Maximum entropy principle for stochastic models in computational sciences

Christian Soize

Maximum Entropy Principle for Stochastic Models in Computational Sciences

C. Soize
Université Paris-Est, 77454 Marne-la-Vallée, France
christian.soize@univ-paris-est.fr

Abstract

The construction of probabilistic models in computational sciences such as 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.

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 the uncertain parameters to be taken into account consists in modeling the uncertain parameters of the computational model by random variables, stochastic processes or random fields. Such an approach has extensively been developed in the two last decades (see for instance Schueller, 2005). In particular, the use of the Gaussian Chaos representation for stochastic processes and random fields (Ghanem and Spanos, 1991) has been used to introduce and to develop useful and efficient tools for analyzing stochastic systems using stochastic finite elements. Model uncertainties introduced by the mathematical-mechanical process used to construct the computational model of complex systems are much more difficult to take into account because the parametric probabilistic approach (considered as a prior probabilistic model of uncertain parameters constructed using the Maximum Entropy Principle or Chaos representations) cannot address such model uncertainties. In this context, a nonparametric probabilistic approach of model uncertainties (considered as a prior probabilistic model for which the uncertain parameters are the operators of the computational model and not its usual parameters) has recently been proposed (Soize, 2000, 2001 and 2005) using the Maximum Entropy (MaxEnt) principle as a possible way to construct a prior probabilistic model which allows model uncertainties to be taken into account.

In general, the response of a computational model is a nonlinear mapping of the uncertain parameters (usual parameters or operators of the problem) and consequently, a complete probability model of these uncertain parameters has to be constructed. This means that the prior 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 prior probability distributions. It is well known that the MaxEnt principle (Shannon, 1948) is certainly one of the most efficient method allowing an explicit construction of such prior 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 to construct the prior probability model and consequently, to construct the probability model when statistical data are either not available or partially available. Note that the MaxEnt principle has been used for different cases and in many applications.

Complete developments of the theory and of the results presented in this paper can be found in (Soize, 2008).

Construction of probability distributions using the MaxEnt principle

In this section, we summarize the MaxEnt principle introduce by Shannon in the context of the Information Theory (Shannon, 1948). Let $\mathbf{a} = (a_1, \ldots, a_N)$ be any vector in $\mathbb{R}^N$. Let $\mathbf{A} = (A_1, \ldots, A_N)$ 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 $d\mathbf{a} = da_1 \ldots da_N$ and which has to verify the following normalization condition,

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

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 Application 3. 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 $\mathbb{R}^\mu$,

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

in which $\mathbf{f} = (f_1, \ldots, f_\mu)$ is a given vector in $\mathbb{R}^\mu$ with $\mu \geq 1$, where $\mathbf{a} \mapsto \mathbf{g}(\mathbf{a}) = (g_1(\mathbf{a}), \ldots, g_\mu(\mathbf{a}))$ 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

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

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 MaxEnt principle consists in constructing $p_{\mathbf{A}} \in \mathcal{C}$ such that

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

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

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

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

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

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), $< \mathbf{x}\,, \mathbf{y} >_\mu = x_1 y_1 + \ldots + x_\mu y_\mu$ is the Euclidean inner product on $\mathbb{R}^\mu$. For $\boldsymbol{\lambda}$ fixed in $\mathcal{L}_\mu$, let $\mathbf{B}_{\boldsymbol{\lambda}}$ 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

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

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

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

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

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

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 $\boldsymbol{\lambda}$,

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

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

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

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). 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 dodleg algorithm which is a variant of the Powell dogleg method described in the work by Coleman and Li (1994 and 1996) (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.

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 $E\{\mathbf{g}(\mathbf{B}_{\boldsymbol{\lambda}})\}$ defined by Eq. (11) can be performed using the Markov Chain Monte Carlo method (MCMC) (Kaipo and Somersalo, 2005; Spall, 2003; MacKeown, 1997). The transition kernel of the homogeneous Markov chain of the MCMC method can be constructed using the Metropolis-Hastings algorithm (Hastings, 1970) or the Gibbs sampling (Geman and Geman, 1984) 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 independent realizations of random variable $\mathbf{B}_{\boldsymbol{\lambda}}$ 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. The construction of the transition kernel is performed by using 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 can be used to estimate $E\{\mathbf{g}(\mathbf{B}_{\boldsymbol{\lambda}})\}$ in order to calculate $\boldsymbol{\lambda}^{\text{sol}}$.

Construction of the probability distribution of $\mathbf{B}_{\boldsymbol{\lambda}}$ *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

$$\Phi(\mathbf{u}, \boldsymbol{\lambda}) = < \boldsymbol{\lambda},\, \mathbf{g}(\mathbf{u}) >_\mu \quad . \tag{12}$$

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

$$d\mathbf{U}(r) = \mathbf{V}(r)\, dr \quad , \tag{13}$$

$$d\mathbf{V}(r) = -\boldsymbol{\nabla}_{\mathbf{u}} \Phi(\mathbf{U}(r), \boldsymbol{\lambda})\, dr - \frac{1}{2} f_0 \mathbf{V}(r)\, dr + \sqrt{f_0}\, d\mathbf{W}(r) \quad , \tag{14}$$

with the initial condition

$$\mathbf{U}(0) = \mathbf{U}_0 \quad , \quad \mathbf{V}(0) = \mathbf{V}_0 \quad a.s. \quad , \tag{15}$$

where $\mathbf{W} = (W_1, \ldots, W_N)$ 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. (14), 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.

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. (13) to (15) has a unique solution defined almost surely for all $r \geq 0$ (see Theorems 4 and 5 in pages 154 to 157 of (Soize, 1994)) 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

$$\mathbf{b}(\mathbf{u}, \mathbf{v}) = \begin{bmatrix} \mathbf{v} \\ -\boldsymbol{\nabla}_{\mathbf{u}}\Phi(\mathbf{u}, \boldsymbol{\lambda}) - \frac{1}{2} f_0 \mathbf{v} \end{bmatrix} \quad , \quad [\,\sigma\,] = \begin{bmatrix} 0_N & 0_N \\ 0_N & f_0 I_N \end{bmatrix} \quad , \tag{16}$$

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 $P_s(d\mathbf{u}, d\mathbf{v}, \boldsymbol{\lambda})$ be an invariant measure of Eqs. (13) and (14). 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 $\mathbf{u} \mapsto \|\boldsymbol{\nabla}_{\mathbf{u}}\Phi(\mathbf{u}, \boldsymbol{\lambda})\|_N$ 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

$$\inf_{\|\mathbf{u}\|_N > R} \Phi(\mathbf{u}, \boldsymbol{\lambda}) \to +\infty \quad \text{if} \quad R \to +\infty \quad , \tag{17}$$

$$\inf_{\mathbf{u} \in \mathbb{R}^N} \Phi(\mathbf{u}, \boldsymbol{\lambda}) = \Phi_{\min} \quad \text{with} \quad \Phi_{\min} \in \mathbb{R} \quad , \tag{18}$$

$$\int_{\mathbb{R}^N} \|\boldsymbol{\nabla}_{\mathbf{u}}\Phi(\mathbf{u}, \boldsymbol{\lambda})\|_N \, p(\mathbf{u}, \boldsymbol{\lambda})\, d\mathbf{u} < +\infty \quad . \tag{19}$$

