Sparse Representations and Compressive Sampling for enhancing the computational efficiency of the Wiener Path Integral technique

Apostolos F. Psaros[a], Ioannis A. Kougioumtzoglou[a,*], Ioannis Petromichelakis[a]

[a]*Department of Civil Engineering and Engineering Mechanics, Columbia University, 500 W 120th St, New York, NY 10027, United States*

Abstract

The computational efficiency of the Wiener path integral (WPI) technique for determining the stochastic response of diverse dynamical systems is enhanced by exploiting recent developments in the area of sparse representations. Specifically, an appropriate basis for expanding the system joint response probability density function (PDF) is utilized. Next, only very few PDF points are determined based on the localization capabilities of the WPI technique. Further, compressive sampling procedures in conjunction with group sparsity concepts and appropriate optimization algorithms are employed for efficiently determining the coefficients of the system response PDF expansion. It is shown that the herein developed enhancement renders the technique capable of treating readily relatively high-dimensional stochastic systems. Two illustrative numerical examples are considered. The first refers to a single-degree-of-freedom Duffing oscillator exhibiting a bimodal response PDF. In the second example, the 20-variate joint response transition PDF of a 10-degree-of-freedom nonlinear structural system under stochastic excitation is determined. Comparisons with pertinent Monte Carlo simulation data demonstrate the accuracy of the enhanced WPI technique.

Keywords: path integral, nonlinear system, stochastic dynamics, sparse representations, compressive sampling

1. Introduction

Response determination methodologies based on Monte Carlo simulation (MCS) and its variants (e.g., [1, 2]) are considered among the most versatile tools in the area of stochastic engineering dynamics. However, for large scale complex systems, these approaches can be computationally prohibitive. Extensive research in the field during the past few decades has shown that alternative

*Corresponding author
Email address: ikougioum@columbia.edu (Ioannis A. Kougioumtzoglou)

approximate analytical and/or numerical schemes offer efficient ways to address a broad class of problems. State-of-the-art semi-analytical techniques for determining the response of stochastic dynamical systems include moments equations and statistical linearization [3–5], stochastic averaging schemes [6], probability density evolution methodologies [7], Fokker-Planck equation solution techniques [8], as well as numerical schemes based on discretized versions of the Chapman-Kolmogorov equation [9–11]. Additional well-established methodologies relate to stochastic reduced order models, stochastic Galerkin and collocation schemes (e.g., [12, 13]), as well as techniques based on dynamically orthogonal field equations [14]. Nevertheless, solving high-dimensional nonlinear stochastic differential equations (SDEs) remains a persistent challenge in the field of engineering dynamics.

One of the recently developed promising techniques in stochastic engineering dynamics relates to the concept of the Wiener path integral (WPI) [15]. Path integral techniques have proven to be potent tools in theoretical physics, with applications ranging from superfluidity to quantum chromodynamics (e.g., [16]). The notion of path integral, which generalizes integral calculus to functionals, was introduced by Wiener [17] and by Feynman [18], independently. Recently, an approximate WPI based technique has been developed for determining the stochastic response of nonlinear and/or hysteretic multi-degree-of-freedom (MDOF) structural systems [19]. The technique exhibits significant versatility and can account even for systems endowed with fractional derivative terms [20]. Furthermore, it has been extended for addressing certain one-dimensional mechanics problems with random material/media properties [21], while preliminary results towards an error quantification analysis can be found in [22]. From a computational efficiency perspective, recent work by Kougioumtzoglou et al. [23] reduced the computational complexity by, potentially, several orders of magnitude as compared to the original formulation and numerical implementation of the technique.

The objective of this paper is to further enhance the computational efficiency of the WPI technique by exploiting recent developments in the area of sparse representations. Indicatively, sparse expansions of multivariate polynomials have been recently used for numerically solving stochastic (partial) differential equations [24–26]. In this paper, compressive sampling procedures are employed in conjunction with group sparsity concepts and appropriate optimization algorithms for decreasing drastically the computational cost associated with determining the system response probability density function (PDF). It is shown that the herein developed enhancement renders the technique capable of treating readily relatively high-dimensional stochastic systems. Two illustrative numerical examples are considered. The first refers to a single-degree-of-freedom Duffing oscillator exhibiting a bimodal response PDF. In the second example, the 20-variate joint response transition PDF of a 10-DOF nonlinear structural system under stochastic excitation is determined. Comparisons with pertinent MCS data demonstrate the accuracy of the enhanced WPI technique.

2. Wiener Path Integral Technique

2.1. Wiener Path Integral formalism

A wide range of problems in engineering mechanics and dynamics can be described by stochastic equations of the form

$$\boldsymbol{F}\left[\boldsymbol{x}\right] = \boldsymbol{w} \tag{1}$$

where $\boldsymbol{F}\left[.\right]$ represents an arbitrary nonlinear differential operator; $\boldsymbol{w}$ denotes the external excitation; and $\boldsymbol{x}$ is the system response to be determined. It is noted that Kougioumtzoglou [21] has shown recently that the WPI technique can address not only problems subject to stochastic excitation $\boldsymbol{w}(t)$, but also a certain class of one-dimensional mechanics problems with stochastic media properties; that is, stochasticity is embedded in the operator $\boldsymbol{F}\left[.\right]$. Nevertheless, for the purpose of this paper, and without loss of generality, an m-DOF nonlinear dynamical system with stochastic external excitation is considered herein in the form

$$\boldsymbol{M}\ddot{\boldsymbol{x}} + \boldsymbol{C}\dot{\boldsymbol{x}} + \boldsymbol{K}\boldsymbol{x} + \boldsymbol{g}(\boldsymbol{x}, \dot{\boldsymbol{x}}) = \boldsymbol{w}(t) \tag{2}$$

where $\boldsymbol{x}$ is the displacement vector process ($\boldsymbol{x}^T = [x_1 \dots x_m]$); $\boldsymbol{M}$, $\boldsymbol{C}$, $\boldsymbol{K}$ correspond to the $m \times m$ mass, damping and stiffness matrices, respectively; $\boldsymbol{g}(\boldsymbol{x}, \dot{\boldsymbol{x}})$ denotes an arbitrary nonlinear vector function; and $\boldsymbol{w}(t)$ is a white noise stochastic vector process with the power spectrum matrix

$$\boldsymbol{S}_{\boldsymbol{w}} = \begin{bmatrix} S_0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & S_0 \end{bmatrix} \tag{3}$$

Next, relying on the mathematical framework of path integrals [16], the transition PDF $p(\boldsymbol{x}_f, \dot{\boldsymbol{x}}_f, t_f | \boldsymbol{x}_i, \dot{\boldsymbol{x}}_i, t_i)$ can be written as [19]

$$p(\boldsymbol{x}_f, \dot{\boldsymbol{x}}_f, t_f | \boldsymbol{x}_i, \dot{\boldsymbol{x}}_i, t_i) = \int\limits_{\mathcal{C}\{\boldsymbol{x}_i, \dot{\boldsymbol{x}}_i, t_i; \boldsymbol{x}_f, \dot{\boldsymbol{x}}_f, t_f\}} W[\boldsymbol{x}(t)][\mathrm{d}\boldsymbol{x}(t)] \tag{4}$$

with $\{\boldsymbol{x}_i, \dot{\boldsymbol{x}}_i, t_i\}$ denoting the initial state and $\{\boldsymbol{x}_f, \dot{\boldsymbol{x}}_f, t_f\}$ the final state, and $\boldsymbol{x}_i = \boldsymbol{x}(t_i)$, $\boldsymbol{x}_f = \boldsymbol{x}(t_f)$, $\dot{\boldsymbol{x}}_i = \dot{\boldsymbol{x}}(t_i)$ and $\dot{\boldsymbol{x}}_f = \dot{\boldsymbol{x}}(t_f)$. The integral of Eq. (4) represents a functional integration over the space of all possible paths $\mathcal{C}\{\boldsymbol{x}_i, \dot{\boldsymbol{x}}_i, t_i; \boldsymbol{x}_f, \dot{\boldsymbol{x}}_f, t_f\}$, $W[\boldsymbol{x}(t)]$ denotes the probability density functional of the stochastic process in the path space and $[\mathrm{d}\boldsymbol{x}(t)]$ is a functional measure. Further, the probability density functional for the stochastic process $\boldsymbol{x}(t)$ pertaining to the MDOF system of Eq. (2) is defined as (e.g., [19])

$$W[\boldsymbol{x}(t)] = \exp\left(-\int\limits_{t_i}^{t_f} L\left(\boldsymbol{x}, \dot{\boldsymbol{x}}, \ddot{\boldsymbol{x}}\right) \mathrm{d}t\right) \tag{5}$$

where $L\left(\boldsymbol{x}, \dot{\boldsymbol{x}}, \ddot{\boldsymbol{x}}\right)$ denotes the Lagrangian functional given as

$$L\left(\boldsymbol{x}, \dot{\boldsymbol{x}}, \ddot{\boldsymbol{x}}\right) = \frac{1}{2}\left(\boldsymbol{M}\ddot{\boldsymbol{x}} + \boldsymbol{C}\dot{\boldsymbol{x}} + \boldsymbol{K}\boldsymbol{x} + \boldsymbol{g}(\boldsymbol{x}, \dot{\boldsymbol{x}})\right)^T \boldsymbol{B}^{-1} \times \ldots \\ \left(\boldsymbol{M}\ddot{\boldsymbol{x}} + \boldsymbol{C}\dot{\boldsymbol{x}} + \boldsymbol{K}\boldsymbol{x} + \boldsymbol{g}(\boldsymbol{x}, \dot{\boldsymbol{x}})\right) \tag{6}$$

where

$$\boldsymbol{B} = \begin{bmatrix} 2\pi S_0 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & 2\pi S_0 \end{bmatrix} \tag{7}$$

Note that Eq. (5) can be loosely interpreted as the probability assigned to each and every possible path starting from $\{\boldsymbol{x}_i, \dot{\boldsymbol{x}}_i, t_i\}$ and ending at $\{\boldsymbol{x}_f, \dot{\boldsymbol{x}}_f, t_f\}$.

