New Partial Differential Equations Governing the Joint, Response-Excitation, Probability Distributions of Nonlinear Systems, under General Stochastic Excitation. II: Numerical Solution

Themistoklis P. Sapsis
Graduate student, National Technical University of Athens, Dept. of Naval Architecture and Marine Engineering

Gerassimos A. Athanassoulis
Professor, National Technical University of Athens, Dept. of Naval Architecture and Marine Engineering

ABSTRACT: In a companion paper (Athanassoulis & Sapsis 2006, *this Conference*) the problem of determining the probabilistic structure of the dynamical response of nonlinear systems subjected to general, external, stochastic excitation has been considered, and new partial differential equations have been derived, governing the joint, response-excitation, characteristic function and probability density function. These new equations are supplemented by a marginal compatibility condition (with respect to the known probability distribution of the forcing), which is of non-local character and, thus, difficult to implement. This is the price paid for discarding the assumption that the forcing is a process of independent increments, which implies that the response is now non-Markovian. In the present paper a method for the numerical solution of these new equations is introduced and illustrated through its application to a specific, simple, nonlinear problem. The solution method is based on the representation of the joint probability density function (or the joint characteristic function) by means of a convex superposition of kernel functions, which permits to satisfy *a priori* the non-local marginal compatibility condition. On the basis of this representation, the partial differential equation that governs the joint, response-excitation, probability density function (or the joint characteristic function) is eventually transformed to a system of ordinary differential equations for the kernel parameters. Numerical results are presented for the case of a first order dynamical system, with cubic nonlinearity, under smooth stochastic excitations. An important feature of the proposed numerical solution method is its suitability to be implemented using parallel or distributed computing schemes.

KEYWORDS: Stochastic Dynamics, Numerical solution of Stochastic Differential Equations, Correlated Stochastic Excitation, Generalized Fokker-Planck-Kolmogorov Equation, Non-Markovian Responses, Kernel Density Functions.

1 INTRODUCTION

In a companion paper (Athanassoulis & Sapsis 2006, *this Conference*, subsequently referred to as [I]) we have considered a simple dynamical system described by the following Stochastic Ordinary Differential Equation (SODE(*)),

$$\dot{x}(t)+k\,x(t)+a\,x^3(t)=y(t,\omega), \tag{1.1a}$$

$$x(t_0)=x_0(\omega), \tag{1.1b}$$

where k,a are deterministic constants, while the forcing term $y(\bullet,\omega)$ is a given, real-valued, stochastic process and the initial value $x_0(\omega)$ is a given, real-valued, random variable, with characteristic function $\phi_0(\upsilon)$, $\upsilon\in\mathbb{R}$.

The underlying probability space is denoted by $(\Omega,\mathscr{B}(\Omega),\mathscr{P}_\Omega)$, where Ω is the sample space, $\mathscr{B}(\Omega)$ is the family of Borel sets of Ω, and $\mathscr{P}_\Omega$ is the corresponding probability measure over Ω. The stochastic process $y(\bullet,\omega)$ is a measurable map $y(\bullet,\omega):\Omega\to\mathscr{Y}$, which defines the induced probability space $(\mathscr{Y}=C^k(I),\mathscr{B}(\mathscr{Y}),\mathscr{P}_y)$, $I=[t_0,T]\subseteq\mathbb{R}$, $k=0$ or $k>0$. Eq. (1.1), provided it has a strong solution, defines another stochastic process $x(\bullet,\omega):\Omega\to\mathscr{X}$ with induced probability space $(\mathscr{X}=C^{k+1}(I),\mathscr{B}(\mathscr{X}),\mathscr{P}_x)$, $\mathscr{P}_x$ being another probability measure to be determined. As we have already seen in [I], use is also made of the joint, response-excitation, process $x(\bullet,\omega)\times y(\bullet,\omega):\Omega\to\mathscr{X}\times\mathscr{Y}$ with induced joint, response-excitation, probability space $(\mathscr{X}\times\mathscr{Y},\mathscr{B}(\mathscr{X}\times\mathscr{Y}),\mathscr{P}_{xy})$. $\mathscr{P}_{xy}$ is another probability measure to be determined. The determination of $\mathscr{P}_{xy}$ is based on the following two condi-

(*) The definitions of most of the abbreviations used in this paper are given at the end of the Introduction of [I]. Some additional abbreviations which will be used herein will be defined in the text, at their first appearance.

tions: i) (Almost) every sample-function couple $\big(x(t,\omega),y(t,\omega)\big)$, $t\in I=[t_0,T]$, satisfies Eqs. (1.1a,b), and ii) The $y-$marginal of $\mathscr{P}_{xy}$ coincides with the known probability measure $\mathscr{P}_y$, associated with the forcing process. The latter condition can be written formally as

$$\int_{\mathscr{X}} \mathscr{P}_{xy}(dx)=\mathscr{P}_y\,. \tag{1.1c}$$

As we have shown in [I], a Hopf-type FDE governing the evolution of the joint, response-excitation, Ch.Fl $\mathscr{F}_{xy}(u,v)$ can be obtained, by exploiting the Volterra functional calculus. From this FDE we have managed to obtain the following linear PDE, governing the evolution of the joint ch.f of the pair of random variables $\big(x(t;\omega),y(t;\omega)\big)$, at each time instance $t>t_0$:

$$\frac{1}{\upsilon}\frac{\partial\phi_{x(t)y(s)}(\upsilon,\nu)}{\partial t}\bigg|_{s=t}+k\frac{\partial\phi_{x(t)y(t)}(\upsilon,\nu)}{\partial\upsilon}-a\frac{\partial^3\phi_{x(t)y(t)}(\upsilon,\nu)}{\partial\upsilon^3}-\frac{\partial\phi_{x(t)y(t)}(\upsilon,\nu)}{\partial\nu}=0\,, \tag{1.2a}$$

where $\phi_{x(t)y(s)}(\upsilon,\nu)$ is the joint ch.f of $x(t;\omega)$ and $y(s;\omega)$. Let it be noted that the function $\phi_{x(t)y(s)}(\upsilon,\nu)$ is dependent on 4 arguments, namely υ,t,ν,s.

Since the stochastic function $y(\bullet,\omega)$, $\omega\in\Omega$, is given, its ch.f $\phi_{y(t)}(\nu)$, $\nu\in\mathbb{R}$, $t\in I$, is known. Hence, the *y*-marginal of the joint ch.f $\phi_{x(t)y(s)}(\upsilon,\nu)$, has to coincide with the excitation marginal $\phi_{y(s)}(\nu)$. Thus, we have the following ***marginal compatibility condition***:

$$\phi_{x(t)y(s)}(0,\nu)=\phi_{y(s)}(\nu),\quad \nu\in\mathbb{R},\quad t,s\ge t_0\,. \tag{1.2b}$$

In addition, the initial condition (1.1b) implies the following ***initial condition*** for the joint characteristic function $\phi_{x(t)y(s)}(\upsilon,\nu)$:

$$\phi_{x(t_0)y(s)}(\upsilon,0)=\phi_{x(t_0)}(\upsilon)=\phi_0(\upsilon)\,,\quad \upsilon\in\mathbb{R}\,. \tag{1.2c}$$

Finally, two obvious, yet essential, conditions that the sought-for function $\phi_{x(t)y(s)}(\upsilon,\nu)$ should obey are the following:

$$\phi_{x(t)y(s)}(0,0)=1\,,\qquad t,s\ge t_0\,, \tag{1.2d}$$

$$\phi_{x(t)y(s)}(\upsilon,\nu)\quad \textit{is non-negative definite w.r.t. its arguments } \upsilon,\nu\,,\quad \textit{for any } t,s\ge t_0\,, \tag{1.2e}$$

which come directly from the fact that it is a characteristic function. The last two conditions will be referred to as ***constitutive conditions***.