Under these above hypotheses and using Theorems 4 to 7 in pages 211 to 216 of (Soize, 1994) 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. (13) and (14) have a unique invariant measure $P_s(d\mathbf{u}, d\mathbf{v}, \boldsymbol{\lambda}) = \rho_s(\mathbf{u}, \mathbf{v}, \boldsymbol{\lambda})\, d\mathbf{u}\, d\mathbf{v}$ which is such that

$$\rho_s(\mathbf{u}, \mathbf{v}, \boldsymbol{\lambda}) = c'_{\boldsymbol{\lambda}}\, \exp\{-\frac{1}{2}\|\mathbf{v}\|_N^2 - \Phi(\mathbf{u}, \boldsymbol{\lambda})\} \quad , \quad \forall\, (\mathbf{u}, \mathbf{v}) \in \mathbb{R}^N \times \mathbb{R}^N \quad , \tag{20}$$

in which $c'_{\boldsymbol{\lambda}}$ is the constant of normalization. It should be noted that the conditions defined by Eqs. (17) to (19) 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), (12) and (20), it can be deduced that the probability density function $p(\mathbf{b}, \boldsymbol{\lambda})$ of random variable $\mathbf{B}_{\boldsymbol{\lambda}}$ is related to the invariant measure $\rho_s(\mathbf{u}, \mathbf{v}, \boldsymbol{\lambda})\, d\mathbf{u}\, d\mathbf{v}$ by the following equation,

$$p(\mathbf{b}, \boldsymbol{\lambda}) = \int_{\mathbb{R}^N} \rho_s(\mathbf{b}, \mathbf{v}, \boldsymbol{\lambda})\, d\mathbf{v} \quad , \quad \forall\, \mathbf{b} \in \mathbb{R}^N \quad . \tag{21}$$

Let us consider $\mathbf{U}_0 = \mathbf{u}_0$ and $\mathbf{V}_0 = \mathbf{v}_0$ as initial condition defined by Eq. (15) with $\mathbf{u}_0$ and $\mathbf{v}_0$ two given vectors in $\mathbb{R}^N$. Thus the probability distribution $P_{\mathbf{U}_0,\mathbf{V}_0}(d\mathbf{u}, d\mathbf{v})$ of $(\mathbf{U}_0, \mathbf{V}_0)$ 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. (13) and (14) with the initial condition

$$\mathbf{U}(0)=\mathbf{u}_0 \quad , \quad \mathbf{V}(0)=\mathbf{v}_0 \quad a.s. \quad . \tag{22}$$

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

$$\lim_{r\to+\infty}\mathbf{U}(r)=\mathbf{B}_{\boldsymbol{\lambda}} \quad \text{in probability distribution.} \tag{23}$$

As explained above, the free parameter $f_0>0$ introduced in Eq. (14), 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. (23) and corresponding to the stationary solution associated with the invariant measure.

Random generator of independent realizations

We propose a random generator of n_s independent realizations $\mathbf{B}_{\boldsymbol{\lambda}}(\theta_1),\ \ldots,\mathbf{B}_{\boldsymbol{\lambda}}(\theta_{n_s})$ of random variable $\mathbf{B}_{\boldsymbol{\lambda}}$ whose probability distribution is $p(\mathbf{b},\boldsymbol{\lambda})\,d\mathbf{b}$. For $\theta_1,\ldots,\theta_{n_s}$ in Θ, let $\{\mathbf{W}(r,\theta_1),r\geq 0\},\ldots,\{\mathbf{W}(r,\theta_{n_s}),r\geq 0\}$ be n_s independent realizations of the normalized Wiener stochastic process $\mathbf{W}$ previously introduced. For all ℓ fixed in $\{1,\ldots,n_s\}$, let $\{(\mathbf{U}(r,\theta_\ell),\mathbf{V}(r,\theta_\ell)),r\geq 0\}$ be the unique solution of the following equation (see Eqs. (13) and (14)) defined for all $r\geq 0$ by

$$d\mathbf{U}(r,\theta_\ell)=\mathbf{V}(r,\theta_\ell)\,dr \quad , \tag{24}$$

$$d\mathbf{V}(r,\theta_\ell)=-\boldsymbol{\nabla}_{\mathbf{u}}\Phi(\mathbf{U}(r,\theta_\ell),\boldsymbol{\lambda})\,dr-\frac{f_0}{2}\mathbf{V}(r,\theta_\ell)\,dr+\sqrt{f_0}\,d\mathbf{W}(r,\theta_\ell) \quad , \tag{25}$$

with the initial condition

$$\mathbf{U}(0,\theta_\ell)=\mathbf{u}_0 \quad , \quad \mathbf{V}(0,\theta_\ell)=\mathbf{v}_0 \quad . \tag{26}$$

From Eq. (23), we deduce that each independent realization $\mathbf{B}_{\boldsymbol{\lambda}}(\theta_\ell)$ can be constructed by

$$\mathbf{B}_{\boldsymbol{\lambda}}(\theta_\ell)\simeq\mathbf{U}(r,\theta_\ell) \quad \text{for } r \text{ sufficiently large.} \tag{27}$$

Estimation of mathematical expectations

We propose two estimations of the mathematical expectation $E\{\mathbf{g}(\mathbf{B}_{\boldsymbol{\lambda}})\}$ 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 θ, let $\{\mathbf{U}(r,\theta),r\geq 0\}$ be the solution of Eqs. (24) to (26). Then using the ergodic theorem (Yosida, 1995), we can estimate $E\{\mathbf{g}(\mathbf{B}_{\boldsymbol{\lambda}})\}$ by

$$E\{\mathbf{g}(\mathbf{B}_{\boldsymbol{\lambda}})\}=\lim_{R\to+\infty}\frac{1}{R}\int_0^R\mathbf{g}(\mathbf{U}(r,\theta))\,dr \quad . \tag{28}$$

(ii) Use of the Monte Carlo method. Let $\mathbf{B}_{\boldsymbol{\lambda}}(\theta_1), \ldots, \mathbf{B}_{\boldsymbol{\lambda}}(\theta_{n_s})$ be n_s independent realizations of random variable $\mathbf{B}_{\boldsymbol{\lambda}}$ constructed using the random generator previously presented. Then the mathematical expectation $E\{\mathbf{g}(\mathbf{B}_{\boldsymbol{\lambda}})\}$ can be estimated by

$$E\{\mathbf{g}(\mathbf{B}_{\boldsymbol{\lambda}})\} = \lim_{n_s \to +\infty} \frac{1}{n_s} \sum_{\ell=1}^{n_s} \mathbf{g}(\mathbf{B}_{\boldsymbol{\lambda}}(\theta_\ell)) \quad . \tag{29}$$

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 (Soize, 2008). Since $\mathbf{A} = \mathbf{B}_{\boldsymbol{\lambda}^{\text{sol}}}$ for the convergence in probability distribution, we then propose to use the Monte Carlo method which yields the following estimation

$$E\{\mathbf{Y}\} = \lim_{n_s \to +\infty} \frac{1}{n_s} \sum_{\ell=1}^{n_s} \mathbf{q}(\mathbf{B}_{\boldsymbol{\lambda}^{\text{sol}}}(\theta_\ell)) \quad . \tag{30}$$

Discretization of the Itô stochastic differential equation

The discretization of the Itô stochastic differential equation (ISDE) defined by Eqs. (13) to (15) 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 $\mathbb{R}^N$ must be carried out. We then have to construct an approximation of the solution $\{(\mathbf{U}(r), \mathbf{V}(r)), r \geq 0\}$ of this ISDE. Two integration schemes are proposed. The first one is an explicit Euler scheme (Talay and Taburo, 1990; Kloeden and Platen, 1992) and the second one is a semi-implicit scheme (Kloeden and Platen, 1992; Soize, 1994; Talay, 2002). We cannot give here the details of these integration schemes and we refer the reader to (Soize, 2008).