Clearly, the largest contribution to the functional integral of Eq. (4) comes from the trajectory $\boldsymbol{x}_c(t)$ for which the integral in the exponential of Eq. (5) (also known as the stochastic action) becomes as small as possible; see, for instance, [16]. According to calculus of variations (e.g., [27]) this trajectory $\boldsymbol{x}_c(t)$ with fixed endpoints satisfies the extremality condition

$$\delta \int_{t_i}^{t_f} L(\boldsymbol{x}_c, \dot{\boldsymbol{x}}_c, \ddot{\boldsymbol{x}}_c)\mathrm{d}t = 0 \tag{8}$$

which yields the system of Euler-Lagrange (E-L) equations

$$\begin{aligned} \frac{\partial L}{\partial x_{c,1}} - \frac{\partial}{\partial t}\frac{\partial L}{\partial \dot{x}_{c,1}} + \frac{\partial^2}{\partial t^2}\frac{\partial L}{\partial \ddot{x}_{c,1}} &= 0 \\ &\vdots \\ \frac{\partial L}{\partial x_{c,m}} - \frac{\partial}{\partial t}\frac{\partial L}{\partial \dot{x}_{c,m}} + \frac{\partial^2}{\partial t^2}\frac{\partial L}{\partial \ddot{x}_{c,m}} &= 0 \end{aligned} \tag{9}$$

together with $4 \times m$ boundary conditions

$$\begin{aligned} &x_{c,1}(t_i) = x_{1,i}, \quad \dot{x}_{c,1}(t_i) = \dot{x}_{1,i}, \quad x_{c,1}(t_f) = x_{1,f}, \quad \dot{x}_{c,1}(t_f) = \dot{x}_{1,f} \\ &\qquad\qquad\qquad\qquad \vdots \\ &x_{c,m}(t_i) = x_{m,i}, \quad \dot{x}_{c,m}(t_i) = \dot{x}_{m,i}, \quad x_{c,m}(t_f) = x_{m,f}, \quad \dot{x}_{c,m}(t_f) = \dot{x}_{m,f} \end{aligned} \tag{10}$$

Next, solving Eqs. (9)-(10) yields the m-dimensional most probable path, $\boldsymbol{x}_c(t)$, and thus, a single point of the system response transition PDF can be determined as [19]

$$p(\boldsymbol{x}_f, \dot{\boldsymbol{x}}_f, t_f | \boldsymbol{x}_i, \dot{\boldsymbol{x}}_i, t_i) \approx C \ \exp\left(-\int_{t_i}^{t_f} L(\boldsymbol{x}_c, \dot{\boldsymbol{x}}_c, \ddot{\boldsymbol{x}}_c)\mathrm{d}t\right) \tag{11}$$

In Eq. (11), the normalization constant C can be determined by utilizing the condition

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(\boldsymbol{x}_f, \dot{\boldsymbol{x}}_f, t_f | \boldsymbol{x}_i, \dot{\boldsymbol{x}}_i, t_i)\mathrm{d}x_{1,f}\mathrm{d}\dot{x}_{1,f} \ldots \mathrm{d}x_{m,f}\mathrm{d}\dot{x}_{m,f} = 1 \tag{12}$$

It can be readily seen by comparing Eqs. (4) and (11) that in the approximation of Eq. (11) only one trajectory, i.e., the most probable path $\boldsymbol{x}_c(t)$, is considered in evaluating the path integral of Eq. (4). Regarding the degree of this approximation, direct comparisons of Eq. (11) with pertinent MCS data related to various engineering dynamical systems [19, 20] have demonstrated satisfactory accuracy; see also [22].

Further, note that instead of solving the derived E-L Eqs. (9)-(10), an alternative solution approach can be applied for determining the most probable path $\boldsymbol{x}_c(t)$. Specifically, a more direct functional optimization formulation for the expression $\int_{t_i}^{t_f} L(\boldsymbol{x}_c, \dot{\boldsymbol{x}}_c, \ddot{\boldsymbol{x}}_c)\mathrm{d}t$ can be applied, which can be readily combined with a standard Rayleigh-Ritz solution approach; see [20, 21] for more details. Overall, considering fixed initial conditions $(\boldsymbol{x}_i, \dot{\boldsymbol{x}}_i)$ typically (i.e., system initially at rest), both approaches require the solution of a functional minimization problem for determining a single point of the joint response PDF. In the ensuing analysis, adopting a data analysis perspective, this procedure will be referred to as obtaining a measurement of the joint response PDF.

2.2. Numerical Implementation

Although the boundary value problem (BVP) of Eqs. (9)-(10) is amenable to a closed-form analytical solution for a linear dynamical system, i.e., $\boldsymbol{g}(\boldsymbol{x}, \dot{\boldsymbol{x}}) = \boldsymbol{0}$, unfortunately this is not the case, in general, for nonlinear systems. Therefore, a numerical solution scheme needs to be implemented. In this regard, adopting a brute-force numerical solution approach, for each time instant t_f an effective domain of values is considered for the joint response PDF $p(\boldsymbol{x}_f, \dot{\boldsymbol{x}}_f, t_f | \boldsymbol{x}_i, \dot{\boldsymbol{x}}_i, t_i)$. Next, discretizing the effective domain using N points in each dimension, the joint response PDF values are obtained corresponding to the points of the mesh. More specifically, for an m-DOF system corresponding to $2m$ stochastic dimensions (m displacements and m velocities) the number of measurements required is N^{2m}. Clearly, this demonstrates the high computational cost related to a brute force solution scheme implementation, especially for high-dimensional systems.

To address the above computational limitations, Kougioumtzoglou et al. [23] employed a polynomial expansion for the joint response PDF; thus, yielding the required number of PDF measurements equal to the number of the expansion coefficients. Further, it was shown that the computational cost follows a power-law function of the form $\sim (2m)^l / l!$ (where l is the degree of the polynomial), which can be orders of magnitude smaller than N^{2m}. Indicatively, the joint response PDF of a 10-DOF nonlinear dynamical system can be obtained with only 10,626 measurements by utilizing the polynomial approximation, whereas a brute force PDF domain discretization scheme would require 30^{20} measurements (for $N = 30$). However, even with the enhancement in computational efficiency proposed in [23], the related computational cost as a power law function of the number of stochastic dimensions still restricts the applicability of the methodology to relatively low-dimensional systems. In this paper, further

enhancement in the computational efficiency of the WPI technique is achieved by employing sparse representations for the response PDF in conjunction with appropriate optimization algorithms.

3. PDF Approximation and Sparse Representations

3.1. Joint response PDF approximation

The solution approach proposed by Kougioumtzoglou et al. [23] can be construed as a special case of expanding the joint response PDF by employing an appropriate basis. Specifically, without loss of generality and considering fixed initial conditions, the only variables describing the PDF at a time instant t_f are $\boldsymbol{x}_f$ and $\dot{\boldsymbol{x}}_f$. Next, dropping the subscript f for simplicity, the joint response PDF is considered to be a square-integrable function, i.e., $p(\boldsymbol{x}, \dot{\boldsymbol{x}}) \in \mathbb{L}^2(\mathbb{R}^{2m})$. In this regard, $p(\boldsymbol{x}, \dot{\boldsymbol{x}})$ is approximated as

$$p(\boldsymbol{x}, \dot{\boldsymbol{x}}) \approx \exp\left(\sum_{i=1}^{n} c_i d_i(\boldsymbol{x}, \dot{\boldsymbol{x}})\right) \tag{13}$$

where c_i and $d_i(\boldsymbol{x}, \dot{\boldsymbol{x}})$, for $i \in \{1, \ldots, n\}$, denote the expansion coefficients and the basis functions, respectively. Note that Eq. (13) can be written, alternatively, as

$$\log\left(p(\boldsymbol{x}, \dot{\boldsymbol{x}})\right) \approx \sum_{i=1}^{n} c_i d_i(\boldsymbol{x}, \dot{\boldsymbol{x}}) \tag{14}$$

Further, following the selection of n points to perform the approximation, Eq. (14) takes the form of a linear system of n equations, i.e.,

$$\boldsymbol{y}_0 = \boldsymbol{D}\boldsymbol{c} \tag{15}$$

where $\boldsymbol{y}_0 \in \mathbb{R}^{n\times 1}$ is a vector of n points (measurements) of $\log\left(p(\boldsymbol{x}, \dot{\boldsymbol{x}})\right)$, $\boldsymbol{D} \in \mathbb{R}^{n\times n}$ is the basis matrix and $\boldsymbol{c} = [c_1, \ldots, c_n]^T \in \mathbb{R}^{n\times 1}$ is the expansion coefficient vector. A WPI solution approach coupled with Eq. (15) has proved to drastically increase the computational efficiency of the WPI technique [23], as only $n \ll N^{2m}$ BVPs of the form of Eqs. (9)-(10) need to be solved for determining the joint response PDF.