The above problem (1.2a-e) can be equivalently reformulated in terms of the corresponding joint, response-excitation, pdf $f_{x(t)y(s)}(x,y)$. Recalling that $f_{x(t)y(s)}(x,y)$ and $\phi_{x(t)y(s)}(\upsilon,\nu)$ constitute a Fourier transform pair, i.e.

$$\phi_{x(t)y(s)}(\upsilon,\nu)=\mathcal{F}_{\substack{x\to\upsilon\\ y\to\nu}}\Big\{f_{x(t)y(s)}(x,y)\Big\}\,,$$

and applying the inverse Fourier transformation $\mathcal{F}^{-1}_{\substack{\upsilon\to x\\ \nu\to y}}\{\bullet\}$ to both sides of conditions (1.2a,b,c,d,e), we readily obtain:

A new linear PDE, with respect to $f_{x(t)y(t)}(x,y)$, of the form

$$\frac{\partial f_{x(t)y(s)}(x,y)}{\partial t}\bigg|_{s=t}+\frac{\partial}{\partial x}\Big[\big(kx+ax^3\big)f_{x(t)y(t)}(x,y)\Big]+\frac{\partial}{\partial x}\big[y\,f_{x(t)y(t)}(x,y)\big]=0\,, \tag{1.2a$'$}$$

the follow new forms of the ***marginal compatibility condition*** and of the ***initial condition***

$$\int_{\mathbb{R}} f_{x(t)y(s)}(x,y)\,dx=f_{y(s)}(y)\,,\quad y\in\mathbb{R}\,,\quad t,s\ge t_0\,, \tag{1.2b$'$}$$

$$\int_{\mathbb{R}} f_{x(t_0)y(s)}(x,y)\,dy=f_{x(t_0)}(x)=f_0(x)\,,\quad x\in\mathbb{R}\,, \tag{1.2c$'$}$$

and the corresponding new forms of the ***constitutive conditions***

$$\int_{\mathbb{R}\times\mathbb{R}} f_{x(t)y(s)}(x,y)\,dxdy=1,\qquad t,s\ge t_0 \tag{1.2d$'$}$$

$$f_{x(t)y(s)}(x,y)\ge 0\,,\quad \textit{for any } x,y\in\mathbb{R} \textit{ and any } t,s\ge t_0\,. \tag{1.2e$'$}$$

Clearly, problem (1.2) –either in the form (1.2a,b,c,d,e) or in the form (1.2a$'$b$'$c$'$d$'$e$'$)– exhibits some peculiarities making it distinctly different from the usual initial-boundary value problems for PDEs, coming from problems of Mathematical Physics. These peculiarities reflect the probabilistic origin of the present problem.

In the present work an original method for the numerical solution of problem (1.2) is developed, and some first, illustrative, numerical results are presented. The main tool, on which the formulation of the numerical scheme relies, is the representation of

the sought-for pdf and ch.f by means of a convex superposition of kernel density functions (kdfs) and their Fourier transformation, the kernel characteristic functions (kch.fs), respectively. Then, taking advantage of the linearity of the PDE (1.2a or 1.2a′), we derive a simple system of ODEs that govern the evolution of the kernel parameters, ***locally in time***. In the ***longer-time scale***, the evolution of the probability distribution is approximately described by updating the representation of the sought-for pdf, in discrete time, using the updated kernels. Possible improvements of this scheme are also discussed.

A short presentation of the basic facts about kdfs is given in the next section.

2 KERNEL DENSITY FUNCTIONS

Kernel density functions constitute a key notion/tool within the framework of nonparametric statistical estimation. See, e.g., Scott 1992. In our approach, a kdf $K(x;x_*,h)$ is mainly thought of as a generalized (non-symmetric) summability kernel, appropriate to represent pdfs (Gavriliadis 2005). The defining properties of an $M-$variate kdf are the following:

Pr.1) $K(x;x_*,h)$ is a continuous, real-valued function defined in a domain of the form $\mathscr{D}_K = A\times A\times \mathscr{M}_{M\times M}^{NonNegDef}$, where $A\subseteq \mathbb{R}^M$ is taken to be contained in (or to be equal to) the support of a target pdf, say $f(x)$, which is to be represented (see Lemma 2.1 and Theorem 2.2, below), and $\mathscr{M}_{M\times M}^{NonNegDef}$ is the set of non-negative definite, $M\times M-$ matrices, which can serve as covariance matrices,

Pr.2) $K(x;x_*,h)\geq 0$, for $(x;x_*,h)\in \mathscr{D}_K$,

Pr.3) $\int_A K(x;x_*,h)\,dx_* = 1$, for $(x,h)\in A\times \mathscr{M}_{M\times M}^{NonNegDef}$,

Pr.4) $\lim_{\|h\|\to 0} \int_{\|x-x_*\|>\delta} K(x;x_*,h)\,dx_* = 0$,

for any $x_*\in A$ *and* $\delta>0$.

Clearly, properties Pr.2), Pr.3) ensure that each kdf is a pdf on its own. The ***shape*** of the kernel function $K(x;x_*,h)$ is controlled by its covariance matrix h, also called *bandwidth* (or *shape*) *parameter*. h quantifies the spreading of the kernel probability mass around its "***center***" x_*. Another –simpler and in many cases adequate– choice of the shape parameter is the *M*-variate vector of the eigenvalues of the covariance matrix. In this sense, the domain $\mathscr{D}_K = A\times A\times \mathscr{M}_{M\times M}^{NonNegDef}$ can be (and will be) simplified as $A\times A\times [0,\infty)^M$.

Using the defining properties Pr.1) – Pr.4), and only these, it is not difficult to prove the following

Lemma 2.1: If $f(x)$ is a continuous pdf and $K(\cdot;\cdot,\cdot)$ is any kernel function satisfying Pr.1) – Pr.4), then, for any x,

$$\lim_{\|h\|\to 0}\int_A K(x;x_*,h)\,f(x_*)\,dx_* = f(x). \quad \blacksquare \qquad (2.1)$$

That is, as the bandwidth decreases, the kernel function shrinks around its "center" x_*, having the weak asymptotic limit

$$K(x;x_*,h)\xrightarrow{\|h\|\to 0}\delta(x-x_*). \qquad (2.2)$$

On the other hand, as the bandwidth increases the kernel function spreads out.

Theorem 2.2: The set of all convex finite superpositions of the form $\sum_{n=1}^{N} p_n\,K(x;x_n,h_n)$, where $p_1+p_2+\ldots+p_N=1$, $p_n\geq 0$ for all n, and $K(\cdot;\cdot,\cdot)$ is any kernel function satisfying Pr.1) – Pr.4), is dense within the set of all continuous pdfs supported in A. That is, given any continuous pdf $f(x)$, a specific kernel function $K(x;x_*,h)$, and an arbitrary (small) number $\varepsilon>0$, there exist a bandwidth parameter h_*, a finite set of centers $\{x_n\}_{n=1}^N$ in A, and a vector $\mathbf{p}=(p_1,p_2,\ldots,p_N)$ lying in the positive cone of $\mathbb{R}^N$, such that

$$\max_{x\in A}\left|f(x)-f^N(x)\right| < \varepsilon, \qquad (2.3a)$$

where $f^N(x)=\sum_{n=1}^{N} p_n K(x;x_n,h_n)$. $\blacksquare$ (2.3b)

The clue of the proof of this theorem is Lemma 2.1, in conjunction with the properties of the Riemann sum approximation of the integral $\int_A K(x;x_*,h)\,f(x_*)\,dx_*$ (Athanassoulis and Gavriliadis 2002). The technical details are omitted.