Let m and M two integers such that $m < M$. The Itô stochastic differential equation is solved on the finite interval $[0\,, (M-1)\,\Delta r]$ in which Δr is the sampling step of the continous index parameter r. The integration scheme is based on the use of the M sampling points r_k such that $r_k = (k-1)\,\Delta r$ for $k = 1, \ldots, M$ with $\Delta r = \beta/m$ in which β is a given positive real number. Consequently, the two parameters for studying the convergence of the approximation will be m and M. We then introduce the following notation $\mathbf{U}^k = \mathbf{U}(r_k)$, $\mathbf{V}^k = \mathbf{V}(r_k)$ and $\mathbf{W}^k = \mathbf{W}(r_k)$ for all k in $\{1, \ldots, M\}$.

Estimation of the mathematical expectations and random generator of independent realizations

Estimation of the mathematical expectation using ergodic method

For θ in Θ, let $\{\mathbf{U}^k(\theta), k = 1, \ldots, M\}$ be any realization of the family of vector-valued random variables $\{\mathbf{U}^k, k = 1, \ldots, M\}$ calculated by using the discretized version of Eqs. (13) to (15). From Eq. (28), for m and M_0 fixed, and for M sufficiently large, $E\{\mathbf{g}(\mathbf{B}_{\boldsymbol{\lambda}})\}$ can be estimated by

$$E\{\mathbf{g}(\mathbf{B}_{\boldsymbol{\lambda}})\} \simeq \frac{1}{M - M_0 + 1} \sum_{k=M_0}^{M} \mathbf{g}(\mathbf{U}^k(\theta)) \quad . \tag{31}$$

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

Random generator of independent realizations

For all $\boldsymbol{\lambda}$ (or for $\boldsymbol{\lambda} = \boldsymbol{\lambda}_{\text{sol}}$ and then $\mathbf{A} = \mathbf{B}_{\boldsymbol{\lambda}^{\text{sol}}}$) and for all ℓ in $\{1, \ldots, n_s\}$, let $\{\mathbf{U}^k(\theta_\ell), k = 1, \ldots, M\}$ be n_s independent realizations of the family of vector-valued random variables $\{\mathbf{U}^k, k = 1, \ldots, M\}$ calculated with the discretized version of Eqs. (24) to (26). From Eq. (23) and for m and M sufficiently large, we can write

$$\mathbf{B}_{\boldsymbol{\lambda}}(\theta_\ell) \simeq \mathbf{U}^M(\theta_\ell) \quad , \quad \forall\, \ell \in \{1, \ldots, n_s\} \quad . \tag{32}$$

Consequently, $\mathbf{B}_{\boldsymbol{\lambda}}(\theta_1), \ldots, \mathbf{B}_{\boldsymbol{\lambda}}(\theta_{n_s})$ are n_s independent realizations of random variable $\mathbf{B}_{\boldsymbol{\lambda}}$ constructed using Eq. (32).

Estimation of the mathematical expectations using the Monte Carlo method

For all $\boldsymbol{\lambda}$ (or for $\boldsymbol{\lambda} = \boldsymbol{\lambda}_{\text{sol}}$), let $\mathbf{B}_{\boldsymbol{\lambda}}(\theta_1), \ldots, \mathbf{B}_{\boldsymbol{\lambda}}(\theta_{n_s})$ be n_s independent realizations of random variable $\mathbf{B}_{\boldsymbol{\lambda}}$ constructed with the random generator previously introduced. Then, from Eqs. (29) and (30) and for n_s sufficiently large, it can be deduced that the mathematical expectations $E\{\mathbf{g}(\mathbf{B}_{\boldsymbol{\lambda}})\}$ and $E\{\mathbf{Y}\}$ can be estimated by

$$E\{\mathbf{g}(\mathbf{B}_{\boldsymbol{\lambda}})\} \simeq \frac{1}{n_s} \sum_{\ell=1}^{n_s} \mathbf{g}(\mathbf{B}_{\boldsymbol{\lambda}}(\theta_\ell)) \quad , \quad E\{\mathbf{Y}\} \simeq \frac{1}{n_s} \sum_{\ell=1}^{n_s} \mathbf{q}(\mathbf{B}_{\boldsymbol{\lambda}_{\text{sol}}}(\theta_\ell)) \quad . \tag{33}$$

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.

Application 1. Probability model for a non-stationary 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 non-stationary 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. We then propose a new formulation to solve this stochastic inverse problem.

Definition of the time series and its available information

Let $\{\Gamma_1, \ldots, \Gamma_N\}$ be a real-valued random time series in which $N = 128$. Let $\boldsymbol{\Gamma} = (\Gamma_1, \ldots, \Gamma_N)$ 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 $E\{\mathbf{\Gamma}\} = 0 \in \mathbb{R}^N$ is zero.

(ii) For all j in $\{1, \ldots, N\}$, since the random variable Γ_j is centered, the variance of Γ_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 $\{1, \ldots, N\}$ into $\mathbb{R}+$.

(iii) Let $\{\mathbb{S}_k(\mathbf{\Gamma}), k = 1, \ldots, \nu\}$ be the random VRS with $\nu = 40$ in which index k is associated with the given frequencies $\omega_1, \ldots, \omega_\nu$ and where $\boldsymbol{\gamma} \mapsto \boldsymbol{\mathcal{S}}(\boldsymbol{\gamma}) = (\mathbb{S}_1(\boldsymbol{\gamma}), \ldots, \mathbb{S}_\nu(\boldsymbol{\gamma}))$ is a given nonlinear mapping from $\mathbb{R}^N$ into $\mathbb{R}^\nu$ such that

$$\mathbb{S}_k(\boldsymbol{\gamma}) = \omega_k \, \max\{|x_1^k|, \ldots, |x_N^k|\} \quad , \quad x_j^k = \sum_{j'=1}^N [B^k]_{jj'} \, \gamma_{j'} \quad , \tag{34}$$

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

$$\boldsymbol{\mathcal{S}}(\mathbf{\Gamma}) = (\mathbb{S}_1(\mathbf{\Gamma}), \ldots, \mathbb{S}_\nu(\mathbf{\Gamma})) \quad , \tag{35}$$

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

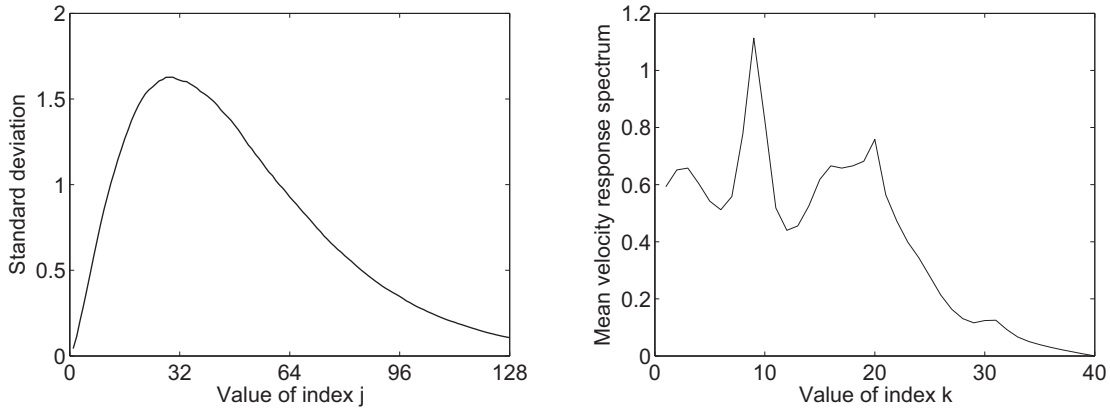

Figure 1. Graph of $j \mapsto \sigma_j = \sqrt{E\{\Gamma_j^2\}}$ (left figure) and graph of $k \mapsto \underline{\mathbb{S}}_k$ (right figure).

