these vectors can be expressed using an indicial notation: $G_i = \sum_{j,k=1}^n G_{ijk} D_j D_k$, $H_i = \sum_{j,k,l=1}^n H_{ijkl} D_j D_k D_l$, $i = 1, \dots, n$.

Eq. (33) represents the high-fidelity full-order model (FOM) which depends on the input parameters β and Ω . It should be recalled that in resonating MEMS FRCs are an important design tool in which a selected quantity, like the midspan deflection of a beam or the rotation of a micromirror, is plotted versus Ω for different β . Indeed, resonators operate close to a reference frequency where the behavior should be predictable. However, approximating and reconstructing different FRCs for high dimensional dynamical system is computationally demanding, making the use of full-order models for this task largely impractical.

5.2. Proper orthogonal decomposition

To substantially accelerate the computationally expensive simulations and obtain a more interpretable and parsimonious set of governing equations, we propose a reduction technique that simultaneously decreases the problem dimensionality and introduces a new set of coordinates better adapted to capture the system dynamics within only a few dominant modes. The proposed reduction technique is based on Proper Orthogonal Decomposition (POD), a model reduction method that constructs an ordered set of basis functions according to their energy content. By projecting the full-order data onto a limited number of dominant high-energy modes, a reduced-order representation of the system can be efficiently obtained [66]. The snapshots of the FOM solutions are collected to generate a matrix $\mathbf{X} \in \mathbb{R}^{p \times k}$, whose p denotes the time stamps of solutions.

Next, the Singular Value Decomposition (SVD) of the matrix $\mathbf{X}$ is computed: $\mathbf{X} = \mathbf{U}\Sigma\mathbf{V}^T$, where the columns of the orthonormal matrix $\mathbf{U} \in \mathbb{R}^{k \times k}$ are left singular vectors, often called Proper Orthogonal Modes (POMs) [67,68] or spatial modes, the columns of $\mathbf{V} \in \mathbb{R}^{p \times p}$ give the temporally evolving coefficients, and $\Sigma \in \mathbb{R}^{k \times p}$ is a diagonal matrix representing the singular values of the matrix $\mathbf{X}$ ordered from the largest to the smallest conventionally. In particular, the rank of $\mathbf{X}$ is equal to the number of nonzero singular values, and the optimal rank- $\hat{k}$ approximation $\hat{\mathbf{X}}$ of $\mathbf{X}$ is given by the rank- $\hat{k}$ SVD truncation $\hat{\mathbf{X}} = \sum_{i=1}^{\hat{k}} \sigma_i \mathbf{U}_i \mathbf{V}_i^T$, in which σ_i is the i th singular value contained in the diagonal of Σ , and $\mathbf{U}_i, \mathbf{V}_i$ are the i th column of $\mathbf{U}$ and $\mathbf{V}$, respectively. In this work, $\hat{k} = 3$, which gives a reduced-order model of dimension 3. The reduced order dynamical coordinates $\Lambda_i = \sigma_i \mathbf{V}_i^T$, $i = 1, 2, 3$.

5.3. High-dimensional case 1: Beam MEMS resonator

The doubly clamped beam is investigated as a representative example of a high-dimensional nonlinear structural system, where a straight beam-type MEMS resonator is excited near its resonance frequency. This configuration captures the essential dynamics of micro-resonators [69] in which geometric nonlinearities play a significant role in shaping the system's response.

The doubly clamped beam in this case has length $L = 1000 \mu\text{m}$ with rectangular cross-section of dimensions $10 \mu\text{m} \times 24 \mu\text{m}$, as shown in Fig. 10a. The beam is fabricated from isotropic polysilicon [70], with density $\rho = 2330 \text{ kg/m}^3$, Young modulus $E = 167 \text{ GPa}$ and Poisson coefficient $\nu = 0.22$. Dirichlet boundary conditions are enforced on the two opposite sides of the beam. The quality factor is assumed equal to $Q = 50$. The device is excited by a body load $\mathbf{F} = \mathbf{M}\phi_1 \beta \cos(\Omega t)$ proportional to the first eigenmode ϕ_1 , with $\mathbf{M}$ mass matrix and β load multiplier. The device vibrates according to its first bending mode at $\omega_0 = 0.5475 \text{ rad}/\mu\text{s}$. In order to keep the FOM computational time at a reasonable level, a rather coarse mesh with 2607 nodes has been employed. To effectively infer the underlying dynamics of beam MEMS resonator, the POD modes are built retaining in the linear trial space the first three most energetic POD modes, depicted in Fig. 10b–d, respectively. This number of POD modes represents the minimum required to achieve a good accuracy while maintaining a computationally efficient model [39].

To infer the underlying dynamics of the MEMS resonator, training data are collected from a single simulation case with a load multiplier $\beta = 0.25$ and excitation frequency $\Omega = 0.551$. Fig. 11 illustrates the characteristics of the training set in the frequency, time and evolutionary domains. Specifically, Fig. 11a displays the time–history responses of the first three POD modes, while Fig. 11b shows their corresponding frequency spectra. The red boxes in the spectra highlight the effective harmonic orders, confirming

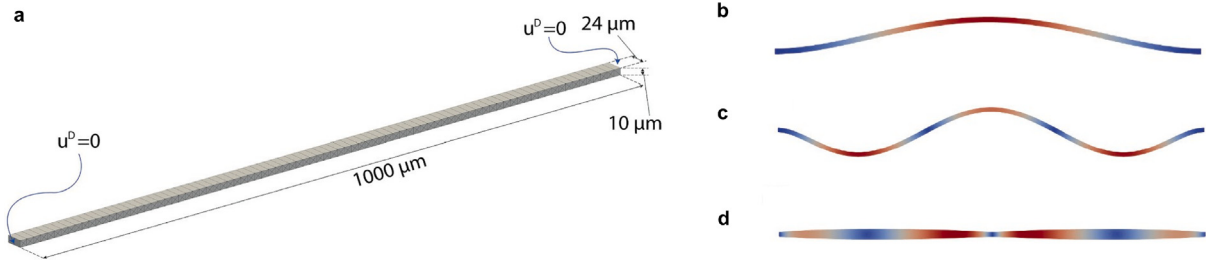


Fig. 10. Schematic representation of the beam MEMS resonator. (a) Geometry and mesh of doubly clamped beam MEMS resonator, (b–d) First three most energetic POD modes obtained from SVD.

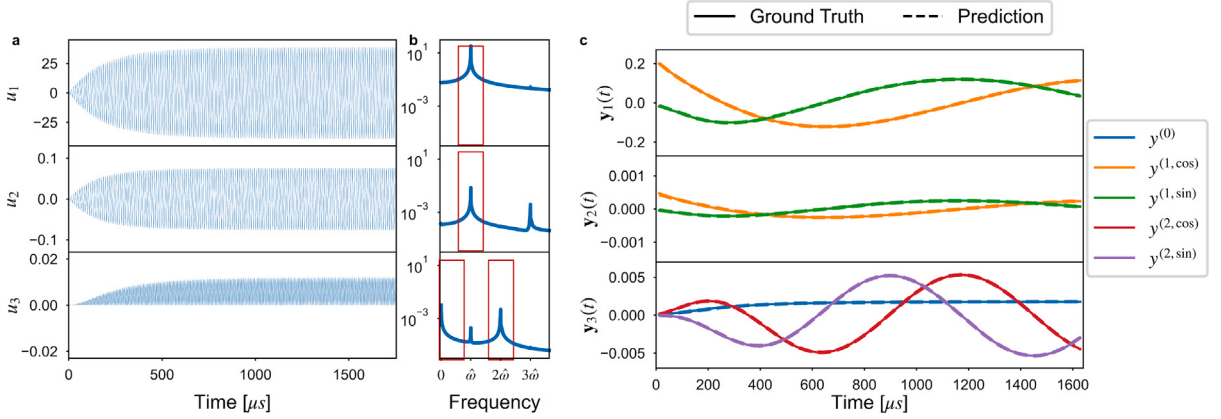


Fig. 11. Beam MEMS resonator training data in time, frequency and evolutionary domain. The model is trained using a single parameter set ($\beta = 0.25, \Omega = 0.551$). (a) Time series of the POD modal coordinates; (b) Frequency spectra, where red boxes indicate the identified effective harmonic orders; (c) Comparison of slowly varying variables for the first three POD modes, showing excellent agreement between the ground truth (solid) and prediction by Eq. (34) (dashed).

