

Adaptive Learning of Periodic Solutions for Nonlinear Systems by Physics-informed Gaussian Process

Yixin Li

Sun Yat-sen University

Zhong-Rong Lu

Sun Yat-sen University

Dahao Yang

Sun Yat-sen University

Zechang Zheng

University of Liege

Jike Liu

Sun Yat-sen University

Li Wang

wangli75@mail.sysu.edu.cn

Sun Yat-sen University

Research Article

Keywords: Physics-informed Gaussian process (PIGP), Nonlinear dynamics, Adaptive collocation, Periodic kernel function, Harmonic balance method (HBM)

Posted Date: September 3rd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10615567/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Adaptive Learning of Periodic Solutions for Nonlinear Systems by Physics-informed Gaussian Process

Yixin Li¹, Zhong-Rong Lu¹, Dahao Yang¹, Zechang Zheng², Jike Liu¹, and Li Wang^{*,1}

¹School of Aeronautics and Astronautics, Sun Yat-sen University, Shenzhen Campus, Shenzhen 518107, P.R. China

²Laboratory Vibration of Turbomachines, University of Liege, Belgium

Abstract

This paper presents a novel physics-informed Gaussian process (PIGP) for the adaptive computation of periodic solutions to nonlinear dynamical systems. The proposed approach is built upon two key steps. First, the Gaussian process (GP), as a Bayesian non-parametric machine learning framework, is employed to approximate the periodic solution without requiring a predetermined set of basis functions. To this end, tailored physics-informed periodic kernel and mean functions are constructed for GP regression. The nonlinear governing equation is then collocated to generate observational data, and in conjunction with an iterative linearization scheme, the periodic solution is obtained with both mean and covariance information. Second, recognizing that the performance of GP regression critically depends on the selection of collocation points, an adaptive collocation strategy based on variance- or residual-based criterion is developed to enable accurate and efficient solving of nonlinear systems. Numerical examples are conducted to validate the accuracy, robustness, and efficiency of the proposed PIGP. Notably, under the same level of accuracy, the PIGP demonstrates superior computational efficiency compared to the widely-recognized harmonic balance method (HBM).

Keywords: Physics-informed Gaussian process (PIGP); Nonlinear dynamics; Adaptive collocation; Periodic kernel function; Harmonic balance method (HBM)

1 Introduction

Nonlinear dynamical systems are ubiquitous in modern science and engineering [1]. In structural engineering, for instance, geometric nonlinearity, material nonlinearity, and interfacial friction or contact are prevalent sources of nonlinearity that pose significant challenges to analysis and design [2]. Unlike linear systems, nonlinear systems often exhibit complex dynamic behaviors, including instability, bifurcations, and multiple coexisting solutions. In particular, obtaining periodic solutions is essential for characterizing these behaviors and thus constitutes a central task in uncovering the intrinsic laws governing nonlinear dynamical systems.

Existing methods for computing periodic solutions of nonlinear systems can be broadly classified into two categories: analytical and numerical (including semi-analytical) approaches. On the one hand, well-known analytical methods include the perturbation method, the Lindstedt–Poincaré method, and the method of multiple scales [3, 4]. These techniques are generally effective for weakly nonlinear systems [5], but often fail to produce satisfactory results when applied to strongly nonlinear problems. On the other hand, numerical methods, such as the Runge-Kutta (RK) method [6], the shooting method [7] and the harmonic balance method (HBM) [8, 9], are capable of handling strongly

*Correspondence to: Li Wang, Email: wangli75@mail.sysu.edu.cn

nonlinear systems with high robustness and have therefore been widely employed in practical engineering computations. Among them, the RK method is versatile and widely used in various nonlinear dynamic problems; however, it cannot capture unstable solutions and typically requires prohibitively long integration times to reach the steady-state response. To improve efficiency, the shooting method incorporates the periodicity condition into the RK process over a single prescribed period and updates the initial state to obtain the periodic solution. While this method can capture unstable solutions and is well-suited for non-smooth systems [10, 11], it often necessitates repeated time integrations, which become computationally intensive, particularly for large-scale systems. In contrast, the HBM approximates the periodic solution using Fourier series and solves for the Fourier coefficients within a Galerkin regime [12]. The HBM is typically computationally inexpensive for smooth systems, since low-order Fourier approximations often suffice to obtain accurate solutions. Further enhancements have been introduced to advance the applicability of the HBM to non-smooth systems: the alternating frequency–time (AFT) scheme [13] enables fast computation of Fourier coefficients for general and complex nonlinear terms, and the event-driven Gauss quadrature [14] further accelerates the AFT process, especially for non-smooth systems. The main limitation of the HBM, however, is that its convergence rate deteriorates significantly for strongly non-smooth systems [15], requiring a large number of Fourier terms to achieve the desired accuracy, which becomes computationally costly.

In recent years, a new data-driven and machine-learning paradigm has emerged as an alternative tool for solving ordinary and partial differential equations, including dynamical systems [16, 17]. Within this paradigm, physics-informed machine learning methods have been progressively developed and have demonstrated unique advantages across multiple fields, including fluid mechanics [18], solid mechanics [19, 20] and linear and nonlinear differential equations [21–23]. Representative examples include the physics-informed neural network (PINN) and the physics-informed Gaussian process (PIGP) [18, 24, 25]. In principle, the PINN directly incorporates physical constraint equations into its training process and constructs a loss function based on both physical information (e.g., differential governing equations) and observational data [23]. This allows the network to find solutions that adhere to physical laws, even with limited data [26]. However, the PINN suffers from issues such as unstable training, high computational cost, and a lack of interpretability. By contrast, the PIGP is a non-parametric Bayesian method [18], that inherently integrates physical constraints into the Gaussian mean and kernel functions, and subsequently combines them with training data to yield the posterior mean and covariance analytically. Consequently, the PIGP offers several distinctive advantages:

1. *Interpretability*: The PIGP is formulated within the Bayesian framework, and the posterior mean and covariance are derived analytically [24].
2. *Uncertainty Quantification*: It naturally quantifies the solution uncertainty by providing the posterior covariance, which is crucial for robust decision-making [27].
3. *Fewer parameters*: Compared to the PINN, the PIGP significantly reduces the number of trainable parameters. For typical problems, it involves only a handful of hyperparameters, primarily from the kernel and mean functions, greatly simplifying the optimization process [28, 29].
4. *Tractable Optimization*: Leveraging its Bayesian foundation, optimization can be efficiently performed via maximum likelihood estimation (MLE), which is easily monitored and exhibits strong stability [30].

Owing to these advantages, the PIGP has attracted increasing attention in various theoretical and engineering fields [31, 32], including the solution of differential equation. However, current applications of the PIGP are primarily focused on linear problems [33, 34], and remain limited to nonlinear dynamical systems [35], let alone for periodic solutions of these systems.

In this work, the PIGP is extensively developed for computing periodic solutions of general nonlinear dynamical systems. In contrast to conventional analytical and numerical methods, the PIGP is

a data-driven approach in which the nonlinear governing equations, evaluated at selected collocation points, are treated as observational data, and the solution is then iteratively inferred via Bayesian formulation. The realization of the PIGP relies on three key components. First, a tailored periodic kernel function is constructed to enable accurate representation of periodic solutions. Additionally, a mean function that incorporates Fourier terms at the fundamental frequency is introduced to facilitate direct extraction of the corresponding Fourier coefficients. Second, the nonlinear system is solved iteratively, with the system linearized at each iteration so that the PIGP can provide the desired posterior mean and covariance of the solution. The parameters in the mean and kernel are simultaneously optimized at each iteration via maximum likelihood estimation. Third, since either the posterior covariance or the expectation of the squared residual can effectively serve as an indicator of solution error, the collocation points can be adaptively and incrementally selected based on these error metrics, thereby enabling an efficient and adaptive learning process. Consequently, compared to the well-established HBM, the proposed PIGP is found to offer superior efficiency for a given level of accuracy, along with the distinct advantage of inherent uncertainty quantification for the obtained solution.

The remainder of this paper is organized as follows. Section 2 briefly states the problem for periodic solution of nonlinear dynamical systems and reviews the HBM with addressing of its limitations. Section 3 presents the detailed formulation of the proposed PIGP, including the linearization procedure, the construction of periodic GP kernel, the Bayesian inference framework, and the adaptive collocation strategy. Section 4 provides comprehensive numerical examples to demonstrate the accuracy, robustness, and efficiency of the proposed method, along with comparisons to the HBM. Finally, conclusions are drawn in Section 5.

2 Problem statement

Consider the nonlinear dynamical system that was governed by the following equation,

$$\mathbf{R}(\mathbf{x}(t), t) = \mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{C}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) + \mathbf{f}_{\text{nl}}(\mathbf{x}(t), \dot{\mathbf{x}}(t)) - \mathbf{F}(t) = \mathbf{0} \quad (1)$$

where $\mathbf{x}(t) \in \mathbb{R}^{n \times 1}$ denotes the displacement response, $\dot{\mathbf{x}}(t)$ and $\ddot{\mathbf{x}}(t)$ represent the respective velocity and acceleration responses and n denotes the number of degrees of freedom (DOFs) of the system. $\mathbf{M}, \mathbf{C}, \mathbf{K} \in \mathbb{R}^{n \times n}$ are the matrices of mass, damping and linear stiffness, respectively. $\mathbf{f}_{\text{nl}}(\mathbf{x}, \dot{\mathbf{x}}) \in \mathbb{R}^{n \times 1}$ is the nonlinear restoring force, encompassing all the nonlinear terms of the system. $\mathbf{F}(t)$ represents the periodic external excitation with period T , i.e., $\mathbf{F}(t + T) = \mathbf{F}(t)$, and therefore, the steady-state response $\mathbf{x}(t)$ the system is generally periodic with the fundamental frequency $\omega = \frac{2\pi}{T}$.

To get the periodic solution, the HBM is usually adopted. It employs trigonometric functions as basis functions [37] and approximates the solution by a truncated Fourier series $\mathbf{x}_H(t)$, i.e.,

$$\mathbf{x}_H(t) = \mathbf{a}_0 + \sum_{k=1}^H [\mathbf{a}_k \cos(k\omega t) + \mathbf{b}_k \sin(k\omega t)] \quad (2)$$

where $\hat{\mathbf{x}}_H = [\mathbf{a}_0; \mathbf{a}_1; \mathbf{b}_1; \dots; \mathbf{a}_H; \mathbf{b}_H]$ is the vector of unknown Fourier coefficients and H denotes the truncation order, giving a total of $N_H = 2H + 1$ harmonic terms. Then, by substituting this approximation into the governing equation (1) and applying the Galerkin procedure, the following harmonic balance equation can be obtained,

$$\begin{aligned} \frac{1}{T} \int_0^T \mathbf{R}(\mathbf{x}_H(t), t) dt &= 0, \\ \frac{2}{T} \int_0^T \cos(k\omega t) \mathbf{R}(\mathbf{x}_H(t), t) dt &= 0, \frac{2}{T} \int_0^T \sin(k\omega t) \mathbf{R}(\mathbf{x}_H(t), t) dt = 0, \quad k = 1, \dots, H, \end{aligned} \quad (3)$$

or in the concise form,

$$\mathbf{S}(\omega) \hat{\mathbf{x}}_H + \hat{\mathbf{f}}_{\text{nl}, H} - \hat{\mathbf{F}}_H = \mathbf{0} \quad (4)$$

where $\mathbf{S}(\omega)$ is the dynamic stiffness matrix of the linear terms in the equation, $\hat{\mathbf{f}}_{\text{nl},H}$ and $\hat{\mathbf{F}}_H$ denote the Fourier coefficients of the nonlinear term $\mathbf{f}_{\text{nl}}(\mathbf{x}_H, \dot{\mathbf{x}}_H)$ and the external excitation $\mathbf{F}(t)$. Specifically, $\hat{\mathbf{f}}_{\text{nl},H}$ is expressed as

$$\hat{\mathbf{f}}_{\text{nl},H} = \left[\frac{1}{T} \int_0^T \mathbf{f}_{\text{nl}}(\mathbf{x}_H, \dot{\mathbf{x}}_H) dt; \dots; \frac{2}{T} \int_0^T \cos(k\omega t) \mathbf{f}_{\text{nl}}(\mathbf{x}_H, \dot{\mathbf{x}}_H) dt; \right. \\ \left. \frac{2}{T} \int_0^T \sin(k\omega t) \mathbf{f}_{\text{nl}}(\mathbf{x}_H, \dot{\mathbf{x}}_H) dt; \dots \right]. \quad (5)$$

The solution can then be obtained by iteratively solving the algebraic equation (4). As was established in Ref. [15], the HBM converges exponentially with respect to H for smooth systems, while the convergence rate degrades significantly to polynomial order for non-smooth systems. In this way, the HBM is quite efficient for smooth systems, as a low order H suffices to achieve accurate solution. However, the following limitations exist for the HBM:

- For non-smooth systems, achieving the desired level of accuracy often requires a prohibitively large number of harmonic terms, making the computation costly;
- There is still a lack of clear criteria to select an appropriate harmonic order H when dealing with different problems [36], especially for non-smooth systems. Though in practice H can be determined based on extensive trial calculations or prior experience, the process is often costly and almost no explicit indicator is available to measure the solution error.

