Mechanical Systems and Signal Processing 153 (2021) 107534

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

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

Wiener path integral most probable path determination: A computational algebraic geometry solution treatment

Ioannis Petromichelakis [a], Rúbia M. Bosse [b], Ioannis A. Kougioumtzoglou [a,*], André T. Beck [b]

[a] Department of Civil Engineering and Engineering Mechanics, Columbia University, 500W 120th St, New York, NY 10027, United States
[b] Department of Structural Engineering, São Carlos School of Engineering, University of São Paulo, 13566-590 São Carlos, SP, Brazil

ARTICLE INFO

Article history:
Received 17 July 2020
Received in revised form 23 September 2020
Accepted 6 December 2020

Keywords:
Path integral
Stochastic dynamics
Nonlinear systems
Gröbner basis
Numerical optimization

ABSTRACT

The recently developed Wiener path integral (WPI) technique for determining the stochastic response of diverse nonlinear systems relies on solving a functional minimization problem for the most probable path, which is then utilized for evaluating a specific point of the system joint response probability density function (PDF). However, although various numerical optimization algorithms can be employed for determining the WPI most probable path, there is generally no guarantee that the selected algorithm converges to a global extremum.

In this paper, first, a Newton's optimization scheme is proposed for determining the most probable path, and various convergence behavior aspects are elucidated. Second, the existence of a unique global minimum and the convexity of the objective function of the considered nonlinear system are demonstrated by resorting to computational algebraic geometry concepts and tools, such as Gröbner bases. Several numerical examples pertaining to diverse nonlinear oscillators are considered, where it is proved that the associated objective functions are convex, and that the proposed Newton's scheme converges to the globally optimum most probable path. Comparisons with pertinent Monte Carlo simulation data are included as well for demonstrating the reliability of the WPI technique.

1. Introduction

Monte Carlo simulation (MCS) has been, undoubtedly, one of the most versatile techniques for addressing stochastic dynamics problems and for determining response statistics of complex dynamical systems (e.g., [1–3]). Nevertheless, the associated computational cost can become prohibitive eventually, and thus, there is merit in developing alternative efficient semi-analytical solution techniques. In this regard, indicative techniques developed over the past few decades include statistical linearization, stochastic averaging, perturbation approaches, discrete Chapman-Kolmogorov equation schemes, Fokker–Planck equation solution techniques, probability density evolution methods, and polynomial chaos expansions. The interested reader is directed to various standard books in the field for a detailed presentation (e.g., [4–9]).

Recently, relying on pioneering work by Wiener [10] and on fundamental contributions in the field of theoretical physics (e.g., [11–23]; see also the books in [24,25] for a broader perspective), a semi-analytical technique based on the concept of Wiener path integral (WPI) has been developed in the field of stochastic engineering dynamics for determining the stochastic response of diverse nonlinear structural and mechanical systems (e.g., [26,27]). In fact, the technique, which relies on func-

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

https://doi.org/10.1016/j.ymssp.2020.107534
0888-3270/

tional integration concepts and on calculus of variations tools, exhibits both computational efficiency and satisfactory accuracy in evaluating the system joint response probability density function (PDF) (e.g., [28–31]). Further, the WPI technique exhibits versatility in addressing diverse system behaviors, including hysteresis and fractional derivative modeling (e.g., [33,32]), and in accounting for various rather sophisticated descriptions of stochastic excitations [34].

An integral part of the standard implementation of the WPI technique relates to a variational treatment for deriving a functional minimization problem. Solving this optimization problem yields the most probable path, which is used for evaluating a specific point of the joint response PDF (e.g., [31,35]). Clearly, a wide range of numerical optimization schemes can be employed for determining the WPI most probable path (e.g., [33,36]). However, there is generally no guarantee that the selected optimization algorithm converges to the global minimum (instead of a local minimum). Of course, it can be argued that the relatively high accuracy degree exhibited by the WPI technique, based on comparisons with pertinent MCS data in a plethora of numerical examples (e.g., [28,34]), can be construed as an indication of determining successfully the optimal most probable path. Nevertheless, it becomes clear that there is a need for pursuing the challenging task of proving the existence of a unique global minimum and/or the convexity of the objective function corresponding to an arbitrary nonlinear system under consideration.

In this paper, first, a Newton's numerical optimization scheme is developed for determining the most probable path. The rationale relates to the fact that, for the special case of linear systems, the objective function is not only convex, but also quadratic; and thus, a Newton's scheme appears to be an ideal choice as it converges in only one iteration to the unique global extremum (e.g., [36]). This convergence behavior indicates that a Newton's scheme can be a suitable choice also for nonlinear systems, since their response behavior can be construed as a perturbation (not necessarily small) from the linear regime. Further, certain convergence properties of the scheme are derived and discussed. Second, demonstrating the potential convexity (and thus, the existence of a global extremum) of the functional to be minimized is addressed by resorting to computational algebraic geometry concepts and tools such as Gröbner bases (e.g., [37–41]). Various numerical examples pertaining to diverse nonlinear oscillators are considered, where it is proved that the associated objective functions are convex, and that the proposed Newton's scheme converges to the globally optimum most probable path. Comparisons with MCS-based estimates are included as well for demonstrating the reliability of the WPI technique.

2. Wiener path integral technique overview

2.1. Wiener path integral representation and most probable path approximation

In this section, the basic elements of a recently developed stochastic response determination technique based on the concept of Wiener path integral are presented for completeness; see also [33,34] for a more detailed discussion. In this regard, consider a stochastically excited nonlinear multi-degree-of-freedom (MDOF) system, whose governing equation is given by

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

In Eq. (1), $\boldsymbol{x} = \left[x_j(t)\right]_{n\times 1}$ represents the n-dimensional response displacement vector, $\mathbf{M}$ denotes the $n \times n$ mass matrix considered to be diagonal according to standard modeling in structural dynamics (e.g., [6]), and $\mathbf{g} = \left[g_j(\boldsymbol{x}, \dot{\boldsymbol{x}}, t)\right]_{n\times 1}$ is an arbitrary nonlinear n-dimensional vector-valued function, which can account also for hysteretic response behaviors (e.g., [33]). Further, $\boldsymbol{w}$ is a white noise stochastic excitation vector process with $\mathbb{E}[\boldsymbol{w}(t)] = 0$ and $\mathbb{E}\left[\boldsymbol{w}(t)\boldsymbol{w}^T(t-\tau)\right] = \mathbf{D}\delta(\tau)$, where $\mathbf{D} \in \mathbb{R}^{n\times n}$ is a deterministic coefficient matrix.

As shown in [33,34], the joint response transition PDF corresponding to the system of Eq. (1) can be expressed as a functional integral (or WPI) over the space of possible paths with fixed boundary conditions that the response process can follow; that is,

$$p(\boldsymbol{x}_f, \dot{\boldsymbol{x}}_f, t_f | \boldsymbol{x}_0, \dot{\boldsymbol{x}}_0, t_0) = \int_{\mathcal{C}\left\{\boldsymbol{x}_0, \boldsymbol{v}_0, t_0, \boldsymbol{x}_f, \boldsymbol{v}_f, t_f\right\}} \exp\left(-\int_{t_0}^{t_f} \mathcal{L}(\boldsymbol{x}, \dot{\boldsymbol{x}}, \ddot{\boldsymbol{x}})\mathrm{d}t\right)\mathcal{D}[\boldsymbol{x}(t)] \tag{2}$$

where

$$\mathcal{L}(\boldsymbol{x}, \dot{\boldsymbol{x}}, \ddot{\boldsymbol{x}}) = \frac{1}{2}[\mathbf{M}\ddot{\boldsymbol{x}} + \boldsymbol{g}(\boldsymbol{x}, \dot{\boldsymbol{x}})]^{\mathrm{T}}\mathbf{D}^{-1}[\mathbf{M}\ddot{\boldsymbol{x}} + \boldsymbol{g}(\boldsymbol{x}, \dot{\boldsymbol{x}})] \tag{3}$$

denotes the Lagrangian of the system and $\mathcal{D}[\boldsymbol{x}(t)]$ represents a functional measure.

However, evaluating analytically the WPI of Eq. (2) is a significantly challenging task. Therefore, various approximations have been developed in the literature with the most probable path approach being among the most popular ones (e.g., [24]). Specifically, it is seen that the path maximizing the exponential term in Eq. (2) contributes the most to the determination of the WPI. According to calculus of variations (e.g., [42]), this trajectory $\boldsymbol{x}_c(t)$ with fixed endpoints, also known as most probable path, satisfies the extremality condition

$$\delta \int_{t_0}^{t_f} \mathcal{L}(\boldsymbol{x}, \dot{\boldsymbol{x}}, \ddot{\boldsymbol{x}})\mathrm{d}t = 0 \tag{4}$$

which leads to the Euler–Lagrange (E-L) equations

$$\frac{\partial \mathcal{L}}{\partial x_j} - \frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial \dot{x}_j} + \frac{\partial^2}{\partial t^2}\frac{\partial \mathcal{L}}{\partial \ddot{x}_j} = 0,\ j = 1, \ldots, n \tag{5}$$

in conjunction with the set of boundary conditions

$$\begin{aligned} x_j(t_0) &= x_{j,0}\ \dot{x}_j(t_0) = \dot{x}_{j,0} \\ x_j(t_f) &= x_{j,f}\ \dot{x}_j(t_f) = \dot{x}_{j,f} \end{aligned} \tag{6}$$

Next, solving Eqs. (5) and (6) yields the n-dimensional most probable path $\boldsymbol{x}_c(t)$, and thus, a specific point of the system response transition PDF is determined as [27]

$$p(\boldsymbol{x}_f, \dot{\boldsymbol{x}}_f, t_f | \boldsymbol{x}_0, \dot{\boldsymbol{x}}_0, t_0) \approx C \exp\left(- \int_{t_0}^{t_f} \mathcal{L}(\boldsymbol{x}_c, \dot{\boldsymbol{x}}_c, \ddot{\boldsymbol{x}}_c) \mathrm{d}t \right) \tag{7}$$

where C is a normalization constant. Although it is clear by comparing Eqs. (2) and 7 that only one trajectory (i.e., the most probable path $\boldsymbol{x}_c(t)$) is accounted for in the evaluation of the WPI, it has been shown in various diverse applications (e.g., [28,29,32–35]) that the accuracy degree exhibited by this kind of approximation is relatively high. In passing, note that the accuracy degree of the WPI technique has been further enhanced recently by considering a quadratic approximation to account also for fluctuations around the most probable path; see [31] for details.

2.2. Rayleigh–Ritz solution technique for the most probable path

In general, the boundary value problem (BVP) of Eqs. (5) and (6) is not amenable to an analytical solution treatment. In this regard, the following Rayleigh–Ritz numerical solution scheme can be employed in conjunction with the variational problem

$$\text{minimize } \mathcal{J}(\boldsymbol{x}, \dot{\boldsymbol{x}}, \ddot{\boldsymbol{x}}) = \int_{t_0}^{t_f} \mathcal{L}(\boldsymbol{x}, \dot{\boldsymbol{x}}, \ddot{\boldsymbol{x}}) dt \tag{8}$$

for determining the most probable path $\boldsymbol{x}_c(t)$ (see also [33,43]). In other words, the extremality condition of Eq. (4) can be alternatively expressed as a direct functional minimization problem in the form of Eq. (8). Next, $\boldsymbol{x}(t)$ is approximated by

$$\boldsymbol{x}(t) \approx \hat{\boldsymbol{x}}(t) = \boldsymbol{\psi}(t) + \mathbf{Z}h(t) \tag{9}$$

where $\boldsymbol{\psi}(t)$ is appropriately selected to satisfy the boundary conditions, and the trial functions $\boldsymbol{h}(t) = [h_l(t)]_{L\times 1}$ vanish at the boundaries, i.e, $\boldsymbol{h}(t_0) = \boldsymbol{h}(t_f) = 0; \mathbf{Z} \in \mathbb{R}^{n\times L}$ is a coefficient matrix and L is the number of trial functions considered. Utilizing a vectorized form of $\mathbf{Z}$, Eq. (9) becomes

$$\hat{\boldsymbol{x}}(t) = \boldsymbol{\psi}(t) + \mathbf{H}(t)\boldsymbol{z} \tag{10}$$

with

$$\boldsymbol{z} = \begin{bmatrix} Z_1^{\mathrm{T}} \\ Z_2^{\mathrm{T}} \\ \vdots \\ Z_L^{\mathrm{T}} \end{bmatrix} \in \mathbb{R}^{nL} \text{ and } \mathbf{H}(t) = \begin{bmatrix} \boldsymbol{h}^{\mathrm{T}}(t) & 0 & \ldots & 0 \\ 0 & \boldsymbol{h}^{\mathrm{T}}(t) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \boldsymbol{h}^{\mathrm{T}}(t) \end{bmatrix} \tag{11}$$

where Z_l denotes the l^{th} row of matrix $\mathbf{Z}$ and $\mathbf{H}(t)$ represents an $n \times nL$ time-dependent matrix. Clearly, there is a wide range of choices for functions $\boldsymbol{\psi}$ and $\boldsymbol{h}$. In the ensuing analysis, the Hermite interpolation polynomials

$$\psi_j(t) = \sum_{k=0}^{3} a_{j,k} t^k \tag{12}$$

are adopted, i.e., $\boldsymbol{\psi}(t) = \left[\psi_j(t)\right]_{n\times 1}$, where the $n \times 4$ coefficients $a_{j,k}$ are determined by the $n \times 4$ boundary conditions in Eq. (6). For the trial functions, the shifted Legendre polynomials given by the recursive formula

$$\ell_{q+1}(t) = \frac{2q+1}{q+1}\left(\frac{2t - t_0 - t_f}{t_f - t_0}\right)\ell_q(t) - \frac{q}{q+1}\ell_{q-1}(t),\ q = 1, \ldots, L-1 \tag{13}$$

are employed, which are orthogonal in the interval $[t_0, t_f]$, with $\ell_0(t) = 1$ and $\ell_1(t) = (2t - t_0 - t_f)/(t_f - t_0)$. In this regard, the trial functions take the form

$$h_l(t) = (t - t_0)^2 (t - t_f)^2 \ell_l(t) \tag{14}$$

where the factor $(t-t_0)^2(t-t_f)^2$ multiplies the l^{th}-order Legendre polynomial $\ell_l(t)$ to yield the l^{th} trial function $h_l(t)$. Note that $h_l(t)$ is a polynomial of order $l+4$ and vanishes at the boundaries. Clearly, each component $\hat{x}_j(t)$ of $\hat{\boldsymbol{x}}(t)$ in Eq. (9) is a polynomial of order up to $L+4$ in t.

It can be readily seen that a significant advantage of the Rayleigh–Ritz solution technique relates to the fact that the variational problem of Eq. (8) degenerates to an ordinary minimization problem of a function that depends on a finite number of variables (e.g. [28,33,35,43]). Specifically, the functional $\mathcal{J}$, dependent on the n functions $\boldsymbol{x}(t)$ (and their time derivatives), is cast in the form

$$J(\boldsymbol{z}) := \mathcal{J}\left(\hat{\boldsymbol{x}}, \dot{\hat{\boldsymbol{x}}}, \ddot{\hat{\boldsymbol{x}}}\right) \tag{15}$$

which depends on a finite number of nL coefficients $\boldsymbol{z}$. The corresponding optimization problem takes the form

$$\min_{\boldsymbol{z}} J(\boldsymbol{z}) \tag{16}$$

whereas the extremality condition in Eq. (4) is replaced by the first-order optimality condition

$$\nabla J(\boldsymbol{z}) = \boldsymbol{0} \tag{17}$$

Eq. (17) represents a set of nL nonlinear algebraic equations to be solved numerically. Once the solution $\boldsymbol{z}^*$ of the optimization problem in Eq. (16) is obtained, the most probable path is determined via Eq. (10).

2.3. Linear oscillator: A closed-form exact solution case

It has been shown recently in [30] that, for the special case of linear systems under Gaussian white noise, the WPI most probable path approach summarized in Section 2.1 is amenable to analytical treatment. In fact, the E-L Eqs. (5) and (6) become linear and can be solved analytically for the most probable path, which is substituted into Eq. (7) yielding a closed-form expression for the joint response transition PDF. Most importantly, it has been shown in [30] that the expression of Eq. (7) corresponding to linear systems is exact, and involves no approximations.

