Probabilistic Engineering Mechanics 67 (2022) 103193

Contents lists available at ScienceDirect

Probabilistic Engineering Mechanics

journal homepage: www.elsevier.com/locate/probengmech

Uncertainty quantification of nonlinear system stochastic response estimates based on the Wiener path integral technique: A Bayesian compressive sampling treatment

Maria I. Katsidoniotaki [a], Apostolos F. Psaros [b], Ioannis A. Kougioumtzoglou [a,*]

[a] Department of Civil Engineering and Engineering Mechanics, Columbia University, USA
[b] Division of Applied Mathematics, Brown University, USA

ARTICLE INFO

Keywords:
Path integral
Nonlinear system
Stochastic dynamics
Sparse representations
Compressive sampling

ABSTRACT

The Wiener path integral (WPI) technique for determining the stochastic response of diverse nonlinear systems is enhanced herein based on a Bayesian compressive sampling (CS) treatment. Specifically, first, sparse expansions for the system response joint probability density function (PDF) are employed. Next, exploiting the localization capabilities of the WPI technique for direct evaluation of specific PDF points leads to an underdetermined linear system of equations for the expansion coefficients. Further, relying on a Bayesian CS solution formulation yields a posterior distribution for the expansion coefficient vector. In this regard, a significant advantage of the herein developed methodology relates to the fact that the uncertainty of the response PDF estimates obtained by the WPI technique is quantified. Furthermore, an adaptive scheme is proposed based on the quantified uncertainty of the estimates for optimal selection of PDF sample points. This yields considerably fewer boundary value problems to be solved as part of the WPI technique, and thus, the associated computational cost is significantly reduced. Two indicative numerical examples pertaining to a Duffing nonlinear oscillator and to an oscillator with asymmetric nonlinearities are considered for demonstrating the capabilities of the developed technique. Comparisons with pertinent Monte Carlo simulation data are included as well.

1. Introduction

Developing mathematical techniques for treating the problem of uncertainty propagation in the field of stochastic engineering dynamics has been a persistent challenge for more than six decades. In fact, complex nonlinear behaviors, high-dimensionality, and sophisticated excitation modeling require novel potent tools for solving the governing equations of motion accurately and in a computationally efficient manner. In this regard, diverse solution methodologies have been developed for determining system response statistics with varying degrees of success; see, indicatively, [1–3] for a broad perspective.

One of the promising stochastic engineering dynamics techniques, recently developed by Kougioumtzoglou and co-workers, pertains to the notion of path integral. From a mathematics point of view, the path integral concept refers to the generalization of integral calculus to functionals. It was first introduced by Wiener [4] (see also preliminary work by Daniell [5]), and was reinvented in a different form by Feynman [6] leading eventually to a reformulation of quantum mechanics [7]. In the field of stochastic engineering dynamics, the Wiener path integral (WPI) technique has demonstrated a high degree of accuracy [8], is capable of handling diverse system and excitation modeling [9–14], and can readily treat high-dimensional systems at a relatively low computational cost [15].

Further, the computational efficiency of the WPI technique was enhanced recently. This was done by resorting to compressive sampling (CS) concepts and tools in conjunction with appropriate expansions for the joint response probability density function (PDF) [16,17]; see also [18] for a broad perspective. Specifically, first, an appropriately selected basis was considered for expanding the joint response PDF by utilizing only a few nonzero terms. Next, a relatively small number of PDF points were determined directly by relying on the localization capabilities of the WPI technique. In this regard, an underdetermined linear system of equations was formulated for the sparse expansion coefficient vector. Furthermore, CS procedures in conjunction with group sparsity concepts and appropriate optimization algorithms were employed for solving efficiently the underdetermined system of equations and computing the coefficients of the joint response PDF expansion.

In this paper, the WPI technique is extended based on a Bayesian CS treatment (e.g., [19–22]). In particular, compared to the deterministic coefficient vector estimate obtained in the standard CS framework,

* Corresponding author.
E-mail address: ikougioum@columbia.edu (I.A. Kougioumtzoglou).

https://doi.org/10.1016/j.probengmech.2021.103193
Received 4 November 2021; Accepted 6 December 2021
Available online 10 December 2021

Bayesian CS yields a posterior distribution for the expansion coefficient vector. Clearly, this provides a tool for uncertainty quantification associated with the estimated system response PDF. Moreover, this additional information regarding the estimates is exploited herein for reducing further the computational cost associated with the WPI technique. Indeed, based on the quantified uncertainty of the estimates, an adaptive scheme is developed for optimal selection of PDF sample points; thus, yielding fewer boundary value problems (BVPs) to be solved as part of the WPI technique. Two indicative numerical examples pertaining to a Duffing nonlinear oscillator and to a nonlinear oscillator with an asymmetric response PDF are considered for demonstrating the capabilities of the developed technique. Comparisons with pertinent Monte Carlo simulation (MCS) data are included as well.

2. Wiener path integral technique: selected aspects

2.1. Wiener path integral solution formulation

The dynamics of a q-degree-of-freedom (q-DOF) nonlinear system subject to external stochastic excitation is governed by the equation

$$\mathbf{M}\ddot{\mathbf{x}} + \mathbf{C}\dot{\mathbf{x}} + \mathbf{K}\mathbf{x} + \mathbf{g}(\dot{\mathbf{x}}, \mathbf{x}, t) = \mathbf{w}(t) \tag{1}$$

where $\mathbf{x}(t) = \left[x_1(t), \ldots, x_q(t)\right]^T$ is the response displacement vector and $\mathbf{M}$, $\mathbf{C}$, $\mathbf{K}$ denote the mass, damping and stiffness matrices, respectively. Also, $\mathbf{g}(\dot{\mathbf{x}}, \mathbf{x}, t)$ represents an arbitrary nonlinear function. The excitation $\mathbf{w}(t) = \left[w_1(t), \ldots, w_q(t)\right]^T$ is modeled next, for notation simplicity, as a white noise vector process with $E\left[\mathbf{w}(t_l)\right] = E\left[\mathbf{w}(t_{l+1})\right] = 0$ and $E\left[\mathbf{w}(t_l)\mathbf{w}^T(t_{l+1})\right] = \mathbf{S}_w\delta(t_{l+1} - t_l)$, where t_l, t_{l+1} are two arbitrary time instants and $\mathbf{S}_w \in R_+^{q\times q}$ denotes a constant power spectrum matrix. Note that alternative, more complex, excitation modeling as a non-white and non-Gaussian stochastic vector process can be also readily accounted for by the WPI technique; see [11] for more details.

As shown in [11,23] (see also [7] for a broader perspective), the joint response transition PDF $p(\mathbf{x}_f, \dot{\mathbf{x}}_f, t_f|\mathbf{x}_i, \dot{\mathbf{x}}_i, t_i)$ corresponding to the system of Eq. (1) can be expressed as a functional integral in the form

$$p(\mathbf{x}_f, \dot{\mathbf{x}}_f, t_f|\mathbf{x}_i, \dot{\mathbf{x}}_i, t_i) = \int_C \exp(-\int_{t_i}^{t_f} \mathcal{L}\left(\mathbf{x}, \dot{\mathbf{x}}, \ddot{\mathbf{x}}\right) \mathrm{d}t)[\mathrm{d}\mathbf{x}(t)] \tag{2}$$

where $C = \{\mathbf{x}_i, \dot{\mathbf{x}}_i, t_i; \mathbf{x}_f, \dot{\mathbf{x}}_f, t_f\}$ is the set of all possible paths with initial condition $\{\mathbf{x}_i, \dot{\mathbf{x}}_i, t_i\}$ and final condition $\{\mathbf{x}_f, \dot{\mathbf{x}}_f, t_f\}$, $\mathrm{d}\mathbf{x}(t)$ denotes a functional measure, and $\mathcal{L}$ represents the Lagrangian of the system given by

$$\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}}, \ddot{\mathbf{x}}) = \frac{1}{2}\{\mathbf{M}\ddot{\mathbf{x}} + \mathbf{C}\dot{\mathbf{x}} + \mathbf{K}\mathbf{x} + \mathbf{g}(\dot{\mathbf{x}}, \mathbf{x}, t)\}^T\mathbf{S}_w^{-1}\{\mathbf{M}\ddot{\mathbf{x}} + \mathbf{C}\dot{\mathbf{x}} + \mathbf{K}\mathbf{x} + \mathbf{g}(\dot{\mathbf{x}}, \mathbf{x}, t)\} \tag{3}$$

Due to considerable difficulties in evaluating analytically the functional integral in Eq. (2), researchers resort routinely to approximate schemes involving the concept of the "most probable path" [7]. This is the trajectory $\mathbf{x}_c(t)$ for which the integral of the Lagrangian $\int_{t_i}^{t_f} \mathcal{L}\left(\mathbf{x}, \dot{\mathbf{x}}, \ddot{\mathbf{x}}\right) \mathrm{d}t$, known also as stochastic action, is minimized. This leads to the Euler–Lagrange equations

$$\mathcal{L}_{x_j} - \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{L}_{\dot{x}_j} + \frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{L}_{\ddot{x}_j} = 0, \quad j \in \{1, \ldots, q\} \tag{4}$$

subject to $4 \times q$ boundary conditions

$$\mathbf{x}_j(t_i) = \mathbf{x}_{j,i}, \quad \dot{\mathbf{x}}_j(t_i) = \dot{\mathbf{x}}_{j,i}, \quad \mathbf{x}_j(t_f) = \mathbf{x}_{j,f} \quad \text{and} \quad \dot{\mathbf{x}}_j(t_f) = \dot{\mathbf{x}}_{j,f} \tag{5}$$

Eqs. (4) and (5) constitute a BVP to be solved for determining the most probable path $\mathbf{x}_c(t)$. Following solution of Eq. (4) and substituting $\mathbf{x}_c(t)$ into Eq. (2), a specific point of the joint response transition PDF is determined approximately as

$$p(\mathbf{x}_f, \dot{\mathbf{x}}_f, t_f|\mathbf{x}_i, \dot{\mathbf{x}}_i, t_i) \approx \mathrm{C}\exp(-\int_{t_i}^{t_f} \mathcal{L}\left(\mathbf{x}_c, \dot{\mathbf{x}}_c, \ddot{\mathbf{x}}_c\right) \mathrm{d}t) \tag{6}$$

where C is a normalization constant. Although it is clear by comparing Eq. (2) and (6) that only one trajectory (i.e., the most probable path $\mathbf{x}_c(t)$) is accounted for in the evaluation of the WPI, it has been shown in various diverse applications (e.g., [9–14]) that the accuracy degree exhibited by this kind of approximation is relatively high. In fact, as proved in [24], for the case of linear multi-DOF systems the most probable path approximation yields the exact joint response PDF. Further, note that the accuracy degree of the WPI technique has been enhanced recently by considering a quadratic approximation to account also for fluctuations around the most probable path; see [8] for details.

Clearly, in the general case, the Euler–Lagrange Eqs. (4) and (5) are not amenable to an analytic solution treatment, and therefore, numerical schemes are required. In this regard, adopting a brute-force solution approach, for a specific time instant t_f the values of the joint response PDF are computed based on Eq. (6) over a discretized PDF domain of N points in each dimension. This yields N^{2q} BVPs to be solved for a q-DOF system.