To address the above limitations, a novel data-driven approach, named physics-informed Gaussian process (PIGP), is proposed, as elaborated in Section 3. It directly provides the covariance information of the solution, which can serve as an explicit metric to indicate the accuracy. Benefit from the adaptivity feature of the PIGP, accurate solution is acquired by using as few data points as possible, leading to improved efficiency.

3 Physics-informed Gaussian process (PIGP) for periodic solution

3.1 Linearization for iterative solution

Note that the Gaussian process (GP) is well applicable to solve linear equations [28]. Thus, before applying the GP, the nonlinear problem (1) should be linearized. This linearization coincides perfectly with the Newton iteration for solving nonlinear dynamical equations [39, 40]. Specifically, at the $(k+1)$ th iteration, the solution $\mathbf{x}_k(t)$ is readily obtained and to get $\mathbf{x}_{k+1}(t)$, the nonlinear term $\mathbf{f}_{\text{nl}}(\mathbf{x}, \dot{\mathbf{x}})$ is linearized through first-order Taylor expansion at $(\mathbf{x}_k, \dot{\mathbf{x}}_k)$, leading to

$$\mathbf{f}_{\text{nl}}(\mathbf{x}, \dot{\mathbf{x}}) \approx \mathbf{f}_{\text{nl}}(\mathbf{x}_k, \dot{\mathbf{x}}_k) + \mathbf{J}_x^{(k)}(t) \cdot (\mathbf{x} - \mathbf{x}_k) + \mathbf{J}_{\dot{x}}^{(k)}(t) \cdot (\dot{\mathbf{x}} - \dot{\mathbf{x}}_k) \quad (6)$$

where $\mathbf{J}_x^{(k)}, \mathbf{J}_{\dot{x}}^{(k)}$ are the Jacobian matrices of the following form

$$\mathbf{J}_x^{(k)}(t) = \left. \frac{\partial \mathbf{f}_{\text{nl}}}{\partial \mathbf{x}} \right|_{(\mathbf{x}, \dot{\mathbf{x}}) = (\mathbf{x}_k, \dot{\mathbf{x}}_k)} \in \mathbb{R}^{n \times n}, \\ \mathbf{J}_{\dot{x}}^{(k)}(t) = \left. \frac{\partial \mathbf{f}_{\text{nl}}}{\partial \dot{\mathbf{x}}} \right|_{(\mathbf{x}, \dot{\mathbf{x}}) = (\mathbf{x}_k, \dot{\mathbf{x}}_k)} \in \mathbb{R}^{n \times n}. \quad (7)$$

Substituting Eq. (6) into the original equation (1), a linearized time-varying dynamical equation is obtained as follows,

$$\mathbf{M}\ddot{\mathbf{x}} + \mathbf{C}_k(t)\dot{\mathbf{x}} + \mathbf{K}_k(t)\mathbf{x} = \mathbf{f}_k(t) \quad (8)$$

where the terms $\mathbf{C}_k(t)$, $\mathbf{K}_k(t)$, $\mathbf{f}_k(t)$ are defined as,

$$\begin{aligned}\mathbf{C}_k(t) &= \mathbf{C} + \mathbf{J}_{\dot{\mathbf{x}}}^{(k)}(t), \\ \mathbf{K}_k(t) &= \mathbf{K} + \mathbf{J}_{\mathbf{x}}^{(k)}(t), \\ \mathbf{f}_k(t) &= \mathbf{F}(t) - \left[\mathbf{f}_{nl}(\mathbf{x}_k, \dot{\mathbf{x}}_k) - \mathbf{J}_{\mathbf{x}}^{(k)}(t)\mathbf{x}_k - \mathbf{J}_{\dot{\mathbf{x}}}^{(k)}(t)\dot{\mathbf{x}}_k \right].\end{aligned}\tag{9}$$

For notational convenience in the subsequent analysis, define the linear operator as

$$\mathcal{L}_t^{(k)}[\cdot] = \mathbf{M} \frac{d^2}{dt^2}(\cdot) + \mathbf{C}_k(t) \frac{d}{dt}(\cdot) + \mathbf{K}_k(t)(\cdot)\tag{10}$$

and then, the linearized dynamical equation (8) is expressed compactly as

$$\mathcal{L}_t^{(k)}[\mathbf{x}(t)] = \mathbf{f}_k(t).\tag{11}$$

3.2 PIGP for solving linearized dynamical equation

The GP is a non-parametric Bayesian machine learning method built upon a rigorous statistical foundation. It is particularly well-suited for handling high-dimensional data and small-sample scenarios, offering high accuracy and strong robustness [27]. For an arbitrary scalar (random) GP $y(t)$, its behavior is fully characterized by the mean function $\mu_y(t)$ and the kernel function $k_y(t, t')$, namely,

$$y(t) \sim \mathcal{GP}(\mu_y(t), k_y(t, t'))\tag{12}$$

where $\mu_y(t) = \mathbb{E}[y(t)]$, $k_y(t, t') = \text{Cov}(y(t), y(t')) = \mathbb{E}[(y(t) - \mu_y(t))(y(t') - \mu_y(t'))]$, $\mathbb{E}[\cdot]$ is the expectation operator and $\text{Cov}(\cdot, \cdot)$ means the covariance. If $y(t)$ is constrained to be periodic with fundamental frequency ω , the mean function $\mu_y(t)$ should be also periodic and the following kernel function is frequently used [28],

$$k_y(t, t') = \sigma^2 \exp \left(- \frac{4 \sin^2 \left(\frac{\omega(t - t')}{2} \right)}{l^2} \right)\tag{13}$$

where σ^2, l are the hyperparameters of variance and characteristic length scale. A detailed derivation of this periodic kernel function (13) is referred to [Appendix A](#).

Now, turn to the linearized dynamical equation (11). To get the periodic solution, $\mathbf{x}(t)$ is assumed pertaining to the following GP prior,

$$\mathbf{x}(t) \sim \mathcal{GP}(\boldsymbol{\mu}(t), \mathbf{k}(t, t'))\tag{14}$$

where $\boldsymbol{\mu}(t) \in \mathbb{R}^n = \mathbb{E}[\mathbf{x}(t)]$ is the mean function and $\mathbf{k}(t, t') = \text{Cov}(\mathbf{x}(t), \mathbf{x}(t')) \in \mathbb{R}^{n \times n}$ is the periodic kernel function. Herein, the mean function is set to encompass the fundamental frequency terms,

$$\boldsymbol{\mu}(t) = \mathbf{a} \cos(\omega t) + \mathbf{b} \sin(\omega t)\tag{15}$$

with $\mathbf{a}, \mathbf{b}$ being unknown parameters that directly encode the amplitude information at the fundamental frequency. Moreover, different DOFs of $\mathbf{x}(t)$ are assumed to be uncorrelated. Consequently, $\mathbf{k}(t, t')$ takes a diagonal form, with each diagonal entry given by the periodic kernel in Eq. (13),

$$\begin{aligned}\mathbf{k}(t, t') &= \text{diag} [k_1(t, t'), k_2(t, t'), \dots, k_n(t, t')]; \\ k_i(t, t') &= \sigma_i^2 \exp \left(- \frac{4 \sin^2 \left(\frac{\omega(t - t')}{2} \right)}{l_i^2} \right), i = 1, 2, \dots, n\end{aligned}\tag{16}$$

where $\sigma_i^2, l_i, i = 1, 2, \dots, n$ are the hyperparameters associated with the i th DOF, remaining to be determined.

Next, as a data-driven approach, the GP requires observational data for regression. Herein, the data is obtained by collocating the linearized governing equation (11) at a set of N time points $\mathbb{T}_c = \{t_1 < t_2 < \dots < t_N\}$ with the periodic interval $[0, T]$, i.e.,

$$\mathbf{f}_k(\mathbb{T}_c) := \begin{pmatrix} \mathbf{f}_k(t_1) \\ \mathbf{f}_k(t_2) \\ \vdots \\ \mathbf{f}_k(t_N) \end{pmatrix} = \mathcal{L}_t^{(k)}[\mathbf{x}](\mathbb{T}_c) + \boldsymbol{\epsilon} := \begin{pmatrix} \mathcal{L}_t^{(k)}[\mathbf{x}]|_{t=t_1} \\ \mathcal{L}_t^{(k)}[\mathbf{x}]|_{t=t_2} \\ \vdots \\ \mathcal{L}_t^{(k)}[\mathbf{x}]|_{t=t_N} \end{pmatrix} + \boldsymbol{\epsilon} \quad (17)$$

where $\mathbf{f}_k(\mathbb{T}_c)$ is the observational data and $\mathcal{L}_t^{(k)}[\mathbf{x}](\mathbb{T}_c)$ is the corresponding model prediction. Since $\mathbf{x}(t)$ is generally an approximate solution, a discrepancy/error $\boldsymbol{\epsilon}$ arises between the data and the model prediction. This error $\boldsymbol{\epsilon}$ is assumed independent from the solution $\mathbf{x}$ and is modelled as zero-mean Gaussian noise, i.e.,

$$\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \sigma_\epsilon^2 \mathbf{I}) \quad (18)$$

where σ_ϵ^2 is the noise variance and $\mathbf{I}$ is the identity matrix. Then, to infer the solution from the observational data, the following joint Gaussian distribution between $\mathbf{x}(t)$ and $\mathbf{f}_k(\mathbb{T}_c)$ is established

$$\begin{bmatrix} \mathbf{x}(t) \\ \mathbf{f}_k(\mathbb{T}_c) \end{bmatrix} \sim \mathcal{N} \left(\begin{pmatrix} \boldsymbol{\mu}(t) \\ \boldsymbol{\mu}_f(\mathbb{T}_c) \end{pmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{x}\mathbf{x}}(t, t) & \boldsymbol{\Sigma}_{\mathbf{x}\mathbf{f}}(t, \mathbb{T}_c) \\ \boldsymbol{\Sigma}_{\mathbf{f}\mathbf{x}}(\mathbb{T}_c, t) & \boldsymbol{\Sigma}_{\mathbf{f}\mathbf{f}}(\mathbb{T}_c, \mathbb{T}_c) \end{bmatrix} \right) \quad (19)$$

where the cross-covariance satisfies $\boldsymbol{\Sigma}_{\mathbf{x}\mathbf{f}}(t, \mathbb{T}_c) = \boldsymbol{\Sigma}_{\mathbf{f}\mathbf{x}}(\mathbb{T}_c, t)^\top$. The constituent mean and covariance components are derived from the GP prior in Eq. (14) and are detailed as follows:

- Mean of the solution,

$$\boldsymbol{\mu}(t) = \mathbb{E}[\mathbf{x}(t)] = \mathbf{a} \cos(\omega t) + \mathbf{b} \sin(\omega t).$$

- Mean of the observational data,

$$\boldsymbol{\mu}_f(\mathbb{T}_c) = \mathbb{E}[\mathbf{f}_k(\mathbb{T}_c)] = \mathbb{E}[\mathcal{L}_t^{(k)}[\mathbf{x}](\mathbb{T}_c)] = \mathcal{L}_t^{(k)}[\boldsymbol{\mu}](\mathbb{T}_c) = \mathbf{D}_c(\mathbb{T}_c)\mathbf{a} + \mathbf{D}_s(\mathbb{T}_c)\mathbf{b}$$

where $\mathbf{D}_c(\mathbb{T}_c) = \mathcal{L}_t^{(k)}[\cos(t)](\mathbb{T}_c)$ and $\mathbf{D}_s(\mathbb{T}_c) = \mathcal{L}_t^{(k)}[\sin(t)](\mathbb{T}_c)$.

