

## Polynomial chaos expansion of a multimodal random vector

C. Soize<sup>†</sup>

---

**Abstract.** A methodology and algorithms are proposed for constructing the polynomial chaos expansion (PCE) of multimodal random vectors. An algorithm is developed for generating independent realizations of any multimodal multivariate probability measure that is constructed from a set of independent realizations using the Gaussian kernel-density estimation method. The PCE is then performed with respect to this multimodal probability measure, for which the realizations of the polynomial chaos are computed with an adapted algorithm. Finally, a numerical application is presented for analyzing the convergence properties.

**Key words.** multimodal probability distribution, polynomial chaos, random vector, high dimension

**AMS subject classifications.** 62H10, 62H12, 62H20, 60E10, 60E05, 60G35, 60G60

**1. Introduction.** In 1991, R. Ghanem [16] has proposed (1) an efficient construction of the polynomial chaos expansion (PCE) [8] for representing second-order stochastic processes and random fields, and (2) to use it for solving boundary value problems with uncertain parameters by a spectral approach and the stochastic finite elements. Since 1991, numerous works have been published in the area of the PCE and of its use in the spectral approaches for solving linear and nonlinear stochastic boundary value problems, and some associated statistical inverse problems (see for instance [1, 9, 10, 11, 13, 17, 18, 27, 31, 34, 42, 44]). Several extensions have been proposed concerning generalized chaos expansions, the PCE for an arbitrary probability measure, the PCE with random coefficients [14, 28, 38, 40, 49, 50], and recently, the construction of a basis adaptation in homogeneous chaos spaces [48]. Although several works have been devoted to the acceleration of stochastic convergence of the PCE (see for instance [19, 24, 29, 48]), the question relative to the speed of convergence (which can be very low) of the PCE for a multimodal probability distribution on  $\mathbb{R}^n$  has been little addressed. Recently, a procedure through mixtures of PCE has been proposed in [33] for the one-dimension case. In this paper, we propose a methodology for the PCE of a multimodal  $\mathbb{R}^m$ -valued random variable. This problem belongs to the class of the PCE with respect to an arbitrary probability measure. The framework of the developments presented in the paper is motivated by the difficulty encountered for the PCE of a random vector for which its probability density function is multimodal, and for which it is known that the speed of convergence of the PCE can be low. Nevertheless, the method proposed is very general and goes beyond multimodality. In the context of the statistical inverse problem related to the identification of a PCE of a random vector, one does not know if the unknown probability density function is unimodal or is multimodal. So the method proposed allows for accelerating the speed of convergence in all the cases. We propose an algorithm for generating independent realizations of the multimodal probability measure on  $\mathbb{R}^n$ , which is constructed from a set of realizations using the Gaussian kernel-density estimation method from the nonparametric statistics. Then, the PCE of the  $\mathbb{R}^m$ -valued random variable is performed with respect to the constructed multimodal probability measure on  $\mathbb{R}^n$ , for which the realizations of the polynomial chaos are computed with an adapted algorithm, recently introduced. Finally, a numerical application is presented for

---

<sup>†</sup>Université Paris-Est, Laboratoire Modélisation et Simulation Multi Echelle, MSME UMR 8208 CNRS, 5 bd Descartes, 77454 Marne-la-Vallée, France (christian.soize@univ-paris-est.fr).

analyzing the convergence properties. This new class of algorithms for the multimodal case can be useful in the context of uncertainty quantification for direct and inverse problems, and in particular, for the approaches devoted to dimension reduction in chaos expansions for nonlinear coupled problems, when an iterative solver is used (see for instance [2, 3, 4]).

**2. Construction of a representation of a multimodal random vector in high dimension from a set of realizations.** In the first part of this paper, we propose a construction of a stochastic model of a multimodal random vector  $\mathbf{X}$  with values in  $\mathbb{R}^N$ , defined on a probability space  $(\Theta, \mathcal{T}, \mathcal{P})$  using only a set of  $\nu \gg 1$  independent realizations of  $\mathbf{X}$ . The construction consists (i) in introducing the usual empirical estimations of the mean vector and the covariance matrix of  $\mathbf{X}$ , (ii) in constructing a reduced-order statistical model  $\mathbf{X}^{(n)}$  (with values in  $\mathbb{R}^N$ ) of  $\mathbf{X}$  (with dimension  $n < N$ ) using the classical principal component analysis, and yielding a multimodal reduced random vector  $\mathbf{H}$  with values in  $\mathbb{R}^n$  (which is assumed to be in high dimension, that is to say, with  $N > n \gg 1$ ), (iii) in constructing a multimodal probability density function  $\boldsymbol{\eta} \mapsto p_{\mathbf{H}}(\boldsymbol{\eta})$  on  $\mathbb{R}^n$  of  $\mathbf{H}$  introducing an adapted nonparametric statistical estimator, (iv) in developing a generator of independent realizations of  $\mathbf{H}$  that follows the multimodal probability distribution  $p_{\mathbf{H}}(\boldsymbol{\eta}) d\boldsymbol{\eta}$  on  $\mathbb{R}^n$ , (v) for fixed  $n$ , in computing a sequence  $\{\hat{w}_M\}_{M \geq 1}$  of statistical estimations of  $\{w(\mathbf{H})\}$  using this generator, in which  $w$  is a given measurable mapping defined on  $\mathbb{R}^n$ . At the end of this section, a numerical investigation is presented in order to illustrate the construction and the algorithms.

### 2.1. Data description and usual empirical estimations of second-order moments.

Let  $\mathbf{X} = (X_1, \dots, X_N)$  be a  $\mathbb{R}^N$ -valued second-order random vector defined on a probability space  $(\Theta, \mathcal{T}, \mathcal{P})$ , whose probability distribution is represented by an unknown probability density function  $\mathbf{x} \mapsto p_{\mathbf{X}}(\mathbf{x})$  with respect to the Lebesgue measure  $d\mathbf{x}$  on  $\mathbb{R}^N$ . Let  $E$  be the mathematical expectation. It is assumed that  $\nu$  (with  $\nu \gg 1$ ) independent realizations  $\mathbf{x}^{\text{exp},1}, \dots, \mathbf{x}^{\text{exp},\nu}$  of  $\mathbf{X}$  are known (coming from experimental data or from numerical simulations). Let  $\hat{\mathbf{m}}_{\mathbf{X}}$  and  $[\hat{C}_{\mathbf{X}}]$  be the empirical estimations of the mean vector  $\mathbf{m}_{\mathbf{X}} = E\{\mathbf{X}\}$  and the covariance matrix  $[C_{\mathbf{X}}] = E\{(\mathbf{X} - \mathbf{m}_{\mathbf{X}})(\mathbf{X} - \mathbf{m}_{\mathbf{X}})^T\}$ , such that

$$\hat{\mathbf{m}}_{\mathbf{X}} = \frac{1}{\nu} \sum_{\ell=1}^{\nu} \mathbf{x}^{\text{exp},\ell} \quad , \quad [\hat{C}_{\mathbf{X}}] = \frac{1}{\nu-1} \sum_{\ell=1}^{\nu} (\mathbf{x}^{\text{exp},\ell} - \hat{\mathbf{m}}_{\mathbf{X}})(\mathbf{x}^{\text{exp},\ell} - \hat{\mathbf{m}}_{\mathbf{X}})^T. \quad (2.1)$$

Note that  $[\hat{C}_{\mathbf{X}}]$  can be written as

$$[\hat{C}_{\mathbf{X}}] = [\hat{R}_{\mathbf{X}}] - \frac{\nu+1}{\nu-1} \hat{\mathbf{m}}_{\mathbf{X}} \hat{\mathbf{m}}_{\mathbf{X}}^T \quad , \quad [\hat{R}_{\mathbf{X}}] = \frac{1}{\nu-1} \sum_{\ell=1}^{\nu} \mathbf{x}^{\text{exp},\ell} (\mathbf{x}^{\text{exp},\ell})^T, \quad (2.2)$$

in which  $[\hat{R}_{\mathbf{X}}]$  is the estimation of the second-order moment matrix,  $[R_{\mathbf{X}}] = E\{\mathbf{X}\mathbf{X}^T\}$ .

**2.2. Reduced-order statistical model of  $\mathbf{X}$ .** The eigenvalues  $\lambda_{X,1} \geq \lambda_{X,2} \geq \dots \geq \lambda_{X,N} \geq 0$ , and the associated orthonormal eigenvectors  $\phi_X^1, \dots, \phi_X^N$ , such that  $(\phi_X^i)^T \phi_X^j = \delta_{ij}$  (the Kronecker symbol), are such that  $[\hat{C}_{\mathbf{X}}] \phi_X^i = \lambda_{X,i} \phi_X^i$ . The principal component analysis allows a reduced-order statistical model,  $\mathbf{X}^{(n)}$ , of  $\mathbf{X}$  to be constructed, such that

$$\mathbf{X}^{(n)} = \hat{\mathbf{m}}_{\mathbf{X}} + \sum_{i=1}^n \sqrt{\lambda_{X,i}} H_i \phi_X^i. \quad (2.3)$$

Consequently, random vector  $\mathbf{X}$  is such that  $\mathbf{X} = \mathbf{X}^{(N)}$ , and the value of  $n$  is fixed in  $[1, N]$  for that  $\text{err}(n) \leq \varepsilon$  in which  $\varepsilon$  is such that  $0 \leq \varepsilon \ll 1$ , and where  $\text{err}$  is the error function defined by

$$\text{err}(n) = 1 - \frac{\sum_{i=1}^n \lambda_{X,i}}{\text{tr}[\widehat{C}_{\mathbf{X}}]}. \quad (2.4)$$