that the system response is predominantly composed of multiple discrete frequencies. Using the proposed filtering procedure (Section 2.4), the slow varying evolutionary variables for the POD modes are extracted, as shown in Fig. 11c. Building upon these extracted variables, our approach successfully identifies a parsimonious set of governing equations that govern the reduced-order space from a candidate library Θ consisting of ten pure polynomial terms up to the fifth order and trigonometric functions ($\cos(\Omega t)$ and $\sin(\Omega t)$), as explicitly formulated in Eq. (34). The identified model is capable of accurately reconstructing the slow-varying evolutionary variables (Fig. 11c dashed line), confirming that the identified nonlinear terms correctly capture the weakly nonlinear effects across different harmonic orders. Specifically, these identified equations capture the linear restorative forces alongside the geometric nonlinearities inherent in the large-deflection regime of the MEMS beam.

$$\begin{cases} \ddot{u}_1 + 2.998 \times 10^{-1} u_1 + 2.730 \times 10^{-6} u_1^3 + 1.100 \times 10^{-2} \dot{u}_1 &= 1.902 \beta \cos(\Omega t) \\ \ddot{u}_2 + 3.006 \times 10^{-1} u_2 - 1.468 \times 10^{-6} u_1 + 5.453 \times 10^{-9} u_1^3 + 2.450 \times 10^{-5} \dot{u}_1 &= 4.202 \times 10^{-3} \beta \cos(\Omega t) \\ \ddot{u}_3 + 3.006 \times 10^{-1} u_3 + 2.402 \times 10^{-6} u_1^2 - 1.574 \times 10^{-5} \dot{u}_1^2 &= 0 \end{cases} \quad (34)$$

The resulting model provides a compact yet precise representation of the resonator complex dynamics, enabling efficient prediction of its complete dynamic portrait while significantly reducing the computational burden compared to the full-order finite element model. Specifically, to evaluate the generalization capability of the identified equations, we predict the global FRCs of the beam MEMS resonator under various excitation intensities. As illustrated in Fig. 12, the predicted FRCs (dashed line) demonstrate remarkable agreement with the reference FOM results (solid line) across entire investigated range of forcing amplitudes ($\beta \in [0.125, 0.750] \mu\text{N}$), with the left and right panels presenting the amplitude–frequency and frequency–phase curves, respectively. In particular, the identified model successfully captures the nonlinear hardening behavior, characterized by the rightward tilting of the resonance peaks as the excitation amplitude increases. Notably, while the governing equations were inferred using data from only a single training point (the red dot at $\beta = 0.25, \Omega = 0.551$), the model accurately extrapolates the complex bifurcation characteristics and phase transitions for forcing levels significantly higher than the training condition. This high-fidelity match underscores the physical consistency and robustness of the proposed identification framework for high-dimensional nonlinear structural systems.

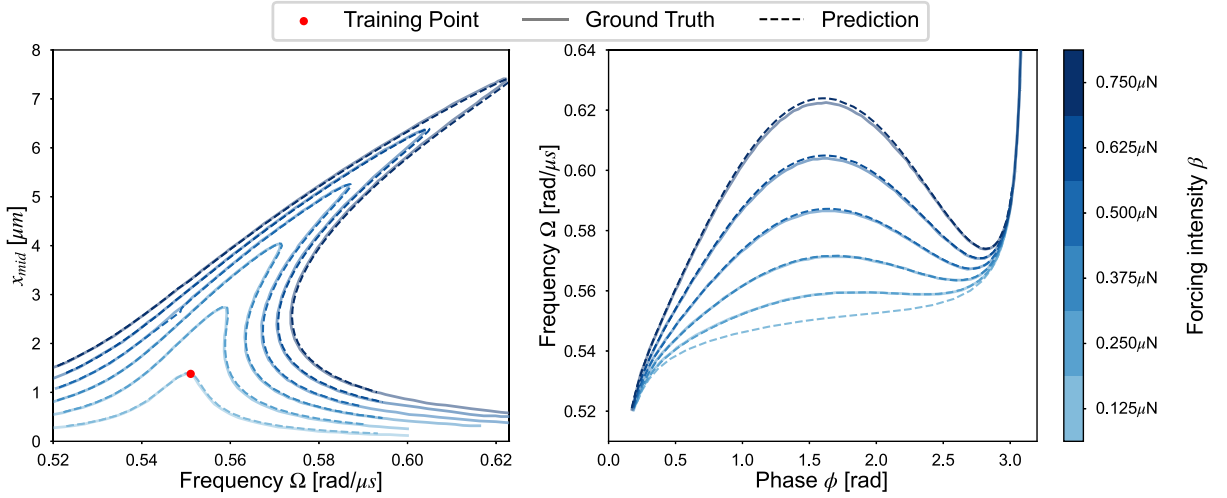


Fig. 12. FRFs of beam MEMS resonator. Comparison between predicted and reference frequency response functions under different excitation amplitudes (left: amplitude; right: phase). The proposed method accurately captures the hardening dynamical performance of beam MEMS resonator.

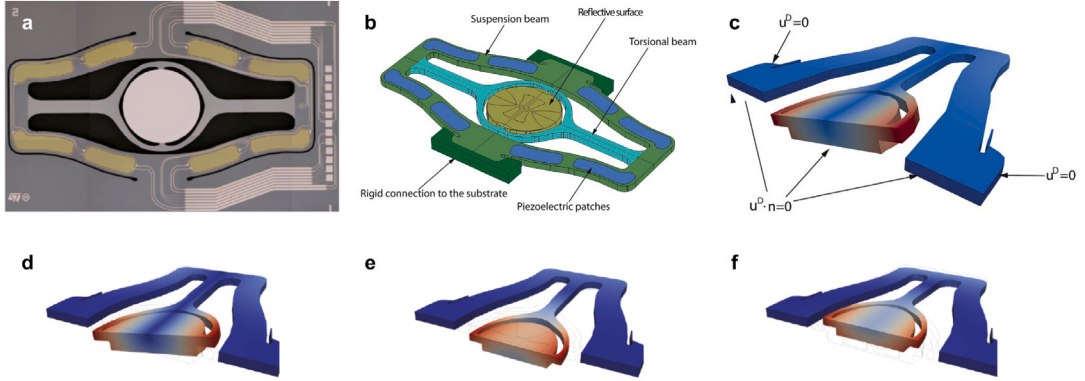


Fig. 13. Schematic representation of the MEMS micromirror: (a) Optical picture of the Micromirror; (b) Schematic view of the layout with few details on the device components; (c) Third (torsional) eigenmode that is actuated during operations and the boundary conditions used in the simulation of half of the device; (d–f) First three most energetic POD modes obtained from SVD.