Nevertheless, it is demonstrated herein that further significant decrease in the computational cost is possible, if $r \ll n$ measurements (or, in other words, BVPs to be solved) are utilized in Eq. (15). As shown in the following section, this yields an underdetermined system of equations that can be solved by relying on potent sparse representation concepts and tools.

3.2. Sparse Representations and Compressive Sampling

Compressive sampling (or compressive sensing) procedures are currently revolutionizing the signal processing field [28, 29]. In this section it is shown that by relying on compressive sampling concepts, and by exploiting additional information regarding $p(\boldsymbol{x}, \dot{\boldsymbol{x}})$, the approximation scheme of Eq. (13) can become even

more efficient computationally. The rationale of the herein proposed enhancement relates to using the least amount of joint response PDF measurements (i.e., $r \ll n$ measurements obtained using the WPI technique) for computing the coefficient vector $\boldsymbol{c}$.

If only $r < n$ measurements are obtained, Eq. (14) takes the form of an underdetermined linear system, which can be written as

$$\boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{y}_0 = \boldsymbol{\Phi}\boldsymbol{D}\boldsymbol{c} = \boldsymbol{A}\boldsymbol{c} \tag{16}$$

In Eq. (16) $\boldsymbol{\Phi}$ is an $r \times n$ matrix, also known as compressive sampling matrix [30] as it randomly deletes rows of $\boldsymbol{y}_0$ and $\boldsymbol{D}$. The underdetermined system of Eq. (16) has either no solution, or an infinite number of solutions. Nevertheless, in many cases there is additional information available concerning the coefficient vector $\boldsymbol{c}$. For instance, if only a small number of its components, say k out of n components, are nonzero, then the problem can be regularized and there has been extensive research during the past decade on solution procedures [31]. In particular, the sufficiently sparse ($k \ll n$) coefficient vector $\boldsymbol{c}$ is typically referred to as k-sparse. For such cases, searching for the vector $\hat{\boldsymbol{c}}$ with the least amount of elements that satisfies the condition $\boldsymbol{y} = \boldsymbol{A}\hat{\boldsymbol{c}}$ constitutes a non-convex optimization problem. Although this problem has a unique solution if $\boldsymbol{A}$ has certain desired properties and the number of measurements, r, is sufficiently large (e.g., [29]), it is known to be NP-hard (where NP stands for nondeterministic polynomial time), or in other words, there is no known algorithm for solving it efficiently (e.g., [32]).

To address the above challenge, greedy algorithms can be used to find an approximate solution of the original non-convex problem [31]. Alternatively, the regularization constraint can be relaxed. For example, instead of seeking for the solution with the least amount of elements (or in other words, with the minimum ℓ_0-norm), the solution with the minimum ℓ_1-norm is sought for, alternatively. The problem becomes, therefore, convex and can be readily solved via standard numerical algorithms. However, the price to be paid for such a relaxation approach relates to increasing the number of measurements, r, required for a unique solution [29]; see also [33, 34].

The main question in such problems relates to the properties that $\boldsymbol{A}$ should have in order for the aforementioned minimization problem to have a unique solution. Also, depending on the type of $\boldsymbol{A}$ selected, knowledge of the number of measurements for nearly exact recovery of the coefficient vector $\boldsymbol{c}$ is required in an a priori manner. The latter is known in the sparse representations literature as measurement bound, as a lower bound of r measurements guaranteeing nearly exact recovery of $\boldsymbol{c}$ is sought for; see, e.g., [35] for an introduction to the topic. In this regard, theoretical measurement bounds exist only for certain classes of matrices, e.g., for Gaussian matrices $\boldsymbol{A}$, or random submatrices of Bounded Orthonormal Systems (BOS), such as Fourier, Wavelet and Legendre bases (see [28, 36, 37]). These bounds typically show how the order of magnitude of the required number of measurements r changes with increasing dimension n, and sparsity k. Therefore, they are mainly useful for comparing the performances

of various optimization algorithms and for providing with an indicative number of measurements. In Section 3.5, a more general approach is described, which is often used in practical applications.

3.3. Sparse polynomial approximation and group sparsity

Although approximation strategies based on univariate functions are considered a well-developed topic, there is still active research in approximation schemes utilizing multivariate polynomials (see for example [38]). In the ensuing analysis, the monomial basis (e.g., [39]) is adopted for approximating the exponent of the joint response PDF in Eq. (13), and therefore a polynomial approximation is constructed. The rationale for selecting the above basis relates to the fact that in cases of linear systems (i.e., $\boldsymbol{g}(\boldsymbol{x}, \dot{\boldsymbol{x}}) = 0$) the joint response PDF is Gaussian, or, in other words, the function $\log\left(p(\boldsymbol{x}, \dot{\boldsymbol{x}})\right)$ can be expressed exactly as a second-order polynomial. In the general case, where $\boldsymbol{g}(\boldsymbol{x}, \dot{\boldsymbol{x}}) \neq 0$, $p(\boldsymbol{x}, \dot{\boldsymbol{x}})$ can be construed as a "perturbation" (small or large) from the Gaussian PDF, and thus, more monomials are required to enhance the approximation accuracy. The resulting polynomial is, consequently, of higher order.

Further, to determine the polynomial approximation coefficients, $n = \binom{l+2m}{2m}$ points from $\mathbb{R}^{2m}$ need to be chosen, for an l-degree polynomial. These are the points at which the joint response PDF is sampled using the WPI technique and can be selected either randomly, or based on some kind of optimality criterion to enhance the robustness and accuracy of the approximation (see, e.g., [40]). As noted by Sommariva and Vianello [41], choosing "optimal" approximation points can, also, overcome certain numerical issues that typically accompany the monomial basis, such as the handling of resulting ill-conditioned Vandermonde matrices.

Next, the monomials are ordered based on the graded lexicographical order, which for a 10-DOF dynamical system, for instance, would take the form

$$1 \prec x_1 \prec \cdots \prec \dot{x}_{10} \prec \underbrace{x_1^2 \prec 2x_1x_2 \prec x_2^2 \prec 2x_1x_3 \prec 2x_2x_3 \prec x_3^2 \cdots \prec \dot{x}_{10}^2}_{\text{monomials of order 2}} \prec \ldots \tag{17}$$

Interestingly, this ordering scheme becomes important in the context of sparse polynomial approximation. Numerical examples involving arbitrary nonlinear systems of the form of Eq. (2) have demonstrated that the coefficients corresponding to the Gaussian part of the exponent, i.e., monomials of order 2, are always nonzero, whereas only few of the higher order coefficients are nonzero. In particular, the fact that Gaussian coefficients form a group, which is always active, serves as an additional piece of information that can be exploited. In the framework of sparse representations, this corresponds to group (or structured) sparsity, which is a term describing any kind of structure that the coefficient vector is known to have [42]. For the group sparsity to be considered and exploited, the standard compressive sampling algorithms need to be modified as delineated in the following section. In this regard there are both convex (e.g., [43]) and non-convex approaches (e.g., [44]).

3.4. Optimization Algorithm

In this paper, the StructOMP greedy algorithm proposed by Huang et al. [44] is adopted for addressing the original non-convex problem. It can be construed as a generalization of the widely used Orthogonal Matching Pursuit (OMP) algorithm [45] and is preferred in the ensuing numerical examples over alternative convex approaches, such as Group-LASSO [46]. In fact, for various typical stochastic dynamics problems of the form of Eq. (2), StructOMP has exhibited superior performance, both in terms of convergence rate and of approximation accuracy.

Specifically, the input to StructOMP is the r-length measurement vector $\boldsymbol{y}$, the $r \times n$ matrix $\boldsymbol{A}$ and the group structure (in the form of blocks) that the coefficient vector is anticipated to exhibit. In the herein considered applications the coefficient vector is separated into blocks, with every block corresponding to a single monomial, except for the second-order monomials that are grouped together. In standard sparse vectors, each component of the coefficient vector is considered to have complexity 1. This means that if this coefficient is active, then the coefficient vector will be less sparse by 1. In group sparse vectors each block is assigned a value that describes its complexity, which depends on its coding length (see the original paper by Huang et al. [44] for more details). Obviously, all the single monomials are assigned the same complexity value, whereas the grouped monomials are assigned higher complexity values than the single ones. Additionally, the total complexity of the coefficient vector, s, is the sum of the individual complexities of the blocks used to construct it.

