A parallel framework of physics-informed neural networks for model identification of linear and nonlinear systems

Nikhil Mahar, Subhamoy Sen, Laurent Mevel

A Parallel Framework of Physics-Informed Neural Networks for Model Identification of Linear and Nonlinear Systems

Nikhil Mahar[a], Subhamoy Sen[a,1], Laurent Mevel[b]

[a] *i4S Laboratory, Indian Institute of Technology Mandi, Mandi, HP, India*
[b] *Univ. Gustave Eiffel, Inria, Cosys-SII, I4S, Campus de Beaulieu, Rennes, France*

Abstract

Recent advances in sensor and communication technologies have enabled widespread monitoring of complex structures using both model-based and data-driven methods for Structural Health Monitoring (SHM). While model-assisted approaches struggle with the complexity of real-world systems, purely data-driven methods offer simplicity and speed but lack physical interpretability. Physics-Informed Neural Networks (PINNs) offer a promising middle ground by integrating data with governing physical laws. However, most existing PINN applications focus on solving forward problems, with limited attention to inverse parameter estimation relevant to SHM.

This study presents a parallel PINN framework for efficient state and parameter estimation in complex structural systems. To address the computational burden of high-dimensional systems, the global dynamics are reformulated into a set of coupled lower-dimensional subproblems, one for each degree of freedom, without any physical partitioning of the domain. This algebraic decoupling retains structural continuity through shared global matrices, enabling localized learning while preserving global consistency. The approach is validated on both linear and nonlinear systems using harmonic and real earthquake excitations, demonstrating robustness under high noise, sparse data, and varying system complexity. An experimental study on a scaled shear frame further supports its practical relevance for real-world SHM.

Keywords: *Parameter estimation, Coupled dynamical system, Parallel physics-informed neural networks.*

1. Introduction

Within the realm of structural engineering, endeavors to model and predict structural behavior can be broadly classified into two primary approaches: data-driven and physics-driven methods. Data-driven methods, encompassing statistical techniques defined in time [1–4] and frequency domains [5–7], along with latent space tools like blind source separation (BSS) [8–11] have been extensively studied. These approaches

*Corresponding author; *E-mail address: subhamoy@iitmandi.ac.in*

typically rely on empirical data for model construction but often exhibit constraints when dealing with non-linear systems. In contrast, physics-driven methods assume a predefined model structure, either in the form of a parameterized model or a set of governing differential equations, which are subsequently updated through optimization or Bayesian approaches [12–17].

In contrast to model-based approaches for parameter estimation, such as Bayesian filtering algorithms [18–20], model-independent techniques [8–11] generally rely on certain fault-sensitive statistical distances and may not leverage the known physics to inculcate inherent spatial correlation in the data into the estimation process. Conversely, the inclusion of a support model, while theoretically advantageous, can introduce significant computational burdens and compromise qualitative precision due to the sensitivity of statistical distances concerning the model, making them susceptible to errors.

The adoption of a comprehensive optimization-based approach instead of a simple hypothesis framework may inadvertently slow down the overall process and reduce efficiency. Additionally, integrating data with physical models for estimation can introduce significant computational and modeling constraints, potentially resulting in suboptimal performance. However, a breakthrough is emerging in the realm of machine learning (ML) and artificial intelligence (AI)-based algorithms.

Continuing in the domain of vibration-based structural health monitoring (SHM) and system identification, Artificial Neural Networks (ANNs) have gained considerable attention in the past decade [21, 22]. Researchers have successfully utilized ANNs to ascertain various parameters of structural systems, including mass, stiffness, damping coefficients, eigenvalues, and eigenmodes, utilizing acceleration and vibration data as inputs [23–26]. Nevertheless, the accuracy, efficiency, and generalizability of such networks heavily depend on the availability of extensive training data, limiting their use for sparsely measured systems. Moreover, their "black box" nature poses challenges in interpreting the underlying physics of the system. These limitations eventually impede their integration into real-life structural system identification, where data collection is often expensive, time-consuming, and uncertain [27, 28]. Additionally, the fixed architecture of ANNs may lead to errors when the system properties undergo changes, necessitating assumptions and a substantial amount of input data capturing a wide spectrum of system variability.

To overcome inherent limitations in modeling complex physical systems, researchers have introduced PINNs as a viable solution [29]. PINNs integrate the physics of the system, represented by governing equations, into the loss function, combining the strengths of ANNs with physical laws. This integration allows for more precise modeling in situations with limited data and simplified system representations. PINNs have showcased remarkable adaptability, as seen in innovations such as Conservative PINNs (cPINNs) and Variational PINNs (VPNNs), customized for specific problem domains [30, 31]. [32] introduced physics-guided CNNs and physics-informed Long Short-Term Memory (LSTM) networks, showcasing the utility of PINNs in predicting building responses to earthquake excitation. Similarly [33, 34] introduced an uncertainty-oriented, physics-informed LSTM neural network and a hybrid model and data-driven framework for dynamic load

identification. These advancements underscore PINNs' versatility in diverse fields, spanning fluid mechanics and solid mechanics [35, 36]. Their proficiency in handling differential equations, commonly used to describe physical system behavior, further enhances their utility [37].

While PINNs have been extensively used for inverse problems, limited studies have explored their application in state-parameter estimation, specifically for inverse problems in structural dynamics. [38] proposed physics-informed neural ODEs to address linear and nonlinear structural identification problems in discrete systems. However, a significant limitation is their reliance on full state measurements at all relevant degrees of freedom, constraining their real-world application. Recently, researchers have utilized PINNs to their full potential, solving both forward and inverse problems in discrete structural systems governed by Ordinary Differential Equations (ODEs). For example, [39], [40], and [41] concurrently estimated the states and parameters of governing linear and nonlinear ODEs representing multi-degree-of-freedom (MDOF) systems under forced vibration. PINNs have demonstrated adaptability in addressing changes in structural properties, proving advantageous for diagnosing and controlling damaged systems [42, 43].

However, these studies primarily focused on modal parameter identification, limiting their application to nonlinear systems. The conventional PINNs approach, with a single neural network as a global regularizer, is perceived to be computationally challenging due to over-parameterization, impeding its application for high-dimensional systems. This aspect will be demonstrated through numerical experimentation later in this article. The proposed method tackles these challenges by breaking down the problem into smaller algebraic segments using a parallelized PINN framework. This innovative architecture estimates both observable and hidden variables, as well as health parameters, within SHM by focusing on state and parameter estimation over time. Training optimizes the network and health parameters by applying a physics-informed loss function that leverages actual data and integrates spatial correlations in dynamic responses and boundary conditions.

The proposed parallel and interconnected PINN framework enables localized enforcement of physics-based constraints without physically partitioning the structure, improving computational efficiency and practical feasibility. Its performance is demonstrated through numerical and experimental studies involving both linear and nonlinear systems, where the estimated time histories and parameter trajectories closely match true values, confirming the approach's accuracy and responsiveness in system identification. The theoretical background on ANNs and PINNs is introduced in Section 2, followed by a detailed description of the proposed methodology in Section 3. Numerical and experimental validations are presented in Sections 4 and 5, respectively, with discussion and concluding remarks summarized in Section 6 and Section 7.

2. Theoretical background

Originally inspired by biological neurons, ANNs are universal function approximators, available in shallow or deep forms depending on the number of hidden layers [44, 45]. Deep Neural Networks (DNNs), with

multiple layers, enhance model expressiveness and performance by leveraging large datasets and computational power [46, 47]. Their success across domains such as computer vision, natural language processing, and medical imaging has led to widespread adoption [48]. For a given input-measurement data $\{\mathbf{I}, \mathbf{y}\}$, a generalized DNN with L layers is defined as:

$$\tilde{\mathbf{y}} = \mathcal{N}(\mathbf{I}; \theta) \tag{1}$$

where, $\mathcal{N}(\mathbf{I}; \theta) = \sigma_L(\mathbf{W}_L \cdot \sigma_{L-1}(\ldots \sigma_1(\mathbf{W}_1 \cdot \mathbf{I} + \mathbf{b}_1) + \ldots) + \mathbf{b}_L)$. In supervised learning, the network is trained using n input-measurement data pairs $\{\mathbf{I}_i, \mathbf{y}_i\}$ by optimizing the weights $\mathbf{W}_l$ and biases $\mathbf{b}_l$, collectively denoted as θ, to minimize data/measurement loss as:

$$\hat{\theta} = \arg\min_{\theta} \left\{ \mathcal{L}(\theta) \coloneqq \frac{1}{n} \sum_{i=1}^{n} |\mathbf{y}_i - \tilde{\mathbf{y}}_i|^2 \right\} \tag{2}$$

2.1. Physics-Informed Neural Network

Yet, both deep and shallow neural networks face the ongoing hurdle of needing extensive training data, especially in engineering, where acquiring such measurement data through experiments or numerical simulations can be prohibitively expensive. This shortage of data can undermine performance, particularly in systems with sparse sampling. Researchers address this by utilizing accumulated information, such as differential equations, to fill data gaps. Raissi et al. [29] proposed PINNs, which integrate the physical laws represented by partial differential equations into supervised learning tasks. These networks can be used for solving PDEs directly (forward problems) or estimating parameters when solutions are available (inverse problems). Essentially, PINNs apply the broad function approximation attributes of neural networks to articulate solutions for both linear and non-linear PDEs governing system dynamics in contexts expressed as:

$$\begin{aligned} \mathcal{F}(\mathbf{u}(\mathbf{x}, t); \lambda) &= 0, \quad \mathbf{x} \in \Omega, \quad t \in [0, T] \\ \mathcal{B}(\mathbf{u}(\mathbf{x}, t)) &= 0, \quad \mathbf{x} \in \partial\Omega, \quad t \in [0, T] \end{aligned} \tag{3}$$

Here, $\mathcal{F}[\bullet; \lambda]$ denotes a linear or nonlinear governing differential operator (*gde*) defining the dynamics of a latent variable, $\mathbf{u}$, defined in the space-time continuum ($\{\mathbf{x},\ t\}$). The differential setup is characterized by a set of parameters, λ. $\mathcal{B}(\bullet)$ denotes arbitrary boundary operator (e.g., Dirichlet, Neumann, Robin, etc.) that the latent variable $\mathbf{u}(\mathbf{x}, t)$ needs to adhere to in addition to the *gde*.

2.1.1. Forward PINN

Recent applications of PINN aimed to substitute expensive differential equation simulations with a forward model guided by physics. Herein, the latent variable $\mathbf{u}(\mathbf{x}, t)$ can be consistently approximated by the DNN model $\mathcal{N}(\mathbf{x}, t; \theta)$, where $\mathbf{x}$ and t define the input $\mathbf{I}$ (see Equation (1)) through spatial and temporal

coordinates. The network $\mathcal{N}(\mathbf{x}, t; \theta)$ is characterized by parameters θ, which can otherwise be trained using an extensive input ($\mathbf{I}$) and corresponding measurement data ($\mathbf{y}$) in the absence of any known physics as detailed in Equation (3). To reduce reliance on these extensive datasets, physical laws can be incorporated into the training process by adding additional constraint terms to the loss function $\mathcal{L}(\theta)$ introduced in Equation (2). These constraints, derived from various elements (such as *gde*, boundary conditions, etc.), ensure the training aligns with the known system physics.

$$\begin{aligned}
\mathcal{L}_{int}(\theta) &= \frac{1}{N_1}\sum_{i=1}^{N_{int}}\sum_{j=1}^{N_{t1}}\left|\mathcal{F}\left(\mathcal{N}(\mathbf{x}_i, t_j; \theta); \lambda\right)\right|^2, \text{with } \{\mathbf{x}_i, t_j\} \in \Omega_{int} \\
\mathcal{L}_{bc}(\theta) &= \frac{1}{N_2}\sum_{i=1}^{N_{bc}}\sum_{j=1}^{N_{t2}}\left|\mathcal{B}\left(\mathcal{N}(\mathbf{x}_i, t_j; \theta)\right)\right|^2, \text{with } \{\mathbf{x}_i, t_j\} \in \Omega_{bc} \\
\mathcal{L}_{data}(\theta) &= \frac{1}{N_3}\sum_{i=1}^{N_{data}}\sum_{j=1}^{N_{t3}}\left|\mathcal{H}(\mathcal{N}(\mathbf{x}_i, t_j; \theta)) - \mathbf{y}(\mathbf{x}_i, t_j)\right|^2, \quad \text{with } \{\mathbf{x}_i, t_j\} \in \Omega_{data} \\
\mathcal{L}(\theta) &= \alpha * \mathcal{L}_{int}(\theta) + \beta * \mathcal{L}_{bc}(\theta) + \gamma * \mathcal{L}_{data}(\theta)
\end{aligned} \tag{4}$$

Here, the set $\{\mathbf{x}_i, t_j\}$ represents all space-time points sampled within the internal domain (Ω_{int}), boundary domain (Ω_{bc}) and measurement data domain (Ω_{data}) . These points are referred to as collocation points (for the internal domain), boundary points (for the boundary domain) and measurement points (for the measurement data domain). The quantities N_{int}, N_{bc} and N_{data} denote the number of spatial locations in the internal, boundary and measurement domains, respectively, while N_{t1}, N_{t2} and N_{t3} represent the number of time instances associated with each. Consequently, the total number of collocation points, boundary points and measurement points are given by $N_1 = N_{int} \times N_{t1}$, $N_2 = N_{bc} \times N_{t2}$ and $N_3 = N_{data} \times N_{t3}$ respectively. Collocation and boundary points are used to enforce the physical laws in addition to data/measurement loss introduced in Equation (2). $\mathcal{H}(\bullet)$ is the measurement function mapping the field variable $\mathbf{u}(\mathbf{x}_i, t_j)$ to corresponding measurement data $\mathbf{y}(\mathbf{x}_i, t_j)$. For the context of the forward framework, the system parameter λ is assumed to be known. The weighted loss function $\mathcal{L}(\theta)$ can be assumed to be a linear combination of residual loss, $\mathcal{L}_{int}$, boundary loss, $\mathcal{L}_{bc}$, and data loss, $\mathcal{L}_{data}$ as defined in Equation (4). The network optimization protocol is then enforced to estimate θ, under the influence of a configurable set of hyperparameters, α, β and γ, acting as penalty terms, that optimally minimize $\mathcal{L}(\theta)$. It's noteworthy that the network parameters are shared by both $\mathcal{N}(\mathbf{x}, t; \theta)$ and $\mathcal{L}(\theta)$. Once the optimum θ, which minimizes $\mathcal{L}$, is determined, the optimal solution $\mathbf{u}(\mathbf{x}, t)$ is also established.