5.4. High-dimensional case 2: MEMS micromirror

Scanning micromirrors have experienced rapid growth in recent years, driven by their successful implementation in a wide range of applications, from pico-projectors for Augmented Reality (AR) displays to three-dimensional (3D) scanning systems for Light Detection and Ranging (LiDAR) technologies. Because of the inertial and geometrical effects triggered by large rotations, micromirrors are intrinsically nonlinear and the prediction of their dynamic behavior is essential to guarantee a proper design and control of the mirror during the online stage. Specifically, maintaining motion stability is critical for line scanning performance, making the correct characterization of hardening and softening effects a priority.

The mirror considered, fabricated by STMicroelectronics, is illustrated in Fig. 13. The mirror plate is suspended to a gimbal connected with a torsional beam along the rotation axis and two suspension beams on each side. The mirror is assumed to be made of isotropic polysilicon [70], with density $\rho = 2330 \text{ kg/m}^3$, Young modulus $E = 167 \text{ GPa}$ and Poisson coefficient $\nu = 0.22$. Exploiting structural symmetry, only half of the micromirror is modeled with the FEM and a total of 9723 dofs. The Dirichlet boundary conditions are imposed in the substrate (the dark green areas in Fig. 13b) and on the symmetry plane to enforce the symmetry conditions. Given the high stiffness of the central mirror plate, its rotation angle is adopted as the primary state variable for both the time-domain response and the frequency response function (FRF) analysis.

To characterize the complex nonlinearities of the scanning micromirror, the training set of the MEv-SINDy framework is derived from a single simulation trajectory with a load multiplier $\beta = 3$ and an excitation frequency $\Omega = 0.183841$. Fig. 14 illustrates the sampled data across the time, frequency and evolutionary domains. Unlike the beam case, the torsional tilting of the mirror triggers a richer multi-harmonic spectrum, as evidenced by the discrete frequency peaks in Fig. 14b. The filtering procedure effectively

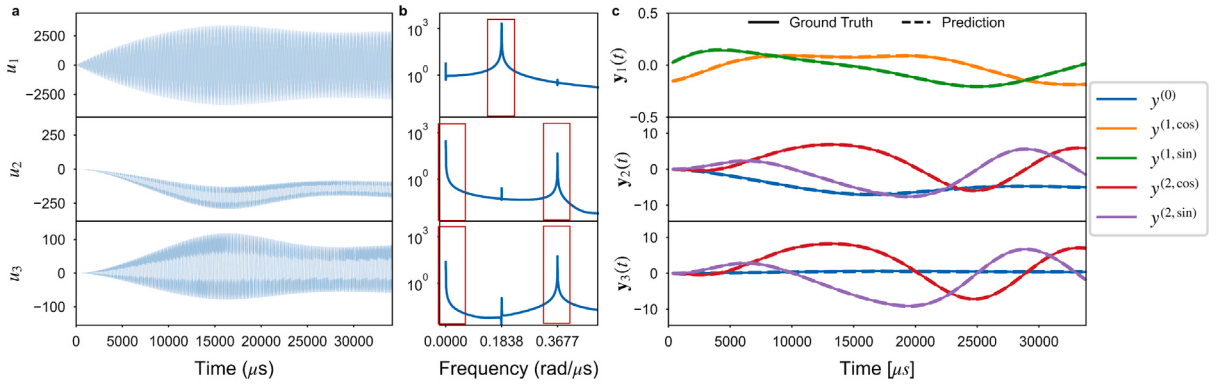


Fig. 14. MEMS micromirror training data in time, frequency and evolutionary domains: The model is trained using a single parameter set ($\beta = 3, \Omega = 0.183841$). (a) Time series of the POD modal coordinates; (b) Frequency spectra, where red boxes indicate the identified effective harmonic orders; (c) Comparison of slowly varying variables for the first three POD modes, showing excellent agreement between the ground truth (solid) and prediction by Eq. (35) (dashed).

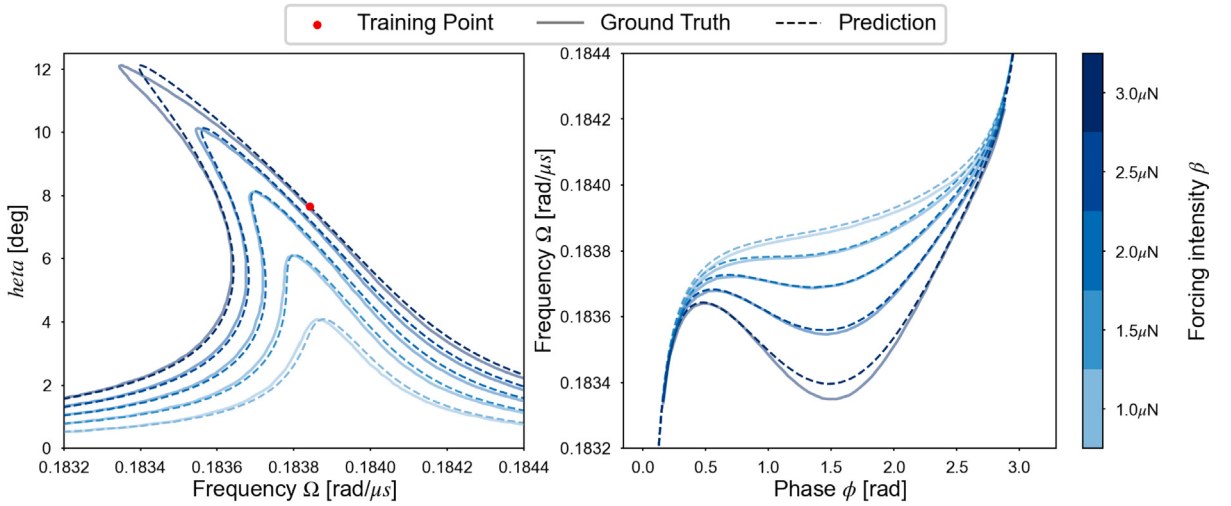


Fig. 15. FRCs of MEMS micromirror. Comparison between predicted and reference frequency response functions under different excitation amplitudes (left: amplitude; right: phase). The proposed method accurately captures the softening dynamical performance of MEMS micromirror.

isolates the evolutionary variables for the underlying manifold (Fig. 14c), providing a clean foundation for sparse regression. The algorithm yields a compact set of reduced-order equations (Eq. (35)) that faithfully reproduce the training signals (see the comparison between continuous and dashed lines in Fig. 14c). Notably, the identified model successfully decouples the inertial nonlinearities and cross-modal interactions inherent in the gimbal-suspension assembly.

$$\begin{cases} \ddot{u}_1 + 3.383 \times 10^{-2} u_1 - 1.180 \times 10^{-11} u_1^3 + 1.865 \times 10^{-4} \dot{u}_1 = 1.076 \times 10^{-1} \beta \cos(\Omega t) \\ \ddot{u}_2 + 3.374 \times 10^{-2} u_2 - 1.103 \times 10^{-7} u_1^2 + 3.912 \times 10^{-5} \dot{u}_1^2 = 0 \\ \ddot{u}_3 + 3.374 \times 10^{-2} u_3 - 9.127 \times 10^{-7} u_1^2 + 2.392 \times 10^{-5} \dot{u}_1^2 = 0 \end{cases} \quad (35)$$

The robustness of the identified model is further validated by predicting the global Frequency Response Functions (FRCs). As shown in Fig. 15, the MEv-SINDy framework accurately extrapolates the system behavior across a wide range of forcing intensities. Of particular interest is the model ability to capture the softening behaviors, which is a critical factor for ensuring scanning stability in LiDAR and AR applications. The seamless agreement between the ROM predictions and the FOM ground truth — despite the model being trained on a single localized point — underscores the physical consistency of the discovered dynamics.