2.2. Joint response PDF sparse representations

It can be readily seen that, by relying on the aforementioned brute-force solution approach, the WPI technique becomes computationally prohibitive with increasing number of dimensions. This is due to the fact that the number N^{2q} of BVPs to be solved increases exponentially with respect to q. To address this challenge, a polynomial expansion of the log-PDF was employed in [25] of the form

$$log(p(\mathbf{x}, \dot{\mathbf{x}})) \approx \sum_{i=1}^{n} c_i b_i(\mathbf{x}, \dot{\mathbf{x}}) \tag{7}$$

where c_i are the polynomial expansion coefficients to be determined and b_i are the basis functions, with $i \in \{1, \ldots, n\}$. Further, following the selection of n locations to perform the approximation, Eq. (7) takes the form of a linear system of n equations, i.e.,

$$\mathbf{y}_0 = \mathbf{B}\mathbf{c} \tag{8}$$

where $\mathbf{y}_0 \in \mathbb{R}^{n\times 1}$ is a vector of n samples of $log(p(\mathbf{x}, \dot{\mathbf{x}}))$, $\mathbf{B} \in \mathbb{R}^{n\times n}$ is the basis matrix and $\mathbf{c} = [c_1, \ldots, c_n]^T \in \mathbb{R}^{n\times 1}$ is the expansion coefficient vector. Therefore, in comparison to the brute-force approach, which requires N^{2q} BVPs to be solved numerically, the approach based on the PDF representation of Eq. (7) requires the solution of only $n = \frac{(p+2q)!}{(2q)!p!}$ BVPs, where p is the order of the polynomial approximation. For the vast majority of applications, $n \ll N^{2q}$, and thus, the gain in terms of computational efficiency is significant.

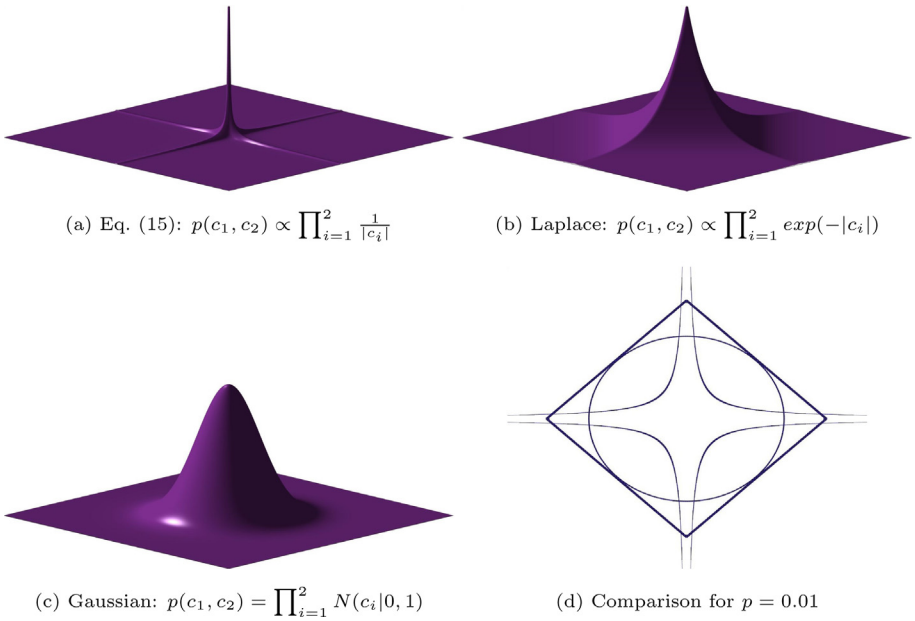
In [16], it was shown that the computational efficiency of the WPI technique can be further enhanced by resorting to sparse representations and compressive sampling concepts and tools; see also [18] for a broad perspective. Specifically, it was shown that for a large class of nonlinear systems the joint PDF expansion of Eq. (7) exhibits group sparsity, and thus, only $m < n$ PDF samples are required via the WPI technique. In this regard, Eq. (8) takes the form

$$\mathbf{y} = \mathbf{D}\mathbf{y}_0 = \mathbf{D}\mathbf{B}\mathbf{c} = \boldsymbol{\Phi}\mathbf{c} \tag{9}$$

where $\mathbf{y} \in \mathbb{R}^{m\times 1}$ is a vector of $m < n$ samples of $log(p(\mathbf{x}, \dot{\mathbf{x}}))$ and the matrix $\mathbf{D} \in \mathbb{R}^{m\times n}$ deletes rows randomly from the polynomial basis $\mathbf{B}$. The product $\mathbf{DB}$ yields the matrix $\boldsymbol{\Phi} \in \mathbb{R}^{m\times n}$ and $\mathbf{c} \in \mathbb{R}^{n\times 1}$ is the sparse polynomial coefficient vector to be determined. Eq. (9) constitutes an underdetermined linear system of equations, which can be solved based on an l_1-norm minimization formulation. This leads to an unconstrained minimization problem in the form

$$\hat{\mathbf{c}} = arg\ min_{\mathbf{c}}\{\|\mathbf{y} - \boldsymbol{\Phi}\mathbf{c}\|_2^2 + \lambda\|\mathbf{c}\|_1\} \tag{10}$$

where λ is a hyperparameter. Solving the problem of Eq. (10) yields the sparse coefficient vector to be substituted into the expansion of Eq. (9) for approximating the system joint response PDF. The interested reader is also directed to [16–18] for a more detailed presentation.



(a) Eq. (15): $p(c_1, c_2) \propto \prod_{i=1}^{2} \frac{1}{|c_i|}$ (b) Laplace: $p(c_1, c_2) \propto \prod_{i=1}^{2} exp(-|c_i|)$

(c) Gaussian: $p(c_1, c_2) = \prod_{i=1}^{2} N(c_i|0, 1)$ (d) Comparison for $p = 0.01$

Fig. 1. Comparison between (a) prior of Eq. (15) with uniform hyperprior (b) Laplace prior (c) Gaussian distribution and (d) their contours for $p = 0.01$.

3. A Bayesian compressive sampling solution treatment

In this section, the WPI technique is enhanced by developing a solution framework based on Bayesian CS (e.g., [18,19]). Specifically, a significant advantage of the herein developed technique relates to the fact that not only an estimate is obtained for the coefficient vector $\mathbf{c}$ in Eq. (9), but also the uncertainty of this estimate is quantified. Further, this capability for uncertainty quantification of the system response PDF estimates motivates in this section also the development of an adaptive scheme for optimal selection of PDF sample points. In this regard, the associated computational cost can be significantly reduced. In fact, it is shown that fewer BVPs need to be solved as part of the WPI technique without compromising, in general, the exhibited accuracy.

3.1. Bayesian modeling

Considering noise in the samples vector $\mathbf{y}$, Eq. (9) is expressed as

$$\mathbf{y} = \boldsymbol{\Phi}\mathbf{c} + \mathbf{e} \tag{11}$$

where the components of the noise vector $\mathbf{e} \in \mathbb{R}^{m\times 1}$ are modeled as independent and identically distributed Gaussian random variables with zero mean and unknown variance σ_e^2. Equivalently, $\mathbf{y} \in \mathbb{R}^{m\times 1}$ follows a Gaussian likelihood model given by

$$p(\mathbf{y}|\mathbf{c}, \sigma_e^2) = (2\pi\sigma_e^2)^{-\frac{m}{2}} exp(-\frac{\|\mathbf{y} - \boldsymbol{\Phi}\mathbf{c}\|_2^2}{2\sigma_e^2}) \tag{12}$$

Next, the objective of the Bayesian solution treatment relates to obtaining the coefficient vector $\mathbf{c}$ based on the available samples vector $\mathbf{y}$. To this aim, first, a sparsity-promoting prior distribution is selected for $\mathbf{c}$. In this regard, employing a hierarchical modeling approach (e.g., [26]), a Gaussian multivariate prior is considered for the coefficients $\mathbf{c}$ in the form

$$p(\mathbf{c}|\boldsymbol{\sigma}_c^2) = \prod_{i=1}^{n} N(c_i|0, \sigma_{ci}^2) \tag{13}$$

where $\boldsymbol{\sigma}_c^2 = \{\sigma_{ci}^2\}_{i=1}^n$ is a hyperparameter vector with n independent weights. Note that modeling approaches based on such Gaussian hierarchical priors are typically referred to as Gaussian automatic relevance determinators (ARD) in the Bayesian literature (e.g., [27]). ARDs adopt the following rationale for identifying the "relevant" parameters c_i. During the tuning process of $\boldsymbol{\sigma}_c^2$ (see Sections 3.2–3.3), if a non-zero value is obtained for the unknown hyperparameter σ_{ci}^2, then the corresponding parameter c_i is considered active in the approximation of $\mathbf{y}$, whereas if $\sigma_{ci}^2 \to 0$, c_i becomes inactive.

Further, hyperpriors are introduced for the hyperparameters $\boldsymbol{\sigma}_c^2$ and σ_e^2. Specifically, the inverse Gamma hyperprior constitutes a convenient choice for $p(\boldsymbol{\sigma}_c^2)$ from an analytical treatment perspective, since it is the conjugate prior of the Gaussian distributions in Eq. (13). Consequently, the convolution integral

$$p(\mathbf{c}) = \int p(\mathbf{c}|\boldsymbol{\sigma}_c^2)p(\boldsymbol{\sigma}_c^2)d\boldsymbol{\sigma}_c^2 \tag{14}$$

leads to a Student-t prior distribution for $\mathbf{c}$. Alternatively, the hyperpriors can become non-informative by setting inverse uniform distributions as the priors of $\boldsymbol{\sigma}_c^2$ and σ_e^2, i.e., $p(\sigma_{ci}^{-2})$, $p(\sigma_e^{-2}) \propto 1$. In this case, Eq. (14) yields

$$p(\mathbf{c}) \propto \prod_{i=1}^{n} \frac{1}{|c_i|} \tag{15}$$

In fact, selecting uniform hyperpriors for $\boldsymbol{\sigma}_c^2$ and σ_e^2 not only simplifies the computations in the ensuing analysis, but also renders the predictions scale-invariant; i.e., independent of the measurement units of the data. Furthermore, the distribution of the coefficient vector $p(c_i) \propto \frac{1}{|c_i|}$ exhibits a sharp peak at the origin, similarly to the Laplace prior $p(c_i) \propto exp(-|c_i|)$ that is widely used as a sparse prior in the literature [28]. Therefore, although $p(\mathbf{c}|\boldsymbol{\sigma}_c^2)$ is Gaussian, and thus, relatively less sparsity-promoting, the resulting $p(\mathbf{c})$ following integration over the hyperparameters via Eq. (14) exhibits significant sparsity-promoting behavior as shown in Fig. 1.

3.2. Posterior inference

In this section, the posterior distribution $p(\mathbf{c}|\mathbf{y})$ defined as

$$p(\mathbf{c}|\mathbf{y}) = \int p(\mathbf{c}|\mathbf{y}, \boldsymbol{\sigma}_c^2, \sigma_e^2)p(\boldsymbol{\sigma}_c^2, \sigma_e^2|\mathbf{y})\mathrm{d}\boldsymbol{\sigma}_c^2\mathrm{d}\sigma_e^2 \tag{16}$$

is obtained analytically. Applying Bayes' theorem, $p(\mathbf{c}|\mathbf{y}, \boldsymbol{\sigma}_c^2, \sigma_e^2)$ in Eq. (16) is expressed as

$$p(\mathbf{c}|\mathbf{y}, \boldsymbol{\sigma}_c^2, \sigma_e^2) = \frac{p(\mathbf{y}|\mathbf{c}, \sigma_e^2)p(\mathbf{c}|\boldsymbol{\sigma}_c^2)}{p(\mathbf{y}|\boldsymbol{\sigma}_c^2, \sigma_e^2)} \tag{17}$$