Normalization

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 $\mathbf{A} = (A_1, \ldots, A_N)$ be the $\mathbb{R}^N$-valued random variable defined, for all j in

$\{1,\ldots,N\}$, by $\Gamma_j = \sqrt{N}\,\sigma_j\,A_j$. We can then rewrite $\mathbf{\Gamma}$ as

$$\mathbf{\Gamma} = \sqrt{N}\,[\sigma]\,\mathbf{A} \quad , \quad [\,\sigma\,]_{jj'} = \sigma_j\,\delta_{jj'} \quad , \tag{36}$$

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

(i) The mean value $E\{\mathbf{A}\} = 0 \in \mathbb{R}^N$ because $E\{\mathbf{A}\} = N^{-1/2}\,[\sigma]^{-1}\,E\{\mathbf{\Gamma}\}$. Consequently, $\mathbf{A}$ is a centered random variable,

$$E\{\mathbf{A}\} = 0 \quad . \tag{37}$$

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

$$E\{A_j^2\} = \frac{1}{N} \quad , \quad \forall\, j \in \{1,\ldots,N\} \quad , \tag{38}$$

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

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

$$s_k(\mathbf{a}) = \frac{\mathcal{S}_k(\sqrt{N}\,[\sigma]\,\mathbf{a})}{\underline{\mathcal{S}}_k} \quad , \quad \forall\, k = 1,\ldots,\nu \quad . \tag{39}$$

It can then easily be deduced that

$$E\{\mathbf{s}(\mathbf{A})\} = \underline{\mathbf{s}} \;\in\; \mathbb{R}^\nu \quad . \tag{40}$$

Therefore, the available information which allows the probability distribution of random vector $\mathbf{A}$ to be constructed is made up of Eqs. (37), (38) and (40).

Unknown Lagrange multipliers and probability density function

Taking into account the normalization condition for the probability density function $p_{\mathbf{A}}$ and the available information defined by Eqs. (37), (38) and (40), the use of the MaxEnt principle yields, for all $\mathbf{a}$ in $\mathbb{R}^N$,

$$p_{\mathbf{A}}(\mathbf{a}) = c_0^{\text{sol}} \exp\{- < \boldsymbol{\lambda}_1^{\text{sol}},\, \mathbf{a} >_N - < \boldsymbol{\lambda}_2^{\text{sol}},\, \mathbf{a}^2 >_N - < \boldsymbol{\lambda}_3^{\text{sol}},\, \mathbf{s}(\mathbf{a}) >_\nu)\} \quad , \tag{41}$$

in which $\mathbf{a}^2$ denotes the vector $(a_1^2,\ldots,a_N^2)$ 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. (37), (38) and Eq. (40) respectively. Since $\boldsymbol{\mathcal{S}}(-\boldsymbol{\gamma}) = \boldsymbol{\mathcal{S}}(\boldsymbol{\gamma})$ and from Eqs. (37) and (41), it can be proven that $\boldsymbol{\lambda}_1^{\text{sol}} = 0$. Therefore, the Lagrange multiplier $\boldsymbol{\lambda}$ 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 $\mu = 168$. For all $\boldsymbol{\lambda}_2$ in $(]0\,,+\infty[)^N \subset \mathbb{R}^N$ and for all $\boldsymbol{\lambda}_3$ in $\mathbb{R}^\nu$, Eqs. (17) to (19) are satisfied.

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) with the interior-reflective Newton method. Figure 2 (right) displays the graph of the function iter $\mapsto$ convALG(iter) $= \|E\{\mathbf{g}(\mathbf{B}_{\boldsymbol{\lambda}(\text{iter})})\} - \mathbf{f}\|_\mu^2$ in which iter is the number of iteration of the optimization algorithm. The mathematical expectation defined by Eq. (11) is estimated by using the ergodic method (see Eq. (31)) with $M_0 = 300$ and the semi-implicit scheme is used with $\beta = 2\pi/(\sqrt{2\lambda_2^{\max}})$ in which $\lambda_2^{\max} = \max\{(\boldsymbol{\lambda}_2)_1, \ldots, (\boldsymbol{\lambda}_2)_N\}$ and with $m = 5$. Parameter f_0 has been fixed to 1. 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 θ. The convergence is then reasonably reached for $M = 20000$. 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).

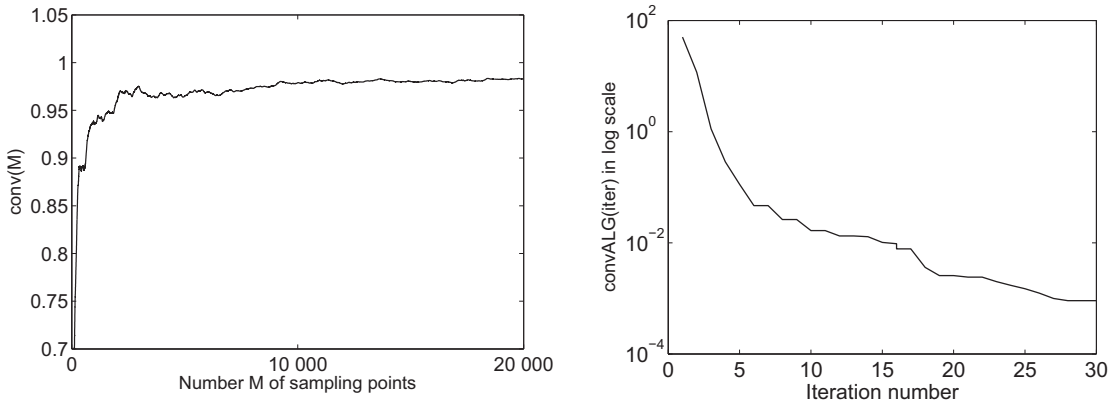

Figure 2. Graphs of $M \mapsto$ conv(M) (left figure) and of iter $\mapsto$ convALG(iter) (right figure).