Let  $\mathbf{H} = (H_1, \dots, H_n)$  be the  $\mathbb{R}^n$ -valued second-order random variable. The  $\nu$  independent realizations  $\boldsymbol{\eta}^{\text{exp},1}, \dots, \boldsymbol{\eta}^{\text{exp},\nu}$ , with  $\boldsymbol{\eta}^{\text{exp},\ell} = (\eta_1^{\text{exp},\ell}, \dots, \eta_n^{\text{exp},\ell}) \in \mathbb{R}^n$ , of the  $\mathbb{R}^n$ -valued random variable  $\mathbf{H}$  are computed, for  $\ell = 1, \dots, \nu$  and  $i = 1, \dots, n$ ,

$$\eta_i^{\text{exp},\ell} = \frac{1}{\sqrt{\lambda_{X,i}}} (\mathbf{x}^{\text{exp},\ell} - \widehat{\mathbf{m}}_{\mathbf{X}})^T \boldsymbol{\phi}_X^i. \quad (2.5)$$

By construction, it can easily be verified that

$$\widehat{\mathbf{m}}_{\mathbf{H}} = \frac{1}{\nu} \sum_{\ell=1}^{\nu} \boldsymbol{\eta}^{\text{exp},\ell} = \mathbf{0} \quad , \quad [\widehat{R}_{\mathbf{H}}] = \frac{1}{\nu-1} \sum_{\ell=1}^{\nu} \boldsymbol{\eta}^{\text{exp},\ell} (\boldsymbol{\eta}^{\text{exp},\ell})^T = [I_n]. \quad (2.6)$$

### 2.3. Construction of the multimodal probability density function of random vector

**H.** As explained in the introduction of Section 2, the probability density function  $p_{\mathbf{H}}$  with respect to the Lebesgue measure  $d\boldsymbol{\eta}$  on  $\mathbb{R}^n$  of the  $\mathbb{R}^n$ -valued second-order random variable  $\mathbf{H} = (H_1, \dots, H_n)$ , must be constructed. The unknown probability density function  $p_{\mathbf{X}}$  has been assumed to be multimodal. Due to the reduced representation introduced by Eq. (2.3), probability density function  $p_{\mathbf{H}}$  (that differs from  $p_{\mathbf{X}}$ ), could be not multimodal. The objective of the present work is to develop a methodology adapted to the case for which  $p_{\mathbf{H}}$  is multimodal, and consequently, it is assumed that  $p_{\mathbf{H}}$  can be multimodal. It should be noted that the method proposed is very general and can be used for a probability density function  $p_{\mathbf{H}}$  that is or that is not multimodal. We propose to choose for  $p_{\mathbf{H}}$  its estimation carried out by using the Gaussian kernel-density estimation method on the basis of the knowledge of the  $\nu$  independent realizations,  $\boldsymbol{\eta}^{\text{exp},1}, \dots, \boldsymbol{\eta}^{\text{exp},\nu}$  computed with Eq. (2.5). A modification of the classical Gaussian kernel-density estimation method [5] is used in order that Eq. (2.6),  $[\widehat{R}_{\mathbf{H}}] = [I_n]$ , be preserved. The positive valued function  $p_{\mathbf{H}}$  on  $\mathbb{R}^n$  is then defined, for all  $\boldsymbol{\eta}$  in  $\mathbb{R}^n$ , by

$$p_{\mathbf{H}}(\boldsymbol{\eta}) = \frac{1}{\nu} \sum_{\ell=1}^{\nu} \mu_{n,\widehat{s}_n} \left( \frac{\widehat{s}_n}{s_n} \boldsymbol{\eta}^{\text{exp},\ell} - \boldsymbol{\eta} \right), \quad (2.7)$$

in which  $\mu_{n,\widehat{s}_n}$  is the positive function from  $\mathbb{R}^n$  into  $]0, +\infty[$  defined, for all  $\boldsymbol{\eta}$  in  $\mathbb{R}^n$ , by

$$\mu_{n,\widehat{s}_n}(\boldsymbol{\eta}) = \frac{1}{(\sqrt{2\pi} \widehat{s}_n)^n} \exp\left\{-\frac{1}{2\widehat{s}_n^2} \|\boldsymbol{\eta}\|_n^2\right\}, \quad (2.8)$$

in which  $\|\boldsymbol{\eta}\|_n^2 = \eta_1^2 + \dots + \eta_n^2$ , and where the positive parameters  $s_n$  and  $\widehat{s}_n$  are defined by

$$s_n = \left\{ \frac{4}{\nu(2+n)} \right\}^{1/(n+4)}, \quad \widehat{s}_n = \frac{s_n}{\sqrt{s_n^2 + \frac{\nu-1}{\nu}}}. \quad (2.9)$$

Parameter  $s_n$  is the usual multidimensional optimal Silverman bandwidth (in taking into account that the empirical estimation of the standard deviation of each component is unity, and parameter  $\widehat{s}_n$  has been introduced in order that the second equation in Eq. (2.6) holds. It should be noted that, for  $n$  fixed, parameters  $s_n$  and  $\widehat{s}_n$  go to  $0^+$ , and  $\widehat{s}_n/s_n$  goes to  $1^-$ , when  $\nu$  goes to  $+\infty$ . Using Eqs. (2.7) to (2.9), it can easily be verified that

$$E\{\mathbf{H}\} = \int_{\mathbb{R}^n} \boldsymbol{\eta} p_{\mathbf{H}}(\boldsymbol{\eta}) d\boldsymbol{\eta} = \frac{\widehat{s}_n}{s_n} \widehat{\mathbf{m}}_{\mathbf{H}} = \mathbf{0}, \quad (2.10)$$

$$E\{\mathbf{H}\mathbf{H}^T\} = \int_{\mathbb{R}^n} \boldsymbol{\eta} \boldsymbol{\eta}^T p_{\mathbf{H}}(\boldsymbol{\eta}) d\boldsymbol{\eta} = \widehat{s}_n^2 [I_n] + \left(\frac{\widehat{s}_n}{s_n}\right)^2 \frac{(\nu - 1)}{\nu} [\widehat{R}_{\mathbf{H}}] = [I_n]. \quad (2.11)$$

**Remark 1.**

(i) The multimodal probability density function  $p_{\mathbf{H}}$  that is constructed by Eqs. (2.7) to (2.9), depends on  $\nu$ . Such a construction is correct in the framework of the hypotheses that have been introduced, because the set of realizations of  $\mathbf{X}$  is given, and consequently,  $\nu$  is fixed (See Section 2.1).

(ii) In Eqs. (2.7) and (2.8), if  $\widehat{s}_n$  was chosen as  $s_n$  (usual nonparametric statistical estimator with Gaussian kernel-density estimation method), then Eq. (2.11) would not hold; we propose such a construction of the multimodal probability distribution of random vector  $\mathbf{H}$  in order that  $E\{\mathbf{H}\mathbf{H}^T\} = [I_n]$ .

(iii) In the framework of the proposed developments, no assumptions (in particular concerning the support) are introduced concerning the unknown multivariate probability distribution that has to be estimated from a set of data by the nonparametric statistics using the kernel-density estimation method (kernel smoothing). Consequently, a first arbitrary choice must be done concerning the kernel smoothing, and the kernel proposed is the multivariate Gaussian kernel for centered and uncorrelated random variables (see Eq. (2.6)) (see Eq. (2.6)). The second arbitrary choice concerns the bandwidth. Again, since no information is available, the multidimensional Silverman bandwidth that is optimal for a Gaussian distribution is chosen. It should be noted that this choice for the bandwidth is consistent with the usual choice of the empirical estimators used for estimating the mean vector and the covariance matrix, which are optimal (unbiased, efficient and consistent estimators) for a Gaussian distribution. In addition, the methodology proposed in the paper is general and is independent of the choice of the kernel smoothing, provided that the assumptions used by the generator of realizations are satisfied. Therefore, if additional information are available concerning the unknown multivariate probability distribution that has to be estimated from the set of data, the Gaussian kernel  $\mu_{n,\widehat{s}_n}(\boldsymbol{\eta})$  and the Silverman bandwidth  $s_n$  can be replaced by a kernel and bandwidth that are better adapted. Nevertheless, it should be noted that the nonparametric estimation yields a probability density function  $p_{\mathbf{H}}$  that is only used for constructing a random vector  $\mathbf{H}$  (whose probability distribution is defined by this pdf), which is used as a germ for performing the PCE in order to increase the speed of convergence of the truncated PCE.

**Remark 2.**

For  $i$  fixed in  $\{1, \dots, n\}$ , the probability density function  $p_{H_i}$  on  $\mathbb{R}$  of the random variable  $H_i$  is

calculated in integrating Eq. (2.7) over  $\mathbb{R}^{n-1}$ , and yields

$$p_{H_i}(\eta_i) = \frac{1}{\nu} \sum_{\ell=1}^{\nu} \mu_{1, \hat{s}_n} \left( \frac{\hat{s}_n}{s_n} \eta_i^{\text{exp}, \ell} - \eta_i \right). \quad (2.12)$$

For  $i$  and  $j$  fixed in  $\{1, \dots, n\}$ , the joint probability density function  $p_{H_i H_j}$  on  $\mathbb{R}^2$  of the random variables  $H_i$  and  $H_j$ , is calculated in integrating Eq. (2.7) over  $\mathbb{R}^{n-2}$ , and yields

$$p_{H_i, H_j}(\eta_i, \eta_j) = \frac{1}{\nu} \sum_{\ell=1}^{\nu} \mu_{2, \hat{s}_n} \left( \frac{\hat{s}_n}{s_n} \boldsymbol{\eta}^{\text{exp}, \ell, ij} - \boldsymbol{\eta}^{ij} \right), \quad (2.13)$$

in which  $\boldsymbol{\eta}^{\text{exp}, \ell, ij} = (\eta_i^{\text{exp}, \ell}, \eta_j^{\text{exp}, \ell})$  and  $\boldsymbol{\eta}^{ij} = (\eta_i, \eta_j)$  belong to  $\mathbb{R}^2$ .