The above theorem makes clear that any (continuous) pdf can be approximated, as closely as it is required, by a representation of the form (2.3b).

3 REFORMULATION OF THE PROBLEM BY USING KERNEL DENSITY REPRESENTATIONS

We shall now apply the representation (2.3), Theorem 2.2, in order to reformulate problem (1.2) in a

way facilitating its numerical solution. Again here and subsequently, as in the Introduction, x, y are scalars, and $f_{xy} = f_{x(t)y(s)}(x,y)$, $\phi_{xy} = \phi_{x(t)y(s)}(\upsilon,\nu)$ are four-argument, two-variate, joint, response-excitation pdf and ch.f, respectively. For clarity, in the present and the subsequent sections, vector or matrix quantities will be explicitly denoted by using bold letters.

Applying the representation (2.3) for the pdf, and the corresponding one for the ch.f, obtained by means of a Fourier transformation, we define the approximants

$$f^N_{x(t)y(s)}(x,y) = \sum_{k=1}^{N} p_k(t,s) K\big(x,y;\mathbf{m}^k(t,s),\mathbf{h}^k(t,s)\big), \tag{3.1}$$

$$\begin{aligned}\phi^N_{x(t)y(s)}(\upsilon,\nu) &= \mathcal{F}_{\substack{x\to\upsilon \\ y\to\nu}}\Big\{ f^N_{x(t)y(s)}(x,y)\Big\} = \\ &= \sum_{k=1}^{N} p_k(t,s)\tilde{K}\big(\upsilon,\nu;\mathbf{m}^k(t,s),\mathbf{h}^k(t,s)\big).\end{aligned} \tag{3.2}$$

Here $\tilde{K}\big(\upsilon,\nu;\mathbf{m}^k,\mathbf{h}^k\big) = \tilde{K}\big(\upsilon,\nu;\mathbf{m}^k(t,s),\mathbf{h}^k(t,s)\big)$ is the kch.f obtained by Fourier transformation of the kdf $K\big(x,y;\mathbf{m}^k,\mathbf{h}^k\big) = K\big(x,y;\mathbf{m}^k(t,s),\mathbf{h}^k(t,s)\big)$, $\mathbf{m}^k = \big(m^k_x, m^k_y\big)$ is the location parameter, namely the position of the most probable (highest) value of the kdf, and $\mathbf{h}^k$ is the shape parameter, represented either by the $2\times 2-$ covariance matrix of the kdf or by the two eigenvalues of the latter (both pictures will be applied to the numerical treatment). For the numerical computations $K\big(x,y;\mathbf{m}^k,\mathbf{h}^k\big)$ is taken to be a Gaussian pdf. (See, e.g., Härdle 1990, Sec. 2.9).

Our main goal now is to exploit the representations (3.1), (3.2), in order to solve the system (1.2a,b,c) or the equivalent (1.2 a′b′c′). Conditions (1.2d,e), or the equivalent (1.2 d′e′), are automatically satisfied since the approximants (3.1), (3.2) are by construction pdfs and ch.fs, respectively.

To facilitate the discussion, let us denote eq. (1.2 a′) by

$$\mathcal{L}\big[f_{xy}\big](x,y,t) = 0\ , \qquad (x,y)\in\mathbb{R}^2,\ \ t\ge t_0\,, \tag{3.3a}$$

where $\mathcal{L}[\bullet](x,y,t)$ is the linear differential operator

$$\mathcal{L}\bullet = \left.\frac{\partial \bullet}{\partial t}\right|_{s=t} + k\frac{\partial[x\bullet]}{\partial x} + a\frac{\partial\big[x^3\bullet\big]}{\partial x} - \frac{\partial[y\bullet]}{\partial x}. \tag{3.3b}$$

Also, equation (1.2a), obtained by a Fourier transformation of the above, will be denoted by

$$\tilde{\mathcal{L}}\big[\phi_{xy}\big](\upsilon,\nu,t) = 0\ , \qquad (\upsilon,\nu)\in\mathbb{R}^2,\ \ t\ge t_0\,, \tag{3.4a}$$

where $\tilde{\mathcal{L}}[\bullet](\upsilon,\nu,t)$ is the linear differential operator

$$\tilde{\mathcal{L}}\bullet = \left.\frac{\partial \bullet}{\partial t}\right|_{s=t} + k\upsilon\frac{\partial\bullet}{\partial\upsilon} - a\upsilon\frac{\partial^3\bullet}{\partial\upsilon^3} - \upsilon\frac{\partial\bullet}{\partial\nu}. \tag{3.4b}$$

It is interesting to note here that the two equivalent formulations –(1. 2a′) or (3.3) in terms of the pdf, and (1.2a) or (3.4) in terms of the ch.f– are both useful and they will be considered in parallel, since the conceptual arguments are better stated using the pdf formulation, while the numerical analysis is better developed using the ch.f formulation.

Substituting the approximation (3.1) into eq. (3.3), we obtain

$$\sum_{k=1}^{N}\mathcal{L}\Big[p_k K\big(x,y;\mathbf{m}^k,\mathbf{h}^k\big)\Big] = 0\ , \qquad (x,y)\in\mathbb{R}^2. \tag{3.5}$$

Let us denote by $\varepsilon\big(\mathbf{h}^j\big)$ the radius of the effective support of $K\big(x,y;\mathbf{m}^j,\mathbf{h}^j\big)$. $\varepsilon\big(\mathbf{h}^j\big)$ will be taken and always kept to be small. Since, thus, each kernel function $K\big(x,y;\mathbf{m}^j,\mathbf{h}^j\big)$ is taken to be concentrated around its center $\mathbf{m}^j = \big(m^j_x, m^j_y\big)$ and it is positive there, eq. (3.5), restricted in a neighborhood $\mathcal{N}\big(\mathbf{m}^j,\varepsilon\big(\mathbf{h}^j\big)\big)$, is locally equivalent with the equation

$$\begin{aligned}&\sum_{k=1}^{N}\mathcal{L}\Big[p_k K\big(x,y;\mathbf{m}^k,\mathbf{h}^k\big)\Big]\cdot K\big(x,y;\mathbf{m}^j,\mathbf{h}^j\big) = 0, \\ &\qquad (x,y)\in\mathcal{N}\big(\mathbf{m}^j,\varepsilon\big(\mathbf{h}^j\big)\big)\end{aligned} \tag{3.6}$$