6. Robustness experiments

To provide a concrete, literature-wide benchmark regarding the data acquisition burden for nonlinear frequency-response curve (FRC) prediction in MEMS structures, Table 2 systematically itemizes the required training data volume (expressed as the number

Table 2

Quantitative comparison of required training data volume (number of excitation trials) for global frequency-response curve (FRC) prediction in MEMS structures.

Reference/Methodology	Target system	Training volume (N_{trial})
AE + SINDy [39]	MEMS MicroBeam	56
SINDy + Kalman Filtering [71]	MEMS MicroArch	40
Deep Learning (NN-ROM) [15]	MEMS MicroBeam	309
Proper Orthogonal Decomposition (POD)[8]	MEMS Micromirror	50
MEv-SINDy	Beam/Micromirror	1 (One-Shot)

of independent parametric excitation trials, N_{trial}) across several state-of-the-art nonlinear reduced-order modeling (ROM) and system identification frameworks. As quantified by the empirical data, existing paradigms — ranging from standard intrusive proper orthogonal decomposition (POD) and black-box neural networks (NN) to advanced adaptive SINDy variants — strictly require dense parametric sampling ($40 \leq N_{\text{trial}} \leq 309$) across various excitation frequencies and amplitudes to construct a generalizable model capable of tracing the global FRC. This heavy data dependency exists because their mathematical formulations operate in the instantaneous state space, where the $\mathcal{O}(\epsilon)$ weakly nonlinear signatures are readily discarded as numerical residuals unless mapped across extensive boundary configurations. In sharp contrast, by embedding the generalized harmonic balance method to project the transient physics onto the slow-varying evolutionary manifold, the proposed MEv-SINDy framework successfully recovers the full invariant physical operators using exactly $N_{\text{trial}} = 1$. This order-of-magnitude reduction in data volume definitively demonstrates that the one-shot learning capability represents a fundamental mathematical advancement in sample efficiency rather than an artifact of trial-and-error hyperparameter tuning.

While the cross-literature comparison establishes the theoretical advancement of the MEv-SINDy framework, a systematic assessment of its robustness under diverse empirical conditions is essential to define its practical operational boundaries. Accordingly, this chapter investigates the sensitivity and reliability of the proposed method under varying data constraints, focusing on two critical operational aspects. First, we examine the influence of training point selection on global predictive accuracy, specifically investigating whether the location or specific regime of the single training sample affects the model's ability to extrapolate the system's global nonlinear response. Second, we evaluate the scaling of prediction precision with the number of training points. By comparing the performance of models trained on sparse versus enriched datasets across a large experimental database, we highlight our approach's primary advantage: the ability to achieve high-dimensional ROM discovery with minimal data requirements, thereby significantly reducing the data acquisition and computational overhead inherent in traditional data-driven modeling.

6.1. Influence of training point selection

To provide a comprehensive assessment of MEv-SINDy's reliability, we conduct extensive testing across two benchmark datasets: the beam MEMS resonator (56 distinct operating points, featuring specific forcing frequency and amplitude) and the MEMS micromirror (22 distinct operating points). For each individual point in these datasets, the governing equations are first identified via the proposed framework. Subsequently, the full-range frequency responses are reconstructed using MATCONT (as detailed in Section 3.5), a sophisticated parameter continuation solver. To quantify the fidelity of the identified models, we utilize the Mean Chamfer Distance of Frequency Response Curves (MCDRC) as our primary metric (as detailed in Section 3.6). Specifically, a score below $\text{MCDRC} < 0.1$ indicates that the identified model has successfully captured the essential nonlinear characteristics across the entire dynamical regime.

In particular, the reliability of the MEv-SINDy framework is systematically evaluated by analyzing the influence of training point selection on global predictive accuracy. As illustrated in Fig. 13, the framework exhibits remarkable robustness across diverse operating regimes. For the beam MEMS resonator, 51 out of 56 candidate training points yield successful global inferences (Fig. 16a), while for the MEMS micromirror, the framework achieves a 100% success rate across all 22 tested points (Fig. 16b). Notably, accurate global ROMs are identified when the training data is sampled from the strongly nonlinear regimes characterized by significant resonance tilting.

This consistency is further quantified in Fig. 17a, where the MCDRC distribution for all successful cases is shown to be highly concentrated near zero. The narrow spread of these violin plots underscores that the identification fidelity is largely invariant to the specific choice of training conditions. Qualitatively, Figs. 17b–e confirm this high fidelity; the FRC curves reconstructed from various localized training points demonstrate a nearly perfect overlap with the ground-truth simulations. The framework faithfully reproduces complex nonlinear phenomena, including the softening behavior of the beam and the dual hardening–softening transitions in the micromirror. Collectively, these results demonstrate that MEv-SINDy effectively extracts the underlying physical operators rather than merely overfitting localized data, enabling reliable global extrapolation from minimal, localized measurements.

To further enhance the reliability of the MEv-SINDy framework within a one-shot learning context, it is imperative to investigate the specific conditions under which identification fails. The five failure cases observed in the beam MEMS resonator (Fig. 17a) can be categorized into two distinct modes:

1. **Limited nonlinear effects:** Several failed points are located at frequencies far from the system natural frequency. In these near-linear regimes, the nonlinear contributions to the dynamical response are extremely small. Consequently, the sparse

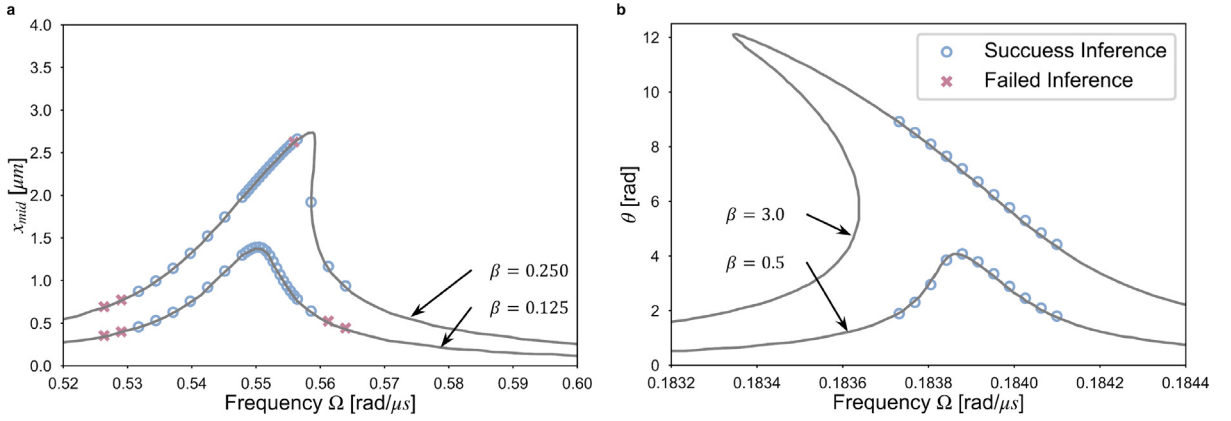


Fig. 16. Robustness evaluation of proposed approach across various training conditions. The predictive performance is assessed by identifying the model using a single training point and evaluating its global extrapolation capability. Blue circles indicate a successful inference (MCDRF < 0.1), while red crosses denote a failed inference. (a) Beam MEMS resonator: results for 56 different training points, covering two excitation amplitudes $\beta \in \{0.125, 0.250\}$ and 28 frequencies Ω ranging from 0.526 to 0.564 rad/ μ s. A total of 51 training points yield successful global predictions. (b) MEMS micromirror: results for 22 distinct training points, covering two excitation amplitudes $\beta \in \{0.5, 3.0\}$ and 11 frequencies Ω ranging from 0.1837 to 0.1841 rad/ μ s. All 22 training points achieve successful inference.