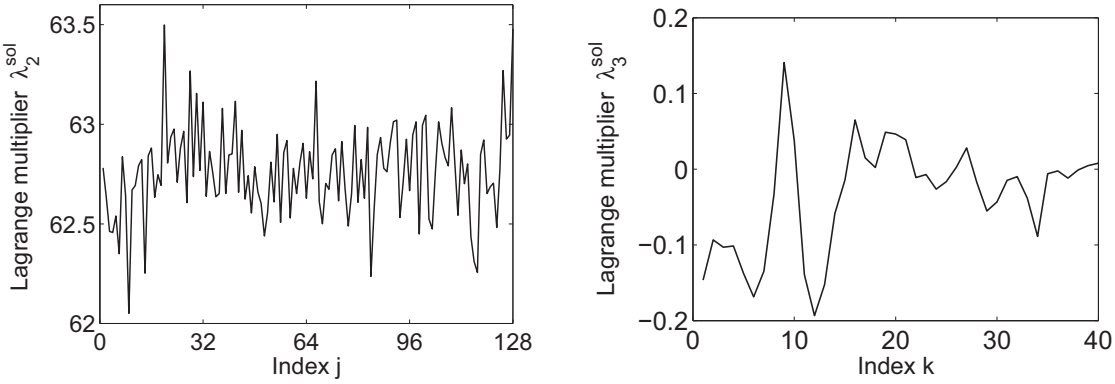

Figure 3. Graph of $j \mapsto (\boldsymbol{\lambda}_2^{\text{sol}})_j$ (left figure) and graph of $k \mapsto (\boldsymbol{\lambda}_3^{\text{sol}})_k$ (right figure).

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. The estimation of the mathematical expectation $E\{\mathbf{g}(A)\}$ of

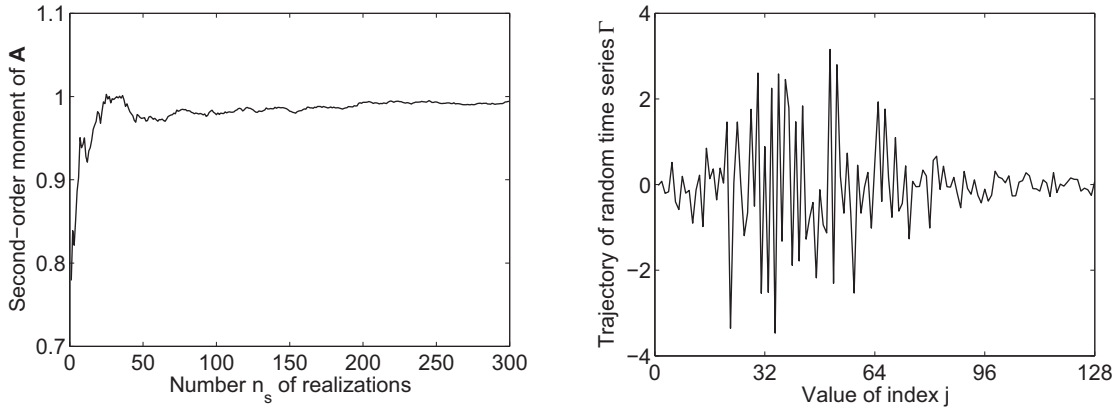

Figure 4. Graphs of the function $n_s \mapsto \text{convMC}(n_s)$ (left figure) and of the function $j \mapsto \Gamma_j(\theta_\ell)$ for a given realization θ_ℓ (right figure).

the constraint is carried out with the Monte Carlo method (see Eq. (33)). Computation is performed with $m = 5$, $M = 400$ and $n_s = 300$. Equation (32) is satisfied in the

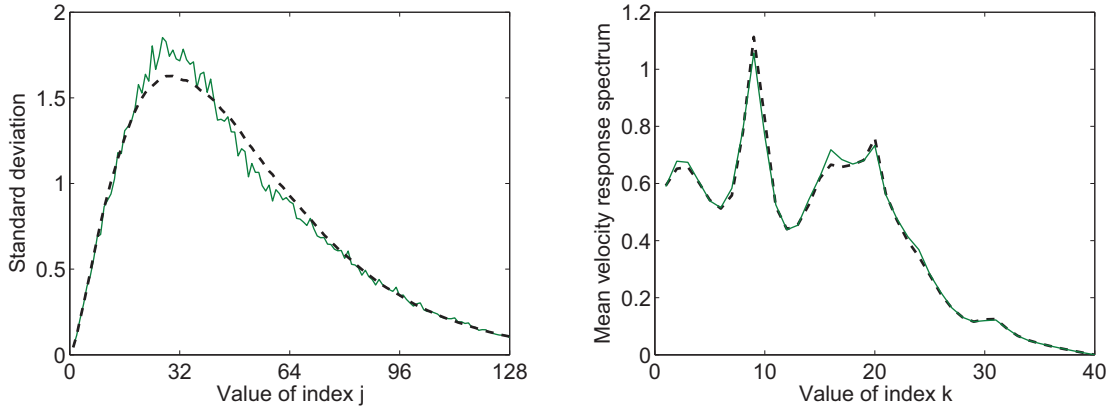

Figure 5. Graph of $j \mapsto \sigma_j = E\{\Gamma_j^2\}^{1/2}$ (left figure) and graph of $k \mapsto \underline{\mathcal{S}}_k$ (right figure). Reference (dashed lines). Estimation with the random generator (solid lines).

mean-square sense for $M = 400$. Figure 4 (right) displays 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 $\mathbf{A}(\theta_\ell) \simeq \mathbf{U}^M(\theta_\ell)$. Concerning the value of n_s, Figure 4 (left) 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 $E\{\|\mathbf{A}\|_N^2\} = E\{\|\mathbf{B}_{\boldsymbol{\lambda}^{\text{sol}}}\|_N^2\}$ of the random variable $\|\mathbf{A}\|_N$. Figure 5 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 5 (left) displays the graph of the standard-deviation function $j \mapsto \sigma_j$ from $\{1, \ldots, N\}$ into $\mathbb{R}^+$. Figure 5 (right) displays the graph of the mean velocity response spectrum $k \mapsto \underline{\mathcal{S}}_k$ from $\{1, \ldots, \nu\}$ 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.

Application 2. 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 (Soize, 2000, 2001 and 2005). 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.

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

$$[\,\mathbf{G}\,] = \begin{bmatrix} G_{11} & G_{12} & 0 & 0 \\ G_{12} & G_{22} & G_{23} & 0 \\ 0 & G_{23} & G_{33} & G_{34} \\ 0 & 0 & G_{34} & G_{44} \end{bmatrix} \quad , \tag{42}$$

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

$$E\{[\,\mathbf{G}\,]\} = [\,I_n\,] \,,\; \frac{E\{\|[\,\mathbf{G}\,] - [\,I_n\,]\|_F^2\}}{\|[\,I_n\,]\|_F^2} = \delta^2 < +\infty \,,\; E\{\|[\,\mathbf{G}\,]^{-1}\|_F^2\} = \alpha < +\infty \;, \tag{43}$$

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 [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

$$[\,\mathbf{G}\,] = [\,\mathbf{L}\,]^T\,[\,\mathbf{L}\,] \quad , \quad [\,\mathbf{L}\,] = \begin{bmatrix} A_1^2 & A_2 & 0 & 0 \\ 0 & A_3^2 & A_4 & 0 \\ 0 & 0 & A_5^2 & A_6 \\ 0 & 0 & 0 & A_7 \end{bmatrix} \quad , \tag{44}$$

in which $\mathbf{A} = (A_1, \ldots, A_N)$ is a $\mathbb{R}^N$-valued random vector with $N = 7$. Clearly, Eq. (44) 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. (43) and can be rewritten as

$$E\{\mathbf{e}(\mathbf{A})\} = \underline{\mathbf{e}} \in \mathbb{R}^N, E\{\|[\,\mathcal{G}(\mathbf{A})\,]\|_F^2\} = n(\delta^2+1)\,,\; E\{\|[\,\mathcal{G}(\mathbf{A})\,]^{-1}\|_F^2\} = \alpha < +\infty\,, \quad (45)$$