Assuming that the system of neighborhoods $\Big\{\mathcal{N}\big(\mathbf{m}^j,\varepsilon\big(\mathbf{h}^j\big)\big),\ j=1,\ldots,N\Big\}$ covers the essential support of the sought-for density function f_{xy}, we can assert that the global equation (3.5) is equivalent with the system of local equations

$$\begin{aligned}&\sum_{k=1}^{N}\mathcal{L}\Big[p_k K\big(x,y;\mathbf{m}^k,\mathbf{h}^k\big)\Big]\cdot K\big(x,y;\mathbf{m}^j,\mathbf{h}^j\big) = 0, \\ &\qquad \forall\ j\in\{1,\ldots,N\},\ \text{and} \\ &\forall\,(x,y)\in\bigcup_{j=1,\ldots,N}\mathcal{N}\big(\mathbf{m}^j,\varepsilon\big(\mathbf{h}^j\big)\big).\end{aligned} \tag{3.7}$$

By taking a Fourier transformation, eq. (3.7) is equivalently rewritten as

$$\begin{aligned}&\sum_{k=1}^{N}\tilde{\mathcal{L}}\Big[p_k\tilde{K}\big(\upsilon,\nu;m^k,h^k\big)\Big] * \tilde{K}\big(\upsilon,\nu;m^j,h^j\big) = 0, \\ &\qquad \forall\ j\in\{1,\ldots,N\},\ \text{and} \\ &\forall\,(\upsilon,\nu)\in\mathbb{R}^2,\end{aligned} \tag{3.8}$$

where $*$ denotes the convolution operator. Although the latter equation could be considered as being more complicated than eq. (3.7), an efficient numerical solution scheme will be based on it.

4 A TWO-LEVEL NUMERICAL SOLUTION SCHEME OF THE SET OF EQs. (3.8)

To proceed with use will be made of a specific choice of the kdf. Assuming a Gaussian density as the kdf, we have

$$K\left(x,y;\mathbf{m}^k,\mathbf{C}^k\right)= \frac{1}{2\pi\sqrt{\left|\det\left[\boldsymbol{C}\right]\right|}}\exp\left[-\frac{1}{2}\begin{pmatrix}x-m_x^k\\ y-m_y^k\end{pmatrix}^T\left[\boldsymbol{C}^k\right]^{-1}\begin{pmatrix}x-m_x^k\\ y-m_y^k\end{pmatrix}\right], \tag{4.1a}$$

and corresponding characteristic function

$$\tilde{K}\left(\upsilon,\nu;\mathbf{m}^k,\mathbf{C}^k\right)=\exp\left(i\begin{bmatrix}m_x^k\\ m_y^k\end{bmatrix}^T\cdot\begin{bmatrix}\upsilon\\ \nu\end{bmatrix}-\frac{1}{2}\begin{bmatrix}\upsilon\\ \nu\end{bmatrix}^T\cdot\left[\boldsymbol{C}^k\right]\cdot\begin{bmatrix}\upsilon\\ \nu\end{bmatrix}\right), \tag{4.1b}$$

where

$$\mathbf{m}^k=\mathbf{m}^k\left(t,s\right)=\left(m_x^k\left(t\right),\, m_y^k\left(s\right)\right) \tag{4.2a}$$

is the mean vector, and

$$\boldsymbol{C}^k=\boldsymbol{C}^k\left(t,s\right)=\begin{pmatrix}C_{xx}^k\left(t,t\right) & C_{xy}^k\left(t,s\right)\\ C_{yx}^k\left(s,t\right) & C_{yy}^k\left(s,s\right)\end{pmatrix} \tag{4.2b}$$

is the covariance matrix of our Gaussian kdf. As we have already mentioned above, a dual realization of the shape parameter will be considered herewith. Apart from the covariance matrix $\boldsymbol{C}^k$, the vector $\mathbf{h}^k=\left(h_x^k,h_y^k\right)$ having as elements the two eigenvalues of the matrix $\boldsymbol{C}^k$, will also be used in this case.

Our numerical solution scheme will be implemented by restricting the kdf to be highly concentrated, so that the effective supports of any pair of two different kernels to be practically ***non overlapping***. This permits us to neglect the interaction between any pair of Gaussian kernels, i.e. to disregard the summation in the right-hand side of eq. (3.7), and its equivalent eq. (3.8). Thus, under the above assumption, which is equivalent with the condition $\left\|\mathbf{h}^k\right\|<\varepsilon_1$, for all $k\in\{1,\ldots,N\}$, where ε_1 is an appropriate (small) constant, eq. (3.8) simplifies to

$$\tilde{\mathcal{L}}\left[p_j\tilde{K}\left(\upsilon,\nu;\mathbf{m}^j,\mathbf{C}^j\right)\right]=0,\quad t\geq s\geq t_0,\ \textit{and}\quad \forall\ j\in\{1,\ldots,N\},\ \textit{and}\quad \forall\ (\upsilon,\nu)\in\mathbb{R}^2. \tag{4.3a}$$

Furthermore, assuming the amplitudes p_j positive and piecewise constant, the above equation is further simplified to

$$\tilde{\mathcal{L}}\left[\tilde{K}\left(\upsilon,\nu;\mathbf{m}^j,\mathbf{C}^j\right)\right]=0,\ \textit{within each time interval}\ \tau^{(\ell)}\leq s\leq t\leq\tau^{(\ell+1)},\ \textit{and}\ \forall\ j\in\{1,\ldots,N\},\ \textit{and}\ \forall\ (\upsilon,\nu)\in\mathbb{R}^2. \tag{4.3b}$$

On the basis of the above discussion, a two-level (two-time scale) approach comes into the scene:

a. Solve the set of independent equations (4.3b) within each interval $\tau^{(\ell)}\leq s\leq t\leq\tau^{(\ell+1)}$ (this is the ***short-time phase*** or ***inner-cycle phase***), and then
b. Come back to the complete representation and update the values of the amplitudes p_j, passing from the interval $\left[\tau^{(\ell)},\,\tau^{(\ell+1)}\right]$ to the interval $\left[\tau^{(\ell+1)},\,\tau^{(\ell+2)}\right]$ (this is the ***coarse-time phase*** or ***the outer-cycle phase***).

The ***criterion*** for defining the sequence of coarse updating times $\tau^{(\ell)}$, $\ell=1,2,3,\ldots$, is formulated as a sufficient condition for the validity of the assumptions underlying the derivation of the set of independent equations (4.3b). It turns out that the most critical assumption is the restriction of each kdf to be highly concentrated around its center. As expected, because of the diffusive character of the problem, it has been found that, during the short-time phase solution, kernel parameters evolve in a way leading to a continuous increase of the variance parameter $\left\|\mathbf{h}^k\right\|$. (See, for example, Figures 2c, 3c, in Section 5, below, and the discussion therein). The growth of the quantity $\left\|\mathbf{h}^k\right\|$ leads to the spreading of the mass of the corresponding kdf, which results in the violation of the assumption of negligible interaction between the kernels.