- Covariance of the solution,

$$\boldsymbol{\Sigma}_{\mathbf{x}\mathbf{x}}(t, t) = \text{Cov}(\mathbf{x}(t), \mathbf{x}(t)) = \mathbf{k}(t, t).$$

- Cross-covariance between the data and the solution,

$$\boldsymbol{\Sigma}_{\mathbf{f}\mathbf{x}}(\mathbb{T}_c, t) = \text{Cov}(\mathbf{f}_k(\mathbb{T}_c), \mathbf{x}(t)) = \begin{pmatrix} \text{Cov}(\mathcal{L}_t^{(k)}[\mathbf{x}]|_{t=t_1}, \mathbf{x}(t)) \\ \text{Cov}(\mathcal{L}_t^{(k)}[\mathbf{x}]|_{t=t_2}, \mathbf{x}(t)) \\ \vdots \\ \text{Cov}(\mathcal{L}_t^{(k)}[\mathbf{x}]|_{t=t_N}, \mathbf{x}(t)) \end{pmatrix}$$

where each entry is given by

$$\begin{aligned} \text{Cov}(\mathcal{L}_t^{(k)}[\mathbf{x}]|_{t=t_i}, \mathbf{x}(t)) &= \text{Cov}([\mathbf{x}(s)], \mathbf{x}(t))|_{s=t_i} = \mathcal{L}_s^{(k)} \text{Cov}(\mathbf{x}(s), \mathbf{x}(t))|_{s=t_i} = \mathcal{L}_s^{(k)} \mathbf{k}(s, t)|_{s=t_i} \\ &= [\mathbf{M} \frac{d^2}{ds^2} \mathbf{k}(s, t) + \mathbf{C}_k(s) \frac{d}{ds} \mathbf{k}(s, t) + \mathbf{K}_k(s) \mathbf{k}(s, t)]|_{s=t_i}. \end{aligned}$$

- Covariance of the observational data,

$$\Sigma_{\mathbf{f}\mathbf{f}}(\mathbb{T}_c, \mathbb{T}_c) = \text{Cov}(\mathbf{f}_k(\mathbb{T}_c), \mathbf{f}_k(\mathbb{T}_c)) = \begin{pmatrix} \mathbf{A}(t_1, t_1) & \mathbf{A}(t_1, t_2) & \cdots & \mathbf{A}(t_1, t_N) \\ \mathbf{A}(t_2, t_1) & \mathbf{A}(t_2, t_2) & \cdots & \mathbf{A}(t_2, t_N) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}(t_N, t_1) & \mathbf{A}(t_N, t_2) & \cdots & \mathbf{A}(t_N, t_N) \end{pmatrix} + \sigma_\epsilon^2 \mathbf{I}$$

where

$$\begin{aligned} \mathbf{A}(s, t) &= \text{Cov}(\mathcal{L}_s^{(k)}[\mathbf{x}(s)], \mathcal{L}_t^{(k)}[\mathbf{x}(t)]) = \mathcal{L}_s^{(k)} \text{Cov}(\mathbf{x}(s), \mathbf{x}(t)) (\mathcal{L}_t^{(k)})^\top, \\ &= \mathcal{L}_s^{(k)} \mathbf{K}(s, t) (\mathcal{L}_t^{(k)})^\top = (a_{ij}(s, t))_{(i=1:n, j=1:n)}, \\ a_{ij}(s, t) &= \sum_{p=1}^n \left(M_{ip} \frac{d^2}{ds^2} + C_{k,ip}(s) \frac{d}{ds} + K_{k,ip}(s) \right) \left(M_{jp} \frac{d^2}{dt^2} + C_{k,jp}(t) \frac{d}{dt} + K_{k,jp}(t) \right) k_p(s, t) \end{aligned}$$

with $M_{ij}, C_{k,ij}(t), K_{k,ij}(t)$ denoting the (i, j) th entry of the respective matrices $\mathbf{M}, \mathbf{C}_k(t), \mathbf{K}_k(t)$ and $k_p(s, t)$ readily defined in Eq. (16).

With the joint distribution established in Eq. (19), the conditional Gaussian distribution of the solution $\mathbf{x}_{k+1}(t)$ under the observational data $\mathbf{f}_k(\mathbb{T}_c)$ is obtained via Bayesian formulation [38],

$$\mathbf{x}(t) \mid \mathbf{f}_k(\mathbb{T}_c) \sim \mathcal{N}(\boldsymbol{\mu}_{k+1}(t), \Sigma_{k+1}(t, t)) \quad (20)$$

where $\boldsymbol{\mu}_{k+1}(t)$ and $\Sigma_{k+1}(t, t)$ are the posterior mean and variance of the target solution with the following expressions,

$$\begin{aligned} \boldsymbol{\mu}_{k+1}(t) &= \boldsymbol{\mu}(t) + \Sigma_{\mathbf{x}\mathbf{f}}(t, \mathbb{T}_c) \Sigma_{\mathbf{f}\mathbf{f}}(\mathbb{T}_c, \mathbb{T}_c)^{-1} (\mathbf{f}_k(\mathbb{T}_c) - \boldsymbol{\mu}_{\mathbf{f}}(\mathbb{T}_c)), \\ \Sigma_{k+1}(t, t) &= \Sigma_{\mathbf{x}\mathbf{x}}(t, t) - \Sigma_{\mathbf{x}\mathbf{f}}(t, \mathbb{T}_c) \Sigma_{\mathbf{f}\mathbf{f}}(\mathbb{T}_c, \mathbb{T}_c)^{-1} \Sigma_{\mathbf{f}\mathbf{x}}(\mathbb{T}_c, t). \end{aligned} \quad (21)$$

The posterior mean $\boldsymbol{\mu}_{k+1}(t)$ represents the most probable solution and is set to the updated solution at the $(k+1)$ th iteration, i.e., $\mathbf{x}_{k+1}(t) = \boldsymbol{\mu}_{k+1}(t)$. Concurrently, the posterior covariance $\Sigma_{k+1}(t, t)$ quantifies the uncertainty of the solution and can serve as a useful error indicator. As is noteworthy, the physical constraints of the nonlinear dynamical problem are directly embedded in GP-based solution process: the periodicity condition is enforced via the selection of the periodic GP kernel (16), while the governing equation is incorporated as the observational data. This integration naturally motivates the term *physics-informed Gaussian process (PIGP)*. However, before obtaining the posterior mean and covariance in Eq. (21), two issues must be addressed:

- Determination of the hyperparameters $\mathbf{h} = \{\sigma_i^2, l_i, i = 1, 2, \dots, n; \sigma_\epsilon^2\}$ which consist of the kernel variance and length-scale parameters as well as the noise variance, and estimation of the mean-function parameters $\mathbf{a}, \mathbf{b}$. These are accomplished by maximizing the marginal likelihood derived from Eq. (19),

$$\begin{aligned} \mathbf{f}_k(\mathbb{T}_c) &\sim \mathcal{N}(\boldsymbol{\mu}_{\mathbf{f}}(\mathbb{T}_c), \Sigma_{\mathbf{f}\mathbf{f}}(\mathbb{T}_c, \mathbb{T}_c)) \quad \text{or} \\ p(\mathbf{f}_k(\mathbb{T}_c)) &= \frac{1}{(2\pi)^{nN/2} |\Sigma_{\mathbf{f}\mathbf{f}}(\mathbb{T}_c, \mathbb{T}_c)|^{1/2}} e^{-\frac{1}{2} (\mathbf{f}_k(\mathbb{T}_c) - \boldsymbol{\mu}_{\mathbf{f}}(\mathbb{T}_c))^\top \Sigma_{\mathbf{f}\mathbf{f}}(\mathbb{T}_c, \mathbb{T}_c)^{-1} (\mathbf{f}_k(\mathbb{T}_c) - \boldsymbol{\mu}_{\mathbf{f}}(\mathbb{T}_c))}. \end{aligned} \quad (22)$$

Maximizing the log-marginal likelihood in Eq. (22) with respect to the hyperparameters $\mathbf{h}$ and the mean parameters $\mathbf{a}, \mathbf{b}$ yields their optimal estimates. Specifically, the two set of parameters are tackled in an alternating manner. For the hyperparameters, there is

$$\begin{aligned} \mathbf{h}^{opt} &= \arg \max_{\mathbf{h}} p(\mathbf{f}_k(\mathbb{T}_c)) = \arg \min_{\mathbf{h}} -\ln p(\mathbf{f}_k(\mathbb{T}_c)) \\ &= \arg \min_{\mathbf{h}} \left\{ \frac{1}{2} \ln |\Sigma_{\mathbf{f}\mathbf{f}}(\mathbb{T}_c, \mathbb{T}_c)| + \frac{1}{2} (\mathbf{f}_k(\mathbb{T}_c) - \boldsymbol{\mu}_{\mathbf{f}}(\mathbb{T}_c))^\top \Sigma_{\mathbf{f}\mathbf{f}}(\mathbb{T}_c, \mathbb{T}_c)^{-1} (\mathbf{f}_k(\mathbb{T}_c) - \boldsymbol{\mu}_{\mathbf{f}}(\mathbb{T}_c)) \right\}, \end{aligned} \quad (23)$$

which be quickly solved by the conventional search method, e.g., the `fminsearch` function in MATLAB. Note that a good choice of initial hyperparameters $\mathbf{h}$ can well accelerate the searching process and in this way, those readily obtained in the preceding iteration, i.e., the k th iteration, are well utilized as the initial guess for minimization in the $(k+1)$ th iteration. While for the mean parameters, the maximization can be performed analytically

$$\begin{aligned}
[\mathbf{a}; \mathbf{b}]^{opt} &= \arg \max_{[\mathbf{a}; \mathbf{b}]} p(\mathbf{f}_k(\mathbb{T}_c)) \\
&= \arg \min_{[\mathbf{a}; \mathbf{b}]} \left\{ \frac{1}{2} (\mathbf{f}_k(\mathbb{T}_c) - (\mathbf{D}_c(\mathbb{T}_c)\mathbf{a} + \mathbf{D}_s(\mathbb{T}_c)\mathbf{b}))^\top \boldsymbol{\Sigma}_{\mathbf{f}\mathbf{f}}(\mathbb{T}_c, \mathbb{T}_c)^{-1} (\mathbf{f}_k(\mathbb{T}_c) - (\mathbf{D}_c(\mathbb{T}_c)\mathbf{a} + \mathbf{D}_s(\mathbb{T}_c)\mathbf{b})) \right\} \\
&= \left(\begin{bmatrix} \mathbf{D}_c(\mathbb{T}_c)^\top \\ \mathbf{D}_s(\mathbb{T}_c)^\top \end{bmatrix} \boldsymbol{\Sigma}_{\mathbf{f}\mathbf{f}}(\mathbb{T}_c, \mathbb{T}_c)^{-1} [\mathbf{D}_c(\mathbb{T}_c), \mathbf{D}_s(\mathbb{T}_c)] \right)^{-1} \begin{bmatrix} \mathbf{D}_c(\mathbb{T}_c)^\top \\ \mathbf{D}_s(\mathbb{T}_c)^\top \end{bmatrix} \boldsymbol{\Sigma}_{\mathbf{f}\mathbf{f}}(\mathbb{T}_c, \mathbb{T}_c)^{-1} \mathbf{f}_k(\mathbb{T}_c).
\end{aligned} \tag{24}$$

- Selection of the collocation points $\mathbb{T}_c$. Though a set of uniformly distributed collocation points may give rise to satisfactory results, an *adaptive collocation strategy* by further considering the posterior covariance or residual information will be proposed in what follows to improve the accuracy as much as possible under the prescribed number of collocation points.