**2.4. Generator for multimodal random vector  $\mathbf{H}$ .** Let  $w$  be a mapping from  $\mathbb{R}^n$  into an Euclidean space such that  $w(\mathbf{H})$  is a second-order random variable. The estimation of  $E\{w(\mathbf{H})\} = \int_{\mathbb{R}^n} w(\boldsymbol{\eta}) p_{\mathbf{H}}(\boldsymbol{\eta}) d\boldsymbol{\eta}$  requires a generator of independent realizations of random vector  $\mathbf{H}$  for which the multimodal probability distribution  $p_{\mathbf{H}}(\boldsymbol{\eta}) d\boldsymbol{\eta}$  is defined by Eq. (2.7) with Eqs. (2.8) to (2.9). Such a generator can be performed using the Markov Chain Monte Carlo method (MCMC) [23, 36, 43]. The transition kernel of the homogeneous Markov chain of the MCMC method can be constructed using the Metropolis-Hastings algorithm [30, 22] (that requires the definition of a good proposal distribution), the Gibbs sampling [15] (that requires the knowledge of the conditional distribution) or the slice sampling [32] (that can exhibit difficulties related to the general shape of the probability distribution, in particular for multimodal distributions). In general, these algorithms are efficient, but can also be not efficient if there exist attraction regions which do not correspond to the invariant measure under consideration and tricky even in high dimension. These cases cannot easily be detected and are time consuming. The method proposed in [39] is very robust, has recently been applied with success for complex problems in high dimension [6, 20] and is reused hereinafter. It looks similar to the Gibbs approach but corresponds to a more direct construction of a random generator of realizations for random variable  $\mathbf{H}$  whose probability distribution is  $p_{\mathbf{H}}(\boldsymbol{\eta}) d\boldsymbol{\eta}$  and is multimodal. 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 [39], the construction of the transition kernel by using the detailed balance equation is replaced by the construction of an Itô Stochastic Differential Equation (ISDE), which admits  $p_{\mathbf{H}}(\boldsymbol{\eta}) d\boldsymbol{\eta}$  (defined by Eqs. (2.7) to (2.9)) as a unique invariant measure. The ergodic method or the Monte Carlo method can be used for estimating  $E\{w(\mathbf{H})\}$ .

It should be noted that the main ideas presented in this paper are not related to a specific MCMC algorithm for constructing a set of realizations. The alternative MCMC algorithm proposed hereinafter can be replaced by any traditional MCMC algorithm. Nevertheless, this alternative algorithm is very robust and very rich in terms of control based on the use of mathematical results for the Itô stochastic differential equations.

**2.4.1. Interpretation of the multimodal probability distribution  $p_{\mathbf{H}}$  as the invariant measure of an Itô stochastic differential equation (ISDE).** Let  $\boldsymbol{\eta} \mapsto \Phi(\boldsymbol{\eta})$  be the function from  $\mathbb{R}^n$  into  $\mathbb{R}$  such that

$$p_{\mathbf{H}}(\boldsymbol{\eta}) = c_n e^{-\Phi(\boldsymbol{\eta})}. \quad (2.14)$$

From Eqs. (2.7), (2.8) and (2.14), it can be deduced that

$$c_n = \frac{1}{(\sqrt{2\pi} \widehat{s}_n)^n}, \quad \Phi(\boldsymbol{\eta}) = -\log\{q(\boldsymbol{\eta})\}, \quad (2.15)$$

in which  $\boldsymbol{\eta} \mapsto q(\boldsymbol{\eta})$  is the continuously differentiable function from  $\mathbb{R}^n$  into  $\mathbb{R}^{+*} = ]0, +\infty[$  defined by

$$q(\boldsymbol{\eta}) = \frac{1}{\nu} \sum_{\ell=1}^{\nu} \exp\left\{-\frac{1}{2\widehat{s}_n^2} \left\| \frac{\widehat{s}_n}{s_n} \boldsymbol{\eta}^{\text{exp},\ell} - \boldsymbol{\eta} \right\|^2\right\}, \quad (2.16)$$

and where  $s_n$  and  $\widehat{s}_n$  are given by Eq. (2.9). It can then be deduced that  $\boldsymbol{\eta} \mapsto \Phi(\boldsymbol{\eta})$  is a continuously differentiable function on  $\mathbb{R}^n$ .

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

$$d\mathbf{U}(r) = \mathbf{V}(r) dr, \quad (2.17)$$

$$d\mathbf{V}(r) = -\nabla_{\mathbf{u}}\Phi(\mathbf{U}(r)) dr - \frac{1}{2}f_0\mathbf{V}(r) dr + \sqrt{f_0}d\mathbf{W}(r), \quad (2.18)$$

with the initial condition

$$\mathbf{U}(0) = \mathbf{u}_0, \quad \mathbf{V}(0) = \mathbf{v}_0 \quad a.s., \quad (2.19)$$

in which  $\mathbf{u}_0$  and  $\mathbf{v}_0$  are given vectors in  $\mathbb{R}^n$  (that will be taken as zero in the application presented later), and 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  $\mathbb{R}^n$ . 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  $[R_{\mathbf{W}}(r, r')] = \min(r, r')[I_n]$ . In Eq. (2.18), the free parameter  $f_0 > 0$  will allow a dissipation term to be introduced in the nonlinear second-order dynamical system (formulated in the Hamiltonian form with an additional dissipative term) in order to kill the transient part of the response and consequently, to get more rapidly the stationary solution corresponding to the invariant measure. It can easily be proved that function  $\mathbf{u} \mapsto \Phi(\mathbf{u})$ : (i) is continuous on  $\mathbb{R}^n$ , (ii) is such that  $\mathbf{u} \mapsto \|\nabla_{\mathbf{u}}\Phi(\mathbf{u})\|$  is a locally bounded function on  $\mathbb{R}^n$  (*i.e.* is bounded on all compact sets in  $\mathbb{R}^n$  that is the case because  $\mathbf{u} \mapsto \Phi(\mathbf{u})$  is a continuously differentiable function on  $\mathbb{R}^n$ ), and (iii) is such that,

$$\inf_{\|\mathbf{u}\|>R} \Phi(\mathbf{u}) \rightarrow +\infty \quad \text{if} \quad R \rightarrow +\infty, \quad (2.20)$$

$$\inf_{\mathbf{u} \in \mathbb{R}^n} \Phi(\mathbf{u}) = \Phi_{\min} \quad \text{with} \quad \Phi_{\min} \in \mathbb{R}, \quad (2.21)$$

$$\int_{\mathbb{R}^n} \|\nabla_{\mathbf{u}}\Phi(\mathbf{u})\| p_{\mathbf{H}}(\mathbf{u}) d\mathbf{u} < +\infty \quad \text{with} \quad p_{\mathbf{H}}(\mathbf{u}) = c_n e^{-\Phi(\mathbf{u})}. \quad (2.22)$$

Under hypotheses (i) to (iii), and using Theorems 4 to 7 in pages 211 to 216 of Ref. [37], in which the Hamiltonian is taken as  $H(\mathbf{u}, \mathbf{v}) = \|\mathbf{v}\|^2/2 + \Phi(\mathbf{u})$ , and using [12, 25] for the ergodic property, it can be deduced that the problem defined by Eqs. (2.17) to (2.19) admits a unique solution. This solution is a second-order diffusion stochastic process  $\{(\mathbf{U}(r), \mathbf{V}(r)), r \in \mathbb{R}^+\}$ , which converges to a stationary and ergodic diffusion stochastic process  $\{(\mathbf{U}_{\text{st}}(r_{\text{st}}), \mathbf{V}_{\text{st}}(r_{\text{st}})), r_{\text{st}} \geq 0\}$ , when  $r$  goes to infinity, associated with the invariant probability measure  $P_{\text{st}}(d\mathbf{u}, d\mathbf{v}) = \rho_{\text{st}}(\mathbf{u}, \mathbf{v}) d\mathbf{u} d\mathbf{v}$ . The probability density function  $(\mathbf{u}, \mathbf{v}) \mapsto \rho_{\text{st}}(\mathbf{u}, \mathbf{v})$  on  $\mathbb{R}^n \times \mathbb{R}^n$  is the unique solution of the steady-state Fokker-Planck equation with the normalization condition, associated with Eqs. (2.17) and (2.18), and is written (see Propositions 8 and 9 in pages 120 to 123 of Ref. [37]), as

$$\rho_{\text{st}}(\mathbf{u}, \mathbf{v}) = c_0 \exp\left\{-\frac{1}{2}\|\mathbf{v}\|^2 - \Phi(\mathbf{u})\right\}, \quad (2.23)$$

in which  $c_0$  is the constant of normalization. From Eqs. (2.14) and (2.23), it can be deduced that

$$p_{\mathbf{H}}(\boldsymbol{\eta}) = \int_{\mathbb{R}^n} \rho_{\text{st}}(\boldsymbol{\eta}, \mathbf{v}) d\mathbf{v} \quad , \quad \forall \boldsymbol{\eta} \in \mathbb{R}^n. \quad (2.24)$$

It can therefore be concluded that random variable  $\mathbf{H}$  for which the multimodal probability density function is  $p_{\mathbf{H}}$ , can be defined, for any  $r_{\text{st}} > 0$  as

$$\mathbf{H} = \mathbf{U}_{\text{st}}(r_{\text{st}}) = \lim_{r \rightarrow +\infty} \mathbf{U}(r) \quad \text{in probability distribution.} \quad (2.25)$$

As explained above, the free parameter  $f_0 > 0$  introduced in Eq. (2.18), 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 corresponding to the stationary and ergodic solution associated with the invariant measure.