Thus, the set of kernel parameters $\mathbf{m}^k\left(t,s\right)$ and $\boldsymbol{C}^k\left(t,s\right)$ evolve in accordance with the simplified dynamical equations (4.3b) from time $\tau^{(\ell)}$, until the spreading index $\left\|\mathbf{h}^k\left(t\right)\right\|$, of some kernel, exceeds a certain critical value, say $\varepsilon_1>0$. This value of t is taken to be the next updating time $\tau^{(\ell+1)}$. At that time instant, the inner-cycle (short-time) solution phase is interrupted, and an approximation of the to-

tal joint pdf $f^{N}_{x\left(\tau^{(\ell+1)}\right)y\left(\tau^{(\ell+1)}\right)}\left(x,y\right)$ is calculated by means of eq. (3.1), in the specific form:

$$f^{N}_{x\left(\tau^{(\ell+1)}\right)y\left(\tau^{(\ell+1)}\right)}\left(x,y\right)= \\ = \sum_{k=1}^{N} p_k\left(\tau^{(\ell)}\right) K\left(x,y\,;\mathbf{m}^k\left(\tau^{(\ell+1)},\tau^{(\ell+1)}\right),\mathbf{h}^k\left(\tau^{(\ell+1)},\tau^{(\ell+1)}\right)\right). \tag{4.4}$$

Then, the calculated pdf (4.4) is re-approximated, by using a new set of kdfs, satisfying the concentration condition $\left\|\mathbf{h}^k\right\| = \varepsilon_2 < \varepsilon_1$, with different amplitudes $p_k\left(\tau^{(\ell+1)}\right)$. The latter are calculated by means of an optimization algorithm (used also for the set up of the initial conditions), which is described in the Appendix. After the updating of the amplitudes, the next inner-cycle begins, and the procedure continues as described above.

During each time interval $\left[\tau^{(\ell)},\,\tau^{(\ell+1)}\right]$ the amplitudes are considered constant and, thus, globally, p_j are piecewise constant functions of time. In fact, the evolution of the amplitudes p_j is much slower than the evolution of the kernel parameters $\mathbf{m}^k$ and $\mathbf{h}^k$, and this is what justifies the piece-wise constant assumption for p_j in our numerical scheme. Cf Sofi et al 2006 (*this Conference*), where a similar finding came out when the sought-for pdf is represented as series with respect to appropriate, time-dependent, basis functions. An improved numerical solution, taking also into account the evolution of p_j in a continuous fashion, can be constructed and will be published elsewhere.

It should be stressed that the accuracy of the method proposed and developed herewith is critically dependant on the threshold value ε_1 for the variance parameter (spreading index) $\left\|\mathbf{h}^k\left(t\right)\right\|$.

5 A LOCAL-MOMENT METHOD FOR THE NUMERICAL SOLUTION OF EQs. (4.3b)

We are now focusing on the numerical treatment of equations (4.3b). For each value of $j\in\left\{1,\ldots,N\right\}$, eq. (4.3b) contains three unknown functions, namely the response mean value $m_x^j\left(t\right)$, and covariances $C_{xx}^j\left(t,t\right)$ and $C_{xy}^j\left(t,s\right)$, which should be determined, and two known functions, namely the excitation mean value $m_y^j\left(s\right)$ and the autocovariance $C_{yy}^j\left(s,s\right)$, introducing the appropriate, inner-cycle, forcing. Thus, any solution scheme of eq. (4.3b) should provide us with a number of equations (hopefully three) governing the evolution of the three unknown functions, along with the evidence that introducing the obtained solution in the operator $\tilde{\mathcal{L}}\left[\tilde{K}\left(\upsilon,\nu;\mathbf{m}^j,\mathbf{C}^j\right)\right]$ will result in 0 (at least approximately) for all values of $\left(\upsilon,\nu\right)\in\mathbb{R}^2$.

Since the (Gaussian) kernel $\tilde{K}\left(\upsilon,\nu;\mathbf{m}^j,\mathbf{C}^j\right)$ is $C^{\infty}\left(\mathbb{R}^2\right)$ in $\left(\upsilon,\nu\right)$ and dies out as $\left\|\left(\upsilon,\nu\right)\right\|\to\infty$, eq. (4.3b) is equivalent to the following system of localized moment equations:

$$\frac{\partial^{\,p+q}}{\partial^p\upsilon\,\partial^q\nu}\tilde{\mathcal{L}}\left[\tilde{K}\left(\upsilon,\nu;\mathbf{m}^j,\mathbf{C}^j\right)\right]\Bigg|_{\substack{\upsilon=0\\ \nu=0}} = 0\,, \tag{5.1}$$

for any $\left(p,q\right)\in\mathbb{N}_0\times\mathbb{N}_0$, *where* $\mathbb{N}_0=\left\{0,1,2,3,\ldots\right\}$.

Exploiting the specific (Gaussian) form of the kernel, and considering the cases $\left(p,q\right)=\left(1,0\right)$, $\left(2,0\right)$ and $\left(1,1\right)$, the following three (nonlinear) ODEs are obtained from (5.1):

$$m_{x,t}^k\left(t\right)+k\,m_x^k\left(t\right)+3am_x^k\left(t\right)C_{xx}^k\left(t,t\right)+a\left[m_x^k\left(t\right)\right]^3=m_y^k\left(t\right) \tag{5.2a}$$

$$C_{xx,t}^k\left(t,s\right)+kC_{xx}^k\left(t,s\right)+3am_x^k\left(t\right)C_{xx}^k\left(t,s\right)m_x^k\left(t\right)+ \\ +3aC_{xx}^k\left(t,t\right)C_{xx}^k\left(t,s\right)=C_{xy}^k\left(t,s\right) \tag{5.2b}$$

$$C_{xy,t}^k\left(t,s\right)+kC_{xy}^k\left(t,s\right)+3aC_{xx}^k\left(t,t\right)C_{xy}^k\left(t,s\right)+ \\ +3am_x^k\left(t\right)C_{xy}^k\left(t,s\right)m_x^k\left(t\right)=C_{yy}^k\left(t,s\right) \tag{5.2c}$$

These equations involve the three unknown functions $m_x^j\left(t\right)$, $C_{xx}^j\left(t,t\right)$ and $C_{xy}^j\left(t,s\right)$, and they are differential equations with respect to t, parametrically dependent on s. (No derivatives with respect to s appear.) They should be satisfied for all values of $\left(t,s\right)$ such that $\tau^{(\ell)}\le s\le t\le\tau^{(\ell+1)}$. We are especially interesting in the solution of system (5.2) on the diagonal $s=t$.

It has been found that if the three moment equations (5.2a,b,c) holds true, then various other –but not all– moments, corresponding to other values of $\left(p,q\right)$, are also zero. Besides, there are also values of $\left(p,q\right)$, corresponding to higher-order moments, for which eqs. (5.1) are not satisfied. In any case, the system (5.2a,b,c) is closed and can be efficiently solved, providing us with a reasonable approximation of the evolution of the kernel parameters $m_x^j\left(t\right)$, $C_{xx}^j\left(t,\,s=t\right)$ and $C_{xy}^j\left(t,\,s=t\right)$. When the value of $\left\|\mathbf{h}^k\left(t\right)\right\|$ exceeds the threshold value ε_1, the current inner-cycle phase is finished and the procedure switches to outer-cycle phase.

The numerical solution of the set of nonlinear ODEs (5.2) is implemented by using the method of

the ***quasilinearization*** (Bellman 1973, Lakshmikantham and Malek 1994). Taking advantage of the symmetry properties of the correlation matrix, the equations can be solved on the 'diagonal', that is around $s=t$. The sequence of time instants for the numerical scheme has the form

$$(t,s):\ (t_i,t_i) \rightarrow (t_{i+1},t_i) \rightarrow (t_{i+1},t_{i+1}).$$

An important aspect of the present method is its suitability for parallel computation. Parallelization techniques can be applied both to the dynamical evolution of the kernels and to the optimization algorithm (see Appendix). In the first case the algorithm can take advantage of the independent evolution of each kernel. For the parallelization of the optimization algorithm we can split the group of kdfs into subgroups and then independently approximate each subgroup by new kernels with small variance. Hence, we can probably succeed fast computations for systems of higher dimensions, subjected to general (smooth) excitation.