Algorithm 1: PIGP for Periodic Solution of Nonlinear Systems:

$[\mathbf{x}_{\text{sol}}(t), \boldsymbol{\Sigma}(t, t), \mathbf{h}] = \text{PIGP}(\mathbf{R}(\mathbf{x}, t), \mathbf{x}_0, \mathbf{h}_0, \mathbb{T}_c, \text{Tol})$

Input: Model equation $\mathbf{R}(\mathbf{x}, t)$;

Initial solution $\mathbf{x}_0(t)$;

Initial hyperparameters $\mathbf{h}_0$;

Set of collocation points $\mathbb{T}_c$;

Error tolerance for convergence criterion Tol (e.g., $=10^{-5}$).

Output: The solution $\mathbf{x}_{\text{sol}}(t)$, the covariance $\boldsymbol{\Sigma}(t, t)$ and eventual hyperparameters $\mathbf{h}$

```

1 Initialize  $k = 0$ , initial solution  $\mathbf{x}_0(t)$  and hyperparameters  $\mathbf{h} = \mathbf{h}_0$ ;
2 repeat
3   Prescribe the GP kernel  $\mathbf{k}(t, t')$  with the hyperparameters  $\mathbf{h}$ ;
4   Get the linearized equation by Taylor expansion around  $\mathbf{x}_k$  to obtain  $\mathcal{L}_t^{(k)}[\mathbf{x}] = \mathbf{f}_k(t)$ ;
5   Collect the observational data  $\mathbf{f}_k(\mathbb{T}_c)$ ;
6   Compute the mean matrices  $\mathbf{D}_c(\mathbb{T}_c)$ ,  $\mathbf{D}_s(\mathbb{T}_c)$  and the covariance matrices  $\boldsymbol{\Sigma}_{\mathbf{x}\mathbf{x}}(t, t)$ ,
    $\boldsymbol{\Sigma}_{\mathbf{x}\mathbf{f}}(t, \mathbb{T}_c)$ ,  $\boldsymbol{\Sigma}_{\mathbf{f}\mathbf{f}}(\mathbb{T}_c, \mathbb{T}_c)$  as defined in the derivations following Eq. (19);
7   Acquire the mean parameters  $[\mathbf{a}; \mathbf{b}]$  via Eq. (24) and compute the mean of the data
    $\boldsymbol{\mu}_{\mathbf{f}}(\mathbb{T}_c) = \mathbf{D}_c(\mathbb{T}_c)\mathbf{a} + \mathbf{D}_s(\mathbb{T}_c)\mathbf{b}$  as well as the mean function  $\boldsymbol{\mu}(t) = \mathbf{a} \cos(\omega t) + \mathbf{b} \sin(\omega t)$ ;
8   Update the solution  $\mathbf{x}_{k+1}(t) = \boldsymbol{\mu}(t) + \boldsymbol{\Sigma}_{\mathbf{x}\mathbf{f}}(t, \mathbb{T}_c) \boldsymbol{\Sigma}_{\mathbf{f}\mathbf{f}}(\mathbb{T}_c, \mathbb{T}_c)^{-1} (\mathbf{f}_k(\mathbb{T}_c) - \boldsymbol{\mu}_{\mathbf{f}}(\mathbb{T}_c))$ ;
9   Update the hyperparameters  $\mathbf{h}$  via maximum marginal likelihood in Eq. (23);
10  Compute convergence indicator  $\text{RelErr} = \frac{\|\mathbf{x}_{k+1} - \mathbf{x}_k\|}{\max(1, \|\mathbf{x}_k\|)}$ ;
11  Update  $k \leftarrow k + 1$ ;
12 until  $\text{RelErr} < \text{Tol}$  or  $k \geq 1000$ ;
13 return  $\begin{cases} \mathbf{x}_{\text{sol}}(t) = \mathbf{x}_k(t), \\ \boldsymbol{\Sigma}(t, t) = \boldsymbol{\Sigma}_{\mathbf{x}\mathbf{x}}(t, t) - \boldsymbol{\Sigma}_{\mathbf{x}\mathbf{f}}(t, \mathbb{T}_c) \boldsymbol{\Sigma}_{\mathbf{f}\mathbf{f}}(\mathbb{T}_c, \mathbb{T}_c)^{-1} \boldsymbol{\Sigma}_{\mathbf{f}\mathbf{x}}(\mathbb{T}_c, t), \\ \mathbf{h}. \end{cases}$ 

```

Consequently, given the set of collocation points $\mathbb{T}_c$, the algorithmic details of the proposed PIGP for periodic solution of nonlinear systems are presented in Algorithm 1. Herein, the norm operator in the convergence indicator is approximately computed as $\|\mathbf{y}(t)\| = \sqrt{\sum_{k=1}^m \mathbf{y}(\frac{(k-1)T}{m})^\top \mathbf{y}(\frac{(k-1)T}{m})} / m$ with

$m = 1000$. The initial solution $\mathbf{x}_0(t)$ can be obtained by numerical continuation or directly solving the approximately linearized equation in the form of Eq. (11) but with $(\mathbf{x}_k, \dot{\mathbf{x}}_k) = (\mathbf{0}, \mathbf{0})$. As with the initial hyperparameters, it is set to $\mathbf{h} = \{\sigma_i^2 = 0.25, l_i = 0.5, i = 1, 2, \dots, n; \sigma_\epsilon^2 = 10^{-6}\}$ in this paper.

3.3 Adaptive collocation strategy

To further enhance the accuracy and efficiency of the proposed PIGP, an adaptive collocation strategy is developed. The core idea is to gradually enrich the set of collocation points by adding new points where the current solution is most deficient, as indicated by either the posterior covariance or the residual of the governing equation. Mathematically, adaptive collocation gradually expands the set of collocation points $\mathbb{T}_c$ through certain criterion, i.e.,

$$\mathbb{T}_c \leftarrow \mathbb{T}_c \cup \{t^*\} \quad (25)$$

where t^* is determined based on the solution in Eq. (21) under the collocation points $\mathbb{T}_c$. Two distinct criteria are proposed for this purpose:

1. **Variance-based:** The posterior covariance $\Sigma_{k+1}(t, t)$ provides a local measure of the solution uncertainty at time t . To improve the overall accuracy or reduce the maximum uncertainty, the new collocation point t^* is selected at the instant where the trace of the covariance matrix attains its maximum,

$$t^* = \arg \max_{t \in [0, T]} \text{tr}(\Sigma_{k+1}(t, t)) \quad (26)$$

where $\text{tr}(\cdot)$ denotes the trace of a matrix.

Algorithm 2: Adaptive PIGP for Nonlinear Periodic Systems

Input: Model equation $\mathbf{R}(\mathbf{x}, t)$;
Initial solution $\mathbf{x}_0(t)$;
Initial hyperparameters $\mathbf{h}_0$;
Initial set of collocation points $\mathbb{T}_{c0}$;
Error tolerance Tol (e.g., 10^{-5}) and desired number of collocation points N_c .

Output: The solution $\mathbf{x}_{\text{sol}}(t)$ and the covariance $\Sigma(t, t)$

- 1 Initialize collocation points $\mathbb{T}_c = \mathbb{T}_{c0}$;
 - 2 **repeat**
 - 3 Get the solution $[\mathbf{x}_{\text{sol}}(t), \Sigma(t, t), \mathbf{h}] = \text{PIGP}(\mathbf{R}(\mathbf{x}, t), \mathbf{x}_0, \mathbf{h}_0, \mathbb{T}_c, \text{Tol})$ via the PIGP in Algorithm 1;
 - 4 Expand the set of collocation points $\mathbb{T}_c \leftarrow \mathbb{T}_c \cup \{t^*\}$ via variance-based criterion (26) or residual-based criterion (27);
 - 5 Update $\mathbf{x}_0(t) = \mathbf{x}_{\text{sol}}(t)$ and $\mathbf{h}_0 = \mathbf{h}$.
 - 6 **until** $|\mathbb{T}_c| \geq N_c$ (or other termination criterion);
 - 7 **return** $\mathbf{x}_{\text{sol}}(t), \Sigma(t, t)$.
-

2. **Residual-based:** The residual of the governing equation serves as a direct and physically meaningful indicator of solution fidelity. Accordingly, t^* is chosen at the point where the expected squared residual norm is maximized

$$\begin{aligned} t^* &= \arg \max_{t \in [0, T]} \mathbb{E}[||\mathbf{R}(\mathbf{x}(t), t)||^2] \\ &\approx \arg \max_{t \in [0, T]} \mathbb{E}[||\mathbf{R}(\boldsymbol{\mu}_{k+1}(t), t) + \mathcal{L}_t^{(k+1)}[\mathbf{x}(t) - \boldsymbol{\mu}_{k+1}(t)]||^2] \\ &= \arg \max_{t \in [0, T]} \left\{ ||\mathbf{R}(\boldsymbol{\mu}_{k+1}(t), t)||^2 + \text{tr}(\mathcal{L}_t^{(k+1)} \Sigma_{k+1}(t, s) (\mathcal{L}_s^{(k+1)})^\top |_{s=t}) \right\} \end{aligned} \quad (27)$$

where $\Sigma_{k+1}(t, s)$ is in the similar form to Eq. (21),

$$\Sigma_{k+1}(t, s) = \Sigma_{ff}(t, s) - \Sigma_{xf}(t, \mathbb{T}_c) \Sigma_{ff}(\mathbb{T}_c, \mathbb{T}_c)^{-1} \Sigma_{fx}(\mathbb{T}_c, s). \quad (28)$$

In practice, the maximization over the continuous interval $[0, T)$ in Eqs. (26) and (27) is performed approximately by evaluating the objective function on a sufficiently fine uniform grid, for instance $\{kT/10^3, k = 0, 1, \dots, 10^3 - 1\}$ and selecting the grid point that yields the maximum value. By considering the adaptive collocation strategy, the PIGP enables adaptive solution of the nonlinear problem and the details are depicted in Algorithm 2. Herein, the initial collocation points $\mathbb{T}_{c0}$ are typically chosen as a small number of uniformly distributed points are usually designated as a small number of uniformly distributed points, i.e., $\mathbb{T}_{c0} = \{kT/m, k = 0, 1, \dots, m - 1\}$ with m generally ranging from 5 to 10. This adaptive strategy ensures that the PIGP progressively focuses computational resources on the most critical regions of the time domain, leading to an efficient and accurate solution process.

4 Numerical examples

In this section, several representative nonlinear dynamical systems are investigated to validate the feasibility and performance of the proposed PIGP method. The selected examples encompass both single DOF (SDOF) and multiple DOF (MDOF) systems, as well as smooth and non-smooth systems, thereby providing a comprehensive assessment of the proposed approach. Specifically, the governing equations of the considered systems are given below:

N1. A smooth SDOF Duffing oscillator,

$$\ddot{x} + 0.5\dot{x} + x + x^3 = 0.3 \cos t. \quad (29)$$

N2. A SDOF non-smooth system,

$$x'' + 0.25x' + x + f_{nl}(x) = \cos(3.2t), \quad (30)$$

where the non-smooth restoring force is defined as

$$f_{nl}(x) = \begin{cases} 15x, & x \geq 0, \\ 0, & x < 0. \end{cases} \quad (31)$$

For this system, the T -periodic solution is unstable and a period-doubling bifurcation occurs, meaning that a stable $2T$ -periodic solution also exists.