Nevertheless, despite the available exact analytical solution for the joint response PDF derived in [30], the Rayleigh–Ritz numerical solution approach discussed in Section 2.2 is also considered in detail in the following for the case of linear systems. This is done intentionally as it provides the motivation and elucidates the rationale for developing a Newton's numerical optimization scheme in Section 3. Specifically, consider a linear system whose dynamics is described by Eq. (1) with

$$\boldsymbol{g}(\boldsymbol{x}, \dot{\boldsymbol{x}}, t) = \boldsymbol{g}_{lin}(\boldsymbol{x}, \dot{\boldsymbol{x}}) := \mathbf{C}\dot{\boldsymbol{x}} + \mathbf{K}\boldsymbol{x} \tag{18}$$

where $\mathbf{C}$ and $\mathbf{K}$ denote the system damping and stiffness matrices, respectively. In this regard, the left hand-side of Eq. (1) can be represented by the linear differential operator $\mathbf{G}[\cdot]$ defined as

$$\mathbf{G} = \mathbf{M}\frac{\partial^2}{\partial t^2} + \mathbf{C}\frac{\partial}{\partial t} + \mathbf{K} \tag{19}$$

Next, for simplicity and without loss of generality, consider $\mathbf{D} = 2\pi S_0\mathbf{I}$, where $\mathbf{I}$ denotes the identity matrix. Substituting the expansion of Eq. (10) into the Lagrangian of Eq. (3), and taking into account Eq. (19), yields

$$\mathcal{L}(\boldsymbol{x}, \dot{\boldsymbol{x}}, \ddot{\boldsymbol{x}}) \approx L(\boldsymbol{z}, t) = \frac{1}{2}\frac{1}{2\pi S_0}[\mathbf{G}[\boldsymbol{\psi}] + \mathbf{G}[\mathbf{H}]\boldsymbol{z}]^{\mathrm{T}}[\mathbf{G}[\boldsymbol{\psi}] + \mathbf{G}[\mathbf{H}]\boldsymbol{z}] \tag{20}$$

Further, expanding Eq. (20) and substituting into Eq. (15), the objective function takes the form

$$J(\boldsymbol{z}) = J_{lin}(\boldsymbol{z}) := \frac{1}{2\pi S_0}\left[\frac{1}{2}\boldsymbol{z}^{\mathrm{T}}\mathbf{Q}\boldsymbol{z} + \mathbf{b}^{\mathrm{T}}\boldsymbol{z}\right] + c \tag{21}$$

where the symmetric matrix $\mathbf{Q} \in \mathbb{R}^{nL\times nL}$ is given by

$$[\mathbf{Q}]_{kl} = \int_{t_0}^{t_f}\sum_{j=1}^{n}[\mathbf{G}[\mathbf{H}]]_{kj}[\mathbf{G}[\mathbf{H}]]_{jl}\mathrm{d}t,\ \ k,l = 1,..,nL \tag{22}$$

the vector $\mathbf{b} \in \mathbb{R}^{nL}$ is determined as

$$[\mathbf{b}]_l = \int_{t_0}^{t_f}\sum_{j=1}^{n}[\mathbf{G}[\psi]]_j[\mathbf{G}[\mathbf{H}]]_{jl}\mathrm{d}t,\ \ l = 1,..,nL \tag{23}$$

and the constant term c (i.e., independent of $\boldsymbol{z}$) is equal to

$$c = \frac{1}{2}\frac{1}{2\pi S_0}\int_{t_0}^{t_f}\mathbf{G}[\psi]^T\mathbf{G}[\psi]\mathrm{d}t \tag{24}$$

Clearly, for the optimization problem of Eq. (16), the multiplicative factor $\frac{1}{2\pi S_0}$ and the constant term c in the definition of the objective function of Eq. (21) do not affect the solution $\boldsymbol{z}^*$. Thus, Eq. (16) becomes, equivalently,

$$\min_{\boldsymbol{z}} \frac{1}{2}\boldsymbol{z}^{\mathrm{T}}\mathbf{Q}\boldsymbol{z} + \mathbf{b}^{\mathrm{T}}\boldsymbol{z} \tag{25}$$

Note that the objective function of Eq. (21) (or, alternatively, Eq. (25)) is not only quadratic, but also convex for positive definite symmetric matrices $\mathbf{Q}$ and its unique global minimizer is given by

$$\boldsymbol{z}^* = -\mathbf{Q}^{-1}\mathbf{b}. \tag{26}$$

Substituting this result into Eq. (10) yields a closed-form expression for the most probable path, i.e.,

$$\hat{\boldsymbol{x}}_c(t) = \boldsymbol{\psi}(t) - \mathbf{H}(t)\mathbf{Q}^{-1}\mathbf{b} \tag{27}$$

Further, it is worth pointing out that the time-dependent matrix $\mathbf{Q}$ is a function only of the initial and final time points $(t_0 \text{ and } t_f)$, and is independent of the boundary values $\boldsymbol{x}(t_0), \dot{\boldsymbol{x}}(t_0), \boldsymbol{x}(t_f)$ and $\dot{\boldsymbol{x}}(t_f)$; the latter are involved only in the evaluation of vector $\mathbf{b}$ through the Hermite polynomials $\boldsymbol{\psi}$ (see Eq. (23)). The interested reader is also directed to Appendix A, where, for tutorial effectiveness, the positive definiteness of matrix $\mathbf{Q}$ is demonstrated for the case of a single-degree-of-freedom (SDOF) linear oscillator.

Finally, it has been shown in this section that a Rayleigh–Ritz numerical solution treatment for determining the most probable path yields an objective function to be minimized, which is both quadratic and convex for linear systems. Obviously, for such cases a Newton's optimization scheme for determining the most probable path converges to the global extremum in only one iteration [36]. Thus, taking into account that nonlinear response behaviors can be construed generally as perturbations (not necessarily small) from the linear regime, it can be argued that a Newton's optimization scheme (such as the one developed in the following section) serves as a natural choice for addressing general cases involving arbitrary nonlinearities.

3. A Newton's numerical optimization scheme for Wiener path integral most probable path determination

3.1. Numerical scheme formulation

In this section, a Newton's iterative algorithm is developed for solving the optimization problem in Eq. (16) corresponding to an arbitrary nonlinear oscillator. In this regard, as highlighted in Section 2.3, the rationale for developing a Newton's scheme relates to the form of the objective function of Eq. (21) referring to linear systems, which is both convex and quadratic; and thus, a Newton's scheme appears to be an ideal choice as it converges in only one iteration to the unique global extremum [36]. This convergence behavior suggests that a Newton's scheme can be a suitable choice also for nonlinear systems governed by Eq. (1) with

$$\mathbf{g}(\boldsymbol{x}, \dot{\boldsymbol{x}}, t) = \mathbf{g}_{lin}(\boldsymbol{x}, \dot{\boldsymbol{x}}) + \varepsilon \mathbf{g}_{nl}(\boldsymbol{x}, \dot{\boldsymbol{x}}) \tag{28}$$

where $\varepsilon > 0$ is a parameter indicating the intensity of the nonlinearity degree and $\mathbf{g}_{nl}(\boldsymbol{x}, \dot{\boldsymbol{x}})$ is an arbitrary nonlinear function. Obviously, in the limiting case, as $\varepsilon \to 0$ the nonlinear function $\mathbf{g}(\boldsymbol{x}, \dot{\boldsymbol{x}}, t)$ becomes linear, i.e., $\mathbf{g} \to \mathbf{g}_{lin}$ (see also Eq. (18)), and the objective function in Eq. (16) approaches the quadratic form of Eq. (21), i.e., $J(\boldsymbol{z}) \to J_{lin}(\boldsymbol{z})$. This asymptotic behavior of $J(\boldsymbol{z})$ suggests that a suitable optimization scheme relates to starting from an initial point $\boldsymbol{z}^{(0)}$ and to successively minimizing a quadratic function J_q^k, which approximates J locally at $\boldsymbol{z}^{(k)}$, i.e.,

$$J_q^k(\boldsymbol{z}) = J(\boldsymbol{z}^{(k)}) + \nabla J(\boldsymbol{z}^{(k)})(\boldsymbol{z} - \boldsymbol{z}^{(k)}) + \frac{1}{2}(\boldsymbol{z} - \boldsymbol{z}^{(k)})^T \nabla^2 J(\boldsymbol{z}^{(k)})(\boldsymbol{z} - \boldsymbol{z}^{(k)}) \tag{29}$$

In Eq. (29), ∇J and $\nabla^2 J$ denote the gradient vector and the Hessian matrix of J, respectively. The next point $\boldsymbol{z}^{(k+1)}$ of the iterative scheme is obtained by minimizing $J_q^k(\boldsymbol{z})$ and setting $\nabla J = \boldsymbol{0}$. This yields

$$\boldsymbol{z}^{(k+1)} = \boldsymbol{z}^{(k)} - \left[\nabla^2 J(\boldsymbol{z}^{(k)})\right]^{-1} \nabla J(\boldsymbol{z}^{(k)}) \tag{30}$$

which is the update formula of the standard Newton's iterative optimization scheme (e.g., [36]).

It is worth noting that for the case of linear oscillators, i.e., $\varepsilon = 0$, and considering Eq. (21), the Hessian matrix becomes $\nabla^2 J = \mathbf{Q}$, which is constant with respect to $\boldsymbol{z}$. Thus, the update formula in Eq. (30) becomes

$$\boldsymbol{z}^{(k+1)} = \boldsymbol{z}^{(k)} - \mathbf{Q}^{-1}\left[\mathbf{Q}\boldsymbol{z}^{(k)} + \mathbf{b}\right]^{-1} = -\mathbf{Q}^{-1}\mathbf{b} \tag{31}$$

which is equal to the closed-form solution derived in Eq. (26). In other words, as anticipated for linear systems, the Newton's optimization scheme converges to the exact solution in only one iteration for any arbitrarily selected starting point $\boldsymbol{z}^{(0)}$. Further, the optimal (for linear systems) point $\boldsymbol{z}^* = -\mathbf{Q}^{-1}\mathbf{b}$ is expected to be a reasonable choice to be used as a starting point in the optimization scheme for the general case of nonlinear systems.

3.2. Convergence analysis aspects

In this section, certain convergence analysis aspects are elucidated pertaining to the herein proposed Newton's scheme in conjunction with the general class of dynamical systems governed by Eqs. (1) and (28).

Specifically, as shown in [36], provided that the Hessian matrix $\nabla^2 J$ is Lipschitz continuous in the neighborhood of the solution $\boldsymbol{z}^*$ and that the initial point $\boldsymbol{z}^{(0)}$ is sufficiently close to $\boldsymbol{z}^*$, the Newton's iterative scheme given by Eq. (30) converges to $\boldsymbol{z}^*$ at a quadratic rate, i.e.,

$$\|\boldsymbol{z}^{(k+1)} - \boldsymbol{z}^*\| \leqslant L\|\left[\nabla^2 J(\boldsymbol{z}^*)\right]^{-1}\|\|\boldsymbol{z}^{(k)} - \boldsymbol{z}^*\|^2 \tag{32}$$

where L is the Lipschitz constant of $\nabla^2 J(\boldsymbol{z})$ for $\boldsymbol{z}$ near $\boldsymbol{z}^*$, i.e., L is a positive real constant defined as

$$\frac{\|\nabla^2 J(\boldsymbol{z}_2) - \nabla^2 J(\boldsymbol{z}_1)\|}{\|\boldsymbol{z}_2 - \boldsymbol{z}_1\|} \leqslant L \tag{33}$$

for all $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ in a neighborhood of $\boldsymbol{z}^*$.

Next, substituting Eq. (28) into Eq. (3), and considering Eq. (21), the gradient vector of $J(\boldsymbol{z})$ becomes

$$\nabla J(\boldsymbol{z}) = \mathbf{Q}\boldsymbol{z} + \mathbf{b} + \varepsilon\nabla g_1(\boldsymbol{z}) + \frac{\varepsilon}{2}\nabla g_2(\boldsymbol{z}) \tag{34}$$

and the Hessian matrix of $J(\boldsymbol{z})$ takes the form

$$\nabla^2 J(\boldsymbol{z}) = \mathbf{Q} + \varepsilon\nabla^2 g_1(\boldsymbol{z}) + \frac{\varepsilon^2}{2}\nabla^2 g_2(\boldsymbol{z}) \tag{35}$$

where

$$g_1(\boldsymbol{z}) = \int_{t_0}^{t_f} \left[\mathbf{M}\ddot{\hat{\boldsymbol{x}}} + \mathbf{C}\dot{\hat{\boldsymbol{x}}} + \mathbf{K}\hat{\boldsymbol{x}}\right]^{\mathrm{T}} \mathbf{g}_{nl}\left(\hat{\boldsymbol{x}}, \dot{\hat{\boldsymbol{x}}}\right) \mathrm{d}t \tag{36}$$

and

$$g_2(\boldsymbol{z}) = \int_{t_0}^{t_f} \mathbf{g}_{nl}\left(\hat{\boldsymbol{x}}, \dot{\hat{\boldsymbol{x}}}\right)^{\mathrm{T}} \mathbf{g}_{nl}\left(\hat{\boldsymbol{x}}, \dot{\hat{\boldsymbol{x}}}\right) \mathrm{d}t \tag{37}$$

Further, substituting Eq. (35) into Eq. (32) leads to

$$\|\boldsymbol{z}^{(k+1)} - \boldsymbol{z}^*\| \leqslant L\|\left[\mathbf{Q} + \varepsilon\nabla^2 g_1(\boldsymbol{z}^*) + \frac{\varepsilon^2}{2}\nabla^2 g_2(\boldsymbol{z}^*)\right]^{-1}\|\|\boldsymbol{z}^{(k)} - \boldsymbol{z}^*\|^2 \tag{38}$$

Moreover, substituting Eq. (35) into Eq. (33) yields

$$\frac{\|\varepsilon\nabla^2 g_1(\boldsymbol{z}_2) + \frac{\varepsilon^2}{2}\nabla^2 g_2(\boldsymbol{z}_2) - \varepsilon\nabla^2 g_1(\boldsymbol{z}_1) - \frac{\varepsilon^2}{2}\nabla^2 g_2(\boldsymbol{z}_1)\|}{\|\boldsymbol{z}_2 - \boldsymbol{z}_1\|} \leqslant L \tag{39}$$

Next, applying the triangle inequality to the left hand-side of Eq. (39) leads to

$$\begin{aligned} &\frac{\|\varepsilon\nabla^2 g_1(\boldsymbol{z}_2) + \frac{\varepsilon^2}{2}\nabla^2 g_2(\boldsymbol{z}_2) - \varepsilon\nabla^2 g_1(\boldsymbol{z}_1) - \frac{\varepsilon^2}{2}\nabla^2 g_2(\boldsymbol{z}_1)\|}{\|\boldsymbol{z}_2 - \boldsymbol{z}_1\|} \quad \leqslant \\ &\varepsilon\frac{\|\nabla^2 g_1(\boldsymbol{z}_2) - \nabla^2 g_1(\boldsymbol{z}_1)\|}{\|\boldsymbol{z}_2 - \boldsymbol{z}_1\|} + \frac{\varepsilon^2}{2}\frac{\|\nabla^2 g_2(\boldsymbol{z}_2) - \nabla^2 g_2(\boldsymbol{z}_1)\|}{\|\boldsymbol{z}_2 - \boldsymbol{z}_1\|} \quad \leqslant \varepsilon L_1 + \frac{\varepsilon^2}{2}L_2 \end{aligned} \tag{40}$$

where L_1 and L_2 represent Lipschitz constants of $\nabla^2 g_1(\boldsymbol{z})$ and $\nabla^2 g_2(\boldsymbol{z})$, respectively, and are independent of ε. Further, considering Eqs. (39) and (40), it is readily seen that the term $\varepsilon L_1 + \frac{\varepsilon^2}{2}L_2$ represents a Lipschitz constant of $\nabla^2 J(\boldsymbol{z})$ for $\boldsymbol{z}$ in the neighborhood of $\boldsymbol{z}^*$, which decreases with decreasing ϵ. Also, as $\varepsilon \to 0$ the term $\|\left[\mathbf{Q} + \varepsilon\nabla^2 g_1(\boldsymbol{z}^*) + \frac{\varepsilon^2}{2}\nabla^2 g_2(\boldsymbol{z}^*)\right]^{-1}\|$ in Eq. (38) approaches the constant positive term $\|\boldsymbol{Q}^{-1}\|$ (see also Appendix A for the positive definiteness of $\boldsymbol{Q}$). Thus, as anticipated, the convergence rate shown in Eq. (38) is increasing for decreasing nonlinearity degree. This is demonstrated further in the numerical examples of Section 5.

Finally, it is important to note that although it has been shown in Section 3 that a Newton's optimization scheme for determining the most probable path appears to be a suitable choice, the convergence rate shown in Eq. (38) can be construed as local (see also [36]). In other words, there is no guarantee about existence and convergence to a global minimum. In fact, proving the potential convexity of $J(\boldsymbol{z})$ (and thus, the existence of a global extremum) is addressed in the following section by resorting to computational algebraic geometry concepts and tools such as Gröbner bases.