The notable benefits of PINN in addressing the forward problem, as illustrated in [29] and [49], encompass its freedom from extensive dependence on the richness of simulation data, thereby ensuring computational efficiency. Additionally, PINN, trained with *gde*s, inherently adheres to essential physical principles such as invariance and symmetry.

2.1.2. Inverse PINN

The inverse method for implementing PINN in SHM and state-parameter estimation lacks extensive exploration. In system estimation's inverse domain, conventional data-driven methods focus on approximating $\mathbf{u}(\mathbf{x}, t)$ via a DNN model $\mathcal{N}(\mathbf{x}, t; \{\theta; \lambda\})$, conditioned on system parameters λ. To facilitate this,

λ is augmented with θ, forming an expanded parameter set $\{\theta;\lambda\}$, which is inferred from existing data. Furthermore, to integrate physics-based understanding, physical constraints are embedded in the training process through loss functions:

$$\begin{aligned}
\mathcal{L}_{int}(\{\theta;\lambda\}) &= \tfrac{1}{N_1}\sum_{i=1}^{N_{int}}\sum_{j=1}^{N_{t1}}\left|\mathcal{F}\left(\mathcal{N}(\mathbf{x}_i,t_j;\{\theta;\lambda\});\lambda\right)\right|^2, \quad \text{with } \{\mathbf{x}_i,t_j\}\in\Omega_{int},\\
\mathcal{L}_{bc}(\{\theta;\lambda\}) &= \tfrac{1}{N_2}\sum_{i=1}^{N_{bc}}\sum_{j=1}^{N_{t2}}\left|\mathcal{B}\left(\mathcal{N}(\mathbf{x}_i,t_j;\{\theta;\lambda\})\right)\right|^2, \quad \text{with } \{\mathbf{x}_i,t_j\}\in\Omega_{bc},\\
\mathcal{L}_{data}(\{\theta;\lambda\}) &= \tfrac{1}{N_3}\sum_{i=1}^{N_{data}}\sum_{j=1}^{N_{t3}}\left|\mathcal{H}(\mathcal{N}(\mathbf{x}_i,t_j;\{\theta;\lambda\}))-\mathbf{y}(\mathbf{x}_i,t_j)\right|^2, \quad \text{with } \{\mathbf{x}_i,t_j\}\in\Omega_{data}
\end{aligned} \tag{5}$$

wherein $\{\mathbf{x}_i,t_j\}\in\Omega_{data}$ denotes the space-time coordinates for which measurement ($\mathbf{y}(\mathbf{x}_i,t_j)$) is available. These samples are typically obtained from real-world sources, such as experimental data. Additionally, the overall loss function is defined as a weighted sum of the component loss functions.

$$\mathcal{L}(\{\theta;\lambda\}) = \alpha*\mathcal{L}_{int}(\{\theta;\lambda\}) + \beta*\mathcal{L}_{bc}(\{\theta;\lambda\}) + \gamma*\mathcal{L}_{data}(\{\theta;\lambda\}) \tag{6}$$

While a converged PINN simultaneously fits training data and satisfies physical laws via a physics-based loss, its use in inverse problems for structural dynamics, especially with coupled differential equations, remains underexplored. This study bridges that gap by proposing a novel parallel PINN framework.

3. Proposed Framework

The proposed framework focuses on the inverse estimation of a mechanical model, which is represented in a discretized form using Finite Element Modelling (FEM). The system is characterized by mass $\mathbf{M}$, damping $\mathbf{C}$, and stiffness $\mathbf{K}(\lambda)$ matrices, where only the stiffness matrix is modeled as a function of the health parameter λ. In this context, the index x refers to the spatial positions of the nodal degrees of freedom (DOFs), with the corresponding system responses: displacement $\mathbf{u}$, velocity $\dot{\mathbf{u}}$, and acceleration $\ddot{\mathbf{u}}$, represented as vectors defined over the discretized DOFs. The dynamic behavior of the system under external excitation $\mathbf{f}(x,t)$ is governed in continuous time by the following equation:

$$\begin{aligned}
\mathbf{M}\ddot{\mathbf{u}}(\mathbf{x},t) + \mathbf{C}\dot{\mathbf{u}}(\mathbf{x},t) + \mathbf{K}(\lambda)\mathbf{u}(\mathbf{x},t) - \mathbf{f}(\mathbf{x},t) = 0\\
\text{or in compact form,} \quad \mathcal{F}(\mathbf{u}(\mathbf{x},t),\lambda) = 0
\end{aligned} \tag{7}$$

The inverse estimation of a mechanical model, based on discrete-time response signals, inevitably leads to time-discrete formulations that remain spatially coupled. Such coupled MDOF formulations can pose significant computational challenges, potentially rendering direct estimation approaches impractical. To overcome this complexity, a parallel PINN framework is introduced, wherein the system's *gde* is algebraically decomposed into smaller, tractable subdomains. This is achieved by isolating each row of the *gde*, effectively creating a set of lower-dimensional systems, each governed by a reduced number of state and parameter

variables. A dedicated PINN is employed for each of these subsystems, allowing for simultaneous estimation across the entire model. Each PINN operates on an independent subset of the global health parameters and is trained concurrently with the others. Despite their parallel operation, these networks are intrinsically linked through a composite loss function that incorporates the outputs from all subsystems, ensuring consistency and coherence with the global dynamics. This modular approach not only enhances computational efficiency but also enables scalable and robust estimation of the underlying health parameters.

By parameterizing the system dynamics using a set of health indicators λ, the goal is to estimate the spatially distributed health state of the structural system. To this end, the parameter set $\boldsymbol{\lambda}$ is defined as a collection of localized health or damage parameters, denoted as $\boldsymbol{\lambda}_i$ for $i = 1, \cdots, n_s$, where n_s represents the total number of substructures resulting from the discretization of the system. Each individual parameter λ_i corresponds to a specific discretized element of the FEM and quantifies its health condition. Practically, $\boldsymbol{\lambda}_i$ reflects the degree of stiffness degradation at the associated location, where a value of 0 represents complete damage, and a value of 1 indicates an undamaged or fully healthy state. The key steps of the proposed estimation procedure are outlined as follows: Step 1: Define a set of space-time collocation points, denoted as $\Omega_{int} = \{(\mathbf{x}_i, t_j)\}$, which are distributed across spatial and temporal coordinates. When focusing on an individual substructure at a specific spatial location x_i, the neural network approximator for the latent variable $\mathbf{u}$ simplifies to a time-dependent form: $\mathbf{u}(t)|_{x=x_i} \approx \mathcal{N}_i(t_j; \{\theta_i, \lambda_i\})$, for all $j \in \{1, \ldots, N_{t1}\}$, thus removing explicit spatial dependency. Simillarly, measurement points for each isolated substructure, the relevant subset $\Omega_{i,data}$—containing measurements at space coordinate x_i is used to train the corresponding network $\mathcal{N}_i(t_j; \{\theta_i, \lambda_i\})$, for all $j \in \{1, \ldots, N_{t3}\}$.

Step 2: Subsequently, to train networks for the i^{th} substructure independently, it is necessary to define the physics-informed loss function to quantify the physics loss as follows:

$$\mathcal{L}_{i,int}(\{\theta; \lambda\})|_{\mathbf{x_i}} = \frac{1}{N_{t1}} \sum_{j=1}^{N_{t1}} \left| R_i(t_j; \{\bar{\theta}_i; \bar{\lambda}_i\}) \right|^2 \quad \text{with} \quad \{\mathbf{x}_i; t_j\} \in \Omega_{int} \tag{8}$$

where $R_i(t_j, \{\bar{\theta}; \bar{\lambda}_i\})$ represents the residual in the physical equation depicting the substructure dynamics for i^{th} substructure (i.e., $\mathcal{F}_i(\bullet)$), conditioned only on a subset, $\{\bar{\theta}; \bar{\lambda}\}$, of the complete sets of network parameters $\theta = (\theta_i)$ and system parameters $\lambda = (\lambda_i)$. This can further be elaborated concerning Equation (3) as,

$$R_i(t_j, \{\bar{\theta}_i; \bar{\lambda}_i\}) = \mathcal{F}_i\left(\mathbb{N}_i(t_j; \{\bar{\theta}_i; \bar{\lambda}_i\}); \bar{\lambda}_i\right) \tag{9}$$

wherein $\bar{\theta}_i$ and $\bar{\lambda}_i$ are the relevant subset of θ and λ to the i^{th} residual function $R_i(\bullet)$. $\bar{\theta}_i$ and $\bar{\lambda}_i$ can be defined as:

$$\begin{aligned}
\bar{\theta}_1 &= \{\theta_1, \theta_2\} \quad \text{and} \quad \bar{\lambda}_1 = \{\lambda_1, \lambda_2\}; \quad \text{for} \quad i = 1; \\
\bar{\theta}_{n_s} &= \{\theta_{n_s-1}, \theta_{n_s}\} \quad \text{and} \quad \bar{\lambda}_{n_s} = \{\lambda_{n_s-1}, \lambda_{n_s}\}; \quad \text{for} \quad i = n_s; \\
\bar{\theta}_i &= \{\theta_{i-1}, \theta_i, \theta_{i+1}\}; \quad \text{and} \quad \bar{\lambda}_i = \{\lambda_{i-1}, \lambda_i, \lambda_{i+1}\} \quad \text{elsewhere};
\end{aligned} \tag{10}$$

Similarly,

$$\begin{aligned}
\mathbb{N}_1(t_j; \{\bar{\theta}_1; \bar{\lambda}_1\}) &= [\mathcal{N}_1(t_j; \{\theta_1; \lambda_1\}), \mathcal{N}_2(t_j; \{\theta_2; \lambda_2\})] \quad \text{for} \quad i = 1; \\
\mathbb{N}_{n_s}(t_j; \{\bar{\theta}_{n_s}; \bar{\lambda}_{n_s}\}) &= [\mathcal{N}_{n_s-1}(t_j; \{\theta_{n_s-1}; \lambda_{n_s-1}\}), \mathcal{N}_{n_s}(t_j; \{\theta_{n_s}; \lambda_{n_s}\})] \quad \text{for} \quad i = n_s; \\
\mathbb{N}_i(t_j; \{\bar{\theta}_i; \bar{\lambda}_i\}) &= [\mathcal{N}_{i-1}(t_j; \{\theta_{i-1}; \lambda_{i-1}\}), \mathcal{N}_i(t_j; \{\theta_i; \lambda_i\}), \mathcal{N}_{i+1}(t_j; \{\theta_{i+1}; \lambda_{i+1}\})] \quad \text{elsewhere};
\end{aligned} \tag{11}$$

In the framework proposed in this study, the physics outlined by $\mathcal{F}_i(\bullet)$ essentially illustrates the dynamics of the substructure, which then functions as constraints within the residual function during parameter set optimization. Through effective optimization, the refined parameter sets guarantee adherence to the system dynamics, even at the substructure level.

Step 3: Boundary condition loss:

$$\mathcal{L}_{i,bc}(\{\theta_i; \lambda_i\}) = \left|\mathcal{B}_i(\mathcal{N}_i(t_0; \{\theta_i; \lambda_i\})) - \mathbf{y}(\mathbf{x}_i, t_0)\right|^2 \quad \text{with } \{\mathbf{x}_i, t_0\} \in \Omega_{i,bc} \tag{12}$$

Step 4: The measurement loss is then defined by comparing network output to the measurement data $\mathbf{y}(\mathbf{x}_i, t_j)$ sampled at $\Omega_{i,data}$.

$$\mathcal{L}_{i,data}(\{\theta_i; \lambda_i\}) = \frac{1}{N_{t3}} \sum_{j=1}^{N_{t3}} \left|\mathcal{H}_i(\mathcal{N}_i(t_j; \{\theta_i; \lambda_i\})) - \mathbf{y}(\mathbf{x}_i, t_j)\right|^2 \quad \text{with } \{\mathbf{x}_i, t_j\} \in \Omega_{i,data} \tag{13}$$

Step 5: The overall loss is defined as aggregating losses from all parallel networks as,

$$\mathcal{L}(\{\theta, \lambda\}) = \sum_{i=1}^{n_s} (\alpha \mathcal{L}_{i,int} + (1-\alpha)\mathcal{L}_{i,bc} + (1-\alpha)\mathcal{L}_{i,data}) \tag{14}$$

where α is the weight term. It has to be chosen wisely as the issue of competing terms in the loss function can arise in deep learning training, particularly when dealing with multiple physical constraints or a large number of subdomains. Each term in the loss function represents a different physical constraint or data fidelity term, and thus, they might converge at different rates, meaning one or some terms might dominate the training process. In the proposed framework, we assign a weight α and $(1-\alpha)$ to each term in the loss function, as described in Equation (14), to balance their influences.

Step 6: Finally, the optimization of the neural network and structural parameters $\{\theta, \lambda\}$

$$\{\theta^*, \lambda^*\} = arg \min_{\{\theta,\lambda\}} \mathcal{L}(\{\theta, \lambda\}) \tag{15}$$

Figure 1 illustrates the architecture of the proposed framework for joint state and parameter identification in structural systems. The schematic progresses from left to right. On the far left, the physical system is

modeled as MDOF structure equipped with sensors. Each DOF is associated with an independent neural network $\mathcal{N}_i$, forming the parallel neural network (Parallel NNs) block. Sensor data is organized into three categories, as shown in the "Inputs" block: (i) measurement points $\Omega_{i,\text{data}}$ (pink), (ii) boundary condition points $\Omega_{i,\text{bc}}$ (yellow), and (iii) collocation points for enforcing physics-based constraints $\Omega_{i,\text{int}}$ (green). Each neural network $\mathcal{N}_i$ receives time t as input and outputs the corresponding displacement prediction $\tilde{u}_i(t)$. Although only displacement is directly predicted, velocity $\dot{\tilde{u}}_i(t)$ and acceleration $\ddot{\tilde{u}}_i(t)$ are computed via automatic differentiation (AD) or finite differences (FD), as shown in the "AD/FD" block. This differentiation process establishes a physical continuity chain that inherently links displacement, velocity, and acceleration. The resulting displacement, velocity, and acceleration are then used to evaluate (i) physics-informed residual loss $\mathcal{F}_i$ (green), in addition to (ii) data loss $\mathcal{H}_i$ (pink), based on measured acceleration, and (iii) the boundary/initial condition loss $\mathcal{B}_i$ (yellow). Since velocities and acceleration are derived from predicted displacements, which are implicitly consistent using AD/FD, and used altogether in a physics-informed loss function $\mathcal{F}_i$ (green), it eliminates the need to predict them separately and apply additional scaling factors. Finally, physics loss, data loss, and boundary condition loss are combined using weighting factors and aggregated into a total loss $\mathcal{L}(\theta, \lambda)$. These weighting factors are taken as per user experience and can be automated in the future using self-adaptive tuning techniques [50]. Optimization is further carried out via a feedback loop that simultaneously updates both the neural network parameters θ and the unknown physical parameters λ.