N3. A smooth two-DOF coupled system,

$$\begin{cases} \ddot{x}_1 + 0.1\dot{x}_1 + x_1 + 0.3x_1^3 + 0.5(x_1 - x_2) = 2 \cos(t) + 0.5 \cos(2t), \\ \ddot{x}_2 + 0.1\dot{x}_2 + x_2 + x_2^3 + 0.5(x_2 - x_1) = 0. \end{cases} \quad (32)$$

N4. A nonlinear three-story shear-structure model (see Fig. 1) with the nonlinear component placed at the first floor. The relative displacements of the three floors with respect to the ground are chosen as the generalized coordinates, i.e., $\mathbf{q}(t) = [q_1(t), q_2(t), q_3(t)]^T$. The nonlinear restoring force is characterized by $r = r(q_1, \dot{q}_1)$, and a harmonic ground acceleration $a(t) = a_0 \cos(\omega t)$ is applied to the structure. As a result, the governing equation is given by,

$$\mathbf{M}\ddot{\mathbf{q}}(t) + \mathbf{C}\dot{\mathbf{q}}(t) + \mathbf{K}\mathbf{q}(t) + \mathbf{f}_{nl}(\mathbf{q}, \dot{\mathbf{q}}) = \mathbf{f}_{ex}(t) \quad (33)$$

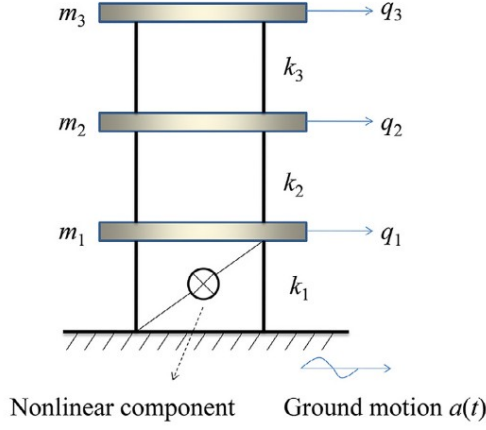


Fig. 1. A nonlinear three-story shear structure model.

where the mass and stiffness matrices are

$$\mathbf{M} = \begin{bmatrix} 3500 & 0 & 0 \\ 0 & 3500 & 0 \\ 0 & 0 & 3500 \end{bmatrix}, \mathbf{K} = \begin{bmatrix} 3.0 \times 10^6 & -1.5 \times 10^6 & 0 \\ -1.5 \times 10^6 & 3.0 \times 10^6 & -1.5 \times 10^6 \\ 0 & -1.5 \times 10^6 & 1.5 \times 10^6 \end{bmatrix}. \quad (34)$$

The damping matrix is modeled using Rayleigh damping,

$$\mathbf{C} = \alpha_R \mathbf{M} + \beta_R \mathbf{K} \quad (35)$$

with $\alpha_R = 0.1358, \beta_R = 5.7097 \times 10^{-4}$. The external excitation by ground acceleration is

$$\mathbf{f}_{ex}(t) = -\mathbf{M} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} a_0 \cos(\omega t) \quad (36)$$

with $a_0 = 1, \omega = 7$. As with the nonlinear restoring force, two representative non-smooth scenarios are considered,

Scenario I. Bilinear-stiffness restoring force e.g., caused by breathing crack,

$$\mathbf{f}_{nl}(\mathbf{q}, \dot{\mathbf{q}}) = \begin{bmatrix} k_{bl}|q_1| \\ 0 \\ 0 \end{bmatrix} \quad (37)$$

with $k_{bl} = 5.0 \times 10^5$.

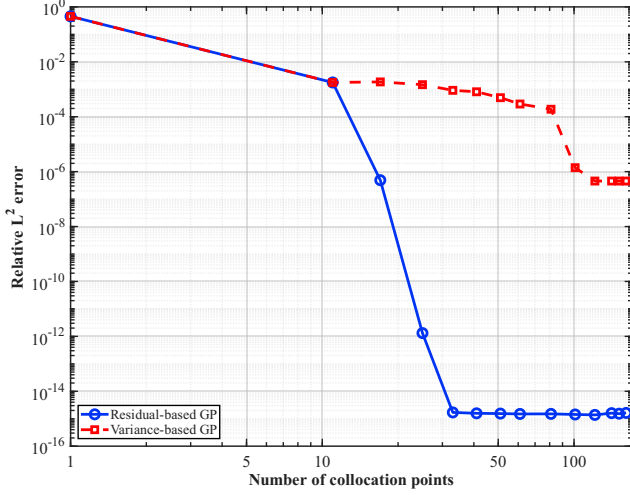
Scenario II. Nonlinear viscous damper

$$\mathbf{f}_{nl}(\mathbf{q}, \dot{\mathbf{q}}) = \begin{bmatrix} c_{nl}\dot{q}_1|\dot{q}_1| \\ 0 \\ 0 \end{bmatrix} \quad (38)$$

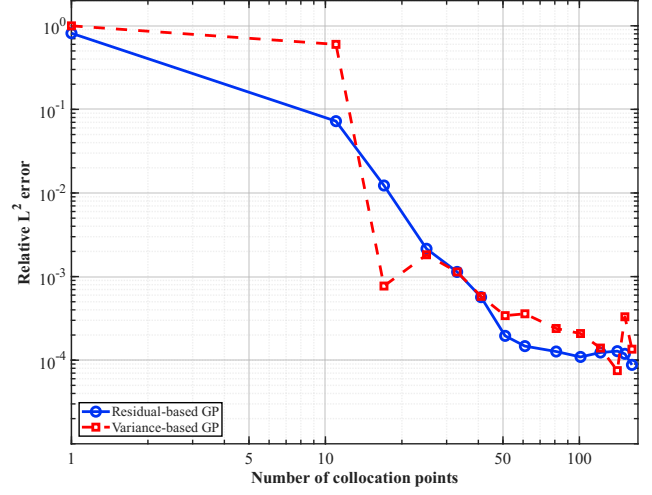
with $c_{nl} = 1.0 \times 10^5$.

4.1 Results by PIGP

The proposed PIGP with adaptive collocation (in Algorithm 2) is employed to get the solution of the above nonlinear dynamical systems. Herein, the high-fidelity solutions obtained by the shooting



(a) System N1



(b) System N2 (with unstable T -periodic solution)

Fig. 2. Convergence of relative L^2 error with the number of collocation points N_c by residual-based and variance-based adaptive criteria for: (a) system N1 and (b) system N2 (with unstable T -periodic solution).

method with very small error tolerance are used as reference. To quantify the accuracy of the solutions, the following relative L^2 error is defined

$$Error_x = \frac{\|x - x_{\text{shooting}}\|_2}{\|x_{\text{shooting}}\|_2} = \frac{\left(\sum_{i=1}^N |x(t_i) - x_{\text{shooting}}(t_i)|^2\right)^{1/2}}{\left(\sum_{i=1}^N |x_{\text{shooting}}(t_i)|^2\right)^{1/2}} \quad (39)$$

where $x(t)$ is the solution by the proposed or other methods, $x_{\text{shooting}}(t)$ is the reference solution obtained through the shooting method, $t_i = (i-1)T/N$ and N is set to 1000. Note that in the above error metric, x can be replaced by a vector response (from an MDOF system) so that the relative L^2 error well indicates the total accuracy of the MDOF system solution. In addition, since the proposed method inherently provides uncertainty information, the maximum posterior covariances of the displacement, velocity, and acceleration responses over the period $[0, T]$ are recorded and denoted by $(\sigma_x)_{\max}$, $(\sigma_{\dot{x}})_{\max}$, $(\sigma_{\ddot{x}})_{\max}$, respectively.

To begin with, the effect of the variance-based and residual-based adaptive criteria is evaluated using systems N1 and N2 (with T -periodic unstable solution). The relative L^2 errors under different numbers of collocation points N_c are depicted in Fig. 2 for the two systems. The results indicate that the residual-based criterion consistently achieves higher accuracy and more stable convergence with respect to N_c than its variance-based counterpart. This superiority is attributed to the fact that the residual-based criterion incorporates both the physical information of the governing equation and the covariance information, thereby leading to improved adaptive performance. Consequently, the residual-based criterion is adopted exclusively for the adaptive collocation strategy in all subsequent analyses.

Next, the solutions of the four systems N1–N4 are obtained by the proposed PIGP and the results are presented below:

- Fig. 3 displays the displacement, velocity and acceleration responses for system N1, along with the corresponding 95% confidence intervals (CIs), under $N_c = 15$ and $N_c = 35$. As N_c increases from 15 to 35, the relative L^2 error and maximum covariances decreases significantly from $Error_x = 5.28 \times 10^{-3}$, $(\sigma_x)_{\max} = 2.14 \times 10^{-1}$, $(\sigma_{\dot{x}})_{\max} = 4.69 \times 10^{-1}$, $(\sigma_{\ddot{x}})_{\max} = 1.67$ to respective values 8.32×10^{-10} , 1.24×10^{-4} , 2.21×10^{-4} , 1.49×10^{-3} . Notably, even with $N_c = 15$,

the displacement solution is already quite accurate, with $Error_x = 5.28 \times 10^{-3}$. It is also observed that for a fixed N_c , the covariance increases with the order of the time-derivative operator, becoming more pronounced from displacement to acceleration. This behavior is consistent with the fact that displacement responses converge faster than acceleration responses [15].

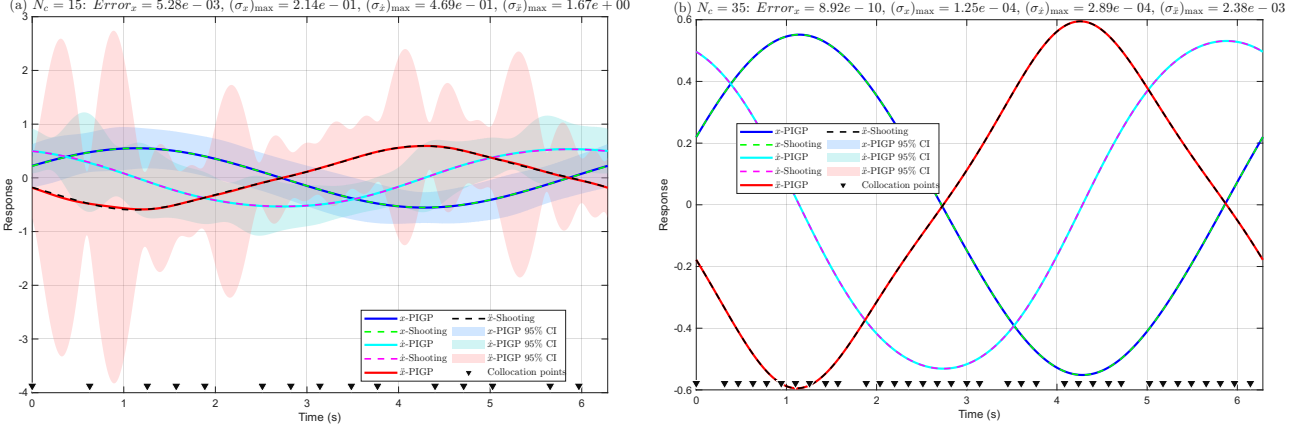


Fig. 3. Results of the displacement, velocity and acceleration responses by the proposed PIGP and the shooting method (for reference) for system N1 and under different numbers of collocation points N_c .

- For system N2, both the unstable T -periodic and the stable $2T$ -periodic solutions are computed by the proposed PIGP, with the results shown in Figs. 4 and 5, respectively. For the $2T$ -periodic solution, the frequency ω in the kernel (16) is correspondingly halved. In both cases, the PIGP accurately reproduces the reference solutions, with the relative L^2 error below 3×10^{-2} . As Furthermore, the adaptively selected collocation points are reasonably concentrated around the non-smooth points (where $x(t) = 0$), verifying the effectiveness of the adaptive strategy.

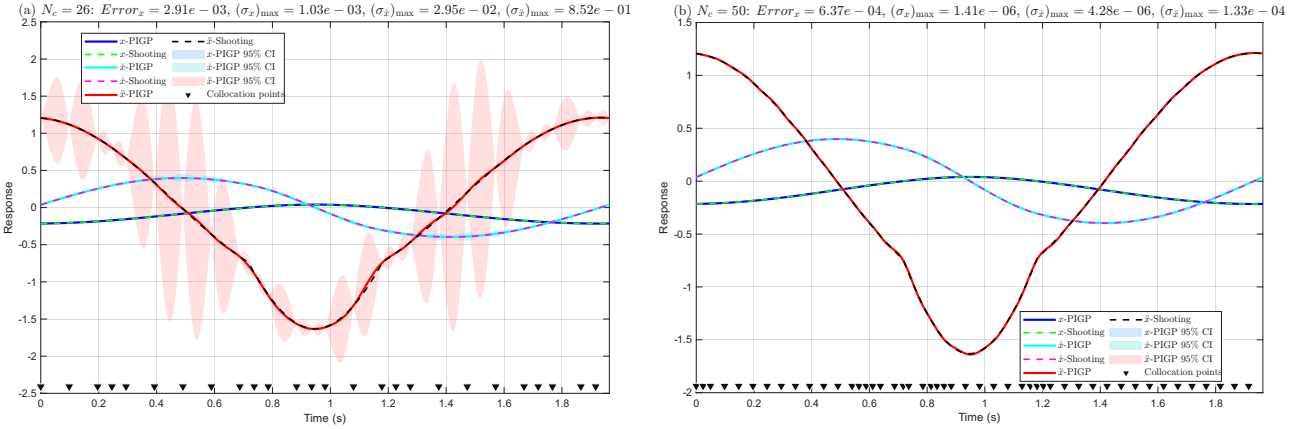


Fig. 4. Results of the displacement, velocity and acceleration responses by the proposed PIGP and the shooting method (for reference) for unstable T -periodic solution of system N2 and under different numbers of collocation points N_c .