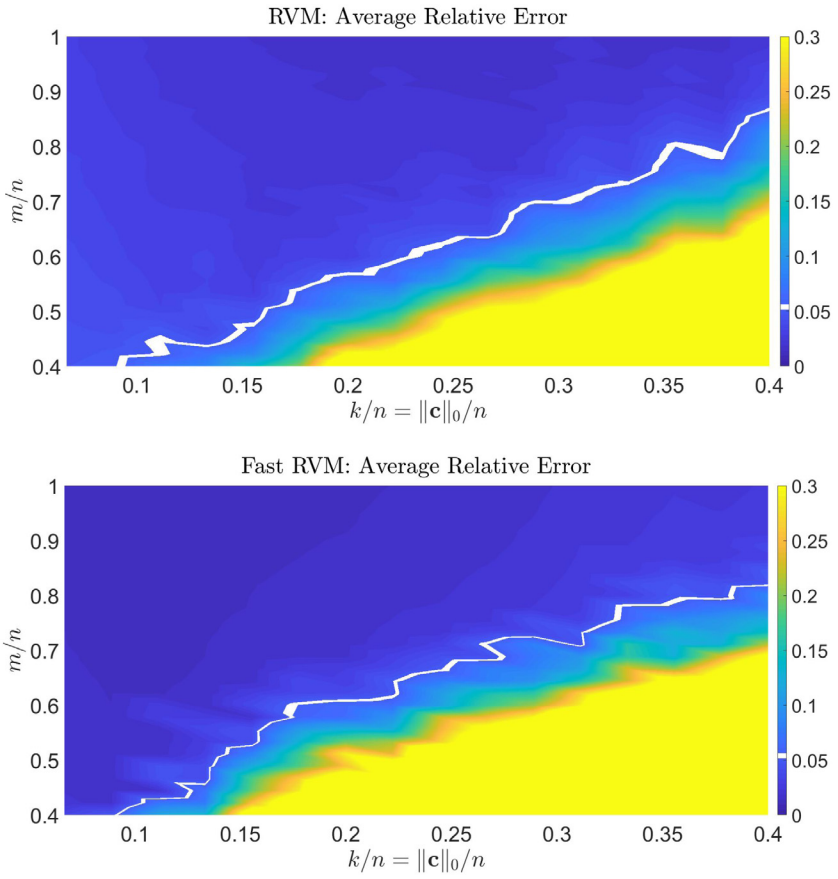


Fig. 2. Average relative error of coefficient vector $\mathbf{c}$ using RVM (top) and Fast RVM (bottom). The white line indicates 5% relative error.

where $p(\mathbf{y}|\boldsymbol{\sigma}_c^2, \sigma_e^2)$ is the marginal likelihood of $\mathbf{y}$. Obviously, once the distribution $p(\mathbf{c}|\mathbf{y})$ is known, the evaluation of $p(\mathbf{y}_0|\mathbf{y})$ is straightforward via the linear relationship of Eq. (8). Thus, the uncertainty of the WPI-based joint response PDF estimate can be quantified.

Next, according to the Laplace asymptotic approximation [29], $p(\mathbf{c}|\mathbf{y}) \approx p(\mathbf{c}|\mathbf{y}, \hat{\boldsymbol{\sigma}}_c^2, \hat{\sigma}_e^2)$, where $\hat{\boldsymbol{\sigma}}_c^2, \hat{\sigma}_e^2$ are obtained by maximizing $p(\boldsymbol{\sigma}_c^2, \sigma_e^2|\mathbf{y})$. Nevertheless, due to the proportionality relationship $p(\boldsymbol{\sigma}_c^2, \sigma_e^2|\mathbf{y}) \propto p(\mathbf{y}|\boldsymbol{\sigma}_c^2, \sigma_e^2)p(\boldsymbol{\sigma}_c^2)p(\sigma_e^2)$ and the uniform distribution modeling of $p(\boldsymbol{\sigma}_c^2)$ and $p(\sigma_e^2)$, this is equivalent to maximizing $p(\mathbf{y}|\boldsymbol{\sigma}_c^2, \sigma_e^2)$. This scheme is referred to in the literature as type-II maximum likelihood estimation (MLE).

Further, considering Eq. (17), the posterior distribution $p(\mathbf{c}|\mathbf{y}, \boldsymbol{\sigma}_c^2, \sigma_e^2)$ is evaluated analytically and takes the form of a multivariate Gaussian distribution, i.e., $p(\mathbf{c}|\mathbf{y}, \boldsymbol{\sigma}_c^2, \sigma_e^2) = N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where the mean vector and covariance matrix are expressed as [28]

$$\boldsymbol{\mu} = \sigma_e^{-2}\boldsymbol{\Sigma}\boldsymbol{\Phi}^T y \quad \text{and} \quad \boldsymbol{\Sigma} = (\sigma_e^{-2}\boldsymbol{\Phi}^T\boldsymbol{\Phi} + \mathbf{A})^{-1} \tag{18}$$

respectively. In Eq. (18), σ_e^{-2} and $\mathbf{A} = \mathbf{I}\cdot\boldsymbol{\sigma}_c^{-2} = diag\{\sigma_{c1}^{-2}, \ldots, \sigma_{cn}^{-2}\} \in \mathbb{R}^{n\times n}$ are the hyperparameters to be evaluated.

Furthermore, due to the linear relationship between $\mathbf{c}$ and $\mathbf{y}_0$ in Eq. (8), the distribution of $\mathbf{y}_0$ is also Gaussian given by [19]

$$p(\mathbf{y_0}|\mathbf{y}, \boldsymbol{\sigma}_c^2, \sigma_e^2) = N(\mathbf{B}\boldsymbol{\mu}, \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^T) \tag{19}$$

Equivalently, the distribution of the WPI-based estimate $p(\mathbf{x}, \dot{\mathbf{x}})$ at an arbitrary point $(\mathbf{x}, \dot{\mathbf{x}})$ is log-normal based on the relation $\mathbf{y}_0 = log(p(\mathbf{x}, \dot{\mathbf{x}}))$ with

$$E\{p(\mathbf{x}, \dot{\mathbf{x}})\} = exp(\mathbf{B}\boldsymbol{\mu} + \frac{1}{2}diag(\mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^T)) \tag{20}$$

$$Var\{p(\mathbf{x}, \dot{\mathbf{x}})\} = exp(2\mathbf{B}\boldsymbol{\mu} + diag(\mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^T))(exp(diag(\mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^T)) - 1). \tag{21}$$

Obviously, the uncertainty of the response PDF estimate obtained by the WPI technique is quantified herein based on the log-normal distribution described by Eqs. (20) and (21).

3.3. Performance assessment of selected relevance vector machine (RVM) algorithms

Considering Eqs. (12) and (13), a type-II MLE scheme for obtaining the optimal values $\hat{\boldsymbol{\sigma}}_c^2, \hat{\sigma}_e^2$ pertains to maximizing the logarithm of $p(\mathbf{y}|\boldsymbol{\sigma}_c^2, \sigma_e^2)$ expressed as

$$log\ p(\mathbf{y}|\boldsymbol{\sigma}_c^2, \sigma_e^2) = -\frac{1}{2}[m\cdot log2\pi + log|\mathbf{C}| + \mathbf{y}^T\mathbf{C}^{-1}\mathbf{y}] \tag{22}$$

where $\mathbf{C} = \sigma^2\mathbf{I} + \boldsymbol{\Phi}\mathbf{A}^{-1}\boldsymbol{\Phi}^T$. This hyperparameter learning problem can be solved iteratively by employing a relevance vector machine (RVM) scheme [28,30]. Specifically, a relatively standard implementation dictates that the values $\boldsymbol{\sigma}_c^2$ and σ_e^2 are obtained at each step as

$$\sigma_{ci}^{2\ (new)} = \frac{\mu_i^2}{\gamma_i} \text{ and } \sigma_e^{2(new)} = \frac{\|\mathbf{y} - \boldsymbol{\Phi}\mathbf{c}_s\|^2}{m - \Sigma_i\gamma_i} \tag{23}$$

where $i \in \{1, 2, \ldots, n\}$, and $\gamma_i = 1 - \Sigma_{ii}/\sigma_{ci}^2 \in [0, 1]$ with Σ_{ii} being the variance of the ith element. Notably, the set of coefficients c_i corresponding to large values of σ_{ci}^2 are the ones contributing to the representation of $\mathbf{y}$ the most, whereas small values of σ_{ci}^2 correspond to rather inactive coefficients [31].

Further, the Fast RVM scheme [32] represents an enhancement of the standard RVM that is computationally more efficient and promotes sparser, in general, solutions. According to Fast RVM, the expression of $p(\mathbf{y}|\boldsymbol{\sigma}_c^2, \sigma_e^2)$ in Eq. (22) and the matrix $\mathbf{C}$ are related to a single hyperparameter σ_{ci}^2 as

$$log\ p(\mathbf{y}|\boldsymbol{\sigma}_c^2, \sigma_e^2) = log\ p(\mathbf{y}|\boldsymbol{\sigma}_{c-i}^2, \sigma_e^2) + \frac{1}{2}[log\ \sigma_{ci}^{-2} - log(\sigma_{ci}^{-2} + s_i) + \frac{q_i^2}{\sigma_{ci}^{-2} + s_i}] \tag{24}$$

where $q_i = \phi_i^T C_i^{-1} y$ and $s_i = \phi_i^T C_i^{-1} \phi_i$. The matrix $\mathbf{C}$ is given by $\mathbf{C} = \mathbf{C}_{-i} + \sigma_{ci}^2 \phi_i \phi_i^T$, implying that the ith row is neglected in $\mathbf{C}_{-i}$. Furthermore, the maximization of Eq. (24) according to [33] yields the expressions

$$\sigma_{ci}^2 = \begin{cases} \frac{q_i^2 - s_i}{s_i^2}, & \text{if } q_i^2 > s_i, \\ 0, & \text{if } q_i^2 \leq s_i \end{cases} \tag{25}$$

Based on the value of σ_{ci}^2, the basis function ϕ_i is added to or removed from the basis matrix $\boldsymbol{\Phi}$. Note that the Fast RVM yields m significant basis functions ϕ_i to be accounted for in Eq. (18) compared to the n functions ϕ_i used in the standard RVM. A more detailed presentation of the scheme can be found in [32].

Next, to assess the performance of the two RVM schemes, an experiment is conducted involving reconstruction of 100 synthetically generated coefficient vectors $\mathbf{c}$. In Fig. 2, the average relative errors of the RVM and the Fast RVM schemes are shown. Note that the error is defined as $\|\hat{\mathbf{c}} - \mathbf{c}\|_2 / \|\hat{\mathbf{c}}\|_2$, where $\hat{\mathbf{c}}$ is the reference coefficient vector to be reconstructed and $\mathbf{c}$ is the mean of the Bayesian posterior. Comparisons are made for various values of samples m and sparsity k, for a given number of coefficients n. Also, the basis matrix $\boldsymbol{\Phi}$ is randomly generated and a zero-mean Gaussian noise with standard deviation $\sigma_e = 0.005$ is added to each of the m samples.

Remarkably, for an indicative sparsity ratio of $k/n = 0.2$ (a value commonly used in problems considered herein), utilizing a relatively low sampling ratio of $m/n \approx 0.6$ yields a reconstruction error smaller than 5% for both schemes.

3.4. Optimal selection of sample points

According to Section 2.2, the samples vector $\mathbf{y}$ in Eq. (9) can be obtained based on a random selection of points over the effective domain of the system response log-PDF. However, such an approach can be further optimized based on a judicious selection of the samples.

In this regard, a scheme is proposed in this section for optimal selection of points by minimizing the differential entropy of the total samples vector $\mathbf{y}_0$ [19]. Specifically, the differential entropy of $\mathbf{y}_0$ when adding a new point in $\mathbf{y}$ takes the form

$$h_{new}(\mathbf{y}_0) = h(\mathbf{y}_0) - \frac{1}{2} log(1 + (\phi_{m+1}^T \boldsymbol{\Sigma} \phi_{m+1})/\sigma_e^2) \tag{26}$$

where ϕ_{m+1} is the new row added to the matrix $\boldsymbol{\Phi}$ of Eq. (9). The minimization of Eq. (26) is equivalent to maximizing the variance of the new sample vector $\mathbf{y}_{m+1}$ based on equation

$$\phi_{m+1}^T \boldsymbol{\Sigma} \phi_{m+1} = \phi_{m+1}^T Cov(\mathbf{c}) \phi_{m+1} = Var(\mathbf{y}_{m+1}) \tag{27}$$

In practice, this is achieved by adding to the previous sample vector $\mathbf{y}_m$ the point of the log-PDF $\mathbf{y}_0$ with the largest variance, or equivalently, the point of the PDF with the largest relative variance $(\boldsymbol{\sigma}/\boldsymbol{\mu})^2$. Indeed, considering Eqs. (20) and (21), as well as Eq. (19), yields

$$\frac{Var\{p(\bar{\mathbf{x}}, \dot{\bar{\mathbf{x}}})\}}{E\{p(\bar{\mathbf{x}}, \dot{\bar{\mathbf{x}}})\}^2} = exp(diag(\mathbf{B} \boldsymbol{\Sigma} \mathbf{B}^T)) - 1 = exp(Var(\mathbf{y}_0)) - 1 \tag{28}$$

The performance of the proposed scheme for optimal selection of sample points is assessed in the numerical examples of Section 4, where it is shown that the same relative error of the reconstructed coefficient vector $\mathbf{c}$ can be obtained by using fewer sample points compared to the standard implementation.