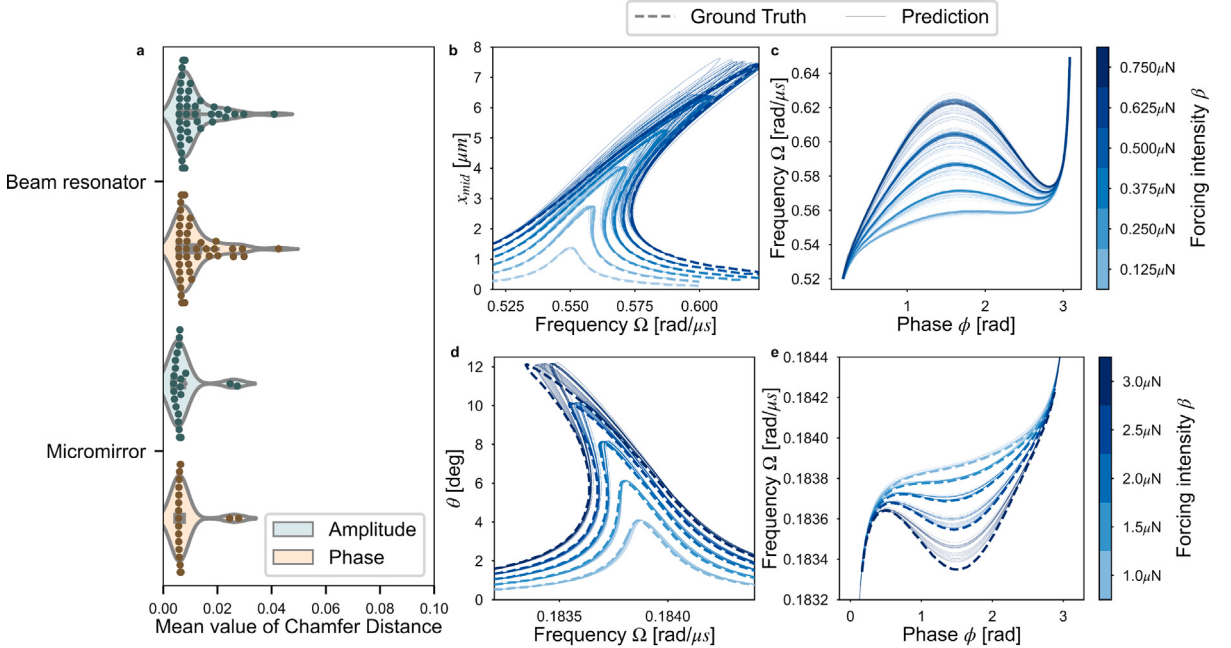


Fig. 17. Quantitative and qualitative error analysis of FRF predictions. (a) MCDRF distribution across all successful inferences; violin plots show high error concentration near zero for both devices. (b–c) Qualitative comparison between Ground Truth (dashed) and predictions (solid) for the beam resonator, and (d–e) for the micromirror. The consistent overlap across various training conditions highlights the framework's minimal sensitivity to training data selection.

regression algorithm only discovers the linear damping for the first POD mode. While the resulting model can predict a precise dynamical response within the local linear regime, it lacks the cubic stiffness nonlinearity required for global extrapolation. This underscores that for robust one-shot learning, the training point must be selected from a region where nonlinear effects are sufficiently prominent to be accounted by the sparse identification process.

2. **Incorrect identified dynamics:** Occasionally, identification failure occurs because the discovered model converges to an incorrect dynamical equation, leading to an erroneous global FRF prediction. Fortunately, such discrepancies are readily detectable during the validation phase; by comparing the model's predicted time-history against the original sampled trajectory, one can easily identify mismatches in time domain. To rectify this, two primary strategies can be employed:

fine-tuning the hyperparameters (such as the sparsity threshold) to re-infer the true underlying dynamics or re-selecting a more informative training point that better characterizes the system's nonlinear landscape.

Standard regression methods fail with a single trial because their instantaneous formulation discards the $\mathcal{O}(\epsilon)$ weak nonlinear features as numerical residuals. By mapping the dynamics onto the slow-varying evolutionary manifold, MEv-SINDy isolates the weak nonlinear effects, allowing the invariant physical coefficients to be fully extracted from the informational content of a single transient curve. This uniform statistical distribution definitively demonstrates that the one-shot predictive capability is an inherent, generalized mathematical attribute of the MEv-SINDy algorithm, rather than a fluke resulting from a trial-and-error data selection process. Moreover, the very few failures are generally limited to very low amplitudes where nonlinearities are not activate and hence non-identifiable.

6.2. Multi-point training and frequency normalization

While the MEv-SINDy framework demonstrates high reliability in one-shot learning, incorporating data from multiple excitation conditions can further stabilize the identification process. However, a fundamental mathematical discrepancy arises when attempting to directly merge datasets from different operating points. In the physical domain, the underlying system parameters (Eq. (1)) are intrinsic constants across different excitation cases. The linear frequency term requires special consideration. Its physical coefficient ω_0^2 is split into two distinct parts. The first part is the observed frequency $\hat{\omega}$ obtained via FFT (Eq. (15)). The second part is a small fine-tuning coefficient identified through the regression process (Eq. (21)). So the governing equations in our identification framework (Eq. (21)) are constructed based on the observed response frequency $\hat{\omega}$, which is determined via FFT for each specific excitation case. Because $\hat{\omega}$ varies across different forcing conditions, the mathematical representation of the linear frequency in $\Theta(t)$ becomes coupled with the local operating point. Consequently, Eq. (21) yields inconsistent coefficients for the linear frequency term across datasets, preventing a straightforward least-squares concatenation.