- Fig. 6 presents the responses and 95% CIs for system N3. Again, the relative L^2 errors and covariances for both DOFs decrease consistently with increasing N_c . Similar to system N1, the adaptively collected points for system N3 tend to uniformly distributed, as both systems are smooth and exhibit smooth responses.
- For system N4, the solutions and their 95% CIs under scenarios I and II are displayed in Figs. 7 and 8, respectively. Accurate solutions are achieved with appropriate choices of N_c , and

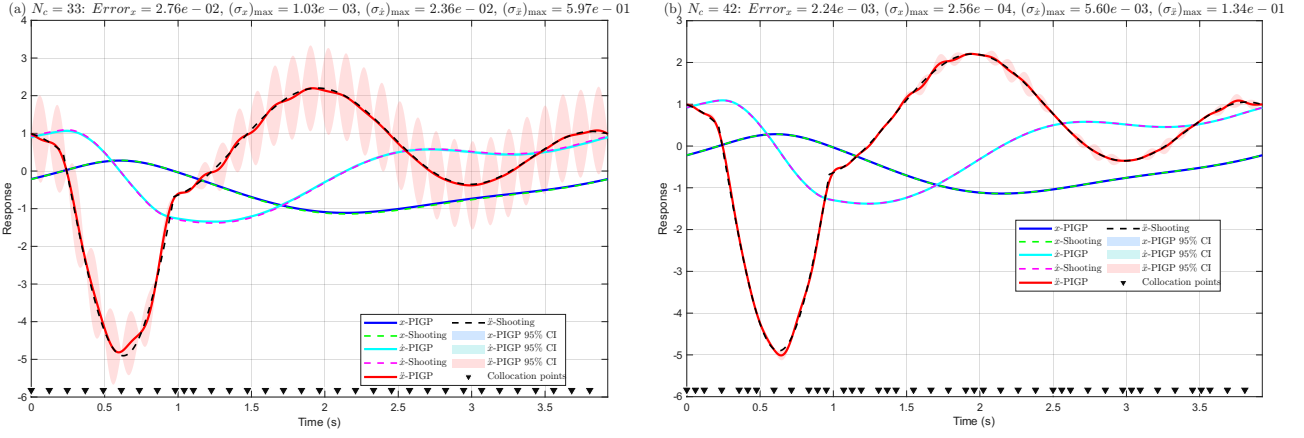


Fig. 5. Results of the displacement, velocity and acceleration responses by the proposed PIGP and the shooting method (for reference) for stable $2T$ -periodic solution of system N2 and under different numbers of collocation points N_c .

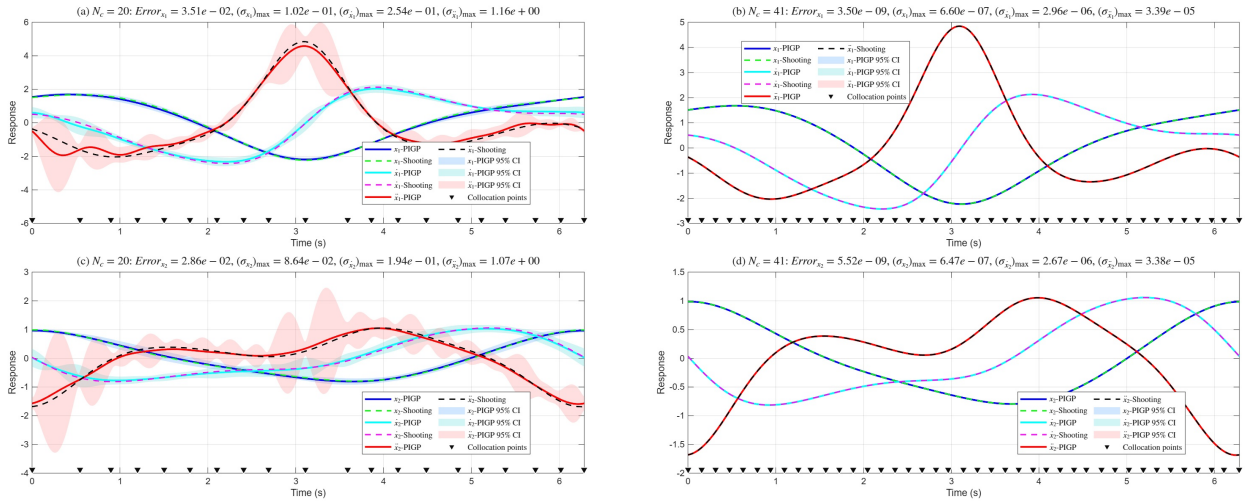


Fig. 6. Results of the displacement, velocity and acceleration responses by the proposed PIGP and the shooting method (for reference) for system N3 and under different numbers of collocation points N_c : (a) (b) correspond to the first DOF x_1 , and (c) (d) correspond to the second DOF x_2 .

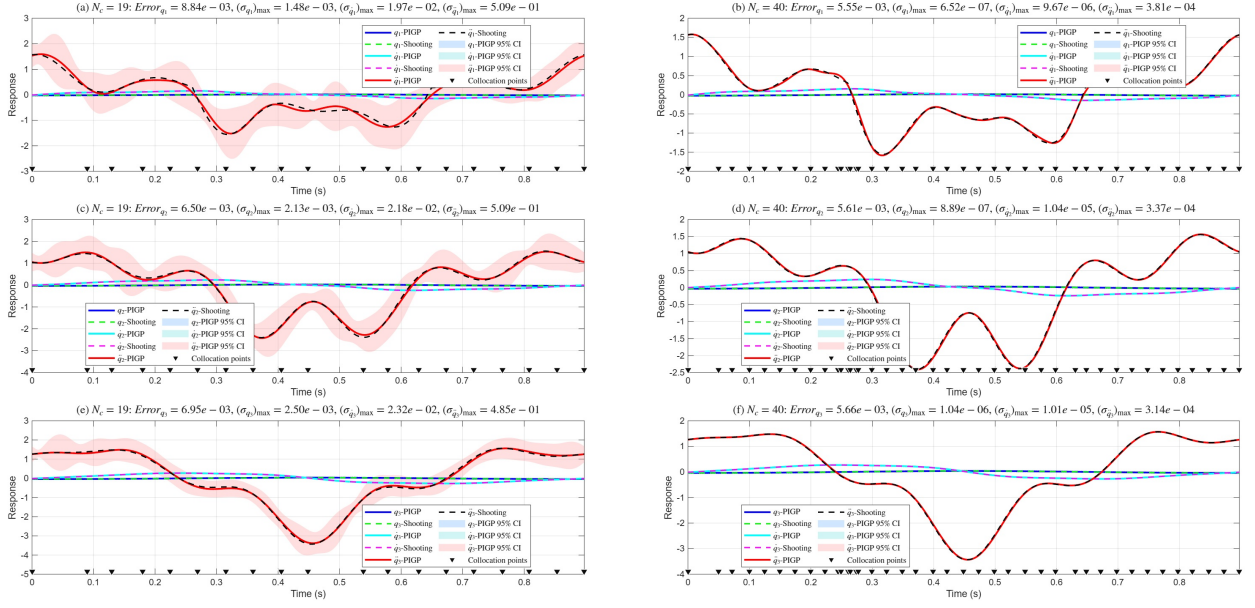


Fig. 7. Results of the displacement, velocity and acceleration responses by the proposed PIGP and the shooting method (for reference) for scenario I of system N4 and under different numbers of collocation points N_c : (a) (b) correspond to the first DOF q_1 , (c) (d) correspond to the second DOF q_2 and (e) (f) correspond to the third DOF q_3 .

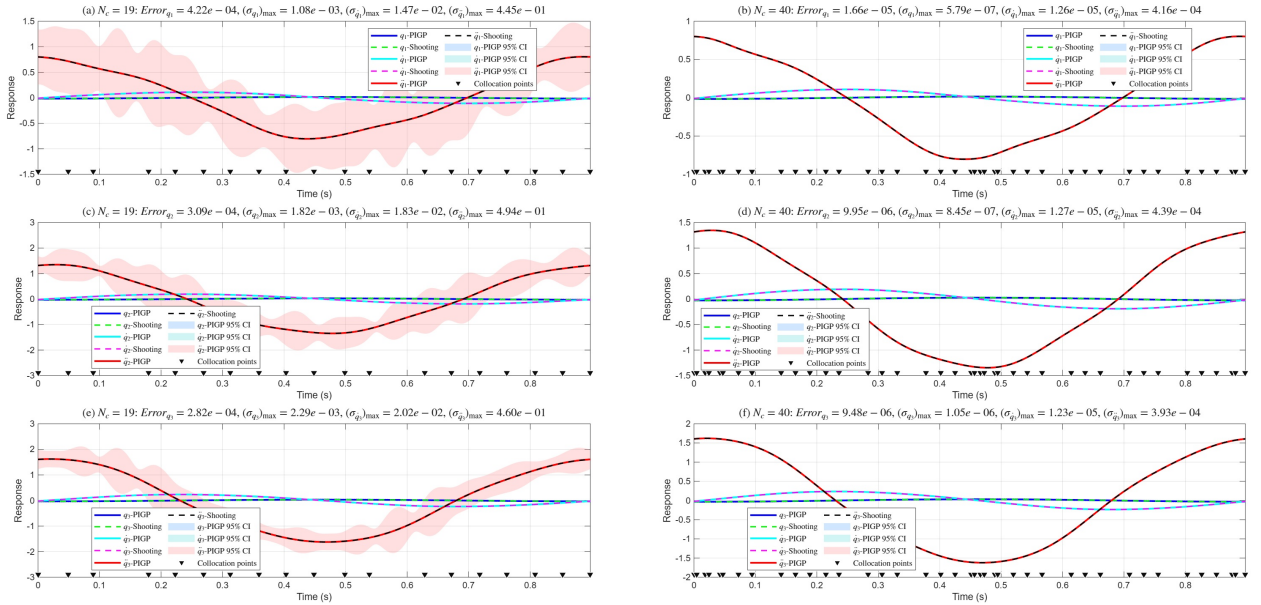


Fig. 8. Results of the displacement, velocity and acceleration responses by the proposed PIGP and the shooting method (for reference) for scenario II of system N4 and under different numbers of collocation points N_c : (a) (b) correspond to the first DOF q_1 , (c) (d) correspond to the second DOF q_2 and (e) (f) correspond to the third DOF q_3 .

both errors and covariances diminish as N_c increases. Again, the adaptively selected points are observed to cluster densely around the non-smooth points, specifically where $q_1(t) = 0$ in scenario I and $\dot{q}_1(t) = 0$ in scenario II.

In summary, the proposed PIGP is capable of accurately and adaptively solving smooth and non-smooth, SDOF and MDOF, stable and unstable periodic problems. Both the errors and the covariances converge with increasing N_c , confirming that the posterior covariance serves as a reliable indicator of solution accuracy. Moreover, the adaptively selected points are uniformly distributed for smooth systems but densely concentrated around non-smooth features for non-smooth systems.

4.2 Comparison of uniform and adaptive collocation strategies

The effects of uniform and adaptive collocation strategies are comparatively investigated using systems N3 and N4 (under scenario II). For uniform collocation, the time points are selected as $\{kT/N_c, k = 0, 1, \dots, N_c\}$ and will change with N_c so that the solution process varies substantially with the number of points. In contrast, the adaptive strategy incrementally enriches the collocation set, thereby enabling efficient and adaptive computation.

The variation of relative L^2 error with N_c under both collocation strategies is depicted in Fig. 9. As is observed, both collocation strategies yield satisfactory results, as the relative L^2 error drops below 0.01 for $N_c \geq 12$ in both systems. Nevertheless, the adaptive collocation strategy clearly outperforms the uniform one in terms of accuracy under the same N_c for both smooth and non-smooth systems and the superiority becomes more pronounced for larger N_c . Consequently, the adaptive collocation strategy is naturally recommended for use with the proposed PIGP.