6 NUMERICAL EXAMPLES

We shall now apply the above described numerical scheme to the numerical determination of the response pdf of a dynamical system (1.1), excited by a known stochastic process (see below), with system parameters a and k having the values given in Table 1, under Case I and Case II.

Table 1: System Parameters

System parameters	Case I	Case II
k	1	1
a	1	-1

By performing a stability analysis to problem (1.1) we found that for $a>0$ (and thus also for the Case I in Table 1), the nonlinear system has one stable fixed point located at zero. A pitchfork bifurcation occurs at $a=0$, and the fixed point at zero becomes unstable in the semi-axis $a<0$. In the same region $(a<0)$ two symmetric stable points appear at $\pm\sqrt{k/|a|}$. Hence, we have the bifurcation diagram shown in Figure 1.

On the basis of the above described dynamical features of the studied problem, it is natural to expect that, in Case I, the evolved pdf will become eventually a unimodal distribution centered at zero, while in Case II, the probability will concentrate around the pair of the two symmetric stable fixed points $\pm\sqrt{k/|a|}$, hence ultimately a bimodal distribution will appear. Since the stable fixed points are global attractors we expect to attain these results after some time, independently of the initial density. The numerical results to be presented and discussed below clearly comply with this behavior, dictated by the qualitative analysis of the studied system.

Consider first the Case I, with a bimodal initial pdf, defined as a convex superposition of two Gaussians with parameter values $m_1=0,\ m_2=0.6$, $\sigma_1=0.1,\ \sigma_2=0.6$, and amplitudes $p_1=0.4$ and $p_2=0.6$, respectively. This initial pdf is shown in Figure 2b, at the section $t=0$. The excitation process is taken to be, in this case, a Gaussian stationary random function with zero mean and covariance function given by

$$C_{YY}(\tau)=\frac{1}{2}\cos^2(2\tau). \qquad (6.1)$$

Numerical results are presented in Figure 2. More specifically, in the two upper plots of this figure (Figures 2a and 2b), the evolution of the probability density $f_{x(t)}(x)$ is shown, for the time interval $0<t<1.4\,\text{sec}$, enough to get the steady state response pdf. Also, in the same figure (Figure 2a) the orbits of $m_x^k(t)$ are plotted by using thick black lines. The apparent discontinuities every 0.2sec are due to the re-approximation of the calculated density by means of a new convex superposition of kdfs with smaller variance, every time the concentration parameters h^k exceeds the critical value ε_1 (which in this example was set to $\varepsilon_1=0.3$). In Figure 2c the evolution of the variance for some kdfs of the response density is shown. The diffusive character of the evolution (strictly increasing variances with respect to time) is clearly seen in the numerical results. Again the apparent discontinuities are due to the redistribution of the response pdf by using kdfs of smaller variances.

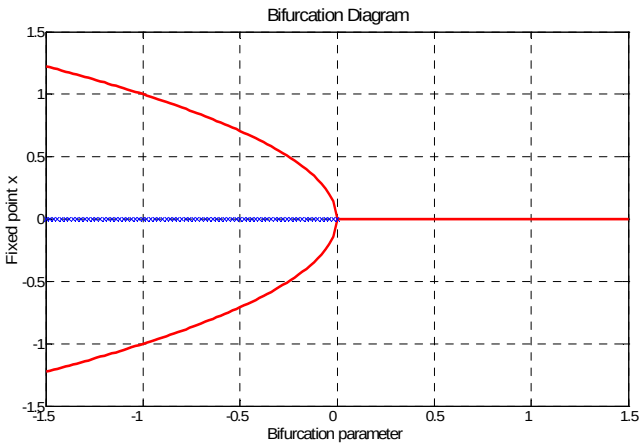

Figure 1: Bifurcation diagram for system (1.1).

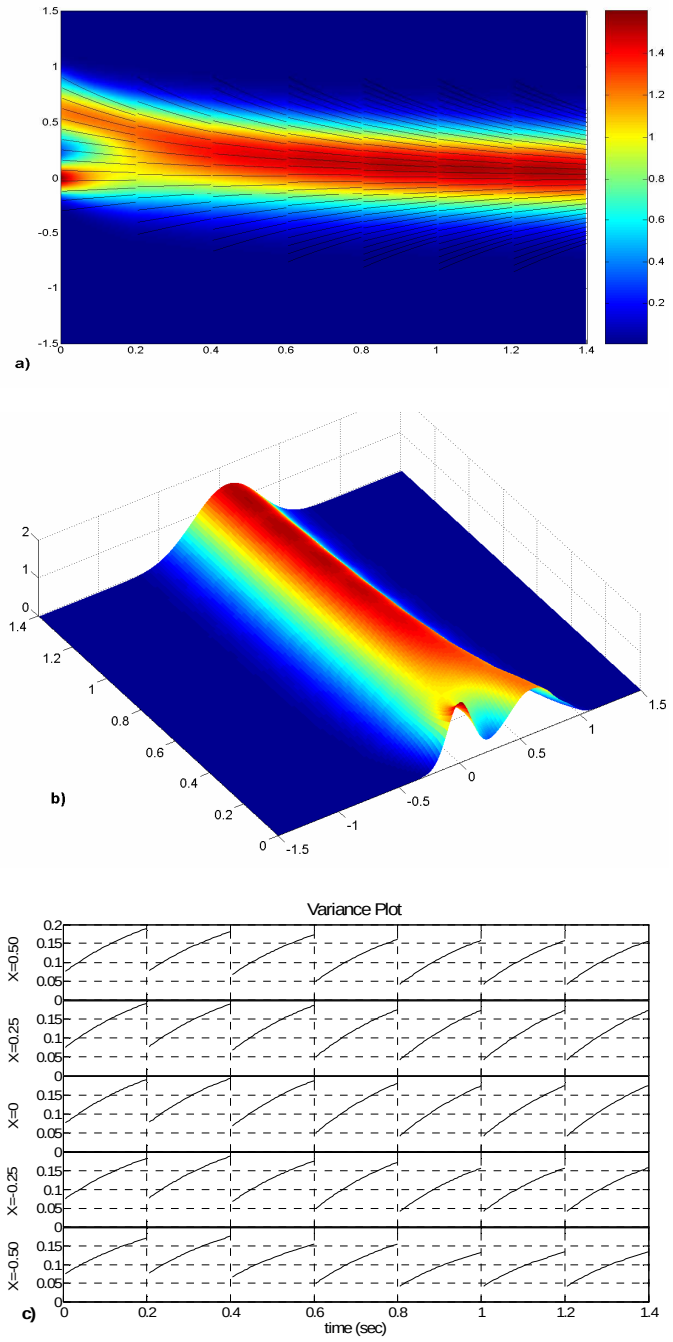

Figure 2: a) Response pdf $f_{x(t)}(x)$ and $m_x^k(t)$ curves for Case I with stationary excitation. b) 3D plot of the response pdf. c) Variance plots for some kdfs.