**Remark 3.** Instead of Eq. (2.18), the following equation  $d\mathbf{V}(r) = -\nabla_{\mathbf{u}}\Phi(\mathbf{U}(r)) dr - \frac{1}{2}f_0[D_0]\mathbf{V}(r) dr + \sqrt{f_0}[S_0]d\mathbf{W}(r)$  could be used, in which  $[S_0]$  would belong to  $\mathbb{M}_n(\mathbb{R})$  and where  $[D_0]$  would be a positive symmetric matrix such that  $[D_0] = [S_0][S_0]^T$  with  $1 \leq \text{rank}[D_0] \leq n$ . If such an equation were used, then the invariant measure would always be given by Eq. (2.23) (see page 244 of Ref. [37]). In particular, a diagonal positive-definite damping matrix  $\frac{1}{2}f_0[D_0]$  could be chosen in order trying to increase the speed of convergence towards the stationary and ergodic solution of the Itô equation. However, in order not to complicate too much setting data parameters of the algorithm, while maintaining good control of the speed of convergence towards the stationary and ergodic solution, the simpler form defined by Eq. (2.18) is proposed.

**2.4.2. Discretization scheme of the ISDE.** A discretization scheme must be used for numerically solving the ISDE defined by Eqs. (2.17) to (2.19). For general surveys on discretization schemes for Itô stochastic differential equations, we refer the reader to [26, 45, 46]. Concerning the particular cases related to Hamiltonian dynamical systems (which have also been analyzed in [47] using an implicit Euler scheme), we propose to use the Störmer-Verlet scheme, which is a very efficient scheme that preserves energy for nondissipative Hamiltonian dynamical systems (see [21] for reviews about this scheme in the deterministic case, and see [7] and the references therein for the stochastic case). The Störmer-Verlet scheme has been validated in [20], for solving an ISDE of the type defined

by Eqs. (2.17) to (2.19), and corresponding to a weakly dissipative Hamiltonian dynamical system. We then propose to reuse hereinafter the Störmer-Verlet scheme proposed in [20].

Let  $M \geq 1$  be an integer. The Itô stochastic differential equation defined by Eqs. (2.17) and (2.18) with the initial condition defined by Eq. (2.19), is solved on the finite interval  $\mathcal{R} = [0, (M-1)\Delta r]$ , in which  $\Delta r$  is the sampling step of the continuous 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  $k = 1, \dots, M$ . The following notations are introduced:  $\mathbf{U}^k = \mathbf{U}(r_k)$ ,  $\mathbf{V}^k = \mathbf{V}(r_k)$ , and  $\mathbf{W}^k = \mathbf{W}(r_k)$ , for  $k = 1, \dots, M$ , with

$$\mathbf{U}^1 = \mathbf{u}_0 \quad , \quad \mathbf{V}^1 = \mathbf{v}_0 \quad , \quad \mathbf{W}^1 = \mathbf{0} \quad a.s. \quad (2.26)$$

Let  $\{\Delta \mathbf{W}^{k+1} = \mathbf{W}^{k+1} - \mathbf{W}^k, k = 1, \dots, M-1\}$  be the family of independent Gaussian second-order centered  $\mathbb{R}^n$ -valued random variables such that  $E\{\Delta \mathbf{W}^{k+1} (\Delta \mathbf{W}^{k+1})^T\} = \Delta r [I_n]$ . For  $k = 1, \dots, M-1$ , the Störmer-Verlet scheme applied to Eqs. (2.17) and (2.18) is written as

$$\begin{aligned} \mathbf{U}^{k+\frac{1}{2}} &= \mathbf{U}^k + \frac{\Delta r}{2} \mathbf{V}^k \quad , \\ \mathbf{V}^{k+1} &= \frac{1-b}{1+b} \mathbf{V}^k + \frac{\Delta r}{1+b} \mathbf{L}^{k+\frac{1}{2}} + \frac{\sqrt{f_0}}{1+b} \Delta \mathbf{W}^{k+1} \quad , \\ \mathbf{U}^{k+1} &= \mathbf{U}^{k+\frac{1}{2}} + \frac{\Delta r}{2} \mathbf{V}^{k+1} \quad , \end{aligned} \quad (2.27)$$

with the initial condition defined by (2.26), where  $b = f_0 \Delta r / 4$ , and where  $\mathbf{L}^{k+\frac{1}{2}}$  is the  $\mathbb{R}^n$ -valued random variable such that

$$\mathbf{L}^{k+\frac{1}{2}} = -\{\nabla_{\mathbf{u}} \Phi(\mathbf{u})\}_{\mathbf{u}=\mathbf{U}^{k+\frac{1}{2}}} \quad . \quad (2.28)$$

From Eqs. (2.15) and (2.16), it can be deduced that,

$$\nabla_{\mathbf{u}} \Phi(\mathbf{u}) = -\frac{1}{q(\mathbf{u})} \nabla_{\mathbf{u}} q(\mathbf{u}) \quad , \quad (2.29)$$

$$\nabla_{\mathbf{u}} q(\mathbf{u}) = \frac{1}{\hat{s}_n^2} \frac{1}{\nu} \sum_{\ell=1}^{\nu} \left( \frac{\hat{s}_n}{s_n} \boldsymbol{\eta}^{\text{exp}, \ell} - \mathbf{u} \right) \exp\left\{-\frac{1}{2\hat{s}_n^2} \left\| \frac{\hat{s}_n}{s_n} \boldsymbol{\eta}^{\text{exp}, \ell} - \mathbf{u} \right\|^2\right\} \quad . \quad (2.30)$$

**2.4.3. Choosing the parameters for numerical integration of the ISDE.** In this section, we construct the values of the parameters  $\Delta r, M_0, M, f_0, \mathbf{u}_0$  and  $\mathbf{v}_0$ , which are used in the discretization scheme of the ISDE, presented in Section 2.4.2. First, we associate with the nonlinear Hamiltonian dynamical system, a linearized diagonal second-order dynamical system in  $\mathbf{U}^{\text{lin}} = (U_1^{\text{lin}}, \dots, U_n^{\text{lin}})$  (the components are not coupled) such that, for all  $i$  in  $\{1, \dots, n\}$ ,  $\ddot{U}_i^{\text{lin}}(r) + \frac{1}{2} f_0 \dot{U}_i^{\text{lin}}(r) + \mathbb{K}_i U_i^{\text{lin}}(r) = \sqrt{f_0} \dot{W}_i(r)$ , in which  $\dot{\mathbf{W}} = (\dot{W}_1, \dots, \dot{W}_n)$  is the generalized Gaussian white stochastic process (the generalized derivative of  $\mathbf{W}$ ). The behavior of the nonlinear stiffness force  $\mathbf{F}(\mathbf{u}) = -\mathbf{L}(\mathbf{u})$ , with  $\mathbf{u} = (u_1, \dots, u_n)$  and  $\mathbf{F}(\mathbf{u}) = (F_1(\mathbf{u}), \dots, F_n(\mathbf{u}))$ , can have some fluctuations in the neighborhood of  $\mathbf{u} = \mathbf{0}$ , and is such that  $\mathbf{F}(\mathbf{0}) \neq \mathbf{0}$ . Consequently, we cannot calculate  $\mathbb{K}_i$  in writing  $\mathbb{K}_i u_i = \{\partial F_i(\mathbf{u}) / \partial u_i\}_{\mathbf{u}=\mathbf{0}}$ . We propose to replace this equation by an incremental equation on a symmetric interval  $[-\Delta, \Delta]$  for a sufficiently large increment  $\Delta > 0$  (typically, the problem under consideration being normalized to 1,  $\Delta$  can, for instance, be chosen as 5), which yields  $\mathbb{K}_i = (F_i(\mathbf{u}^i) - F_i(-\mathbf{u}^i)) / (2\Delta)$  in which  $\mathbf{u}^i = (u_1^i, \dots, u_n^i)$  with  $u_j^i = \Delta \delta_{ij}$ . Let  $0 < \lambda_1 \leq \dots \leq \lambda_n$  be the eigenvalues of matrix  $[\mathbb{K}]$ , and let be  $\omega_{\min} = \sqrt{\lambda_1}$  and  $\omega_{\max} = \sqrt{\lambda_n}$ .

(i) A first estimation of  $\Delta r$  is chosen as  $\Delta r_0 = \pi/(10\omega_{\max})$ . An oversampling  $m_{\text{overs}} > 1$  is introduced to get a sufficient accuracy of the Störmer-Verlet scheme (for instance,  $m_{\text{overs}} = 10$ ), and yields  $\Delta r = \Delta r_0/m_{\text{overs}}$  (a convergence analysis with respect to  $m_{\text{overs}}$  must be carried out).

(ii) The minimum damping rate,  $\zeta_{\min}$  of the linear second-order dynamical system is such that  $2\zeta_{\min}\omega_{\min} = f_0/2$  that yields  $f_0 = 4\zeta_{\min}\omega_{\min}$ . The damping rate  $\zeta_{\min}$  is chosen, for instance, as 0.7 to rapidly kill the transient response induced by the initial conditions (that are not distributed following the invariant measure).

(iii) The larger relaxation "time" of the linear second-order dynamical system can be defined as  $r_0$  such that  $\exp\{-\zeta_{\min}\omega_{\min}r_0\} = \varepsilon_0$  with  $\varepsilon_0 \ll 1$ , which yields  $r_0 = -\log(\varepsilon_0)/(\zeta_{\min}\omega_{\min})$  (for instance, the value of  $\varepsilon_0$  can be chosen as  $1/200$ ). The parameter  $M_0$  is then defined as  $r_0 = M_0\Delta r$  that yields  $M_0 = 1 + \text{fix}(r_0/\Delta r)$  in which  $\text{fix}(x)$  rounds  $x$  to the nearest integer towards zero. Value  $r_0$  (and then integer  $M_0$ ) corresponds to reaching the stationary response. Integer  $M_0$  will be used for calculating independent realizations of  $\mathbf{H}$ .

(iv) The initial conditions are chosen as  $\mathbf{u}_0 = \mathbf{v}_0 = \mathbf{0}$ .