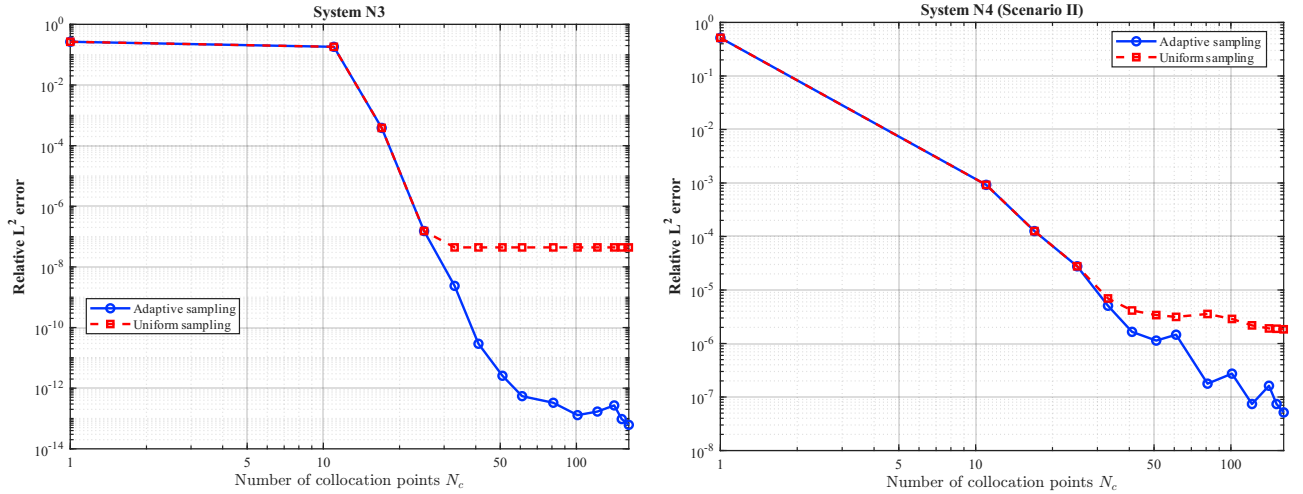


Fig. 9. Convergence of relative L^2 error with the number of collocation points N_c under uniform and adaptive collocation strategies for system N3 and system N4 (scenario II).

4.3 Comparison with the HBM

The performance of the proposed PIGP is further benchmarked against the classical HBM in terms of both accuracy and computational efficiency, using systems N3 and N4 (under scenario II). For the HBM, the number of unknowns is linearly related to $N_H = 2H + 1$, where H is the truncation order. Comparisons are therefore conducted under equal numbers of harmonic terms N_H and collocation points N_c . To render the HBM applicable to general smooth and non-smooth systems, the alternating frequency–time (AFT) scheme is adopted. Following the analysis in Ref. [14], the number of uniform sampling points is set to $2H + 1$ for smooth system N3 and $2H^2 + 1$ for non-smooth system N4. The HBM implementation is based on the complete program package '*HBNLvib*' provided in the

book *Harmonic Balance for Nonlinear Vibration Problems* [12]. To ensure a fair comparison, all the computations are conducted under the same hardware, equipped with an Intel Core i5-12490F CPU and 16 GB RAM at 3200 MT/s.

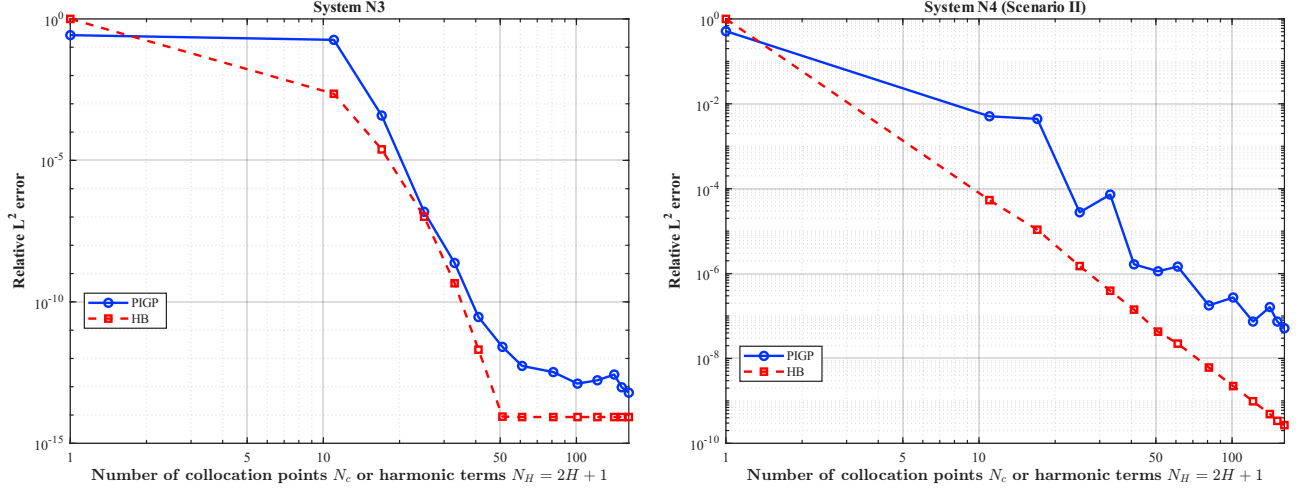


Fig. 10. Comparison of the convergence results of the relative L^2 error by the proposed PIGP and HBM for system N3 and system N4 (scenario II).

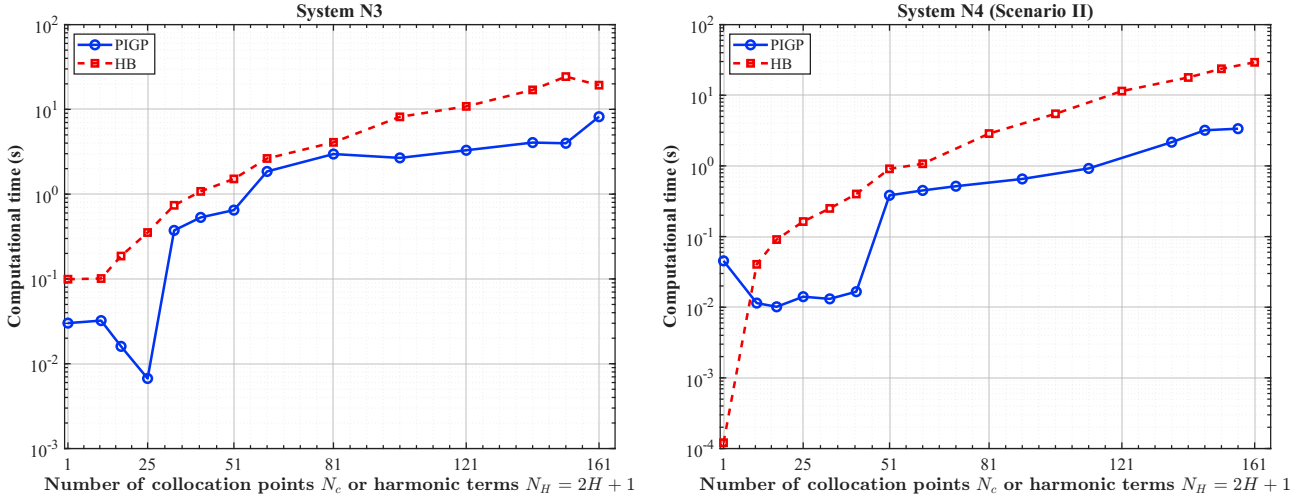


Fig. 11. Comparison of the computational time (in second) under different N_c (or N_H) by the proposed PIGP and HBM for system N3 and system N4 (scenario II).

Fig. 10 compares the convergence of the relative L^2 errors for both methods. The PIGP exhibits a relatively larger error than the HBM for the same N_c (or N_H), which is reasonable since the Fourier truncation in the HBM is optimal with respect to the L^2 -norm error. Nevertheless, the convergence rates of the two methods exhibit the same qualitative behavior: exponential convergence for the smooth system N3 and polynomial convergence (with nearly identical polynomial orders, as reflected by the slopes in the figure) for the non-smooth system N4 (scenario II). Fig. 11 compares the computational time required by the two methods. Except for cases with very small N_c (or N_H), the proposed PIGP is consistently less time-consuming than the HBM under the same number of unknowns, which can be attributed to the adaptivity of the PIGP.

To further compare the efficiency, Tables 1 and 2 list the computational times required by both methods to achieve the same accuracy levels for systems N3 and N4 (scenario II), respectively. Here,

t_{PIGP} denotes the time consumed by the proposed PIGP and t_{HBM} that by the HBM. Across almost all accuracy levels, the time ratio verifies $t_{\text{PIGP}}/t_{\text{HBM}} \leq 0.6$ for both systems, indicating that the PIGP requires less time to attain the same accuracy. The only exception occurs at a very low accuracy level for system N4 (scenario II), where the PIGP requires more iterations to achieve convergence. Overall, the proposed PIGP is found to be more efficient than the HBM under the same level of accuracy.

Table 1. Comparison of computational time between PIGP and HBM at the same levels of accuracy (or relative L^2 error) for system N3.

Accuracy level	PIGP		HBM		$t_{\text{PIGP}}/t_{\text{HBM}}$
	Relative L^2 error	Time (s)	Relative L^2 error	Time (s)	
Low accuracy	2.6731×10^{-1}	0.0301	1.0000×10^0	0.0993	0.30
10^{-3} – 10^{-4}	3.8595×10^{-4}	0.0160	2.2609×10^{-3}	0.1012	0.16
10^{-7}	1.5358×10^{-7}	0.0067	1.0237×10^{-7}	0.3527	0.02
10^{-9} – 10^{-10}	2.3776×10^{-9}	0.3735	4.5456×10^{-10}	0.7390	0.51
10^{-12}	2.5712×10^{-12}	0.6484	2.0449×10^{-12}	1.0785	0.60

Table 2. Comparison of computational time between PIGP and HBM at the same levels of accuracy (or relative L^2 error) for system N4 (scenario II).

Accuracy level	PIGP		HBM		$t_{\text{PIGP}}/t_{\text{HBM}}$
	Relative L^2 error	Time (s)	Relative L^2 error	Time (s)	
Low accuracy	5.1586×10^{-1}	0.1132	1.0000×10^0	0.0004	2.92×10^2
10^{-5}	2.7735×10^{-5}	0.0182	5.3722×10^{-5}	0.0538	0.34
10^{-6}	1.6508×10^{-6}	0.0279	1.5025×10^{-6}	0.1791	0.16
10^{-7}	1.7762×10^{-7}	0.2385	1.4164×10^{-7}	0.4506	0.53
10^{-8}	5.1576×10^{-8}	0.4186	4.2845×10^{-8}	0.9280	0.45

5 Conclusion

This paper proposes a novel data-driven framework, named the physics-informed Gaussian process (PIGP), for computing periodic solutions of nonlinear dynamical systems. The core idea lies in embedding the physical information directly into the GP formulation: the periodic boundary condition is enforced through tailored periodic kernel and mean functions, while the linearized governing equation is treated as observational data. By establishing a joint Gaussian distribution between the observational data and the solution, the periodic solution is obtained as the posterior mean, accompanied by the posterior covariance, which serves as a reliable indicator of solution accuracy. Furthermore, an adaptive collocation strategy is incorporated into the PIGP framework to enable an efficient and self-adaptive solution process. Several numerical examples are conducted to evaluate the performance of the proposed PIGP. The results demonstrate the following key findings:

- The proposed PIGP is capable of delivering accurate solutions for smooth and non-smooth, SDOF and MDOF, as well as stable and unstable periodic systems, provided that a sufficient number of collocation points is employed. The inherent covariance information effectively reflects the accuracy of the solution, as both the covariance and the relative error consistently decrease with an increasing number of collocation points.