4. Convexity and convergence to the global minimum of the Wiener path integral most probable path optimization problem: A computational algebraic geometry approach based on Gröbner bases

In this section, a computational algebraic technique based on Gröbner bases is developed, which is capable of determining the entire set of solutions corresponding to an algebraic system of coupled multivariate polynomial equations. Thus, for a wide range of dynamical systems (e.g., systems with a nonlinearity function in Eq. (28) of polynomial form), not only the entire set of solutions corresponding to the first-order optimality condition $\nabla J(\boldsymbol{z}) = \boldsymbol{0}$ in Eq. (17) can be determined, but also convexity of $J(\boldsymbol{z})$ is implied if the technique yields only one solution. In the latter case, clearly, the Newton's optimization scheme of the previous section converges to the same global minimum determined by the herein proposed computational algebraic technique. In fact, it is noted that, in its standard implementation, the technique determines the values of J corresponding to all real solutions of the system in Eq. (41), i.e., J is evaluated on the entire set of points defined by $\nabla J(\boldsymbol{z}) = \boldsymbol{0}$. Of course, it is possible to extend the technique for determining the actual solutions $\boldsymbol{z}^*$ as well (and not only the values $J(\boldsymbol{z}^*)$) (e.g., [41]). However, this implies significant additional computational cost. In this regard, coupling the Newton's optimization scheme of Section 3 with the herein proposed computational algebraic technique for showing that the $\boldsymbol{z}^*$ provided by Newton's scheme corresponds, indeed, to the global minimum of $J(\boldsymbol{z})$, appears to be an efficient alternative. Indicatively, the interested reader is also directed to papers [44–46] and to books [37–41] for a more detailed presentation of the topic.

Next, considering a polynomial nonlinearity function $\boldsymbol{g}_{nl}$ in Eq. (28) of degree d, the term $g_1(\boldsymbol{z})$ in Eq. (36) becomes a multivariate polynomial of degree $d+1$ in $p := nL$ variables, whereas the term $g_2(\boldsymbol{z})$ in Eq. (37) becomes a multivariate polynomial of degree $2d$ in p variables. Therefore, the objective function $J(\boldsymbol{z})$ takes the form of a multivariate polynomial of degree $2d$ in p variables and the first-order optimality condition of Eq. (17) leads to an algebraic system of p equations of the form

$$\begin{aligned} f_1(z_1,\ldots,z_p) &= 0 \\ &\vdots \\ f_s(z_1,\ldots,z_p) &= 0 \end{aligned} \tag{41}$$

In Eq. (41), each f_i is a polynomial of degree at most $2d-1$ with coefficients in $\mathbb{R}$. In this regard, the convexity of the objective function $J(\boldsymbol{z})$ can be proved by showing that the system of Eq. (41) has a unique real solution. Also, $s = p$ is considered in the ensuing analysis, although this is not a necessary requirement for the technique (e.g., [41]).

4.1. Computational algebraic geometry: Selected basic elements and concepts

In this section, fundamental results related to computer implementations of algebraic geometry concepts are presented for determining the entire set of solutions of Eq. (41) in an algebraic symbolic manner. In the following, $\mathbb{K}[z_1,\ldots,z_p]$ denotes the polynomial ring over the field $\mathbb{K}$, which can be construed as the set of all polynomials in p variables with coefficients in $\mathbb{K}$. Further, an ideal I is a subset of the polynomial ring, usually generated by a finite collection of polynomials as $I = \langle f_1,\ldots,f_s\rangle = \{\sum_{i=1}^{s} h_i f_i \mid h_1,\ldots,h_s \in \mathbb{K}[z_1,\ldots,z_p]\}$. In this regard, an arbitrary polynomial $f \in I$ vanishes on any solution of the polynomial system defined by $f_1,\ldots,f_s$ (see Eq. (41)). The entire set of solutions of Eq. (41) is referred to as the affine variety $\boldsymbol{V}(f_1,\ldots,f_s)$, and thus, for $I = \langle f_1,\ldots,f_s\rangle$ it holds $\boldsymbol{V}(f_1,\ldots,f_s) = \boldsymbol{V}(I)$. The interested reader is also directed to standard books in the fields of algebraic geometry and commutative algebra [37–41] for further details and definitions of the related concepts.

First, defining a precise monomial order denoted by $>$ is a prerequisite for the ensuing analysis. Indicative standard monomial orders include the lexicographical order (*lex*), the graded lexicographical order (*grlex*) and the reverse graded lexicographical order (*grevlex*). Following the selection of an order $>$, $LT(f)$ is defined as the leading term of a polynomial $f \in \mathbb{K}[z_1,\ldots,z_p]$ with respect to the order $>$. Similarly, $LM(f)$ denotes the leading monomial of f, i.e., the monomial of $LT(f)$. Moreover, for an ideal $I \subseteq \mathbb{K}[z_1,\ldots,z_p]$, $LT(I)$ denotes the ideal generated by the leading terms of every element in I, i.e., $LT(I) = \langle LT(f) | f \in I\rangle$. Also, for an ideal $I = \langle f_1,\ldots,f_s\rangle$, the following property is satisfied; that is,

$$\langle LT(f_i),\ldots,LT(f_s)\rangle \subseteq \langle LT(I)\rangle \tag{42}$$

Moreover, the degree of a polynomial, denoted as $\deg(f)$, is defined as the maximum among the sums of powers of all monomials in f and is independent of the monomial order.

Next, a division algorithm is presented, which can be construed as a multivariate generalization of the Euclidean division of univariate polynomials (e.g., [41]).

Definition 4.1 (*Division Algorithm*). Let $>$ be a monomial order and $F = \{f_1,\ldots,f_s\}$ be an ordered s-tuple of polynomials in $\mathbb{K}[z_1,\ldots,z_p]$. Then, every $f \in \mathbb{K}[z_1,\ldots,z_p]$ can be written as

$$f = q_1 f_1 + \ldots + q_s f_s + r$$

where $q_i, r \in \mathbb{K}[z_1,\ldots,z_p]$, and either $r = 0$, or r is a linear combination of monomials, none of which is divisible by any of $LT(f_1),\ldots,LT(f_s)$.

It is noted that not only different monomial orders, but also different arrangements of the f_i's lead, in general, to different q_i's and r. Nevertheless, in the latter case, it is possible to use the Division Algorithm for dividing polynomials $f \in \mathbb{K}[z_1,\ldots,z_p]$ by s-tuples of polynomials $F = \{f_1,\ldots,f_s\}$, yielding unique remainders r, unaffected by the arrangement of the f_i's in F. This leads to the concept of a Gröbner basis.

Definition 4.2 (*Gröbner basis*). Define a monomial order $>$ on $\mathbb{K}[z_1,\ldots,z_p]$ and let $I \subset \mathbb{K}[z_1,\ldots,z_p]$ be an ideal. Next, two equivalent definitions for a Gröbner basis are presented (see also [40]).

1. A finite subset $G = \{g_1,\ldots,g_t\}$ of I is a Gröbner basis if
$$\langle LT(g_1),\ldots,LT(g_t)\rangle = \langle LT(I)\rangle$$
2. A Gröbner basis for I (w.r.t. $>$) is a finite collection of polynomials $G = \{g_1,\ldots,g_t\} \subset I$ with the property that for every $f \in I, LT(f)$ is divisible by $LT(g_i)$ for some i.

The construction of a Gröbner basis G for the ideal $I = \langle f_1,\ldots,f_s\rangle$, provided an arbitrary generating set $\{f_1,\ldots,f_s\}$, can be achieved by Buchberger's algorithm (e.g., [41,47,48]), and in this case the property $I = \langle f_1,\ldots,f_s\rangle = \langle g_1,\ldots,g_t\rangle$ holds. It is noted that a Gröbner basis $G = \{g_1,\ldots,g_t\}$ of $I \in \mathbb{K}[z_1,\ldots,z_p]$ is not unique for a given monomial order $>$. Also, division of any $f \in \mathbb{K}[z_1,\ldots,z_p]$ by G using the Division Algorithm, yields a representation in the form

$$f = q_1 g_1 + \ldots + q_t g_t + \bar{f}^G \tag{43}$$

In the above representation, although the q_i's are non-unique elements of I, (i.e., they depend on the arrangement of the g_i's in G), the remainder $\bar{f}^G$ is unique for a given $>$. Further, $\bar{f}^G$ is not divisible by any $LT(g_i)$, and thus, it is not divisible by any element in $LT(I)$. In this regard, $\bar{f}^G$, represents a uniquely determined normal form for modulo I, and is given as a linear combination of the monomials $x^a \notin \langle LT(I)\rangle$, where $x^a = z_1^{a_1} z_2^{a_2} \ldots z_n^{a_n}$ for some n-tuple $a = (a_1,\ldots,a_n)$ of non-negative integers. Furthermore, the remainders $\bar{f}^G$ and $\bar{g}^G$ generated by the divisions of polynomials $f, g \in I \subset \mathbb{K}[z_1,\ldots,z_p]$ by a Gröbner basis, respectively, exhibit the following properties (e.g., [41]), i.e.,

$$f \in I \iff \bar{f}^G = 0 \tag{44}$$
$$\bar{f}^G = \bar{g}^G \iff f - g \in I \tag{45}$$
$$\bar{f}^G + \bar{g}^G = \overline{f+g}^G \tag{46}$$
$$\overline{\bar{f}^G \cdot \bar{g}^G}^G = \overline{f \cdot g}^G \tag{47}$$

Next, the uniqueness of the remainder $\bar{f}^G$ determined by division of a polynomial $f \in \mathbb{K}[z_1,\ldots,z_p]$ by a Gröbner basis G of $I \subset \mathbb{K}[z_1,\ldots,z_p]$ enables the definition of the coset $[f]$.

Definition 4.3 (*Coset*). Given $f \in \mathbb{K}[z_1,\ldots,z_p]$, the coset $[f]$ is defined as the set

$$[f] = f + I = \{f + h \mid h \in I\}$$

In essence, the coset $[f]$ groups together all polynomials in $\mathbb{K}[z_1,\ldots,z_p]$ that yield the same remainder when divided by G. Hence, this implies a one-to-one correspondence between remainders and cosets ($\bar{f}^G \leftrightarrow [f]$), and thus, $\bar{f}^G$ can be construed as a representative of its coset $[f]$. Next, the quotient ring $\mathbb{K}[z_1,\ldots,z_p]/I$ is defined formally, which represents the set of all cosets of polynomials in $\mathbb{K}[z_1,\ldots,z_p]$ with respect to an ideal $I \in \mathbb{K}[z_1,\ldots,z_p]$.

Definition 4.4 (*Quotient Ring*). The quotient ring $\mathbb{K}[z_1,\ldots,z_p]/I$ is defined as the set

$$\mathbb{K}[z_1,\ldots,z_p]/I = \{[f] \mid f \in \mathbb{K}[z_1,\ldots,z_p]\}$$

According to [41], Eqs. (46) and (47) for the remainder apply also on $\mathbb{K}[z_1,\ldots,z_p]/I$, i.e., $\bar{f}^G + \bar{g}^G \leftrightarrow [f] + [g]$ and $\overline{\bar{f}^G \cdot \bar{g}^G}^G \leftrightarrow [f] \cdot [g]$. In this regard, the quotient ring $\mathbb{K}[z_1,\ldots,z_p]/I$ constitutes also an algebra, which is denoted by A in the ensuing analysis.

It is important to note that the remainders $\bar{f}^G$ are linear combinations of the monomials

$$B = \{x^a \mid x^a \notin LT(I)\} \tag{48}$$

which form a basis of A (also known as the basis of standard monomials in the literature). The role of algebra A in obtaining the entire set of solutions of Eq. (41) is catalytic. However, an important requirement relates to A being finite-dimensional, which is ensured by the following theorem; see also [41] for a detailed proof.

Theorem 1 (*Finiteness Theorem*). *Let $I \subseteq \mathbb{K}[z_1,\ldots,z_p]$ be an ideal. Then, the algebra*

$$A = \mathbb{K}[z_1, \ldots, z_p]/I$$

is finite-dimensional, if and only if, the variety $\boldsymbol{V}(I) \subset \mathbb{K}^n$ *is a finite set.*

Next, considering $\mathbb{K} = \mathbb{C}, f \in \mathbb{C}[z_1, \ldots, z_p]$ and $I \subset \mathbb{C}[z_1, \ldots, z_p]$, the multiplication defined in Eq. (47) can be used to define a linear map m_f from the algebra $A = \mathbb{C}[z_1, \ldots, z_p]/I$ to itself. Specifically, for $[g] \in A, m_f : A \to A$ is defined as

$$m_f([g]) = [f] \cdot [g] = [f \cdot g] \in A \tag{49}$$

Relying on the vector space structure of the algebra A, it can be shown that m_f is, indeed, a linear map [41]. Further, assuming that the system of polynomial equations in Eq. (41) has a finite number of solutions, Theorem 1 implies that A is a finite-dimensional algebra. This, enables the representation of the linear map m_f by a matrix M_f associated with a basis of A. This is precisely the basis B of standard monomials defined in Eq. (48). More importantly, Theorem 2 shows that the multiplication matrix M_f can be utilized to evaluate f on $\boldsymbol{V}(I)$; see [41] for a detailed proof.

Theorem 2. *Let* $I \subset \mathbb{C}[z_1, \ldots, z_p]$ *be a zero-dimensional ideal and* $A = \mathbb{C}[z_1, \ldots, z_p]/I$. *Further, let* $f \in \mathbb{C}[z_1, \ldots, z_p]$ *with* M_f *being the multiplication matrix corresponding to the linear map* $m_f : A \to A$ *defined in Eq. (49). Then, the eigenvalues of* M_f *are equal to the values of* f *on* $\boldsymbol{V}(I)$.

In other words, according to Theorem 2, if the system in Eq. (41) has a finite number of μ solutions (i.e., $\boldsymbol{V}(I)$ is a finite set of size μ), substituting these solutions into any polynomial $f \in \mathbb{C}[z_1, \ldots, z_p]$ yields μ values that are equal to the eigenvalues of M_f. In passing, it is worth noting that the (Strong) Nullstellensatz, which is a consequence of Hilbert's original result and is used in the proof of Theorem 2, is of paramount importance to efforts attempting to associate ideals I with the corresponding varieties $\boldsymbol{V}(I)$. In this regard, it can be argued that it provides the tools for establishing a "dictionary" between geometry and algebra [41].

It is important to note that matrix M_f can only be defined if the standard basis B of Eq. (48) is finite. This is true if the variety $\boldsymbol{V}(f_1, \ldots, f_s)$ is a finite set, i.e., the system of polynomial equations in Eq. (41) has a finite number of μ solutions in $\mathbb{C}$. Thus, it becomes clear that the total number μ of (complex) solutions of the system in Eq. (41) is equal to the number of monomials in the standard basis B, i.e., $\mu = \text{length}(B)$. If f is a dense polynomial of even degree $2d$, then it follows from Bézout's Theorem that $\mu = (2d-1)^n$ [46].

4.2. Algorithmic aspects and mechanization of the technique

The steps for determining the multiplication matrix M_f of an arbitrary $f \in \mathbb{K}[z_1, \ldots, z_p]$ corresponding to the system of polynomials in Eq. (41) are presented in Algorithm 1, which is based on the following three main subroutines:

- `Groebner`$(f_1, \ldots, f_s)$: This subroutine computes a Gröbner basis $G = \{g_1, \ldots, g_t\}$ of the ideal generated by $f_1, \ldots, f_s$ based on Buchberger's algorithm (e.g., [47,48]). The implementation can be found in most symbolic mathematical computation languages (see for instance `gbasis(.)` built-in function in Matlab).
- `StandardBasis`(G): This subroutine computes the basis of standard monomials B defined in Eq. (48) corresponding to the Gröbner basis $G = \{g_1, \ldots, g_t\}$. Indicatively, B is constructed by selecting all monomials that are not divisible by $LT(g_i)$ for any $i = 1, \ldots, t$.
- `NormalForm`(h, G): This subroutine computes the unique remainder of the division of an arbitrary polynomial h by the Gröbner basis G via the Division Algorithm in Definition 4.1. The column vector of coefficients of this remainder with respect to basis B, is denoted as $[\texttt{NormalForm}(h, G)]_B$. The interested reader is directed to [45] for an indicative implementation in a symbolic language system.