(v-1) *Ergodic method.* The integer  $M$  is defined as  $M = m_{\text{ergo}}M_0$  where  $m_{\text{ergo}} \gg 1$  is an integer which has to be chosen in order to reach a reasonable convergence for estimating  $E\{w(\mathbf{H})\}$  using the ergodic method (see Eq. (2.33) that we introduce later). For instance, an initial value for  $m_{\text{ergo}}$  can be chosen as 200 or 400, but a convergence analysis must be carried out with respect to  $m_{\text{ergo}}$ .

(v-2) *Monte Carlo method.* An integer  $\widehat{M}_0$  is introduced such that  $\widehat{M}_0 = i_0M_0$ . The integer  $i_0$  is chosen in order that the sequence  $\{\mathbf{U}^k, k \geq \widehat{M}_0\}$  corresponds to the stationary solution. Taking into account the construction of integer  $M_0$ , the integer  $i_0$  could be chosen to 1 or 2. We then introduce the sequence of integers  $\{\widehat{M}_\ell, \ell = 1, \dots, \nu_s\}$  such that  $\widehat{M}_\ell = (1 + \ell)\widehat{M}_0$ . Integer  $M$  is then defined by  $M = \widehat{M}_{\nu_s} = (1 + \nu_s)\widehat{M}_0$ . By construction of  $\widehat{M}_0$ , the vectors  $\mathbf{U}^{\widehat{M}_\ell}(\theta)$  and  $\mathbf{U}^{\widehat{M}_{\ell+1}}(\theta)$  can be considered as two independent realizations of the random vector  $\mathbf{U}(r)$  for any fixed  $r$  such that  $r > r_0$  (stationary part of the response). The integer  $\nu_s$  is chosen in order to reach a reasonable convergence for estimating  $E\{w(\mathbf{H})\}$  using the Monte Carlo method (see Eq. (2.34) that we introduce later). For instance, an initial value for  $\nu_s$  can be chosen as 200, but a convergence analysis with respect to  $\nu_s$  must be carried out.

**2.4.4. Random generator of independent realizations of  $\mathbf{H}$ .** A random generator of  $\nu_s$  independent realizations,  $\mathbf{H}(\theta_1), \dots, \mathbf{H}(\theta_{\nu_s})$ , of random vector  $\mathbf{H}$  whose multimodal probability distribution is  $p_{\mathbf{H}}(\boldsymbol{\eta})d\boldsymbol{\eta}$ , is constructed as follows. Let  $\{\mathbf{U}^k(\theta), k = 1, \dots, M\}$  be constructed using the algorithm presented in Section 2.4.2. Using the sequence  $\{\widehat{M}_\ell, \ell = 1, \dots, \nu_s\}$  defined in Section 2.4.3-(v-2), each independent realization  $\mathbf{H}(\theta_\ell)$  can then be obtained, for  $\ell$  in  $\{1, \dots, \nu_s\}$ , as

$$\boldsymbol{\eta}^{\text{sim},\ell} = \mathbf{H}(\theta_\ell) = \mathbf{U}^{\widehat{M}_\ell}(\theta) \quad , \quad \widehat{M}_\ell = (1 + \ell)\widehat{M}_0. \quad (2.31)$$

**2.5. Estimating  $E\{w(\mathbf{H})\}$ .** In this section, two methods are proposed for estimating  $E\{w(\mathbf{H})\}$ : (i) and ergodic method, and (ii) a Monte Carlo method.

(i) *Estimating  $E\{w(\mathbf{H})\}$  with the ergodic method.* For any realization  $\theta$  in  $\Theta$ , let  $\{(\mathbf{U}(r, \theta), \mathbf{V}(r, \theta)), r \geq 0\}$  be the solution of Eqs. (2.17) to (2.19). Then using the ergodic property,  $E\{w(\mathbf{H})\} = \int_{\mathbb{R}^n} w(\boldsymbol{\eta}) p_{\mathbf{H}}(\boldsymbol{\eta}) d\boldsymbol{\eta}$  can be estimated (see [12, 26, 46]) by

$$E\{w(\mathbf{H})\} = \lim_{R \rightarrow +\infty} \frac{1}{R} \int_{r_{M_0}}^{r_{M_0}+R} w(\mathbf{U}(r, \theta)) dr \quad \text{with probability 1,} \quad (2.32)$$

in which  $r_{M_0} = (M_0 - 1) \Delta r$  with  $M_0$  a fixed integer greater than 1. The parameter  $M_0$  (estimated as explained in Section 2.4.3-(iii)) allows us to remove the transient part of the response induced by the initial condition. Let  $\{\mathbf{U}^k(\theta), k = 1, \dots, M\}$  be the corresponding realization of the discretized solution constructed as explained in Section 2.4.2, in which  $M$  has been estimated in Section 2.4.3-(iv). Thus, the numerical approximation of Eq. (2.32) is written as

$$E\{w(\mathbf{H})\} = \lim_{M \rightarrow +\infty} \widehat{w}_M^{\text{ER}}, \quad \widehat{w}_M^{\text{ER}} = \frac{1}{M - M_0 + 1} \sum_{k=M_0}^M w(\mathbf{U}^k(\theta)). \quad (2.33)$$

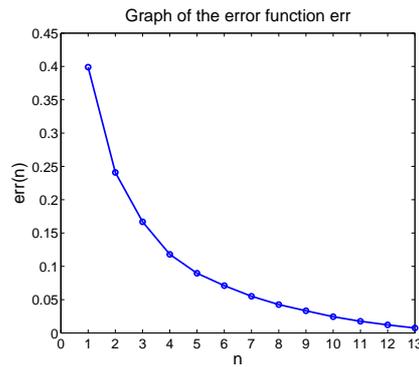
(ii) *Estimating  $E\{w(\mathbf{H})\}$  with the Monte Carlo method.* Using Eq. (2.31), the Monte Carlo method for estimating  $E\{w(\mathbf{H})\}$  yields

$$E\{w(\mathbf{H})\} = \lim_{\nu_s \rightarrow +\infty} \widehat{w}_{\nu_s}^{\text{MC}}, \quad \widehat{w}_{\nu_s}^{\text{MC}} = \frac{1}{\nu_s} \sum_{\ell=1}^{\nu_s} w(\mathbf{U}^{\widehat{M}_\ell}(\theta)), \quad \widehat{M}_\ell = (1 + \ell) \widehat{M}_0, \quad (2.34)$$

## 2.6. Numerical application.

**2.6.1. Data generation for the numerical application.** We consider the data description introduced in Section 2.1 for  $N = 100$  and  $\nu = 500$ . The algorithm that has been used for generating the  $\nu$  independent realizations  $\mathbf{x}^{\text{exp},1}, \dots, \mathbf{x}^{\text{exp},\nu}$  of random vector  $\mathbf{X}$  with values in  $\mathbb{R}^N$ , is described in Appendix A. The reader can then simulate the "data description" used in this numerical application.

**2.6.2. Defining the optimal values of parameters.** The estimations  $\widehat{\mathbf{m}}_{\mathbf{X}}$  and  $[\widehat{C}_{\mathbf{X}}]$  are computed using Eqs. (2.1) and (2.2). Concerning the construction of the reduced-order statistical model, Fig. 2.1 displays the graph of the error function defined by Eq. (2.4). We choose  $n = 5$  corresponding to an error of 0.089. With the values of the main parameters,  $N = 100$ ,  $\nu = 500$  and  $n = 5$ , for the numerical integration of the ISDE (see Section 2.4.3), the optimal value of the damping rate has been found to  $\zeta_{\min} = 0.7$  that yields  $f_0 = 5.0565$ . The optimal value of the relaxation "time" parameter has been found to  $\varepsilon_0 = 1/200$  (see Section 2.4.3-(iii)), and  $\Delta r_0 = 0.1670$ . The others values of the parameters are analyzed in the next section.



**Figure 2.1.** Reduced-order statistical model: graph of the error function  $n \mapsto \text{err}(n)$

**2.6.3. Convergence analysis.** The convergence analysis is carried out with the ergodic method and with the Monte Carlo method.

(i) *Ergodic method.* The parameters defined in Section 2.6.2 are fixed. In order to analyze the convergence of Eq. (2.33) with respect to  $M$  for the ergodic estimation, using the integration scheme defined in Section 2.4.2, two error functions,  $\text{err}_1(M)$  and  $\text{err}_2(M)$  are introduced. The first one is related to the estimation of the mean value  $E\{\mathbf{H}\} = \widehat{\mathbf{m}}_{\mathbf{H}}$  that must be equal to  $\mathbf{0}$  (see Eq. (2.10)). In Eq. (2.33), we choose  $w(\mathbf{H}) = \mathbf{H}$ , and then

$$\text{err}_1^{\text{ER}}(M) = \frac{1}{n} \sum_{i=1}^n |\{\widehat{\mathbf{m}}_{\mathbf{H}} - \widehat{\mathbf{w}}_M^{\text{ER}}\}_i|, \quad \widehat{\mathbf{w}}_M^{\text{ER}} = \frac{1}{M - M_0 + 1} \sum_{k=M_0}^M \mathbf{U}^k(\theta). \quad (2.35)$$

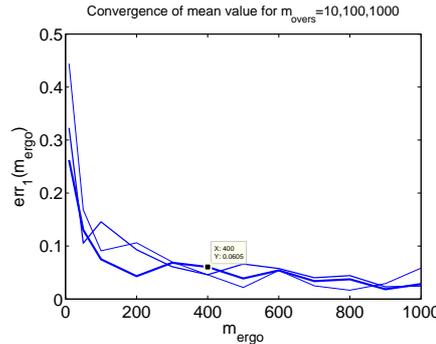
The second one is related to the estimation of the second-order moment matrix  $E\{\mathbf{H}\mathbf{H}^T\} = [\widehat{R}_{\mathbf{H}}]$  that must be equal to  $[I_n]$  (see Eq. (2.11)). In Eq. (2.33), we choose  $w(\mathbf{H}) = \mathbf{H}\mathbf{H}^T$ , and then

$$\text{err}_2^{\text{ER}}(M) = \frac{\|[\widehat{R}_{\mathbf{H}}] - [\widehat{w}_M^{\text{ER}}]\|_F}{\|[\widehat{R}_{\mathbf{H}}]\|_F}, \quad [\widehat{w}_M^{\text{ER}}] = \frac{1}{M - M_0 + 1} \sum_{k=M_0}^M \mathbf{U}^k(\theta) \mathbf{U}^k(\theta)^T, \quad (2.36)$$

in which  $\|\cdot\|_F$  is the Frobenius norm. For  $m_{\text{overs}} = 10; 100; \text{ and } 1,000$  (with  $\Delta r = \Delta r_0/m_{\text{overs}}$ ), we have  $M_0 = 260; 2,600; \text{ and } 26,000$ . Figures. 2.2 and 2.3 display the graphs of function  $m_{\text{ergo}} \mapsto \text{err}_1^{\text{ER}}(m_{\text{ergo}} M_0)$  and  $m_{\text{ergo}} \mapsto \text{err}_2^{\text{ER}}(m_{\text{ergo}} M_0)$  (with  $M = m_{\text{ergo}} M_0$ ).