As in Section 3.2, $\boldsymbol{c}$ denotes the original coefficient vector that solves the system of Eq. (15) and $\hat{\boldsymbol{c}}$ the estimated one that solves the system of Eq. (16) using StructOMP. The algorithm selects which block reduces the approximation error

$$err = \|\boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{c}}\|_2 \tag{18}$$

per unit increase of complexity the most (this block is considered to provide the maximum progress to the algorithm), and then assigns values to the coefficients of the selected block via least squares regression. Subsequently, the algorithm finds the next block with the maximum progress and terminates either when err becomes smaller than a prescribed threshold or when the complexity of $\hat{\boldsymbol{c}}$ becomes larger than a prescribed value. For the Performance Analysis in Section 3.5 the latter is used, because the recovery error is measured for fixed complexity s. On the contrary, in the numerical examples in Section 4 the former is used, since the goal is to minimize the recovery error even if a less sparse (or more complex) coefficient vector is used in the expansion.

3.5. Performance Analysis

As noted in Section 3.4, the input to the StructOMP algorithm is the r-length measurement vector $\boldsymbol{y}$, the $r \times n$ matrix $\boldsymbol{A}$ (where $\boldsymbol{A} = \boldsymbol{\Phi D}$) and the group structure that the coefficient vector $\boldsymbol{c}$ is anticipated to have. Thus, a decision has to be made a priori regarding the number r of measurements, the degree of the multivariate polynomial to be used and the group structure

provided as input to StructOMP. First, the degree of the polynomial expansion is selected and the basis matrix $\boldsymbol{D}$, and thus, $\boldsymbol{A}$ is constructed. Next, the group structure is formed using the group of second-order monomials, while the remaining monomials are considered separately as single monomials. Based on the rationale explained in Sections 3.3 and 3.4, since the group of second-order monomials is always active, the complexity of the coefficient vector is directly related only to the number of single monomials (NSM). In addition, given that the more complex the coefficient vector is the more measurements are needed for its accurate recovery, the number of measurements r depends solely on NSM. Therefore, the anticipated NSM has to be decided a priori and a tool is needed to find the corresponding required number of joint response PDF measurements r.

In the absence of theoretical results, novel algorithms are typically tested with the aid of synthetic data before being used in practical applications [45, 47–49]. In this regard, based on the experimental set-up described below, empirical measurement bounds are determined, guaranteeing coefficient vector estimates with bounded error. In particular, for a monomial basis, coefficient vectors with synthetic data are created, with varying numbers of single monomials, and hence, with varying total complexity, s. Next, a value is assigned randomly (e.g., from a Gaussian distribution; see [45]) to each nonzero component, and recovery of these vectors is attempted using StructOMP with only $r < n$ measurements and coefficient vector complexity s. Finally, the average recovery error

$$\frac{\|\boldsymbol{c} - \hat{\boldsymbol{c}}\|_2}{\|\boldsymbol{c}\|_2} \tag{19}$$

is measured over 100 independent runs of the algorithm for each pair $(r/n,\, s/r)$, and the result is shown in Fig. 1. It is observed that for every r/n there is a value of s/r above which sparse approximation becomes relatively inaccurate, or in other words, it changes *phase* (e.g., [47]). This is the reason why the plot in Fig. 1, illustrating the transition from highly accurate recovery (blue) to recovery with significant error (red), is commonly called *Phase Diagram* (e.g., [47]).

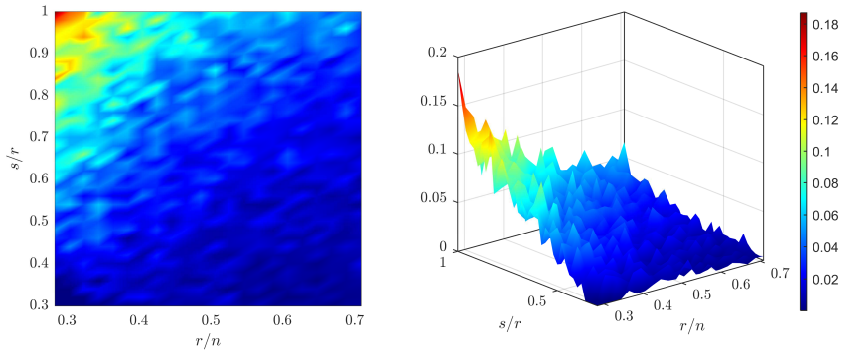

Fig. 1. Phase Diagram for StructOMP using the Monomial Basis. The z-axis corresponds to the average normalized ℓ_2 recovery error, $\frac{\|\boldsymbol{c}-\hat{\boldsymbol{c}}\|_2}{\|\boldsymbol{c}\|_2}$, over 100 runs; the x-axis corresponds to the ratio showing how much underdetermined the problem is, whereas the y-axis corresponds to the ratio showing the level of complexity of the coefficient vector.

The quantities r/n and s/r in Fig. 1 are non-dimensional. Therefore, to use Fig. 1 for creating a measurement bounds plot for an m-DOF system, the actual dimension of the coefficient vector, n, is substituted into r/n. In this regard, the x-axis corresponds to the required number of measurements r, while the y-axis corresponds to the NSM of the coefficient vector. Specifically, for a 10-DOF dynamical system of the form of Eq. (2) with 20 stochastic dimensions and considering a fourth-order polynomial expansion, n becomes 10,626. Fig. 2 shows the estimated measurement bounds for $n = 10{,}626$ with the complexity s represented by the NSM of the coefficient vector. Indicatively, for a 10-DOF linear dynamical system of the form of Eq. (2), only the group of second-order monomials is active, because the joint response PDF is Gaussian, and thus, NSM is equal to zero. Therefore, as shown in Fig. 2 the coefficient vector for such a system can be recovered with less than $r = 3{,}000$ measurements of the joint response PDF using the WPI technique and with average normalized error less than 3%. For a 10-DOF nonlinear dynamical system of the form of Eq. (2), with a non-Gaussian response PDF, NSM is nonzero and as shown in Fig. 2 the number of measurements r has to increase accordingly. Further, a significant additional advantage of employing a sparse approximation treatment relates to the a priori knowledge about the sensitivity of the technique. As shown in Fig. 2 an estimate of the expected increase of the error is readily available in case the coefficient vector sparsity is not predicted accurately.

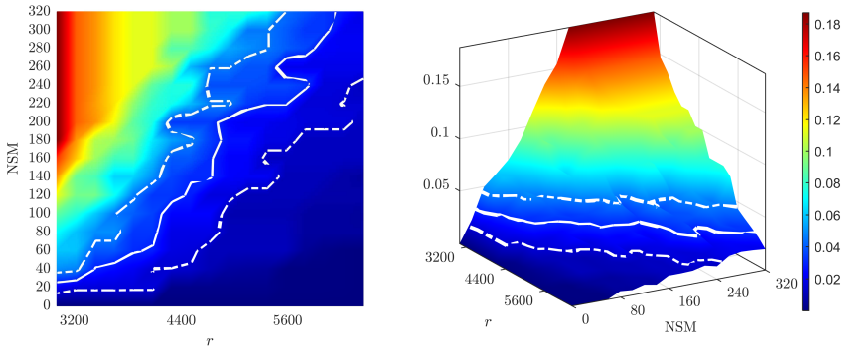

Fig. 2. Measurement bounds for $n = 10{,}626$, corresponding to $m = 10$ and a fourth-order polynomial approximation using StructOMP. The z-axis corresponds to the average normalized ℓ_2 recovery error, $\frac{\|\boldsymbol{c}-\hat{\boldsymbol{c}}\|_2}{\|\boldsymbol{c}\|_2}$, over 100 runs; the x-axis corresponds to the ratio showing how much underdetermined the problem is, whereas the y-axis corresponds to the ratio showing the level of complexity of the coefficient vector. The white solid line indicates the required number of measurements for the error to be smaller than 3%, while the white dashed lines show the deviation of the error by $\pm 1\%$.

3.6. Wiener Path Integral computational efficiency enhancement

For any m-DOF system of the form of Eq. (2), the joint response PDF can be described by Eq. (13) with a length n coefficient vector. Therefore, plots similar to Fig. 2 can be constructed for any dimension m. Such plots are useful for deciding on the number of required measurements and for providing an estimate for the coefficient vector complexity. For instance, for an error less than 3% and selecting the number of single monomials to be 10% of the Gaussian coefficients (see Fig. 2) the required number of measurements can be found for an arbitrary system of m DOFs. In this regard, Fig. 3 shows how the required number of measurements grows with increasing dimension of the system, m. This number is compared with the respective one required for cases where the formulation does not yield an underdetermined problem; that is, the number of measurements is equal to the number of coefficients in the expansion yielding a power law function of the form $\sim (2m)^l / l!$ (see [23]). Further, the number of coefficients corresponding to a linear system response multivariate Gaussian PDF is included as well. It can be readily seen that the proposed approach can be orders of magnitude more efficient than both a brute-force numerical implementation of the WPI [19], and the approximate technique developed by Kougioumtzoglou et al. [23]. Most importantly, as shown in Fig. 3, this enhancement in efficiency becomes even more prevalent as the number of DOFs (or equivalently the number of stochastic dimensions) increases; thus, rendering the herein proposed sparse representation approach indispensable, especially for high-dimensional systems. Of course, it is noted that Fig. 3 shows an indicative rate of growth of r. Systems with complex nonlinearities may require a larger number r. Thus, it is suggested to terminate the StructOMP algorithm only after the addition of a new block does not cause any further reduction of the

approximation error in Eq. (18) (see section 3.4 for more details).