3.5. Mechanization of the technique

The mechanization of the developed WPI technique based on Bayesian CS comprises the following steps:

(a) The expansion basis matrix $\mathbf{B}$ in Eq. (8) is constructed, which is associated with an n-dimensional coefficient vector $\mathbf{c}$ to be determined.

(b) The number $m < n$ of joint response PDF sample points to be used in the underdetermined linear system of Eq. (9) is selected. To this aim, the results shown in Fig. 2 and pertaining to the RVM schemes presented in Section 3.3 can serve as a guide.

(c) The locations of the m samples in Eq. (9) over the effective domain of the system response log-PDF are selected either randomly, or based on the optimal sampling scheme proposed in Section 3.4.

(d) The posterior distribution of the coefficient vector $\mathbf{c}$ is Gaussian with a mean vector and covariance matrix given by Eq. (18). These are determined by employing, for instance, the RVM schemes presented in Section 3.3 in conjunction with the optimal sampling scheme in Section 3.4.

(e) The distribution of the WPI-based estimate $p(\mathbf{x}, \dot{\mathbf{x}})$ at an arbitrary point $(\mathbf{x}, \dot{\mathbf{x}})$ is log-normal with a mean and a variance given by Eqs. (20) and (21), respectively.

4. Numerical examples

4.1. Duffing nonlinear oscillator

Consider a stochastically excited single-DOF Duffing nonlinear oscillator, whose equation of motion is given by

$$m\ddot{x} + c\dot{x} + kx(1 + \epsilon x^2) = w(t) \tag{29}$$

where $m = 1$, $k = 1$, $c = 0.1$, $\epsilon = 0.1$, and $w(t)$ is a white noise excitation with a constant power spectrum value $S_0 = 0.0637$. Note that the exact stationary joint response log-PDF for this oscillator has the closed-form expression (e.g., [34])

$$\mathbf{y}_0^{exact} = log(p(x, \dot{x})) = -\frac{c}{\pi S_0 m}\left(\frac{kx^2}{2m} + \frac{\epsilon k x^4}{4m} + \frac{\dot{x}^2}{2}\right) + \mathrm{C} \tag{30}$$

Clearly, the expression in Eq. (30) represents a 4th order polynomial with 3 non-zero coefficients, and C is a normalization constant. In this regard, selecting a 4th order polynomial ($p = 4$) as an approximating basis in Eq. (8) yields $n = 15$, whereas based on the exact solution of Eq. (30), the sparsity in the stationary phase is $k = 3$.

Next, the two RVM schemes discussed in Section 3.3, as well as their implementations based on the optimal selection of points presented in Section 3.4, are compared in conjunction with the Duffing nonlinear oscillator of Eq. (29). Specifically, the relative error $\|\hat{\mathbf{y}}_0 - \mathbf{y}_0\|_2 / \|\hat{\mathbf{y}}_0\|_2$ and the variance of $\mathbf{y}_0$ at an arbitrarily chosen time instant $t = 20$ s (both averaged over 1000 trials) are plotted in Fig. 3. $\hat{\mathbf{y}}_0$ denotes the estimate obtained by the standard brute-force WPI formulation presented in Section 2.1, whereas $\mathbf{y}_0$ denotes the mean of the Bayesian estimate in Eq. (19). It is readily seen that the Fast RVM coupled with the optimal sampling scheme of Section 3.4 yields the smallest error and exhibits the lowest uncertainty degree compared to the other alternatives, particularly for smaller values of the sampling ratio m/n.

In the following, the Fast RVM with the optimal sampling scheme is employed and the value $m/n = 0.6$ is utilized. This translates into using $m = 9$ points for computing the $n = 15$-dimensional coefficient vector $\mathbf{c}$, and thus, for determining the log-PDF $\mathbf{y}_0$. In this regard, the computed coefficient vector values c_i, with $i = 1, \ldots, 15$, are shown in Fig. 4. Specifically, the deterministic estimates obtained by the standard WPI technique with a polynomial PDF approximation are compared with the estimated distributions obtained by the herein developed Bayesian framework. The target (exact) expansion coefficient values based on Eq. (30) are also included for the stationary phase ($t = 20$ s). In general, it is seen that the mean values of the Bayesian estimates compare well with the respective deterministic estimates, both for the non-stationary ($t = 1$ s) and the stationary ($t = 20$ s) phases.

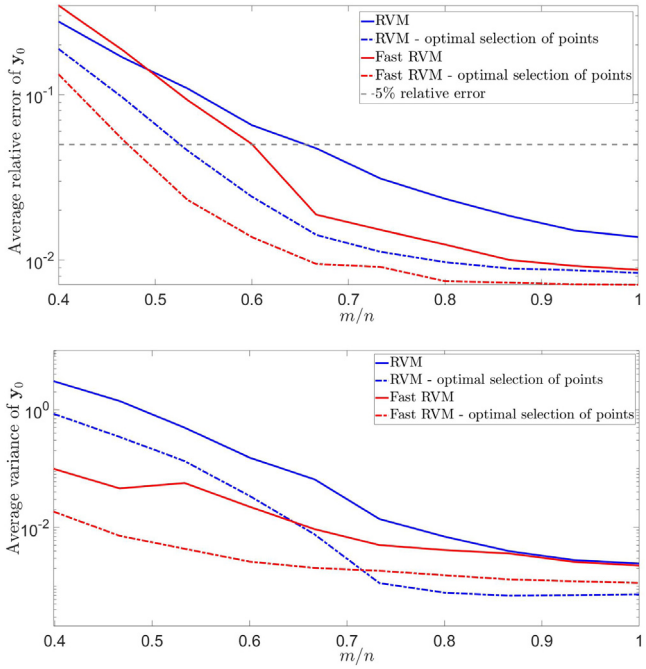


Fig. 3. Assessing the performance of various RVM schemes for determining the joint response PDF of a Duffing nonlinear oscillator at an arbitrarily chosen time instant $t = 20$ s. Average relative error (top) and average variance (bottom) of the Bayesian estimates for various sample ratios m/n.

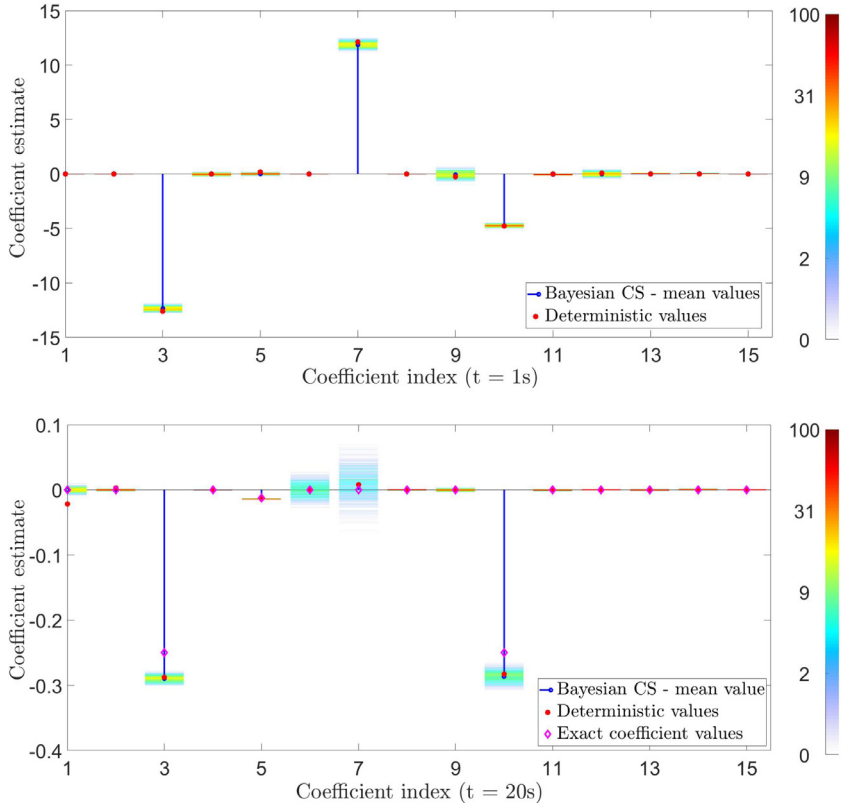


Fig. 4. Uncertainty quantification of Duffing nonlinear oscillator joint response PDF expansion coefficient vectors at $t = 1$ s and $t = 20$ s. Comparisons between the standard WPI technique with a polynomial PDF approximation (deterministic estimates) and the Bayesian WPI formulation (estimates of the coefficient vector distribution based on Eq. (18)). The exact coefficient values based on Eq. (30) are also included for the stationary phase ($t = 20$ s).

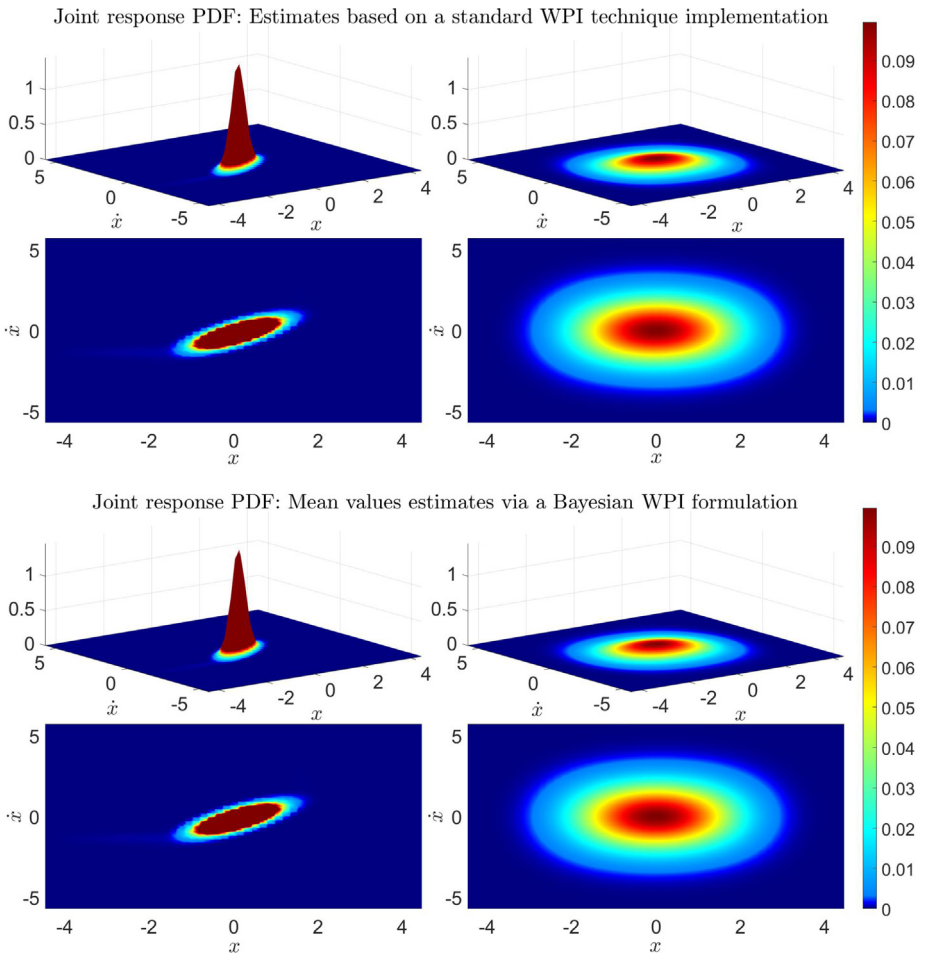


Fig. 5. Joint response PDF of a Duffing nonlinear oscillator at $t = 1$ s (left) and $t = 20$ s (right); standard brute-force WPI formulation (top), and Bayesian formulation — mean values estimates (bottom).