Algorithm 1. `MultMatrix`$(f, f_1, \ldots, f_s)$ - Computation of multiplication matrix M_f

Input: $f, f_1, \ldots, f_s \in \mathbb{R}[z_1, \ldots, z_p]$
Output: M_f
1: $G = \texttt{Groebner}(f_1, \ldots, f_s)$
2: $B = \texttt{StandardBasis}(G)$
3: $\mu =$ length(B)
4: Initialize M_f as an empty $\mu \times \mu$ matrix
5: **for** $i = 1$ to μ **do**
6: $\quad M_f(:, i) = [\texttt{NormalForm}(B(i) \cdot f, G)]_B$ ▷Computation of the i-th column of M_f
7: **end for**
8: **return** M_f

Finally, consider the optimization problem of Eq. (16) and set

$$J^* = \min_{\boldsymbol{z}\in\mathbb{R}^n} J(\boldsymbol{z}) \tag{50}$$

Then, J^* is equal to the smallest real eigenvalue of matrix M_J representing the linear map $m_{J_\epsilon} : A \to A$ (defined in Eq. (49)) with respect to the monomial basis B (defined in Eq. (48)), where $A = \mathbb{C}[z_1, \ldots, z_p]/I$ with $I = \left\langle \frac{\partial J}{\partial z_1}, \ldots, \frac{\partial J}{\partial z_n} \right\rangle$. In other words, the globally minimum value of $J(\boldsymbol{z})$, as well as all its values on its critical points (i.e., points where $\nabla J = \boldsymbol{0}$), can be found by determining matrix M_J. In passing, it is noted that the graded lexicographical order

$$x_1^d > \ldots > x_n^d > x_1^{d-1}x_2 > \ldots > x_1 > \ldots > x_n > 1 \tag{51}$$

is considered in the following numerical examples.

5. Numerical examples

In this section, various numerical examples pertaining to oscillators with diverse nonlinear behaviors are considered for demonstrating the reliability of the WPI technique to evaluate the joint response PDF, in conjunction with the proposed Newton's scheme for determining the most probable path. Further, the herein developed Gröbner basis approach is also employed for demonstrating the existence of a unique solution (and thus, the convexity of the objective function) of the most probable path optimization problem.

Furthermore, according to the standard numerical implementation of the WPI technique, the evaluation of the joint response PDF of Eq. (7) at a given time instant involves the discretization of the PDF effective domain into N^{2n} points (where N is the number of points along each dimension), and thus, requires the solution of N^{2n} BVPs of the form of Eqs. (5) and (6) (or, equivalently, Eq. (17)). Clearly, this leads to an exponential growth of the computational cost as a function of the dimensionality n of the system. This limitation of the brute-force implementation of the WPI technique has been addressed in [49,50] by employing multi-dimensional function approximation techniques in conjunction with compressive sampling concepts and tools for reducing drastically the total number of required BVPs to be solved numerically. To provide an indicative comparison in terms of computational cost between MCS and a brute-force implementation of the WPI technique, and considering $n = 1$ and $N = 31$ in the following examples, the joint PDF obtained based on the solution of $31^2 = 961$ BVPs requires approximately 10 s, whereas a MCS-based PDF estimate using $10,000$ realizations is obtained in approximately 1 h on the same computer.

5.1. Linear oscillator

Consider a SDOF linear oscillator whose governing equation is a scalar version of Eq. (1), i.e.,

$$m\ddot{x} + c\dot{x} + kx = w(t), \tag{52}$$

where $m = 5, c = 0.2, k = 1$, and $E(w(t)w(t+\tau)) = 2\pi S_0\delta(\tau)$ with $S_0 = 0.5$.

As pointed out in Section 3.1, for linear systems the Newton's optimization scheme converges to the exact solution $\boldsymbol{z}^* = -\mathbf{Q}^{-1}\mathbf{b}$ in only one iteration for any arbitrarily selected starting point $\boldsymbol{z}^{(0)}$. In this regard, for an indicative final time instant $t_f = 1s$ and for boundary conditions $(x(0), \dot{x}(0), x(t_f), \dot{x}(t_f)) = (0, 0, -0.5, -1.0)$, the objective function of Eq. (21) is shown in Fig. 1 by utilizing $L = 2$ trial functions. Further, as also stated in Section 3.1 and proved in Appendix A, it is readily seen that the objective function $J(\boldsymbol{z})$ of Eq. (21) is convex, and thus, the Newton's scheme converges to the exact optimal solution $\boldsymbol{z}^* = -\mathbf{Q}^{-1}\mathbf{b} = (0.0173, 0.00014)$ in a single iteration starting from the arbitrarily chosen point $\boldsymbol{z}^{(0)} = (50, 50)$. Further, employing the Gröbner basis approach developed in Section 4 (see also Appendix B) yields a single solution and the corresponding objective function value becomes $J(\boldsymbol{z}^*) = 4.4204$, which coincides practically with the estimate based on Newton's scheme.

Finally, utilizing a brute-force discretization of the PDF effective domain (e.g. [33]), the joint response PDF at a specific time instant is determined via the solution of $N^2 = 961$ boundary value problems. In Fig. 2, the corresponding marginal response displacement and velocity PDFs at various time instants are plotted. It is shown that the WPI-based estimates utilizing Newton's scheme coincide with the estimates based on the Gröbner basis approach. Comparisons with MCS data (10,000 realizations) are included as well demonstrating the high accuracy degree exhibited by the WPI technique.

5.2. Duffing nonlinear oscillator

In this example, consider a SDOF Duffing nonlinear oscillator whose governing equation is a scalar version of Eq. (1), i.e.,

$$m\ddot{x} + c\dot{x} + kx + \varepsilon g_{nl}(x, \dot{x}) = w(t) \tag{53}$$

where

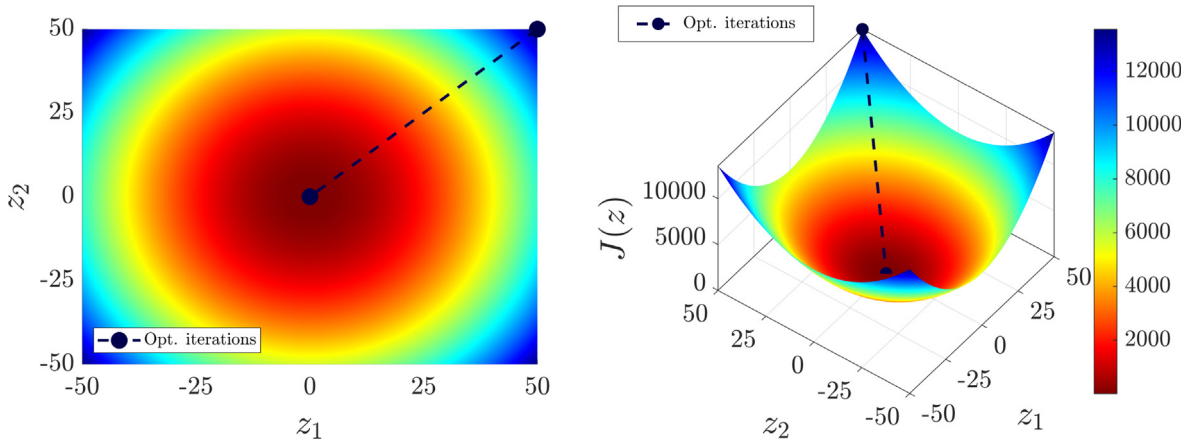


Fig. 1. Most probable path optimization problem objective function using $L = 2$ trial functions and corresponding to a linear oscillator under white noise ($t_f = 1s, x(t_f) = -0.5, \dot{x}(t_f) = -1.0$). The Newton's optimization scheme iterations are also included.

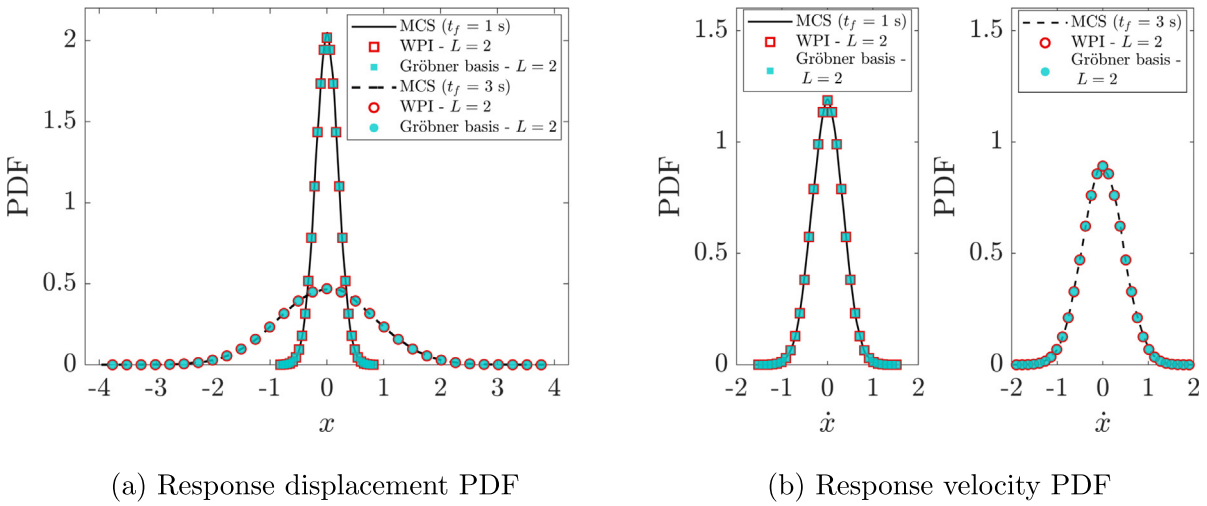


Fig. 2. Response displacement and velocity PDFs at various time instants corresponding to a linear oscillator under white noise. Comparisons between WPI-based estimates utilizing the Newton's scheme and the Gröbner basis approaches for the most probable path determination. MCS-based estimates are also included (10,000 realizations).

$$g_{nl}(x, \dot{x}) = kx^3 \tag{54}$$

In Eqs. (53) and (28), the same parameters values are used for m, c, k and S_0 as in Example 5.1.

Next, for an indicative final time instant $t_f = 1s$ and for boundary conditions $(x(0), \dot{x}(0), x(t_f), \dot{x}(t_f)) = (0, 0, -0.5, -1.0)$, the objective functions $J(\boldsymbol{z})$ of the most probable path optimization problem for $\varepsilon = 1, 10$, and 20 are shown in Figs. 3–5, respectively. The iteration points of the Newton's scheme are also included in the figures corresponding both to an arbitrarily selected starting point (i.e., $\boldsymbol{z}^{(0)} = (50, -50)$), and to a starting point equal to the exact optimum $\boldsymbol{z}^{(0)} = \boldsymbol{z}_{lin} = -\mathbf{Q}^{-1}\mathbf{b}$ corresponding to the associated linear system (i.e., $\varepsilon = 0$). Clearly, as pointed out in Section 3.1, $\boldsymbol{z}^{(0)} = \boldsymbol{z}_{lin}$ is shown to be a reasonable choice as a starting point in the optimization scheme, since for all cases the number of iterations is significantly smaller than the respective one based on an arbitrarily selected starting point. Also, as dictated by Eq. (38), it can be readily seen that the convergence rate increases for smaller values of the nonlinearity parameter ε. Numerical results related to the iterations of the Newton's scheme are summarized in Table 1, which includes also results based on the Gröbner basis approach. It is seen that, for all nonlinearity parameter values, the Gröbner approach yields a single solution and the corresponding objective function value coincides practically with the estimate based on the Newton's scheme (see also Appendix B). This proves the convexity of the objective function and that the Newton's scheme converges, indeed, to the global minimum.

In Fig. 6, the WPI-based marginal response displacement and velocity PDFs are plotted for various nonlinearity magnitude values at two indicative time instants. It is shown that the WPI-based estimates utilizing the Newton's scheme coincide with

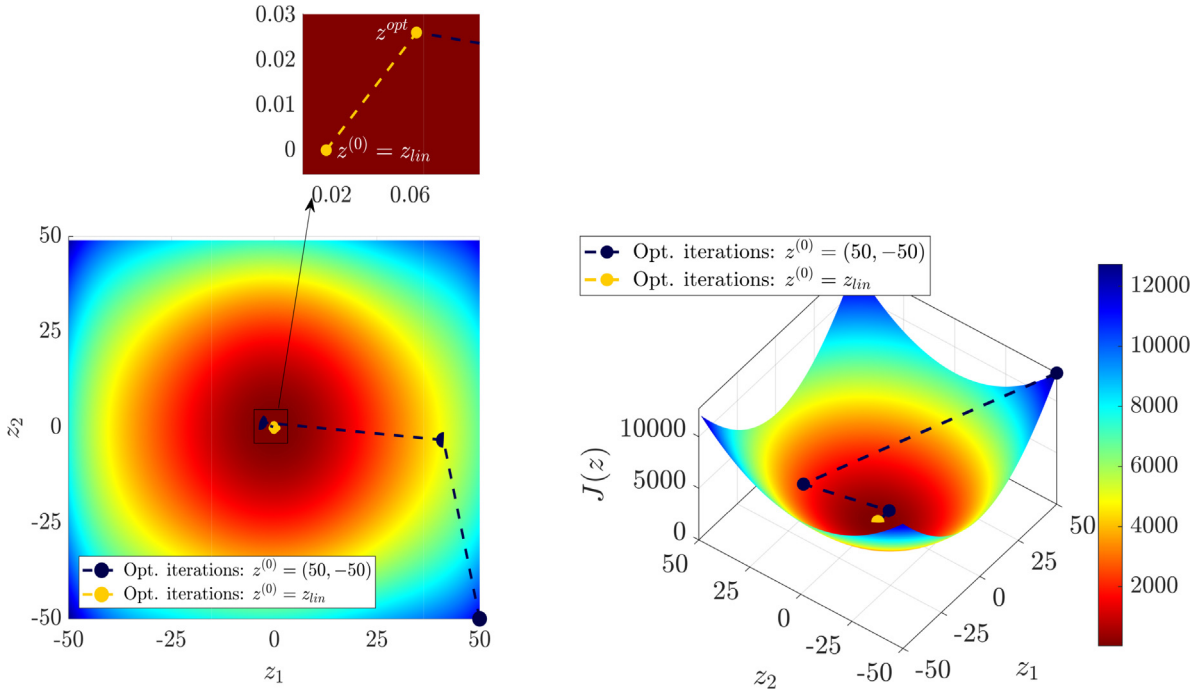


Fig. 3. Most probable path optimization problem objective function using $L = 2$ trial functions and corresponding to a Duffing oscillator with $\varepsilon = 1.0$ under white noise ($t_f = 1s, x(t_f) = -0.5, \dot{x}(t_f) = -1.0$). The Newton's optimization scheme iterations are also included.

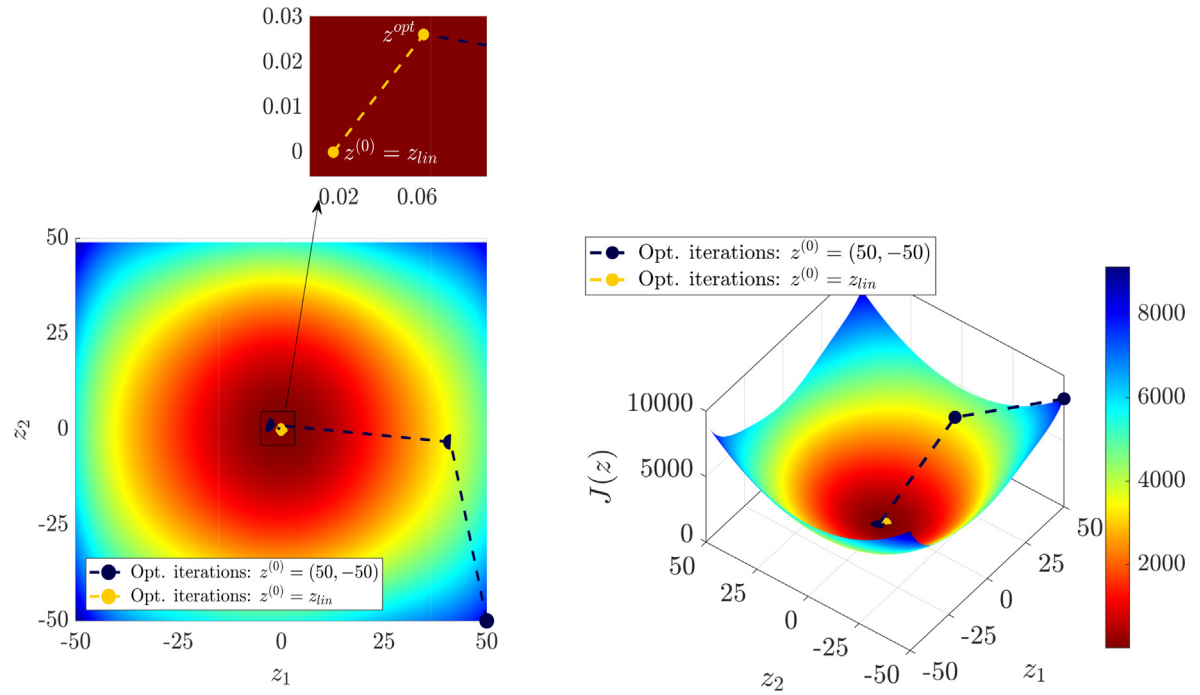


Fig. 4. Most probable path optimization problem objective function using $L = 2$ trial functions and corresponding to a Duffing oscillator with $\varepsilon = 10$ under white noise ($t_f = 1s, x(t_f) = -0.5, \dot{x}(t_f) = -1.0$). The Newton's optimization scheme iterations are also included.

the estimates based on the Gröbner basis approach. Comparisons with MCS data (10,000 realizations) are included as well demonstrating the high accuracy degree exhibited by the WPI technique.