To resolve this, we introduce a frequency-normalization step after reformulate the evolutionary library. By selecting a reference frequency $\bar{\omega}$ (e.g., the mean of all $\hat{\omega}$), we apply a compensation term to the target vector $y(t)$. This ensures the linear frequency term is mapped to a consistent basis, effectively decoupling the physical parameters from the local observation frequency. The corrected target vector $y_{corr}(t)$ replaces the original vector y in Eq. (21). This substitution ensures that the regression targets are consistent across all training sets.

$$y_{corr}(t) = y(t) - (\hat{\omega}^2 - \bar{\omega}^2)\Theta_x(t) \quad (36)$$

where Θ_x stands for the candidate function x in the evolutionary library:

$$\Theta_x(t) = \begin{bmatrix} \Theta_x^{(0)}(t) \\ \Theta_x^{(1,\cos)}(t) \\ \Theta_x^{(1,\sin)}(t) \\ \Theta_x^{(2,\cos)}(t) \\ \Theta_x^{(2,\sin)}(t) \\ \vdots \end{bmatrix} = \begin{bmatrix} -\frac{\hat{\omega}_i}{2\pi} \int_{t-\pi/\hat{\omega}_i}^{t+\pi/\hat{\omega}_i} x(s) ds \\ -\frac{\hat{\omega}_i}{\pi} \int_{t-\pi/\hat{\omega}_i}^{t+\pi/\hat{\omega}_i} x(s) \cos(\hat{\omega}_i s) ds \\ -\frac{\hat{\omega}_i}{\pi} \int_{t-\pi/\hat{\omega}_i}^{t+\pi/\hat{\omega}_i} x(s) \sin(\hat{\omega}_i s) ds \\ -\frac{\hat{\omega}_i}{\pi} \int_{t-\pi/\hat{\omega}_i}^{t+\pi/\hat{\omega}_i} x(s) \cos(2\hat{\omega}_i s) ds \\ -\frac{\hat{\omega}_i}{\pi} \int_{t-\pi/\hat{\omega}_i}^{t+\pi/\hat{\omega}_i} x(s) \sin(2\hat{\omega}_i s) ds \\ \vdots \end{bmatrix} = \begin{bmatrix} -a_0(t) \\ -b_1(t) \\ -c_1(t) \\ -b_2(t) \\ -c_2(t) \\ \vdots \end{bmatrix}, \quad (37)$$

The efficacy of this multi-point approach is demonstrated in Fig. 18. A comparison of the single-point training results reveals a trade-off in predictive performance: the model trained on Point 1 provides a better fit for low-amplitude excitation cases, whereas the model trained on Point 2 exhibits higher fidelity for high-amplitude regimes. This suggests that localized data, while accurate within its own regime, may not fully capture the system sensitivity across the entire forcing range. By implementing the frequency-normalization terms defined in Eq. (36), we can combine the complementary advantages of both points. The composite model (Training on Points 1 and 2) effectively fuses the “low-amplitude” and “high-amplitude” information into a single, unified ROM. As a result, the identified model achieves a superior and more consistent match with the ground-truth FRC under all tested conditions. This shows that our framework can successfully integrate disparate datasets to overcome the limitations of localized sampling, ensuring stable and high-fidelity global predictions.

7. Conclusion

Learning the complete dynamic portrait of weakly forced oscillators typically requires large training datasets. This is true especially when increasing popular deep learning are applied. In this study, we proposed a one-shot method to learn frequency-response curves from a single excitation configuration or, in other words, from a single training point. To infer weakly nonlinear effects, we modified the Sparse Identification of Nonlinear Dynamic (SINDy) algorithm accounting for an evolution-based learning framework (MEv-SINDy). We shifted the focus from autonomous single-frequency dynamics to non-autonomous multi-frequency dynamics. This advancement was achieved through the Generalized Harmonic Balance (GHB) method. The core of our approach is the identification of explicit governing equations. Once these equations are learned, they provide strong physical extrapolation capabilities. This analytical representation allows the model to predict complex nonlinear responses far beyond the training point. We can then use standard numerical tools to perform rapid parametric sweeps. This strategy effectively bridges the gap between

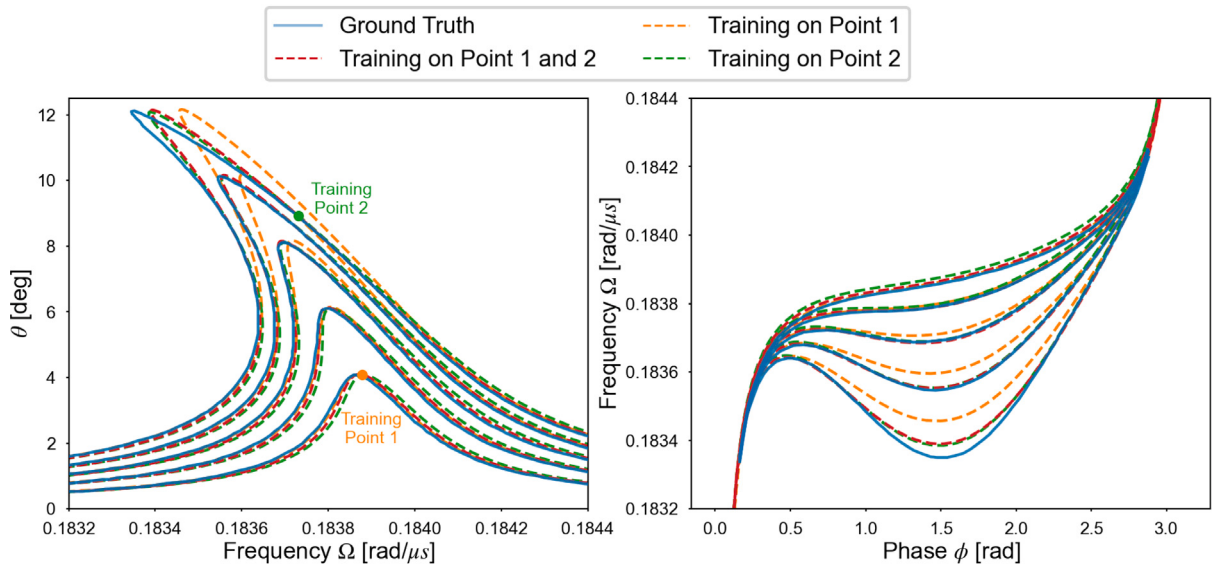


Fig. 18. Comparison of Frequency Response Curves (FRCs) identified from different training sets. (left) Amplitude–frequency response; the forcing amplitudes increase from bottom to top. (right) Phase–frequency response; the forcing amplitudes decrease from bottom to top.

limited data observations and global nonlinear response prediction, addressing the major challenge of data acquisition in real-world MEMS applications.

We applied the proposed MEv-SINDy framework considering two high-dimensional MEMS applications. These cases included a beam resonator and a MEMS micromirror. In both examples, our method achieved highly accurate one-shot learning by predicting complex nonlinear responses from only a single training point. To ensure the reliability of the approach, we performed extensive robustness tests. We investigated the influence of training point selection on the identification accuracy. This analysis helped us define effective strategies for choosing optimal data points. Furthermore, we extended the framework to include multi-point training thanks to a frequency normalization preprocessing. The results demonstrate the use of multiple training points can refine the MEv-SINDy outcomes across different excitation levels and frequency ranges. These findings confirm the practical utility of our method for real-world engineering design. The framework significantly reduces the computational burden while preserving the complex physics of forced oscillators.

CRediT authorship contribution statement

Teng Ma: Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Luca Rosafalco:** Writing – review & editing, Investigation, Conceptualization. **Wei Cui:** Supervision, Project administration, Investigation, Funding acquisition. **Lin Zhao:** Supervision, Project administration, Investigation, Funding acquisition. **Attilio Frangi:** Writing – review & editing, Supervision, Investigation, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

L.R. and A.F. acknowledge support by the Joint Research Platform “Sensor sysTEms and Advanced Materials” (STEAM) between Politecnico di Milano and STMicroelectronics. T.M, W.C., and L.Z. are supported by the National Key Research and Development Program of China (grant no. 2022YFC3005301), the National Natural Science Foundation of China (grant no. 52478552 and no. 52378527), the Natural Science Foundation of Shanghai (grant no. 23ZR1464900), the Fundamental Research Funds for the Central Universities (grant no. 22120240363) and China Scholarship Council (grant no. 20230626015).

Data availability

All data are provided in the main text and the codes are available in the public GitHub repository: <https://github.com/TengMa25/MEv-SINDy.git>.