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

Unknown Lagrange multipliers and probability density function

We then have $\mu = N+2 = 9$. The Lagrange multiplier $\boldsymbol{\lambda}$ is written as $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_1, \lambda_2, \lambda_3) \in \mathcal{L}_\mu \subset \mathbb{R}^\mu = \mathbb{R}^N \times \mathbb{R} \times \mathbb{R}$ with $\mathcal{L}_\mu = \mathbb{R}^N \times]0\,,+\infty[\times]0\,,+\infty[$. The explicit Euler scheme is used. For all $\boldsymbol{\lambda}_1$ in $\mathbb{R}^N$, for all λ_2 and λ_3 in $]0\,,+\infty[$, Eqs. (17) to (19) are satisfied. 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). The probability density function defined by Eq. (6) is then written, for all $\mathbf{a}$ in $\mathbb{R}^N$, as

$$p_{\mathbf{A}}(\mathbf{a}) = c_0^{\text{sol}} \exp\{- < \boldsymbol{\lambda}_1^{\text{sol}},\, \mathbf{e}(\mathbf{a}) >_N - \boldsymbol{\lambda}_2^{\text{sol}} \|[\,\mathcal{G}(\mathbf{a})\,]\|_F^2 - \boldsymbol{\lambda}_3^{\text{sol}} \|[\,\mathcal{G}(\mathbf{a})\,]^{-1}\|_F^2\} \quad , \quad (46)$$

Computation of the vector-valued Lagrange multipliers

The interior-reflective Newton method used to solve Eq. (10). The mathematical expectation defined by Eq. (11) is estimated by using the Monte Carlo method (see Eq. (33)) and the random generator (see Eq. (32)) uses the Explicit Euler scheme. Parameter f_0 is fixed to 1 and the sampling step Δr is such that $\beta = 1$ and $m = 5$. Equation (32) is satisfied in the mean-square sense for $M = 5000$. Figures 6 (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 $n_s = 600$. Concerning solution $\boldsymbol{\lambda}^{\text{sol}}$, we get $\boldsymbol{\lambda}_1^{\text{sol}} = (0.7381, 4.1697, 1.2465, -0.9248, 0.8998, 4.2584, 0.8714)$, $\lambda_2^{\text{sol}} = 2.2293$ and $\lambda_3^{\text{sol}} = 1.7749$.

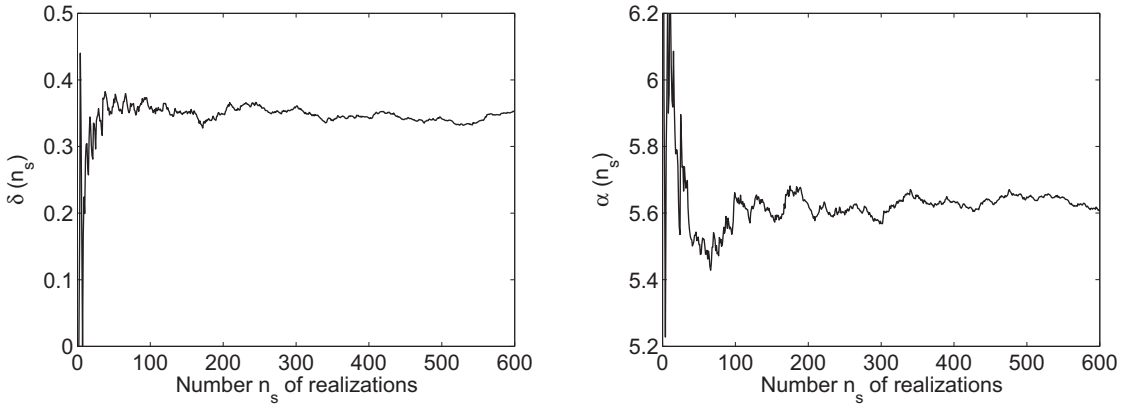

Figure 6. Graph of $n_s \mapsto \delta(n_s)$ (left figure) and graph of $n_s \mapsto \alpha(n_s)$ (right figure).

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. The estimations of the mathematical expectations $E\{\mathbf{g}(A)\}$ of the constraints are calculated by the Monte Carlo method (see Eq. (33)). Computation is performed with $f_0 = 1$, $m = 5$, $M = 5000$ and $n_s = 600$. Concerning the estimation of the constraints, we obtain

$$E\{[\,\mathbf{G}\,]\} \simeq \begin{bmatrix} 1.0001 & 0.0051 & 0 & 0 \\ 0.0051 & 0.9910 & 0.0057 & 0 \\ 0 & 0.0057 & 1.0222 & 0.0077 \\ 0 & 0 & 0.0077 & 1.0056 \end{bmatrix} \quad , \tag{47}$$

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.

Application 3. 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 the work by Soize (2006).

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

Let $\mathcal{A}$ be any part of $\mathbb{R}^N$, $\mathbf{x} = (x_1, \ldots, x_N)$ be any vector in $\mathbb{R}^N$ and let $d\mathbf{x} = dx_1 \ldots dx_N$ be the Lebesgue measure on $\mathbb{R}^N$. Let $\mathbf{X} = (X_1, \ldots, X_N)$ 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

$$\text{Supp } p_{\mathbf{X}} = \mathcal{A} \quad , \quad \int_{\mathbb{R}^N} p_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x} = \int_{\mathcal{A}} p_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x} = 1 \quad . \tag{48}$$

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

$$E\{\mathbf{g}(\mathbf{X})\} = \mathbf{f} \quad , \tag{49}$$

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

$$p_{\mathbf{X}}(\mathbf{x}) = \mathbb{1}_{\mathcal{A}}(\mathbf{x})\, c_X^{\text{sol}}\, \exp(- < \boldsymbol{\lambda}^{\text{sol}}, \mathbf{g}(\mathbf{x}) >_\mu) \quad , \quad \forall \mathbf{x} \in \mathbb{R}^N \quad , \tag{50}$$

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. (48). It should be noted that we cannot directly used the previous theory in introducing the function $\Phi(\mathbf{u}, \boldsymbol{\lambda}) = -\ln(\mathbb{1}_{\mathcal{A}}(\mathbf{u})) + < \boldsymbol{\lambda}, \mathbf{g}(\mathbf{u}) >_\mu$ (see Eq. (12)) because the function $\mathbf{u} \mapsto \|\boldsymbol{\nabla}_{\mathbf{u}} \Phi(\mathbf{u}, \boldsymbol{\lambda})\|_N$ 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. (49) 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 $\mathbf{A} = (A_1, \ldots, A_N)$ defined in the second section. We then have

$$E\{\mathbf{h}(\mathbf{X})\} = \int_{\mathbb{R}^N} \mathbf{h}(\mathbf{x})\, \mathbb{1}_{\mathcal{A}}(\mathbf{x})\, c_X^{\text{sol}}\, \exp(- < \boldsymbol{\lambda}^{\text{sol}}, \mathbf{g}(\mathbf{x}) >_\mu)\, d\mathbf{x} \quad , \tag{51}$$

which can be rewritten as

$$E\{\mathbf{h}(\mathbf{X})\} = \frac{c_X^{\text{sol}}}{c_0^{\text{sol}}} \int_{\mathbb{R}^N} \mathbf{h}(\mathbf{a})\, \mathbb{1}_{\mathcal{A}}(\mathbf{a})\, p_{\mathbf{A}}(\mathbf{a})\, d\mathbf{a} \quad , \tag{52}$$

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

$$E\{\mathbf{h}(\mathbf{X})\} = \frac{E\{\mathbf{h}(\mathbf{A})\, \mathbb{1}_{\mathcal{A}}(\mathbf{A})\}}{E\{\mathbb{1}_{\mathcal{A}}(\mathbf{A})\}} \quad , \tag{53}$$

where the mathematical expectations in the right-hand side of Eq. (53) are calculated by the theory presented in the paper. 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. We refer the reader to the work by Soize (2008) for a complete discussion concerning this approach and in particular on the method proposed to analyze the case of a compact support.