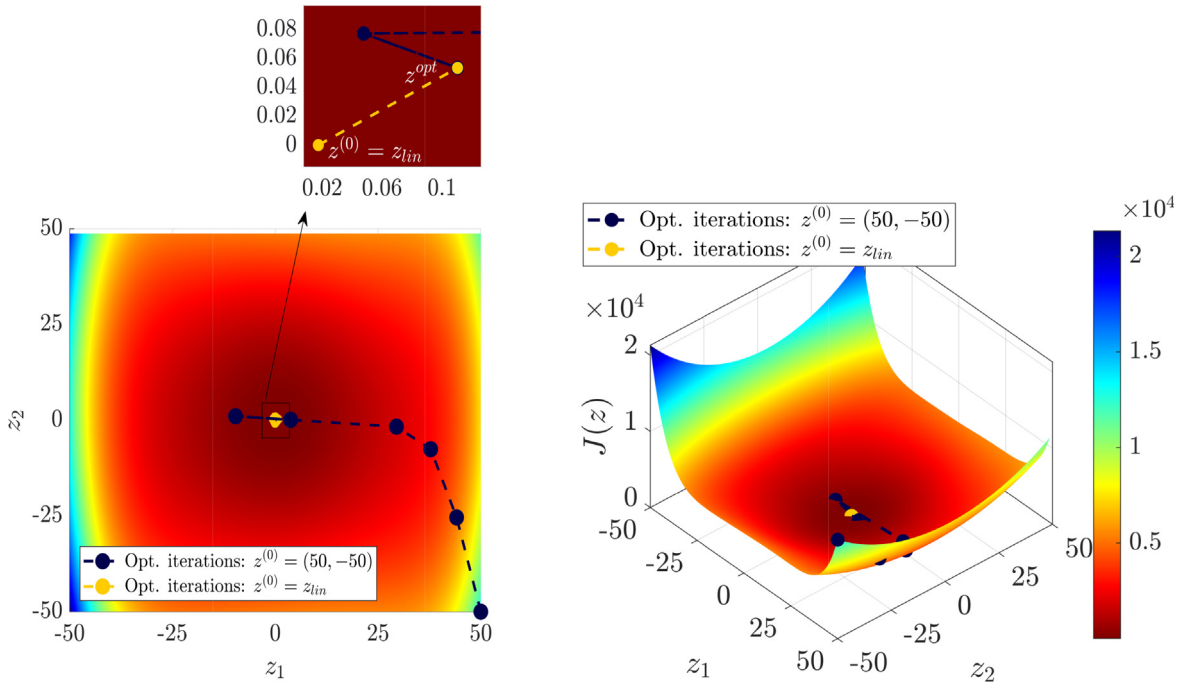


Fig. 5. Most probable path optimization problem objective function using $L = 2$ trial functions and corresponding to a Duffing oscillator with $\varepsilon = 20$ under white noise ($t_f = 1s, x(t_f) = -0.5, \dot{x}(t_f) = -1.0$). The Newton's optimization scheme iterations are also included.

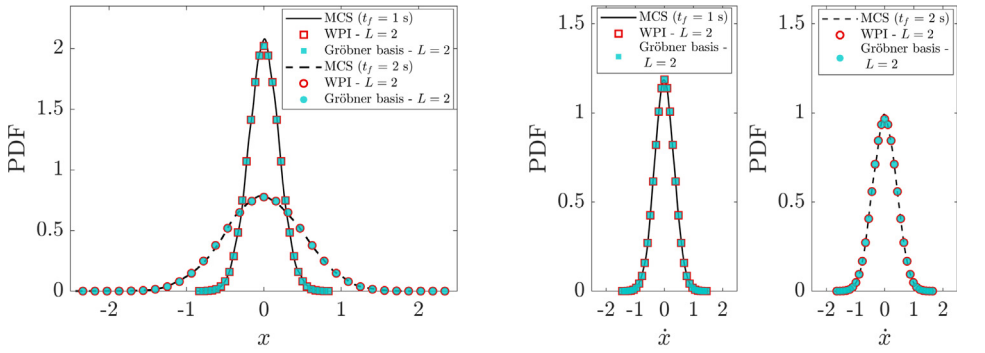
Table 1
Convergence rate and objective function values for a Duffing oscillator.

$\varepsilon = 1.0$		$\varepsilon = 10$		$\varepsilon = 20$		$\varepsilon = 100$	
$\|z^{(k)} - z^*\|$	$J(\boldsymbol{z}^{(k)})$	$\|z^{(k)} - z^*\|$	$J(\boldsymbol{z}^{(k)})$	$\|z^{(k)} - z^*\|$	$J(\boldsymbol{z}^{(k)})$	$\|z^{(k)} - z^*\|$	$J(\boldsymbol{z}^{(k)})$
Newton's numerical optimization scheme - Arbitrarily selected starting point $\boldsymbol{z}^{(0)}$							
70.697	1.26E4	70.685	8.23E3	70.671	1.30E4	70.138	8.81E4
37.666	4.14E3	44.568	3.97E3	50.835	5.26E3	58.718	2.34E4
7.735	1.92E2	18.447	9.96E2	38.458	2.81E3	49.450	6.92E3
4.63E-2	4.458	12.425	4.59E2	29.451	1.83E3	40.993	2.52E3
5.37E-7	4.452	3.571	45.024	9.694	2.76E2	31.862	1.10E3
1.08E-16	4.452	2.93E-2	4.745	3.467	43.157	23.047	6.10E2
		1.56E-6	4.743	5.588E-2	5.096	17.922	4.28E2
		7.03E-15	4.743	1.08E-5	5.086	14.318	1.04E2
				6.44E-13	5.086	5.787	28.117
						2.459	8.534
						0.118	8.491
						1.32E-4	8.491
						2.97E-10	8.491
						1.41E-15	8.491
Newton's numerical optimization scheme - Starting point $\boldsymbol{z}^{(0)} = -\mathbf{Q}^{-1}\mathbf{b}$ corresponding to a linear oscillator with $\varepsilon = 0$							
5.24E-3	4.452	5.29E-2	4.751	1.07E-1	5.120	5.74E-1	9.446
8.49E-9	4.452	8.49E-6	4.743	6.77E-5	5.086	7.68E-3	8.491
		2.15E-13	4.743	2.617E-11	5.086	1.11E-6	8.491
						2.28E-14	8.491
Computational algebraic geometry approach based on Gröbner bases							
	4.452		4.743		5.086		8.490

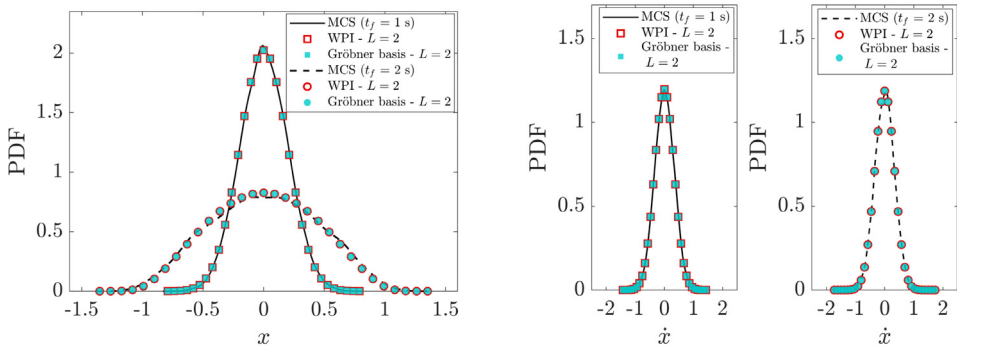
5.3. Nonlinear oscillator with an asymmetric response PDF

In this example, consider a SDOF nonlinear oscillator with an asymmetric response PDF, whose governing equation is given by Eqs. (55) and (56), i.e.,

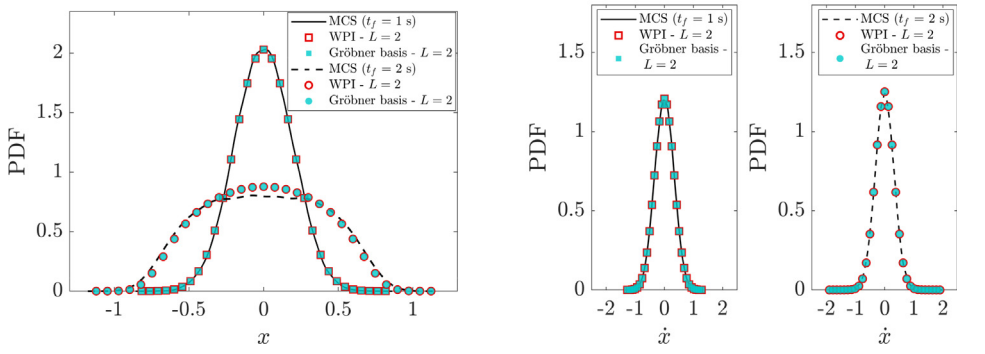
$$m\ddot{x} + c\dot{x} + kx + \varepsilon g_{nl}(x, \dot{x}) = w(t) \tag{55}$$



(a) $\varepsilon = 1$: Response displacement PDF (b) $\varepsilon = 1$: Response velocity PDF



(c) $\varepsilon = 10$: Response displacement PDF (d) $\varepsilon = 10$: Response velocity PDF



(e) $\varepsilon = 20$: Response displacement PDF (f) $\varepsilon = 20$: Response velocity PDF

Fig. 6. Response displacement and velocity PDFs at various time instants corresponding to a Duffing oscillator under white noise. Comparisons between WPI-based estimates utilizing the Newton's scheme and the Gröbner basis approaches for the most probable path determination. MCS-based estimates are also included (10,000 realizations).

where

$$g_{nl}(x, \dot{x}) = ax^2 + x^3 \tag{56}$$

In Eq. (55) and (56), the parameters values used are, $m = 1, c = 0.2, k = 1, E(w(t)w(t+\tau)) = 2\pi S_0 \delta(\tau)$ with $S_0 = 0.05$, and a is constant.

Further, for an indicative final time instant $t_f = 1s$ and for boundary conditions $(x(0), \dot{x}(0), x(t_f), \dot{x}(t_f)) = (0, 0, -0.3, -0.8)$, the objective functions $J(\boldsymbol{z})$ of the most probable path optimization problem for $\varepsilon = 1$ and $a = 1.5, \varepsilon = 10$ and $a = \frac{3\sqrt{10}}{20}$, and $\varepsilon = 50$ and $a = \frac{3\sqrt{2}}{20}$ are shown in Figs. 7–9, respectively. The iteration points of the Newton's scheme are also included in the

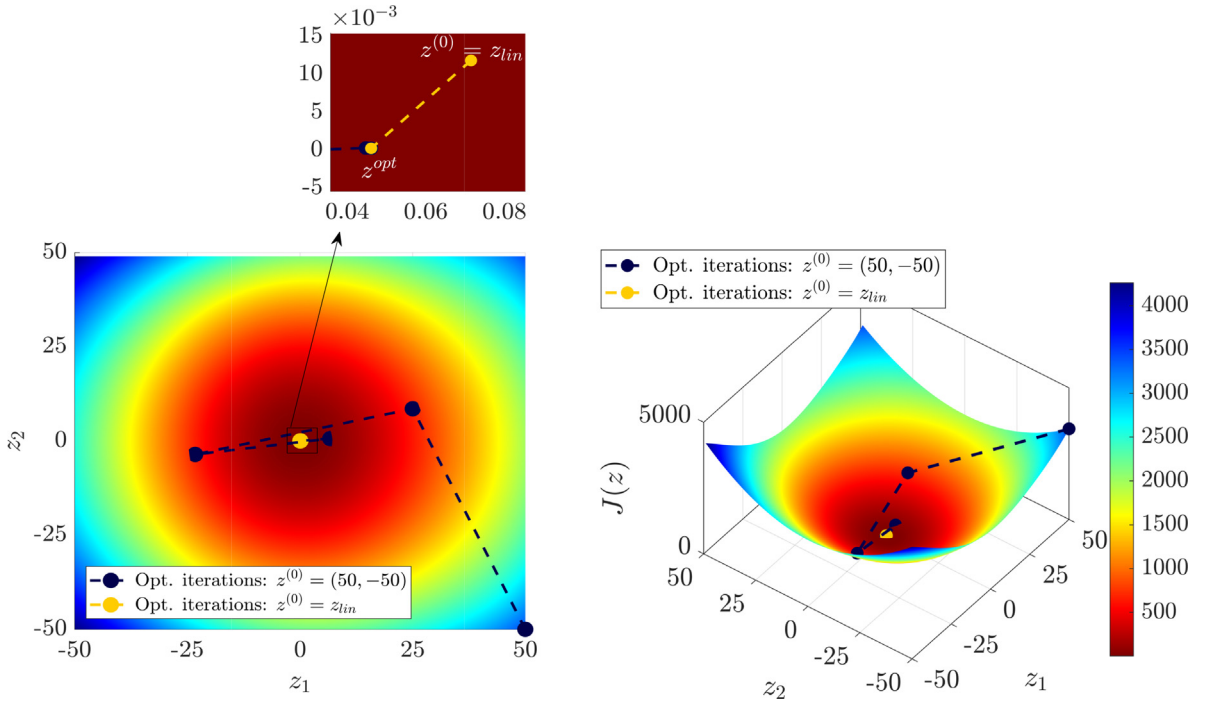


Fig. 7. Most probable path optimization problem objective function using $L = 2$ trial functions and corresponding to a nonlinear oscillator with an asymmetric response PDF with $\varepsilon = 1$ and $a = 1.5$ under white noise ($t_f = 1s, x(t_f) = -0.3, \dot{x}(t_f) = -0.8$). The Newton's optimization scheme iterations are also included.

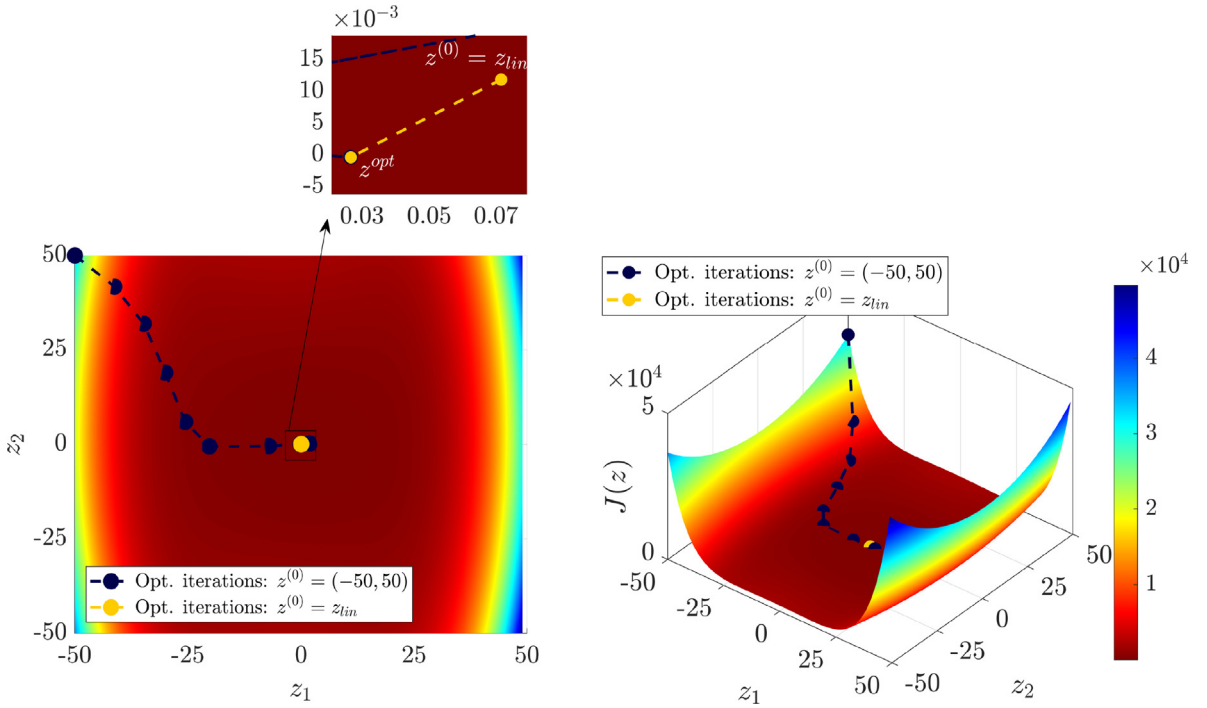


Fig. 8. Most probable path optimization problem objective function using $L = 2$ trial functions and corresponding to a nonlinear oscillator with an asymmetric response PDF with $\varepsilon = 10$ and $a = \frac{3\sqrt{10}}{20}$ under white noise ($t_f = 1s, x(t_f) = -0.3, \dot{x}(t_f) = -0.8$). The Newton's optimization scheme iterations are also included.