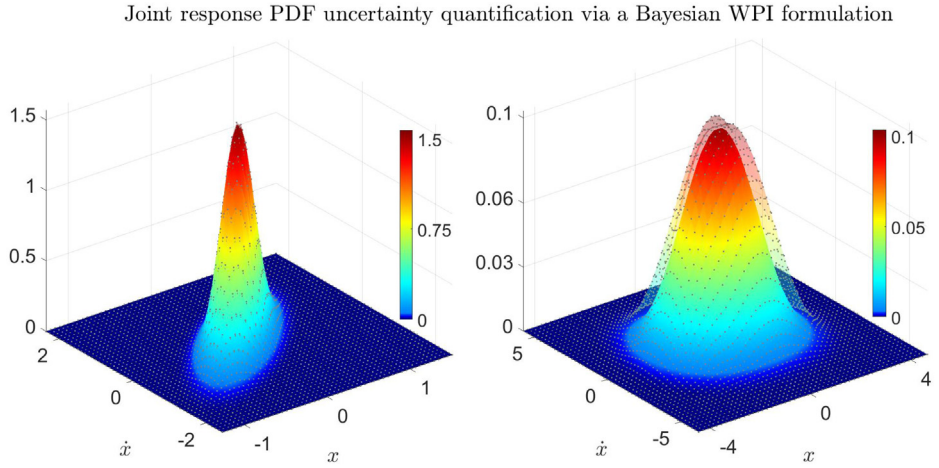


Fig. 6. Uncertainty quantification of the Duffing nonlinear oscillator joint response PDF estimates based on the proposed Bayesian WPI formulation for $t = 1$ s (left) and $t = 20$ s (right).

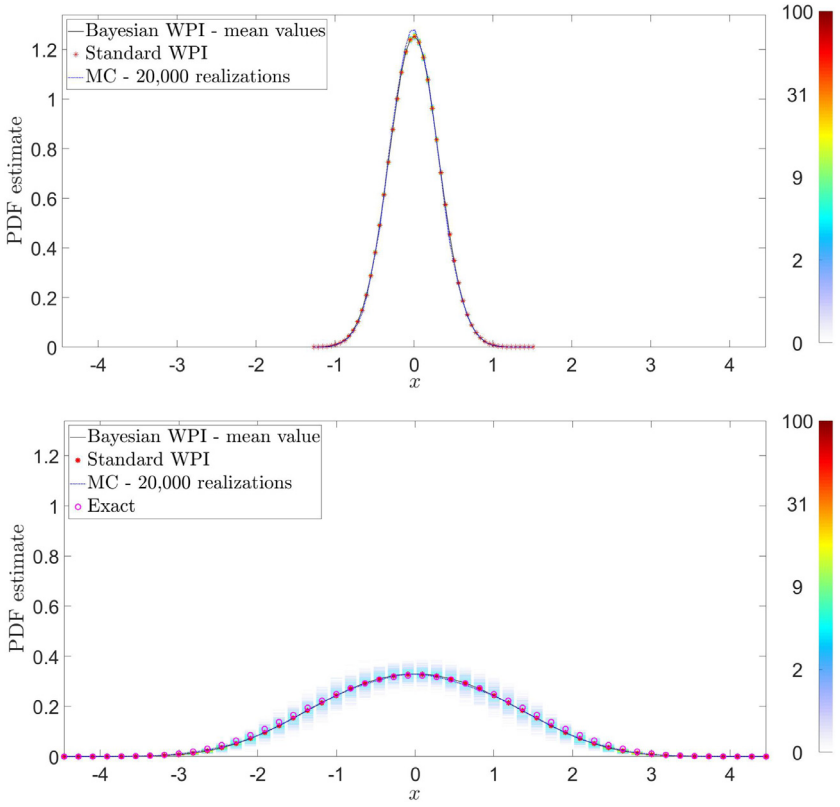


Fig. 7. WPI-based estimates and uncertainty quantification of a Duffing nonlinear oscillator response displacement PDF at $t = 1$ s (top) and at $t = 20$ s - stationary phase (bottom). Comparisons with Monte Carlo simulations (20,000 realizations) and the exact solution for the stationary phase.

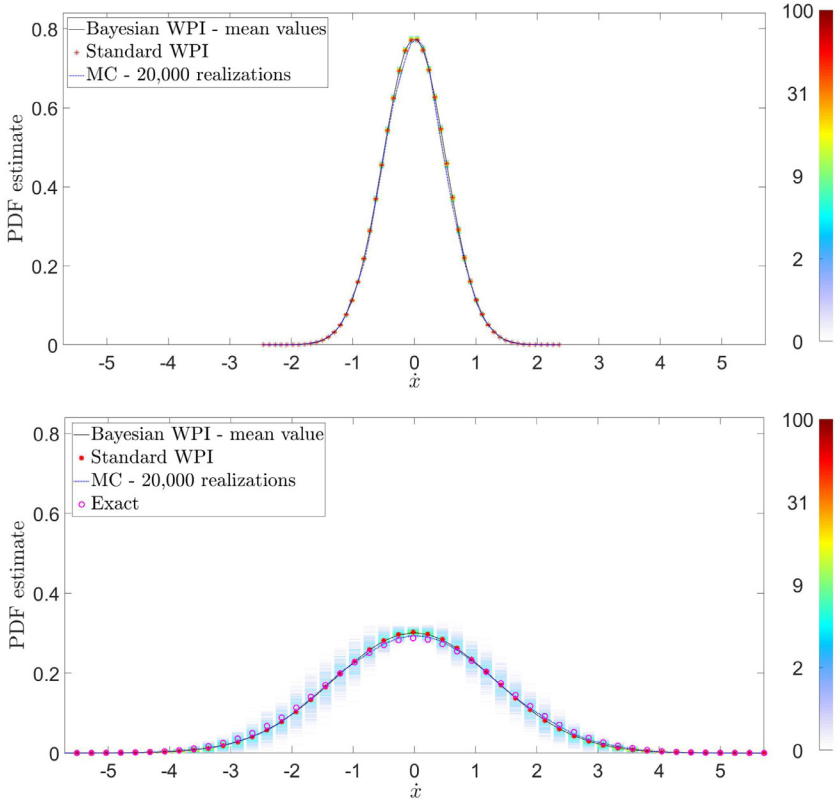


Fig. 8. WPI-based estimates and uncertainty quantification of a Duffing nonlinear oscillator response velocity PDF at $t = 1$ s (top) and at $t = 20$ s - stationary phase (bottom). Comparisons with Monte Carlo simulations (20,000 realizations) and the exact solution for the stationary phase.

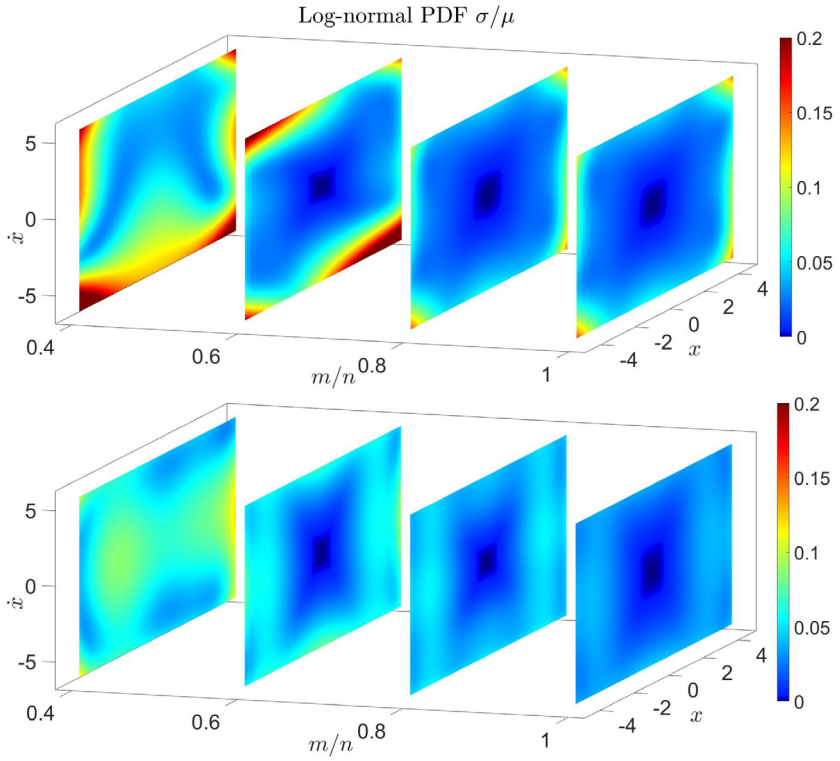


Fig. 9. Reduction of the relative variance σ/μ of the Bayesian WPI estimates with increasing sample ratio m/n. The results refer to the joint response PDF of a Duffing nonlinear oscillator at $t = 20$ s. Comparisons between the standard Fast RVM (top) and the Fast RVM coupled with the proposed optimal sampling scheme of Section 3.4 (bottom).

Further, the joint response PDFs at $t = 1$ s and $t = 20$ s obtained by a standard brute-force implementation of the WPI are shown in Fig. 5 (top). This refers to a discretized PDF domain of $N^2 = 2601$ points. Furthermore, the log-normal distribution of the joint response PDF estimate at an arbitrary point is obtained by utilizing Eqs. (20) and (21). In Fig. 5 (bottom), the mean values of the log-normal distribution are plotted demonstrating excellent agreement with the deterministic estimates in Fig. 5 (top). Moreover, Fig. 6 shows the log-normal distributions corresponding to each and every point of the joint response PDF domain. It becomes clear that a significant advantage of the herein developed technique relates to its capability of quantifying the uncertainty of the response PDF estimates.

Next, the response PDF estimates at two indicative time instants $t = 1$ s and $t = 20$ s (stationary phase), obtained by the standard WPI technique, are shown in Fig. 7 for the displacement x and in Fig. 8 for the velocity $\dot{x}$. These are compared both with pertinent MCS-based estimates (20,000 realizations) and with the closed-form exact stationary PDF of Eq. (30) for $t = 20$ s. Also, the mean values of the Bayesian estimates are included as well. Clearly, these are in good agreement with the respective deterministic WPI-based estimates.

Interestingly, it is observed that the variance, and thus the uncertainty of the joint response PDF estimates, is larger for $t = 20$ s compared to $t = 1$ s. This can be attributed, at least partly, to the larger time interval considered in the BVP of Eq. (4). In other words, it is anticipated that the uncertainty of the PDF estimate increases for larger values of the final time instant t_f, or equivalently, for a final configuration less correlated with the known deterministic initial state $(x_i, \dot{x}_i)$ at t_i. Further, the uncertainty of the estimates depends, obviously, also on the amount of missing data. This is seen in Fig. 9, where the relative standard deviation σ/μ of the estimates obtained by the Bayesian WPI decreases with increasing sample ratio m/n. Furthermore, the superior performance of the Fast RVM coupled with the optimal sampling scheme of Section 3.4 compared to the standard Fast RVM is demonstrated in Fig. 9. In fact, the former appears to converge faster to smaller values of relative variance than the latter for increasing m/n. Note that this is in agreement with similar findings observed in Fig. 3.