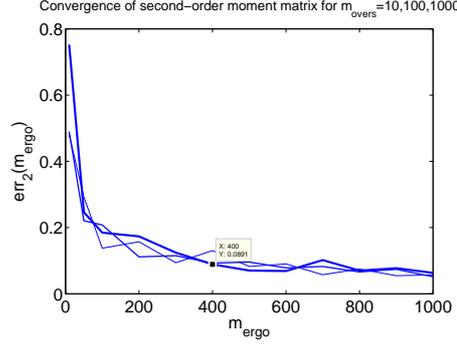
In all the paper, the following parameters are fixed to the values:  $f_0 = 5.0565$ ,  $\varepsilon_0 = 1/200$ ,  $m_{\text{overs}} = 10$  that yields  $M_0 = 260$ ,  $\Delta r = 0.01638$  and  $r_0 = 4.26$ .

A reasonably good accuracy is obtained in using  $m_{\text{ergo}} = 400$  that yields  $M = 104,000$  for which  $\text{err}_1^{\text{ER}}(M) = 0.060$  and  $\text{err}_2^{\text{ER}}(M) = 0.089$ . The value  $m_{\text{ergo}} = 10,000$  or  $100,000$  ( $M = 2,600,000$  or  $26,000,000$ ) yields  $\text{err}_1^{\text{ER}}(M) = 0.0080$  and  $\text{err}_2^{\text{ER}}(M) = 0.0264$ , or  $\text{err}_1^{\text{ER}}(M) = 0.0021$  and  $\text{err}_2^{\text{ER}}(M) = 0.0061$ .



**Figure 2.2.** Graph of error function  $m_{\text{ergo}} \mapsto \text{err}_1^{\text{ER}}(m_{\text{ergo}} M_0)$  for  $m_{\text{overs}} = 10$  (thick solid line),  $m_{\text{overs}} = 100$  (mid solid line),  $m_{\text{overs}} = 1000$  (thin solid line)

(ii) *Monte Carlo method.* Similarly to the ergodic method presented before, two error functions are introduced. The first one is related to the estimation of the mean value  $E\{\mathbf{H}\} = \widehat{\mathbf{m}}_{\mathbf{H}}$  that must be



**Figure 2.3.** Graph of error function  $m_{\text{ergo}} \mapsto \text{err}_2^{\text{ER}}(m_{\text{ergo}} M_0)$  for  $m_{\text{overs}} = 10$  (thick solid line),  $m_{\text{overs}} = 100$  (mid solid line),  $m_{\text{overs}} = 1000$  (thin solid line)

equal to  $\mathbf{0}$  (see Eq. (2.10)), and yields

$$\text{err}_1^{\text{MC}}(\nu_s) = \frac{1}{n} \sum_{i=1}^n |\{\widehat{\mathbf{m}}_{\mathbf{H}} - \widehat{\mathbf{w}}_{\nu_s}^{\text{MC}}\}_i|, \quad \widehat{\mathbf{w}}_{\nu_s}^{\text{MC}} = \frac{1}{\nu_s} \sum_{\ell=1}^{\nu_s} \mathbf{U}^{\widehat{M}_\ell}(\theta), \quad \widehat{M}_\ell = (1 + \ell) \widehat{M}_0. \quad (2.37)$$

The second one is related to the estimation of the second-order moment matrix  $E\{\mathbf{H}\mathbf{H}^T\} = [\widehat{R}_{\mathbf{H}}]$  that must be equal to  $[I_n]$  (see Eq. (2.11)), and yields, for  $\widehat{M}_\ell = (1 + \ell) \widehat{M}_0$ ,

$$\text{err}_2^{\text{MC}}(\nu_s) = \frac{\|[\widehat{R}_{\mathbf{H}}] - [\widehat{w}_{\nu_s}^{\text{MC}}]\|_F}{\|[\widehat{R}_{\mathbf{H}}]\|_F}, \quad [\widehat{w}_{\nu_s}^{\text{MC}}] = \frac{1}{\nu_s - 1} \sum_{\ell=1}^{\nu_s} \mathbf{U}^{\widehat{M}_\ell}(\theta) (\mathbf{U}^{\widehat{M}_\ell}(\theta))^T. \quad (2.38)$$

A numerical study has been performed with  $f_0 = 5.0565$ ,  $\varepsilon_0 = 1/200$ ,  $m_{\text{overs}} = 10$  that yields  $M_0 = 260$ ,  $\Delta r = 0.01638$  and  $r_0 = 4.26$ . A sensitivity analysis has been carried out with respect to the value of  $i_0$  that is such that  $\widehat{M}_0 = i_0 M_0$ . The values 1, 2, and 3 for  $i_0$  yield very close results and the differences are not significant. Consequently,  $i_0$  is fixed to the value 1. For these values of the parameters, the graphs of the error functions  $\nu_s \mapsto \text{err}_1^{\text{MC}}(\nu_s)$  and  $\nu_s \mapsto \text{err}_2^{\text{MC}}(\nu_s)$  have been constructed and yield similar results to those shown in Figs. 2.2 and 2.3. A reasonably good accuracy is obtained in using  $\nu_s = 400$  that yields  $M = 104,000$  for which  $\text{err}_1^{\text{MC}}(\nu_s) = 0.0692$  and  $\text{err}_2^{\text{MC}}(\nu_s) = 0.1118$ . The value  $\nu_s = 10,000$  ( $M = 2,600,000$ ) yields  $\text{err}_1^{\text{MC}}(\nu_s) = 0.0081$  and  $\text{err}_2^{\text{MC}}(\nu_s) = 0.034$ , while for  $\nu_s = 100,000$  ( $M = 26,000,000$ ) yields  $\text{err}_1^{\text{MC}}(\nu_s) = 0.0016$  and  $\text{err}_2^{\text{MC}}(\nu_s) = 0.0079$ . In comparing these errors with those given by the ergodic method, it can be concluded that, for a same value of the CPU time, the ergodic method is slightly more efficient. Nevertheless, if we consider the estimation of statistical quantities related to the random variable  $\Xi$  that results from a nonlinear transformation  $\mathfrak{h}$  of  $\mathbf{H}$  such that  $\Xi = \mathfrak{h}(\mathbf{H})$ , transformation  $\mathfrak{h}$  has to be evaluated  $M$  times with the ergodic method, but only  $\nu_s$  times with the Monte Carlo method. Since  $M/\nu_s = \widehat{M}_0$ , the gain is large enough when using the Monte Carlo method with respect to the ergodic method.

**2.6.4. Estimation of the probability density function of  $\mathbf{H}$ .** In order to numerically validate the generator (see Sections 2.4.2 and 2.5-(i)) of random vector  $\mathbf{H}$  whose probability density

function  $p_{\mathbf{H}}$  is defined by Eqs. (2.7) to (2.9), we propose to compare  $p_{\mathbf{H}}$  with its estimation  $\widehat{p}_{\mathbf{H}}$  constructed with the generator. This construction used the Gaussian kernel-density estimation method presented in Section 2.3. For all  $\boldsymbol{\eta}$  in  $\mathbb{R}^n$ , we have  $p_{\mathbf{H}}(\boldsymbol{\eta}) = E\{\delta_0(\mathbf{H} - \boldsymbol{\eta})\}$  that can be rewritten as

$$p_{\mathbf{H}}(\boldsymbol{\eta}) = \lim_{\sigma_n \rightarrow 0^+} E\{\mu_{n,\sigma_n}(\mathbf{H} - \boldsymbol{\eta})\}, \quad (2.39)$$

in which  $\boldsymbol{\eta} \mapsto \mu_{n,\sigma_n}(\boldsymbol{\eta})$  is the following Gaussian function from  $\mathbb{R}^n$  into  $]0, +\infty[$ , defined, for all  $\boldsymbol{\eta}$  in  $\mathbb{R}^n$ , by

$$\mu_{n,\sigma_n}(\boldsymbol{\eta}) = \frac{1}{(\sqrt{2\pi} \sigma_n)^n} \exp\left\{-\frac{1}{2\sigma_n^2} \|\boldsymbol{\eta}\|_n^2\right\}, \quad (2.40)$$

in which  $\|\boldsymbol{\eta}\|_n^2 = \eta_1^2 + \dots + \eta_n^2$ .

