



**HAL**  
open science

# Construction of probability distributions in high dimension using the maximum entropy principle: Applications to stochastic processes, random fields and random matrices

Christian Soize

► **To cite this version:**

Christian Soize. Construction of probability distributions in high dimension using the maximum entropy principle: Applications to stochastic processes, random fields and random matrices. International Journal for Numerical Methods in Engineering, 2008, 76 (10), pp.1583-1611. 10.1002/nme.2385 . hal-00684517

**HAL Id: hal-00684517**

**<https://hal.science/hal-00684517>**

Submitted on 2 Apr 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Construction of probability distributions in high dimension using the maximum entropy principle. Applications to stochastic processes, random fields and random matrices

Christian Soize<sup>1</sup>

<sup>1</sup>*Université Paris-Est, Laboratoire Modélisation et Simulation Multi Echelle, MSME FRE3160 CNRS, 5 Bd Descartes, 77454 Marne la Vallée, France*

## SUMMARY

The construction of probabilistic models in computational mechanics requires the effective construction of probability distributions of random variables in high dimension. This paper deals with the effective construction of the probability distribution in high dimension of a vector-valued random variable using the maximum entropy principle. The integrals in high dimension are then calculated in constructing the stationary solution of an Itô stochastic differential equation associated with its invariant measure. A random generator of independent realizations is explicitly constructed in the paper. Three fundamental applications are presented. The first one is a new formulation of the stochastic inverse problem relative to the construction of the probability distribution in high dimension of an unknown non-stationary random time series (random accelerograms) for which the Velocity Response Spectrum is given. The second one is also a new formulation related to the construction of the probability distribution of positive-definite band random matrices. Finally, we present an extension of the theory when the support of the probability distribution is not all the space but is any part of the space. The third application is then a new formulation related to the construction of the probability distribution of the Karhunen-Loeve expansion of Non-Gaussian positive-valued random fields. Copyright © 2008 John Wiley & Sons, Ltd.

KEY WORDS: Maximum entropy principle; High dimension; Stochastic process; Random matrix, Karhunen-Loeve expansion, Random fields

## 1. Introduction

The probabilistic modeling of uncertainties in computational sciences such as in computational mechanics is a great challenge. For instance, the parametric probabilistic approach which allows data uncertainties to be taken into account consists in modeling uncertain parameters of the computational model by random variables, stochastic processes and random fields. Such approaches have extensively been developed in the two last decades [1, 2]. In particular, the use

---

\*Correspondence to: Christian Soize, Université Paris-Est, Laboratoire Modélisation et Simulation Multi Echelle, MSME FRE3160 CNRS, 5 Bd Descartes, 77454 Marne la Vallée, France

of the Gaussian Chaos representation for stochastic processes and random fields [3] has been used to introduce and to develop useful and very efficient tools for analyzing stochastic systems using stochastic finite elements (see [4, 5, 6]). More recently, additional developments have been proposed to construct Chaos representations with arbitrary probability measures [7]. Model uncertainties introduced by the mathematical-mechanical process used during the construction of the computational model of complex systems are much more difficult to take into account because the parametric probabilistic approach cannot address such model uncertainties. In this context, a nonparametric probabilistic approach of model uncertainties has recently been proposed as a possible way to circumvent these difficulties [8, 9, 10] and is based on the use of the random matrix theory.

In general, the response of a computational model is a nonlinear mapping of the uncertain parameters and consequently, a complete probability model of these uncertain parameters has to be constructed. This means that the probability distributions of the random quantities of interest such as vector-valued random variables, random matrices, etc, have to be constructed and random generators of independent realizations have to be derived from the knowledge of the probability distributions. It is well known that the Maximum Entropy (MaxEnt) Principle [11, 12] is certainly one of the most efficient method allowing an explicit construction of such probability distributions to be performed using only the available information. This powerful method developed by Shannon in the context of the Information Theory is extremely useful in many situations for which statistical data related to the random variable of interest are either partially available or not available at all. We are then considering the case for which the number of independent realizations (which constitute the available statistical data obtained from measurements) is too small to obtain a good convergence of the statistical estimator of the probability density function using nonparametric statistics [13]. The MaxEnt principle has been used for different cases and in many applications (see for instance: [14] for a simple overview concerning the MaxEnt Principle with applications, [15] for advanced developments in Physics, [16] for the use of the MaxEnt principle in stochastic dynamics). It should be noted that this paper does not deal with the maximum entropy in the moment problem [17, 18, 19, 20] but is devoted to general nonlinear constraints in high dimension which are not polynomials, that is to say, which are not expressed as a linear combination of moments.

Let us consider a  $\mathbb{R}^N$ -valued random variable  $\mathbf{A} = (A_1, \dots, A_N)$  in which  $N$  is large (high dimension). For instance, the set  $\{A_1, \dots, A_N\}$  can represent a random time series constructed from the time sampling  $\mathbb{A}(t_1), \dots, \mathbb{A}(t_N)$  of a stochastic process  $\{\mathbb{A}(t), t \in T\}$ . Random vector  $\mathbf{A}$  can also be used to generate a sparse random matrix  $[\mathbf{G}]$  whose non zero random elements are expressed in function of  $\mathbf{A}$ . Finally,  $\mathbf{A}$  can be the vector constituted of the random coordinates in the Karhunen-Loeve expansion of a random field. The use of the MaxEnt principle allows the probability density function  $\mathbf{a} \mapsto p_{\mathbf{A}}(\mathbf{a})$  on  $\mathbb{R}^N$  (with respect to the Lebesgue measure  $d\mathbf{a}$ ) of random variable  $\mathbf{A}$  to be constructed using the available information. Introducing the Lagrange multiplier  $\boldsymbol{\lambda} \in \mathcal{L}_{\mu} \subset \mathbb{R}^{\mu}$  associated with the constraints defined by the available information, in which  $\mathcal{L}_{\mu}$  is the subset of  $\mathbb{R}^{\mu}$  of all the admissible values of  $\boldsymbol{\lambda}$ , it can be proven that  $p_{\mathbf{A}}$  depends on  $\boldsymbol{\lambda}$ . Lagrange multiplier  $\boldsymbol{\lambda}$  is then calculated solving a nonlinear algebraic equation or equivalently, solving an optimization problem for a convex cost function. For each given value of  $\boldsymbol{\lambda}$  in  $\mathcal{L}_{\mu} \subset \mathbb{R}^{\mu}$ , an integral  $\mathbf{i}_N(\boldsymbol{\lambda})$  on  $\mathbb{R}^N$  must be computed. In high dimension, such a calculation is very difficult to perform and in general, induces a high numerical cost. Thus there are two main problems. The first one is related to the effective calculation of Lagrange multiplier  $\boldsymbol{\lambda}$  using an adated algorithm which requires the evaluation

of a large number of integral  $\mathbf{i}_N(\boldsymbol{\lambda})$  in high dimension. The second one is to construct a random generator of independent realizations of random variable  $\mathbf{A}$  whose  $p_{\mathbf{A}}$  has been constructed with the MaxEnt principle. It should be noted that these two problems are not trivial in high dimension. In this paper, we propose a method to solve these two main problems. Integral  $\mathbf{i}_N(\boldsymbol{\lambda})$  is calculated in constructing the stationary solution associated with the invariant measure of a nonlinear Itô stochastic differential equation (ISDE) depending on  $\boldsymbol{\lambda}$ . This proposed method is an alternative to the Metropolis-Hastings algorithm or to the Gibbs sampling (see Section 4). Such a nonlinear ISDE is solved by using either an explicit Euler scheme or a semi-implicit scheme. Integral  $\mathbf{i}_N(\boldsymbol{\lambda})$  is then calculated using (1) either the ergodic method or (2) the Monte Carlo method with the usual estimator of the mathematical expectation with  $n_s$  independent realizations constructed with a generator. This generator of independent realizations of  $\mathbf{A}$  is then constructed solving  $n_s$  times the ISDE with  $n_s$  independent Wiener stochastic processes. Three fundamental applications are presented. The first one is a new formulation of the stochastic inverse problem relative to the construction of the probability distribution in high dimension and of its generator for a vector-valued random variable corresponding to an unknown non-stationary random time series (random accelerograms) for which the Velocity Response Spectrum is given (see Section 7). The second one is also a new formulation related to the construction of the probability distribution in high dimension and of its generator for positive-definite band random matrices (see Section 8). Finally, we present an extension of the theory corresponding to the case for which the support of the probability distribution in high dimension of random variable  $\mathbf{A}$  is not  $\mathbb{R}^N$  but is any part  $\mathcal{A}$  of  $\mathbb{R}^N$  (see Section 9). The third application is then a new formulation related to the construction of the probability distribution in high dimension for the Karhunen-Loeve expansion of Non-Gaussian positive-valued random fields.

## 2. Construction of probability distributions using the maximum entropy principle

Let  $\mathbf{a} = (a_1, \dots, a_N)$  be any vector in  $\mathbb{R}^N$ . Let  $\mathbf{A} = (A_1, \dots, A_N)$  be a  $\mathbb{R}^N$ -valued second-order random variable for which the probability distribution  $P_{\mathbf{A}}(d\mathbf{a})$  on  $\mathbb{R}^N$  is unknown but is represented by a probability density function  $\mathbf{a} \mapsto p_{\mathbf{A}}(\mathbf{a})$  from  $\mathbb{R}^N$  into  $\mathbb{R}^+ = [0, +\infty[$  with respect to the Lebesgue measure  $d\mathbf{a} = da_1 \dots da_N$  and which has to verify the following normalization condition,

$$\int_{\mathbb{R}^N} p_{\mathbf{A}}(\mathbf{a}) d\mathbf{a} = 1 \quad . \quad (1)$$

Presently, it is assumed that the support of the probability density function  $p_{\mathbf{A}}$  is  $\mathbb{R}^N$ . The case for which the support of  $p_{\mathbf{A}}$  is any part  $\mathcal{A}$  of  $\mathbb{R}^N$  will be treated in Section 9.

The problem to be solved is the construction of the unknown probability density function  $p_{\mathbf{A}}$  by using the MaxEnt principle for which the constraints associated with the available information are assumed to be defined by the following equation on  $\mathbb{R}^{\mu}$ ,

$$E\{\mathbf{g}(\mathbf{A})\} = \mathbf{f} \quad , \quad (2)$$

in which  $\mathbf{f} = (f_1, \dots, f_{\mu})$  is a given vector in  $\mathbb{R}^{\mu}$  with  $\mu \geq 1$ , where  $\mathbf{a} \mapsto \mathbf{g}(\mathbf{a}) = (g_1(\mathbf{a}), \dots, g_{\mu}(\mathbf{a}))$  is a given measurable mapping from  $\mathbb{R}^N$  into  $\mathbb{R}^{\mu}$  and where  $E$  is the

mathematical expectation. Equation (2) can then be rewritten as

$$\int_{\mathbb{R}^N} \mathbf{g}(\mathbf{a}) p_{\mathbf{A}}(\mathbf{a}) d\mathbf{a} = \mathbf{f} \quad . \quad (3)$$

Let  $\mathcal{C}$  be the set of all the probability density functions  $\mathbf{a} \mapsto p_{\mathbf{A}}(\mathbf{a})$  defined on  $\mathbb{R}^N$  with values in  $\mathbb{R}^+$  such that Eqs. (1) and (3) hold. The maximum entropy principle [11, 12] consists in constructing  $p_{\mathbf{A}} \in \mathcal{C}$  such that

$$p_{\mathbf{A}} = \arg \max_{p \in \mathcal{C}} S(p) \quad , \quad (4)$$

in which the entropy  $S(p)$  of probability density function  $p$  is defined by

$$S(p) = - \int_{\mathbb{R}^N} p(\mathbf{a}) \log(p(\mathbf{a})) d\mathbf{a} \quad , \quad (5)$$

where  $\log$  is the Neperian logarithm. In order to solve the optimization problem defined by Eq. (4), a Lagrange multiplier  $\lambda_0 \in \mathbb{R}^+$  associated with the constraint defined by Eq. (1) and a Lagrange multiplier  $\boldsymbol{\lambda} \in \mathcal{L}_\mu \subset \mathbb{R}^\mu$  associated with the constraint defined by Eq. (3) are introduced, in which  $\mathcal{L}_\mu$  is the subset of  $\mathbb{R}^\mu$  of all the admissible values of  $\boldsymbol{\lambda}$ . It can then be proven that the solution of Eq. (4) can be written as

$$p_{\mathbf{A}}(\mathbf{a}) = c_0^{\text{sol}} \exp(- \langle \boldsymbol{\lambda}^{\text{sol}}, \mathbf{g}(\mathbf{a}) \rangle_\mu) \quad , \quad \forall \mathbf{a} \in \mathbb{R}^N \quad , \quad (6)$$

with  $c_0^{\text{sol}} = \exp(-\lambda_0^{\text{sol}})$  in which  $(\lambda_0^{\text{sol}}, \boldsymbol{\lambda}^{\text{sol}}) \in \mathbb{R}^+ \times \mathcal{L}_\mu$  is such that Eqs. (1) and (3) are verified. In Eq. (6),  $\langle \mathbf{x}, \mathbf{y} \rangle_\mu = x_1 y_1 + \dots + x_\mu y_\mu$  is the Euclidean inner product on  $\mathbb{R}^\mu$ .

For  $\boldsymbol{\lambda}$  fixed in  $\mathcal{L}_\mu$ , let  $\mathbf{B}_\lambda$  be the  $\mathbb{R}^N$ -valued random variable whose probability density function  $\mathbf{b} \mapsto p(\mathbf{b}, \boldsymbol{\lambda})$  from  $\mathbb{R}^N$  into  $\mathbb{R}^+$  (with respect to the Lebesgue measure  $d\mathbf{b}$  on  $\mathbb{R}^N$ ) is written as

$$p(\mathbf{b}, \boldsymbol{\lambda}) = c_\lambda \exp(- \langle \boldsymbol{\lambda}, \mathbf{g}(\mathbf{b}) \rangle_\mu) \quad , \quad \forall \mathbf{b} \in \mathbb{R}^N \quad , \quad (7)$$

in which  $c_\lambda$  is a finite positive constant depending on  $\boldsymbol{\lambda}$  defined by the following normalization condition

$$\int_{\mathbb{R}^N} p(\mathbf{b}, \boldsymbol{\lambda}) d\mathbf{b} = 1 \quad . \quad (8)$$

Taking  $c_{\boldsymbol{\lambda}^{\text{sol}}} = c_0^{\text{sol}}$ , Eqs. (6) and (7) yield

$$p_{\mathbf{A}}(\mathbf{a}) = p(\mathbf{a}, \boldsymbol{\lambda}^{\text{sol}}) \quad , \quad \forall \mathbf{a} \in \mathbb{R}^N \quad , \quad . \quad (9)$$

which means that we have the following equality  $\mathbf{A} = \mathbf{B}_{\boldsymbol{\lambda}^{\text{sol}}}$  of random variables for the convergence in probability distribution. From Eqs. (3), (6), (7) and (9), it can then be deduced that  $\boldsymbol{\lambda}^{\text{sol}}$  is a solution of the following equation in  $\boldsymbol{\lambda}$ ,

$$E\{\mathbf{g}(\mathbf{B}_\lambda)\} = \mathbf{f} \quad , \quad (10)$$

in which the integral  $E\{\mathbf{g}(\mathbf{B}_\lambda)\}$  which depends on  $\boldsymbol{\lambda}$  is such that

$$E\{\mathbf{g}(\mathbf{B}_\lambda)\} = \int_{\mathbb{R}^N} \mathbf{g}(\mathbf{b}) p(\mathbf{b}, \boldsymbol{\lambda}) d\mathbf{b} \quad . \quad (11)$$

We must then construct a solution  $\boldsymbol{\lambda}^{\text{sol}}$  in  $\mathcal{L}_\mu \subset \mathbb{R}^\mu$  of Eq. (10) in  $\boldsymbol{\lambda}$ . By construction, the constraints associated with the available information (see Eq. (3)) are such that the algebraic

equation in  $\boldsymbol{\lambda}$  (defined by Eq. (10)) admits a unique solution in  $\mathcal{L}_\mu \subset \mathbb{R}^\mu$  (it should be noted that, if it was not the case, it would mean that the available information defined was not consistent and consequently, should be re-examined and then modified). We will denote any one of this by  $\boldsymbol{\lambda}^{\text{sol}} \in \mathbb{R}^\mu$ . Consequently, for such a solution, Eqs. (1) and (3) are verified and the probability density function  $p_{\mathbf{A}}$  is given by Eq. (6) with  $c_0^{\text{sol}} = c_{\boldsymbol{\lambda}^{\text{sol}}}$ . Equation (10) can be solved in  $\boldsymbol{\lambda}$  with an appropriate algorithm such that the interior-reflective Newton method or the trust-region dogleg algorithm which is a variant of the Powell dogleg method described in [21, 22] (as used in Matlab for large-scale or medium-scale algorithm). It should be noted that  $\boldsymbol{\lambda}^{\text{sol}}$  could also be calculated in solving a convex optimization problem but the experience proves that there is no numerical gain with respect to the previous one.

### 3. Difficulties of the construction in high dimension

The vector-valued Lagrange multiplier  $\boldsymbol{\lambda}^{\text{sol}}$  must be computed in solving Eq. (10) which requires to evaluate the integral on  $\mathbb{R}^N$  defined by Eq. (11). For the high-dimension case, that is to say for a large value of  $N$ , this problem is very difficult. Below, for a given value of  $\boldsymbol{\lambda}$  in  $\mathcal{L}_\mu \subset \mathbb{R}^\mu$ , we present some explanations concerning the difficulties of the calculation of  $E\{\mathbf{g}(\mathbf{B}_\lambda)\}$ .