Table 1
L_2-norm errors of Duffing nonlinear oscillator joint response PDF estimates compared to MCS data (20,000 realizations). Results refer to various WPI technique formulations and correspond to indicative time instants.

Relative error (%) of WPI compared to MCS	t = 1 s	t = 10 s	t = 15 s	t = 20 s
Brute-force implementation (N^2 points)	5.3	6.9	6.9	7.6
Polynomial approximation (Bayesian CS with $m/n = 0.6$)	8.4	8.4	8.6	8.8
Polynomial approximation (Bayesian CS with $m/n = 1$)	5.5	6.6	6.5	7.3
Polynomial approximation (Standard linear system of equations with $m/n = 0.6$)	5.4	11.5	12.3	11.3

Moreover, in Table 1, the performance of the various WPI technique formulations is assessed in terms of accuracy based on comparisons with MCS data (20,000 realizations). In this regard, l_2-norm errors of the estimated joint response PDFs are reported. Specifically, it is seen that the Bayesian formulation with $m/n = 1$ exhibits practically the same high accuracy degree as the brute-force implementation with $N^2 = 2601$ points. Further, the Bayesian formulation with $m/n = 0.6$ exhibits a robust behavior with relatively constant error values with increasing time instants. The largest error values are reported for the standard WPI formulation coupled with a polynomial PDF approximation. This can be attributed, at least partly, to the fact that the resulting linear system of equations becomes ill-conditioned in many cases, particularly for larger time instants. Notably, the Bayesian WPI formulation appears capable of meliorating this effect to a certain extent.

4.2. Oscillator with asymmetric nonlinearities

Consider next a single-DOF oscillator with asymmetric nonlinearities, whose equation of motion is given by

$$m\ddot{x} + c\dot{x} + kx(1 + \epsilon_0 x) = w(t) \tag{31}$$

where $m = 1$, $k = 1$, $c = 0.2$, $\epsilon_0 = 0.5$ and $w(t)$ is a white noise excitation with a constant power spectrum value $S_0 = 0.0637$. Further, it can be argued that a polynomial expansion used for approximating the response PDF of the oscillator in Eq. (31) is anticipated to be less sparse than the respective one used for the Duffing oscillator in Section 4.1. This is due to the fact that the form of the nonlinearity function in Eq. (31) produces a response PDF that is asymmetric, and thus, an increased number of non-zero terms in the polynomial expansion is required for approximating the PDF shape accurately. Based on this observation, higher-order polynomial bases are employed next. In particular, a 6th order polynomial ($p = 6$) with $n = 28$ is utilized in the following for the arbitrarily selected time instants $t = 1$ s and $t = 2$ s, whereas a 12th order polynomial ($p = 12$) with $n = 91$ is used for $t = 3$ s.

Next, the performance of the various RVM schemes described in Sections 3.3 and 3.4 is assessed in conjunction with the nonlinear oscillator of Eq. (31). Specifically, the relative error $\|\hat{\mathbf{y}}_0 - \mathbf{y}_0\|_2/\|\hat{\mathbf{y}}_0\|_2$ and the variance of $\mathbf{y}_0$ at an indicative time instant $t = 2$ s (both averaged over 1000 trials) are plotted in Fig. 10. $\hat{\mathbf{y}}_0$ denotes the estimate obtained by the standard brute-force WPI technique, whereas $\mathbf{y}_0$ represents the mean of the Bayesian estimate in Eq. (19). Obviously, the RVM schemes coupled with the optimal sampling scheme yield smaller errors and exhibit a lower uncertainty degree compared to their standard counterparts.

In the ensuing analysis, the RVM with the optimal sampling scheme is used with $m/n = 0.6$. This translates into utilizing $m = 17$ points for determining the $n = 28$-dimensional coefficient vector $\mathbf{c}$ for $t = 1$ s and $t = 2$ s. Further, $m = 55$ points are used for determining the $n = 91$-dimensional coefficient vector $\mathbf{c}$ for $t = 3$ s. In this regard, the coefficient vector values obtained by the standard WPI technique are compared in Fig. 11 with the distributions of the estimates obtained by the herein developed Bayesian CS framework. In general, it is seen that the mean values of the Bayesian estimates agree well with the respective deterministic coefficient vector estimates.

Further, the joint response PDFs at $t = 1, 2$ and $3s$ obtained by a standard brute-force implementation of the WPI technique with $N^2 = 2601$ points are shown in Fig. 12 (top). Furthermore, the log-normal distribution of the joint response PDF estimate at an arbitrary point is obtained by utilizing Eqs. (20) and (21). In Fig. 12 (bottom), the mean values of the log-normal distribution are plotted demonstrating excellent agreement with the deterministic estimates in Fig. 12 (top). Moreover, Fig. 13 shows the log-normal distributions corresponding to each and every point of the joint response PDF domain.

Next, the response displacement and velocity PDFs obtained by a standard brute-force implementation of the WPI technique are plotted in Figs. 14 and 15, respectively, for various indicative time instants.

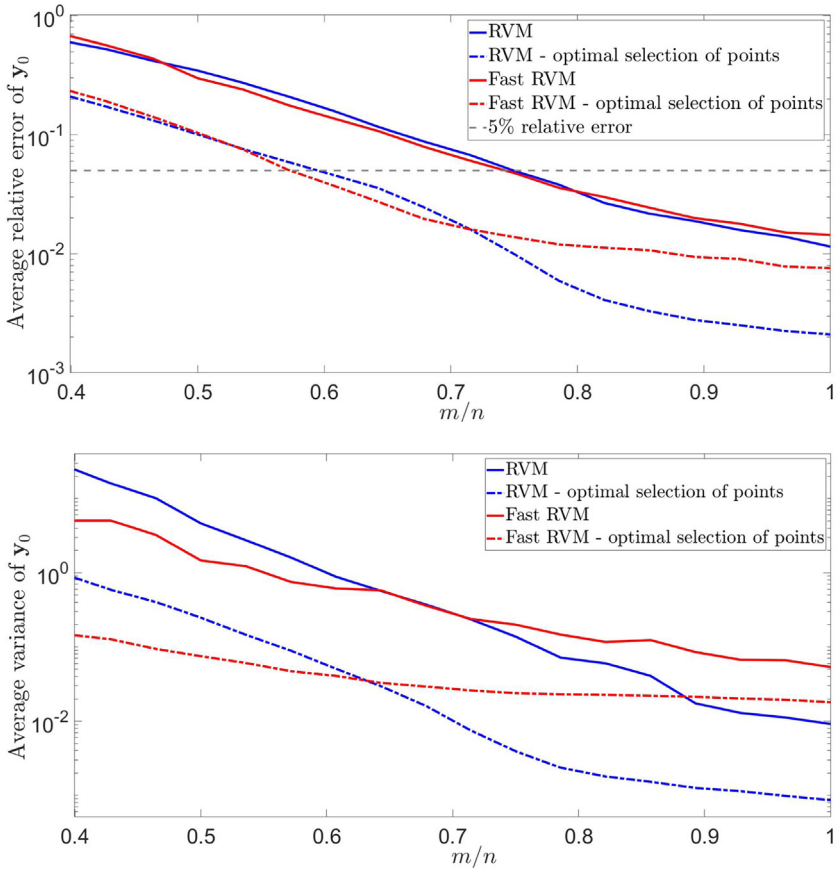


Fig. 10. Assessing the performance of various RVM schemes for determining the joint response PDF of an oscillator with asymmetric nonlinearities at an arbitrarily chosen time instant $t = 2$ s. Average relative error (top) and average variance (bottom) of the Bayesian estimates for various sample ratios m/n.

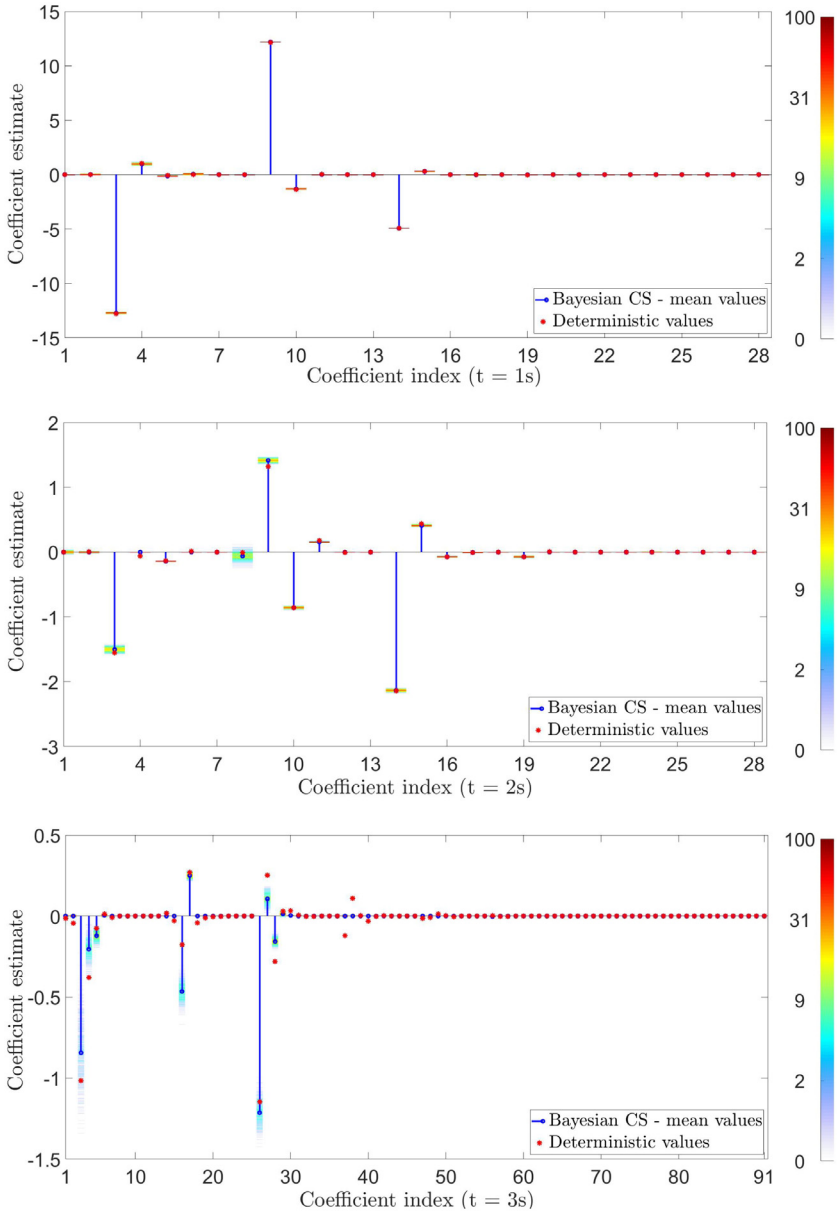


Fig. 11. Uncertainty quantification of joint response PDF expansion coefficient vectors at $t = 1$, 2 and $3s$ for an oscillator with asymmetric nonlinearities. Comparisons between the standard WPI technique in conjunction with a polynomial PDF approximation (deterministic estimates) and the Bayesian WPI formulation (estimates of the coefficient vector distribution based on Eq. (18)).

These are compared both with the mean values of Bayesian estimates and with pertinent MCS data (20,000 realizations), demonstrating a high degree of agreement. Further, Fig. 16 shows the relative standard deviation σ/μ of the estimates obtained by the Bayesian WPI technique decreasing with increasing sample ratio m/n. Note that the RVM coupled with the optimal sampling scheme of Section 3.4 converges faster to smaller values of σ/μ than the standard RVM scheme. This enhanced behavior due to the optimal sampling scheme is anticipated taking also into account the findings in Fig. 10.