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

$$[\mathbf{K}]\, \mathbf{Y} = \mathbf{r} \quad , \quad [\mathbf{K}] = \sum_{j=1}^{n} Z_j\, [k^j] \quad , \tag{54}$$

in which $\mathbf{Z} = (Z_1, \ldots, Z_n)$ is the spatial sampling of a positive-valued random field (for instance, the Young modulus in linear isotropic elasticity for a heterogeneous material), where $[k^1], \ldots, [k^n]$ 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 $\underline{\mathbf{z}} = (\underline{z}_1, \ldots, \underline{z}_n)$ 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 $j \in \{1, \ldots, n\}$, it is assumed that $\underline{z}_j > 0$ and $Z_j > 0$ almost surely. We then introduce the normalized random vector $\mathbf{G} = (G_1, \ldots, G_n)$ with values in $\mathbb{R}^n$ such that $Z_j = \underline{z}_j\, G_j$ for all $j \in \{1, \ldots, n\}$. Therefore, we have $G_j > 0$ almost surely for all $j \in \{1, \ldots, n\}$. 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 (Soize, 2006) and taking into account that for all $g_1 > 0, \ldots, g_n > 0$, we have

$$(\max\{\frac{1}{g_1}, \ldots, \frac{1}{g_n}\})^2 \leq \frac{1}{g_1^2} + \ldots + \frac{1}{g_n^2} \quad , \tag{55}$$

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

$$E\{\frac{1}{G_1^2} + \ldots + \frac{1}{G_n^2}\} = c_G < +\infty \quad . \tag{56}$$

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 $\mathbf{G}$,

$$\mathbf{G}^N = \underline{\mathbf{G}} + \sum_{\alpha=1}^{N} \sqrt{v_\alpha}\, X_\alpha\, \boldsymbol{\varphi}^\alpha \quad , \tag{57}$$

in which $\underline{\mathbf{G}} = (\underline{G}_1, \ldots, \underline{G}_n)$ with $\underline{G}_j = 1$ for all j. In Eq. (57), $\boldsymbol{\varphi}^1, \ldots, \boldsymbol{\varphi}^N$ are the orthonormal eigenvectors ($< \boldsymbol{\varphi}^\alpha, \boldsymbol{\varphi}^\beta >= \delta_{\alpha\beta}$) associated with the N largest eigenvalues $v_1 > v_2 > \ldots > v_N$ 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 $\mathbf{X} = (X_1, \ldots, X_N)$, the available information is the following,

$$\text{Supp } p_{\mathbf{X}} = \mathcal{A} \subset \mathbb{R}^N \quad , \tag{58}$$

$$E\{\mathbf{X}\} = 0 \quad , \tag{59}$$

$$E\{\mathbf{X}\,\mathbf{X}^T\} = [I_N] \quad , \tag{60}$$

$$E\{s(\mathbf{X})\} = \kappa < +\infty \quad , \tag{61}$$

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

$$s(\mathbf{x}) = \sum_{j=1}^{n} (\underline{G}_j + \sum_{\alpha=1}^{N} \sqrt{v_\alpha}\, x_\alpha\, \varphi_j^\alpha)^{-2} \quad . \tag{62}$$

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

$$\mathcal{A} = \{\mathbf{x} \in \mathbb{R}^N \quad \text{such that} \quad \forall j \in \{1, \ldots, n\} \quad , \quad \underline{G}_j + \sum_{\alpha=1}^{N} \sqrt{v_\alpha}\, x_\alpha\, \varphi_j^\alpha > 0\} \quad . \tag{63}$$

The details of the construction are given in (Soize, 2008). We have $\mu = N+N(N+1)/2+1$. The Lagrange multiplier is $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \lambda_3) \in \mathcal{L}_\mu \subset \mathbb{R}^\mu = \mathbb{R}^N \times \mathbb{R}^{N(N+1)/2} \times \mathbb{R}$. The semi-implicit scheme is used to discretize the Itô stochastic differential equation. The probability density function defined by Eq. (6) is then written, for all $\mathbf{a}$ in $\mathbb{R}^N$, as

$$p_{\mathbf{A}}(\mathbf{a}) = c_0^{\text{sol}} \, \exp\{- < \boldsymbol{\lambda}_1^{\text{sol}}, \mathbf{a}) >_N - < \boldsymbol{\lambda}_2^{\text{sol}}, \mathbf{e}(\mathbf{a}) >_{N(N+1)/2} - \lambda_3^{\text{sol}} \, s(\mathbf{a})\} \quad , \tag{64}$$

in which $\mathbf{e}(\mathbf{a})$ 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 $\mathbf{a}\,\mathbf{a}^T$. For the numerical application, $n = 100$ and the values of $\sqrt{v_1}, \ldots, \sqrt{v_{20}}$ 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. The value $N = 10$ implies a reasonable relative error. The mathematical

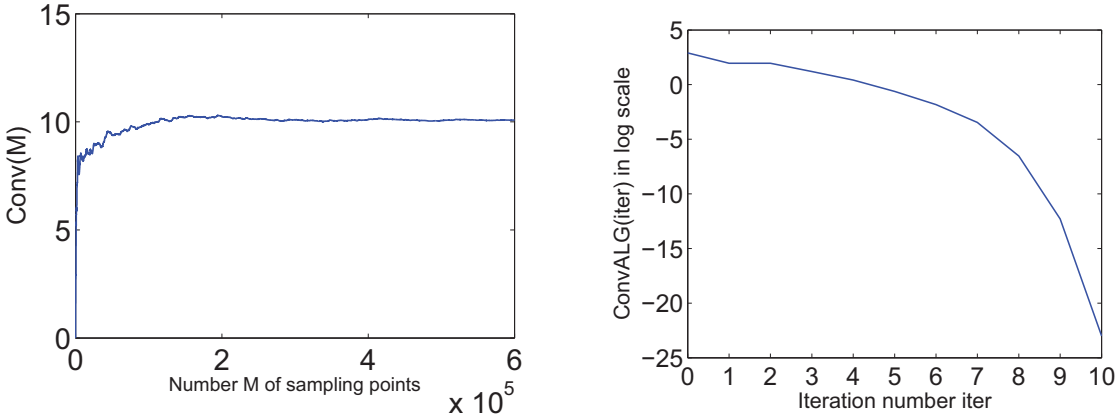

Figure 7. Graphs of $M \mapsto \mathbf{conv}(M)$ (left figure) and of $\mathbf{iter} \mapsto \mathbf{convALG}(\mathbf{iter})$ (right figure).