For  $\boldsymbol{\lambda}$  fixed in  $\mathcal{L}_\mu \subset \mathbb{R}^\mu$ , there exist methods to perform the evaluation of the integral  $E\{\mathbf{g}(\mathbf{B}_\lambda)\}$  with respect to the probability distribution  $p(\mathbf{b}, \boldsymbol{\lambda}) d\mathbf{b}$  in high dimension (see for instance [23]). The first class of methods corresponds to the exact evaluation of the integral using analytical calculation (for instance using integration on Gaussian spaces). It should be noted that integral  $E\{\mathbf{g}(\mathbf{B}_\lambda)\}$  defined by Eq. (11) cannot exactly be calculated and consequently, an approximation must be carried out in order to evaluate it. The second class of methods corresponds to approximate methods. For the high-dimension case, an usual and efficient method consists in using the Monte Carlo method (see for instance [23, 24, 25]).

**Formulation 1.** A first method consists in rewriting  $E\{\mathbf{g}(\mathbf{B}_\lambda)\}$  as the mathematical expectation  $E\{\mathbf{h}(\mathbf{Z}, \boldsymbol{\lambda})\}$  of a  $\mathbb{R}^\mu$ -valued random variable  $\mathbf{h}(\mathbf{Z}, \boldsymbol{\lambda})$  in which  $\mathbf{z} \mapsto \mathbf{h}(\mathbf{z}, \boldsymbol{\lambda})$  is a given measurable mapping from  $\mathbb{R}^N$  into  $\mathbb{R}^\mu$  and where  $\mathbf{Z}$  is a given  $\mathbb{R}^N$ -valued random variable (1) whose probability distribution  $p_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z}$  on  $\mathbb{R}^N$  is known and (2) for which a random generator can easily be constructed. We can then write  $E\{\mathbf{g}(\mathbf{B}_\lambda)\} = E\{\mathbf{h}(\mathbf{Z}, \boldsymbol{\lambda})\}$  and the Monte Carlo method consists in evaluating  $E\{\mathbf{g}(\mathbf{B}_\lambda)\}$  by

$$E\{\mathbf{g}(\mathbf{B}_\lambda)\} \simeq \frac{1}{n_s} \sum_{\ell=1}^{n_s} \mathbf{h}(\mathbf{Z}(\theta_\ell), \boldsymbol{\lambda}) \quad , \quad (12)$$

in which  $\mathbf{Z}(\theta_1), \dots, \mathbf{Z}(\theta_{n_s})$  are  $n_s$  independent realizations of random variable  $\mathbf{Z}$  whose probability distribution is  $p_{\mathbf{Z}}(\mathbf{z}) d\mathbf{z}$ . Clearly, such a computation can be carried out only if a random generator can easily be constructed. For instance, let us assume that, for admissible values of  $\boldsymbol{\lambda}$ , the expression  $\exp(-\langle \boldsymbol{\lambda}, \mathbf{g}(\mathbf{z}) \rangle_\mu + \|\mathbf{z}\|_N^2/2)$  tends to zero when  $\|\mathbf{z}\|_N$  goes to infinity in which  $\|\mathbf{z}\|_N^2 = z_1^2 + \dots + z_N^2$ . Choosing  $\mathbf{Z}$  as a normalized Gaussian  $\mathbb{R}^N$ -valued random variable, we have  $p_{\mathbf{Z}}(\mathbf{z}) = (2\pi)^{-N/2} \exp(-\|\mathbf{z}\|_N^2/2)$  and  $\mathbf{h}(\mathbf{z}, \boldsymbol{\lambda}) = (2\pi)^{N/2} \mathbf{g}(\mathbf{z}) c_\lambda \exp(-\langle \boldsymbol{\lambda}, \mathbf{g}(\mathbf{z}) \rangle_\mu + \|\mathbf{z}\|_N^2/2)$ . If  $p$  defined by Eq. (7) "is not close" to  $p_{\mathbf{Z}}$ , then the speed of convergence of the right-hand side of Eq. (12) can be extremely low with respect to  $n_s$  and consequently, is not efficient at all. This situation corresponds to the case for which (1) the realizations  $\mathbf{Z}(\theta_\ell)$  of  $\mathbf{Z}$  yield very small contributions in the right-hand side of Eq. (12) and (2) the significant

contributions correspond to events having a very small probability due to the presence of the term  $\exp(-\langle \boldsymbol{\lambda}, \mathbf{g}(\mathbf{z}) \rangle_\mu + \|\mathbf{z}\|_N^2/2)$  in  $\mathbf{h}(\mathbf{z}, \boldsymbol{\lambda})$ .

**Formulation 2.** A second method consists in using the Monte Carlo method to evaluate  $E\{\mathbf{g}(\mathbf{B}_\lambda)\}$  by

$$E\{\mathbf{g}(\mathbf{B}_\lambda)\} \simeq \frac{1}{n_s} \sum_{\ell=1}^{n_s} \mathbf{g}(\mathbf{B}_\lambda(\theta_\ell)) \quad , \quad (13)$$

in which  $\mathbf{B}_\lambda(\theta_1), \dots, \mathbf{B}_\lambda(\theta_{n_s})$  are  $n_s$  independent realizations of random variable  $\mathbf{B}_\lambda$  whose probability distribution is  $p(\mathbf{b}, \boldsymbol{\lambda}) d\mathbf{b}$ . Consequently, a random generator must be constructed and this difficult problem is the subject of this paper.

Considering that Formulation 1 cannot be used for the reasons given above, Formulation 2 will be used and the objectives of this paper are to propose (1) a methodology to calculate Lagrange multiplier  $\boldsymbol{\lambda}^{\text{sol}} \in \mathcal{L}_\mu \subset \mathbb{R}^\mu$  and (2) the construction of a random generator of independent realizations for  $\mathbb{R}^N$ -valued random variable  $\mathbf{A}$ .

#### 4. Construction of probability distributions in high dimension using stochastic analysis

For  $\boldsymbol{\lambda}$  fixed in  $\mathcal{L}_\mu \subset \mathbb{R}^\mu$ , the evaluation of  $E\{\mathbf{g}(\mathbf{B}_\lambda)\}$  defined by Eq. (11) can be performed using the Markov Chain Monte Carlo method (MCMC) [26, 27, 24] which is an alternative efficient approach to Formulation 1 presented in Section 3. The transition kernel of the homogeneous Markov chain of the MCMC method can be constructed using the Metropolis-Hastings algorithm [25] or the Gibbs sampling [28] which is a slightly different algorithm for which the kernel is directly deduced from the probability density function and for which the Gibbs samplers are always accepted. These two algorithms allow the transition kernel to be constructed for which the invariant measure is  $p(\mathbf{b}, \boldsymbol{\lambda}) d\mathbf{b}$ . In general, these two algorithms are efficient, but can also be not efficient if there exists attraction regions which do not correspond to the invariant measure under consideration. These cases cannot be easily detected and are time consuming. The method presented below looks like to the Gibbs approach but corresponds to a more direct construction of a random generator of realizations of random variable  $\mathbf{B}_\lambda$  whose probability distribution is  $p(\mathbf{b}, \boldsymbol{\lambda}) d\mathbf{b}$ . The difference between the Gibbs algorithm and the proposed algorithm is that the convergence in the proposed method can be studied with all the mathematical results concerning the existence and uniqueness of Itô stochastic differential equation. In addition, a parameter is introduced which allows the transient part of the response to be killed in order to get more rapidly the stationary solution corresponding to the invariant measure. Thus following [29], the construction of the transition kernel by using the detailed balance equation is replaced by the construction of an Itô Stochastic Differential Equation (ISDE) (depending on  $\boldsymbol{\lambda}$ ) which admits  $p(\mathbf{b}, \boldsymbol{\lambda}) d\mathbf{b}$  defined by Eq. (7) as a unique invariant measure. In addition, either the ergodic method (that we will present below) or the Monte Carlo method (see Eq. (13)) can be used to estimate  $E\{\mathbf{g}(\mathbf{B}_\lambda)\}$  in order to calculate  $\boldsymbol{\lambda}^{\text{sol}}$ .

##### 4.1. Construction of the probability distribution of $\mathbf{B}_\lambda$ as the invariant measure of an ISDE

For  $\boldsymbol{\lambda}$  fixed in  $\mathcal{L}_\mu \subset \mathbb{R}^\mu$ , let  $\mathbf{u} \mapsto \Phi(\mathbf{u}, \boldsymbol{\lambda})$  be the function from  $\mathbb{R}^N$  into  $\mathbb{R}$  defined by

$$\Phi(\mathbf{u}, \boldsymbol{\lambda}) = \langle \boldsymbol{\lambda}, \mathbf{g}(\mathbf{u}) \rangle_\mu \quad . \quad (14)$$

Let  $\{(\mathbf{U}(r), \mathbf{V}(r)), r \in \mathbb{R}^+\}$  be the Markov stochastic process defined on the probability space  $(\Theta, \mathcal{T}, \mathcal{P})$  indexed by  $\mathbb{R}^+ = [0, +\infty[$  with values in  $\mathbb{R}^N \times \mathbb{R}^N$  satisfying, for all  $r > 0$ , the following Itô stochastic differential equation

$$d\mathbf{U}(r) = \mathbf{V}(r) dr \quad , \quad (15)$$

$$d\mathbf{V}(r) = -\nabla_{\mathbf{u}}\Phi(\mathbf{U}(r), \boldsymbol{\lambda}) dr - \frac{1}{2}f_0\mathbf{V}(r) dr + \sqrt{f_0}d\mathbf{W}(r) \quad , \quad (16)$$

with the initial condition

$$\mathbf{U}(0) = \mathbf{U}_0 \quad , \quad \mathbf{V}(0) = \mathbf{V}_0 \quad a.s. \quad , \quad (17)$$

in which  $f_0$  is a free parameter which has to be fixed to any positive value ( $f_0 > 0$ ), where  $\mathbf{W} = (W_1, \dots, W_N)$  is the normalized Wiener process defined on  $(\Theta, \mathcal{T}, \mathcal{P})$  indexed by  $\mathbb{R}^+$  with values in  $\mathbb{R}^N$  and where the random initial condition  $(\mathbf{U}_0, \mathbf{V}_0)$  is a  $\mathbb{R}^N \times \mathbb{R}^N$ -valued second-order random variable independent of the family of random variables  $\{\mathbf{W}(r), r \geq 0\}$ . The probability distribution  $P_{\mathbf{U}_0, \mathbf{V}_0}(d\mathbf{u}, d\mathbf{v})$  on  $\mathbb{R}^N \times \mathbb{R}^N$  of random variable  $(\mathbf{U}_0, \mathbf{V}_0)$  is assumed to be given. The matrix-valued autocorrelation function  $[R_{\mathbf{W}}(r, r')] = E\{\mathbf{W}(r)\mathbf{W}(r')^T\}$  of  $\mathbf{W}$  is then written as  $[R_{\mathbf{W}}(r, r')] = \min(r, r')[I_N]$  with  $[I_N]$  the identity ( $N \times N$ ) matrix. In Eq. (16), the free parameter  $f_0 > 0$  will allow a dissipation term to be introduced in the nonlinear dynamical system in order to kill the transient part of the response and consequently, to get more rapidly the stationary solution corresponding to the invariant measure (see the end of Subsection 4.1).

In a first stage, for an admissible value of  $\boldsymbol{\lambda}$  fixed in  $\mathcal{L}_\mu \subset \mathbb{R}^\mu$ , it is assumed that the problem defined by Eqs. (15) to (17) has a unique solution defined almost surely for all  $r \geq 0$  (no explosion of the solution, see for instance Theorems 4 and 5 in pages 154 to 157 of Ref. [30]) which is a diffusion stochastic process with drift vector  $\mathbf{b}(\mathbf{u}, \mathbf{v}) \in \mathbb{R}^{2N}$  and diffusion matrix  $[\sigma] \in \mathbb{M}_{2N}(\mathbb{R})$  such that

$$\mathbf{b}(\mathbf{u}, \mathbf{v}) = \begin{bmatrix} \mathbf{v} \\ -\nabla_{\mathbf{u}}\Phi(\mathbf{u}, \boldsymbol{\lambda}) - \frac{1}{2}f_0\mathbf{v} \end{bmatrix} \quad , \quad [\sigma] = \begin{bmatrix} 0_N & 0_N \\ 0_N & f_0I_N \end{bmatrix} \quad , \quad (18)$$

in which  $[0_N]$  is the zero ( $N \times N$ ) matrix,  $[I_N]$  is the identity ( $N \times N$ ) matrix and  $\mathbb{M}_{2N}(\mathbb{R})$  is the set of all the square ( $2N \times 2N$ ) real matrices. Let  $0 \leq s < r < +\infty$ ,  $\mathbf{u}$  and  $\mathbf{v}$  in  $\mathbb{R}^N$  and let  $B_{\mathbf{u}}$  and  $B_{\mathbf{v}}$  belonging to the Borel  $\sigma$ -algebra of  $\mathbb{R}^N$ . Since the drift vector and the diffusion matrix are independent of  $r$ , the diffusion stochastic process  $\{(\mathbf{U}(r), \mathbf{V}(r)), r \geq 0\}$  admits a system of homogeneous transition probabilities such that

$$P(\mathbf{u}, \mathbf{v}; r - s, B_{\mathbf{u}}, B_{\mathbf{v}}, \boldsymbol{\lambda}) = \mathcal{P}\{\mathbf{U}(r) \in B_{\mathbf{u}}, \mathbf{V}(r) \in B_{\mathbf{v}} | \mathbf{U}(s) = \mathbf{u}, \mathbf{V}(s) = \mathbf{v}\} \quad , \quad (19)$$

in which  $\mathcal{P}\{\mathbf{U}(r) \in B_{\mathbf{u}}, \mathbf{V}(r) \in B_{\mathbf{v}} | \mathbf{U}(s) = \mathbf{u}, \mathbf{V}(s) = \mathbf{v}\}$  is the conditional probability for that  $\mathbf{U}(r) \in B_{\mathbf{u}}$  and  $\mathbf{V}(r) \in B_{\mathbf{v}}$  if  $\mathbf{U}(s) = \mathbf{u}$  and  $\mathbf{V}(s) = \mathbf{v}$ . Let  $P_s(d\mathbf{u}, d\mathbf{v}, \boldsymbol{\lambda})$  be an invariant measure, *i.e.* a probability measure on  $\mathbb{R}^N \times \mathbb{R}^N$  independent of  $r$  which is a solution of the integral equation

$$P_s(d\mathbf{u}, d\mathbf{v}, \boldsymbol{\lambda}) = \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} P_s(d\mathbf{u}', d\mathbf{v}', \boldsymbol{\lambda}) P(\mathbf{u}', \mathbf{v}'; r, d\mathbf{u}, d\mathbf{v}, \boldsymbol{\lambda}) \quad , \quad \forall r > 0 \quad . \quad (20)$$

If the probability distribution  $P_{\mathbf{U}_0, \mathbf{V}_0}(d\mathbf{u}, d\mathbf{v})$  on  $\mathbb{R}^N \times \mathbb{R}^N$  of the second-order random variable  $(\mathbf{U}_0, \mathbf{V}_0)$  is equal to  $P_s(d\mathbf{u}, d\mathbf{v}, \boldsymbol{\lambda})$ , then the unique solution of the problem defined by Eqs. (15)



to (17) is a stationary diffusion stochastic process on  $\mathbb{R}^+$  for the shift semi-group  $r \mapsto r + s$ ,  $s \geq 0$  and is assumed to be ergodic. This stationary and ergodic solution is then denoted by  $\{(\mathbf{U}_s(r), \mathbf{V}_s(r)), r \geq 0\}$ . In addition, for any probability distribution  $P_{\mathbf{U}_0, \mathbf{V}_0}(d\mathbf{u}, d\mathbf{v})$  (not equal to the invariant measure), the diffusion stochastic process  $\{(\mathbf{U}(r), \mathbf{V}(r)), r \geq 0\}$  converges to the stationary diffusion stochastic process  $\{(\mathbf{U}_s(r), \mathbf{V}_s(r)), r \geq 0\}$  when  $r$  goes to infinity. Let us assume that the invariant measure can be written as  $P_s(d\mathbf{u}, d\mathbf{v}, \boldsymbol{\lambda}) = \rho_s(\mathbf{u}, \mathbf{v}, \boldsymbol{\lambda}) d\mathbf{u} d\mathbf{v}$ . Then the probability density function  $\rho_s(\mathbf{u}, \mathbf{v}, \boldsymbol{\lambda})$  with respect to the Lebesgue measure  $d\mathbf{u} d\mathbf{v}$  on  $\mathbb{R}^N \times \mathbb{R}^N$  is a solution of the steady state Fokker-Planck equation (see Propositions 8 and 9 in pages 120 to 123 of Ref. [30]),

$$\sum_{j=1}^N \frac{\partial}{\partial u_j} \{v_j \rho_s\} + \sum_{j=1}^N \frac{\partial}{\partial v_j} \left\{ \left( -\frac{\partial \Phi(\mathbf{u}, \boldsymbol{\lambda})}{\partial u_j} - \frac{f_0}{2} v_j \right) \rho_s \right\} - \frac{f_0}{2} \sum_{j=1}^N \frac{\partial^2 \rho_s}{\partial v_j^2} = 0, \quad \forall (\mathbf{u}, \mathbf{v}) \in \mathbb{R}^N \times \mathbb{R}^N, \quad (21)$$

with the normalization condition

$$\int_{\mathbb{R}^N} \int_{\mathbb{R}^N} \rho_s(\mathbf{u}, \mathbf{v}, \boldsymbol{\lambda}) d\mathbf{u} d\mathbf{v} = 1. \quad (22)$$