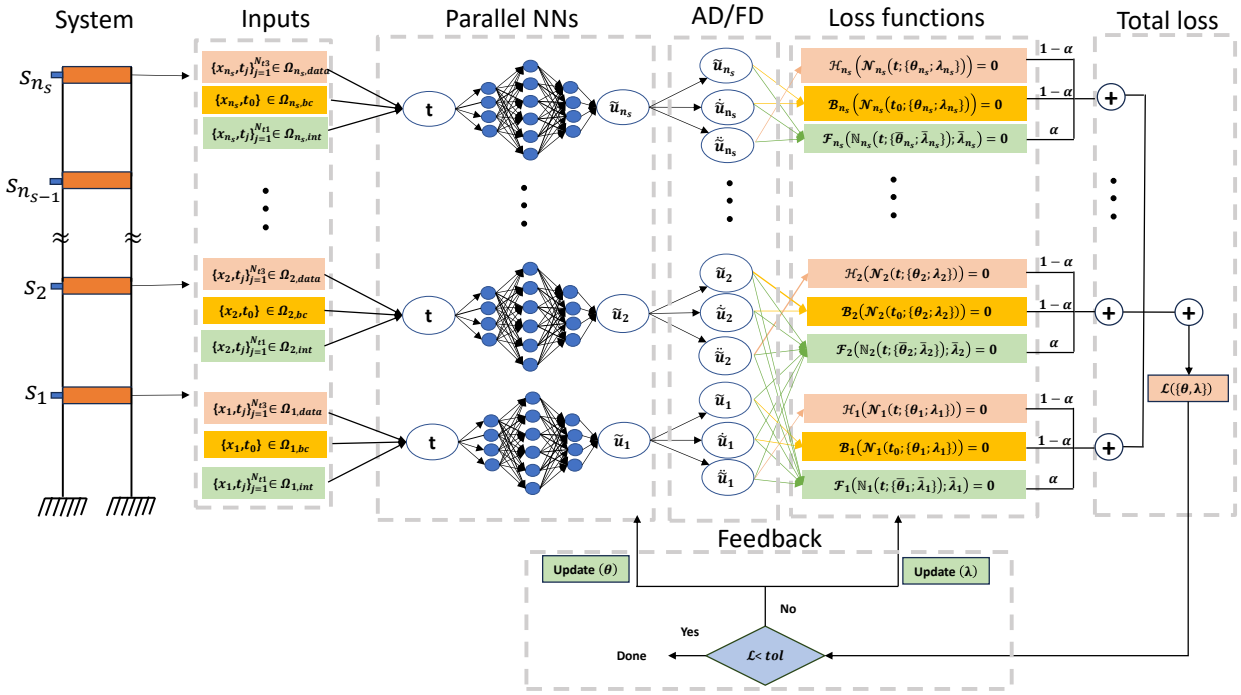


Figure 1: Schematic diagram of the Parallel physics-informed neural networks framework for coupled dynamical systems

The pseudo-code for the above methodology is given in **Algorithm 1**.

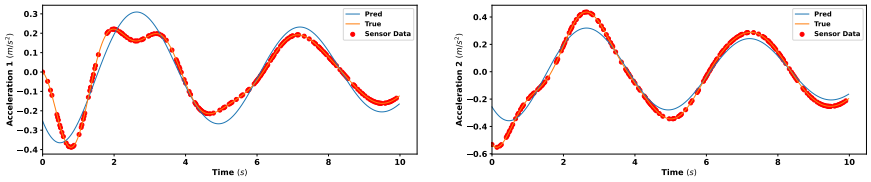
Algorithm 1 Proposed algorithm for Parallel PINNs framework for damage detection, localization and quantification in coupled dynamical system

Initilize:
Divide the structural system into n_s substructures.
Specify the inputs, including measurement points ($\Omega_{i,data} \subset \Omega_{data}$), collocation points ($\Omega_{i,int} \subset \Omega_{int}$), boundary/initial condition points ($\Omega_{i,bc} \subset \Omega_{bc}$).
Specify measurement data $Y_{i,data}$
Create neural network for each substructure ($\mathcal{N}_i(t_j;\{\theta_i;\lambda_i\})$)
Set hyper-parameters (no. of layers, no. of neurons, learning rate) of each neural network.
Initialize the neural network parameters $\{\theta_i, \lambda_i\}$
Set the weight of each component of the total loss in each substructure
Set convergence criteria.
while Convergence not reached **do**
    for $i = 1$ to n_s **do**
        for $j = 1$ to N_{t1} **do**
            Compute $\mathcal{N}_i(t_j;\{\theta_i,\lambda_i\})$, $\dot{\mathcal{N}}_i(t_j;\{\theta_i,\lambda_i\})$, $\ddot{\mathcal{N}}_i(t_j;\{\theta_i,\lambda_i\})$
        end for
        Compute the collocation points loss of i^{th} substructure: $\mathcal{L}_{i,int}$ ▷ Equation (8)
        for $j = 1$ to N_{t2} **do**
            Compute $\mathcal{N}_i(t_j;\{\theta_i,\lambda_i\})$, $\dot{\mathcal{N}}_i(t_j;\{\theta_i,\lambda_i\})$
        end for
        Compute the boundary condition loss of i^{th} substructure: $\mathcal{L}_{i,bc}$ ▷ Equation (12)
        for $j = 1$ to N_{t3} **do**
            Compute $\dot{\mathcal{N}}_i(t_j;\{\theta_i,\lambda_i\})$
        end for
        Compute the data loss of i^{th} substructure: $\mathcal{L}_{i,data}$ ▷ Equation (13)
    end for
    Compute total loss $\mathcal{L}(\theta,\lambda)$ ▷ Equation (14)
    Update parameters $\{\theta,\lambda\}$ by minimizing the total loss $\mathcal{L}(\theta,\lambda)$ using Adam optimizer ▷ Equation (15)
end while
Deployment:
The estimated structural parameters are λ;
The estimated full state of the system $\mathcal{N}_i(t_j;\{\theta_i,\lambda_i\})$, $\dot{\mathcal{N}}_i(t_j;\{\theta_i,\lambda_i\})$, $\ddot{\mathcal{N}}_i(t_j;\{\theta_i,\lambda_i\})$ predicted by trained neural network

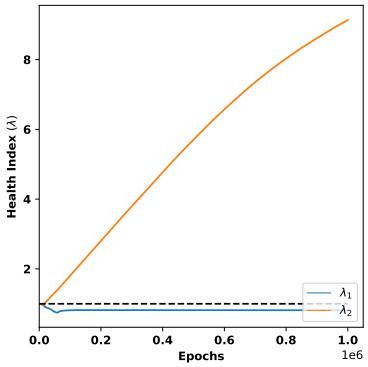
4. Numerical validation

Before exploring the practical application of the proposed approach, it is essential to illustrate the limitations of a conventional PINN approach with a single neural network acting as a global regularizer. To demonstrate this, a 2DOF spring-mass-dashpot linear system is utilized, attempting to estimate the spring stiffnesses from its uncontaminated response to harmonic forcing. It is observed that, even with two coupled equations, the conventional inverse-PINN framework struggles to estimate latent variables and system parameters (refer to Figure 2). This inadequacy can be attributed to overparameterization, highlighting the need for a parallelized framework to address the shortcomings of conventional PINN approaches.

To demonstrate the utility of the developed Parallel PINNs framework, several classical structural systems have been selected as test cases, including one linear and two non-linear MDOF systems. It is important to note that, due to the numerical nature of the validation study, simulated responses are utilized as measurement data, which are intentionally contaminated with realistic levels of noise to mimic real-world scenarios before employing them for estimation. For practical field applications, however, real data would be employed



(a) Linear-2DOF-H-Obs-0



(b) Linear-2DOF-H-Par-0

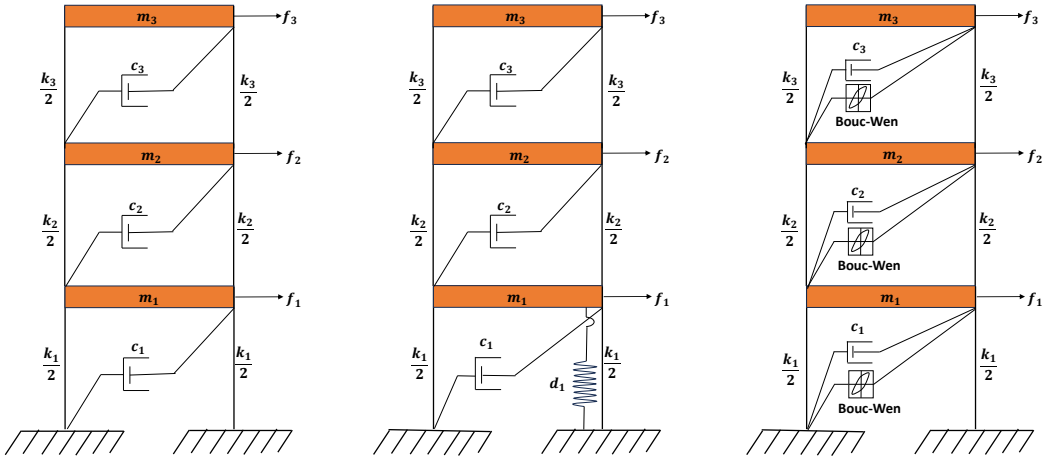
Figure 2: Estimation of observable variable and parameters for linear 2DOF system with conventional inverse PINN *Observable state and parameter estimation are compared against their true values with 0% snr noise*

for system estimation.

The non-linear cases involve the Duffing oscillator and Bouc-Wen model each representing different types of nonlinearity. While all of these models introduce nonlinearity in the system formulation, they differ fundamentally: the Bouc-Wen model excels at simulating asymmetric and rate-dependent hysteresis relevant for structural engineering, while the Duffing oscillator model represent stiffness nonlinearity in various other applications. Incorporating these types of nonlinearity is crucial for demonstrating the comprehensive capabilities of the proposed framework.

As illustrated in the PINN schematics in Figure 1, the proposed framework comprises several components: input data sampled from the substructured domain, parallel neural networks, physics-informed loss functions, and a feedback mechanism. The training process involves the selection of a set of uniformly chosen time instants ($\Omega_{i,int}$ and $\Omega_{i,bc}$), forming a set of time instances $\Omega_{i,phy} := \{\Omega_{i,int}; \Omega_{i,bc}\}$ where the physics of the system (pertinent to the interior and boundary domains, respectively) is enforced during learning. These collocation and boundary points serve as the training inputs for the PINN framework, where the physics and boundary loss for the PINN is determined by assessing the compatibility of the DNN output with the system physics and boundary conditions. While not mandatory for the proposed algorithm, the same sets of timestamps are used for all substructures in this study without affecting the generalization of the framework. Among these timestamps, only a randomly chosen subset ($\Omega_{i,data} \subset \Omega_{i,phy}$) is provided with corresponding measurements $\mathbf{Y}_{i,data}$ ($= \mathbf{y}(\Omega_{i,data})$), resulting in the measurement points and data pair $\mathcal{D} = \{\Omega_{i,data}; \mathbf{Y}_{i,data}\}$.

Before embarking on the numerical experiments, it is important to establish a standardized nomenclature for clarity in presenting plots across three case studies, each corresponding to linear, nonlinear with Duffing oscillator, and nonlinear with Bouc-Wen model, denoted as Linear, Nonlin-D, and Nonlin-BW, respectively. The nomenclature for various plots includes observation (Obs-), Hysteresis (RD-), Phase Portrait (PP-), and Parameter Estimation (Par-) histories, each specific to its corresponding case study. The noise contamination levels are indicated with the *snr* percentages in the plot name, and the type of excitation forces, whether harmonic or earthquake-induced, is denoted with "-H" or "-Eq" in the plots. This systematic approach ensures consistency and facilitates clear interpretation and comparison of results across different scenarios. For e.g., Nonlin-BW-H-PP-5 refers to the Phase Portrait plot of the Bouc-Wen nonlinear system under harmonic excitation with 5% *snr*.



(a) Linear 3DOF (b) Nonlin-D 3DOF (c) Nonlin-BW 3DOF

Figure 3: Schematic diagram of the test structure for (a) 3DOF Linear (b) 3DOF Nonlin-D (c) 3-DOF Nonlin-BW *m, k, c, d and f represent mass, stiffness, damping, cubic nonlinearity and forcing respectively.*

4.1. Numerical case study 1: Linear MDOF system

For the first case study, a linear 3DOF structural system is adopted, represented by a spring-mass-dashpot model. A schematic diagram of the same is presented in Figure 3(a). The generalized equation of motion for the system dynamics can be perceived from the generalized Equation (7) , enforcing the assumed dimensionality. Because of the simplicity of the adopted system, the space can easily be discretized into three substructures without needing a complex FEM. Accordingly, the mass matrix can be idealized with $\mathbf{M} = diag([m_1; m_2; m_3])$, and other system matrices can be defined as follows:

$$\mathbf{K} = k_0 \begin{bmatrix} \lambda_1 + \lambda_2 & -\lambda_2 & 0 \\ -\lambda_2 & \lambda_2 + \lambda_3 & \lambda_3 \\ 0 & -\lambda_3 & \lambda_3 \end{bmatrix}; \mathbf{C} = \begin{bmatrix} c_1 + c_2 & -c_2 & 0 \\ -c_2 & c_2 + c_3 & -c_3 \\ 0 & -c_3 & c_3 \end{bmatrix}; \mathbf{f}(t) = \begin{bmatrix} f_1(t) \\ f_2(t) \\ f_3(t) \end{bmatrix}; \mathbf{u}(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \\ u_3(t) \end{bmatrix}$$

Here, the elements in the mass and damping matrices are considered as $m = m_1 = m_2 = m_3 = 1$ kg, $c_1 = c_2 = c_3 = 0.32$ Ns/m. With an assumption of initial stiffness $k_0 = 30$ N/m, the system is assumed to have deteriorated in its second and third DOF with corresponding values for **HI**s as $\lambda_1 = 1$, $\lambda_2 = 0.833$ $\Big($deterioration of $1/6^{th}$ of original$\Big)$, and $\lambda_3 = 0.667$ $\Big(1/3^{rd}$ of original$\Big)$. The lumped spring-mass-dashpot system is then simulated for $10sec$ with a sampling frequency of $50Hz$ under harmonic force $f_1 = f_2 = f_3 = 0.2\sin(\pi t)$ N.