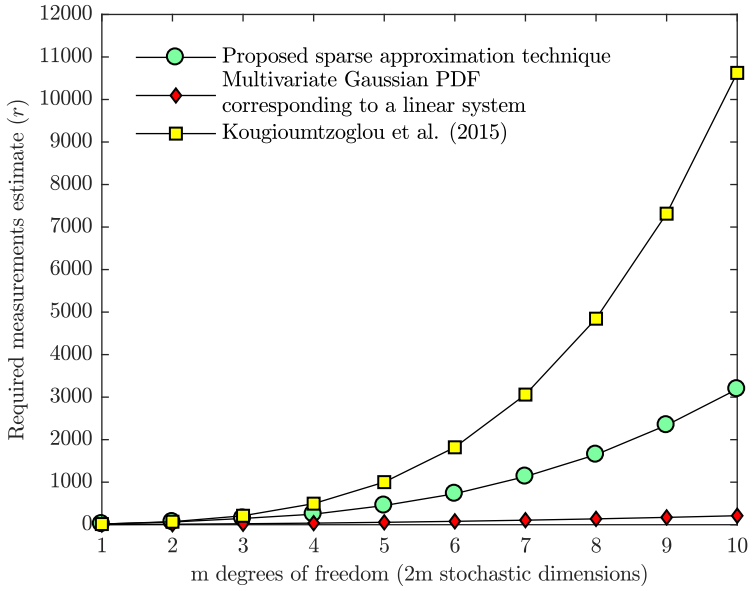

Fig. 3. Required measurements estimate for a general m-DOF system by utilizing the developed sparse approximation technique, and compared with the technique in [23]; the number of measurements required for a multivariate Gaussian PDF is included for completeness.

3.7. Mechanization of the sparse polynomial approximation technique

The mechanization of the developed technique involves the following steps:

(a) Select the polynomial degree l and $n = \binom{l+2m}{2m}$ points of $\mathbb{R}^{2m}$, either randomly (e.g., uniformly distributed), or by employing optimal point selection methodologies (see, e.g., [40]).

(b) Create the basis matrix $\boldsymbol{D}$.

(c) Relying on Fig. 3, select only r out of these n points randomly (e.g., uniformly distributed).

(d) Evaluate $\log\left(p(\boldsymbol{x}, \dot{\boldsymbol{x}})\right)$ at these r points using the WPI technique (Eq. (11)).

(e) Estimate the coefficient vector $\boldsymbol{c}$ using StructOMP (or an alternative appropriate optimization algorithm).

(f) The joint response PDF is given by Eq. (13).

4. Numerical Examples

4.1. SDOF Duffing oscillator with a bimodal response PDF

As shown in Fig. 3 the advantage of the herein developed technique as compared to the implementation of [23] becomes more significant for relatively high-dimensional problems. However, to demonstrate the efficacy of the technique in

determining accurately even relatively complex response PDF shapes, an SDOF Duffing nonlinear oscillator that exhibits a bimodal response PDF is considered first. In this regard, assuming quiescent initial conditions, its equation of motion is given by Eq. (2) with parameter values ($\boldsymbol{M} = 1$; $\boldsymbol{C} = 1$; $\boldsymbol{K} = -0.3$; $\boldsymbol{g} = x^3$; and $S_0 = 0.0637$). It is noted that an exact analytical expression exists for the stationary joint response PDF of this oscillator, given by [50]

$$p(x, \dot{x}) = C \exp\left[\frac{-1}{0.0637\pi}\left(\frac{-0.3x^2}{2} + \frac{x^4}{4} + \frac{\dot{x}^2}{2}\right)\right] \tag{20}$$

where C is a normalization constant. Thus, in addition to utilizing pertinent MCS data, the accuracy degree of the WPI technique can be assessed by direct comparisons with Eq. (20) as well. Next, in implementing the WPI technique summarized in Section 3.7, a 4-th degree polynomial is employed for approximating the response transition PDF $p(x_f, \dot{x}_f, t_f | x_i, \dot{x}_i, t_i)$. Following [23], the number of the expansion coefficients is $n = 15$, however, resorting to the herein proposed technique only $r = 9$ PDF measurements obtained by the WPI are used for determining the joint response PDF of the displacement x and the velocity $\dot{x}$ at a given time instant. In Figs. 4 and 5, the joint PDFs referring to time instants $t = 1$s and $t = 12$s are shown, respectively. For the time instant $t = 1$s, which corresponds to the transient phase of the oscillator dynamics, the high accuracy degree of the technique is demonstrated in Fig. 4 by comparisons with MCS data (50,000 realizations). For the time instant $t = 12$s, which corresponds to the stationary phase of the oscillator dynamics, the high accuracy degree is demonstrated by comparisons with the exact analytical expression given by Eq. (20). The marginal PDFs of x and $\dot{x}$ are shown in Fig. 6 as well. Although the accuracy of the technique depends, in general, on the choice of the polynomial degree, it has been shown in this example that a 4-th degree polynomial is adequate in capturing even relatively complex PDF shapes, such as the bimodal.

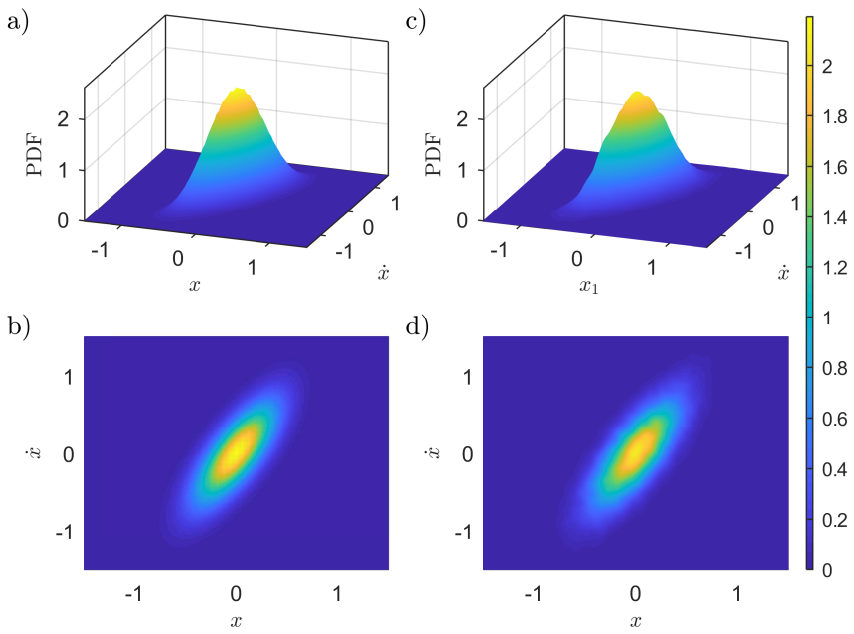

Fig. 4. Joint PDF of $x(t)$ and $\dot{x}(t)$ at time $t = 1$s for a Duffing oscillator with a bimodal response PDF, as obtained via the WPI technique (a - b); comparisons with MCS data - 50,000 realizations (c - d).

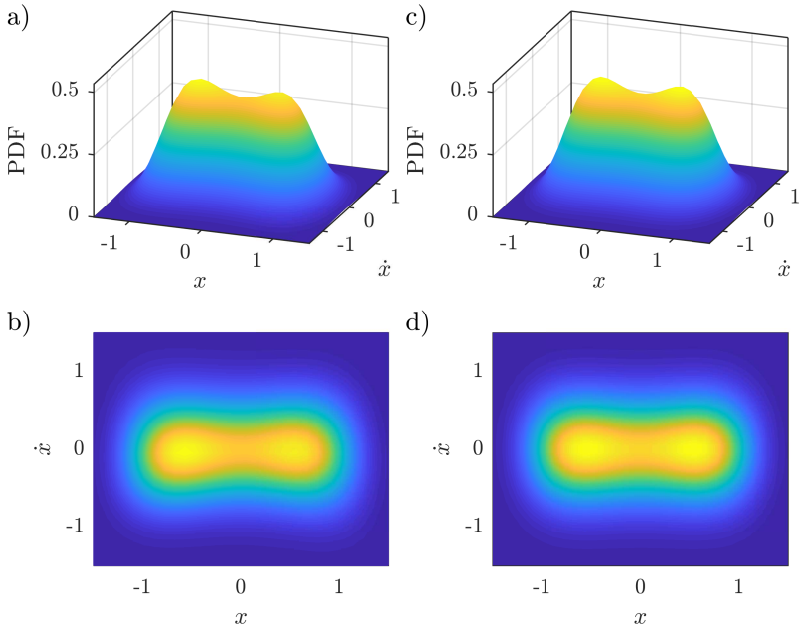

Fig. 5. Joint PDF of $x(t)$ and $\dot{x}(t)$ at time $t = 12$s for a Duffing oscillator with a bimodal response PDF, as obtained via the WPI technique (a - b); comparisons with the exact stationary PDF of Eq. (20) (c - d).