In a second stage, it is assumed that, for all  $\boldsymbol{\lambda} \in \mathcal{L}_\mu \subset \mathbb{R}^\mu$ , function  $\mathbf{u} \mapsto \Phi(\mathbf{u}, \boldsymbol{\lambda})$  is continuous on  $\mathbb{R}^N$  and is such that  $\mathbf{u} \mapsto \|\nabla_{\mathbf{u}} \Phi(\mathbf{u}, \boldsymbol{\lambda})\|_N$  is a locally bounded function on  $\mathbb{R}^N$  (*i.e.* is bounded on all compact set in  $\mathbb{R}^N$ ) and is such that

$$\inf_{\|\mathbf{u}\|_N > R} \Phi(\mathbf{u}, \boldsymbol{\lambda}) \rightarrow +\infty \quad \text{if } R \rightarrow +\infty, \quad (23)$$

$$\inf_{\mathbf{u} \in \mathbb{R}^N} \Phi(\mathbf{u}, \boldsymbol{\lambda}) = \Phi_{\min} \quad \text{with } \Phi_{\min} \in \mathbb{R}, \quad (24)$$

$$\int_{\mathbb{R}^N} \|\nabla_{\mathbf{u}} \Phi(\mathbf{u}, \boldsymbol{\lambda})\|_N p(\mathbf{u}, \boldsymbol{\lambda}) d\mathbf{u} < +\infty. \quad (25)$$

Under these above hypotheses and using Theorems 4 to 7 in pages 211 to 216 of Ref. [30] in which the Hamiltonian is taken as  $H(\mathbf{u}, \mathbf{v}) = \|\mathbf{v}\|_N^2/2 + \Phi(\mathbf{u}, \boldsymbol{\lambda})$ , it can be deduced that Eqs. (21) and (22) have a unique solution which is written as

$$\rho_s(\mathbf{u}, \mathbf{v}, \boldsymbol{\lambda}) = c'_\lambda \exp\left\{-\frac{1}{2}\|\mathbf{v}\|_N^2 - \Phi(\mathbf{u}, \boldsymbol{\lambda})\right\}, \quad \forall (\mathbf{u}, \mathbf{v}) \in \mathbb{R}^N \times \mathbb{R}^N, \quad (26)$$

in which  $c'_\lambda$  is the constant of normalization defined by Eq. (22) (for regular functions  $\Phi$ , the expression (26) of the invariant measure of Eqs. (15) and (16) has been obtained by Caughey in [31]). It should be noted that the conditions defined by Eqs. (23) to (25) are not related to the existence of a unique solution of the optimization problem defined by Eq. (4). These conditions are required in order that this unique solution can be interpreted as the unique invariant measure of an ISDE. From Eqs. (7), (14) and (26), it can be deduced that the probability density function  $p(\mathbf{b}, \boldsymbol{\lambda})$  of random variable  $\mathbf{B}_\lambda$  is related to the invariant measure  $\rho_s(\mathbf{u}, \mathbf{v}, \boldsymbol{\lambda}) d\mathbf{u} d\mathbf{v}$  by the following equation,

$$p(\mathbf{b}, \boldsymbol{\lambda}) = \int_{\mathbb{R}^N} \rho_s(\mathbf{b}, \mathbf{v}, \boldsymbol{\lambda}) d\mathbf{v}, \quad \forall \mathbf{b} \in \mathbb{R}^N. \quad (27)$$

Let us consider  $\mathbf{U}_0 = \mathbf{u}_0$  and  $\mathbf{V}_0 = \mathbf{v}_0$  as initial condition defined by Eq. (17) with  $\mathbf{u}_0$  and  $\mathbf{v}_0$  two given vectors in  $\mathbb{R}^N$ . Thus probability distribution  $P_{\mathbf{U}_0, \mathbf{V}_0}(d\mathbf{u}, d\mathbf{v})$  is then equal to the

measure  $\delta_0(\mathbf{u} - \mathbf{u}_0) \otimes \delta_0(\mathbf{v} - \mathbf{v}_0)$  on  $\mathbb{R}^N \times \mathbb{R}^N$  in which  $\delta_0(\mathbf{u})$  is the Dirac measure at the origin of  $\mathbb{R}^N$ . Let  $\{(\mathbf{U}(r), \mathbf{V}(r)), r \geq 0\}$  be the unique solution of Eqs. (15) and (16) with the initial condition

$$\mathbf{U}(0) = \mathbf{u}_0 \quad , \quad \mathbf{V}(0) = \mathbf{v}_0 \quad a.s. \quad . \quad (28)$$

Let  $\mathbf{B}_\lambda$  be the random variable defined in Section 2 for which the probability density function is  $p(\mathbf{b}, \lambda)$  defined by Eq. (7). Consequently, the random variable  $\mathbf{U}(r)$  converges in probability distribution to the random variable  $\mathbf{B}_\lambda$  when  $r$  goes to infinity. We can then write

$$\lim_{r \rightarrow +\infty} \mathbf{U}(r) = \mathbf{B}_\lambda \quad \text{in probability distribution.} \quad (29)$$

As explained above, the free parameter  $f_0 > 0$  introduced in Eq. (16), allows a dissipation term to be introduced in the nonlinear dynamical system and consequently, allows the transient response generated by the initial conditions  $(\mathbf{u}_0, \mathbf{v}_0)$  to be rapidly killed in order to get more rapidly the asymptotic behavior defined by Eq. (29) and corresponding to the stationary solution associated with the invariant measure.

#### 4.2. Random generator of independent realizations

In this subsection, we propose a random generator of  $n_s$  independent realizations  $\mathbf{B}_\lambda(\theta_1), \dots, \mathbf{B}_\lambda(\theta_{n_s})$  of random variable  $\mathbf{B}_\lambda$  whose probability distribution is  $p(\mathbf{b}, \lambda) d\mathbf{b}$ . For  $\theta_1, \dots, \theta_{n_s}$  in  $\Theta$ , let  $\{\mathbf{W}(r, \theta_1), r \geq 0\}, \dots, \{\mathbf{W}(r, \theta_{n_s}), r \geq 0\}$  be  $n_s$  independent realizations of the normalized Wiener stochastic process  $\mathbf{W}$  introduced in Subsection 4.1. For all  $\ell$  fixed in  $\{1, \dots, n_s\}$ , let  $\{(\mathbf{U}(r, \theta_\ell), \mathbf{V}(r, \theta_\ell)), r \geq 0\}$  be the unique solution of the following equation (see Eqs. (15) and (16)) defined for all  $r \geq 0$  by

$$d\mathbf{U}(r, \theta_\ell) = \mathbf{V}(r, \theta_\ell) dr \quad , \quad (30)$$

$$d\mathbf{V}(r, \theta_\ell) = -\nabla_{\mathbf{u}} \Phi(\mathbf{U}(r, \theta_\ell), \lambda) dr - \frac{f_0}{2} \mathbf{V}(r, \theta_\ell) dr + \sqrt{f_0} d\mathbf{W}(r, \theta_\ell) \quad , \quad (31)$$

with the initial condition

$$\mathbf{U}(0, \theta_\ell) = \mathbf{u}_0 \quad , \quad \mathbf{V}(0, \theta_\ell) = \mathbf{v}_0 \quad . \quad (32)$$

From Eq. (29), we deduce that each independent realization  $\mathbf{B}_\lambda(\theta_\ell)$  can be constructed by

$$\mathbf{B}_\lambda(\theta_\ell) = \mathbf{U}(r, \theta_\ell) \quad \text{for } r \text{ sufficiently large.} \quad (33)$$

#### 4.3. Estimation of mathematical expectations

In this subsection, we propose two estimations of the mathematical expectation  $E\{\mathbf{g}(\mathbf{B}_\lambda)\}$  defined by Eq. (11) and one estimation of  $E\{\mathbf{Y}\}$  in which  $\mathbf{Y} = \mathbf{q}(\mathbf{A})$  is the random response of a large computational model depending on the random parameter  $\mathbf{A}$ .

(i) *Use of the ergodic method.* For any realization  $\theta$ , let  $\{\mathbf{U}(r, \theta), r \geq 0\}$  be the solution of Eqs. (30) to (32). Then using the ergodic theorem [32], we can estimate  $E\{\mathbf{g}(\mathbf{B}_\lambda)\}$  by

$$E\{\mathbf{g}(\mathbf{B}_\lambda)\} = \lim_{R \rightarrow +\infty} \frac{1}{R} \int_0^R \mathbf{g}(\mathbf{U}(r, \theta)) dr \quad . \quad (34)$$

(ii) *Use of the Monte Carlo method.* Let  $\mathbf{B}_\lambda(\theta_1), \dots, \mathbf{B}_\lambda(\theta_{n_s})$  be  $n_s$  independent realizations of random variable  $\mathbf{B}_\lambda$  constructed using the random generator presented in Subsection 4.2. Then the mathematical expectation  $E\{\mathbf{g}(\mathbf{B}_\lambda)\}$  can be estimated (see also Eq. (13)) by

$$E\{\mathbf{g}(\mathbf{B}_\lambda)\} = \lim_{n_s \rightarrow +\infty} \frac{1}{n_s} \sum_{\ell=1}^{n_s} \mathbf{g}(\mathbf{B}_\lambda(\theta_\ell)) \quad . \quad (35)$$

Let us now consider a computational stochastic model for which we are interested in estimating  $E\{\mathbf{Y}\}$  in which  $\mathbf{Y} = \mathbf{q}(\mathbf{A})$  is the random response calculated with a large computational model. Such a mathematical expectation  $E\{\mathbf{Y}\}$  cannot generally be estimated using the ergodic method (see the Remark below). Since  $\mathbf{A} = \mathbf{B}_{\lambda^{\text{sol}}}$  for the convergence in probability distribution, we then propose to use the Monte Carlo method which yields the following estimation

$$E\{\mathbf{Y}\} = \lim_{n_s \rightarrow +\infty} \frac{1}{n_s} \sum_{\ell=1}^{n_s} \mathbf{q}(\mathbf{B}_{\lambda^{\text{sol}}}(\theta_\ell)) \quad . \quad (36)$$

*Remark.* It should be noted that, if the ergodic method can effectively be used to estimate  $E\{\mathbf{g}(\mathbf{B}_\lambda)\}$  in order to calculate  $\lambda^{\text{sol}}$ , it can generally not be used to estimate statistics of random responses of uncertain complex mechanical systems in computational stochastic mechanics as soon as  $\lambda^{\text{sol}}$  has been calculated. Let us consider a computational stochastic model for which we are interested in estimating  $E\{\mathbf{Y}\}$  in which  $\mathbf{Y} = \mathbf{q}(\mathbf{A})$  is the random response calculated with a large computational model and depending on the uncertain parameter  $\mathbf{A}$  whose probability distribution is  $p(\mathbf{b}, \lambda^{\text{sol}}) d\mathbf{b}$ . Note that for large computational models the numerical construction of one evaluation  $\mathbf{y} = \mathbf{q}(\mathbf{a})$  for  $\mathbf{a}$  given in  $\mathbb{R}^N$  is generally very high. In practice, the total number of such evaluations is restricted to a low number (a few hundred but certainly not a few thousands, a ten thousands or more). The MCMC method and then the method proposed consist in constructing an ergodic homogeneous Markov chain  $\{\mathbf{U}(r_k), k = 1, \dots, M\}$  admitting as invariant measure the measure  $p(\mathbf{b}, \lambda^{\text{sol}}) d\mathbf{b}$ . Since  $\mathbf{Y}$  can be rewritten as  $\mathbf{Y} = \mathbf{q}(\mathbf{B}_{\lambda^{\text{sol}}})$ , the use of the ergodic method yields  $E\{\mathbf{Y}\} \simeq 1/M \sum_{k=1}^M \mathbf{q}(\mathbf{U}(r_k, \theta))$  in which  $\{\mathbf{U}(r_k, \theta), k = 1, \dots, M\}$  is a realization of  $\{\mathbf{U}(r_k), k = 1, \dots, M\}$ . In general, convergence is reached for large value of  $M$  (several ten thousands or more) and consequently, would require a very large number  $M$  of evaluations  $\mathbf{y} = \mathbf{q}(\mathbf{a})$  that is not realistic. This is the reason why the ergodic method cannot easily be used to estimate  $E\{\mathbf{Y}\}$  when  $\lambda^{\text{sol}}$  is known but can effectively be used to estimate  $\lambda^{\text{sol}}$ .

## 5. Discretization of the Itô stochastic differential equation

In this section, we construct the discretization of the Itô stochastic differential equation (ISDE) defined by Eqs. (15) to (17) with the initial condition  $\mathbf{U}_0 = \mathbf{u}_0$  and  $\mathbf{V}_0 = \mathbf{v}_0$  in which  $\mathbf{u}_0$  and  $\mathbf{v}_0$  are two given vectors in  $\mathbb{R}^N$ . The objective is to construct an approximation of the solution  $\{(\mathbf{U}(r), \mathbf{V}(r)), r \geq 0\}$  of this ISDE. Two integration schemes will be proposed. The first one will be an explicit Euler scheme and the second one a semi-implicit scheme.

Let  $m$  and  $M$  two integers such that  $m < M$ . The Itô stochastic differential equation will be solved on the finite interval  $[0, (M-1)\Delta r]$  in which  $\Delta r$  is the sampling step of the continuous index parameter  $r$ . The integration scheme will be based on the use of the  $M$  sampling points

$r_k$  such that

$$\Delta r = \frac{\beta}{m} \quad , \quad \{r_k = (k-1)\Delta r \quad , \quad k = 1, \dots, M\} \quad , \quad (37)$$

in which  $\beta$  is a given positive real number. Consequently, the two parameters for studying the convergence of the constructed approximation will be  $m$  and  $M$ . We then introduce the following notation

$$\mathbf{U}^k = \mathbf{U}(r_k) \quad , \quad \mathbf{V}^k = \mathbf{V}(r_k) \quad , \quad \mathbf{W}^k = \mathbf{W}(r_k) \quad \text{for } k = 1, \dots, M \quad . \quad (38)$$

### 5.1. Explicit Euler scheme

The explicit Euler scheme (see for instance [33, 34]) applied to the ISDE defined by Eqs. (15) to (17) yields the following scheme for  $k = 1, \dots, M-1$

$$\mathbf{U}^{k+1} = \mathbf{U}^k + \Delta r \mathbf{V}^k \quad , \quad (39)$$

$$\mathbf{V}^{k+1} = \left(1 - \frac{f_0}{2} \Delta r\right) \mathbf{V}^k + \Delta r \mathbf{L}^k + \sqrt{f_0} \Delta \mathbf{W}^{k+1} \quad , \quad (40)$$

with the initial condition

$$\mathbf{U}^1 = \mathbf{u}_0 \quad , \quad \mathbf{V}^1 = \mathbf{v}_0 \quad . \quad (41)$$

In Eq. (40),  $\Delta \mathbf{W}^{k+1} = \mathbf{W}^{k+1} - \mathbf{W}^k$  is a second-order Gaussian centered  $\mathbb{R}^N$ -valued random variable with covariance matrix  $E\{\Delta \mathbf{W}^{k+1} (\Delta \mathbf{W}^{k+1})^T\} = \Delta r [I_N]$ , with  $\mathbf{W}^1 = 0_N$  and where all the random variables  $\Delta \mathbf{W}^2, \dots, \Delta \mathbf{W}^M$  are mutually independent. We have introduced the  $\mathbb{R}^N$ -valued random variable  $\mathbf{L}^k = (L_1^k, \dots, L_N^k)$  such that for all  $j$ ,  $L_j^k = -\{\partial \Phi(\mathbf{u}, \boldsymbol{\lambda}) / \partial u_j\}_{\mathbf{u}=\mathbf{U}^k}$  which is the partial derivative of  $\Phi(\mathbf{u}, \boldsymbol{\lambda})$  with respect to  $u_j$  in  $\mathbf{u} = \mathbf{U}^k$ . When  $\mathbf{L}^k$  has to be calculated,  $\mathbf{U}^{k+1}$  is known by Eq. (39). Consequently, for each  $j$ , we approximate  $L_j^k$  by the forward finite difference in the direction defined by  $\mathbf{U}^{k+1}$ . Then introducing the  $\mathbb{R}^N$ -valued random variable  $\Delta \mathbf{U}^{k,j}$  such that

$$\Delta \mathbf{U}^{k,j} = (U_1^k, \dots, U_{j-1}^k, U_j^k + \Delta U_j^{k+1}, U_{j+1}^k, \dots, U_N^k) \quad , \quad \Delta U_j^{k+1} = U_j^{k+1} - U_j^k \quad , \quad (42)$$

for all  $j$  in  $\{1, \dots, N\}$  we write

$$L_j^k \simeq -\frac{\Phi(\Delta \mathbf{U}^{k,j}, \boldsymbol{\lambda}) - \Phi(\mathbf{U}^k, \boldsymbol{\lambda})}{U_j^{k+1} - U_j^k} \quad . \quad (43)$$

This scheme is conditionally stable and  $\Delta r$  has to be taken sufficiently small. In practice, convergence of the solution has to be analyzed in function of integer  $m$  which must be taken sufficiently large.