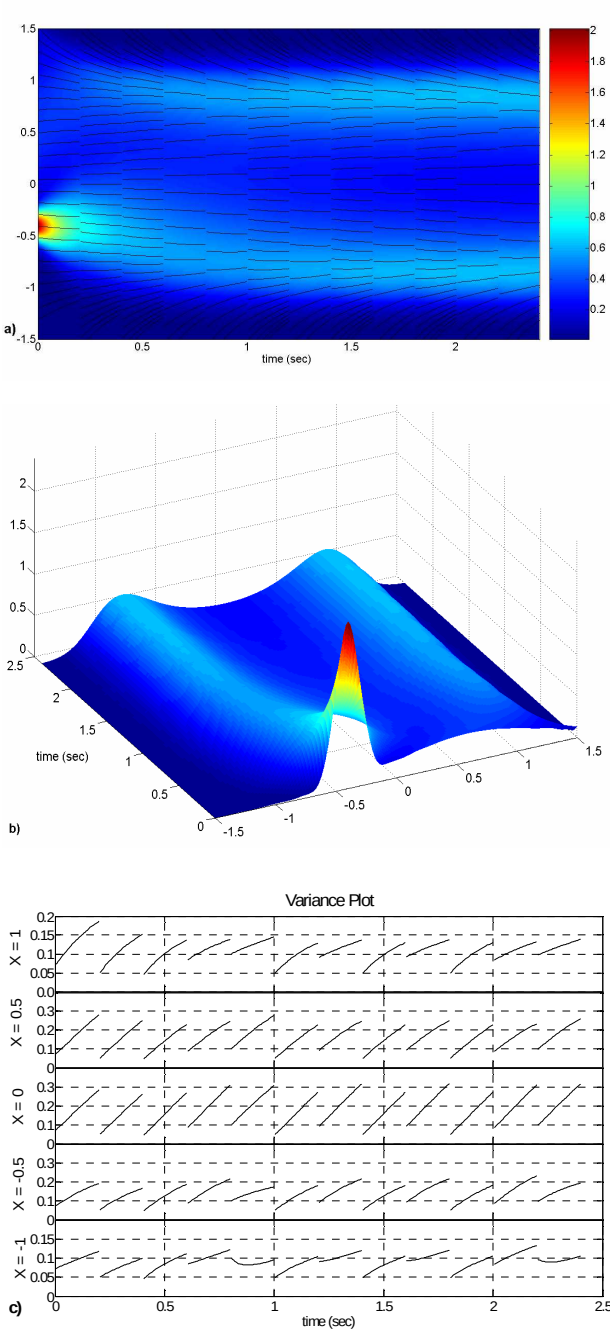

Figure 3: a) Response pdf $f_{x(t)}(x)$ and $m_x^k(t)$ curves for Case II with stationary excitation. b) 3D plot of the response pdf. c) Variance plots for some kdfs.

Let us now consider our system (1.1) with parameter values as in Case II. Two examples of stochastic excitations will be studied. First, the same stationary Gaussian excitation as before, having zero mean and covariance function given by eq (6.1) is applied. The initial distribution is again taken to be bimodal (strongly asymmetric in this case, however), and is defined as a convex superposition of two Gaussian pdfs with parameters $m_1 = -0.4$, $m_2 = 0.6$, $\sigma_1 = 0.1$, $\sigma_2 = 0.6$, and amplitudes $p_1 = 0.4$, $p_2 = 0.6$, respectively.

Numerical results concerning the evolution of the response pdf, for the time interval $0 < t < 2.4\text{sec}$, are presented in Figure 3. Although the initial pdf has taken to be a strongly asymmetric bimodal one, the eventually resulting response density turns to be a symmetric bimodal pdf, with modes exactly at the stable fixed points, located at $\pm\sqrt{k/|a|} = \pm 1$, as expected. The interchange of probability between the kernels (implemented by means of the re-approximation of the response pdf in terms of a new convex superposition of kdfs with smaller variances) takes place approximately every 0.2 seconds. This is

shown in the figure as an apparent discontinuity of the mean-value and variance curves.

From both Figures 2a and 3a (see also Figure 4a, below), we can easily observe a permanent tendency of $m_x^k(t)$-orbits to be attracted by the stable fixed points. This means that there is a continuous inflow of probability mass from the outer region of the phase space ($|x|>1$) to a strip around the locus of the stable fixed points, which is not stopping even after the response pdf has been reached its stationary form. This apparently paradoxical behavior should be addressed to the discrepancy between the tail form of the response pdf $f_{x(t)}(x)$, and the tail form of the Gaussian kernels which are used to represent $f_{x(t)}(x)$. This fact reveals the necessity for an asymptotic study of the tail behavior directly from the differential equation (1.2a′), which will permit the construction and use of the kdfs suitably adapted to the specific system, i.e., exhibiting the correct tail behavior. Such a construction will also facilitate and accelerate the convergence of the numerical solution procedure.

Finally, in Figure 4 we present numerical results for the Case II, with a non-stationary (cyclostationary) Gaussian excitation, with zero mean and covariance function given by

$$C_{YY}(t,s) = \frac{1}{2}\left(1+0.2\cos\left(\frac{\pi t}{2}\right)\right)\cos^2(t-s). \qquad (6.2)$$

Again the initial distribution is constructed as a superposition of two Gaussian pdfs with parameters $m_1=-0.4$, $m_2=0.6$, $\sigma_1=0.3$, $\sigma_2=0.7$, and amplitudes $p_1=0.4$ and $p_2=0.6$, respectively. The evolution of the response probability density function is plotted for the time interval $0<t<8.0\,\text{sec}$, long enough so that the periodic character of the response to become clear.

From Figures 4a, 4b we are able to observe that, after a transient stage ($0<t<1.5\,\text{sec}$), the response density function exhibits a periodic behavior with a period of approximately 4sec, which is the period of the excitation, i.e., the period of the correlation function $C_{YY}(t,s)$, eq. (6.2), with respect to its first argument. Furthermore, it is easily seen that, in this case, a greater amount of kernels is necessary in order to approximate satisfactorily the sought-for pdf, due to the fact that the non-stationary excitation produces a more complicated response.