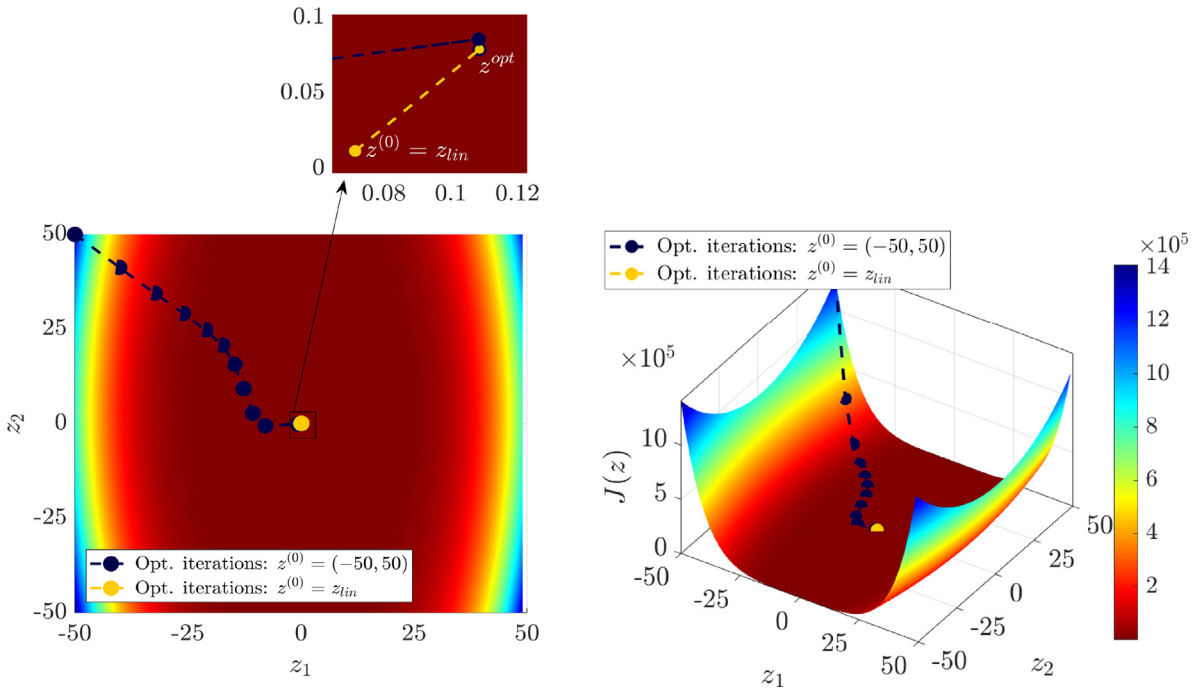


Fig. 9. Most probable path optimization problem objective function using $L = 2$ trial functions and corresponding to a nonlinear oscillator with an asymmetric response PDF with $\varepsilon = 50$ and $a = \frac{3\sqrt{2}}{20}$ under white noise ($t_f = 1s, x(t_f) = -0.3, \dot{x}(t_f) = -0.8$). The Newton's optimization scheme iterations are also included.

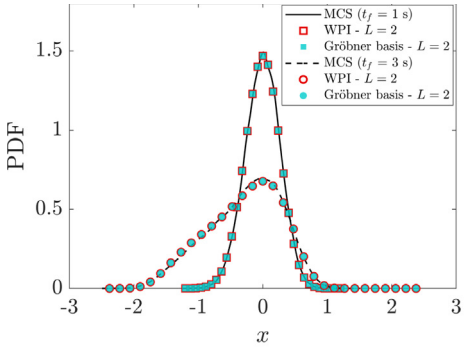
Table 2
Convergence rate and objective function values for a nonlinear oscillator with an asymmetric response PDF.

$\varepsilon = 1, a = 1.50$		$\varepsilon = 10, a = \frac{3\sqrt{10}}{20}$		$\varepsilon = 50, a = \frac{3\sqrt{2}}{20}$		$\varepsilon = 100, a = 0.150$	
$\|z^{(k)} - z^*\|$	$J(\boldsymbol{z}^{(k)})$	$\|z^{(k)} - z^*\|$	$J(\boldsymbol{z}^{(k)})$	$\|z^{(k)} - z^*\|$	$J(\boldsymbol{z}^{(k)})$	$\|z^{(k)} - z^*\|$	$J(\boldsymbol{z}^{(k)})$
Newton's numerical optimization scheme - Arbitrarily selected starting point $\boldsymbol{z}^{(0)}$							
70.678	3.49E+03	70.730	3.30E4	70.732	1.28E6	70.374	5.40E6
25.173	6.21E2	58.642	9.16E3	57.614	3.33E5	57.053	1.41E6
17.770	3.88E2	47.166	2.96E3	47.261	8.62E4	46.490	3.65E5
1.786	5.536	35.498	1.19E3	39.109	2.23E4	38.153	9.45E4
2.31E-2	1.683	26.321	5.92E2	32.603	5.90E3	31.606	2.43E4
3.21E-6	1.683	20.316	3.64E2	27.040	1.68E3	26.449	6.27E3
6.18E-14	1.683	7.104	63.160	21.494	5.72E2	22.213	1.67E3
		1.943	6.109	15.672	2.39E2	18.191	4.96E2
		1.15E-1	1.615	11.019	1.16E2	13.697	1.78E2
		2.08E-4	1.598	7.990	63.940	9.385	71.774
		7.29E-10	1.598	0.782	2.554	6.971	35.490
				6.73E-3	1.794	5.514	23.641
				2.35E-7	1.794	1.570	5.110
				1.07E-15	1.794	0.520	2.616
						6.55E-3	2.271
						1.93E-6	2.271
						1.99E-13	2.271
Newton's numerical optimization scheme - Starting point $\boldsymbol{z}^{(0)} = -\mathbf{Q}^{-1}\mathbf{b}$ corresponding to a linear oscillator with $\varepsilon = 0$							
0.0278	1.684	0.046	1.601	0.077	1.800	0.278	2.357
3.955E-6	1.683	3.205E-5	1.598	2.305E-5	1.794	0.0010	2.2715
9.116E-14	1.683	1.711E-11	1.598	8.485E-12	1.794	1.831E-8	2.271
Computational algebraic geometry approach based on Gröbner bases							
	1.683		1.598		1.793		2.271

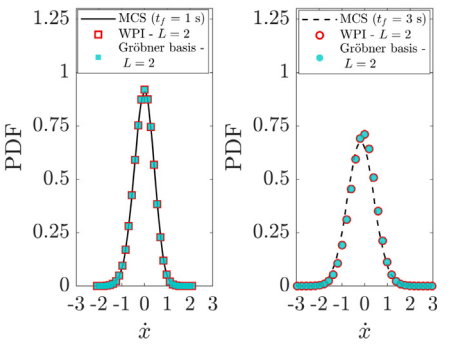
figures corresponding both to an arbitrarily selected starting point (i.e., $\boldsymbol{z}^{(0)} = (50, -50)$ or $\boldsymbol{z}^{(0)} = (-50, 50)$), and to a starting point equal to the exact optimum $\boldsymbol{z}^{(0)} = \boldsymbol{z}_{lin} = -\mathbf{Q}^{-1}\mathbf{b}$ corresponding to the associated linear system (i.e., $\varepsilon = 0$). In a similar

manner as in Example 5.2, $\boldsymbol{z}^{(0)} = \boldsymbol{z}_{lin}$ proves to be an excellent choice as a starting point in the optimization scheme, whereas it is seen that the convergence rate increases for smaller values of the nonlinearity parameter ε. Numerical results related to the iterations of the Newton's scheme are summarized in Table 2, which includes also results based on the Gröbner basis approach. The convexity of the objective function and the convergence of Newton's scheme to the global minimum is demonstrated by noticing that for all nonlinearity parameter values, the Gröbner approach yields a single solution and the corresponding objective function value coincides practically with the estimate based on Newton's scheme.

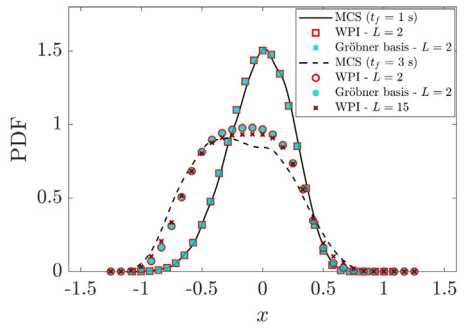
In Fig. 10, the WPI-based marginal response displacement and velocity PDFs are plotted for various nonlinearity magnitude values at two indicative time instants. It is shown that the WPI-based estimates utilizing Newton's scheme coincide



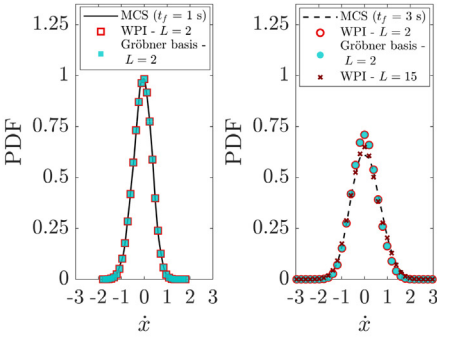
(a) $\varepsilon = 1$: Response displacement PDF



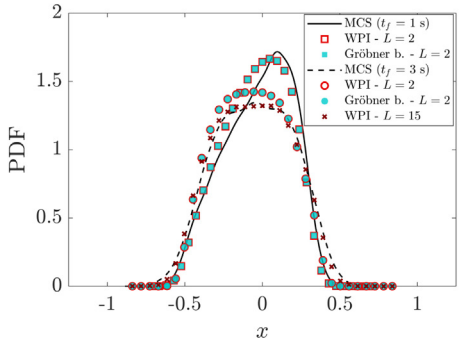
(b) $\varepsilon = 1$: Response velocity PDF



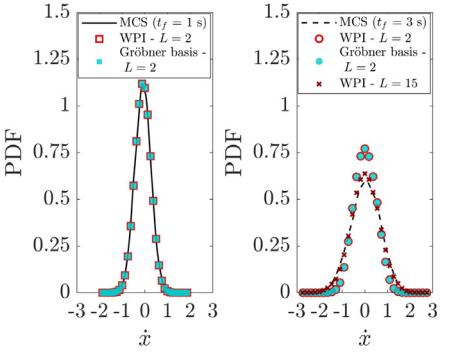
(c) $\varepsilon = 10$: Response displacement PDF



(d) $\varepsilon = 10$: Response velocity PDF



(e) $\varepsilon = 50$: Response displacement PDF



(f) $\varepsilon = 50$: Response velocity PDF

Fig. 10. Response displacement and velocity PDFs at various time instants corresponding to a nonlinear oscillator with an asymmetric response PDF. Comparisons between WPI-based estimates utilizing the Newton's scheme and the Gröbner basis approaches for the most probable path determination. MCS-based estimates are also included (20,000 realizations).

with the estimates based on the Gröbner basis approach. Comparisons with MCS data (20,000 realizations) are included as well demonstrating the high accuracy degree exhibited by the WPI technique.

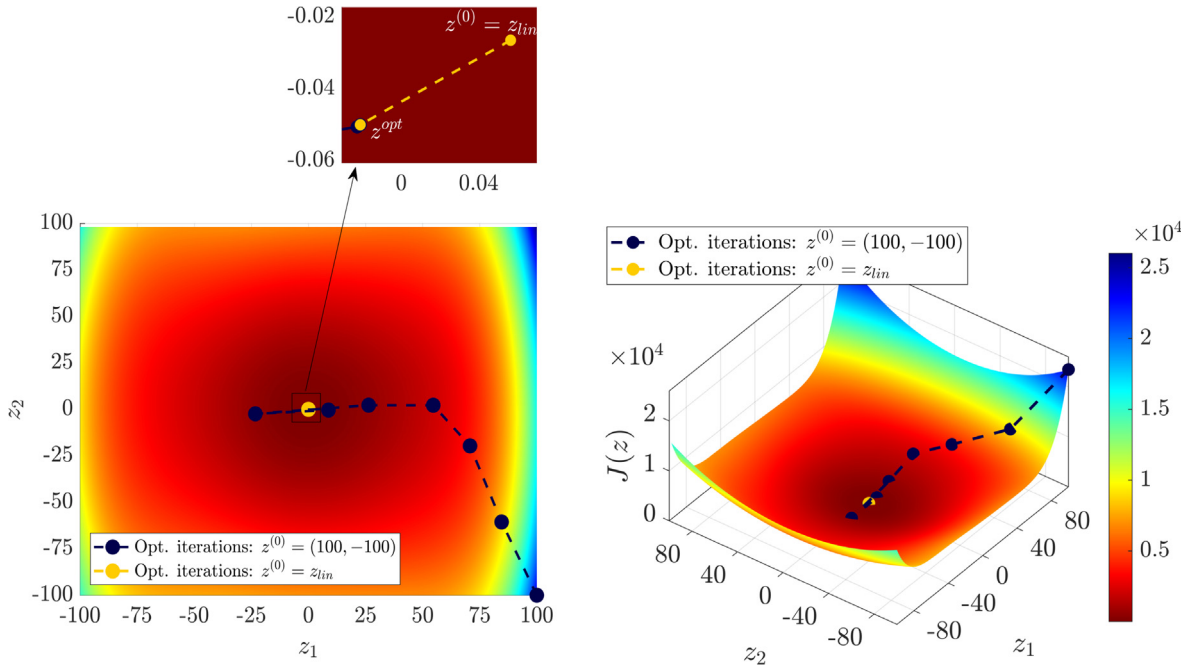


Fig. 11. Most probable path optimization problem objective function using $L = 2$ trial functions and corresponding to a nonlinear oscillator with a bimodal response PDF with $a = 1.3$ and $\varepsilon = 1$ under white noise ($t_f = 1s, x(t_f) = 0.8, \dot{x}(t_f) = 0.9$). The Newton's optimization scheme iterations are also included.

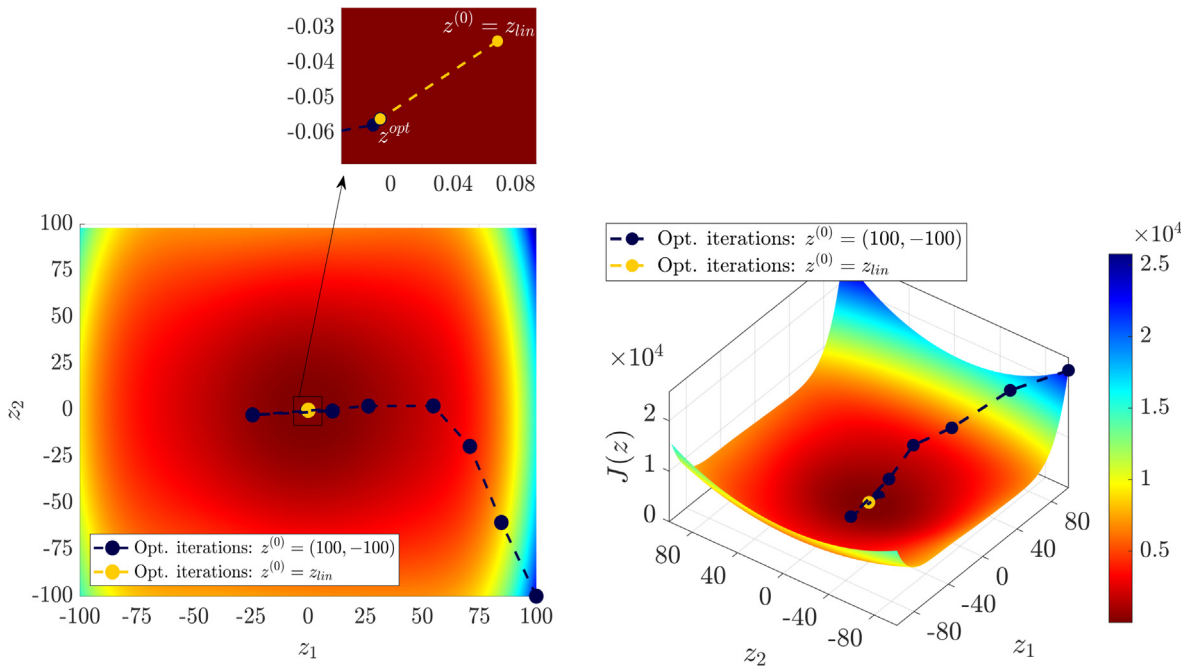


Fig. 12. Most probable path optimization problem objective function using $L = 2$ trial functions and corresponding to a nonlinear oscillator with a bimodal response PDF with $a = 1.5$ and $\varepsilon = 1$ under white noise ($t_f = 1s, x(t_f) = 0.8, \dot{x}(t_f) = 0.9$). The Newton's optimization scheme iterations are also included.