In Table 2, the accuracy degree of the various WPI technique formulations is assessed based on comparisons with MCS data (20,000 realizations). Specifically, based on calculations of the l_2-norm errors referring to joint response PDF estimates, it is seen that the Bayesian approach with $m/n = 1$ yields the smallest error, approximately equal to that of the brute-force implementation with $N^2 = 2601$ points. Further, the Bayesian WPI technique exhibits a robust behavior, since even for $m/n = 0.6$ there is only a slight increase in the reported error compared to the case with $m/n = 1$.

Table 2
L_2-norm errors of joint response PDF estimates compared to MCS data (20,000 realizations) for an oscillator with asymmetric nonlinearities. Results refer to various WPI technique formulations and correspond to indicative time instants.

Relative error (%) of WPI compared to MCS	t = 1 s	t = 2 s	t = 3 s
Brute-force implementation (N^2 points)	5.1	5.5	10.8
Polynomial approximation (Bayesian CS with $m/n = 0.6$)	5.4	5.7	13
Polynomial approximation (Bayesian CS with $m/n = 1$)	5.1	5.4	11
Polynomial approximation (Standard linear system of equations with $m/n = 0.6$)	5.1	5.7	11.6

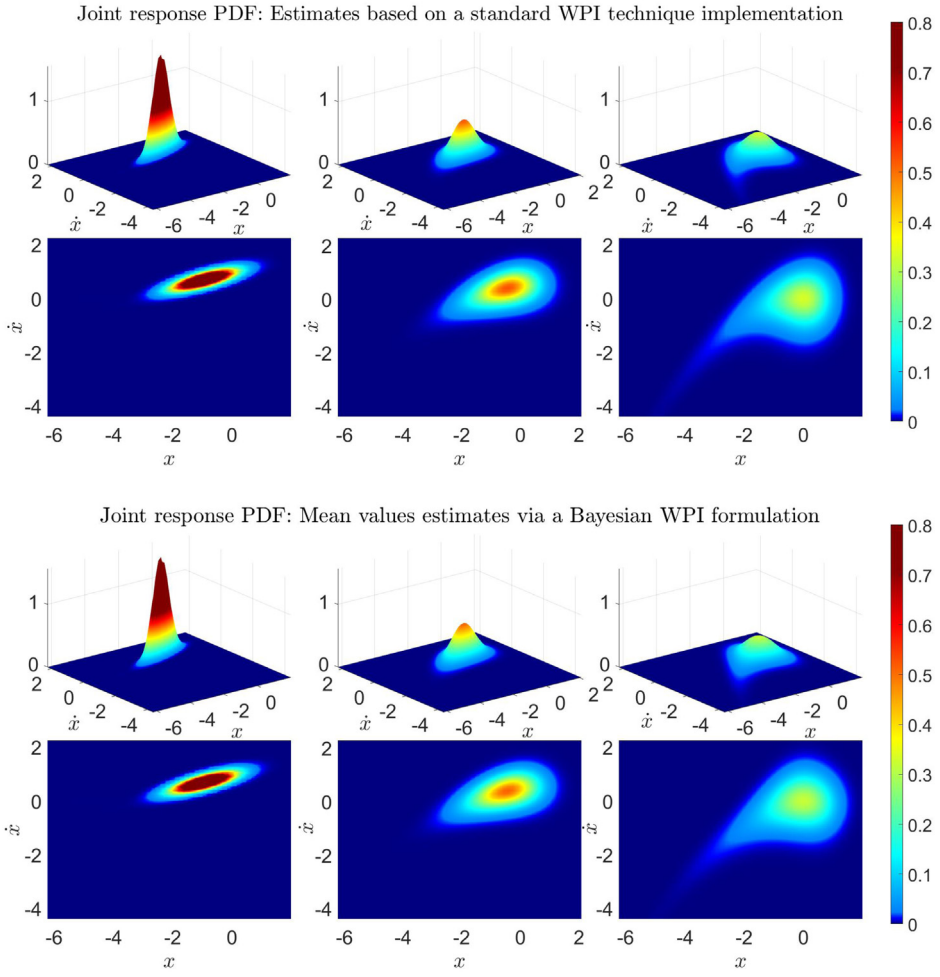


Fig. 12. Joint response PDF of an oscillator with asymmetric nonlinearities at $t = 1$ s (left), $t = 2$ s (middle), and $t = 3$ s (right); standard brute-force WPI formulation (top), and Bayesian formulation — mean values estimates (bottom).

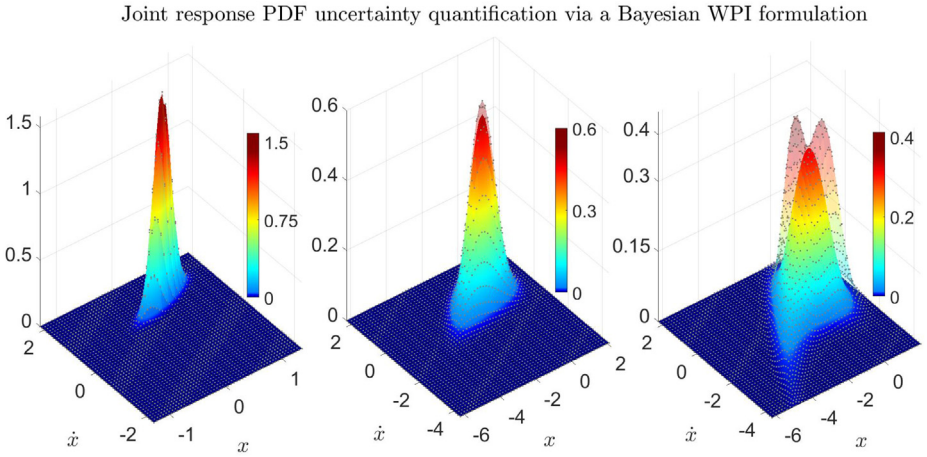


Fig. 13. Uncertainty quantification of joint response PDF estimates based on the developed Bayesian WPI formulation and corresponding to an oscillator with asymmetric nonlinearities for $t = 1$ s (left), $t = 2$ s (middle), $t = 3$ s (right).

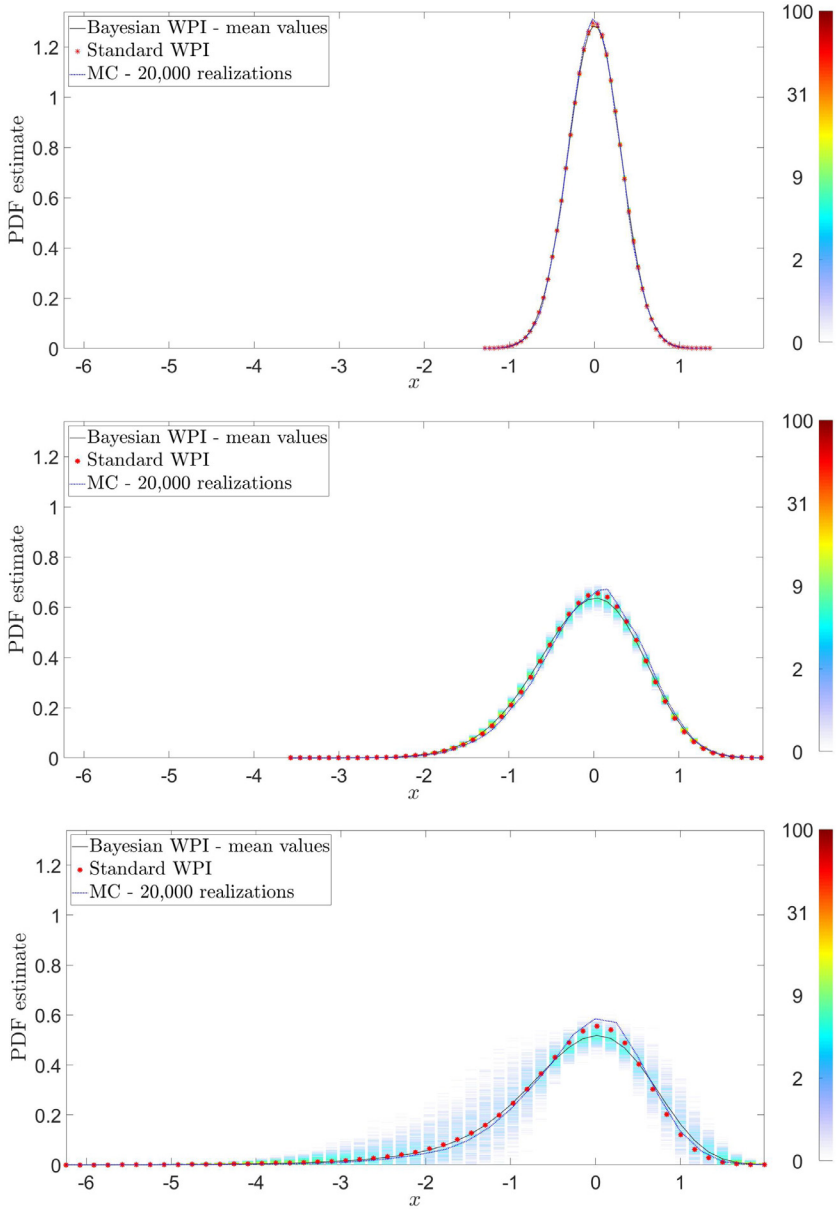


Fig. 14. WPI-based estimates and uncertainty quantification of the response displacement PDF of an oscillator with asymmetric nonlinearities corresponding to (top) $t = 1$ s, (middle) $t = 2$ s, and (bottom) $t = 3$ s. Comparisons with Monte Carlo simulations (20,000 realizations).

5. Concluding remarks

In this paper, the WPI technique for determining the stochastic response of diverse nonlinear dynamical systems has been enhanced based on a Bayesian CS treatment. Specifically, sparse expansions of the polynomial kind have been utilized for representing the system response joint PDF. Next, obtaining PDF values at specific points based on the WPI technique localization capabilities has led to an underdetermined linear system of equations for the expansion coefficients. Further, a solution treatment based on a Bayesian CS formulation has yielded a posterior distribution for the expansion coefficient vector. Clearly, a significant advantage of the herein developed methodology relates to its novel aspect of quantifying the uncertainty of the response PDF estimates obtained by the WPI technique. Furthermore, an adaptive scheme has been proposed based on the quantified uncertainty of

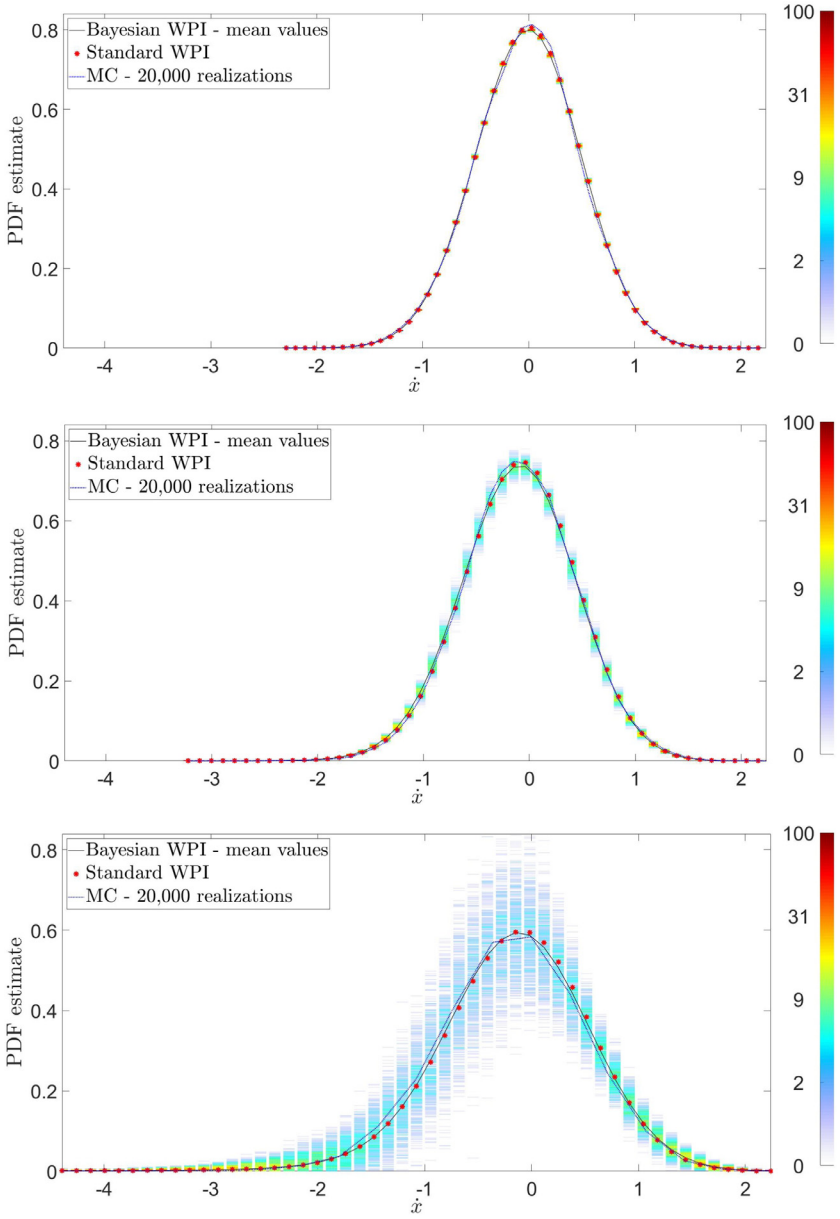


Fig. 15. WPI-based estimates and uncertainty quantification of the response velocity PDF of an oscillator with asymmetric nonlinearities corresponding to (top) $t = 1$ s, (middle) $t = 2$ s, and (bottom) $t = 3$ s. Comparisons with Monte Carlo simulations (20,000 realizations).