Within the parallel PINN framework, the data-driven network is composed of three distinct DNNs, each corresponding to a specific substructure and describing the dynamics of an individual mass block. The training inputs for each DNN are generated using 500 uniformly chosen timestamps for $\Omega_{i,int}$ and further selecting 100 randomly chosen timestamps for $\Omega_{i,data}$ from $\Omega_{i,int}$, respectively. The dataset $\mathcal{D}$ is further constructed by sampling $\mathbf{Y}_{i,data}$ at $\Omega_{i,data}$.

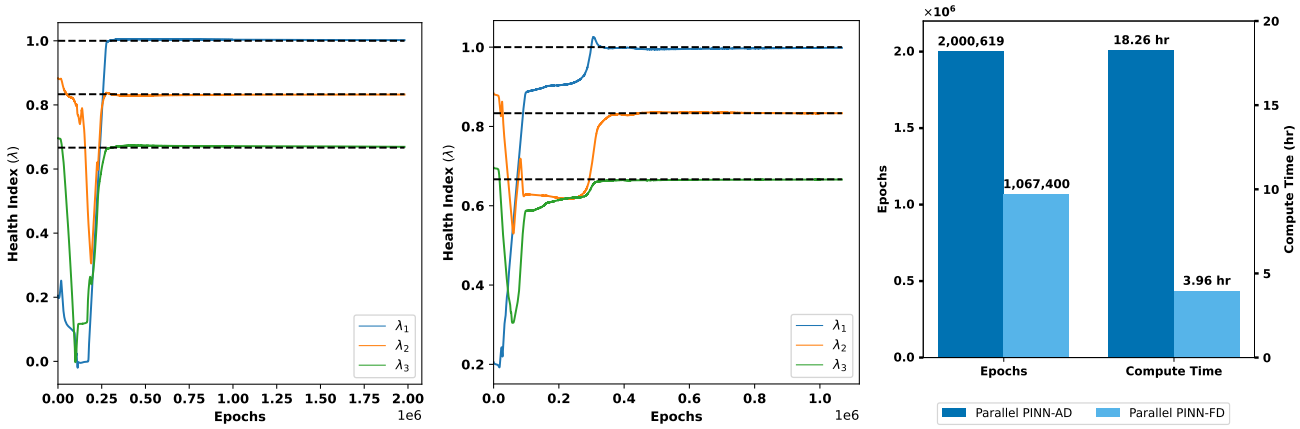
The estimation process initializes λ with arbitrary realizations drawn from a uniform distribution within the range $[0-1]$ for estimating both states and parameters. A neural network architecture with 1 input, 1 output, and 6 hidden layers with 20 neurons in each layer is adopted in all the substructures. This consistency is, however, not mandatory, and the user is free to choose independent networks. A linear and non-linear ($tanh$) activation function is used in the output and remaining layers, respectively. To integrate physics into the inverse estimation process, the latent displacement field $\mathbf{u}(t)$, along with its temporal derivatives $\dot{\mathbf{u}}(t)$ and $\ddot{\mathbf{u}}(t)$, should satisfy the governing physical constraints as defined by Equation (7). This necessitates differentiating the neural network output $\mathbf{u}(t)$ with respect to time, which is achieved through the concurrently trained neural network model. To enforce this constraint in a tractable manner, the derivatives of the displacement field can be obtained either through automatic differentiation (AD) or via numerical approximation such as finite difference (FD) schemes, as shown in Figure 1. Although AD offers machine-precision derivatives and seamless integration into deep learning frameworks, it comes with significant computational cost, as noted in [51]. This aspect has been investigated in this study.

It can be seen from Figure 4(a) and Figure 4(b) that while Parallel PINN-AD and Parallel PINN-FD converge to the true values (indicated by the dashed lines) the FD-based model demonstrates faster convergence. Additionally, Figure 4(c) highlights the substantial computational gain achieved using the Parallel PINN with FD approach, which reduces training time by over 75% compared to the Parallel PINN with AD.

Further, the training procedure utilizes the Adam optimization scheme, incorporating momentum and

adaptive learning rate mechanisms to optimize the parameters of each neural network simultaneously. Parallel PINNs training has been performed in the Python environment using Pytorch on a computer with 12 Intel Core i7-12700K CPUs. The Adam optimizer with a learning rate of 10^{-5} is used to update the parameters of all neural networks. The stopping criteria for the proposed parallel PINNs are set as either reaching a maximum of 5×10^6, epochs, achieving a loss ($\mathcal{L}(\theta, \lambda)$) less than 5×10^{-5}.

Table 1 presents a quantitative evaluation in terms of the mean squared error (MSE) for displacement, velocity, and acceleration predictions, along with the percentage error (PE) in the estimated damage parameter λ. Across all DOFs, the FD-based approach consistently yields lower MSE values and more accurate damage parameter estimates, further validating its efficiency and accuracy.



(a) Parallel PINN-AD (b) Parallel PINN-FD (c) Computational expense

Figure 4: Comparison of computational performance between Parallel PINN-AD and Parallel PINN-FD frameworks

Table 1: Case study: Linear with harmonic forcing — Error analysis of states and parameters for Parallel PINN-AD and Parallel PINN-FD. *(Errors in latent and observable states ($\mathbf{u}$, $\dot{\mathbf{u}}$, $\ddot{\mathbf{u}}$) are reported as Mean Squared Error (MSE), and parameter errors (λ) are reported in percent (PE).)*

DOF	**Parallel PINN-AD**				**Parallel PINN-FD**			
	MSE			**PE**	**MSE**			**PE**
	$\mathbf{u}$	$\dot{\mathbf{u}}$	$\ddot{\mathbf{u}}$	λ	$\mathbf{u}$	$\dot{\mathbf{u}}$	$\ddot{\mathbf{u}}$	λ
1	2.69×10^{-8}	5.45×10^{-7}	1.84×10^{-5}	0.21	5.84×10^{-9}	1.41×10^{-7}	1.05×10^{-5}	0.17
2	4.83×10^{-8}	2.95×10^{-7}	1.71×10^{-5}	0.09	1.16×10^{-8}	1.68×10^{-7}	1.46×10^{-5}	0.04
3	9.46×10^{-8}	6.32×10^{-7}	9.19×10^{-6}	1.36	1.22×10^{-8}	5.58×10^{-8}	5.73×10^{-6}	0.22

4.1.1. Hyperparameter Sensitivity

Although the Parallel PINN-FD framework significantly reduced computational time compared to Parallel PINN-AD, the total training duration of approximately 3.96 hours remains substantial for practical SHM applications. To further accelerate training, a hyperparameter sensitivity study was conducted. The stopping criteria were relaxed by setting the maximum number of training epochs to 5×10^5. Additionally, the learning rate was adjusted to 1×10^{-4} to improve convergence. Among the configurations tested, the

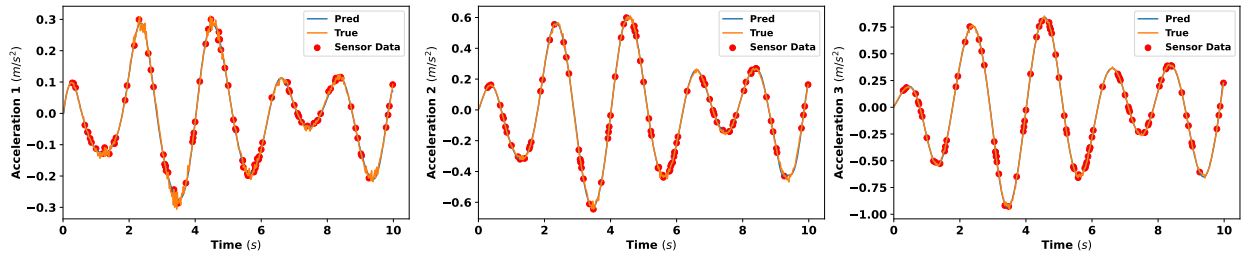
architecture with 3 hidden layers (1−16−32−16−1) yielded the lowest training loss of 1.23×10^{-5} with a significantly reduced training time of approximately 53 minutes. This configuration thus provides an effective balance between training accuracy and computational cost and is adopted for all subsequent studies.

Table 2: Number of layers, network architecture, and corresponding training loss for hyperparameter sensitivity

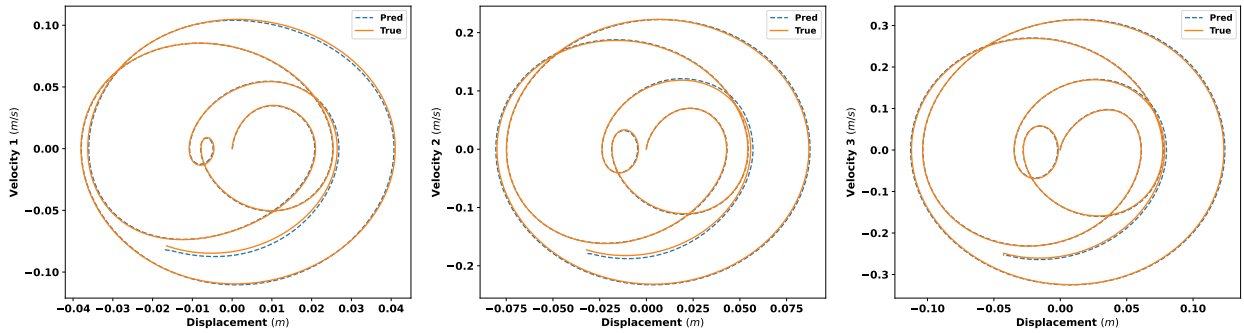
Layers	**Network architecture**	**Training Loss**
2	1-32-32-1	3.92×10^{-5}
3	1-16-32-16-1	1.23×10^{-5}
5	1-16-32-64-32-16-1	4.21×10^{-5}
7	1-16-32-64-128-64-32-16-1	4.16×10^{-5}

4.1.2. Noise sensitivity and forcing variability

The proposed method is further examined for its resilience against noise contamination, wherein the provided data is subjected to contamination with SWGN at signal-to-noise ratios (*snr*) of , 1%, 5%, and 10%. The influence of 5% noise on the quality of the estimation can be assessed based on Figure 5. The parameter estimation for both contaminated signals is additionally shown in Table 3, showcasing the framework's ability to estimate states. Although error in the estimation of state and parameter increases with contamination, the results highlight the effectiveness of the proposed algorithm even in the presence of relatively high noise contamination of 10% *snr*.



(a) Linear-H-Obs-5



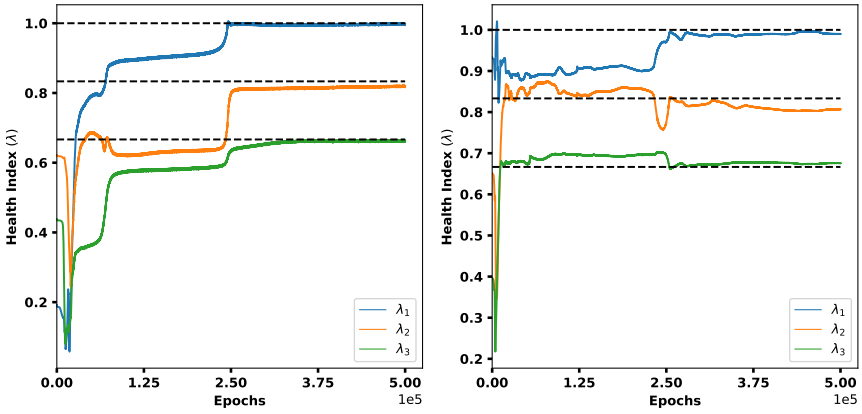
(b) Linear-H-PP-5

Figure 5: Case study: Linear with harmonic forcing- Comparison of actual and Parallel PINN predicted (a) observable, and (b) latent states at 5%

Table 3: Case study: Linear with harmonic forcing — Error analysis of states and parameters for noise levels 1%, 5%, and 10% *snr*. *(Errors in latent and observable states (**u**, **u̇**, **ü**) are reported as Mean Squared Error (MSE), and parameter errors (λ) are reported in percent (PE).)*

DOF	1% Noise				5% Noise				10% Noise			
	MSE			PE	MSE			PE	MSE			PE
	u	**u̇**	**ü**	λ	**u**	**u̇**	**ü**	λ	**u**	**u̇**	**ü**	λ
1	8.82×10^{-5}	1.51×10^{-7}	1.02×10^{-5}	0.19	4.77×10^{-8}	7.72×10^{-7}	8.30×10^{-5}	0.20	1.78×10^{-5}	2.36×10^{-6}	3.00×10^{-4}	2.2
2	4.49×10^{-6}	3.32×10^{-7}	2.19×10^{-5}	1.21	1.89×10^{-7}	1.79×10^{-6}	1.06×10^{-4}	1.58	6.58×10^{-5}	3.68×10^{-6}	3.00×10^{-4}	0.57
3	5.20×10^{-8}	2.52×10^{-7}	1.13×10^{-5}	0.30	2.77×10^{-7}	1.63×10^{-6}	8.18×10^{-5}	0.71	1.21×10^{-4}	4.55×10^{-6}	2.90×10^{-3}	0.82

Furthermore, the system underwent simulation for non-Gaussian, non-stationary force conditions, specifically the EL-Centro earthquake, over a period of 5 seconds at a sampling frequency of 50 Hz. The results corresponding to this numerical case study, presented in Figure 6 and Table 4 demonstrate the quality of state and parameter estimation.



(a) Linear-H-Par-5 (b) Linear-Eq-Par-5

Figure 6: Case study: Linear with (a) harmonic and (b) earthquake excitation - Performance assessment for health parameter estimation
Parameter estimation is conducted under (a) harmonic and (b) earthquake excitation with a 5% snr noise, validating the quality of the estimation within the proposed framework.