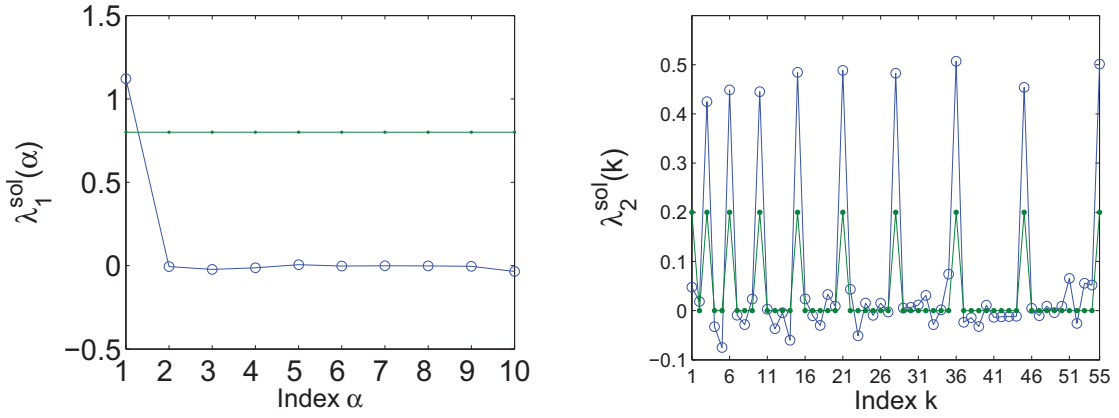

Figure 8. Left figure: Graphs of $\alpha \mapsto (\boldsymbol{\lambda}_1^{\text{sol}})_\alpha$ (circle marker) and of $\alpha \mapsto (\boldsymbol{\lambda}_1^0)_\alpha$ (dot marker). Right figure: Graphs of $k \mapsto (\boldsymbol{\lambda}_2^{\text{sol}})_k$ (circle marker) and of $k \mapsto (\boldsymbol{\lambda}_2^0)_k$ (dot marker).

expectations in Eq. (53) are estimated by using the ergodic theory (see Eq. (31)) and the semi-implicit scheme is used. The sampling step is $\Delta r = 0.01$. Parameter f_0 is equal to 0.5 and $M_0 = 200,000$, $M = 600,000$. The value of λ_3 is fixed to the value $\lambda_3^{\text{sol}} = 0.01$

and Eq. (49) 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. Figure 7 (left) displays the graph of the function $M \mapsto \text{conv}(M)$ showing the convergence of the estimation of $E\{\|\mathbf{B}_{\widetilde{\boldsymbol{\lambda}}^{\text{sol}}}\|_N^2\}$ by using the ergodic method with $\widetilde{\boldsymbol{\lambda}} = (\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2)$. Figure 7 (right) displays the graph of the function $\text{iter} \mapsto \text{convALG}(\text{iter}) = \|E\{\mathbf{g}(\mathbf{B}_{\widetilde{\boldsymbol{\lambda}}(\text{iter})})\} - \mathbf{f}\|_{\mu-1}^2$ showing the convergence of trust-region dogleg algorithm in function of the iteration number. Figure 8 (left) compares the graph $\alpha \mapsto (\boldsymbol{\lambda}_1^{\text{sol}})_\alpha$ of the solution for $\boldsymbol{\lambda}_1$ with the graph $\alpha \mapsto (\boldsymbol{\lambda}_1^0)_k$ of the initial value $\boldsymbol{\lambda}_1^0 = 0.8\,\mathbb{1}_N$. Figure 8 (right) compares the graph $k \mapsto (\boldsymbol{\lambda}_2^{\text{sol}})_k$ of the solution for $\boldsymbol{\lambda}_2$ with the graph $k \mapsto (\boldsymbol{\lambda}_2^0)_k$ of the initial value $\boldsymbol{\lambda}_2^0 = 0.2\,\underline{\mathbf{e}}$. For the solution obtained $\boldsymbol{\lambda}^{\text{sol}} = (\boldsymbol{\lambda}_1^{\text{sol}}, \boldsymbol{\lambda}_2^{\text{sol}}, \lambda_3^{\text{sol}})$ with $\lambda_3^{\text{sol}} = 0.01$, Eq. (59) and Eq. (60) are satisfied at 10^{-11} for each components and Eq. (61) yields $\kappa = 140.027$. The results obtained are thus very good.

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

Coleman, T.F. and Y. Li (1994), "On the convergence of reflective Newton methods for large-scale nonlinear minimization subject to sounds", *Mathematical Programming*, **67**(2), 189-224.

Coleman, T.F. and Y. Li (1996), "An interior, trust region approach for nonlinear minimization subject to bounds", *SIAM Journal on Optimization*, **6**, 418-445.

Geman, S. and D. Geman (1984), "Stochastic relaxation, Gibbs distribution and the Bayesian distribution of images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **PAM I-6**, 721-741.

Ghanem, R. and P.D. Spanos (1991), *Stochastic Finite Elements: A spectral Approach*, Springer-Verlag, New

York.

Hastings, W.K. (1970), "Monte Carlo sampling methods using Markov chains and their applications", *Biometrika*, **109**, 57-97.

Kaipio, J. and E. Somersalo (2005), *Statistical and Computational Inverse Problems*, Springer-Verlag, New York.

Kloeden, P.E. and E. Platen (1992), *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, Berlin.

MacKeown, P.K. (1997), *Stochastic Simulation in Physics*, Springer-Verlag, Singapore.

Schueller, G.I. ed. (2005), "Computational methods in stochastic mechanics and reliability analysis", *Computer Methods in Applied Mechanics and Engineering*, **194**(12-16), 1251-1795.

Shannon, C.E. (1948), "A mathematical theory of communication", *Bell System Technology Journal*, **27**, 379-423 & 623-659.

Soize, C. (1994), *The Fokker-Planck Equation for Stochastic Dynamical Systems and its Explicit Steady State Solutions*, World Scientific, Singapore.

Soize, C. (2000), "A nonparametric model of random uncertainties for reduced matrix models in structural dynamics", *Probabilistic Engineering Mechanics*, **15**(3), 277-294.

Soize, C. (2001), "Maximum entropy approach for modeling random uncertainties in transient elastodynamics", *Journal of the Acoustical Society of America*, **109**(5), 1979-1996.

Soize, C. (2005), "Random matrix theory for modeling uncertainties in computational mechanics", *Computer Methods in Applied Mechanics and Engineering*, **194**(12-16), 1333-1366.

Soize, C. (2006), "Non Gaussian positive-definite matrix-valued random fields for elliptic stochastic partial differential operators", *Computer Methods in Applied Mechanics and Engineering*, **195**, 26-64.

Soize, C. (2008), "Construction of probability distributions in high dimension using the maximum entropy principle. Applications to stochastic processes, random field and random matrices", *International Journal for Numerical Methods in Engineering*, accepted for publication in February 2008.

Spall, J.C. (2003), *Introduction to Stochastic Search and Optimization*, John Wiley and Sons, Hoboken, New Jersey.

Talay, D. and L. Tubaro (1990), "Expansion of the global error for numerical schemes solving stochastic differential equation", *Stoch. Anal. Appl.*, **8**(4), 94-120.

Talay, D. (2002), "Stochastic Hamiltonian system: exponential convergence to the invariant measure and discretization by the implicit Euler scheme", *Markov Processes and Related Fields*, **8**, 163-198.

Yosida, K. (1995), *Functional Analysis*, Springer-Verlag, Berlin.