(i) *Ergodic method.* From Eqs. (2.33) and (2.40), the estimation  $\widehat{p}_{\mathbf{H}}$  of  $p_{\mathbf{H}}$  is written, for all  $\boldsymbol{\eta}$  in  $\mathbb{R}^n$ , as

$$\widehat{p}_{\mathbf{H}}(\boldsymbol{\eta}) = \frac{1}{M - M_0 + 1} \sum_{k=M_0}^M \mu_{n,\widehat{\sigma}_n^{\text{ER}}} \left( \frac{\widehat{\sigma}_n^{\text{ER}}}{\sigma_n^{\text{ER}}} \mathbf{U}^k(\boldsymbol{\theta}) - \boldsymbol{\eta} \right), \quad (2.41)$$

in which the positive parameters  $\sigma_n^{\text{ER}}$  and  $\widehat{\sigma}_n^{\text{ER}}$  are defined by

$$\sigma_n^{\text{ER}} = \left\{ \frac{4}{(M - M_0 + 1)(2 + n)} \right\}^{1/(n+4)}, \quad \widehat{\sigma}_n^{\text{ER}} = \frac{\sigma_n^{\text{ER}}}{\sqrt{(\sigma_n^{\text{ER}})^2 + \frac{M - M_0}{M - M_0 + 1}}}. \quad (2.42)$$

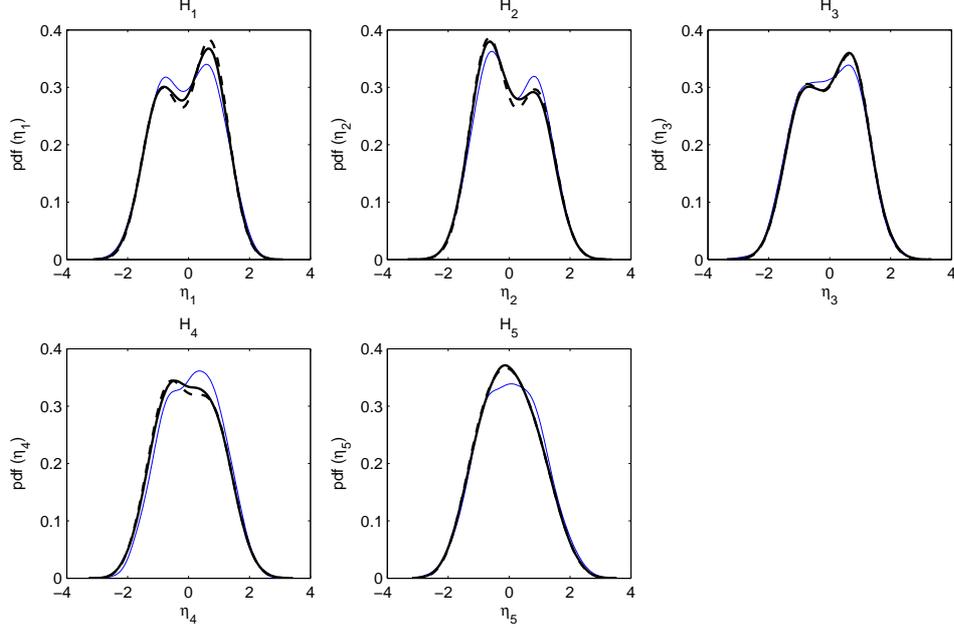
For  $i$  fixed in  $\{1, \dots, n\}$ , the probability density function  $p_{H_i}$  on  $\mathbb{R}$  of random variable  $H_i$  is calculated in using Eq. (2.12). Integrating Eq. (2.41) over  $\mathbb{R}^{n-1}$  yields the estimation  $\eta_i \mapsto \widehat{p}_{H_i}(\eta_i)$  of the probability density function  $p_{H_i}$  on  $\mathbb{R}$ ,

$$\widehat{p}_{H_i}(\eta_i) = \frac{1}{M - M_0 + 1} \sum_{k=M_0}^M \mu_{1,\widehat{\sigma}_n^{\text{ER}}} \left( \frac{\widehat{\sigma}_n^{\text{ER}}}{\sigma_n^{\text{ER}}} U_i^k(\boldsymbol{\theta}) - \eta_i \right). \quad (2.43)$$

For  $i$  in  $\{1, \dots, n\}$ , Fig. 2.4 displays the graph of the probability density function  $p_{H_i}$  calculated with Eq. (2.12), which is compared with  $\widehat{p}_{H_i}$  calculated with Eq. (2.43) for which  $m_{\text{ergo}} = 400$  and  $m_{\text{ergo}} = 10,000$ . It can be seen that convergence is reached for  $m_{\text{ergo}} = 10,000$ , and that a good approximation is obtained for  $m_{\text{ergo}} = 400$ . For  $i$  and  $j$  fixed in  $\{1, \dots, n\}$ , the joint probability density function  $p_{H_i H_j}$  of random variables  $H_i$  and  $H_j$  is calculated in using again Eq. (2.13). Using the ergodic method, and in integrating Eq. (2.41) over  $\mathbb{R}^{n-2}$  yields the estimation  $\widehat{p}_{H_i H_j}$  of the joint probability density function  $p_{H_i H_j}$  on  $\mathbb{R}^2$ , which is written as

$$\widehat{p}_{H_i H_j}(\eta_i, \eta_j) = \frac{1}{M - M_0 + 1} \sum_{k=M_0}^M \mu_{2,\widehat{\sigma}_n^{\text{ER}}} \left( \frac{\widehat{\sigma}_n^{\text{ER}}}{\sigma_n^{\text{ER}}} \mathbf{U}^{k,ij}(\boldsymbol{\theta}) - \boldsymbol{\eta}^{ij} \right), \quad (2.44)$$

in which  $\mathbf{U}^{k,ij}(\boldsymbol{\theta}) = (U_i^k(\boldsymbol{\theta}), U_j^k(\boldsymbol{\theta}))$  and  $\boldsymbol{\eta}^{ij} = (\eta_i, \eta_j)$  belong to  $\mathbb{R}^2$ . The computation has been carried out for the ten couples of indices, and all the obtained results have the same quality. Fig. 2.5



**Figure 2.4.** For  $i$  in  $\{1, \dots, n\}$ , graphs of the probability density functions  $\eta_i \mapsto p_{H_i}(\eta_i)$  (thick dashed line), and  $\eta_i \mapsto \hat{p}_{H_i}(\eta_i)$  for  $m_{\text{ergo}} = 400$  (thin solid line) and for  $m_{\text{ergo}} = 10,000$  (thick solid line).

displays the graph of the joint probability density function  $p_{H_i, H_j}$  calculated with Eq. (2.13), which is compared with  $\hat{p}_{H_i, H_j}$  calculated with Eq. (2.44) for which  $m_{\text{ergo}} = 10,000$ .

(ii) *Monte Carlo method.* From Eqs. (2.34) and (2.40), the estimation  $\hat{p}_{\mathbf{H}}$  of  $p_{\mathbf{H}}$  is written, for all  $\boldsymbol{\eta}$  in  $\mathbb{R}^n$ , as

$$\hat{p}_{\mathbf{H}}(\boldsymbol{\eta}) = \frac{1}{\nu_s} \sum_{\ell=1}^{\nu_s} \mu_{n, \hat{\sigma}_n^{\text{MC}}} \left( \frac{\hat{\sigma}_n^{\text{MC}}}{\sigma_n^{\text{MC}}} \mathbf{U}^{\widehat{M}_\ell}(\boldsymbol{\theta}) - \boldsymbol{\eta} \right) \quad , \quad \widehat{M}_\ell = (1 + \ell) \widehat{M}_0 \quad , \quad (2.45)$$

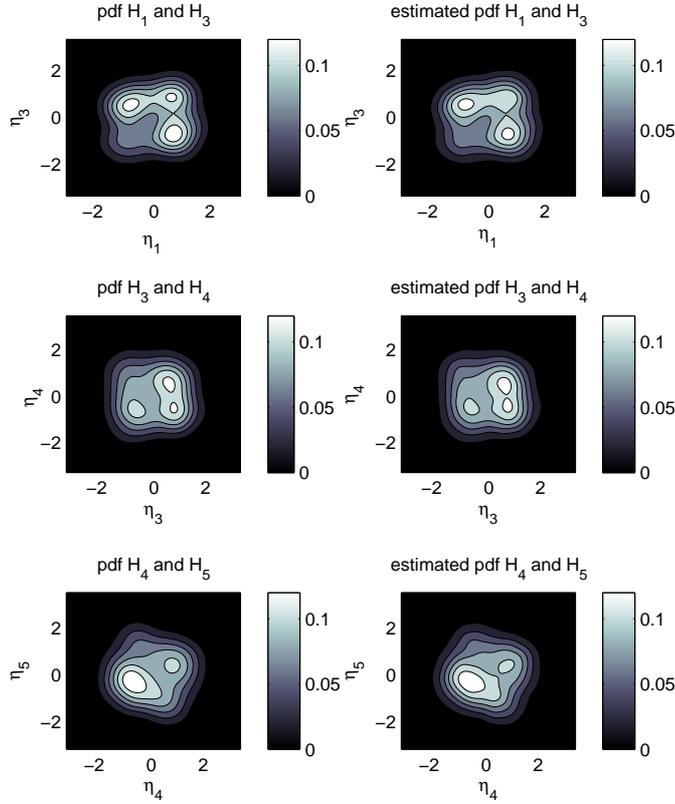
in which the positive parameters  $\sigma_n^{\text{MC}}$  and  $\hat{\sigma}_n^{\text{MC}}$  are defined by

$$\sigma_n^{\text{MC}} = \left\{ \frac{4}{\nu_s(2+n)} \right\}^{1/(n+4)} \quad , \quad \hat{\sigma}_n^{\text{MC}} = \frac{\sigma_n^{\text{MC}}}{\sqrt{(\sigma_n^{\text{MC}})^2 + (\nu_s - 1)/\nu_s}} \quad . \quad (2.46)$$