Table 4: Case study: Linear with (a) harmonic, and (b) earthquake excitation — Error analysis of states and parameters for noise level of 5%

DOF	Harmonic (5% *snr*)				Earthquake (5% *snr*)			
	MSE			PE	MSE			PE
	u	**u̇**	**ü**	λ	**u**	**u̇**	**ü**	λ
1	4.77×10^{-8}	7.72×10^{-7}	8.30×10^{-5}	0.20	2.08×10^{-6}	3.89×10^{-5}	1.90×10^{-2}	1.67
2	1.89×10^{-7}	1.79×10^{-7}	1.06×10^{-5}	1.58	3.70×10^{-6}	7.81×10^{-5}	2.96×10^{-2}	3.96
3	2.77×10^{-7}	1.63×10^{-7}	8.18×10^{-6}	0.71	9.56×10^{-6}	7.91×10^{-5}	1.83×10^{-2}	1.62

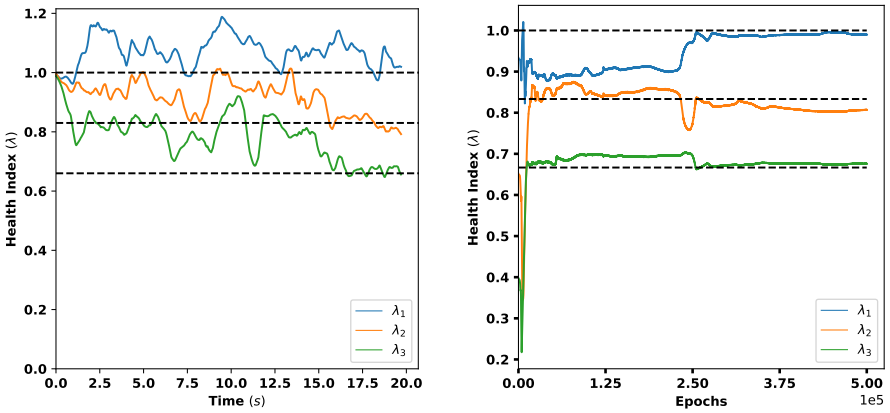
The robustness of the algorithm is evident from the results shown in Figure 5 and Table 3. Despite increased noise levels, the errors in both observable and latent states remain within acceptable bounds. Furthermore, as demonstrated in Figure 6 and Table 4, the algorithm effectively adapts to uncertain external forces. Notably, even under an *snr* as high as 5%, the maximum observed error does not exceed 5%, showing

the algorithm's resilience across a range of forcing conditions.

4.1.3. Comparison with Kalman filtering-based SHM approaches

To benchmark the proposed Parallel PINN-FD framework, a comparison is performed against the classical Kalman filtering-based system identification approach. This serves as a meaningful reference, given the widespread use of Bayesian filtering techniques in handling uncertainty and dynamic and joint state-parameter estimation. Specifically, an Interacting Particle-Ensemble Kalman Filter (IPEnKF) is employed for this comparison.

Linear 3 DOF system is subjected to earthquake force at 50Hz and acceleration data is collected and contaminated with 5% snr SWGN to replicate the real-world situations. The IPEnKF utilizes 20 seconds of measurement data, whereas the proposed Parallel PINN-FD framework requires only 5 seconds of sampled data for training, highlighting its data efficiency. Without delving into the detailed architecture or estimation procedures, the proposed framework exhibits superior stability and smoothness in health index estimation as evident from Figures 7(a) and 7(b). Although the Parallel PINN-FD approach requires slightly more computation time than the IPEnKF, the latter requires extensive tuning of process and measurement noise parameters, posing significant challenges in practical SHM applications. In contrast, the proposed framework demonstrates enhanced robustness to measurement noise and data sparsity, making it more suitable for real-world SHM scenarios.



(a) IPEnKF (b) Parallel PINN-FD

Figure 7: Convergence of health indices λ for: (a) IPEnKF; and (b) Parallel PINN-FD for linear 3-DOF system under earthquake excitation, illustrating the framework's ability to capture changes in system dynamics.

4.2. Numerical case study 2: Duffing MDOF system

The Duffing equation, a nonlinear second-order differential equation, characterizes oscillators exhibiting complex and often chaotic dynamics [52]. Recognized as a canonical model in chaos theory, it continues to serve in diverse applications, such as modeling ship roll dynamics and structural vibrations. In this

study, a Duffing oscillator model is employed (cf. schematic in Figure 3) to demonstrate the efficacy of the proposed estimation framework when applied to higher-order nonlinear systems. Consistent with prior investigations, this nonlinear system is analyzed under both harmonic excitation and stochastic loading, including the El-Centro earthquake. The gde for such systems retains structural similarity to their linear MDOF counterparts, with modifications to account for the system's intrinsic nonlinearity.

$$\mathbf{M}\ddot{\mathbf{u}}(t) + \mathbf{C}\dot{\mathbf{u}}(t) + \mathbf{K}(\lambda)\mathbf{u}(t) + \mathbf{D}\mathbf{u}^3(t) - \mathbf{f}(t) = 0 \tag{16}$$

wherein all other parameters except $\mathbf{D}$ are similar to the earlier case study. $\mathbf{D}$, for the adopted 3DOF nonlinear system, can further be elaborated as follows:

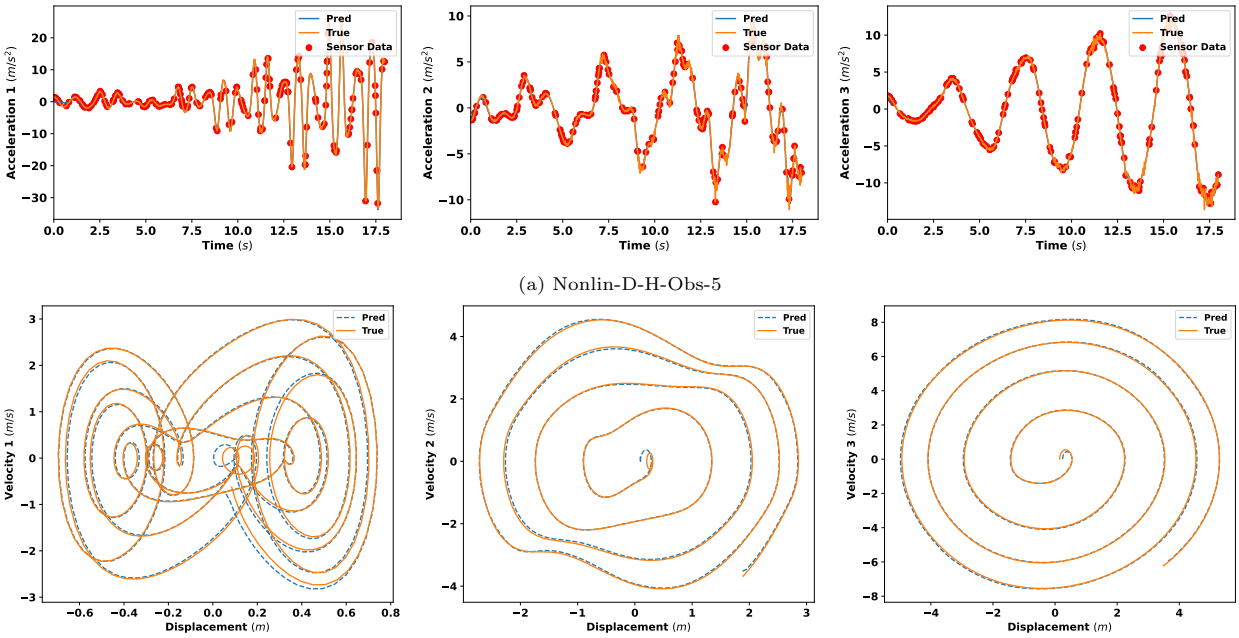
$$\mathbf{D} = \begin{bmatrix} d_1 + d_2 & -d_2 & 0 \\ -d_2 & d_2 + d_3 & -d_3 \\ 0 & -d_3 & d_3 \end{bmatrix};$$

A similar mass matrix used for the linear case is also assumed for the nonlinear case study, while k_0 is assumed to be 20 N/m with $c_1 = c_2 = c_3 = 0.1$ Ns/m. For this case study, the λ is assumed as $\lambda = (1,\ 0.8,\ 0.5)$. For $\mathbf{D}$, cubic nonlinearity coefficients, $d_1,\ d_2$, and d_3, are assumed as $(100,\ 0,\ 0)$. Accordingly, three DNNs are employed in parallel to estimate the systems and parameters from the sampled response.

300 measurement points ($\Omega_{i,data}$) for each of the three DNNs are realized uniformly within an 18-second time window T sampled at a frequency of 50 Hz. For $\Omega_{i,bc}$ initial conditions, pertinent to all DOFs are sampled. $\mathbf{Y}_{i,data}$ is realized by simulating the supporting FEM model for the time window T and subsequently sampling the acceleration responses only at $\Omega_{i,data}$. 5% *snr* SWGN noise is further employed to contaminate the sampled response in order to replicate the real-life noise contamination. This noise contamination is undertaken for all case studies involving different forcing options at a *snr* level of 5% unless otherwise specified. The penalty parameter α is set to 0.1 while the training is performed with the Adam optimizer to estimate λ till the criteria for convergence, i.e. $epochs > 5 \times 10^5$ is achieved. For initialization, λ is randomly sampled from a uniform distribution with the range $[0-1]$. Explicit descriptions of the residual functions for physics and data losses, pertinent $\bar{\theta}$ and $\bar{\lambda}$, are avoided for the sake of brevity. Instead, Equations (8)-(15) defined for a linear MDOF system can be referred to.

Figure 8 offer a detailed insight into the quality of state estimation for the system subjected to harmonic forcing. The impact of noise contamination on estimation quality is perceived to be minimal in the case of harmonic forcing. The corresponding parameter estimates, illustrated in Figure 9 affirm the algorithm's proficiency in promptly and accurately estimating the **HI**s under both harmonic and earthquake forces. Notably, the algorithm demonstrates consistent efficiency even when applied to a higher-order nonlinear system (Nonlin-D) subjected to earthquake excitation. Quantitatively, the percentage error for λ (PE) in parameter estimation remains under 10% for both harmonic and earthquake excitations. This robust performance establishes the methodology's efficacy for generalized, higher-order nonlinear systems under

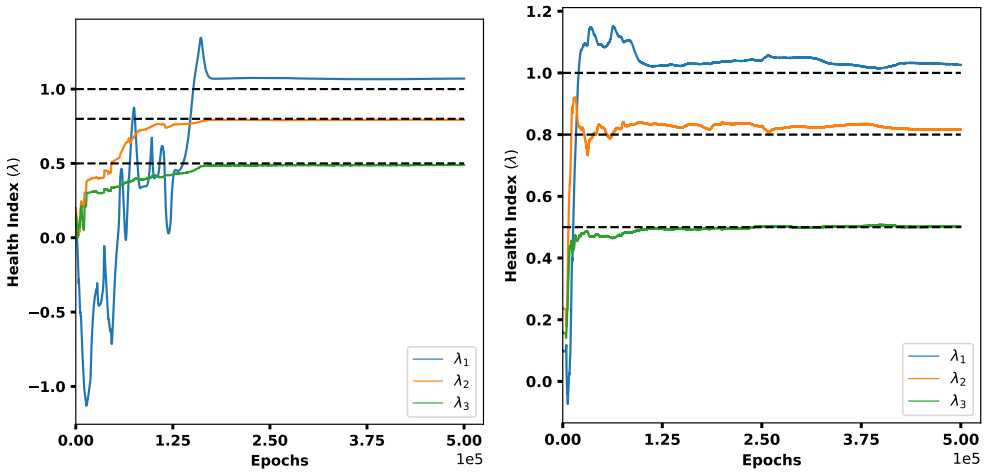
diverse forcing conditions.



(a) Nonlin-D-H-Obs-5

(b) Nonlin-D-H-PP-5

Figure 8: Case study: Nonlin-D with harmonic excitation - Performance assessment for observable and latent state estimation
Observable and latent state estimation is compared against their true values with 5% snr noise, establishing the robustness of the proposed framework against harmonic forcing events.



(a) Nonlin-D-H-Par-5

(b) Nonlin-D-Eq-Par-5

Figure 9: Case study: Nonlin-D with (a) harmonic and (b) earthquake excitation - Performance assessment for health parameter estimation
Parameter estimation is conducted under (a) harmonic and (b) earthquake excitation with a 5% snr noise, validating the quality of the estimation within the proposed framework.

4.3. Bouc-Wen MDOF system

The Bouc-Wen material nonlinear model surpasses the Duffing oscillator in practical applications, particularly in structural engineering. This coupled first-order ODE structure for the Bouc-Wen model is computationally more complex than the single second-order ODE definition of the Duffing oscillator model. Nevertheless, its superior ability to capture a wide range of hysteresis behaviors and complex damping effects makes it more suitable for accurately representing the nonlinear response of structures to dynamic loading. The Bouc-Wen model's flexibility in accommodating diverse material properties makes it a valuable tool for real-world engineering applications, outperforming the Duffing oscillator in practical scenarios. Nevertheless, the intricate representation of nonlinearity in the Bouc-Wen material model requires the proposed method to undergo validation against structures employing this model. This validation is crucial towards evaluating the method's effectiveness in dealing with general nonlinear systems.