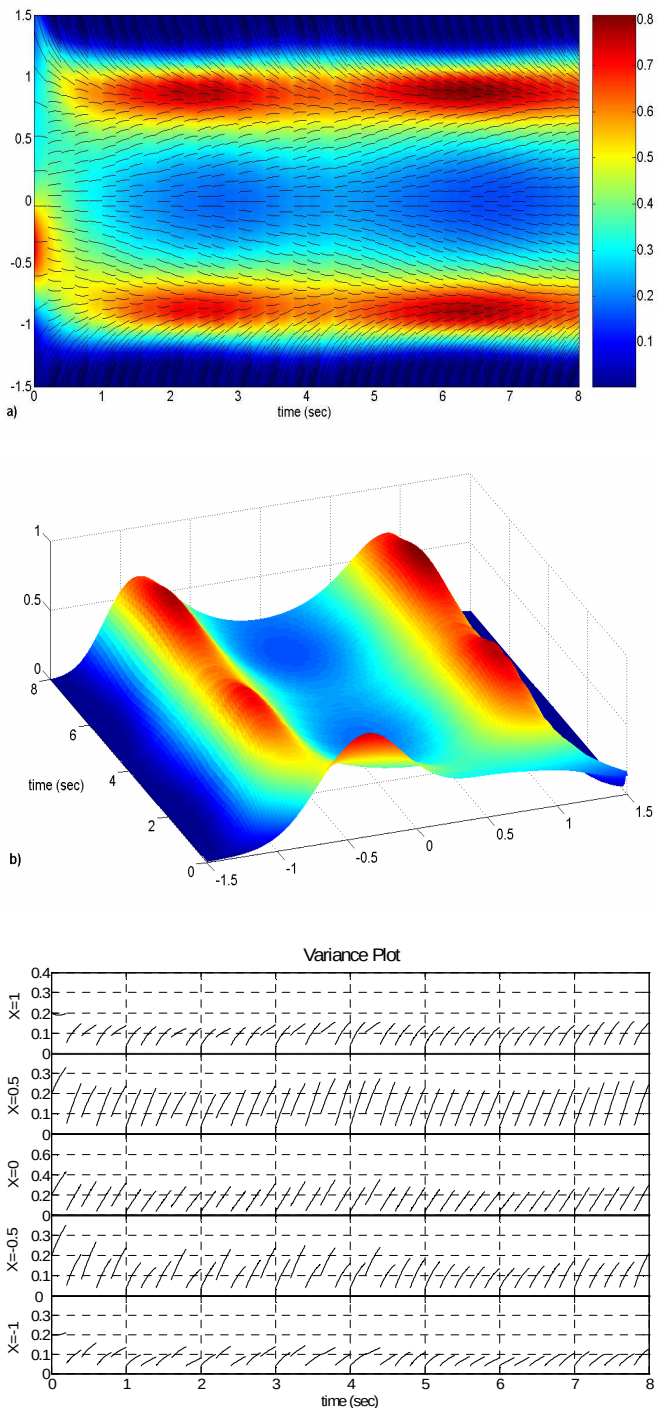

Figure 4: a) Response pdf $f_{x(t)}(x)$ and $m_x^k(t)$ curves for Case II with non-stationary excitation. b) 3D plot of the response pdf. c) Variance plots for some kdfs.

7 CONCLUSIONS

In the present work, a method for the numerical solution of new PDEs (derived in [I]) that govern the evolution of the joint, response-excitation, pdf of nonlinear dynamical systems under general stochastic excitation, is developed and illustrated through its application to a specific, simple, nonlinear system. The key point of the numerical method is the representation of the joint, response-excitation, pdfs and ch.fs by means of appropriate convex superpositions of kernel density and kernel characteristic functions, respectively. In this way, the non-local marginal compatibility condition is satisfied *a priori*, and the PDE governing the evolution of the sought-for pdf

(ch.f) is eventually transformed to a system of nonlinear ODEs for the kernel parameters.

From the results presented in this work we conclude that the proposed method is able to produce quite satisfactory results. Important aspects of the method are (i) It is a *two-level*, *particle-type* method, separating the fast, inner-cycle (short-term) phase, which describes the particle dynamics separately for each particle, from the slow, outer-cycle (long-term) phase, which accounts for the interchange of probability mass between the particles and the evolution of the particles' amplitudes. (ii) It can be improved, keeping its two-level, particle-type character, so that to avoid the piece-wise smoothness assumption for the amplitudes p_j, and to ensure the "exact" satisfaction of the PDE (1.2a′), by solving a linear evolution problem in the outer-cycle phase. (iii) It can be generalized to higher dimensional systems. And (iv) It is plainly suitable for parallelized computations, since the nonlinear ODEs for each particle can be solved independently, and the optimization algorithms is easily parallelizable.

Finally, we emphasize that the method, although in this work has been applied only to problems with Gaussian excitation, can be applied equally well to problems having any kind of external stochastic excitation, characterized by arbitrary –but known– correlation structure and probability distributions.

REFERENCES

Athanassoulis, G.A. & Gavriliadis, P.N., 2002, The Truncated Hausdorff Problem solved by using Kernel Density Functions. Probabilistic Engineering Mechanics, 17, 273-291.

Athanassoulis, G.A. & Gavriliadis, P.N., 2006, The Truncated Hamburger Problem solved by using Kernel Density Functions. (to appear).

Athanassoulis, G.A. & Sapsis, Th.P., 2006, New Partial Differential Equations Governing the Joint, Response-Excitation, Probability Distributions of Nonlinear Systems, under General Stochastic Excitation. I: Derivation. (*Presented in this Conference*)

Bellman, R., 1973, Methods of Nonlinear Analysis, Vol. II, Academic Press, New York.

Di Paola, M. & Sofi, A., 2002, Approximate Solution of the Fokker-Planck-Kolmogorov Equation. Probabilistic Engineering Mechanics, 17, 369-384.

Gavriliadis, P.N., 2005, Theoretical and Numerical Exploitation of the moment problem with applications to the probabilistic prediction of stochastic responces of dynamical systems. PhD Thesis, NTUA.

Härdle ,W., 1990, Smoothing techniques, Springer-Verlag.

Lakshmikantham, V., and Malek, S., 1994, Generalized quasilinearization, Nonlinear World **1**, 59-63.

Lawson C.L., Hanson R.J.,1974, Solving least squares problems. Englewood Cliffs, NJ: Prentice-Hall;

Polidori, D.C. & Beck, J.L. & Papadimitriou, C., 2000, A New Stationary PDF Approximation for Non-Linear Oscillators. International Journal of Non-Linear Mechanics, 35, 657-673.

Pradlwarter, H.J., 2001, Non-Linear Stochastic Response Distributions by Local Statistical Linearization. International Journal of Non-Linear Mechanics, 36, 1135-1151.

Roberts J.B. & Spanos P.D., 2003, Random Vibration and Statistical Linearization. Dover Publications.

Sapsis, Th.P., 2005, Stochastic Analysis with Applications to Dynamical Systems. Diploma Thesis, NTUA.

Scott, D.W., 1992, Multivariate Density Estimation, John Wiley and Sons

Sobczyk, K., 1991, Stochastic Differential Equations. Kluwer Academic Publishers.

Sofi, A., Di Paola, M. & Spanos, P., 2006, A method for determining random response envelope statistics of a class of nonlinear oscillators, *this Conference*.

APPENDIX

In this Appendix a brief outline is given of the optimization algorithm which is used for the construction of appropriate approximants of the sought-for pdf in terms of kdfs exhibiting a small variance (either given or under specific control). This algorithm is used quite often throughout the numerical solution, i.e. each time the solution procedure switches from the inner-cycle to the outer-cycle and the calculated density is re-approximated by means of kdfs of small variance. It is also used for implementing the initialization, by representing the given initial pdf as a convex superposition of appropriate kdfs.

The basic optimization problem is formulated as follows:

Given $f(x)$ and σ_0, find M and $\{p_k, m_k\}_{k=1}^{M}$ such that

$$\int_{-\infty}^{+\infty}\left[f(x)-\sum_{k=1}^{M}\frac{p_k}{\sqrt{2\pi}\sigma_0}\exp\left\{-\frac{1}{2}\left(\frac{x-m_k}{\sigma_0}\right)^2\right\}\right]^2 dx \;=\; \min. \quad \text{(A.1)}$$

under the constraints:

$$p_1 + p_2 + \ldots + p_M = 1, \quad p_k \geq 0, \text{ for all } k.$$

For the inner-cycle/outer-cycle re-approximation of the sought-for pdf, the integrations can be carried out analytically leading to an explicit linear optimization problem, if M is defined. M is obtained by using a variant of an iterative, adaptive procedure, developed by Gavriliadis (2005).

For the initial data representation, the optimization procedure is performed quite similarly. However, in this case, the integrations in (A.1) are performed numerically, since, in general, the initial probability distribution may not be analytically described. A detailed description of the solution algorithm will be presented elsewhere.