- In comparison to the uniform collocation strategy, the adaptive collocation approach yields more accurate solutions, particularly when a relatively large number of collocation points is used. Moreover, for non-smooth systems, the adaptively selected points are reasonably concentrated in the vicinity of non-smooth features, confirming the effectiveness of the adaptive strategy.
- When compared to the classical HBM, the proposed PIGP generally exhibits a larger relative error for the same number of collocation points or harmonic terms, which is expected given the optimality of Fourier truncation in the L^2 -norm sense. Overall, the proposed PIGP proves to be more computationally efficient than the HBM under the same level of accuracy.

In summary, the proposed PIGP constitutes a powerful and efficient data-driven tool for solving nonlinear periodic systems. However, a potential limitation is that, for strongly non-smooth problems, the PIGP solution may exhibit oscillations near non-smooth points or boundaries, leading to reduced accuracy and increased computational cost—a limitation also shared by the HBM [45]. Future work will address this issue by designing new tailored mean and kernel functions that better capture non-smooth behaviors, as the current limitation stems from the absence of non-smooth priors in the present model formulation.

Acknowledgment

Financial supports from National Key R&D Program of China (No. 2024YFB2605600) and Guangdong Basic and Applied Basic Research Foundation (No. 2023A1515240046) are gratefully acknowledged.

Data Availability

Data available on request from the authors.

Compliance with ethical standards

Conflict of interest

The authors declare that they have no conflict of interest.

Appendix A: Derivation of the periodic kernel function

A periodic GP function $y(t)$ with fundamental frequency ω can be expanded by Fourier series,

$$y(t) = a_0 + \sum_{n=1}^{\infty} [a_n \cos(n\omega t) + b_n \sin(n\omega t)]. \quad (40)$$

This expansion can be conveniently reinterpreted as a polynomial function of the two-dimensional variable $\mathbf{u} = (\cos \omega t, \sin \omega t)$, i.e.,

$$y(t) = g(\cos \omega t, \sin \omega t) = g(\mathbf{u}) \quad (41)$$

where the polynomial function $g(\mathbf{u})$ is a GP. Since $g(\mathbf{u})$ is smooth and polynomial, it can be well characterized by the commonly used squared exponential (SE) kernel function [28],

$$k_g(\mathbf{u}, \mathbf{u}') = \sigma^2 \exp\left(-\frac{\|\mathbf{u} - \mathbf{u}'\|^2}{l^2}\right) \quad (42)$$

where $\|\cdot\|$ is the Euclidean norm, σ^2 and l denote the hyperparameters of the output variance and characteristic length scale, respectively. Substituting $\mathbf{u} = (\cos \omega t, \sin \omega t)$ and $\mathbf{u}' = (\cos \omega t', \sin \omega t')$ into Eq. (42) yields the desired periodic kernel function,

$$k_y(t, t') = \sigma^2 \exp \left(-\frac{\|(\cos \omega t, \sin \omega t) - (\cos \omega t', \sin \omega t')\|^2}{l^2} \right) = \sigma^2 \exp \left(-\frac{4 \sin^2 \left(\frac{\omega(t - t')}{2} \right)}{l^2} \right). \quad (43)$$

Thus, the periodic kernel naturally enforces the periodicity constraint by mapping the time domain onto a unit circle in the two-dimensional feature space, as expressed in Eq. (43)

References

- [1] S. Ghosh, L. Xue, A. Mishra, S. Saha, D. Dudkowski, S. K. Dana, T. Kapitaniak, J. Kurths, P. Ji, C. Hens, Universal nonlinear dynamics in damped and driven physical systems: From pendula via josephson junctions to power grids, *Physics Reports* 1147 (2025) 1–112.
- [2] J. Guo, C. Yue, H. Wang, T. Wang, X. Cao, Geometric nonlinearity in dynamic modeling of gyroelastic structure, *Communications in Nonlinear Science and Numerical Simulation* (2025) 109148.
- [3] A. Luongo, Perturbation methods for nonlinear autonomous discrete-time dynamical systems, *Nonlinear Dynamics* 10 (4) (1996) 317–331.
- [4] Z. Rahman, T. Burton, On higher order methods of multiple scales in non-linear oscillations-periodic steady state response, *Journal of Sound and Vibration* 133 (3) (1989) 369–379.
- [5] A. H. Nayfeh, C.-M. Chin, J. Pratt, Perturbation methods in nonlinear dynamics—applications to machining dynamics, *Journal of Manufacturing Science and Engineering Transactions of the ASME*.
- [6] R. Lin, C. Wang, F. Liu, Y. Xu, A new numerical method of nonlinear equations by four order runge-kutta method, in: 2010 IEEE International Conference on Industrial Engineering and Engineering Management, IEEE, 2010, pp. 1295–1299.
- [7] A. Granas, R. Guenther, J. Lee, The shooting method for the numerical solution of a class of nonlinear boundary value problems, *SIAM Journal on numerical analysis* 16 (5) (1979) 828–836.
- [8] A. Grolet, F. Thouverez, On a new harmonic selection technique for harmonic balance method, *Mechanical Systems and Signal Processing* 30 (2012) 43–60.
- [9] W.-J. Kim, N. Perkins, Harmonic balance/galerkin method for non-smooth dynamic systems, *Journal of Sound and Vibration* 261 (2) (2003) 213–224.
- [10] S. Stoykov, S. Margenov. Numerical computation of periodic responses of nonlinear large-scale systems by shooting method. *Computers and mathematics with applications* 67 (2014) 2257–2267.
- [11] R.I. Leine, H. Nijmeijer. *Dynamics and bifurcations of non-smooth mechanical systems*. Springer-Verlag Berlin Heidelberg; 2004.
- [12] M. Krack, J. Gross, *Harmonic Balance for Nonlinear Vibration Problems*, Mathematical Engineering, Springer International Publishing, 2019.

- [13] T.M. Cameron, J.H. Griffin. An alternating frequency/time domain method for calculating the steady-state response of nonlinear dynamic systems. *Journal of Applied Mechanics* 56(1) (1989) 149–154.
- [14] Y. Zhou, J. Huang, L. Wang, Event-driven gauss quadrature and stability analysis for fast alternating frequency–time harmonic balance of non-smooth systems, *Communications in Nonlinear Science and Numerical Simulation* 120 (2023) 107189.
- [15] L. Wang, Z.-R. Lu, J. Liu, Convergence rates of harmonic balance method for periodic solution of smooth and non-smooth systems, *Communications in Nonlinear Science and Numerical Simulation* 99 (2021) 105826.
- [16] Z. Ghahramani, Probabilistic machine learning and artificial intelligence, *Nature* 521 (7553) (2015) 452–459.
- [17] M. I. Jordan, T. M. Mitchell, Machine learning: Trends, perspectives, and prospects, *Science* 349 (6245) (2015) 255–260.
- [18] M. Raissi, P. Perdikaris, G. E. Karniadakis, Machine learning of linear differential equations using gaussian processes, *Journal of Computational Physics* 348 (2017) 683–693.
- [19] E. Haghighat, M. Raissi, A. Moure, H. Gomez, R. Juanes, A physics-informed deep learning framework for inversion and surrogate modeling in solid mechanics, *Computer Methods in Applied Mechanics and Engineering* 379 (2021) 113741.
- [20] C. Zhao, F. Zhang, W. Lou, X. Wang, J. Yang, A comprehensive review of advances in physics-informed neural networks and their applications in complex fluid dynamics, *Physics of Fluids* 36 (10).
- [21] W. He, J. Li, X. Kong, L. Deng, Multi-level physics informed deep learning for solving partial differential equations in computational structural mechanics, *Communications Engineering* 3 (1) (2024) 151.
- [22] C. Song, T. Alkhalifah, U. B. Waheed, Solving the frequency-domain acoustic vti wave equation using physics-informed neural networks, *Geophysical Journal International* 225 (2) (2021) 846–859.
- [23] M. Raissi, P. Perdikaris, G. E. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, *Journal of Computational physics* 378 (2019) 686–707.
- [24] M. Pförtner, I. Steinwart, P. Hennig, J. Wenger, Physics-informed gaussian process regression generalizes linear pde solvers, *arXiv preprint arXiv:2212.12474*.
- [25] H. Liu, Y.-S. Ong, X. Shen, J. Cai, When gaussian process meets big data: A review of scalable gps, *IEEE transactions on neural networks and learning systems* 31 (11) (2020) 4405–4423.
- [26] E. Haghighat, M. Raissi, A. Moure, H. Gomez, R. Juanes, A deep learning framework for solution and discovery in solid mechanics, *arXiv preprint arXiv:2003.02751*.
- [27] U. Braga-Neto, *Fundamentals of Pattern Recognition and Machine Learning*, Springer International Publishing, Cham, 2024, Ch. 10, pp. 271–279.
- [28] C. K. Williams, C. E. Rasmussen, *Gaussian processes for machine learning*, MIT press Cambridge, MA, 2006, Ch. 2, pp. 7–30.

- [29] M. Gulian, M. Raissi, P. Perdikaris, G. Karniadakis, Machine learning of space-fractional differential equations, *SIAM Journal on Scientific Computing* 41 (4) (2019) A2485–A2509.
- [30] U. Braga-Neto, *Fundamentals of Pattern Recognition and Machine Learning*, Springer International Publishing, Cham, 2024, Ch. 12, pp. 311–319.
- [31] A. Amer, F. Kopsaftopoulos, Gaussian process regression for active sensing probabilistic structural health monitoring: experimental assessment across multiple damage and loading scenarios, *Structural Health Monitoring* 22 (2) (2023) 1105–1139.
- [32] É. Roux, P. Balland, C. Elmo, L. Charleux, Physics-informed gaussian processes model for heat source identification, *Materials Today Communications* (2025) 112892.
- [33] M. Harkonen, M. Lange-Hegemann, B. Raita, Gaussian process priors for systems of linear partial differential equations with constant coefficients, in: *International conference on machine learning*, PMLR, 2023, pp. 12587–12615.
- [34] D. Dalton, A. Lazarus, H. Gao, D. Husmeier, Boundary constrained gaussian processes for robust physics-informed machine learning of linear partial differential equations, *Journal of Machine Learning Research* 25 (272) (2024) 1–61.
- [35] M. Raissi, P. Perdikaris, G. E. Karniadakis, Numerical gaussian processes for time-dependent and nonlinear partial differential equations, *SIAM Journal on Scientific Computing* 40 (1) (2018) A172–A198.
- [36] R. Lin, L. Hou, Y. Chen, Y. Jin, N. Saeed, Y. Chen, A novel adaptive harmonic balance method with an asymptotic harmonic selection, *Applied Mathematics and Mechanics* 44 (11) (2023) 1887–1910.
- [37] B. Tian, T. Inoue, L. Fan, A. Heya, H. Takeuchi, Hb-aft analysis of geometrically exact beams considering nodal rotational inertia and gyroscopic force, *Journal of Sound and Vibration* 613 (2025) 119162.
- [38] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, New York, 2006, Ch. 2, pp. 78–97.
- [39] X. Wang, Y. Qin, W. Qian, S. Zhang, X. Fan, A family of newton type iterative methods for solving nonlinear equations, *Algorithms* 8 (3) (2015) 786–798.
- [40] P. Heid, T. Wihler, Adaptive iterative linearization galerkin methods for nonlinear problems, *Mathematics of computation* 89 (326) (2020) 2707–2734.
- [41] E. Hetmaniok, M. Pleszczyński, Comparison of the selected methods used for solving the ordinary differential equations and their systems, *Mathematics* 10 (3) (2022) 306.
- [42] J. Xian, C. Su, Z. Wang, Non-stationary non-gaussian random vibration analysis of duffing systems based on explicit time-domain method, *Structural Safety* 105 (2023) 102368.
- [43] S. W. Shaw, P. J. Holmes, A periodically forced piecewise linear oscillator, *Journal of sound and vibration* 90 (1) (1983) 129–155.
- [44] J. Van Til, F. Alijani, S. Voormeeren, W. Lacarbonara, Frequency domain modeling of nonlinear end stop behavior in tuned mass damper systems under single-and multi-harmonic excitations, *Journal of Sound and Vibration* 438 (2019) 139–152.
- [45] Y. Zhou, J. Huang, L. Wang, Enriching harmonic balance with non-smooth bernoulli bases for accelerated convergence of non-smooth periodic systems, *Theoretical and Applied Mechanics Letters* 15 (1) (2025) 100562.