### 5.2. Semi-implicit scheme

The use of an implicit scheme to solve an ISDE (see for instance [34, 30, 35]) requires to solve a nonlinear algebraic equation for every sampling point  $r_k$  (for instance, using an iteration algorithm). Such an implicit scheme allows the step size  $\Delta r$  to be increased in preserving the stability of the scheme. Nevertheless, such an implicit scheme is time consuming because a very large number of evaluations of  $\mathbf{L}^k$  is required due to the nonlinear algebraic equation which has to be solved at every sampling point. Below we propose a semi-implicit scheme which is

a compromise between the explicit scheme presented in Section 5.1 and an implicit scheme. Such a scheme is restricted to the class of functions  $\Phi(\cdot, \boldsymbol{\lambda})$  for which  $\nabla_{\mathbf{u}}\Phi(\mathbf{u}, \boldsymbol{\lambda})$  is made up of a linear part with respect to  $\mathbf{u}$  and a nonlinear part. The semi-implicit scheme consists in impliciting the linear part and in expliciting the nonlinear part avoiding the nonlinear algebraic equation to be solved at every sampling point. Function  $\Phi$  is then assumed to be written as

$$\Phi(\mathbf{u}, \boldsymbol{\lambda}) = \frac{1}{2} \langle [K_{\boldsymbol{\lambda}}] \mathbf{u}, \mathbf{u} \rangle_N + \Phi_{\text{NL}}(\mathbf{u}, \boldsymbol{\lambda}) \quad , \quad (44)$$

in which the matrix  $[K_{\boldsymbol{\lambda}}]$  depends on  $\boldsymbol{\lambda}$  and is such that, for all admissible values of  $\boldsymbol{\lambda} \in \mathcal{L}_{\mu} \subset \mathbb{R}^{\mu}$ ,  $[K_{\boldsymbol{\lambda}}]$  belongs to the set  $\mathbb{M}_N^+(\mathbb{R})$  of all the positive-definite symmetric  $(N \times N)$  real matrices. Consequently, the random vector  $\mathbf{L}(\mathbf{u}) = -\nabla_{\mathbf{u}}\Phi(\mathbf{u}, \boldsymbol{\lambda})$  can be written as

$$\mathbf{L}(\mathbf{u}) = \mathbf{L}_L(\mathbf{u}) + \mathbf{L}_{\text{NL}}(\mathbf{u}) \quad , \quad \mathbf{L}_L(\mathbf{u}) = -[K_{\boldsymbol{\lambda}}] \mathbf{u} \quad , \quad \mathbf{L}_{\text{NL}}(\mathbf{u}) = -\nabla_{\mathbf{u}}\Phi_{\text{NL}}(\mathbf{u}, \boldsymbol{\lambda}) \quad . \quad (45)$$

The semi-implicit scheme applied to the ISDE defined by Eqs. (15) to (17) yields the following scheme for  $k = 1, \dots, M - 1$

$$\mathbf{U}^{k+1} - \mathbf{U}^k = \frac{\Delta r}{2} (\mathbf{V}^{k+1} + \mathbf{V}^k) \quad , \quad (46)$$

$$\mathbf{V}^{k+1} - \mathbf{V}^k = -\frac{\Delta r}{2} [K_{\boldsymbol{\lambda}}] (\mathbf{U}^{k+1} + \mathbf{U}^k) + \Delta r \mathbf{L}_{\text{NL}}^k - \frac{f_0}{4} \Delta r (\mathbf{V}^{k+1} + \mathbf{V}^k) + \sqrt{f_0} \Delta \mathbf{W}^{k+1} \quad , \quad (47)$$

with the initial condition

$$\mathbf{U}^1 = \mathbf{u}_0 \quad , \quad \mathbf{V}^1 = \mathbf{v}_0 \quad . \quad (48)$$

In Eq. (47),  $\Delta \mathbf{W}^{k+1}$  is defined in Section 5.1,  $\mathbf{L}_{\text{NL}}^k = (L_{\text{NL},1}^k, \dots, L_{\text{NL},N}^k)$  and for all  $j$  in  $\{1, \dots, N\}$ , we have (see Eq. (43)),

$$L_{\text{NL},j}^k \simeq -\frac{\Phi_{\text{NL}}(\Delta \mathbf{U}^{k,j}, \boldsymbol{\lambda}) - \Phi_{\text{NL}}(\mathbf{U}^k, \boldsymbol{\lambda})}{U_j^{k+1} - U_j^k} \quad , \quad (49)$$

in which  $\Delta \mathbf{U}^{k,j}$  is defined by Eq. (42). Equations (46) and (47) can be rewritten as

$$[\mathcal{A}_{\boldsymbol{\lambda}}] \mathbf{V}^{k+1} = [\mathcal{B}_{\boldsymbol{\lambda}}] \mathbf{V}^k - \Delta r [K_{\boldsymbol{\lambda}}] \mathbf{U}^k + \Delta r \mathbf{L}_{\text{NL}}^k + \sqrt{f_0} \Delta \mathbf{W}^{k+1} \quad , \quad (50)$$

$$\mathbf{U}^{k+1} = \mathbf{U}^k + \frac{\Delta r}{2} (\mathbf{V}^{k+1} + \mathbf{V}^k) \quad , \quad (51)$$

in which the matrices  $[\mathcal{A}_{\boldsymbol{\lambda}}]$  and  $[\mathcal{B}_{\boldsymbol{\lambda}}]$  are defined by

$$[\mathcal{A}_{\boldsymbol{\lambda}}] = (1 + \frac{f_0}{4} \Delta r) [I_N] + \frac{\Delta r^2}{4} [K_{\boldsymbol{\lambda}}] \quad , \quad [\mathcal{B}_{\boldsymbol{\lambda}}] = (1 - \frac{f_0}{4} \Delta r) [I_N] - \frac{\Delta r^2}{4} [K_{\boldsymbol{\lambda}}] \quad . \quad (52)$$

First, the linear Eq.(50) is solved to calculate  $\mathbf{V}^{k+1}$  and then Eq.(51) yields  $\mathbf{U}^{k+1}$ .

## 6. Estimation of the mathematical expectations and random generator of independent realizations

In this subsection, we give the estimations of  $E\{\mathbf{g}(\mathbf{B}_{\boldsymbol{\lambda}})\}$  (defined by Eq. (11)) and  $E\{\mathbf{Y}\}$  (defined in Section 4.3-(ii)) and we explicit the random generator of independent realizations of  $\mathbf{B}_{\boldsymbol{\lambda}}$  using Sections 4.3 and 5.

### 6.1. Estimation of the mathematical expectation using ergodic method

For  $\theta$  in  $\Theta$ , let  $\{\mathbf{U}^k(\theta), k = 1, \dots, M\}$  be any realization of the family of vector-valued random variables  $\{\mathbf{U}^k, k = 1, \dots, M\}$  calculated by using Eqs. (39) to (41) (explicit Euler scheme) or Eqs. (50),(51) and (48) semi-implicit scheme). From Eq. (34), for  $m$  and  $M_0$  fixed, and for  $M$  sufficiently large,  $E\{\mathbf{g}(\mathbf{B}_\lambda)\}$  can be estimated by

$$E\{\mathbf{g}(\mathbf{B}_\lambda)\} \simeq \frac{1}{M - M_0 + 1} \sum_{k=M_0}^M \mathbf{g}(\mathbf{U}^k(\theta)) \quad . \quad (53)$$

The parameter  $M_0$  allows to remove the transient part of the response induced by the initial conditions. By definition of  $M_0$ , the stochastic process is stationary for  $k \in \{M_0, \dots, M\}$ . Convergence has to be studied with respect to the two other parameters  $m$  and  $M$ .

### 6.2. Random generator of independent realizations

The random generator is described in Section 4.2. For all  $\lambda$  (or for  $\lambda = \lambda_{\text{sol}}$  and then  $\mathbf{A} = \mathbf{B}_{\lambda_{\text{sol}}}$ ) and for all  $\ell$  in  $\{1, \dots, n_s\}$ , let  $\{\mathbf{U}^k(\theta_\ell), k = 1, \dots, M\}$  be  $n_s$  independent realizations of the family of vector-valued random variables  $\{\mathbf{U}^k, k = 1, \dots, M\}$  calculated by Eqs. (39) to (41) (explicit Euler scheme) or by Eqs. (50),(51) and (48) (semi-implicit scheme). From Eq. (29) and for  $m$  and  $M$  sufficiently large, we can write

$$\mathbf{B}_\lambda(\theta_\ell) \simeq \mathbf{U}^M(\theta_\ell) \quad , \quad \forall \ell \in \{1, \dots, n_s\} \quad . \quad (54)$$

Consequently,  $\mathbf{B}_\lambda(\theta_1), \dots, \mathbf{B}_\lambda(\theta_{n_s})$  are  $n_s$  independent realizations of random variable  $\mathbf{B}_\lambda$  constructed using Eq. (54).

### 6.3. Estimation of the mathematical expectations using the Monte Carlo method

For all  $\lambda$  (or for  $\lambda = \lambda_{\text{sol}}$ ), let  $\mathbf{B}_\lambda(\theta_1), \dots, \mathbf{B}_\lambda(\theta_{n_s})$  be  $n_s$  independent realizations of random variable  $\mathbf{B}_\lambda$  constructed with the random generator as explained in Subsection 6.2. Then, from Eqs. (35) and (36) and for  $n_s$  sufficiently large, it can be deduced that the mathematical expectations  $E\{\mathbf{g}(\mathbf{B}_\lambda)\}$  and  $E\{\mathbf{Y}\}$  can be estimated by

$$E\{\mathbf{g}(\mathbf{B}_\lambda)\} \simeq \frac{1}{n_s} \sum_{\ell=1}^{n_s} \mathbf{g}(\mathbf{B}_\lambda(\theta_\ell)) \quad , \quad E\{\mathbf{Y}\} \simeq \frac{1}{n_s} \sum_{\ell=1}^{n_s} \mathbf{q}(\mathbf{B}_{\lambda_{\text{sol}}}(\theta_\ell)) \quad . \quad (55)$$

Convergence has to be studied with respect to parameters  $m$ ,  $M$  and  $n_s$ . Parameter  $m$  is related to the precision of the approximation. Since the invariance measure cannot be chosen as the probability distribution of the initial conditions,  $M$  must be chosen for that the stationarity of the sequence  $\{\mathbf{U}^k\}_k$  be obtained. Parameter  $n_s$  must be chosen such that the estimator of the mathematical expectation be converged.

## 7. Construction of a probability model for a nonstationary time series and application to the construction of random accelerograms for a given Velocity Response Spectrum

This application is devoted to the construction of a probabilistic model of a nonstationary random time series corresponding to the sampling of a continuous-time stochastic process

(random accelerograms) for which the mean value, the standard deviation and the mean Velocity Response Spectrum (VRS) as a function of the frequency are given (see for instance [36] for the notion of VRS). We then propose a new formulation to solve this stochastic inverse problem.

### 7.1. Definition of the time series and its available information

Let  $\{\Gamma_1, \dots, \Gamma_N\}$  be a real-valued random time series in which  $N = 128$ . Let  $\mathbf{\Gamma} = (\Gamma_1, \dots, \Gamma_N)$  be the  $\mathbb{R}^N$ -valued random vector associated with this random time series. The problem is the construction of the probability distribution on  $\mathbb{R}^N$  of the random vector  $\mathbf{\Gamma}$  using the MaxEnt principle for which the available information is defined below.

- (i) The mean value  $E\{\mathbf{\Gamma}\} \in \mathbb{R}^N$  is zero.
- (ii) For all  $j$  in  $\{1, \dots, N\}$ , since the random variable  $\Gamma_j$  is centered, the variance of  $\Gamma_j$  is written as  $\sigma_j^2 = E\{\Gamma_j^2\} < +\infty$ . Thus  $\mathbf{\Gamma}$  is a second-order random variable because  $E\{\|\mathbf{\Gamma}\|_N^2\} = \sum_{j=1}^N \sigma_j^2 < +\infty$ . Figure 1 (left) displays the graph of the standard-deviation function  $j \mapsto \sigma_j$  from  $\{1, \dots, N\}$  into  $\mathbb{R}_+$ .
- (iii) Let  $\{\mathcal{S}_k(\mathbf{\Gamma}), k = 1, \dots, \nu\}$  be the random VRS with  $\nu = 40$  in which index  $k$  is associated with the given frequencies  $\omega_1, \dots, \omega_\nu$  and where  $\boldsymbol{\gamma} \mapsto \mathcal{S}(\boldsymbol{\gamma}) = (\mathcal{S}_1(\boldsymbol{\gamma}), \dots, \mathcal{S}_\nu(\boldsymbol{\gamma}))$  is a given nonlinear mapping from  $\mathbb{R}^N$  into  $\mathbb{R}^\nu$  such that

$$\mathcal{S}_k(\boldsymbol{\gamma}) = \omega_k \max\{|x_1^k|, \dots, |x_N^k|\} \quad , \quad x_j^k = \sum_{j'=1}^N [B^k]_{jj'} \gamma_{j'} \quad , \quad (56)$$

in which  $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_N)$  and where  $\{[B^k], k = 1, \dots, \nu\}$  is a given family of  $(N \times N)$  real matrices. Consequently, we have  $\mathcal{S}(-\boldsymbol{\gamma}) = \mathcal{S}(\boldsymbol{\gamma})$ . Let  $\mathcal{S}(\mathbf{\Gamma})$  be the  $\mathbb{R}^\nu$ -valued random variable such that

$$\mathcal{S}(\mathbf{\Gamma}) = (\mathcal{S}_1(\mathbf{\Gamma}), \dots, \mathcal{S}_\nu(\mathbf{\Gamma})) \quad , \quad (57)$$

for which the mean value  $E\{\mathcal{S}(\mathbf{\Gamma})\} = \underline{\mathcal{S}}$  is a given vector  $\underline{\mathcal{S}}$  in  $\mathbb{R}^\nu$ . Figure 1 (right) displays the graph of function  $k \mapsto \underline{\mathcal{S}}_k$  from  $\{1, \dots, \nu\}$  into  $\mathbb{R}_+$ .

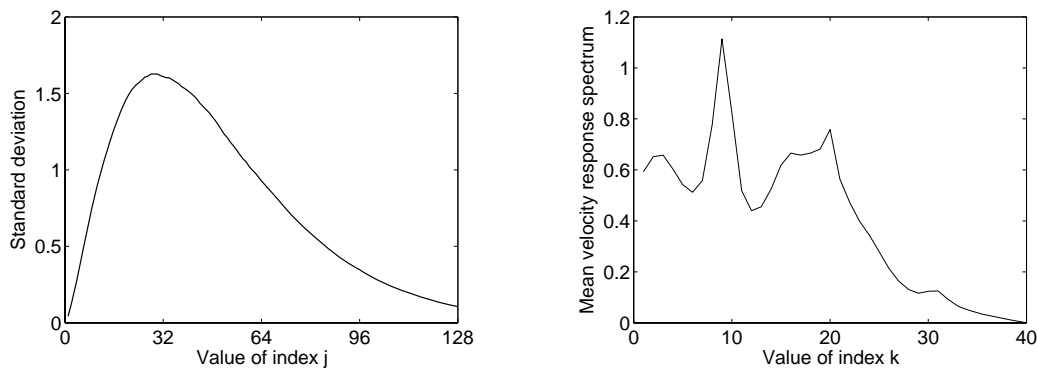


Figure 1. Graph of  $j \mapsto \sigma_j = \sqrt{E\{\Gamma_j^2\}}$  (left figure) and graph of  $k \mapsto \underline{\mathcal{S}}_k$  (right figure).

### 7.2. Normalization

In this section we construct the random vector  $\mathbf{A}$  with values in  $\mathbb{R}^N$  as a normalization of random vector  $\mathbf{\Gamma}$ . We then construct the probability distribution and the random generator of  $\mathbf{A}$  and it will be easy to deduce the probability distribution and the random generator of random vector  $\mathbf{\Gamma}$ . Let  $\mathbf{A} = (A_1, \dots, A_N)$  be the  $\mathbb{R}^N$ -valued random variable defined, for all  $j$  in  $\{1, \dots, N\}$ , by  $\Gamma_j = \sqrt{N} \sigma_j A_j$ . We can then rewrite  $\mathbf{\Gamma}$  as

$$\mathbf{\Gamma} = \sqrt{N} [\sigma] \mathbf{A} \quad , \quad [\sigma]_{jj'} = \sigma_j \delta_{jj'} \quad , \quad (58)$$

in which  $[\sigma]$  is a  $(N \times N)$  real diagonal matrix. The available information introduced in Section 7.1 for  $\mathbf{\Gamma}$  allows the corresponding available information for  $\mathbf{A}$  to be easily deduced.

(i) The mean value  $E\{\mathbf{A}\} \in \mathbb{R}^N$  because  $E\{\mathbf{A}\} = N^{-1/2} [\sigma]^{-1} E\{\mathbf{\Gamma}\}$ . Consequently,  $\mathbf{A}$  is a centered random variable,

$$E\{\mathbf{A}\} = \mathbf{0} \quad . \quad (59)$$

(ii) For all  $j$  in  $\{1, \dots, N\}$ , the second-order moment of random variable  $A_j$  is thus equal to  $1/N$  and consequently, we have

$$E\{A_j^2\} = \frac{1}{N} \quad , \quad \forall j \in \{1, \dots, N\} \quad , \quad (60)$$

and then  $E\{\|\mathbf{A}\|_N^2\} = 1$ .