To illustrate the proposed framework, a numerical case study undertaken on a 3DOF structural system modeled with Bouc-Wen nonlinearity is presented.

$$\mathbf{M}\ddot{\mathbf{u}}(t) + \mathbf{C}\dot{\mathbf{u}}(t) + \alpha^{bw}\mathbf{K}(\lambda)\mathbf{u}(t) + (1-\alpha^{bw})\mathbf{K}(\lambda)\mathbf{z}(t) = \mathbf{f}(t) \tag{17}$$

with the mass, $\mathbf{M}$, $\mathbf{K}$, and the damping $\mathbf{C}$ being of the same structure as in the previous cases (cf. Equation (7)). $\mathbf{z}(t)$, the hysteretic displacement, with an initial condition of $\mathbf{z}(0) = 0$, can further be elaborated as,

$$\dot{\mathbf{z}}_i^r(t) = \dot{\mathbf{u}}_i^r(t)\left\{A_i - \left[\beta_i \operatorname{sign}(\dot{\mathbf{u}}_i^r(t)\mathbf{z}_i^r(t)) + \gamma_i\right]|\mathbf{z}_i^r(t)|^{n_i^{bw}}\right\} \tag{18}$$

where A_i, γ_i, β_i, α^{bw}, and n_i^{bw} are dimensionless parameters controlling the behavior of the model. The superscripts for z^r and u^r denote they are relatively assessed within an MDOF framework. Pertinent to the current case study, they can be defined through inter-story drift and related non-linear hysteretic displacement as,

$$\begin{aligned} \mathbf{z}_i^r &= \mathbf{z}_i & \dot{\mathbf{u}}_i^r &= \dot{\mathbf{u}}_i & \text{for} \quad & i = 1; \\ \mathbf{z}_i^r &= \mathbf{z}_i - \mathbf{z}_{i-1} & \dot{\mathbf{u}}_i^r &= \dot{\mathbf{u}}_i - \dot{\mathbf{u}}_{i-1} & \text{for} \quad & i = \text{elsewhere}; \end{aligned} \tag{19}$$

This study assumed values for $A_1 = A_2 = A_3 = 1$, $\beta_1 = \beta_2 = \beta_3 = 2$, $\gamma_1 = \gamma_2 = \gamma_3 = 2$, and $n_1^{bw} = n_2^{bw} = n_3^{bw} = 1$.

The simulated system assumes equal masses $m_1 = m_2 = m_3 = 1$ kg and damping coefficients $c_1 = c_2 = 1.5, c_3 = 1.2$ Ns/m, along with a stiffness coefficient of $k_0 = 50$ N/m. The model structure for mass, stiffness, and damping matrices for the system remains the same as in the previous configuration. The simulation is conducted at a sampling frequency of 100 Hz over a time series length of 1000, considering λ as $\lambda = (1,\ 0.9,\ 1)$. Noise contamination level is maintained at 5% *snr*.

Moreover, there is an additional necessity to model $\mathbf{z}(t)$ for the Bouc-Wen model (cf. Equation (17)), and this task is undertaken in a manner similar to other latent variables. It involves the adoption of additional network models $\mathcal{Q}_i(t_j;\bar{\eta}_i)$ for each $z_i^r(t_j)$, defined with the relevant network parameters $\bar{\eta}_i$. Consequently, the residual for such a coupled network takes the form $R_i(t_j,\{\bar{\theta}_i;\bar{\eta}_i;\bar{\lambda}_i\})$ (cf. Equation (9)), where the residual exhibits additional dependency on $\bar{\eta}_i$ due to the modeling of z_i^r.

The further elaboration of R_i will eventually involve $z_i^r(t_j)$s from neighboring s, and accordingly, the pertinent network modules for z_i^r that contribute to defining R_i can be listed within $\mathbb{Q}_i(t_j;\bar{\eta}_i)$ as $\mathbb{Q}_i(t_j;\bar{\eta}_i) = [\mathcal{Q}_i(t_j;\bar{\eta}_i)]$.

$$R_i(t_j,\{\bar{\theta}_i;\bar{\eta}_i;\bar{\lambda}_i\}) = \mathcal{F}_i\left(\mathbb{N}_i(t_j;\{\bar{\theta}_i;\bar{\lambda}_i\}),\mathcal{G}_i(\mathbb{Q}_i(t_j;\bar{\eta}_i));\bar{\lambda}_i\right) \tag{20}$$

Here $\mathcal{G}_i(\bullet)$ represents the hysteritic non-linearity governed by Equation (18). Figure 10 presents the modified Parallel PINN architecture for MDOF systems exhibiting Bouc-Wen-type hysteresis. Following the same procedure as previously employed in this study for training the network, latent variables are estimated, and health parameters are initialized with initial assumptions for λ. These assumptions are generated from a uniform distribution within the range of $[0,1]$.

Figure 11, illustrating observable states, phase portraits of latent variables, and hysteresis forces, and Figure 12 for parameter estimation, solidly establishes the efficient performance of the proposed parallel PINN framework. This holds even for complex nonlinear models characterized by hysteretic forces, where conventional modal domain approaches are inapplicable. The estimation process is remarkably swift, effectively utilizing sparse measurement data and leveraging the intrinsic knowledge of the system's physics. The obtained estimates for both states and parameters consistently fall within acceptable bounds. Quantitatively the percentage error (PE) in parameter estimation does not exceed 10%, holding for both harmonic and earthquake excitations.

4.4. Performance assessment for higher dimensional systems

The proposed framework is finally applied to higher-dimensional systems to explore its applicability in such scenarios. In this experiment, 10-degree-of-freedom (DOF) systems are considered under all three modeling approaches: Linear, Nonlin-D, and Nonlin-BW. Using $\lambda = (1,\ 0.66,\ 1,\ 1,\ 0.66,\ 0.33,\ 1,\ 1,\ 0.83,\ 1)$, the systems are simulated, and the estimation approach is applied similarly to the previous cases. The error values for latent and parameter variables are presented in Table 5, demonstrating that even with strongly nonlinear systems and 5% *snr* noise contamination, the maximum deviation for parameter estimates is around 10%, while the deviations for other variables are well below this maximum error limit. This establishes the efficiency of the proposed approach in dealing with larger systems existing in reality.

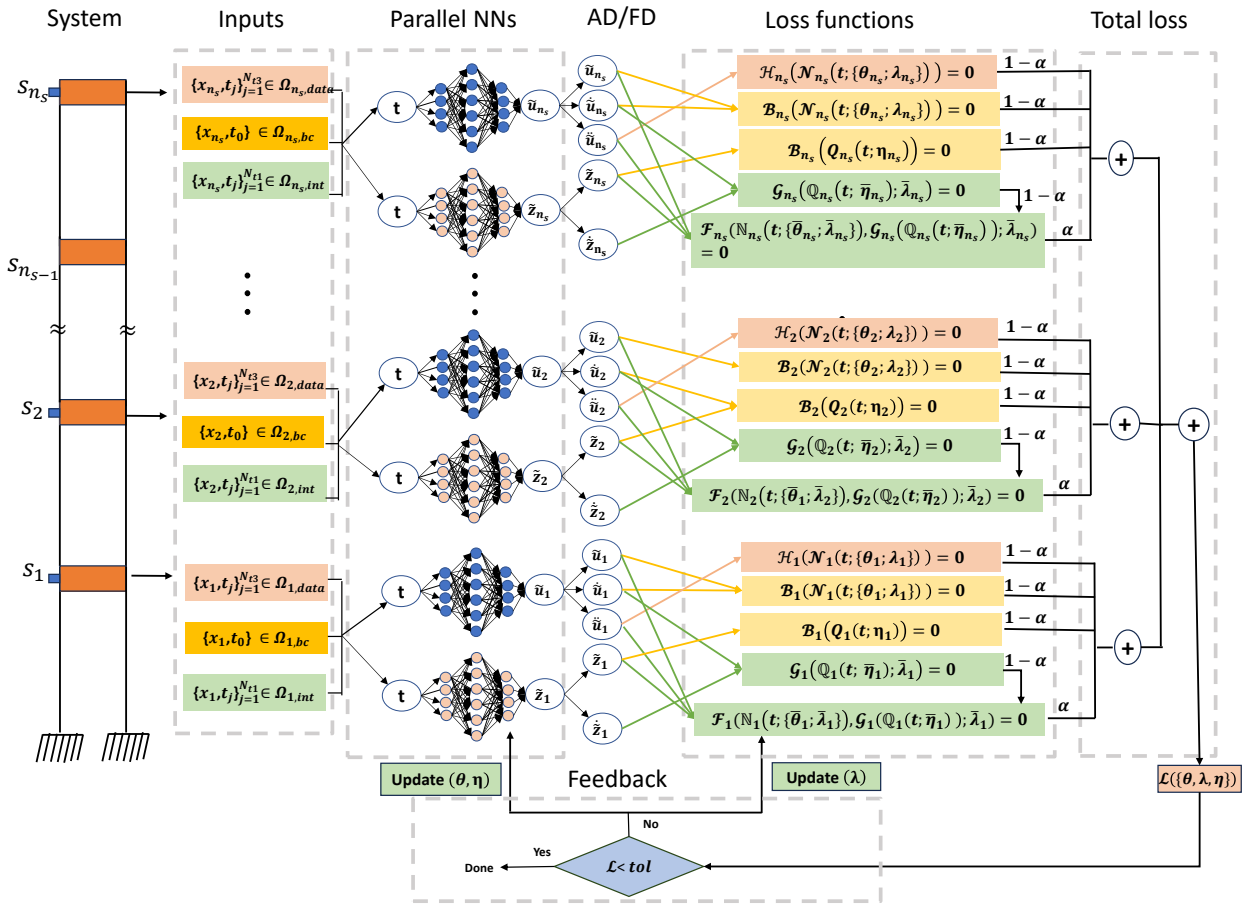


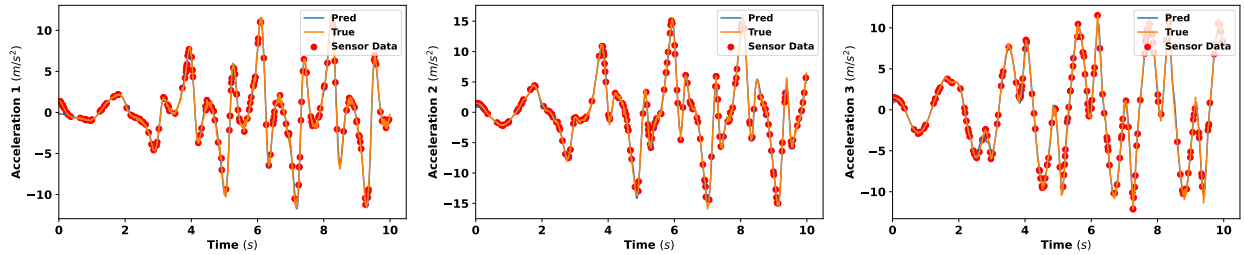
Figure 10: Schematic diagram of the Parallel physics-informed neural networks framework for MDOF system with Bouc-Wen non-linearity

Table 5: Error analysis of states and parameters of Linear 10DOF, Nonlin-D 10DOF, Nonlin-BW 10DOF system with 5% noise subjected to earthquake excitation.

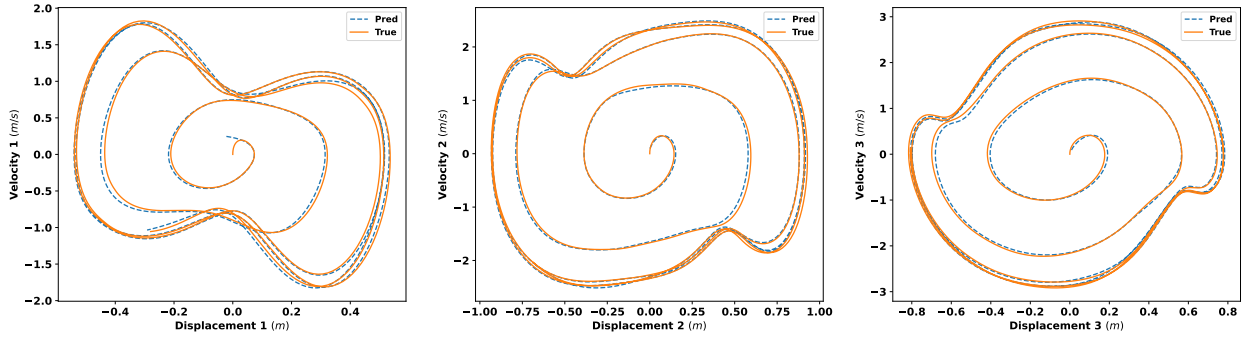
DOF	Linear 10DOF				Nonlin-D 10DOF				Nonlin-BW 10DOF			
	u (m)	$\dot{\mathbf{u}}$ (m/s)	$\ddot{\mathbf{u}}$ (m/s^2)	λ	u (m)	$\dot{\mathbf{u}}$ (m/s)	$\ddot{\mathbf{u}}$ (m/s^2)	λ	u (m)	$\dot{\mathbf{u}}$ (m/s)	$\ddot{\mathbf{u}}$ (m/s^2)	λ
1	4.51×10^{-6}	5.73×10^{-5}	2.41×10^{-2}	1.83	3.04×10^{-3}	3.94×10^{-3}	1.52×10^{-1}	2.21	2.85×10^{-3}	3.20×10^{-3}	1.42×10^{-1}	3.12
2	2.62×10^{-6}	4.98×10^{-5}	1.97×10^{-2}	1.29	2.48×10^{-3}	2.40×10^{-3}	3.70×10^{-1}	1.25	5.17×10^{-4}	1.94×10^{-3}	9.08×10^{-2}	1.89
3	3.86×10^{-6}	6.25×10^{-5}	2.89×10^{-2}	3.12	1.67×10^{-4}	7.99×10^{-4}	1.17×10^{-1}	0.47	4.11×10^{-3}	6.18×10^{-3}	2.73×10^{-1}	4.66
4	6.12×10^{-6}	7.77×10^{-5}	1.91×10^{-2}	1.48	6.88×10^{-3}	5.32×10^{-3}	2.10×10^{-1}	3.85	1.23×10^{-3}	2.91×10^{-3}	1.88×10^{-1}	2.50
5	9.01×10^{-6}	8.22×10^{-5}	2.33×10^{-2}	2.04	4.12×10^{-4}	1.98×10^{-3}	9.77×10^{-2}	0.96	9.94×10^{-4}	4.43×10^{-3}	3.04×10^{-1}	7.35
6	2.95×10^{-6}	3.68×10^{-5}	1.75×10^{-2}	1.21	1.05×10^{-3}	4.75×10^{-3}	1.23×10^{-1}	6.12	6.41×10^{-4}	5.79×10^{-3}	1.10×10^{-1}	2.94
7	7.34×10^{-6}	6.89×10^{-5}	2.64×10^{-2}	2.76	8.91×10^{-4}	6.44×10^{-3}	2.68×10^{-1}	4.48	2.36×10^{-3}	7.81×10^{-3}	1.71×10^{-1}	5.78
8	1.75×10^{-6}	4.13×10^{-5}	2.15×10^{-2}	1.97	3.33×10^{-3}	1.12×10^{-2}	1.79×10^{-1}	9.91	1.01×10^{-3}	2.17×10^{-3}	8.90×10^{-2}	1.43
9	8.20×10^{-6}	7.33×10^{-5}	1.89×10^{-2}	1.58	2.17×10^{-3}	2.58×10^{-3}	6.25×10^{-2}	2.03	3.89×10^{-3}	6.00×10^{-3}	2.21×10^{-1}	6.91
10	5.47×10^{-6}	5.47×10^{-5}	2.74×10^{-2}	3.45	9.78×10^{-4}	1.09×10^{-3}	1.04×10^{-1}	0.83	8.12×10^{-4}	3.57×10^{-3}	1.35×10^{-1}	3.87