the estimates for optimal selection of PDF sample points. In this regard, the total number of BVPs to be solved as part of the WPI technique is reduced, and thus, the associated computational cost decreases. Note that the developed technique can be construed as a generalization and enhancement of earlier efforts in the literature (e.g., [16]) that relied on standard CS tools and provided with deterministic estimates of the response PDF coefficient vector. The efficiency and reliability of the technique have been demonstrated based on comparisons with pertinent MCS data. This has been done in conjunction with two indicative numerical examples pertaining to a Duffing nonlinear oscillator and to an oscillator with asymmetric nonlinearities.

Declaration of competing interest

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

Acknowledgment

I. A. Kougioumtzoglou gratefully acknowledges the support through his CAREER award by the CMMI Division of the National Science Foundation, USA (Award number: 1748537).

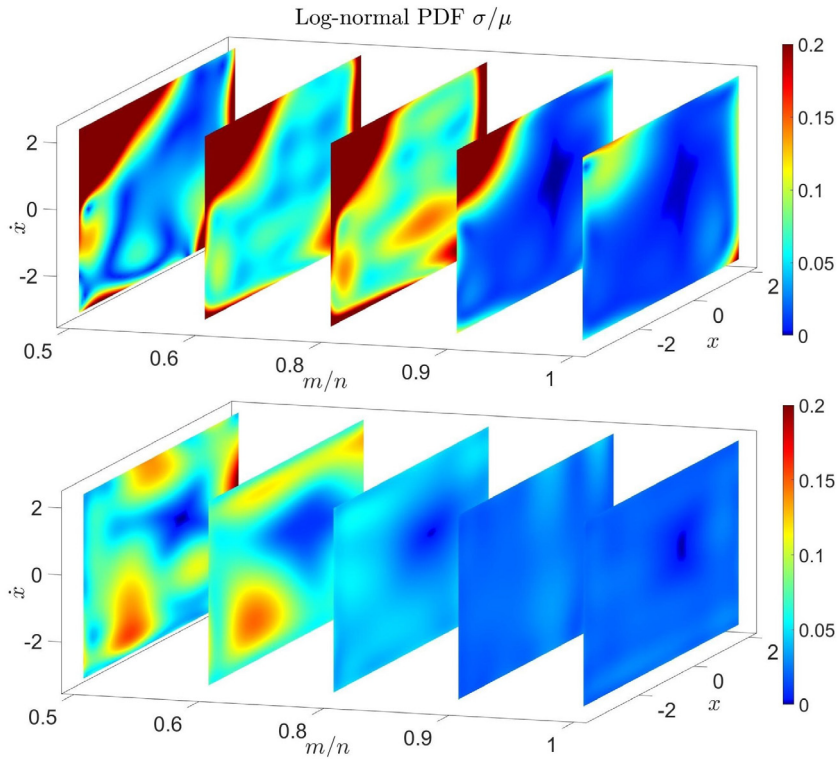


Fig. 16. Reduction of the relative variance σ/μ of the Bayesian WPI estimates with increasing sample ratio m/n. The results refer to the joint response PDF of an oscillator with asymmetric nonlinearities at $t = 2$ s. Comparisons between the standard RVM (top) and the RVM coupled with the proposed optimal sampling scheme of Section 3.4 (bottom).

References

[1] J.B. Roberts, P.D. Spanos, Random Vibration and Statistical Linearization, Courier Corporation, Chelmsford, MA, 2003.

[2] J. Li, J. Chen, Stochastic Dynamics of Structures, Wiley and Sons, 2009.

[3] M. Grigoriu, Stochastic Systems: Uncertainty Quantification and Propagation, Springer, 2012.

[4] N. Wiener, The average of an analytic functional, in: Proceedings of the National Academy of Sciences, USA, Vol. 7, 1921, pp. 253–260.

[5] P.J. Daniell, Integrals in an infinite number of dimensions, Ann. of Math. 20 (1919) 281–288.

[6] R.P. Feynman, Space-time approach to non-relativistic quantum mechanics, Rev. Modern Phys. 20 (1948) 367–387.

[7] M. Chaichian, A. Demichev, Path Integrals in Physics: Stochastic Processes and Quantum Mechanics, Institute of Physics Publishing, Institute of Physics Publishing, Bristol, U.K., 2001.

[8] A.F. Psaros, I.A. Kougioumtzoglou, Functional series expansions and quadratic approximations for enhancing the accuracy of the Wiener path integral technique, ASCE J. Eng. Mech. 146 (7) (2020) 04020065: 1–17.

[9] A. Di Matteo, I.A. Kougioumtzoglou, A. Pirrotta, P.D. Spanos, M. Di Paola, Stochastic response determination of nonlinear oscillators with fractional derivatives elements via the Wiener path integral, Probab. Eng. Mech. 38 (2014) 127–135.

[10] I.A. Kougioumtzoglou, A Wiener path integral solution treatment and effective material properties of a class of one-dimensional stochastic mechanics problems, ASCE J. Eng. Mech. 143 (6) (2017) 04017014: 1–12.

[11] A.F. Psaros, O. Brudastova, G. Malara, I.A. Kougioumtzoglou, Wiener path integral based response determination of nonlinear systems subject to non-white, non-Gaussian, and non-stationary stochastic excitation, J. Sound Vib. 433 (2018) 314–333.

[12] I. Petromichelakis, A.F. Psaros, I.A. Kougioumtzoglou, Stochastic response determination of nonlinear structural systems with singular diffusion matrices: A Wiener path integral variational formulation with constraints, Probab. Eng. Mech. 60 (2020) 103044: 1–15.

[13] I. Petromichelakis, A.F. Psaros, I.A. Kougioumtzoglou, Stochastic response determination and optimization of a class of nonlinear electromechanical energy harvesters: A Wiener path integral approach, Probab. Eng. Mech. 53 (2018) 116–125.

[14] I. Petromichelakis, A.F. Psaros, I.A. Kougioumtzoglou, Stochastic response analysis and reliability-based design optimization of nonlinear electromechanical energy harvesters with fractional derivative elements, ASCE-ASME J. Risk Uncertain. Eng. Syst. B Mech. Eng. 7 (1) (2021) 010901: 1–13.

[15] I. Petromichelakis, I.A. Kougioumtzoglou, Addressing the curse of dimensionality in stochastic dynamics: A Wiener path integral variational formulation with free boundaries, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 476 (2020) 20200385.

[16] A.F. Psaros, I.A. Kougioumtzoglou, I. Petromichelakis, Sparse representations and compressive sampling for enhancing the computational efficiency of the Wiener path integral technique, Mech. Syst. Signal Process. 111 (2018) 87–101.

[17] A.F. Psaros, I. Petromichelakis, I.A. Kougioumtzoglou, Wiener path integrals and multi-dimensional global bases for non-stationary stochastic response determination of structural systems, Mech. Syst. Signal Process. 128 (2020) 551–571.

[18] I.A. Kougioumtzoglou, I. Petromichelakis, F.A. Psaros, Sparse representations and compressive sampling approaches in engineering mechanics: A review of theoretical concepts and diverse applications, Probab. Eng. Mech. 61 (2020) 103082: 1–20.

[19] S. Ji, Y. Xue, L. Carin, Bayesian compressive sensing, IEEE Trans. Signal Process. 56 (6) (2008) 2346–2356.

[20] S. Babacan, R. Molina, K.A. K., Bayesian compressive sensing using Laplace priors, IEEE Trans. Imag. Process 19 (2009) 53–63.

[21] Y. Huang, J.L. Beck, S. Wu, H. Li, Bayesian compressive sensing for approximately sparse signals and application to structural health monitoring signals for data loss recovery, Probab. Eng. Mech. 46 (2016) 62–79.

[22] S. Montoya-Noguera, T. Zhao, Y. Hu, Y. Wang, K.K. Phoon, Simulation of non-stationary non-Gaussian random fields from sparse measurements using Bayesian compressive sampling and Karhunen-Loève expansion, Struct. Saf. 79 (2019) 66–79.

[23] I.A. Kougioumtzoglou, P.D. Spanos, Nonstationary stochastic response determination of nonlinear systems: A Wiener path integral formalism, J. Eng. Mech. 140 (9) (2014) 04014064.

[24] A.F. Psaros, Y. Zhao, I.A. Kougioumtzoglou, An exact closed-form solution for linear multi-degree-of-freedom systems under Gaussian white noise via the Wiener path integral technique, Probab. Eng. Mech. 60 (2020) 103040: 1–12.

[25] I.A. Kougioumtzoglou, A. Di Matteo, P.D. Spanos, A. Pirota, M. Di Paola, An efficient Wiener path integral formulation for stochastic response determination of nonlinear MDOF systems, J. Appl. Mech. 82 (10) (2015) 101005–1.

[26] A. Gelman, J. Hill, Data analysis using regression and multilevel/hierarchical models, Cambridge University Press, 2007.

[27] D.J.C. MacKay, Bayesian methods for backpropagation networks, in: E. Domany, J.L. Van Hemmen, K. Schulten (Eds.), Models of Neural Networks III, Springer, 1994, pp. 211–254.

[28] M.E. Tipping, Sparse Bayesian learning and the relevance vector machine, J. Mach. Learn. Res. 1 (2001) 211–244.

[29] J.L. Beck, L.S. Katafygiotis, Models and their uncertainties I: Bayesian statistical framework, J. Eng. Mech. 124 (1998) 455–461.

[30] M.E. Tipping, The relevance vector machine, in: Advances in Neural Information Processing Systems, Vol. 12, MIT Press, 2000, pp. 652–658.
[31] D. Wipf, J. Palmer, B. Rao, Perspectives on sparse Bayesian learning, in: Advances in Neural Information Processing Systems, NIPS, Vol. 16, 2000.
[32] M.E. Tipping, A. Faul, Fast marginal likelihood maximization for sparse bayesian models, in: 9th International Workshop Artificial Intelligence and Statistics, 2003.
[33] A.C. Faul, M.E. Tipping, Analysis of sparse Bayesian learning. advances in neural information processing systems, in: Advances in Neural Information Processing Systems NIPS 14, 2002, pp. 383–389.
[34] Y.K. Lin, Probabilistic Theory of Structural Dynamics, McGraw-Hill, New York, 1967.