(iii) Let  $\mathbf{s} = (\underline{s}_1, \dots, \underline{s}_\nu) \in \mathbb{R}^\nu$  in which  $\underline{s}_k = 1$  for all  $k = 1, \dots, \nu$  (all the components of vector  $\mathbf{s}$  are equal to 1). Let  $\mathbf{a} \mapsto \mathbf{s}(\mathbf{a}) = (s_1(\mathbf{a}), \dots, s_\nu(\mathbf{a}))$  be the nonlinear mapping from  $\mathbb{R}^N$  into  $\mathbb{R}^\nu$  such that

$$s_k(\mathbf{a}) = \frac{\underline{s}_k(\sqrt{N}[\sigma]\mathbf{a})}{\underline{s}_k} \quad , \quad \forall k = 1, \dots, \nu \quad . \quad (61)$$

It can then easily be deduced that

$$E\{\mathbf{s}(\mathbf{A})\} = \mathbf{s} \in \mathbb{R}^\nu \quad . \quad (62)$$

Therefore, the available information which allows the probability distribution of random vector  $\mathbf{A}$  to be constructed is made up of Eqs. (59), (60) and (62).

### 7.3. Defining the unknown Lagrange multipliers and function $\Phi$

Taking into account the normalization condition for the probability density function  $p_{\mathbf{A}}$  and the available information defined by Eqs. (59), (60) and (62), the use of the MaxEnt principle yields

$$p_{\mathbf{A}}(\mathbf{a}) = c_0^{\text{sol}} \exp\{-\langle \boldsymbol{\lambda}_1^{\text{sol}}, \mathbf{a} \rangle_N - \langle \boldsymbol{\lambda}_2^{\text{sol}}, \mathbf{a}^2 \rangle_N - \langle \boldsymbol{\lambda}_3^{\text{sol}}, \mathbf{s}(\mathbf{a}) \rangle_\nu\} \quad , \quad \forall \mathbf{a} \in \mathbb{R}^N \quad , \quad (63)$$

in which  $\mathbf{a}^2$  denotes the vector  $(a_1^2, \dots, a_N^2)$  in  $\mathbb{R}^N$  and where, for the solution,  $\boldsymbol{\lambda}_1^{\text{sol}} \in \mathbb{R}^N$ ,  $\boldsymbol{\lambda}_2^{\text{sol}} \in \mathbb{R}^N$  and  $\boldsymbol{\lambda}_3^{\text{sol}} \in \mathbb{R}^\nu$  are the values of the Lagrange multipliers associated with the constraints defined by Eq. (59), (60) and Eq. (62) respectively. Since  $\mathcal{S}(-\gamma) = \mathcal{S}(\gamma)$  (see Section 7.1) and from Eqs. (59) and (63), it can be proven that  $\boldsymbol{\lambda}_1^{\text{sol}} = \mathbf{0}$ . Therefore, the Lagrange multiplier  $\boldsymbol{\lambda}$  introduced in Section 2 can be written as  $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_2, \boldsymbol{\lambda}_3) \in \mathcal{L}_\mu \subset \mathbb{R}^\mu = \mathbb{R}^N \times \mathbb{R}^\nu$  in which  $\mu = N + \nu$  and where the admissible set  $\mathcal{L}_\mu$  of  $\boldsymbol{\lambda}$  is such that  $\mathcal{L}_\mu = (]0, +\infty[)^N \times \mathbb{R}^\nu$ . We then have  $\mu = 168$ . The mapping  $\mathbf{a} \mapsto \mathbf{g}(\mathbf{a})$  from  $\mathbb{R}^N$  into  $\mathbb{R}^\mu$  introduced in Section 2 can



then be written as  $\mathbf{g}(\mathbf{a}) = (\mathbf{a}^2, \mathbf{s}(\mathbf{a})) \in \mathbb{R}^N \times \mathbb{R}^\nu = \mathbb{R}^\mu$ . Let  $\mathbf{h} = (h_1, \dots, h_N)$  be the vector in  $\mathbb{R}^N$  such that  $h_j = 1/N$  for all  $j$ . Vector  $\mathbf{f} \in \mathbb{R}^\mu$  introduced in Eq. (2) is then written as  $\mathbf{f} = (\mathbf{h}, \underline{\mathbf{s}}) \in \mathbb{R}^N \times \mathbb{R}^\nu = \mathbb{R}^\mu$ . We use the semi-implicit scheme presented in Section 5.2 for the discretization of the Itô stochastic differential equation. Function  $\Phi(\mathbf{u}, \boldsymbol{\lambda})$  defined by Eq. (14) is then written as  $\Phi(\mathbf{u}, \boldsymbol{\lambda}) = \frac{1}{2} \langle [K_\lambda] \mathbf{u}, \mathbf{u} \rangle_N + \Phi_{\text{NL}}(\mathbf{u}, \boldsymbol{\lambda})$  (see Eq. (44)) in which matrix  $[K_\lambda]$  is such that  $[K_\lambda]_{jj'} = 2\{\boldsymbol{\lambda}_2\}_j \delta_{jj'}$  and where  $\Phi_{\text{NL}}(\mathbf{u}, \boldsymbol{\lambda}) = \langle \boldsymbol{\lambda}_3, \mathbf{s}(\mathbf{u}) \rangle_\nu$ . For all  $j$  fixed in  $\{1, \dots, N\}$ , the generalized partial derivative of function  $\mathbf{s}$  with respect to  $u_j$  is represented by the function  $\mathbf{u} \mapsto \partial \mathbf{s}(\mathbf{u}) / \partial u_j$  which is locally bounded on  $\mathbb{R}^N$ . Therefore, for all  $\boldsymbol{\lambda}_2$  in  $(]0, +\infty[)^N \subset \mathbb{R}^N$  and for all  $\boldsymbol{\lambda}_3$  in  $\mathbb{R}^\nu$ , Eqs. (23) to (25) are satisfied.

#### 7.4. Computation of the vector-valued Lagrange multipliers using ergodic method

Lagrange multiplier  $\boldsymbol{\lambda}^{\text{sol}} = (\boldsymbol{\lambda}_2^{\text{sol}}, \boldsymbol{\lambda}_3^{\text{sol}}) \in \mathcal{L}_\mu \subset \mathbb{R}^\mu = \mathbb{R}^N \times \mathbb{R}^\nu$  are computed in solving Eq. (10) as explained at the end of Section 2.

(i) The interior-reflective Newton method used to solve Eq. (10) is initialized with  $\boldsymbol{\lambda}^0 = (\boldsymbol{\lambda}_2^0, \boldsymbol{\lambda}_3^0) \in (]0, +\infty[)^N \times \mathbb{R}^\nu = \mathcal{L}_\mu$  with  $\boldsymbol{\lambda}_2^0 = 0.5 \eta N \mathbf{1}_N$  and  $\boldsymbol{\lambda}_3^0 = 0.5 \nu^{-1} (1 - \eta) N \mathbf{1}_\nu$  with  $\eta = 0.98$  and where  $\mathbf{1}_N$  and  $\mathbf{1}_\nu$  are the vectors in  $\mathbb{R}^N$  and  $\mathbb{R}^\nu$  for which all the components are equal to 1. The solution has been calculated in four steps (in order to optimize the computer time). For step 1, the parameters are  $M = 600$  and  $\text{iter} = 1, \dots, 16$  and for step 2, are  $M = 2800$  and  $\text{iter} = 17, \dots, 30$ . Figure 2 (right) displays the graph of the function  $\text{iter} \mapsto \text{convALG}(\text{iter}) = \|E\{\mathbf{g}(\mathbf{B}_{\boldsymbol{\lambda}(\text{iter})})\} - \mathbf{f}\|_\mu^2$ . This calculation is completed by two other steps, one for which  $M = 8300$  with 7 iterations and the last one for which  $M = 20000$  with 7 iterations.

(ii) The mathematical expectation defined by Eq. (11) is estimated by using the ergodic method (see Eq. (53)) with  $M_0 = 300$  and the semi-implicit scheme presented in Section 5.2 is used to construct a realization  $\{\mathbf{U}^k(\theta), k = 1, \dots, M\}$  of the sequence of the vector-valued random variables  $\{\mathbf{U}^k, k = 1, \dots, M\}$ . For each value of vector  $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_2, \boldsymbol{\lambda}_3)$  corresponding to an iteration of the algorithm, the sampling step defined by Eq. (37) has been written as  $\Delta r = \beta/m$  with  $\beta = 2\pi/(\sqrt{2\lambda_2^{\text{max}}})$  in which  $\lambda_2^{\text{max}} = \max\{(\boldsymbol{\lambda}_2)_1, \dots, (\boldsymbol{\lambda}_2)_N\}$  and with  $m = 5$ . Parameter  $f_0$  has been fixed to 1. These values of parameters  $\Delta r$  and  $f_0$  have been deduced from a convergence analysis. For instance, Figure 2 (left) displays the graph of the function  $M \mapsto \text{conv}(M) = \frac{1}{M} \sum_{k=1}^M \|\mathbf{U}^k(\theta)\|_N^2$  for  $\boldsymbol{\lambda} = \boldsymbol{\lambda}^{\text{sol}}$  and for realization  $\theta$ . Therefore  $\lim_{M \rightarrow +\infty} \text{conv}(M) = E\{\|\mathbf{B}_{\boldsymbol{\lambda}^{\text{sol}}}\|_N^2\} = E\{\|\mathbf{A}\|_N^2\}$  and we must have  $\lim_{M \rightarrow +\infty} \text{conv}(M) = 1$  (see Eq. (60)). For  $M = 20000$ , we have  $\text{conv}(M) = 0.982$  instead of 1 which corresponds to an error of about 1.8%. The convergence is then reasonably reached for  $M = 20000$ .

(iii) Figure 3 shows solution  $\boldsymbol{\lambda}^{\text{sol}}$  and displays the graph of  $j \mapsto (\boldsymbol{\lambda}_2^{\text{sol}})_j$  (left figure) and the graph of  $k \mapsto (\boldsymbol{\lambda}_3^{\text{sol}})_k$  (right figure).

#### 7.5. Estimation of the constraints using the Monte Carlo method with the random generator

Solution  $\boldsymbol{\lambda}^{\text{sol}}$  of Eq. (10) being known,  $n_s$  independent realizations of random variable  $\mathbf{A} = \mathbf{B}_{\boldsymbol{\lambda}^{\text{sol}}}$  are constructed using the method presented in Section 6.2. The estimation of the mathematical expectation  $E\{\mathbf{g}(A)\}$  of the constraint is carried out with the Monte Carlo method as explained in Section 6.3 (see Eq. (55)). Computation is performed with  $m = 5$  (as in Section 7.4),  $M = 400$  and  $n_s = 300$ . (i) Concerning the value of  $M$ , Figure 4 (left) displays the graph of the function  $M \mapsto \text{conv}_\ell(M) = \frac{1}{M} \sum_{k=1}^M \|\mathbf{U}^k(\theta_\ell)\|_N^2$  for  $\boldsymbol{\lambda} = \boldsymbol{\lambda}^{\text{sol}}$  and for a realization  $\theta_\ell$

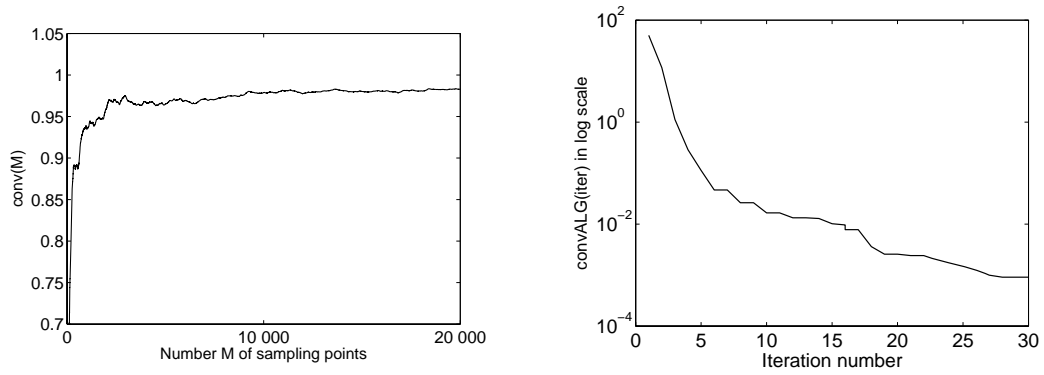


Figure 2. Graph of  $M \mapsto \text{conv}(M)$  (left figure) and graph of  $\text{iter} \mapsto \text{convALG}(\text{iter})$  (right figure).

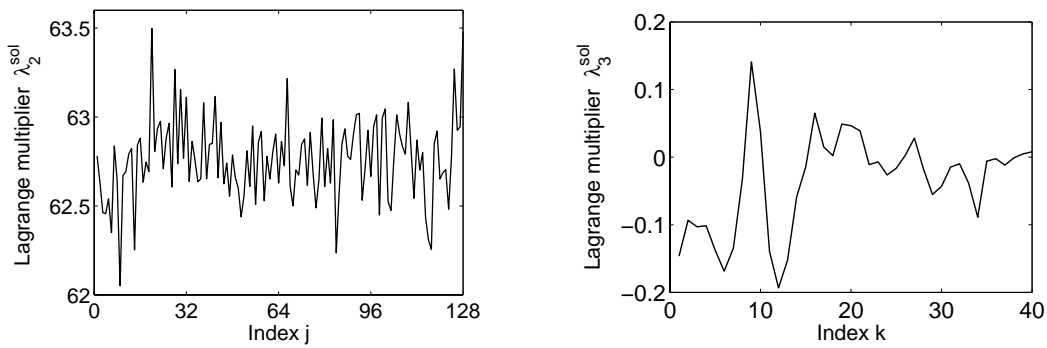


Figure 3. Graph of  $j \mapsto (\lambda_2^{\text{sol}})_j$  (left figure) and graph of  $k \mapsto (\lambda_3^{\text{sol}})_k$  (right figure).

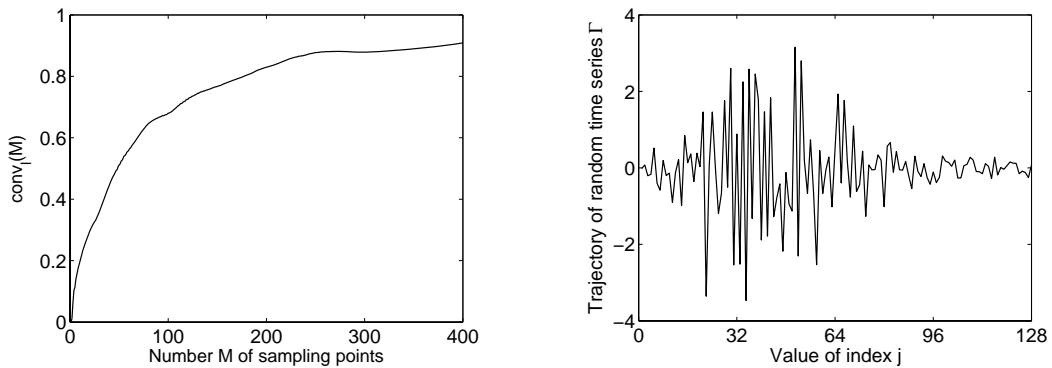
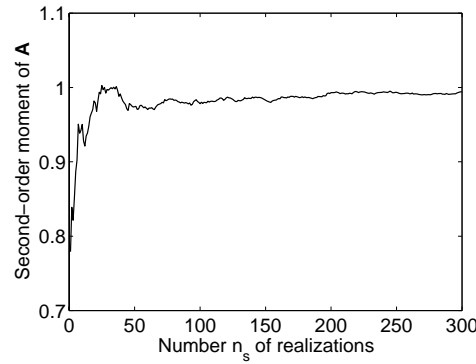
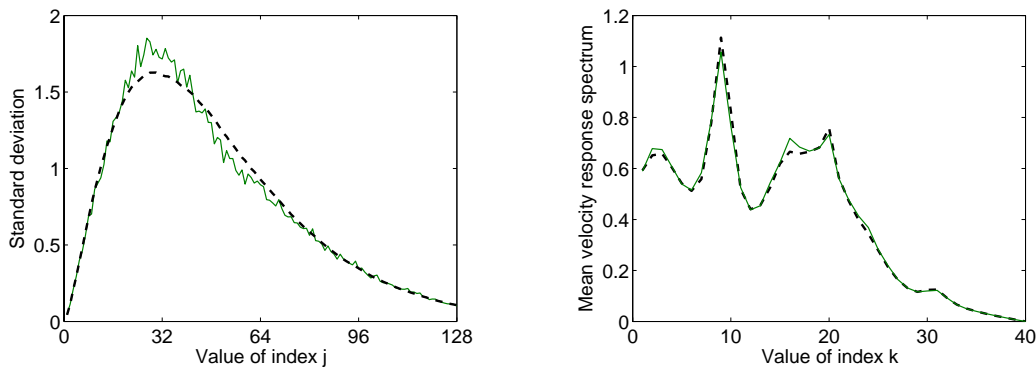


Figure 4. Graph of  $M \mapsto \text{conv}_\ell(M)$  (left figure) and graph of  $j \mapsto \Gamma_j(\theta_\ell)$  for a realization  $\theta_\ell$  (right figure).