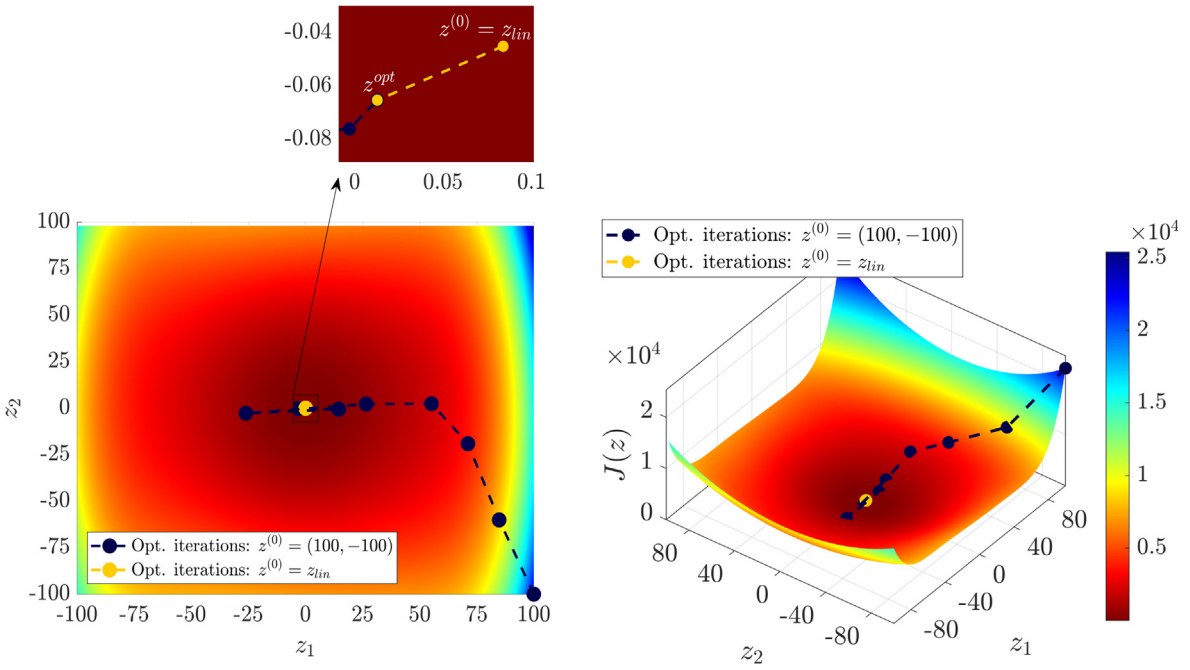


Fig. 13. Most probable path optimization problem objective function using $L = 2$ trial functions and corresponding to a nonlinear oscillator with a bimodal response PDF with $a = 1.8$ and $\varepsilon = 1$ under white noise ($t_f = 1s, x(t_f) = 0.8, \dot{x}(t_f) = 0.9$). The Newton's optimization scheme iterations are also included.

Table 3
Convergence rate and objective function values for a nonlinear oscillator with a bimodal response PDF.

$a = 1.3$		$a = 1.5$		$a = 1.8$	
$\|z^{(k)} - z^*\|$	$J(\boldsymbol{z}^{(k)})$	$\|z^{(k)} - z^*\|$	$J(\boldsymbol{z}^{(k)})$	$\|z^{(k)} - z^*\|$	$J(\boldsymbol{z}^{(k)})$
Newton's numerical optimization scheme - Arbitrarily selected starting point $\boldsymbol{z}^{(0)}$					
141.401	2.35E4	141.386	2.33E4	141.366	2.30E4
104.044	8.73E3	103.987	8.66E3	103.905	8.57E3
73.449	3.79E3	73.522	3.78E3	73.635	3.77E3
54.720	2.02E3	54.932	2.04E3	55.255	2.06E3
26.456	6.13E2	26.538	6.21E2	26.665	6.32E2
23.386	5.47E2	24.484	6.01E2	26.232	6.92E2
8.777	81.474	10.692	1.18E2	14.577	2.13E2
0.742	5.232	1.302	6.274	3.305	15.852
1.75E-3	4.667	4.88E-3	4.514	1.91E-2	4.316
1.15E-8	4.667	9.02E-8	4.514	1.30E-6	4.316
		2.95E-17	4.514	6.38E-15	4.316
Newton's numerical optimization scheme - Starting point $\boldsymbol{z}^{(0)} = -\mathbf{Q}^{-1}\mathbf{b}$ corresponding to a linear oscillator with $\varepsilon = 0$.					
0.072	4.671	0.084	4.520	0.1028	4.325
3.374E-6	4.667	6.193E-6	4.514	1.205E-5	4.316
2.990E-14	4.667	1.058E-13	4.514	4.436E-13	4.316
Computational algebraic geometry approach based on Gröbner bases.					
	4.667		4.514		4.316

5.4. Nonlinear oscillator with a bimodal response PDF

Consider a SDOF nonlinear oscillator exhibiting a bimodal response PDF, whose governing equation is given by

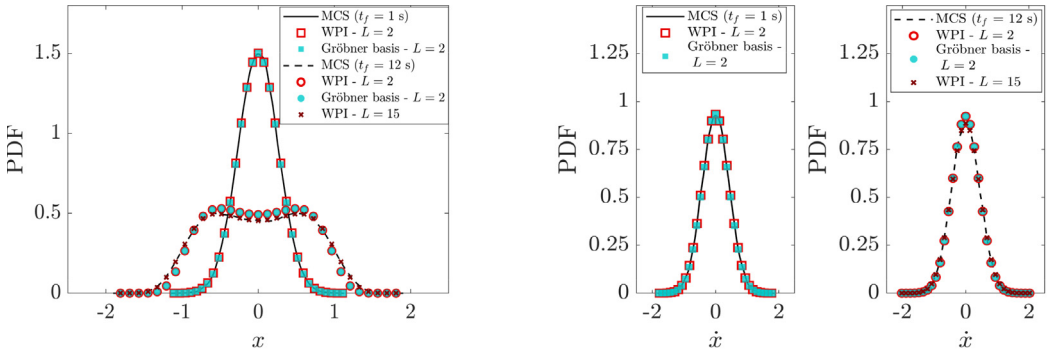
$$m\ddot{x} + c\dot{x} + kx + \varepsilon g_{nl}(x, \dot{x}) = w(t) \tag{57}$$

where

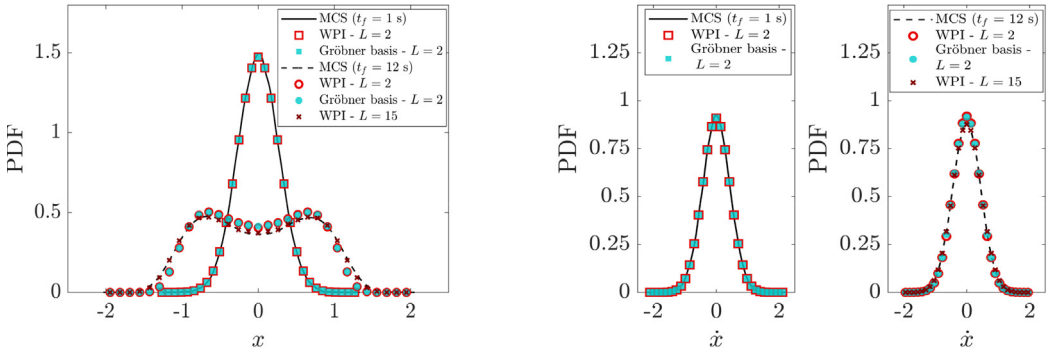
$$g_{nl}(x, \dot{x}) = -ax + x^3 \tag{58}$$

In Eq. (57) and (58), the parameters values utilized are, $m = 1, c = 1.0, k = 1.0, E(w(t)w(t+\tau)) = 2\pi S_0 \delta(\tau)$ with $S_0 = 0.0637$ and a is constant.

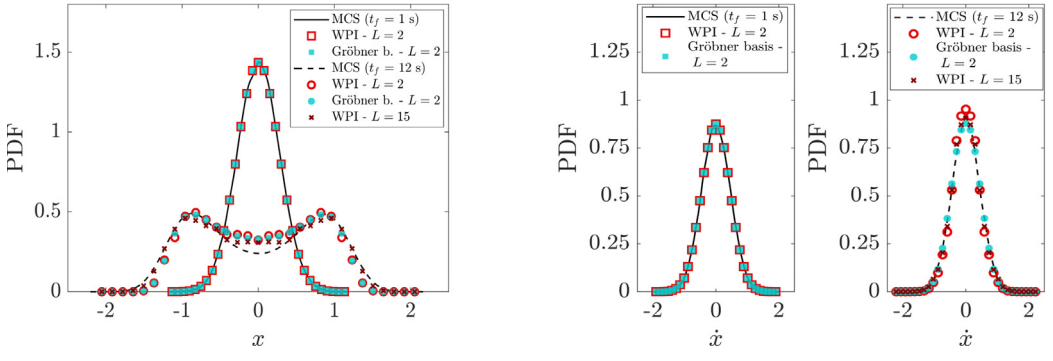
Next, for an indicative final time instant $t_f = 1s$ and for boundary conditions $(x(0), \dot{x}(0), x(t_f), \dot{x}(t_f)) = (0, 0, 0.8, 0.9)$, the objective functions $J(\boldsymbol{z})$ of the most probable path optimization problem for $a = 1.3, 1.5$ and 1.8, considering $\varepsilon = 1$ are shown in Figs. 11–13, respectively. The Newton's scheme iteration points are also included in the figures corresponding both to an arbitrarily selected starting point $\boldsymbol{z}^{(0)} = (-100, 100)$, and to a starting point equal to the exact optimum $\boldsymbol{z}^{(0)} = \boldsymbol{z}_{lin} = -\mathbf{Q}^{-1}\mathbf{b}$ corresponding to the associated linear system (i.e., $\varepsilon = 0$). Obviously, the convergence behavior is highly improved when $\boldsymbol{z}^{(0)} = \boldsymbol{z}_{lin}$ is used. Numerical results related to the iterations of the Newton's scheme are summarized in Table 3, which



(a) $a = 1.3$: Response displacement PDF (b) $a = 1.3$: Response velocity PDF



(c) $a = 1.5$: Response displacement PDF (d) $a = 1.5$: Response velocity PDF



(e) $a = 1.8$: Response displacement PDF (f) $a = 1.8$: Response velocity PDF

Fig. 14. Response displacement and velocity PDFs at various time instants corresponding to a nonlinear oscillator with a bimodal response PDF under white noise. Comparisons between WPI-based estimates utilizing both the Newton's scheme and the Gröbner basis approaches for the most probable path determination. MCS-based estimates are also included (50,000 realizations).

includes also results based on the Gröbner basis approach. In a similar manner as in the previous examples, the single solution obtained by the Gröbner basis approach demonstrates the convexity of the objective function and the existence of a single global minimum.

In Fig. 14, it is shown that the WPI-based estimates utilizing Newton's scheme coincide with the estimates based on the Gröbner basis approach. Comparisons with MCS data (50, 000 realizations) are included as well demonstrating the high accuracy degree exhibited by the WPI technique.

6. Concluding remarks

In this paper, first, a Newton's numerical optimization scheme has been developed for solving the Wiener path integral technique functional minimization problem and for determining the most probable path. Certain numerical aspects have been elucidated and it has been shown that the convergence behavior can be highly improved when the exact solution corresponding to a linear system is used as a starting point in the algorithm. Second, demonstrating the potential convexity (and thus, the existence of a global extremum) of the functional to be minimized has been addressed by resorting to computational algebraic geometry concepts and tools such as Gröbner bases. Various numerical examples pertaining to diverse nonlinear oscillators have been considered, where it has been shown that the associated objective functions are convex, and that the Newton's scheme converges to the globally optimum most probable path. Comparisons with MCS-based estimates have been included as well, demonstrating the reliability of the WPI technique for determining the nonlinear system joint response PDF.

CRediT authorship contribution statement

Ioannis Petromichelakis: Conceptualization, Methodology, Software, Writing - original draft, Visualization. **Rúbia M. Bosse:** Software, Validation, Writing - original draft, Visualization. **Ioannis A. Kougioumtzoglou:** Conceptualization, Methodology, Writing - review & editing, Supervision, Project administration, Funding acquisition. **André T. Beck:** Supervision, Funding acquisition.

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.

Acknowledgement

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

Appendix A. Positive definiteness of matrix Q

In this Appendix, the positive definiteness of matrix **Q** is proved, and thus, convexity of Eq. (25) is also implied. For tutorial effectiveness, the proof is shown hereinafter for a SDOF linear oscillator. In this regard, consider a normalized version of Eq. (52) in the form

$$\ddot{x} + 2\zeta_0\omega_0\dot{x} + \omega_0^2 x = \frac{w(t)}{m} \tag{A.1}$$

where ζ_0 is the damping ratio and ω_0 is the natural frequency of the system. Next, employing two trial functions (i.e., $L = 2$) and considering arbitrary initial and final time instants (t_0 and t_f) in Eq. (22), matrix **Q** is expressed in the form

$$\mathbf{Q} = \begin{bmatrix} Q_{11} & 0 \\ 0 & Q_{22} \end{bmatrix} \tag{A.2}$$

where

$$Q_{11} = \frac{(t_f - t_0)^5}{630}\left(\omega_0{}^4 t_f^4 - 4\omega_0^4 t_f^3 t_0 + 6\omega_0^4 t_f^2 t_0^2 - 4\omega_0^4 t_f t_0^3 + \omega_0^4 t_0^4 + 48\omega_0^2 t_f^2 \zeta_0^2 + -24\omega_0^2 t_f^2 - 96\omega_0^2 t_f t_0 \zeta_0^2 + 48\omega_0^2 t_f t_0 + 48\omega_0^2 t_0^2 \zeta_0^2 - 24\omega_0^2 t_0^2 + 504\right) \tag{A.3}$$

$$Q_{22} = \frac{(t_f - t_0)^5}{6930}\left(\omega_0{}^4 t_f^4 - 4\omega_0^4 t_f^3 t_0 + 6\omega_0^4 t_f^2 t_0^2 - 4\omega_0^4 t_f t_0^3 + \omega_0^4 t_0^4 + 176\omega_0^2 t_f^2 \zeta_0^2 + -88\omega_0^2 t_f^2 - 352\omega_0^2 t_f t_0 \zeta_0^2 + 176\omega_0^2 t_f t_0 + 176\omega_0^2 t_0^2 \zeta_0^2 - 88\omega_0^2 t_0^2 + 3960\right) \tag{A.4}$$

Next, for simplicity and without loss of generality, setting $t_0 = 0$ in Eq. (A.2) yields

$$\mathbf{Q} = \begin{bmatrix} \frac{t_f^5\left(\omega_0^4 t_f^4 + 48\,\omega_0^2 t_f^2 \zeta_0^2 - 24\,\omega_0^2 t_f^2 + 504\right)}{630} & 0 \\ 0 & \frac{t_f^5\left(\omega_0^4 t_f^4 + 176\,\omega_0^2 t_f^2 \zeta_0^2 - 88\,\omega_0^2 t_f^2 + 3960\right)}{6930} \end{bmatrix} \tag{A.5}$$

Since $\mathbf{Q}$ is diagonal, its eigenvalues $\boldsymbol{\lambda}$ are readily determined as

$$\boldsymbol{\lambda} = \begin{bmatrix} \left(\frac{\omega_0^4 t_f^9}{630} + \frac{8\,\omega_0^2 t_f^7 \zeta_0^2}{105} - \frac{4\,\omega_0^2 t_f^7}{105} + \frac{4 t_f^5}{5}\right) \\ \left(\frac{\omega_0^4 t_f^9}{6930} + \frac{8\,\omega_0^2 t_f^7 \zeta_0^2}{315} - \frac{4\,\omega_0^2 t_f^7}{315} + \frac{4 t_f^5}{7}\right) \end{bmatrix} \tag{A.6}$$

Next, setting $y = \omega_0^2$ and considering the most critical case (i.e., $\zeta_0 = 0$) for showing that the eigenvalues $\boldsymbol{\lambda}$ are positive, Eq. (A.6) becomes

$$\boldsymbol{\lambda} = \begin{bmatrix} \left(\frac{y^2 t_f^9}{630} - \frac{4\,y\,t_f^7}{105} + \frac{4 t_f^5}{5}\right) \\ \left(\frac{y^2 t_f^9}{6930} - \frac{4\,y\,t_f^7}{315} + \frac{4 t_f^5}{7}\right) \end{bmatrix} \tag{A.7}$$

Differentiating Eq. (A.6) with respect to y yields

$$\frac{\mathrm{d}\boldsymbol{\lambda}}{\mathrm{d}y} = \begin{bmatrix} \left(\frac{2y t_f^9}{630} - \frac{4 t_f^7}{105}\right) \\ \left(\frac{2y t_f^9}{6930} - \frac{4 t_f^7}{315}\right) \end{bmatrix} \tag{A.8}$$

Setting Eq. (A.8) equal to zero and solving for y leads to

$$y^* = \begin{bmatrix} 12\,t_f^{-2} \\ 44\,t_f^{-2} \end{bmatrix}, \tag{A.9}$$

whereas the eigenvalues evaluated at y^* yield

$$\boldsymbol{\lambda}(y^*) = \begin{bmatrix} 4t_f^5/7 \\ 92t_f^5/315 \end{bmatrix} \tag{A.10}$$

Further, the second derivative of $\boldsymbol{\lambda}$ becomes

$$\frac{\mathrm{d}^2\boldsymbol{\lambda}}{\mathrm{d}y^2} = \begin{bmatrix} t_f^9/315 \\ t_f^9/3465 \end{bmatrix} \tag{A.11}$$

Clearly, since $\mathrm{d}^2\boldsymbol{\lambda}/\mathrm{d}y^2 > 0$, the expression for the eigenvalues $\boldsymbol{\lambda}$ as a function of y is convex, and thus, the points in Eq. (A.9) correspond to minima for $\boldsymbol{\lambda}$. Further, since $\boldsymbol{\lambda}$ is positive at y^*, $\boldsymbol{\lambda}$ is positive for any arbitrary values $\omega_0 > 0$ and $0 < \zeta_0 < 1$. In conclusion, all eigenvalues of matrix $\mathbf{Q}$ are positive, and thus, $\mathbf{Q}$ is positive definite.