5. Experimental validation: Three-story shear frame

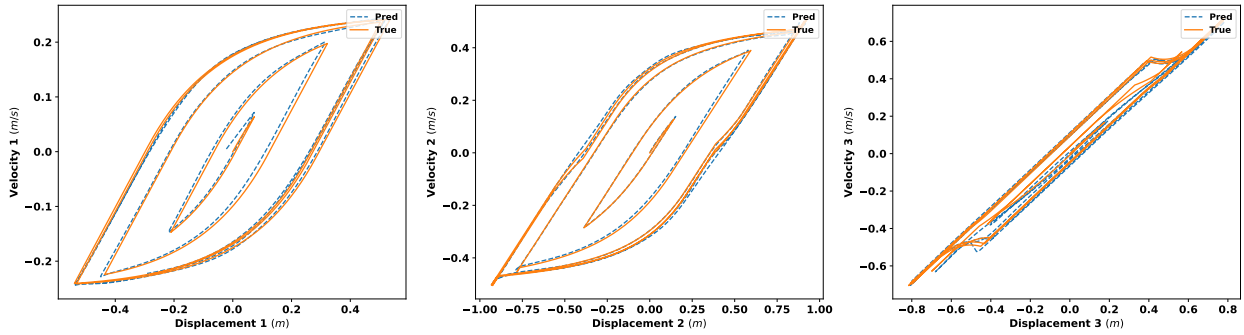
In this study, a laboratory experiment was conducted on a three-story aluminum frame, designed as a shear-building model (see Figure 13), under controlled ambient temperatures ranging from 20°C to 25°C. The



(a) Nonlin-BW-H-Obs-1



(b) Nonlin-BW-H-PP-1



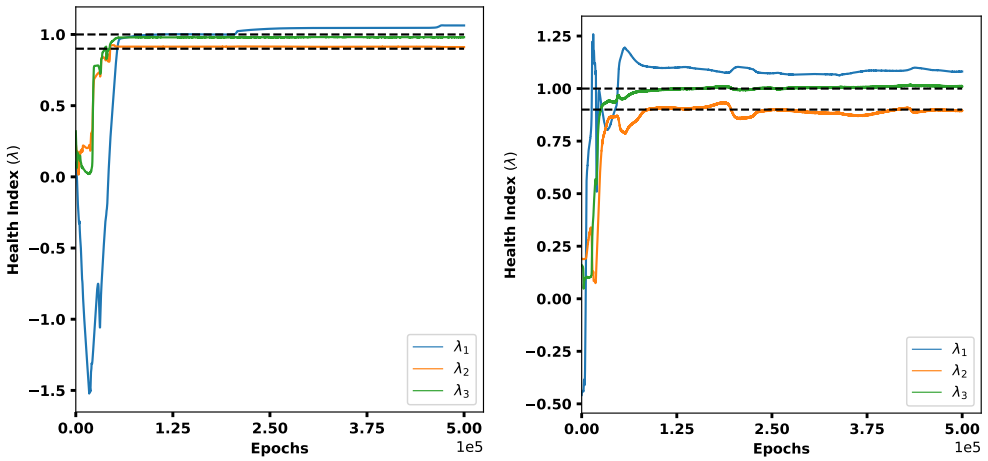
(c) Nonlin-BW-H-RD-1

Figure 11: Case study: Nonlin-BW with harmonic excitation - Performance assessment for observable and latent state estimation

Observable and latent state estimation is compared against their true values with 5% *snr noise establishing the robustness of the proposed framework against harmonic forcing events*

goal was to evaluate the performance of the proposed framework on a scaled structure in a controlled setting. Real-world structures inherently behave as infinite-degree-of-freedom (i-DOFs) systems, where the physics may be only partially known and measurements are often noisy posing a significant challenge for health monitoring. To emulate this complexity, the Parallel PINN framework was employed to infer the behavior of the i-DOFs system using discrete sensor data and a simplified physics-based model. The experiment aimed to assess the framework's ability to produce stable and accurate health parameter estimates, thereby demonstrating its robustness and practical relevance for real-world SHM applications.

The experimental setup replicates an idealized shear building composed of rigid aluminum floors and lightweight columns, designed to emulate real-world structural behavior. The structure is a three-story,



(a) Nonlin-BW-H-Par-5 (b) Nonlin-BW-Eq-Par-5

Figure 12: Case study: Nonlin-BW with (a) harmonic and (b) earthquake excitation - Performance assessment for health parameter estimation
Parameter estimation is conducted under(a) harmonic and (b) earthquake excitation with a 5% snr noise, validating the quality of the estimation within the proposed framework.

single-bay moment frame, standing 1.64 meters tall, with continuous columns (25 mm × 4 mm cross-section) and aluminum floor plates (300 mm x 150 mm × 4 mm) bolted to the columns, forming a lumped-mass system. Floors are spaced at 410 mm intervals. The base is rigidly fixed to a unidirectional shake table using bolts and lock washers.

The shake table, with a platform size of 1000 mm × 1000 mm, supports up to 400 kg, with a displacement stroke of ±100 mm and input excitation up to 1 g across a frequency range of 0–15 Hz. Due to the rigid inter-floor connections, the columns experience fixed-end conditions, enabling a reduced-order 3-DOF lumped mass model. Four uni-axial ICP accelerometers are installed at each floor and the ground level. Earthquake inputs are applied using a servo-hydraulic actuator, and accelerometer data is sampled at 100 Hz via a 4-channel DAQ system using DewesoftX.

To integrate the known physics into the estimation algorithm, a 3-DOF numerical model is developed and calibrated against the real system to marginally match the modal frequencies. This involves tuning material properties (density $\rho = 2700$ kg/m^3, elasticity modulus $E = 68$ GPa). The calibrated model yields mass values $m_1 = m_2 = m_3 = 1.70$ kg and stiffnesses $k_1 = k_2 = k_3 = 3.15$ kN/m, with 1% Rayleigh damping assumed.

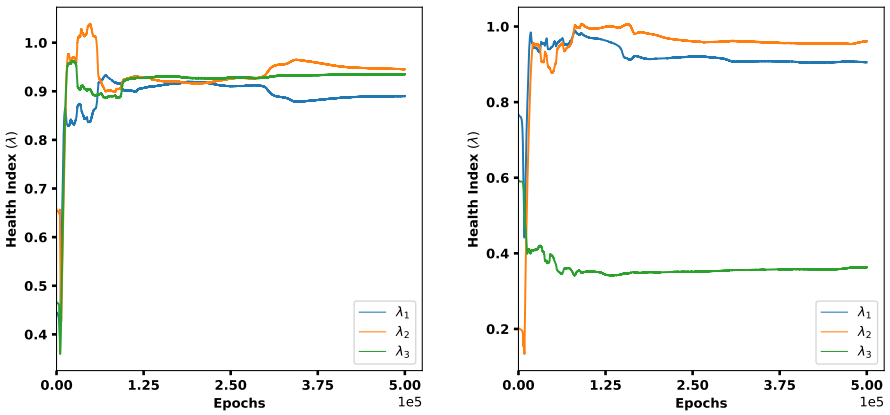
The Parallel PINN framework was then deployed to estimate structural health indices λ_i, initialized randomly. The estimated health indices, $\tilde{\lambda}_i$, are shown in Figure 14(a), and the updated modal properties are reported in Table 6. $\tilde{\lambda}_i$ did not converge to 1, suggesting deviations due to simplifications in the numerical model and boundary assumptions. Consequently, the updated stiffness values reflecting the effective healthy state were identified as $k_1 = 2.84$ kN/m, $k_2 = 2.96$ kN/m, and $k_3 = 2.93$ kN/m.

Figure 13: **Experimental Setup for Damage Detection:** (a) Experimental setup of a 3-story aluminum frame model subjected to EL-Centro earthquake excitation using a shake table. The setup includes a DEWESoft DAQ system with a 4-channel Krypton (ICP) module, accelerometers, and the frame mounted on the shake table. Panels (b) and (c) show close-up views of the frame in the undamaged and damaged conditions, respectively, highlighting the targeted damage location.

Table 6: Comparison of modal frequencies between the experimental (real undamaged) model and the calibrated numerical model. The table presents the first three modal frequencies (ω_1, ω_2, ω_3) in Hz, alongside their corresponding relative errors. The relative error illustrates the accuracy of the numerical model calibration in replicating the real system's dynamic properties.

Frequency (Hz)	**Experimental**	**Numerical**	**Percentage error**
ω_1	3.04	3.05	0.32
ω_2	8.39	8.55	1.90
ω_3	12.69	12.35	2.67

Proposed algorithm's effectiveness in detecting structural damage was validated by analyzing response data from the test structure with damage. Damage was introduced by loosening bolts on diagonal columns at the third story, reducing stiffness to simulate a compromised state. For estimation, structural parameters were maintained as same as the initial numerical model, assuming no prior information on the health state being available for the system, and measurement data from this damaged setup were processed in the Parallel PINN framework. As shown in Figure 14(b), and Table 7 , this induced damage resulted in a 61.43% reduction in third-floor stiffness, while estimation errors for k_1 and k_2 remained below 5% when compared with estimates obtained for updated healthy numerical model, supporting that the estimation results are not contingent on the prior numerical assumptions.



(a) Healthy (b) Damaged

Figure 14: Convergence of health index λ using Parallel PINN framework for shear frame subjected to El-Centro ground motion.

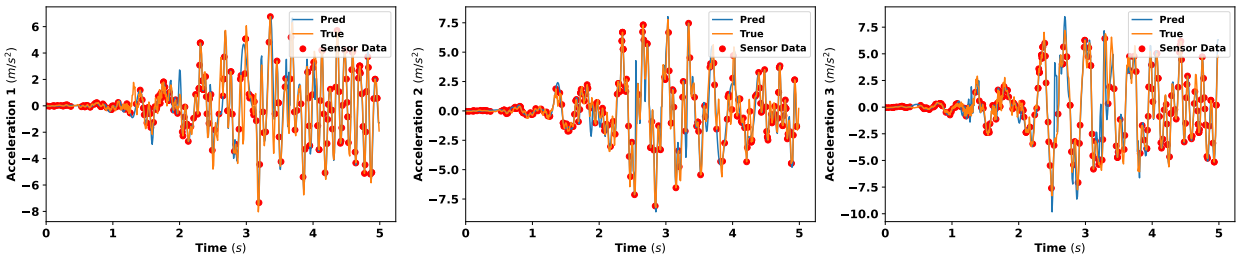


Figure 15: Performance assessment of proposed Parallel PINN-FD for observable state estimation

6. Discussion

The validation results affirm the effectiveness of the proposed parallel PINN framework, which is strategically designed to balance computational efficiency with high-precision parameter estimation. The framework demonstrates consistent performance across a variety of dynamic scenarios, with its accuracy influenced by

Table 7: Comparison of story stiffness values in initial and estimated (healthy and damaged) states for an aluminium frame. The table presents the numerical values of stiffness for each story level (k_1, k_2, and k_3) under two conditions: the estimated (healthy) state and the artificially induced damaged state. The % change in stiffness acts as an indicator of the severity of the damage at each story.

Stiffness (KN/m)	**Numerical (Initial)**	**Estimated (Healthy)**	**Estimated (Damaged)**	**% change in stiffness**
k_1	3.15	2.84	2.82	2
k_2	3.15	2.96	3.00	1.35
k_3	3.15	2.93	1.13	**61.43**

excitation characteristics, measurement noise, system dynamics, and hyperparameter settings. Notably, it exhibits stable convergence and reliable estimations under structured excitations, while maintaining adaptability to more complex input conditions showing its potential for generalized deployment in real-world applications.

To achieve this balance between fidelity and efficiency, the framework integrates finite-difference-based mappings to infer velocity and acceleration from displacement within a parallelized PINN architecture. This structure enables simultaneous training of subnetworks drawing inference from algebraically substructured equations defining the governing physics, each fine-tuned to minimize computational load while accurately estimating local states and health parameters. Global coherence is enforced through shared physics-informed constraints, ensuring that localized learning contributes meaningfully to the full-system solution. A critical factor influencing convergence and model precision is the weighting coefficient α, which balances the data-driven and physics-based loss terms. Empirically, $\alpha = 0.1$ yielded the most favorable results across test cases. While α is currently selected manually, future work may incorporate self-adaptive tuning mechanisms [50] to streamline optimization and further improve training dynamics. Collectively, these architectural and algorithmic choices support a scalable, efficient, and high-fidelity solution for dynamic inverse problems. The practicality of the proposed method is verified experimentally on a three-story aluminum shear frame subjected to seismic-like base excitations. The framework demonstrates precise health assessments even under arbitrary, nonstationary dynamic inputs, with little estimation accuracy loss. These findings highlight the framework's effectiveness in real-world conditions fraught with unpredictable forces and imperfect measurements.

This study focuses on relatively smaller and controlled problems (i.e., lab-scale structures) with the primary objective of assessing the efficiency, robustness, and interpretability of the proposed physics-informed framework. By doing so, it enables developing a clear understanding of the algorithm's capabilities under various levels of uncertainty, such as noise, data sparsity, and damage variability, etc., and how these factors impact estimation accuracy. Such a baseline investigation is essential before deploying the framework in complex real-life scenarios. Authors in [53–57] highlight the pertinent practical challenges encountered in real-world SHM applications when relying solely on data-driven methods. These studies emphasize the need to address issues like environmental variability, sensor malfunctions, asynchronous data, and abnormal

signals. To address these challenges with such data-only approaches often mandates large datasets yielding limited or no interpretability.

As an alternative, the proposed framework, capable of integrating physical laws into the learning process, allows better generalization potential while rendering the overall network more explainable. With its performance and reliability being firmly established through tractable problems in this study, this approach will be extended to more complex, real-world applications in our future research. In this regard, the approaches used for handling data time lag, sensor malfunction, and abnormal signals in [53, 54] can be investigated during the transition to field-scale implementation.