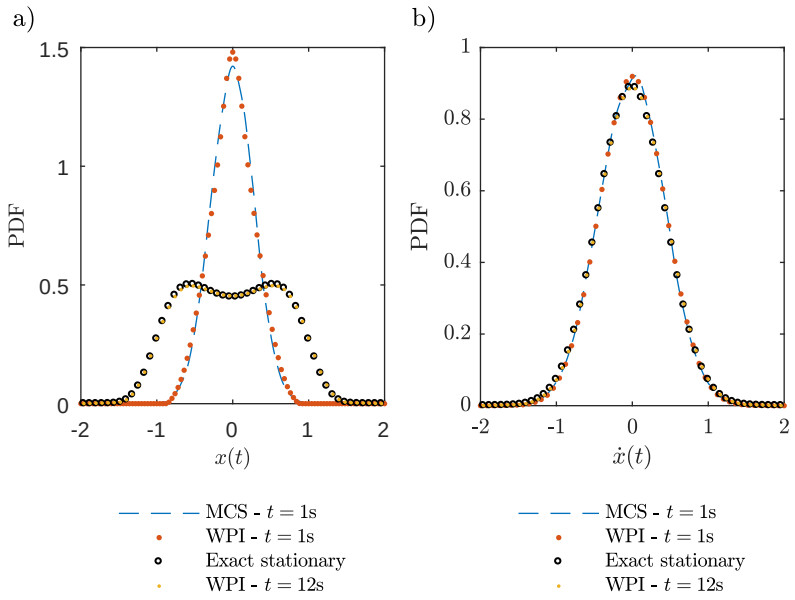

Fig. 6. Marginal PDFs of $x(t)$ and $\dot{x}(t)$ at time instants $t = 1$s and $t = 12$s for a Duffing oscillator with a bimodal response PDF, as obtained via the WPI technique; comparisons with MCS data (50,000 realizations) and the exact stationary PDF of Eq. (20).

4.2. 10-DOF oscillator with damping and stiffness nonlinearities

To demonstrate the accuracy and efficiency of the proposed technique in handling relatively high-dimensional problems, a 10-DOF system of the form of Eq. (2) with cubic damping and stiffness nonlinearites is considered, where

$$\boldsymbol{M} = \begin{bmatrix} m_0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & m_0 \end{bmatrix}, \tag{21}$$

$$\boldsymbol{C} = \begin{bmatrix} 2c_0 & -c_0 & \dots & 0 \\ -c_0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & -c_0 \\ 0 & \dots & -c_0 & 2c_0 \end{bmatrix}, \tag{22}$$

$$\boldsymbol{K} = \begin{bmatrix} 2k_0 & -k_0 & \dots & 0 \\ -k_0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & -k_0 \\ 0 & \dots & -k_0 & 2k_0 \end{bmatrix}, \tag{23}$$

and

$$\boldsymbol{g}(\boldsymbol{x},\dot{\boldsymbol{x}}) = \begin{bmatrix} \epsilon_1 k_0 x_1^3 + \epsilon_2 c_0 \dot{x}_1^3 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{24}$$

The system is excited by a white noise vector process, whose power spectrum matrix is given by Eq. (3), while the parameters values are ($m_0 = 1$; $c_0 = 0.2$; $k_0 = 1$; $\epsilon_1 = 1$; $\epsilon_2 = 1$; and $S_0 = 0.5$). In Figs. 7 and 8, the joint response PDFs for the displacement $x_1(t)$ and velocity $\dot{x}_1(t)$ corresponding to the first DOF obtained by the herein developed efficient WPI technique are plotted for two time instants $t = 1$s and $t = 2$s, respectively. These arbitrarily chosen time instants refer to the non-stationary (transient) phase of the system dynamics. Comparisons with MCS based PDF estimates are included as well. Fig. 9 shows the marginal displacement and velocity PDFs at the above time instants. Figs. 10-12 show the respective results for $x_{10}(t)$ and $\dot{x}_{10}(t)$. In all cases, comparisons with pertinent MCS data demonstrate a high degree of accuracy for the sparse representation based WPI technique.

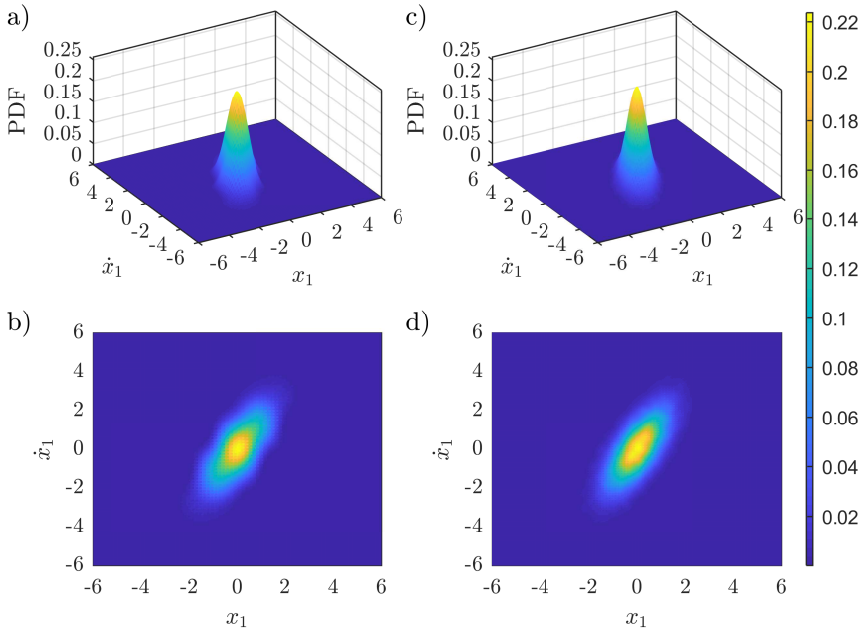

Fig. 7. Joint PDF of $x_1(t)$ and $\dot{x}_1(t)$ at time $t = 1$s, as obtained via the WPI technique (a - b); comparisons with MCS data - 50,000 realizations (c - d).

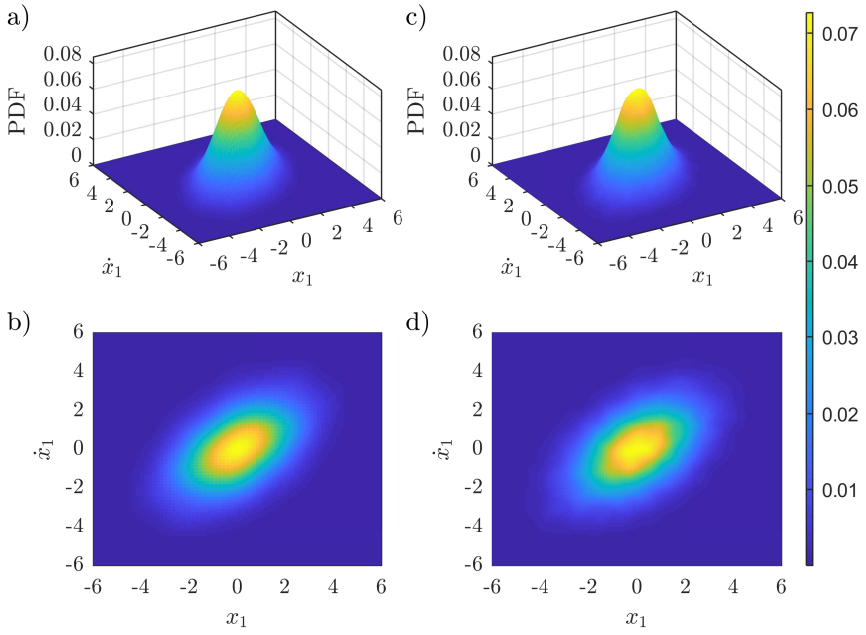

Fig. 8. Joint PDF of $x_1(t)$ and $\dot{x}_1(t)$ at time $t = 2$s, as obtained via the WPI technique (a - b); comparisons with MCS data - 50,000 realizations (c - d).