References

- [1] L. Van Der Maaten, E.O. Postma, H.J. Van Den Herik, et al., Dimensionality reduction: A comparative review, *J. Mach. Learn. Res.* 10 (2009) 1–41.
- [2] D. Huang, H. Abdel-Khalik, C. Rabiti, F. Gleicher, Dimensionality reducibility for multi-physics reduced order modeling, *Ann. Nucl. Energy* 110 (2017) 526–540.
- [3] Y. Maday, E.M. Ronquist, A reduced-basis element method, *J. Sci. Comput.* 17 (2002) 447–459.
- [4] A. Quarteroni, A. Manzoni, F. Negri, Reduced basis methods for partial differential equations: an introduction, vol. 92, Springer, 2015.
- [5] P. Benner, M. Ohlberger, A. Cohen, K. Willcox, *Model Reduction and Approximation: Theory and Algorithms*, SIAM, 2017.
- [6] D. Amsallem, M.J. Zahr, C. Farhat, Nonlinear model order reduction based on local reduced-order bases, *Internat. J. Numer. Methods Engrg.* 92 (2012) 891–916.
- [7] S. Pagani, A. Manzoni, A. Quarteroni, Numerical approximation of parametrized problems in cardiac electrophysiology by a local reduced basis method, *Comput. Methods Appl. Mech. Engrg.* 340 (2018) 530–558.
- [8] G. Gobat, A. Opreni, S. Fresca, A. Manzoni, A. Frangi, Reduced order modeling of nonlinear microstructures through proper orthogonal decomposition, *Mech. Syst. Signal Process.* 171 (2022) 108864.
- [9] P.J. Schmid, L. Li, M.P. Juniper, O. Pust, Applications of the dynamic mode decomposition, *Theor. Comput. Fluid Dyn.* 25 (2011) 249–259.
- [10] S. Le Clainche, J.M. Vega, Higher order dynamic mode decomposition, *SIAM J. Appl. Dyn. Syst.* 16 (2017) 882–925.
- [11] J. Zhai, S. Zhang, J. Chen, Q. He, Autoencoder and its various variants, in: 2018 IEEE International Conference on Systems, Man, and Cybernetics, SMC, IEEE, 2018, pp. 415–419.
- [12] F. Romor, G. Stabile, G. Rozza, Non-linear manifold reduced-order models with convolutional autoencoders and reduced over-collocation method, *J. Sci. Comput.* 94 (2023) 74.
- [13] F.J. Gonzalez, M. Balajewicz, Deep convolutional recurrent autoencoders for learning low-dimensional feature dynamics of fluid systems, 2018, arXiv preprint arXiv:1808.01346.
- [14] S. Fresca, G. Gobat, P. Fedeli, A. Frangi, A. Manzoni, Deep learning-based reduced order models for the real-time simulation of the nonlinear dynamics of microstructures, *Internat. J. Numer. Methods Engrg.* 123 (2022) 4749–4777.
- [15] N. Franco, A. Manzoni, P. Zunino, A deep learning approach to reduced order modelling of parameter dependent partial differential equations, *Math. Comp.* 92 (2023) 483–524.
- [16] S. Fresca, A. Manzoni, POD-DL-ROM: Enhancing deep learning-based reduced order models for nonlinear parametrized PDEs by proper orthogonal decomposition, *Comput. Methods Appl. Mech. Engrg.* 388 (2022) 114181.
- [17] M. Schmidt, H. Lipson, Distilling free-form natural laws from experimental data, *Science* 324 (2009) 81–85.
- [18] K. Ruan, Y. Xu, Z.-F. Gao, Y. Liu, Y. Guo, J.-R. Wen, H. Sun, Discovering physical laws with parallel symbolic enumeration, *Nat. Comput. Sci.* (2025) 1–14.
- [19] S.L. Brunton, J.L. Proctor, J.N. Kutz, Discovering governing equations from data by sparse identification of nonlinear dynamical systems, *Proc. Natl. Acad. Sci.* 113 (2016) 3932–3937.
- [20] T.-T. Gao, G. Yan, Autonomous inference of complex network dynamics from incomplete and noisy data, *Nat. Comput. Sci.* 2 (2022) 160–168.
- [21] S.H. Rudy, S.L. Brunton, J.L. Proctor, J.N. Kutz, Data-driven discovery of partial differential equations, *Sci. Adv.* 3 (2017) e1602614.
- [22] C. Rao, P. Ren, Q. Wang, O. Buyukozturk, H. Sun, Y. Liu, Encoding physics to learn reaction–diffusion processes, *Nat. Mach. Intell.* 5 (2023) 765–779.
- [23] Z. Chen, Y. Liu, H. Sun, Physics-informed learning of governing equations from scarce data, *Nat. Commun.* 12 (2021) 6136.
- [24] S.M. Hirsh, D.A. Barajas-Solano, J.N. Kutz, Sparsifying priors for Bayesian uncertainty quantification in model discovery, *R. Soc. Open Sci.* 9 (2022) 211823.
- [25] R.K. Niven, L. Cordier, A. Mohammad-Djafari, M. Abel, M. Quade, Dynamical system identification, model selection, and model uncertainty quantification by Bayesian inference, *Chaos: Interdiscip. J. Nonlinear Sci.* 34 (2024).
- [26] L. Fung, U. Fasel, M. Juniper, Rapid Bayesian identification of sparse nonlinear dynamics from scarce and noisy data, in: *Proceedings A*, vol. 481, The Royal Society, 2025, 20240200.
- [27] Y. Yuan, X. Tang, W. Zhou, W. Pan, X. Li, H.-T. Zhang, H. Ding, J. Goncalves, Data driven discovery of cyber physical systems, *Nat. Commun.* 10 (2019) 4894.
- [28] T. Ma, W. Cui, T. Gao, S. Liu, L. Zhao, Y. Ge, Data-based autonomously discovering method for nonlinear aerodynamic force of quasi-flat plate, *Phys. Fluids* 35 (2023).
- [29] J.-C. Loiseau, B.R. Noack, S.L. Brunton, Sparse reduced-order modelling: Sensor-based dynamics to full-state estimation, *J. Fluid Mech.* 844 (2018) 459–490.
- [30] N.M. Mangan, S.L. Brunton, J.L. Proctor, J.N. Kutz, Inferring biological networks by sparse identification of nonlinear dynamics, *IEEE Trans. Mol. Biol. Multi-Scale Commun.* 2 (2016) 52–63.
- [31] M. Hoffmann, C. Fröhner, F. Noé, Reactive SINDy: Discovering governing reactions from concentration data, *J. Chem. Phys.* 150 (2019).
- [32] T.-T. Gao, B. Barzel, G. Yan, Learning interpretable dynamics of stochastic complex systems from experimental data, *Nat. Commun.* 15 (2024) 6029.
- [33] Z. Yu, J. Ding, Y. Li, Discover network dynamics with neural symbolic regression, *Nat. Comput. Sci.* (2025) 1–13.
- [34] T. Ma, T.-T. Gao, W. Cui, A. Frangi, G. Yan, L. Zhao, Encoding cumulation to learn perturbative nonlinear oscillatory dynamics, *Adv. Sci.* n/a (2026) e19707, <http://dx.doi.org/10.1002/advs.202519707>.
- [35] L. Fei-Fei, R. Fergus, P. Perona, One-shot learning of object categories, *IEEE Trans. Pattern Anal. Mach. Intell.* 28 (2006) 594–611, <http://dx.doi.org/10.1109/TPAMI.2006.79>.
- [36] A. Ivanov, U. Iben, A. Golovkina, Physics-based polynomial neural networks for one-shot learning of dynamical systems from one or a few samples, 2020, arXiv arXiv:2005.11699. URL: <https://api.semanticscholar.org/CorpusID:218870280>.
- [37] M. Darcy, B. Hamzi, G. Livieri, H. Owahdi, P. Tavallali, One-shot learning of stochastic differential equations with data adapted kernels, *Phys. D: Nonlinear Phenom.* 444 (2023) 133583.
- [38] A. Jiao, H. He, R. Ranade, et al., One-shot learning for solution operators of partial differential equations, *Nat. Commun.* 16 (2025) 8386, <http://dx.doi.org/10.1038/s41467-025-63076-z>, URL: <https://doi.org/10.1038/s41467-025-63076-z>.
- [39] P. Conti, G. Gobat, S. Fresca, A. Manzoni, A. Frangi, Reduced order modeling of parametrized systems through autoencoders and SINDy approach: Continuation of periodic solutions, *Comput. Methods Appl. Mech. Engrg.* 411 (2023) 116072.
- [40] R. Tibshirani, Regression shrinkage and selection via the LASSO, *J. R. Stat. Soc. Ser. B Stat. Methodol.* 58 (1996) 267–288.
- [41] P. Zheng, T. Askham, S.L. Brunton, J.N. Kutz, A.Y. Aravkin, A unified framework for sparse relaxed regularized regression: SR3, *IEEE Access* 7 (2018) 1404–1423.
- [42] P.C. Hansen, Analysis of discrete ill-posed problems by means of the L-curve, *SIAM Rev.* 34 (1992) 561–580.
- [43] E. Van Den Berg, M.P. Friedlander, Probing the Pareto frontier for basis pursuit solutions, *SIAM J. Sci. Comput.* 31 (2009) 890–912.
- [44] A. Cortiella, K.C. Park, A. Doostan, Sparse identification of nonlinear dynamical systems via reweighted ℓ_1 -regularized least squares, *Comput. Methods Appl. Mech. Engrg.* 376 (2021) 113620, <http://dx.doi.org/10.1016/j.cma.2020.113620>.
- [45] A. Cortiella, K.C. Park, A. Doostan, A priori denoising strategies for sparse identification of nonlinear dynamical systems: A comparative study, *J. Comput. Inf. Sci. Eng.* 23 (2023) 011004, <http://dx.doi.org/10.1115/1.4054573>.