7. Conclusion

Conventional health-assessment methods, whether based on physical models or purely data-driven approaches, often face significant limitations due to the complexity of the models and the high computational cost, or the sparsity and noise in the data. To address these challenges, a deep-learning framework utilizing a parallel PINN-FD architecture has been proposed. This framework incorporates physical constraints by embedding governing equations directly into the loss function, allowing for more accurate and reliable predictions. By reformulating the global MDOF system into coupled DOF-level ODEs, each handled by an independent network and coupled through shared system matrices, the framework efficiently estimates both latent states and health parameters with reduced computational demands. This approach leverages both measurement data and established physical models, enhancing prediction accuracy and reducing the dependency on large amounts of data.

The effectiveness of the proposed framework has been demonstrated through its application to both linear and nonlinear systems, including the Duffing and Bouc-Wen models, under harmonic and earthquake excitations. In all cases, the framework exhibited smooth convergence, strong robustness to noisy data, and accurate parameter estimates. Additionally, experimental validation was performed on a three-story aluminum shear frame, where the framework successfully identified undamaged and damaged stiffness profiles.

While the proposed framework demonstrates strong performance in identifying states and parameters across both linear and nonlinear systems, several extensions can further enhance its scope. Current assumptions on known structural topology and boundary conditions can be relaxed to improve generalizability. As system complexity increases, the parallel network design and coupled residuals raise computational costs, an area targeted for future optimization. Additionally, automating hyperparameter selection and integrating techniques like adaptive sampling, model order reduction, and uncertainty quantification will strengthen scalability for real-world SHM applications.

References

[1] Michèle Basseville, Laurent Mevel, and Maurice Goursat. Statistical model-based damage detection and localization: subspace-based residuals and damage-to-noise sensitivity ratios. *Journal of sound and vibration*, 275(3-5):769–794, 2004.

[2] Michele Basseville, Albert Benveniste, Maurice Goursat, Luc Hermans, Laurent Mevel, and Herman Van der Auweraer. Output-only subspace-based structural identification: from theory to industrial testing practice. *Journal of Dynamic Systems Measurement and Control-Transactions of the Asme*, 123(4), 2001.

[3] Michael Döhler, Laurent Mevel, and Falk Hille. Subspace-based damage detection under changes in the ambient excitation statistics. *Mechanical Systems and Signal Processing*, 45(1):207–224, 2014.

[4] Smriti Sharma and Subhamoy Sen. One-dimensional convolutional neural network-based damage detection in structural joints. *Journal of Civil Structural Health Monitoring*, 10(5):1057–1072, 2020.

[5] Rune Brincker, Lingmi Zhang, and Palle Andersen. Modal identification of output-only systems using frequency domain decomposition. *Smart materials and structures*, 10(3):441, 2001.

[6] Jer-Nan Juang and Richard S Pappa. An eigensystem realization algorithm for modal parameter identification and model reduction. *Journal of guidance, control, and dynamics*, 8(5):620–627, 1985.

[7] Smriti Sharma and Subhamoy Sen. Bridge damage detection in presence of varying temperature using two-step neural network approach. *Journal of Bridge Engineering*, 26(6):04021027, 2021.

[8] Adel Belouchrani, Karim Abed-Meraim, J-F Cardoso, and Eric Moulines. A blind source separation technique using second-order statistics. *IEEE Transactions on signal processing*, 45(2):434–444, 1997.

[9] Gaëtan Kerschen, Fabien Poncelet, and J-C Golinval. Physical interpretation of independent component analysis in structural dynamics. *Mechanical Systems and Signal Processing*, 21(4):1561–1575, 2007.

[10] Yongchao Yang and Satish Nagarajaiah. Time-frequency blind source separation using independent component analysis for output-only modal identification of highly damped structures. *Journal of Structural Engineering*, 139(10):1780–1793, 2013.

[11] Satish Nagarajaiah and Yongchao Yang. Modeling and harnessing sparse and low-rank data structure: a new paradigm for structural dynamics, identification, damage detection, and health monitoring. *Structural Control and Health Monitoring*, 24(1):e1851, 2017.

[12] Jann N Yang, Silian Lin, Hongwei Huang, and Li Zhou. An adaptive extended kalman filter for structural damage identification. *Structural Control and Health Monitoring: The Official Journal of the International Association for Structural Control and Monitoring and of the European Association for the Control of Structures*, 13(4):849–867, 2006.

[13] Eleni N Chatzi and Andrew W Smyth. The unscented kalman filter and particle filter methods for nonlinear structural system identification with non-collocated heterogeneous sensing. *Structural Control and Health Monitoring: The Official Journal of the International Association for Structural Control and Monitoring and of the European Association for the Control of Structures*, 16(1):99–123, 2009.

[14] AE Charalampakis and CK Dimou. Identification of bouc–wen hysteretic systems using particle swarm optimization. *Computers & structures*, 88(21-22):1197–1205, 2010.

[15] Saeed Eftekhar Azam, Eleni Chatzi, and Costas Papadimitriou. A dual kalman filter approach for state estimation via output-only acceleration measurements. *Mechanical systems and signal processing*, 60:866–886, 2015.

[16] Neha Aswal, Subhamoy Sen, and Laurent Mevel. Switching kalman filter for damage estimation in the presence of sensor faults. *Mechanical Systems and Signal Processing*, 175:109116, 2022.

[17] Neha Aswal, Baidurya Bhattacharya, and Subhamoy Sen. Joint and dual estimation of states and parameters with extended and unscented kalman filters. In *Recent Developments in Structural Health Monitoring and Assessment–Opportunities and Challenges: Bridges, Buildings and Other Infrastructures*, pages 223–252. World Scientific, 2022.

[18] Neha Aswal, Subhamoy Sen, and Laurent Mevel. Estimation of local failure in tensegrity using interacting particle-ensemble kalman filter. *Mechanical Systems and Signal Processing*, 160:107824, 2021.

[19] Eshwar Kuncham, Neha Aswal, Subhamoy Sen, and Laurent Mevel. Bayesian monitoring of substructures under unknown interface assumption. *Mechanical Systems and Signal Processing*, 193:110269, 2023.

[20] Neha Aswal, Subhamoy Sen, and Laurent Mevel. Strain-based joint damage estimation approach robust to unknown non-stationary input force. *Structural Control and Health Monitoring*, 29(10):e2999, 2022.

[21] Md Sazzad Hossain, Zhi Chao Ong, Zubaidah Ismail, Siamak Noroozi, and Shin Yee Khoo. Artificial neural networks for vibration based inverse parametric identifications: A review. *Applied Soft Computing*, 52:203–219, 2017.

[22] Calebe Paiva Gomes de Souza, Paulo Roberto Gardel Kurka, Romulo Goncalves Lins, and Jose Medeiros de Araujo Junior. Performance comparison of non-adaptive and adaptive optimization algorithms for artificial neural network training applied to damage diagnosis in civil structures. *Applied Soft Computing*, 104:107254, 2021.

[23] Dehao Liu and Yan Wang. Multi-fidelity physics-constrained neural network and its application in materials modeling. *Journal of Mechanical Design*, 141(12), 2019.

[24] Bin Xu, Zhishen Wu, Genda Chen, and Koichi Yokoyama. Direct identification of structural parameters from dynamic responses with neural networks. *Engineering Applications of Artificial Intelligence*, 17(8):931–943, 2004.

[25] Lei Wang, Haoyu Zhang, Yue Wang, and Di Wu. A data-driven force-thermal coupling load identification method considering multi-source uncertainties of structural characteristics and measuring noises. *Computer Methods in Applied Mechanics and Engineering*, 425:116937, 2024.

[26] Yaru Liu, Lei Wang, and Bing Feng Ng. Multitask-transfer-learning method for random-force frequency identification considering multisource uncertainties. *AIAA Journal*, pages 1–16, 2024.

[27] YC Liang, DP Feng, and JE Cooper. Identification of restoring forces in non-linear vibration systems using fuzzy adaptive neural networks. *Journal of sound and vibration*, 242(1):47–58, 2001.

[28] Luca Facchini, Michele Betti, and Paolo Biagini. Neural network based modal identification of structural systems through output-only measurement. *Computers & Structures*, 138:183–194, 2014.

[29] Maziar Raissi, Paris Perdikaris, and George E Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational*

physics, 378:686–707, 2019.
[30] Ameya D Jagtap, Ehsan Kharazmi, and George Em Karniadakis. Conservative physics-informed neural networks on discrete domains for conservation laws: Applications to forward and inverse problems. *Computer Methods in Applied Mechanics and Engineering*, 365:113028, 2020.
[31] Ehsan Kharazmi, Zhongqiang Zhang, and George Em Karniadakis. hp-vpinns: Variational physics-informed neural networks with domain decomposition. *Computer Methods in Applied Mechanics and Engineering*, 374:113547, 2021.
[32] Ruiyang Zhang, Yang Liu, and Hao Sun. Physics-guided convolutional neural network (phycnn) for data-driven seismic response modeling. *Engineering Structures*, 215:110704, 2020.
[33] Lei Wang, Liaoliao Cheng, and Yaru Liu. Uncertainty-oriented physics-informed long short-term memory (uopi-lstm) network framework for dynamic force identification with interval uncertainties. *Expert Systems with Applications*, 274: 127067, 2025.
[34] Haoyu Zhang, Lei Wang, and Yaru Liu. A hybrid model-based and data-driven method for mechanical-thermal dynamic load identification considering multi-source uncertainties. *Computer Methods in Applied Mechanics and Engineering*, 435: 117662, 2025.
[35] Xiaowei Jin, Shengze Cai, Hui Li, and George Em Karniadakis. Nsfnets (navier-stokes flow nets): Physics-informed neural networks for the incompressible navier-stokes equations. *Journal of Computational Physics*, 426:109951, 2021.
[36] Ehsan Haghighat, Maziar Raissi, Adrian Moure, Hector Gomez, and Ruben Juanes. A physics-informed deep learning framework for inversion and surrogate modeling in solid mechanics. *Computer Methods in Applied Mechanics and Engineering*, 379:113741, 2021.
[37] Renato G Nascimento, Kajetan Fricke, and Felipe AC Viana. A tutorial on solving ordinary differential equations using python and hybrid physics-informed neural network. *Engineering Applications of Artificial Intelligence*, 96:103996, 2020.
[38] Zhilu Lai, Charilaos Mylonas, Satish Nagarajaiah, and Eleni Chatzi. Structural identification with physics-informed neural ordinary differential equations. *Journal of Sound and Vibration*, 508:116196, 2021.
[39] Sarvin Moradi, Burak Duran, Saeed Eftekhar Azam, and Massood Mofid. Novel physics-informed artificial neural network architectures for system and input identification of structural dynamics pdes. *Buildings*, 13(3):650, 2023.
[40] Tong Liu and Hadi Meidani. Physics-informed neural networks for system identification of structural systems with a multiphysics damping model. *Journal of Engineering Mechanics*, 149(10):04023079, 2023.
[41] Xin-Yu Guo and Sheng-En Fang. Structural parameter identification using physics-informed neural networks. *Measurement*, 220:113334, 2023.
[42] JR Jain and TK Kundra. Model based online diagnosis of unbalance and transverse fatigue crack in rotor systems. *Mechanics Research Communications*, 31(5):557–568, 2004.
[43] Renato Giorgiani Nascimento and Felipe AC Viana. Fleet prognosis with physics-informed recurrent neural networks. *arXiv preprint arXiv:1901.05512*, 2019.
[44] Myeongsuk Pak and Sanghoon Kim. A review of deep learning in image recognition. In *2017 4th international conference on computer applications and information processing technology (CAIPT)*, pages 1–3. IEEE, 2017.
[45] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. Multilayer feedforward networks are universal approximators. *Neural networks*, 2(5):359–366, 1989.
[46] Tom Young, Devamanyu Hazarika, Soujanya Poria, and Erik Cambria. Recent trends in deep learning based natural language processing. *ieee Computational intelligenCe magazine*, 13(3):55–75, 2018.
[47] Li Deng, Dong Yu, et al. Deep learning: methods and applications. *Foundations and trends® in signal processing*, 7 (3–4):197–387, 2014.
[48] Ajay Shrestha and Ausif Mahmood. Review of deep learning algorithms and architectures. *IEEE access*, 7:53040–53065, 2019.
[49] Souvik Chakraborty. Simulation free reliability analysis: A physics-informed deep learning based approach. *arXiv preprint arXiv:2005.01302*, 2020.
[50] Levi McClenny and Ulisses Braga-Neto. Self-adaptive physics-informed neural networks using a soft attention mechanism. *arXiv preprint arXiv:2009.04544*, 2020.
[51] Kart Leong Lim, Rahul Dutta, and Mihai Rotaru. Physics informed neural network using finite difference method. In *2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 1828–1833. IEEE, 2022.
[52] Hamel. Georg duffing, ingenieur: Erzwungene schwingungen bei veränderlicher eigenfrequenz und ihre technische bedeutung. sammlung vieweg. heft 41/42, braunschweig 1918. vi+ 134 s, 1921.
[53] Xiaonan Zhang, Youliang Ding, Hanwei Zhao, Letian Yi, Tong Guo, Aiqun Li, and Yang Zou. Mixed skewness probability modeling and extreme value predicting for physical system input–output based on full bayesian generalized maximum-likelihood estimation. *IEEE Transactions on Instrumentation and Measurement*, 73:1–16, 2023.
[54] Hanwei Zhao, Youliang Ding, Libo Meng, Zuowei Qin, Fan Yang, and Aiqun Li. Bayesian multiple linear regression and new modeling paradigm for structural deflection robust to data time lag and abnormal signal. *IEEE Sensors Journal*, 23 (17):19635–19647, 2023.
[55] Zi-xiang Yue, You-liang Ding, and Han-wei Zhao. Deep learning-based minute-scale digital prediction model of temperature-induced deflection of a cable-stayed bridge: Case study. *Journal of Bridge Engineering*, 26(6):05021004, 2021.
[56] Manya Wang, Youliang Ding, and Hanwei Zhao. Digital prediction model of temperature-induced deflection for cable-stayed bridges based on learning of response-only data. *Journal of Civil Structural Health Monitoring*, 12(3):629–645, 2022.
[57] Hanwei Zhao, Youliang Ding, Aiqun Li, Wei Sheng, and Fangfang Geng. Digital modeling on the nonlinear mapping between multi-source monitoring data of in-service bridges. *Structural Control and Health Monitoring*, 27(11):e2618,

702 2020.