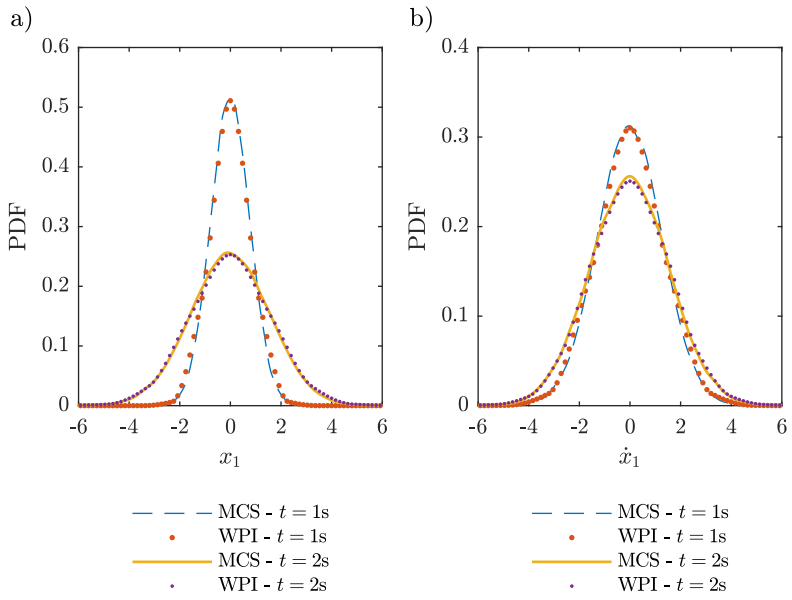

Fig. 9. Marginal PDF of $x_1(t)$ (a) and $\dot{x}_1(t)$ (b) at time instants $t = 1$s and $t = 2$s, as obtained via the WPI technique; comparisons with MCS data (50,000 realizations).

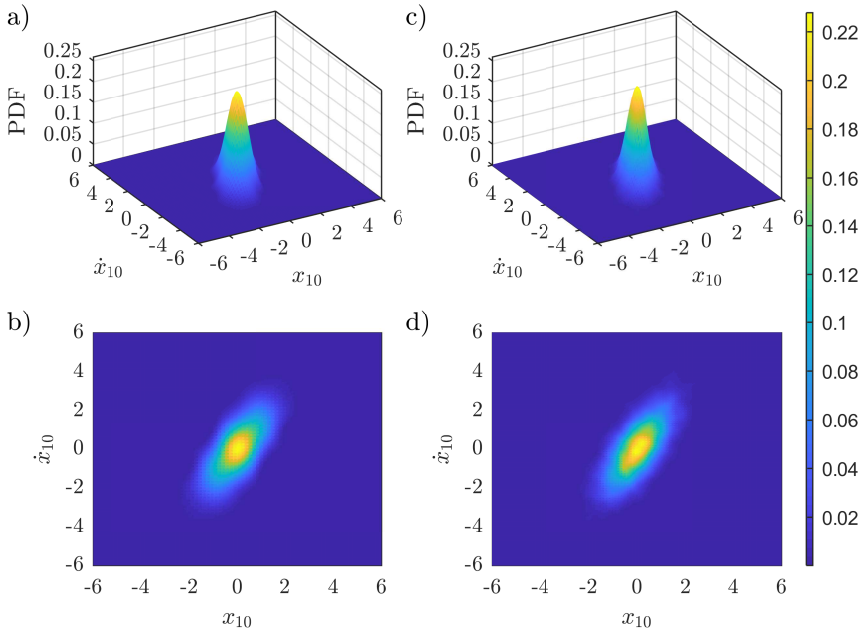

Fig. 10. Joint PDF of $x_{10}(t)$ and $\dot{x}_{10}(t)$ at time $t = 1$s, as obtained via the WPI technique (a - b); comparisons with MCS data - 50,000 realizations (c - d).

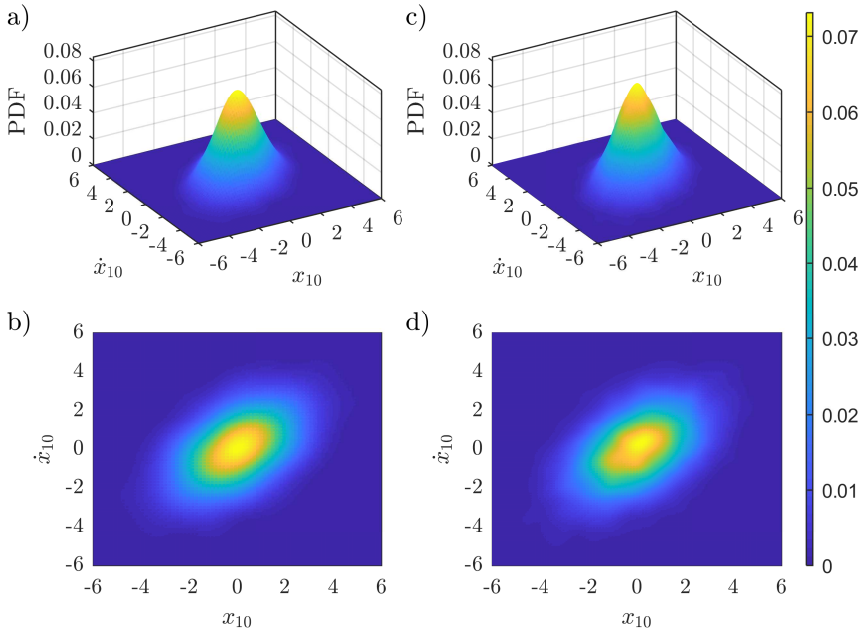

Fig. 11. Joint PDF of $x_{10}(t)$ and $\dot{x}_{10}(t)$ at time $t = 2$s, as obtained via the WPI technique (a - b); comparisons with MCS data - 50,000 realizations (c - d).

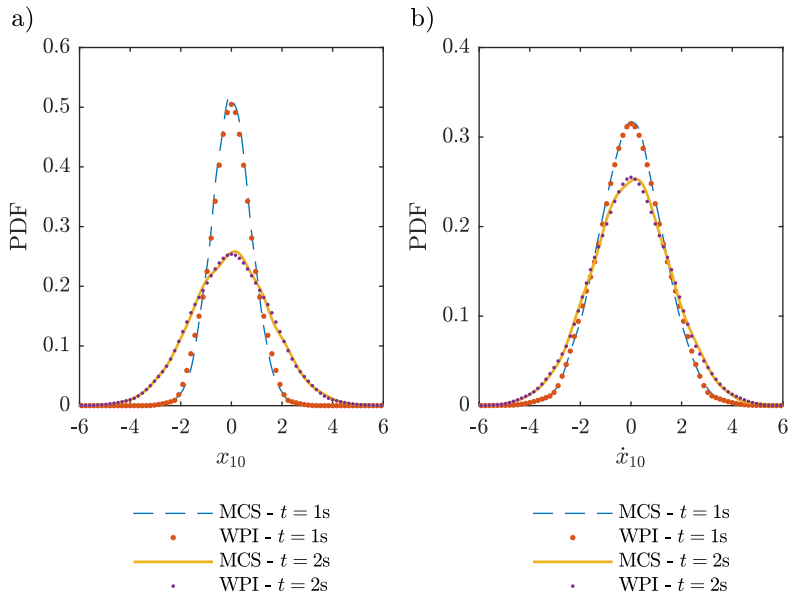

Fig. 12. Marginal PDF of $x_{10}(t)$ (a) and $\dot{x}_{10}(t)$ (b) at time instants $t = 1$s and $t = 2$s, as obtained via the WPI technique; comparisons with MCS data (50,000 realizations).

Regarding computational efficiency, for such a system with 10 DOFs (or in other words, 20 stochastic dimensions), a brute-force WPI numerical implementation requires $\sim 30^{20}$ functional minimization problems of the form of Eqs. (9)-(10) to be solved. Fig. 3 indicates that the polynomial approximation implementation by Kougioumtzoglou et al. [23] requires the solution of only 10,626 functional minimization problems (i.e., measurements of the joint response PDF), whereas resorting to compressive sampling in conjunction with a sparse polynomial approximation technique as developed herein the number of optimization problems to be solved decreases to 3,200. As an indicative order of magnitude, and utilizing a standard PC with up-to-date configurations, the joint response transition PDF of this 10-DOF system is determined in less than an hour by utilizing the herein developed technique. Further, it is noted that according to Fig. 3, the technique becomes even more efficient as compared to the one in [23] for increasing number of DOFs m. In other words, the computational efficiency enhancement becomes even more significant for high-dimensional systems. Of course, note that a relatively accurate MCS based response PDF estimate would require the solution of $\sim 10^6$ deterministic problems; thus, rendering the herein developed WPI technique a significantly more efficient alternative.

5. Conclusion

Although for low-dimensional systems the WPI technique can be significantly more efficient than MCS, its standard numerical implementation has

proven computationally unwieldy for relatively high-dimensional MDOF systems. In this regard, extending the work by Kougioumtzoglou et al. [23] who developed an efficient formulation of the technique, the current paper has proposed an enhanced formulation that decreases the computational cost by potentially several orders of magnitude. Specifically, utilizing an appropriate sparse basis for expanding the system joint response PDF, resorting to the WPI localization features, and employing compressive sampling procedures in conjunction with group sparsity concepts, the response PDF expansion coefficients have been determined efficiently.