- [46] M.J. Wainwright, Sharp thresholds for high-dimensional and noisy sparsity recovery using ℓ_1 -constrained quadratic programs (LASSO), *IEEE Trans. Inform. Theory* 55 (2009) 2183–2202.
- [47] D.L. Donoho, I.M. Johnstone, Ideal spatial adaptation by wavelet shrinkage, *Biometrika* 81 (1994) 425–455.
- [48] J.A. Sanders, F. Verhulst, J. Murdock, *Averaging Methods in Nonlinear Dynamical Systems*, vol. 59, Springer, 2007.
- [49] V.M. Volosov, Averaging in systems of ordinary differential equations, *Russ. Math. Surv.* 17 (1962) 1.
- [50] A.H. Nayfeh, B. Balachandran, *Applied Nonlinear Dynamics: Analytical, Computational, and Experimental Methods*, John Wiley & Sons, 2008.
- [51] A.C.J. Luo, J. Huang, Approximate solutions of periodic motions in nonlinear systems via a generalized harmonic balance, *J. Vib. Control* 18 (2012) 1661–1674.
- [52] L. Vandenberghe, *The CVXOPT linear and quadratic cone program solvers*, 2010, Online: <http://cvxopt.org/documentation/coneprog.pdf> 53.
- [53] Y.K. Cheung, S.H. Chen, S.L. Lau, Application of the incremental harmonic balance method to cubic non-linearity systems, *J. Sound Vib.* 140 (1990) 273–286.
- [54] M.R. Osborne, On shooting methods for boundary value problems, *J. Math. Anal. Appl.* 27 (1969) 417–433.
- [55] T. Carrington Jr., Using collocation to study the vibrational dynamics of molecules, *Spectrochim. Acta Part A: Mol. Biomol. Spectrosc.* 248 (2021) 119158.
- [56] T.J. Ypma, Historical development of the Newton–Raphson method, *SIAM Rev.* 37 (1995) 531–551.
- [57] B. Krauskopf, H.M. Osinga, J. Galán-Vioque, et al., *Numerical Continuation Methods for Dynamical Systems*, vol. 2, Springer, 2007.
- [58] E.J. Doedel, B. Oldeman, Auto-07p: Continuation and bifurcation software, *Montr. QC: Concordia Univ. Can.* (1998).
- [59] L. Guillot, B. Cochelin, C. Vergez, A Taylor series-based continuation method for solutions of dynamical systems, *Nonlinear Dynam.* 98 (2019) 2827–2845.
- [60] M. Krack, J. Gross, *Harmonic balance for nonlinear vibration problems*, 2019.
- [61] H. Dankowicz, F. Schilder, *Recipes for Continuation*, SIAM, 2013.
- [62] A. Dhooge, W. Govaerts, Y.A. Kuznetsov, W. Mestrom, A.M. Riet, B. Sautois, *MATCONT and CL MATCONT: Continuation toolboxes in matlab*, Univ. Gent Belg. Utrecht Univ. Neth. (2006).
- [63] M.A. Butt, P. Maragos, Optimum design of chamfer distance transforms, *IEEE Trans. Image Process.* 7 (1998) 1477–1484.
- [64] N.A. Kudryashov, The generalized Duffing oscillator, *Commun. Nonlinear Sci. Numer. Simul.* 93 (2021) 105526.
- [65] G.A. Holzapfel, *Nonlinear solid mechanics: A continuum approach for engineering science*, 2002.
- [66] C.W. Rowley, T. Colonius, R.M. Murray, Model reduction for compressible flows using POD and Galerkin projection, *Phys. D: Nonlinear Phenom.* 189 (2004) 115–129.
- [67] L.F. Abbott, T.B. Kepler, Model neurons: From Hodgkin-Huxley to hopfield, in: *Statistical Mechanics of Neural Networks: Proceedings of the Xlth Sitges Conference Sitges, Barcelona, Spain, 3–7 June 1990*, Springer, 2005, pp. 5–18.
- [68] K. Lu, Y. Jin, Y. Chen, Y. Yang, L. Hou, Z. Zhang, Z. Li, C. Fu, Review for order reduction based on proper orthogonal decomposition and outlooks of applications in mechanical systems, *Mech. Syst. Signal Process.* 123 (2019) 264–297.
- [69] V. Zega, G. Gattere, S. Koppaka, A. Alter, G.D. Vukasin, A. Frangi, T.W. Kenny, Numerical modelling of non-linearities in MEMS resonators, *J. Microelectromech. Syst.* 29 (2020) 1443–1454.
- [70] A. Corigliano, B. De Masi, A. Frangi, C. Comi, A. Villa, M. Marchi, Mechanical characterization of polysilicon through on-chip tensile tests, *J. Microelectromech. Syst.* 13 (2004) 200–219.
- [71] L. Rosafalco, P. Conti, A. Manzoni, S. Mariani, A. Frangi, Online learning in bifurcating dynamic systems via SINDy and Kalman filtering, 2024, arXiv preprint arXiv:2411.04842.