Figure 5. Graph of  $n_s \mapsto \text{convMC}(n_s)$ .Figure 6. Graph of  $j \mapsto \sigma_j = E\{\Gamma_j^2\}^{1/2}$  (left figure) and graph of  $k \mapsto \underline{\mathcal{S}}_k$  (right figure). Reference (dashed lines). Estimation with the random generator (solid lines).

(it should be noted that all the graphs are similar for  $\ell = 1, \dots, n_s$ ). This graph allows the value of  $M$  to be estimated in order to obtain a realization  $\mathbf{U}^M(\theta_\ell)$  of the stationary solution of the ISDE. Therefore,  $M$  must be such that the graph be flat that is reasonably true for  $M = 400$ . Therefore Eq. (54) is satisfied in the mean-square sense for  $M = 400$ . Figure 4 (right) displays the corresponding trajectory of the random time series  $\mathbf{\Gamma}$ , that is to say the graph of the realization  $j \mapsto \Gamma_j(\theta_\ell)$  in which  $\mathbf{\Gamma}(\theta_\ell) = \sqrt{N}[\sigma] \mathbf{A}(\theta_\ell)$  with  $\mathbf{A}(\theta_\ell) \simeq \mathbf{U}^M(\theta_\ell)$ .

(ii) Concerning the value of  $n_s$ , Figure 5 shows the graph of  $n_s \mapsto \text{convMC}(n_s) = \frac{1}{n_s} \sum_{\ell=1}^{n_s} \|\mathbf{A}(\theta_\ell)\|_N^2$  which is an estimation of the second-order moment  $E\{\|\mathbf{A}\|_N^2\} = E\{\|\mathbf{B}_{\lambda^{\text{sol}}}\|_N^2\}$  of the random variable  $\|\mathbf{A}\|_N$ .

(iii) Figure 6 shows the estimation of the constraints (standard deviation and mean velocity response spectrum) constructed with the random generator and compares this estimation with the references defined in Figure 1. Figure 6 (left) displays the graph of the standard-deviation function  $j \mapsto \sigma_j$  from  $\{1, \dots, N\}$  into  $\mathbb{R}_+$ . Figure 6 (right) displays the graph of the mean velocity response spectrum  $k \mapsto \underline{\mathcal{S}}_k$  from  $\{1, \dots, \nu\}$  into  $\mathbb{R}_+$ . The comparisons are good

and validate the method proposed. The small fluctuations of the estimation of the standard-deviation function computed by the Monte Carlo method using the random generator can be reduced in increasing the value of  $n_s$ .

8. Construction of a probability model for positive-definite band random matrices

This application is devoted to the construction of a probabilistic model of a band random matrix with values in the set of all the symmetric positive-definite  $(n \times n)$  real matrices  $\mathbb{M}_n^+(\mathbb{R})$ , for which the available information is made of the mean value, the norm and the norm of its inverse are given. With such an available information, if the matrix is not a band matrix but a full matrix, an explicit construction can be performed for any value of the matrix dimension (see [8, 9, 10]). If the matrix is a band matrix, such an explicit construction cannot be carried out and a numerical construction must be done. We then propose hereinafter such a numerical construction.

8.1. Definition of the band random matrix and available information

Let  $[\mathbf{G}]$  be the band random matrix with values in  $\mathbb{M}_n^+(\mathbb{R})$  with  $n = 4$ , for which the band structure is such that

$$[\mathbf{G}] = \begin{bmatrix} G_{11} & G_{12} & 0 & 0 \\ G_{12} & G_{22} & G_{23} & 0 \\ 0 & G_{23} & G_{33} & G_{34} \\ 0 & 0 & G_{34} & G_{44} \end{bmatrix}, \tag{64}$$

The problem is the construction of the probability distribution on  $\mathbb{M}_n^+(\mathbb{R})$  of  $[\mathbf{G}]$  using the MaxEnt principle for which the available information is defined by

$$E\{[\mathbf{G}]\} = [I_n] \quad , \quad \frac{E\{\|[\mathbf{G}] - [I_n]\|_F^2\}}{\|[I_n]\|_F^2} = \delta^2 < +\infty \quad , \quad E\{\|[\mathbf{G}]^{-1}\|_F^2\} = \alpha < +\infty \quad , \tag{65}$$

in which  $[I_n]$  is the identity  $(n \times n)$  matrix,  $\|[\text{Mat}]\|_F^2 = \text{tr}([\text{Mat}]^T [\text{Mat}])$  is the square of the Frobenius norm of the real matrix  $[\text{Mat}]$ ,  $\delta = 0.35$  which controls the dispersion of random matrix  $[\mathbf{G}]$  and  $\alpha = 5.6$  which must be a positive and finite real number. The first equation shows that  $[\mathbf{G}]$  is not a centered random variable and its mean value is equal to the identity matrix. The second equation means that  $[\mathbf{G}]$  is a second-order random variable. By construction, band random matrix  $[\mathbf{G}]$  belongs to  $\mathbb{M}_n^+(\mathbb{R})$  almost surely. Therefore,  $[\mathbf{G}]^{-1}$  exists almost surely but, in general, is not a second-order random variable that is to say,  $E\{\|[\mathbf{G}]^{-1}\|_F^2\} = +\infty$ . This is the reason why the third equation is considered as an available information.

Since  $[\mathbf{G}]$  is positive definite almost surely, random matrix  $[\mathbf{G}]$  can be written (Choleski decomposition) as

$$[\mathbf{G}] = [\mathbf{L}]^T [\mathbf{L}] \quad , \quad [\mathbf{L}] = \begin{bmatrix} A_1^2 & A_2 & 0 & 0 \\ 0 & A_3^2 & A_4 & 0 \\ 0 & 0 & A_5^2 & A_6 \\ 0 & 0 & 0 & A_7 \end{bmatrix}, \tag{66}$$

in which  $\mathbf{A} = (A_1, \dots, A_N)$  is a  $\mathbb{R}^N$ -valued random vector with  $N = 7$ . Clearly, Eq. (66) defines a unique nonlinear deterministic mapping  $\mathbf{a} \mapsto [\mathcal{G}(\mathbf{a})]$  from  $\mathbb{R}^N$  into  $\mathbb{M}_n^+(\mathbb{R})$  such that  $[\mathbf{G}] = [\mathcal{G}(\mathbf{A})]$  and a unique nonlinear deterministic mapping  $\mathbf{a} \mapsto \mathbf{e}(\mathbf{a})$  from  $\mathbb{R}^N$  into  $\mathbb{R}^N$  such that  $(G_{11}, G_{12}, G_{22}, G_{23}, G_{33}, G_{34}, G_{44}) = \mathbf{e}(\mathbf{A})$ . The problem above is then equivalent to the construction of the probability distribution on  $\mathbb{R}^N$  of the random vector  $\mathbf{A}$  using the MaxEnt principle for which the available information is deduced from Eq. (65) and can be rewritten as

$$E\{\mathbf{e}(\mathbf{A})\} = \mathbf{e} \in \mathbb{R}^N \quad , \quad E\{\|[\mathcal{G}(\mathbf{A})]\|_F^2\} = n(\delta^2 + 1) \quad , \quad E\{\|[\mathcal{G}(\mathbf{A})]^{-1}\|_F^2\} = \alpha < +\infty \quad , \quad (67)$$

in which  $\mathbf{e} = (1, 0, 1, 0, 1, 0, 1) \in \mathbb{R}^N$ .

### 8.2. Defining the unknown Lagrange multipliers and function $\Phi$

Comparing Eq. (67) with Eq. (2), the nonlinear mapping  $\mathbf{a} \mapsto \mathbf{g}(\mathbf{a})$  from  $\mathbb{R}^N$  into  $\mathbb{R}^\mu$  introduced in Section 2 can be written as  $\mathbf{g}(\mathbf{a}) = (\mathbf{e}(\mathbf{a}), \|[\mathcal{G}(\mathbf{a})]\|_F^2, \|[\mathcal{G}(\mathbf{a})]^{-1}\|_F^2) \in \mathbb{R}^N \times \mathbb{R} \times \mathbb{R} = \mathbb{R}^\mu$  in which  $\mu = N + 2 = 9$ . Vector  $\mathbf{f} \in \mathbb{R}^\mu$  introduced in Eq. (2) is written as  $\mathbf{f} = (\mathbf{e}, n(\delta^2 + 1), \alpha)$ . The Lagrange multiplier  $\boldsymbol{\lambda}$  introduced in Section 2 can then be written as  $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_1, \lambda_2, \lambda_3) \in \mathcal{L}_\mu \subset \mathbb{R}^\mu = \mathbb{R}^N \times \mathbb{R} \times \mathbb{R}$ . The admissible set  $\mathcal{L}_\mu$  is such that  $\mathcal{L}_\mu = \mathbb{R}^N \times ]0, +\infty[ \times ]0, +\infty[$ . We use the explicit Euler scheme presented in Section 5.1 for the discretization of the Itô stochastic differential equation. Function  $\Phi(\mathbf{u}, \boldsymbol{\lambda})$  is then defined by Eq. (14) and for all  $\boldsymbol{\lambda}_1$  in  $\mathbb{R}^N$ , for all  $\lambda_2$  and  $\lambda_3$  in  $]0, +\infty[$ , Eqs. (23) to (25) are satisfied.

### 8.3. Computation of the vector-valued Lagrange multipliers

Lagrange multiplier  $\boldsymbol{\lambda}^{\text{sol}} = (\boldsymbol{\lambda}_1^{\text{sol}}, \lambda_2^{\text{sol}}, \lambda_3^{\text{sol}}) \in \mathcal{L}_\mu \subset \mathbb{R}^\mu$  are computed in solving Eq. (10) as explained at the end of Section 2. The probability density function defined by Eq. (6) is then written as

$$p_{\mathbf{A}}(\mathbf{a}) = c_0^{\text{sol}} \exp\{-\langle \boldsymbol{\lambda}_1^{\text{sol}}, \mathbf{e}(\mathbf{a}) \rangle_N - \lambda_2^{\text{sol}} \|[\mathcal{G}(\mathbf{a})]\|_F^2 - \lambda_3^{\text{sol}} \|[\mathcal{G}(\mathbf{a})]^{-1}\|_F^2\} \quad , \quad \forall \mathbf{a} \in \mathbb{R}^N \quad , \quad (68)$$

(i) The interior-reflective Newton method used to solve Eq. (10) is initialized with  $\boldsymbol{\lambda}^0 = (\boldsymbol{\lambda}_1^0, \lambda_2^0, \lambda_3^0) \in \mathcal{L}_\mu \subset \mathbb{R}^\mu$  with  $\boldsymbol{\lambda}_1^0 = \mathbf{1}_N$  and  $\lambda_2^0 = \lambda_3^0 = 2$  where  $\mathbf{1}_N$  is the vector in  $\mathbb{R}^N$  for which all the components are equal to 1. Figure 7 (right) displays the graph of the function  $\text{iter} \mapsto \text{convALG}(\text{iter}) = \|E\{\mathbf{g}(\mathbf{B}_{\boldsymbol{\lambda}(\text{iter})})\} - \mathbf{f}\|_\mu^2$ .

(ii) The mathematical expectation defined by Eq. (11) is estimated by using the Monte Carlo method (see Eq. (55)), the random generator (see Eq. (54)) with the Explicit Euler scheme presented in Section 5.1. For any value of vector  $\boldsymbol{\lambda}$  corresponding to an iteration of the algorithm, parameter  $f_0$  is fixed to 1 and the sampling step defined by Eq. (37) is deduced from a convergence analysis and is written as  $\Delta r = \beta/m$  with  $\beta = 1$  and  $m = 5$ . Concerning the value of  $M$ , Figure 7 (left) displays the graph of the function  $M \mapsto \text{conv}_\ell(M) = \frac{1}{M} \sum_{k=1}^M \|\mathbf{U}^k(\theta_\ell)\|_N^2$  for  $\boldsymbol{\lambda} = \boldsymbol{\lambda}^{\text{sol}}$  and for a realization  $\theta_\ell$  (it should be noted that all the graphs are similar for  $\ell = 1, \dots, n_s$  and for any admissible value of vector  $\boldsymbol{\lambda}$ ). This graph allows the value of  $M$  to be estimated in order to obtain a realization  $\mathbf{U}^M(\theta_\ell)$  of the stationary solution of the ISDE. Therefore,  $M$  must be such that the graph be flat that is reasonably true for  $M = 5000$ . Therefore Eq. (54) is satisfied in the mean-square sense for  $M = 5000$ .

(iii) The value of  $n_s$  has also been deduced from a convergence analysis. For instance, at solution  $\boldsymbol{\lambda} = \boldsymbol{\lambda}^{\text{sol}}$ , Figures 8 (left and right) show the graphs of  $n_s \mapsto \delta(n_s)$  such that  $\frac{1}{n_s} \sum_{\ell=1}^{n_s} \|\mathcal{G}(\mathbf{A}(\theta_\ell))\|_F^2 = n(\delta(n_s)^2 + 1)$  and  $n_s \mapsto \alpha(n_s) = \frac{1}{n_s} \sum_{\ell=1}^{n_s} \|\mathcal{G}(\mathbf{A}(\theta_\ell))^{-1}\|_F^2 = n(\delta(n_s)^2 + 1)$ . An excellent convergence is obtained for  $n_s = 600$  which is the value used in the computation.

(iv) Concerning solution  $\boldsymbol{\lambda}^{\text{sol}}$ , we get  $\lambda_1^{\text{sol}} = (0.7381, 4.1697, 1.2465, -0.9248, 0.8998, 4.2584, 0.8714)$ ,  $\lambda_2^{\text{sol}} = 2.2293$  and  $\lambda_3^{\text{sol}} = 1.7749$ .

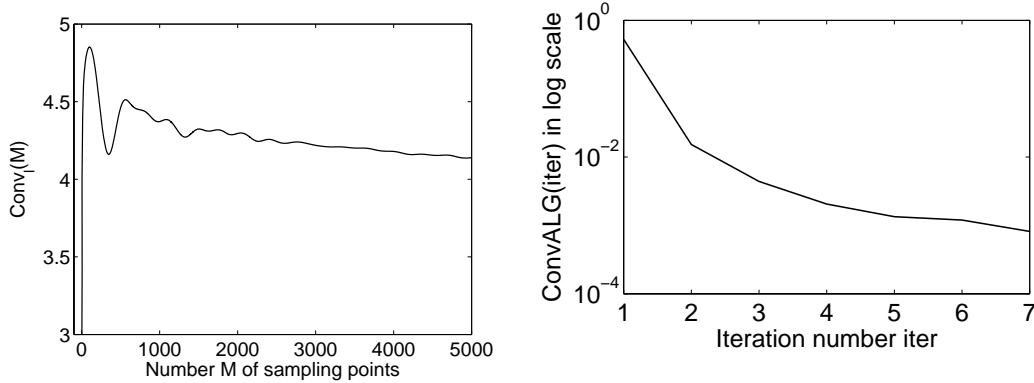


Figure 7. Graph of  $M \mapsto \text{conv}_\ell(M)$  (left figure) and graph of  $\text{iter} \mapsto \text{convALG}(\text{iter})$  (right figure).

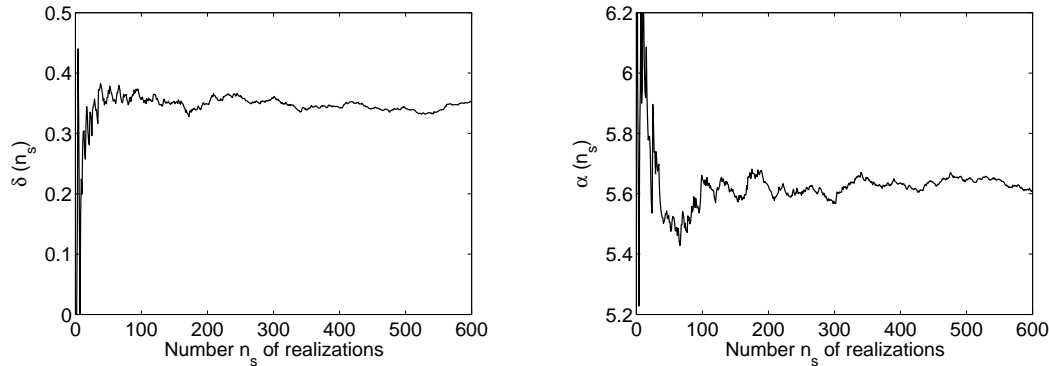


Figure 8. Graph of  $n_s \mapsto \delta(n_s)$  (left figure) and graph of  $n_s \mapsto \alpha(n_s)$  (right figure).

8.4. Estimation of the constraints with the random generator

Solution  $\boldsymbol{\lambda}^{\text{sol}}$  of Eq. (10) being known,  $n_s$  independent realizations of random variable  $\mathbf{A} = \mathbf{B}_{\boldsymbol{\lambda}^{\text{sol}}}$  are constructed using the random generator presented in Section 6.2 (exactly, as we have performed to calculate  $\boldsymbol{\lambda}^{\text{sol}}$  in Section 8.3). The estimations of the mathematical expectations  $E\{\mathbf{g}(\mathbf{A})\}$  of the constraints are calculated by the Monte Carlo method as explained in Section

6.3 (see Eq. (55)) and as performed above in Section 8.3. Computation is then performed with  $f_0 = 1$ ,  $m = 5$ ,  $M = 5000$  and  $n_s = 600$ .