It is worth noting that in comparison to the formulation by Kougioumtzoglou et al. [23], the enhancement in computational efficiency becomes more prevalent as the number of stochastic dimensions increases; thus, rendering the herein proposed sparse representation approach indispensable, especially for high-dimensional systems. Two illustrative numerical examples have been considered. The first refers to a single-degree-of-freedom Duffing oscillator exhibiting a bimodal response PDF. Although the accuracy of the technique depends, in general, on the choice of the polynomial degree for a specific problem, it has been shown that a 4-th degree polynomial is adequate in capturing even relatively complex PDF shapes, such as the bimodal. In the second example, the 20-variate joint response transition PDF of a 10-DOF nonlinear structural system under stochastic excitation has been determined. The high degree of accuracy exhibited has been corroborated by comparisons with pertinent MCS data.

Acknowledgement

The authors gratefully acknowledge the support by the CMMI Division of the National Science Foundation, USA (Award number: 1724930).

References

[1] M. Grigoriu, Stochastic Calculus: Applications in Science and Engineering, Birkhäuser Basel, 2002.

[2] S.-K. Au, Y. Wang, Engineering Risk Assessment with Subset Simulation, Wiley, New York, 2014.

[3] J. B. Roberts, P. D. Spanos, Random Vibration and Statistical Linearization, Dover, New York, 2003.

[4] L. Socha, Linearization Methods for Stochastic Dynamic Systems, vol. 730, Springer Science & Business Media, 2007.

[5] V. C. Fragkoulis, I. A. Kougioumtzoglou, A. A. Pantelous, Statistical linearization of nonlinear structural systems with singular matrices, Journal of Engineering Mechanics 142 (9) (2016) 04016063.

[6] P. D. Spanos, I. A. Kougioumtzoglou, K. R. dos Santos, A. T. Beck, Stochastic Averaging of Nonlinear Oscillators: Hilbert Transform Perspective, Journal of Engineering Mechanics 144 (2) (2017) 04017173.

[7] J. Li, J. Chen, Stochastic Dynamics of Structures, Wiley, New York, 2009.

[8] H. Risken, The Fokker-Planck Equation. Methods of Solution and Applications, Springer-Verlag, Berlin, 1996.

[9] M. F. Wehner, W. Wolfer, Numerical evaluation of path-integral solutions to Fokker-Planck equations, Physical Review A 27 (5) (1983) 2663.

[10] A. Naess, V. Moe, Efficient path integration methods for nonlinear dynamic systems, Probabilistic Engineering Mechanics 15 (2) (2000) 221–231.

[11] I. A. Kougioumtzoglou, P. D. Spanos, Response and first-passage statistics of nonlinear oscillators via a numerical path integral approach, Journal of Engineering Mechanics 139 (9) (2012) 1207–1217.

[12] R. G. Ghanem, P. D. Spanos, Stochastic Finite Elements: A Spectral Approach, Dover, New York, 2003.

[13] R. Field, M. Grigoriu, J. Emery, On the efficacy of stochastic collocation, stochastic Galerkin, and stochastic reduced order models for solving stochastic problems, Probabilistic Engineering Mechanics 41 (2015) 60–72.

[14] T. P. Sapsis, P. F. Lermusiaux, Dynamically orthogonal field equations for continuous stochastic dynamical systems, Physica D: Nonlinear Phenomena 238 (23) (2009) 2347–2360.

[15] I. A. Kougioumtzoglou, P. D. Spanos, An analytical Wiener Path Integral technique for non-stationary response determination of nonlinear oscillators, Probabilistic Engineering Mechanics 28 (2012) 125–131.

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

[17] N. Wiener, The average of an analytic functional and the Brownian movement, Proceedings of the National Academy of Sciences 7 (10) (1921) 294–298.

[18] R. P. Feynman, Space-time approach to non-relativistic quantum mechanics, Reviews of Modern Physics 20 (2) (1948) 367.

[19] I. A. Kougioumtzoglou, P. D. Spanos, Nonstationary stochastic response determination of nonlinear systems: A Wiener Path Integral formalism, Journal of Engineering Mechanics 140 (9) (2014) 04014064.

[20] 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, Probabilistic Engineering Mechanics 38 (2014) 127–135.

[21] I. A. Kougioumtzoglou, A Wiener Path Integral Solution Treatment and Effective Material Properties of a Class of One-Dimensional Stochastic Mechanics Problems, Journal of Engineering Mechanics 143 (6) (2017) 04017014.

[22] A. T. Meimaris, I. A. Kougioumtzoglou, A. A. Pantelous, A closed form approximation and error quantification for the response transition probability density function of a class of stochastic differential equations, Probabilistic Engineering Mechanics (2017) doi:http://dx.doi.org/10.1016/j.probengmech.2017.07.005.

[23] I. A. Kougioumtzoglou, A. Di Matteo, P. D. Spanos, A. Pirrotta, M. Di Paola, An efficient Wiener Path Integral technique formulation for stochastic response determination of nonlinear MDOF systems, Journal of Applied Mechanics 82 (10) (2015) 101005.

[24] A. Cohen, R. Devore, C. Schwab, Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDE's, Analysis and Applications 9 (01) (2011) 11–47.

[25] A. Doostan, H. Owhadi, A non-adapted sparse approximation of PDEs with stochastic inputs, Journal of Computational Physics 230 (8) (2011) 3015–3034.

[26] X. Yang, G. E. Karniadakis, Reweighted $\ell 1$ minimization method for stochastic elliptic differential equations, Journal of Computational Physics 248 (2013) 87–108.

[27] G. M. Ewing, Calculus of Variations with Applications, Dover, New York, 1985.

[28] S. Foucart, H. Rauhut, A Mathematical Introduction to Compressive Sensing, Birkhäuser Basel, 2013.

[29] I. Rish, G. Grabarnik, Sparse Modeling: Theory, Algorithms, and Applications, CRC Press, 2014.

[30] E. J. Candès, Compressive Sampling, in: Proceedings of the International Congress of Mathematicians, vol. 3, Madrid, Spain, 1433–1452, 2006.

[31] Y. C. Eldar, G. Kutyniok, Compressed Sensing: Theory and Applications, Cambridge University Press, 2012.

[32] G. Davis, S. Mallat, M. Avellaneda, Adaptive greedy approximations, Constructive Approximation 13 (1) (1997) 57–98.

[33] L. Comerford, I. A. Kougioumtzoglou, M. Beer, Compressive sensing based stochastic process power spectrum estimation subject to missing data, Probabilistic Engineering Mechanics 44 (2016) 66–76.

[34] Y. Zhang, L. Comerford, I. A. Kougioumtzoglou, M. Beer, Lp-norm minimization for stochastic process power spectrum estimation subject to incomplete data, Mechanical Systems and Signal Processing 101 (2018) 361–376.

[35] E. J. Candès, J. Romberg, T. Tao, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, IEEE Transactions on Information Theory 52 (2) (2006) 489–509.

[36] J.-L. Starck, F. Murtagh, J. M. Fadili, Sparse Image and Signal Processing: Wavelets and Related Geometric Multiscale Analysis, Cambridge University Press, 2016.

[37] H. Rauhut, R. Ward, Sparse Legendre expansions via $\ell 1$-minimization, Journal of Approximation Theory 164 (5) (2012) 517–533.

[38] M. Gasca, T. Sauer, Polynomial interpolation in several variables, Advances in Computational Mathematics 12 (4) (2000) 377–410.

[39] P. J. Olver, On multivariate interpolation, Studies in Applied Mathematics 116 (2) (2006) 201–240.

[40] M. Van Barel, M. Humet, L. Sorber, Approximating optimal point configurations for multivariate polynomial interpolation, Electronic Transactions on Numerical Analysis 42 (2014) 41–63.

[41] A. Sommariva, M. Vianello, Computing approximate Fekete points by QR factorizations of Vandermonde matrices, Computers & Mathematics with Applications 57 (8) (2009) 1324–1336.

[42] R. Jenatton, J.-Y. Audibert, F. Bach, Structured variable selection with sparsity-inducing norms, Journal of Machine Learning Research 12 (Oct) (2011) 2777–2824.

[43] F. Bach, R. Jenatton, J. Mairal, G. Obozinski, Structured sparsity through convex optimization, Statistical Science (2012) 450–468.

[44] J. Huang, T. Zhang, D. Metaxas, Learning with structured sparsity, Journal of Machine Learning Research 12 (Nov) (2011) 3371–3412.

[45] J. A. Tropp, A. C. Gilbert, Signal recovery from random measurements via orthogonal matching pursuit, IEEE Transactions on Information Theory 53 (12) (2007) 4655–4666.

[46] M. Yuan, Y. Lin, Model selection and estimation in regression with grouped variables, Journal of the Royal Statistical Society: Series B (Statistical Methodology) 68 (1) (2006) 49–67.

[47] D. Donoho, J. Tanner, Observed universality of phase transitions in high-dimensional geometry, with implications for modern data analysis and signal processing, Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 367 (1906) (2009) 4273–4293.

[48] P. Jain, A. Tewari, I. S. Dhillon, Orthogonal matching pursuit with replacement, in: Advances in Neural Information Processing Systems, 1215–1223, 2011.

[49] D. L. Donoho, Y. Tsaig, I. Drori, J.-L. Starck, Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit, IEEE Transactions on Information Theory 58 (2) (2012) 1094–1121.

[50] Y.-K. Lin, Probabilistic Theory of Structural Dynamics, McGraw-Hill, New York, 1967.