Appendix B. Computational algebraic geometry approach indicative results

In this section, indicative results pertaining to the computational algebraic geometry approach of Section 4 are presented. These refer to Examples 5.1 and 5.2. Specifically, considering the linear oscillator of Example 5.1 and the BVP corresponding to Fig. 1, the basis B of standard monomials defined in Eq. (48), degenerates to

$$B = [1] \tag{B.1}$$

This implies that there is only one solution to the corresponding system of polynomial equations. Further, the Gröbner basis corresponding to this example has only two elements and takes the form

$$G = [0.031659 z_1 - 0.00053813, z_2 - 0.000083703]^T \tag{B.2}$$

whereas matrix M_f (see also Theorem 2) is simply a scalar, i.e.,

$$M_f = (\,4.4204132568\,) \tag{B.3}$$

The second example relates to the Duffing oscillator considered in Example 5.2, and corresponding to Fig. 3 ($\varepsilon = 1$). In this regard, the basis B of standard monomials defined in Eq. (48) includes 25 elements and takes the form

$$B = \left[z_2^8, z_1z_2^6, z_2^7, z_2^5, z_1^4, z_1^3z_2, z_1^2z_2^4, z_1z_2^5, z_2^6, z_1^3z_2^2, z_1^2z_2^3, z_1z_2^4, z_1^2z_2^2, z_1z_2^3, z_2^4, z_1^3, z_1^2z_2, z_1z_2^2, z_2^3, z_1^2, z_1z_2, z_2^2, z_1, z_2, 1\right]^T \tag{B.4}$$

This implies that the corresponding system of polynomial equations has 25 solutions, i.e., 24 complex solutions and one real solution, which is the solution of interest. Further, the associated Gröbner basis has six elements, which take the form

$$\begin{aligned} G(1) = {} & 1.4091E14z_1^4 - 7.2961E8z_1^3z_2^2 + 7.9242E13z_1^3z_2 - 1.1746E15z_1^3 - 4.8529E6z_1^2z_2^4 + \\ & 6.9587E12z_1^2z_2^3 - 2.0668E13z_1^2z_2^2 + 1.1819E18z_1^2z_2 - 1.9333E18z_1^2 - \\ & 1173.5z_1z_2^6 - 3.8778E7z_1z_2^5 + 1.2085E9z_1z_2^4 - 1.9927E13z_1z_2^3 + 6.9281E13z_1z_2^2 - \\ & 4.5784E18z_1z_2 + 5.9337E18z_1 + z_2^9 - 55.207z_2^8 - 1.2502E6z_2^7 + 4.4144E7z_2^6 \\ & + 5.3703E11z_2^5 - 9.0655E12z_2^4 + 2.6617E16z_2^3 - 1.3989E18z_2^2 - 2.3572E22z_2 \\ & + 6.2157E19 \end{aligned} \tag{B.5}$$

$$\begin{aligned} G(2) = {} & -2.6453z_1^4 - 0.5488z_1^3z_2^2 - 8799.2z_1^3z_2 - 113988.0z_1^3 - 0.000021788z_1^2z_2^4 - \\ & 0.26732z_1^2z_2^3 + 8.2368z_1^2z_2^2 - 9599.7z_1^2z_2 + 124255.0z_1^2 + 5.064E - 8z_1z_2^7 - \\ & 1.8467E - 6z_1z_2^6 - 0.0046385z_1z_2^5 - 0.021661z_1z_2^4 - 13139.0z_1z_2^3 + \\ & 112933.0z_1z_2^2 - 1.4175E9z_1z_2 - 1.5962E9z_1 - 2.7115E - 7z_2^7 + \\ & 0.000011028z_2^6 + 0.076255z_2^5 - 0.84131z_2^4 + 24378.0z_2^3 - 467722.0z_2^2 - \\ & 3.0075E9z_2 + 4.2444E7 \end{aligned} \tag{B.6}$$

$$\begin{aligned} G(3) = {} & -4.1648E - 6z_1^4 - 9.0668E - 11z_1^3z_2^2 - 3.3071E - 6z_1^3z_2 + 0.000059423z_1^3 + \\ & 7.0522E - 13z_1^2z_2^5 - 1.582E - 11z_1^2z_2^4 + 7.0068E - 8z_1^2z_2^3 - 1.7297E - 6z_1^2z_2^2 + \\ & 0.004472z_1^2z_2 + 0.22987z_1^2 - 6.1856E - 12z_1z_2^5 + 1.5016E - 10z_1z_2^4 - \\ & 6.8703E - 7z_1z_2^3 + 0.000021407z_1z_2^2 + 0.21508z_1z_2 - 0.70789z_1 + 2.137E - 14z_2^7 - \\ & 8.7989E - 13z_2^6 - 2.263E - 8z_2^5 + 4.685E - 7z_2^4 + 0.0075597z_2^3 - 0.02578z_2^2 \\ & + 1523.9z_2 - 4.0113 \end{aligned} \tag{B.7}$$

$$\begin{aligned} G(4) = {} & -3.9843E - 15z_1^4 + 1.5684E - 16z_1^3z_2^3 - 1.9188E - 15z_1^3z_2^2 + 2.1578E - 11z_1^3z_2 - \\ & 8.2704E - 12z_1^3 - 1.6752E - 15z_1^2z_2^3 + 2.1891E - 14z_1^2z_2^2 - 1.4947E - 10z_1^2z_2 + \\ & 9.15E - 12z_1^2 + 1.5178E - 17z_1z_2^5 - 4.2105E - 16z_1z_2^4 - 5.4072E - 12z_1z_2^3 + \\ & 4.6095E - 11z_1z_2^2 - 1.2318E - 6z_1z_2 - 2.1857E - 6z_1 - 6.4897E - 17z_2^5 + \\ & 2.0362E - 15z_2^4 + 2.5295E - 11z_2^3 - 2.4669E - 10z_2^2 - 1.5954E - 7z_2 + 4.7618E - 8 \end{aligned} \tag{B.8}$$

$$\begin{aligned} G(5) = {} & 1.4067E - 19z_1^5 - 1.5986E - 18z_1^4 + 5.2101E - 20z_1^3z_2^2 - 5.3285E - 19z_1^3z_2 - \\ & 4.1555E - 15z_1^3 - 4.476E - 19z_1^2z_2^2 + 4.9023E - 18z_1^2z_2 + 1.6962E - 14z_1^2 + \\ & 2.6949E - 21z_1z_2^4 - 6.3943E - 20z_1z_2^3 - 1.3549E - 15z_1z_2^2 + 8.3192E - 15z_1z_2 + \\ & 6.2799E - 11z_1 - 9.473E - 21z_2^4 + 2.4867E - 19z_2^3 + 3.6591E - 15z_2^2 - \\ & 1.6689E - 14z_2 - 1.354E - 12 \end{aligned} \tag{B.9}$$

$$\begin{aligned} G(6) = {} & 3.2303E - 18z_1^4z_2 - 1.6519E - 17z_1^4 - 3.7001E - 17z_1^3z_2 + 2.0263E - 16z_1^3 + \\ & 6.6833E - 19z_1^2z_2^3 - 1.1893E - 17z_1^2z_2^2 - 1.6801E - 13z_1^2z_2 + \\ & 5.1579E - 13z_1^2 - 4.6986E - 18z_1z_2^3 + 9.2504E - 17z_1z_2^2 + 9.0746E - 13z_1z_2 - \\ & 2.0694E - 12z_1 + 1.078E - 20z_2^5 - 3.6708E - 19z_2^4 - 1.2899E - 14z_2^3 + \\ & 1.9492E - 13z_2^2 + 5.5885E - 9z_2 - 1.4722E - 11 \end{aligned} \tag{B.10}$$

References

[1] M. Grigoriu, Applied Non-Gaussian Processes: Examples, Theory, Simulation, Linear Random Vibration, and MATLAB Solutions, PTR Prentice Hall Upper Saddle River, NJ, 1995.
[2] P.D. Spanos, B.A. Zeldin, Monte Carlo treatment of random fields: A broad perspective, Appl. Mech. Rev. (1998).
[3] E. Vanmarcke, Random Fields: Analysis and Synthesis, World Scientific, 2010.
[4] Y. Lin, Probabilistic Theory of Structural Dynamics, Krieger Publishing Company, 1976.
[5] I. Elishakoff, Probabilistic Methods in the Theory of Structures, Dover Publications, Mineola, New York, 1999.
[6] J.B. Roberts, P.D. Spanos, Random Vibration and Statistical Linearization (Dover Civil and Mechanical Engineering), Dover Publications, Mineola, NY, 2003.
[7] R.G. Ghanem, P.D. Spanos, Stochastic Finite Elements: A Spectral Approach, Courier Corporation, 2003.
[8] J. Li, J. Chen, Stochastic Dynamics of Structures, John Wiley & Sons, 2009.

[9] M. Grigoriu, Stochastic Systems: Uncertainty Quantification and Propagation, Springer Science & Business Media, 2012.
[10] N. Wiener, The average of an analytical functional and the Brownian movement, Proc. Natl. Acad. Sci 7 (1921) 294–298.
[11] L. Onsager, S. Machlup, Fluctuations and irreversible processes, Phys. Rev. 91 (6) (1953) 1505.
[12] S. Machlup, L. Onsager, Fluctuations and irreversible process. ii. systems with kinetic energy, Phys. Rev. 91 (6) (1953) 1512.
[13] L. Tisza, I. Manning, Fluctuations and irreversible thermodynamics, Phys. Rev. 105 (6) (1957) 1695.
[14] D. Falkoff, Statistical theory of irrversible processes: Part i. intergral over fluctuation path formulation, Ann. Phys. 4 (3) (1958) 325–346.
[15] R.L. Stratonovich, On the probability functional of diffusion processes, Selected Transl. in Math. Statist. Prob. 10 (1971) 273–286.
[16] P.C. Martin, E.D. Siggia, H.A. Rose, Statistical dynamics of classical systems, Phys. Rev. A 8 (1) (1973) 423.
[17] H.K. Janssen, On a lagrangean for classical field dynamics and renormalization group calculations of dynamical critical properties, Zeitschrift für Physik B Condensed Matter 23 (4) (1976) 377–380.
[18] C. De Dominicis, Techniques de renormalisation de la théorie des champs et dynamique des phénomènes critiques, J. Phys. Colloques 37 (1976) C1-247–C1-253.
[19] R. Graham, Path integral formulation of general diffusion processes, Zeitschrift für Physik B Condensed Matter 26 (3) (1977) 281–290.
[20] A. Bach, D. Dürr, B. Stawicki, Functionals of paths of a diffusion process and the onsager-machlup function, Zeitschrift für Physik B Condensed Matter 26 (2) (1977) 191–193.
[21] H. Dekker, A functional stieltjes measure and generalized diffusion processes, Physica A 87 (2) (1977) 419–425.
[22] H. Haken, Generalized onsager-machlup function and classes of path integral solutions of the fokker-planck equation and the master equation, Zeitschrift für Physik B Condensed Matter 24 (3) (1976) 321–326.
[23] F. Langouche, D. Roekaerts, E. Tirapegui, Functional Integration and Semiclassical Expansions,, vol. 10, Springer Science & Business Media, 2013.
[24] M. Chaichian, A. Demichev, Path Integrals in Physics: Stochastic Processes and Quantum Mechanics, Institute of Physics Publishing, Bristol, U.K., 2001.
[25] H.S. Wio, Path Integrals for Stochastic Processes: An Introduction, World Scientific Pub Co Inc, 2013.
[26] I.A. Kougioumtzoglou, P.D. Spanos, An analytical Wiener path integral technique for non-stationary response determination of nonlinear oscillators, Probab. Eng. Mech. 28 (2012) 125–131.
[27] 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.
[28] 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.
[29] 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.
[30] 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.
[31] A.F. Psaros, I.A. Kougioumtzoglou, Functional series expansions and quadratic approximations for enhancing the accuracy of the Wiener path integral technique, J. Eng. Mech 146 (7) (2020) 04020065.
[32] 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.
[33] 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.
[34] 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.
[35] I.A. Kougioumtzoglou, A Wiener path integral solution treatment and effective material properties of a class one-dimensional stochastic mechanics problems, J. Eng. Mech. 143 (6) (2017) 04017014.
[36] J. Nocedal, S. Wright, Numerical Optimization (Springer Series in Operations Research and Financial Engineering), Springer, New York, 2006.
[37] W.W. Adams, P. Loustaunau, An introduction to Grobner bases, American Mathematical Soc. (1994).
[38] W. Vasconcelos, Computational Methods in Commutative Algebra and Algebraic Geometry, Springer Science & Business Media, 2004.
[39] D. Eisenbud, Commutative Algebra: with a View Toward Algebraic Geometry, Springer Science & Business Media, 2013.
[40] D. Cox, J. Little, D. O'Shea, Ideals, varieties, and algorithms. undergraduate texts in mathematics, 1997.
[41] D.A. Cox, J. Little, D. O'shea, Using Algebraic Geometry, vol. 185, Springer Science & Business Media, 2006.
[42] G.M. Ewing, Calculus of Variations with Applications, Dover, New York, 1985.
[43] O.C. Zienkiewicz, K. Morgan, Finite Elements and Approximation, Dover Publications, 2006.
[44] M. Raghavan, B. Roth, Solving polynomial systems for the kinematic analysis and synthesis of mechanisms and robot manipulators, J. Mech. Des. 117 (B) (1995) 71–79.
[45] L. Gonzalez-Vega, F. Rouillier, M. Roy, Symbolic recipes for polynomial system solving, in: Some Tapas of Computer Algebra, Springer, 1999, pp. 34–65.
[46] P.A. Parrilo, B. Sturmfels, Minimizing polynomial functions. Algorithmic and quantitative real algebraic geometry, DIMACS Series Discr. Mathe. Theoret. Comput. Sci. 60 (2003) 83–99.
[47] B. Buchberger, Ein algorithmus zum auffinden der basiselemente des restklassenringes nach einem nulldimensionalen polynomideal PhD thesis, Universitat Insbruck, 1965.
[48] B. Buchberger, Ein algorithmisches kriterium für die lösbarkeit eines algebraischen gleichungssystems, Aequationes Math 4 (3) (1970) 374–383.
[49] 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.
[50] Petromichelakis I. Psaros, A.F, 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 (2019) 551–571.