(i) Concerning the estimation of the constraints, we obtain

$$E\{\{\mathbf{G}\}\} \simeq \begin{bmatrix} 1.0001 & 0.0051 & 0 & 0 \\ 0.0051 & 0.9910 & 0.0057 & 0 \\ 0 & 0.0057 & 1.0222 & 0.0077 \\ 0 & 0 & 0.0077 & 1.0056 \end{bmatrix}, \quad (69)$$

which has to be compared to the identity matrix,  $\delta \simeq 0.3529$  which has to be compared to 0.3500 and finally,  $\alpha \simeq 5.6053$  which has to be compared to 5.6000. We then have a good comparison.

## 9. Extension of the theory to the case of a probability density function with any support and application to the Karhunen-Loeve expansion of a Non-Gaussian positive-valued random field

In a first subsection, we show how the previous developments can be used for a probability density function for which its support is not  $\mathbb{R}^N$  but is any part  $\mathcal{A}$  of  $\mathbb{R}^N$ . The second subsection will deal with an application devoted to the construction of the Karhunen-Loeve expansion of a subclass of Non-Gaussian positive-valued random fields for which the general class has been introduced and analyzed in [37].

### 9.1. Extension of the theory to a probability density function with any support

Let  $\mathcal{A}$  be any part of  $\mathbb{R}^N$ ,  $\mathbf{x} = (x_1, \dots, x_N)$  be any vector in  $\mathbb{R}^N$  and let  $d\mathbf{x} = dx_1 \dots dx_N$  be the Lebesgue measure on  $\mathbb{R}^N$ . Let  $\mathbf{X} = (X_1, \dots, X_N)$  be a  $\mathbb{R}^N$ -valued second-order random variable for which the probability distribution  $P_{\mathbf{X}}(d\mathbf{x}) = p_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}$  on  $\mathbb{R}^N$  is unknown and is represented by a probability density function  $\mathbf{x} \mapsto p_{\mathbf{X}}(\mathbf{x})$  from  $\mathbb{R}^N$  into  $\mathbb{R}^+$  whose support is  $\mathcal{A} \subset \mathbb{R}^N$  (consequently,  $p_{\mathbf{X}}(\mathbf{x}) = 0$  for all  $\mathbf{x} \notin \mathcal{A}$ ). We then have

$$\text{Supp } p_{\mathbf{X}} = \mathcal{A} \quad , \quad \int_{\mathbb{R}^N} p_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} = \int_{\mathcal{A}} p_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} = 1 \quad . \quad (70)$$

The problem to be solved is the construction of  $p_{\mathbf{X}}$  using the MaxEnt principle for which the constraints associated with the available information (see Eq. (2)) is

$$E\{\mathbf{g}(\mathbf{X})\} = \mathbf{f} \quad , \quad (71)$$

in which  $\mathbf{f} = (f_1, \dots, f_\mu)$  is a given vector in  $\mathbb{R}^\mu$  and where  $\mathbf{x} \mapsto \mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), \dots, g_\mu(\mathbf{x}))$  is a given measurable mapping from  $\mathbb{R}^N$  into  $\mathbb{R}^\mu$ . We then obtain (see Section 2 and Eq. (6)),

$$p_{\mathbf{X}}(\mathbf{x}) = \mathbb{1}_{\mathcal{A}}(\mathbf{x}) c_X^{\text{sol}} \exp(-\langle \boldsymbol{\lambda}^{\text{sol}}, \mathbf{g}(\mathbf{x}) \rangle_\mu) \quad , \quad \forall \mathbf{x} \in \mathbb{R}^N \quad , \quad (72)$$

in which  $\mathbb{1}_{\mathcal{A}}(\mathbf{x}) = 1$  if  $\mathbf{x} \in \mathcal{A}$  and  $\mathbb{1}_{\mathcal{A}}(\mathbf{x}) = 0$  if  $\mathbf{x} \notin \mathcal{A}$  and where  $c_X^{\text{sol}} = \exp(-\lambda_0^{\text{sol}})$  is the constant of normalization calculated with Eq. (70). It should be noted that we cannot directly use the previous theory in introducing the function  $\Phi(\mathbf{u}, \boldsymbol{\lambda}) = -\ln(\mathbb{1}_{\mathcal{A}}(\mathbf{u})) + \langle \boldsymbol{\lambda}, \mathbf{g}(\mathbf{u}) \rangle_\mu$  (see Eq. (14)) because the function  $\mathbf{u} \mapsto \|\nabla_{\mathbf{u}} \Phi(\mathbf{u}, \boldsymbol{\lambda})\|_N$  is not a locally bounded function on  $\mathbb{R}^N$  (it is a distribution or a generalized function). We must then proceed differently.

Note that (1) the calculation of the cumulative distribution function (probability distribution) or the calculation of the moments for the random responses of the computational model or (2) the calculation of the left-hand side of Eq. (71) lead us to calculate quantities of the type  $E\{\mathbf{h}(\mathbf{X})\}$  in which  $\mathbf{x} \mapsto \mathbf{h}(\mathbf{x})$  is a given vector-valued function defined on  $\mathbb{R}^N$ . Let  $p_{\mathbf{A}}$  be the probability density function of the  $\mathbb{R}^N$ -valued random variable  $\mathbf{A} = (A_1, \dots, A_N)$  defined in Section 2 and constructed using the method presented in Sections 3 to 6. We then have

$$E\{\mathbf{h}(\mathbf{X})\} = \int_{\mathbb{R}^N} \mathbf{h}(\mathbf{x}) \mathbb{1}_{\mathcal{A}}(\mathbf{x}) c_X^{\text{sol}} \exp(-\langle \boldsymbol{\lambda}^{\text{sol}}, \mathbf{g}(\mathbf{x}) \rangle_{\mu}) d\mathbf{x} \quad , \quad (73)$$

which can be rewritten as

$$E\{\mathbf{h}(\mathbf{X})\} = \frac{c_X^{\text{sol}}}{c_0^{\text{sol}}} \int_{\mathbb{R}^N} \mathbf{h}(\mathbf{a}) \mathbb{1}_{\mathcal{A}}(\mathbf{a}) p_{\mathbf{A}}(\mathbf{a}) d\mathbf{a} \quad , \quad (74)$$

in which  $p_{\mathbf{A}}$  is defined by Eq. (6). Taking  $\mathbf{h}(\mathbf{a}) = 1$ , it can be deduced that  $c_X^{\text{sol}}/c_0^{\text{sol}} = 1/E\{\mathbb{1}_{\mathcal{A}}(\mathbf{A})\}$ . Finally,  $E\{\mathbf{h}(\mathbf{X})\}$  can be calculated by

$$E\{\mathbf{h}(\mathbf{X})\} = \frac{E\{\mathbf{h}(\mathbf{A}) \mathbb{1}_{\mathcal{A}}(\mathbf{A})\}}{E\{\mathbb{1}_{\mathcal{A}}(\mathbf{A})\}} \quad , \quad (75)$$

where the mathematical expectations in the right-hand side of Eq. (75) are calculated by the theory developed in Sections 2 to 6.

*Remark.* It should be noted that the proposed method consists in solving the problem on an unconstrained support and then in restricting the solution on the desired support  $\mathcal{A}$ , rescaling the probability density function. This implies that the solution on the unrestricted support exists, that is to say that the function  $\mathbf{x} \mapsto \mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), \dots, g_{\mu}(\mathbf{x}))$  from  $\mathbb{R}^N$  into  $\mathbb{R}^{\mu}$  be such that the conditions defined by Eqs. (23) to (25) hold. This is the case for the fundamental application presented below for which the proposed method is very efficient. Nevertheless, such a solution may not exist. In order to explain the difficulties and to give a few ideas to construct the solution in such a case, we then propose to analyze below the "extreme" case for which the probability density function is uniform on a compact support  $\mathcal{A}$ . In this case  $\mathbf{g}$  is zero and Eq. (72) yields  $p_{\mathbf{X}}(\mathbf{x}) = \mathbb{1}_{\mathcal{A}}(\mathbf{x}) c_X^{\text{sol}}$  for all  $\mathbf{x}$  in  $\mathbb{R}^N$ . Clearly, the solution  $p_{\mathbf{A}}$  on the unrestricted support  $\mathbb{R}^N$  does not exist. To analyze such a case, a regularization  $\mathbf{g}^{\varepsilon}$  of the function  $\mathbf{g}$  depending on a parameter  $\varepsilon > 0$  can be introduced such that, for all  $\varepsilon > 0$ , (1) the function  $\mathbf{x} \mapsto \mathbf{g}^{\varepsilon}(\mathbf{x})$  is differentiable on  $\mathbb{R}^N$ , (2) for all  $\mathbf{x}$  in  $\mathcal{A}$ , we have  $\mathbf{g}^{\varepsilon}(\mathbf{x}) = 0$ , (3) the conditions defined by Eqs. (23) to (25) hold for  $\mathbf{g}^{\varepsilon}$  and (4) the support of  $\mathbf{g}^{\varepsilon}(\mathbf{x})$ , which is  $\mathbb{R}^N$  tends to the compact support  $\mathcal{A}$  when  $\varepsilon$  goes to zero. Such a regularization is not always easy to construct, but if it is possible, then this method is very efficient. In order to illustrate such a method, let us consider the simple case for which  $N = 1$  and  $\mathcal{A} = [-1, +1]$ . We can choose the following regularization:  $g^{\varepsilon}(x) = 0.5 \varepsilon^{-2}(x-1)^2$  if  $x > 1$ ,  $g^{\varepsilon}(x) = 0$  if  $x \in \mathcal{A}$  and  $g^{\varepsilon}(x) = 0.5 \varepsilon^{-2}(x+1)^2$  if  $x < -1$ . In this case,  $x \mapsto dg^{\varepsilon}(x)/dx$  is a continuous function on  $\mathbb{R}$  such that  $dg^{\varepsilon}(x)/dx = \varepsilon^{-2}(x-1)$  if  $x > 1$ ,  $dg^{\varepsilon}(x)/dx = 0$  if  $x \in \mathcal{A}$  and  $dg^{\varepsilon}(x)/dx = \varepsilon^{-2}(x+1)$  if  $x < -1$ . Then for  $\varepsilon$  sufficiently small (but not too small), the method would be very efficient.

## 9.2. Application to the Karhunen-Loeve expansion of a Non-Gaussian positive-valued random field

We consider the following computational model resulting from the finite element discretization of an elliptic boundary value problem (for instance, a linear elastostatic problem on a bounded



3D domain) and written as

$$[\mathbf{K}] \mathbf{Y} = \mathbf{r} \quad , \quad [\mathbf{K}] = \sum_{j=1}^n Z_j [k^j] \quad , \quad (76)$$

in which  $\mathbf{Z} = (Z_1, \dots, Z_n)$  is the spatial sampling of a positive-valued random field (for instance, the Young modulus in linear isotropic elasticity for a heterogeneous material), where  $[k^1], \dots, [k^n]$  are  $n$  given symmetric real matrices, where  $\mathbf{r}$  is a given vector and where  $\mathbf{Y}$  is the unknown random vector. The random matrix  $[\mathbf{K}]$  is assumed to be positive-definite almost surely (a.s) and consequently,  $\mathbf{Y} = [\mathbf{K}]^{-1} \mathbf{r}$  almost surely. The mean value  $\underline{\mathbf{z}} = (\underline{z}_1, \dots, \underline{z}_n)$  of  $\mathbf{Z}$  is  $\underline{\mathbf{z}} = E\{\mathbf{Z}\}$ . Since  $\mathbf{Z}$  corresponds to the sampling of a positive-valued random field, then for all  $j \in \{1, \dots, n\}$ , it is assumed that  $\underline{z}_j > 0$  and  $Z_j > 0$  almost surely. We then introduce the normalized random vector  $\mathbf{G} = (G_1, \dots, G_n)$  with values in  $\mathbb{R}^n$  such that  $Z_j = \underline{z}_j G_j$  for all  $j \in \{1, \dots, n\}$ . Therefore, we have  $G_j > 0$  almost surely for all  $j \in \{1, \dots, n\}$ . The probability distribution of random vector  $\mathbf{G}$  must be such that  $\mathbf{Y}$  is a second-order random variable, *i.e.*  $E\{\|\mathbf{Y}\|^2\} = c_Y < +\infty$ . Using similar developments to those given in Ref. [37] and taking into account that for all  $g_1 > 0, \dots, g_n > 0$ , we have

$$\left(\max\left\{\frac{1}{g_1}, \dots, \frac{1}{g_n}\right\}\right)^2 \leq \frac{1}{g_1^2} + \dots + \frac{1}{g_n^2} \quad , \quad (77)$$

it can be proven that  $E\{\|\mathbf{Y}\|^2\} = c_Y < +\infty$  if the following inequality holds,

$$E\left\{\frac{1}{G_1^2} + \dots + \frac{1}{G_n^2}\right\} = c_G < +\infty \quad . \quad (78)$$

The Karhunen-Loeve expansion  $\mathbf{G}^N$  at order  $N$  of the random field  $\mathbf{G}$  yields the following approximation for the  $\mathbb{R}^n$ -valued random vector  $\mathbf{G}$ ,

$$\mathbf{G}^N = \underline{\mathbf{G}} + \sum_{\alpha=1}^N \sqrt{v_\alpha} X_\alpha \boldsymbol{\varphi}^\alpha \quad , \quad (79)$$

in which  $\underline{\mathbf{G}} = (\underline{G}_1, \dots, \underline{G}_n)$  with  $\underline{G}_j = 1$  for all  $j$ . In Eq. (79),  $\boldsymbol{\varphi}^1, \dots, \boldsymbol{\varphi}^N$  are the orthonormal eigenvectors ( $\langle \boldsymbol{\varphi}^\alpha, \boldsymbol{\varphi}^\beta \rangle = \delta_{\alpha\beta}$ ) associated with the  $N$  largest eigenvalues  $v_1 > v_2 > \dots > v_N$  of the covariance matrix  $[C_{\mathbf{G}}] = E\{(\mathbf{G} - \underline{\mathbf{G}})(\mathbf{G} - \underline{\mathbf{G}})^T\}$  of  $\mathbf{G}$  which is assumed to be given.

For the construction by the MaxEnt principle of the probability distribution on  $\mathbb{R}^N$  of the second-order  $\mathbb{R}^N$ -valued random variable  $\mathbf{X} = (X_1, \dots, X_N)$ , the available information is the following,

$$\text{Supp } p_{\mathbf{X}} = \mathcal{A} \subset \mathbb{R}^N \quad , \quad (80)$$

$$E\{\mathbf{X}\} = 0 \quad , \quad (81)$$

$$E\{\mathbf{X}\mathbf{X}^T\} = [I_N] \quad , \quad (82)$$

$$E\{s(\mathbf{X})\} = \kappa < +\infty \quad , \quad (83)$$

in which for all  $\mathbf{x} = (x_1, \dots, x_N) \in \mathcal{A}$ , we have

$$s(\mathbf{x}) = \sum_{j=1}^n (\underline{G}_j + \sum_{\alpha=1}^N \sqrt{v_\alpha} x_\alpha \varphi_j^\alpha)^{-2} \quad . \quad (84)$$

The support  $\mathcal{A}$  of  $p_{\mathbf{x}}$  is defined by

$$\mathcal{A} = \{\mathbf{x} \in \mathbb{R}^N \text{ such that } \forall j \in \{1, \dots, n\} \text{ , } \underline{G}_j + \sum_{\alpha=1}^N \sqrt{v_{\alpha}} x_{\alpha} \varphi_j^{\alpha} > 0\} \text{ .} \quad (85)$$

The nonlinear mapping  $\mathbf{x} \mapsto \mathbf{g}(\mathbf{x})$  from  $\mathbb{R}^N$  into  $\mathbb{R}^{\mu}$  introduced in Section 2 can be written as

$$\mathbf{g}(\mathbf{x}) = (\mathbf{x}, \mathbf{e}(\mathbf{x}), \frac{1}{\kappa} s(\mathbf{x})) \text{ ,} \quad (86)$$

in which  $\mu = N + N(N + 1)/2 + 1$  and where  $\mathbf{e}(\mathbf{x})$  is a vector in  $\mathbb{R}^{N(N+1)/2}$  constituted of the elements (stored column rise) of the upper triangular part (including the diagonal) of the matrix  $\mathbf{x}\mathbf{x}^T$ . The vector  $\mathbf{f}$  in  $\mathbb{R}^{\mu}$  introduced in Eq. (71) is then written as

$$\mathbf{f} = (0_N, \underline{\mathbf{e}}, 1) \text{ ,} \quad (87)$$

in which  $0_N = (0, \dots, 0) \in \mathbb{R}^N$  and where  $\underline{\mathbf{e}} = E\{\mathbf{e}(\mathbf{X})\}$  is the vector in  $\mathbb{R}^{N(N+1)/2}$  constituted of the elements (stored column rise) of the upper triangular part (including the diagonal) of the matrix  $[I_N]$  (and consequently, constituted of 0 and 1).

The Lagrange multiplier  $\boldsymbol{\lambda}$  introduced in Section 2 can then be written as  $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \lambda_3) \in \mathcal{L}_{\mu} \subset \mathbb{R}^{\mu} = \mathbb{R}^N \times \mathbb{R}^{N(N+1)/2} \times \mathbb{R}$ . The semi-implicit scheme defined in Subsection 5.2 is used to discretize the Itô stochastic differential equation. Function  $\Phi(\mathbf{u}, \boldsymbol{\lambda})$  is defined by Eq. (44) in which matrix  $[K_{\boldsymbol{\lambda}}]$  is such that, for all  $\mathbf{u}$  in  $\mathbb{R}^N$ ,

$$\langle \boldsymbol{\lambda}_2, \mathbf{e}(\mathbf{u}) \rangle_{N(N+1)/2} = \frac{1}{2} \langle [K_{\boldsymbol{\lambda}}] \mathbf{u}, \mathbf{u} \rangle_N \text{ ,} \quad (88)$$

and vector  $\mathbf{L}_{\text{NL}}(\mathbf{u})$  is given by

$$\mathbf{L}_{\text{NL}}(\mathbf{u}) = -\boldsymbol{\lambda}_1 - \frac{\lambda_3}{\kappa} \nabla_{\mathbf{u}} s(\mathbf{u}) \text{ .} \quad (89)$$

The gradient  $\nabla_{\mathbf{u}} s(\mathbf{u}) = (\frac{\partial s(\mathbf{u})}{\partial u_1}, \dots, \frac{\partial s(\mathbf{u})}{\partial u_N})$  at point  $\mathbf{u} = (u_1, \dots, u_N)$  is such that, for all  $\alpha \in \{1, \dots, N\}$ ,

$$\frac{\partial s(\mathbf{u})}{\partial u_{\alpha}} = -2 \sum_{j=1}^n \frac{\sqrt{v_{\alpha}} \varphi_j^{\alpha}}{(\underline{G}_j + \sum_{\beta=1}^N \sqrt{v_{\beta}} u_{\beta} \varphi_j^{\beta})^3} \text{ .} \quad (90)$$

It should be noted that the subset of the admissible values of  $\boldsymbol{\lambda}_2$  is such that  $[K_{\boldsymbol{\lambda}}]$  is a positive-definite matrix and the subset of the admissible values of  $\lambda_3$  is  $]0, +\infty[$ . Since the value of  $\kappa$  is arbitrary,  $\lambda_3$  can be fixed to a given value denoted by  $\lambda_3^{\text{sol}}$ . The Lagrange multipliers  $\boldsymbol{\lambda}_1^{\text{sol}}$  and  $\boldsymbol{\lambda}_2^{\text{sol}}$  are computed in solving Eq. (71) by using Eq. (75). The probability density function defined by Eq. (6) is written as

$$p_{\mathbf{A}}(\mathbf{a}) = c_0^{\text{sol}} \exp\{-\langle \boldsymbol{\lambda}_1^{\text{sol}}, \mathbf{a} \rangle_N - \langle \boldsymbol{\lambda}_2^{\text{sol}}, \mathbf{e}(\mathbf{a}) \rangle_{N(N+1)/2} - \lambda_3^{\text{sol}} s(\mathbf{a})\} \text{ , } \forall \mathbf{a} \in \mathbb{R}^N \text{ ,} \quad (91)$$

For the numerical application,  $n = 100$  and the values of  $\sqrt{v_1}, \dots, \sqrt{v_{20}}$  are respectively, 2.78, 1.18, 0.82, 0.48, 0.38, 0.28, 0.23, 0.21, 0.18, 0.17, 0.15, 0.14, 0.13, 0.11, 0.106, 0.102, 0.09, 0.088, 0.081, 0.080. Figure 9 displays the graph of the function  $N \mapsto \text{error}(N) = E\{\|\mathbf{G} - \mathbf{G}^N\|^2\} / E\{\|\mathbf{G}\|^2\} = (\text{tr}\{[C_{\mathbf{G}}]\} - \sum_{\alpha=1}^N v_{\alpha}) / \text{tr}\{[C_{\mathbf{G}}]\}$  which allows the error to be measured when  $\mathbf{G}$  is replaced by the K-L expansion  $\mathbf{G}^N$ . It can be seen that  $N = 10$  implies

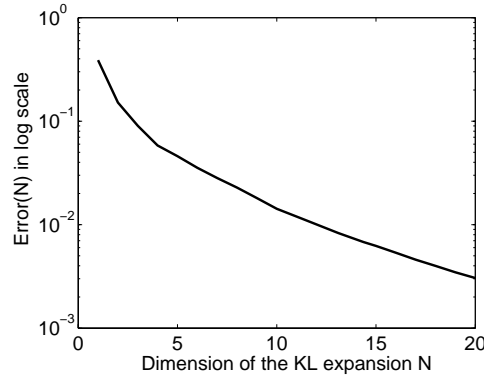


Figure 9. Graph of  $N \mapsto \text{error}(N)$  showing the convergence of Karhunen-Loeve expansion.

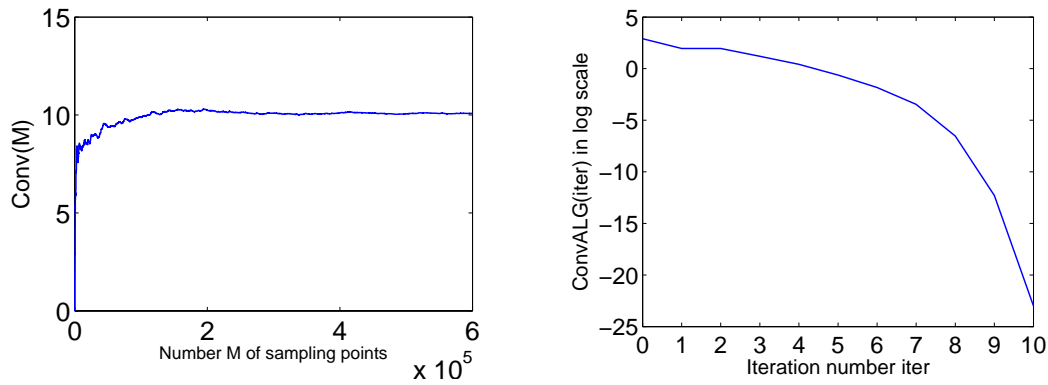


Figure 10. Graph of  $M \mapsto \text{conv}(M)$  (left figure) and graph of  $\text{iter} \mapsto \text{convALG}(\text{iter})$  (right figure).

a reasonable relative error. The mathematical expectations in Eq. (75) are estimated by using the ergodic theory (see Eq. (53)) and the semi-implicit scheme is used to construct a realization  $\{\mathbf{U}^k(\theta), k = 1, \dots, M\}$  of the random variables  $\{\mathbf{U}^k, k = 1, \dots, M\}$ . The sampling step defined by Eq. (37) is taken as  $\Delta r = 0.01$ . Parameter  $f_0$  is equal to 0.5,  $M_0 = 200,000$  and  $M = 600,000$ . These values of parameters  $\Delta r$ ,  $f_0$ ,  $M_0$  and  $M$  have been derived from a convergence analysis. The value of  $\lambda_3$  is fixed to the value  $\lambda_3^{\text{sol}} = 0.01$  and Eq. (71) is solved with respect to  $(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$  by using the trust-region dogleg algorithm which is a variant of the Powell dogleg method. The initial values used are  $\boldsymbol{\lambda}_1^0 = 0.8 \mathbf{1}_N$  and  $\boldsymbol{\lambda}_2^0 = 0.2 \mathbf{e}$ . Figure 10 (left) displays the graph of the function  $M \mapsto \text{conv}(M)$  showing the convergence of the estimation of  $E\{\|\mathbf{B}_{\tilde{\boldsymbol{\lambda}}^{\text{sol}}}\|_N^2\}$  by using the ergodic method with  $\tilde{\boldsymbol{\lambda}} = (\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$ . Figure 10 (right) displays the graph of the function  $\text{iter} \mapsto \text{convALG}(\text{iter}) = \|E\{\mathbf{g}(\mathbf{B}_{\tilde{\boldsymbol{\lambda}}(\text{iter})})\} - \mathbf{f}\|_{\mu-1}^2$  showing the convergence of trust-region dogleg algorithm in function of the iteration number. Figure 11 (left) compares the graph  $\alpha \mapsto (\boldsymbol{\lambda}_1^{\text{sol}})_\alpha$  of the solution for  $\boldsymbol{\lambda}_1$  with the graph  $\alpha \mapsto (\boldsymbol{\lambda}_1^0)_\alpha$  of the initial value  $\boldsymbol{\lambda}_1^0 = 0.8 \mathbf{1}_N$ . Figure 11 (right) compares the graph  $k \mapsto (\boldsymbol{\lambda}_2^{\text{sol}})_k$  of the solution for  $\boldsymbol{\lambda}_2$  with the

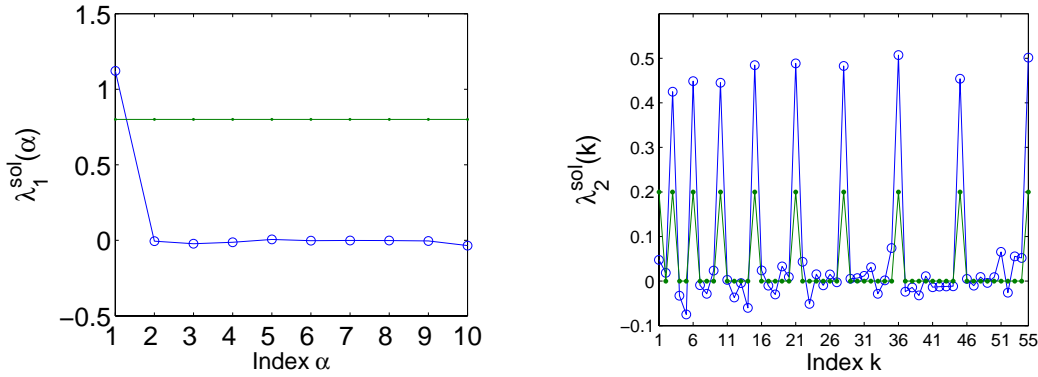


Figure 11. Left figure: Graph of  $\alpha \mapsto (\lambda_1^{\text{sol}})_\alpha$  (circle marker) and graph of  $\alpha \mapsto (\lambda_1^0)_\alpha$  (dot marker). Right figure: Graph of  $k \mapsto (\lambda_2^{\text{sol}})_k$  (circle marker) and graph of  $k \mapsto (\lambda_2^0)_k$  (dot marker).

graph  $k \mapsto (\lambda_2^0)_k$  of the initial value  $\lambda_2^0 = 0.2 \mathbf{e}$ . For the solution obtained  $\lambda^{\text{sol}} = (\lambda_1^{\text{sol}}, \lambda_2^{\text{sol}}, \lambda_3^{\text{sol}})$  with  $\lambda_3^{\text{sol}} = 0.01$ , Eq. (81) and Eq. (82) are satisfied at  $10^{-11}$  for each components and Eq. (83) yields  $\kappa = 140.027$ . The results obtained are thus very good.

## 10. Conclusions

We have proposed a method to effectively construct the probability density function of a random variable in high dimension and for any support of its probability distribution by using the MaxEnt principle. To calculate the integrals of the problem in high dimension and to construct a generator of independent realizations, an alternative algorithm to the Metropolis-Hastings or Gibbs algorithms is proposed. This algorithm is derived from the discretization of an Itô stochastic differential equation for which the stability, the speed of convergence and the transient part can be controlled. The method proposed is validated through three fundamental applications. The first one is a new formulation of the stochastic inverse problem consisting in constructing the probability distribution in high dimension and of its generator for a vector-valued random variable corresponding to an unknown non-stationary random time series (random accelerograms) for which the Velocity Response Spectrum is given. The second one is also a new formulation related to the construction of the probability distribution in high dimension and of its generator for positive-definite band random matrices. Clearly, the method proposed can be used for any sparse random matrix. Finally, we present an extension of the theory when the support of the probability distribution in high dimension of the random variable is not all the space but is any part of the space. The third application is then a new formulation related to the construction of the probability distribution in high dimension for the Karhunen-Loeve expansion of Non-Gaussian positive-valued random fields.

## REFERENCES

1. Schueller GI (editor). A state-of-the-art report on computational stochastic mechanics. *Probabilistic Engineering Mechanics* 1997; **12**(4):197-321.
2. Schueller GI. Computational stochastic mechanics - recent advances. *Computers & Structures* 2001; **79**(22-25):2225-2234.
3. Wiener N. The Homogeneous Chaos. *American Journal of Mathematics* 1938; **60**: 897-936.
4. Spanos PD, Ghanem R. Stochastic finite element expansion for random media. *Journal of Engineering Mechanics, ASCE* 1989; **115**(5):1035-1053.
5. Ghanem R, Spanos PD. *Stochastic Finite Elements: A spectral Approach*. Springer-Verlag: New York, 1991.
6. Ghanem R. Ingredients for a general purpose stochastic finite elements formulation. *Computer Methods in Applied Mechanics and Engineering* 1999; **168**:19-34.
7. Soize C, Ghanem R. Physical systems with random uncertainties : Chaos representation with arbitrary probability measure. *SIAM Journal On Scientific Computing* 2004; **26**(2):395-410.
8. Soize C. A nonparametric model of random uncertainties for reduced matrix models in structural dynamics. *Probabilistic Engineering Mechanics* 2000; **15**(3):277-294.
9. Soize C. Maximum entropy approach for modeling random uncertainties in transient elastodynamics. *Journal of the Acoustical Society of America* 2001; **109**(5):1979-1996.
10. Soize C. Random matrix theory for modeling uncertainties in computational mechanics. *Computer Methods in Applied Mechanics and Engineering* 2005; **194**(12-16):1333-1366.
11. Shannon CE. A mathematical theory of communication. *Bell System Technology Journal* 1948; **27**:379-423 & 623-659.
12. Jaynes ET. Information theory and statistical mechanics. *Physical Review* 1957; **106**(4):620-630 & **108**(2):171-190.
13. Serfling R.J. *Approximation Theorems of Mathematical Statistics*. John Wiley & Sons: New York, 1980.
14. Kapur JN, Kesavan HK. *Entropy Optimization Principles with Applications*. Academic Press: San Diego, 1992.
15. Jumarie G. *Maximum Entropy, Information Without Probability and Complex Fractal*. Kluwer Academic Publishers: Dordrecht, 2000.
16. Sobczyk K, Trebicki J. Maximum entropy principle in stochastic dynamics. *Probabilistic Engineering Mechanics* 1990; **5**(3):102-110.
17. Mead LR., Papanicolaou N. Maximum entropy in the problem of moments. *J. Math. Phys.* 1984; **25**(8):2404-2417.
18. Ciulli S, Mounisif M. On the application of maximum entropy to the moments problem. *J. Math. Phys.* 1991; **32**(7):1717-1719.
19. Tagliani A. Maximum entropy in the Hamburger moments problem. *J. Math. Phys.* 1994; **35**(9):5087-5096.
20. Volpe EV, Baganoff D. Maximum entropy pdfs and the moment problem under near-Gaussian conditions. *Probabilistic Engineering Mechanics* 2003; **18**(3):17-29.
21. Coleman TF, Li Y. On the convergence of reflective newton methods for large-scale nonlinear minimization subject to bounds. *Mathematical Programming* 1994; **67**(2):189-224.
22. Coleman TF, Li Y. An interior, trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on Optimization* 1996; **6**:418-445.
23. Ergorov AD, Sobolevsky PI, Yanovich LA. *Functional Integrals: Approximate Evaluation and Applications*. Kluwer Academic Publisher: Dordrecht, 1993.
24. MacKeown PK. *Stochastic Simulation in Physics*. Springer-Verlag: Singapore, 1997.
25. Hastings WK. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 1970; **109**:57-97.
26. Kaipio J, Somersalo E. *Statistical and Computational Inverse Problems*. Springer-Verlag, New York, 2005.
27. Spall JC. *Introduction to Stochastic Search and Optimization*. John Wiley and Sons: Hoboken, New Jersey, 2003.
28. Geman S, Geman D. Stochastic relaxation, Gibbs distribution and the Bayesian distribution of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1984; Vol PAM I-6:721-741.
29. Mehta ML. *Random Matrices*, Revised and Enlarged Second Edition. Academic Press: New York, 1991.
30. Soize C. *The Fokker-Planck Equation for Stochastic Dynamical Systems and its Explicit Steady State Solutions*. World Scientific: Singapore, 1994.
31. Caughey TK. Derivation and application of the Fokker-Planck equation to discrete nonlinear dynamic systems subjected to white random excitation. *The Journal of The Acoustical Society of America* 1963; **35**(11):1683-1692.
32. Yosida K. *Functional Analysis*. Springer-Verlag: Berlin, 1995.
33. Talay D, Tubaro L. Expansion of the global error for numerical schemes solving stochastic differential equation. *Stoch. Anal. Appl.* 1990; **8**(4):94-120.
34. Kloeden PE, Platen E. *Numerical Solution of Stochastic Differential Equations*. Springer-Verlag: Berlin, 1992.

35. Talay D. Stochastic Hamiltonian system: exponential convergence to the invariant measure and discretization by the implicit Euler scheme. *Markov Processes and Related Fields* 2002; **8**:163-198.
36. Clough RW., Penzien J. *Dynamics of Structures*. McGraw-Hill: New York, 1975.
37. Soize C. Non Gaussian positive-definite matrix-valued random fields for elliptic stochastic partial differential operators. *Computer Methods in Applied Mechanics and Engineering* 2006; **195**:26-64.