For  $i$  fixed in  $\{1, \dots, n\}$ , the estimation  $\hat{p}_{H_i}$  of the probability density function on  $\mathbb{R}$  of  $\mathbb{R}$ -valued random variable  $H_i$  is written as

$$\hat{p}_{H_i}(\eta_i) = \frac{1}{\nu_s} \sum_{\ell=1}^{\nu_s} \mu_{1, \hat{\sigma}_n^{\text{MC}}} \left( \frac{\hat{\sigma}_n^{\text{MC}}}{\sigma_n^{\text{MC}}} U_i^{\widehat{M}_\ell}(\boldsymbol{\theta}) - \eta_i \right) \quad (2.47)$$

For  $i$  in  $\{1, \dots, n\}$ , for  $\nu_s = 400$  and  $\nu_s = 10,000$ , we obtain a similar result that the one displayed in Fig. 2.4, the convergence being reached for  $\nu_s = 10,000$ . For  $i$  and  $j$  fixed in  $\{1, \dots, n\}$ , the estimation  $(\eta_i, \eta_j) \mapsto p_{H_i, H_j}(\eta_i, \eta_j)$  of the joint probability density function on  $\mathbb{R}^2$  of the  $\mathbb{R}^2$ -valued



**Figure 2.5.** For  $(i, j) = (1, 3), (3, 4)$  and  $(4, 5)$ , graphs of the joint probability density functions  $(\eta_i, \eta_j) \mapsto p_{H_i, H_j}(\eta_i, \eta_j)$  (left figure), and  $(\eta_i, \eta_j) \mapsto \hat{p}_{H_i, H_j}(\eta_i, \eta_j)$  for  $m_{\text{ergo}} = 10,000$  (right figure).

random variable  $(H_i, H_j)$  is given by

$$\hat{p}_{H_i, H_j}(\eta_i, \eta_j) = \frac{1}{\nu_s} \sum_{\ell=1}^{\nu_s} \mu_{2, \hat{\sigma}_n^{\text{MC}}} \left( \frac{\hat{\sigma}_n^{\text{MC}}}{\sigma_n^{\text{MC}}} \mathbf{U}^{\widehat{M}_\ell, ij}(\theta) - \boldsymbol{\eta}^{ij} \right), \quad (2.48)$$

in which  $\mathbf{U}^{\widehat{M}_\ell, ij}(\theta) = (U_i^{\widehat{M}_\ell}(\theta), U_j^{\widehat{M}_\ell}(\theta))$  and  $\boldsymbol{\eta}^{ij} = (\eta_i, \eta_j)$  belong to  $\mathbb{R}^2$ . For  $\nu_s = 10,000$ , a similar result to the one shown in Fig. 2.5 is obtained.

**3. Polynomial chaos expansion of a multimodal random vector.** Let  $\mathbf{h}$  be a given measurable mapping from  $\mathbb{R}^N$  into  $\mathbb{R}^M$ , and let  $\mathbf{Q} = (Q_1, \dots, Q_M)$  be the  $\mathbb{R}^M$ -valued random variable such that

$$\mathbf{Q} = \mathbf{h}(\mathbf{X}^{(n)}), \quad (3.1)$$

in which  $\mathbf{X}^{(n)}$  is the multimodal random vector defined by Eq. (2.3). Transformation  $\mathbf{h}$  is assumed to be such that  $\mathbf{Q}$  is a second-order random vector. Possibly, the deterministic mapping  $\mathbf{h}$  transfers

the multimodal character of random vector  $\mathbf{X}^{(n)}$  to random vector  $\mathbf{Q} = \mathbf{h}(\mathbf{X}^{(n)})$ . We are interested in constructing the polynomial chaos expansion (PCE) of random vector  $\mathbf{Q}$ . If such a PCE of  $\mathbf{Q}$  was carried out with respect to the polynomial chaos associated with a unimodal random variable, which is usually the choice done (uniform, Gaussian, etc.), then the speed of convergence could be low. Since the multimodal probability distribution of  $\mathbf{Q}$  is induced by the multimodal probability distribution of  $\mathbf{H}$ , we propose an alternative approach consisting in using the multimodal random vector  $\mathbf{H}$  (for which the probability density function  $p_{\mathbf{H}}$  is defined by Eqs. (2.7) to (2.9)) as the stochastic germ for the polynomial chaos expansion of random vector  $\mathbf{Q}$ .

### 3.1. Constructing independent realizations of $\mathbf{Q}$ and reduced-order statistical model.

Although the contents of this section is very classic and well known, a brief presentation is given in order to define the notations and the mapping  $\mathfrak{h}$ . Let  $\mathbf{q}^{\text{sim},1}, \dots, \mathbf{q}^{\text{sim},\nu_s}$  be  $\nu_s$  independent realizations of random vector  $\mathbf{Q}$ , which are computed, for all  $\ell$  in  $\{1, \dots, \nu_s\}$ , by

$$\mathbf{q}^{\text{sim},\ell} = \mathbf{h}(\widehat{\mathbf{m}}_X + \sum_{i=1}^n \sqrt{\lambda_{X,i}} \eta_i^{\text{sim},\ell} \phi_X^i), \quad (3.2)$$

$$\boldsymbol{\eta}^{\text{sim},\ell} = \mathbf{U}^{\widehat{M}_\ell}(\theta) \quad , \quad \widehat{M}_\ell = (1 + \ell) \widehat{M}_0, \quad (3.3)$$

in which  $\eta^{\text{sim},1}, \dots, \eta^{\text{sim},\nu_s}$  are the  $\nu_s$  independent realizations of the second-order  $\mathbb{R}^n$ -valued random vector  $\mathbf{H}$  whose probability density function is defined by Eqs. (2.7) to (2.9), and for which the generator of independent realizations is detailed in Section 2.4.4 (see Eq. (2.31)). Similarly to Section 2.1, let  $\widehat{\mathbf{m}}_{\mathbf{Q}}$  and  $[\widehat{C}_{\mathbf{Q}}]$  be the empirical estimations of the mean vector  $\mathbf{m}_{\mathbf{Q}} = E\{\mathbf{Q}\}$  and the covariance matrix  $[C_{\mathbf{Q}}] = E\{(\mathbf{Q} - \mathbf{m}_{\mathbf{Q}})(\mathbf{Q} - \mathbf{m}_{\mathbf{Q}})^T\}$ , such that

$$\widehat{\mathbf{m}}_{\mathbf{Q}} = \frac{1}{\nu_s} \sum_{\ell=1}^{\nu_s} \mathbf{q}^{\text{sim},\ell} \quad , \quad [\widehat{C}_{\mathbf{Q}}] = \frac{1}{\nu_s - 1} \sum_{\ell=1}^{\nu_s} (\mathbf{q}^{\text{sim},\ell} - \widehat{\mathbf{m}}_{\mathbf{Q}})(\mathbf{q}^{\text{sim},\ell} - \widehat{\mathbf{m}}_{\mathbf{Q}})^T. \quad (3.4)$$

Let  $\lambda_{Q,1} \geq \lambda_{Q,2} \geq \dots \geq \lambda_{Q,M} \geq 0$  be the eigenvalues, and let  $\phi_Q^1, \dots, \phi_Q^M$ , be the associated orthonormal eigenvectors ( $(\phi_Q^i)^T \phi_Q^j = \delta_{ij}$ ) of the eigenvalue problem  $[\widehat{C}_{\mathbf{Q}}] \phi_Q^j = \lambda_{Q,j} \phi_Q^j$ . The reduced-order statistical model  $\mathbf{Q}^{(m)}$  of  $\mathbf{Q}$  is then written as

$$\mathbf{Q}^{(m)} = \widehat{\mathbf{m}}_{\mathbf{Q}} + \sum_{j=1}^m \sqrt{\lambda_{Q,j}} \Xi_j \phi_Q^j. \quad (3.5)$$

The random vector  $\mathbf{Q}$  is such that  $\mathbf{Q} = \mathbf{Q}^{(M)}$ , and the value of  $m$  is fixed in  $\{1, \dots, M\}$  for that  $\text{err}_{\mathbf{Q}}(m) \leq \varepsilon$  in which  $\varepsilon$  is any positive real number, and where  $\text{err}_{\mathbf{Q}}$  is the error function defined by

$$\text{err}_{\mathbf{Q}}(m) = 1 - \frac{\sum_{j=1}^m \lambda_{Q,j}}{\text{tr}[\widehat{C}_{\mathbf{Q}}]}. \quad (3.6)$$

Let  $\boldsymbol{\xi}^{\text{sim},1}, \dots, \boldsymbol{\xi}^{\text{sim},\nu_s}$  be the  $\nu_s$  independent realizations of the second-order random vector  $\boldsymbol{\Xi} = (\Xi_1, \dots, \Xi_m)$ , computed, for  $\ell = 1, \dots, \nu_s$  and  $j = 1, \dots, m$ , by

$$\xi_j^{\text{sim},\ell} = \frac{1}{\sqrt{\lambda_{Q,j}}} (\mathbf{q}^{\text{sim},\ell} - \widehat{\mathbf{m}}_{\mathbf{Q}})^T \phi_Q^j. \quad (3.7)$$

The estimation  $\widehat{\mathbf{m}}_{\Xi}$  of  $E\{\Xi\}$  and the estimation  $[\widehat{R}_{\Xi}]$  of  $E\{\Xi\Xi^T\}$  are written as

$$\widehat{\mathbf{m}}_{\Xi} = \frac{1}{\nu_s} \sum_{\ell=1}^{\nu_s} \xi^{\text{sim},\ell}, \quad [\widehat{R}_{\Xi}] = \frac{1}{\nu_s - 1} \sum_{\ell=1}^{\nu_s} \xi^{\text{sim},\ell} (\xi^{\text{sim},\ell})^T. \quad (3.8)$$

We then have,

$$\widehat{\mathbf{m}}_{\Xi} = \mathbf{0}, \quad [\widehat{R}_{\Xi}] = [I_m]. \quad (3.9)$$

Finally, Eqs. (3.2) and (3.7) define a mapping  $\mathfrak{h} = (\mathfrak{h}_1, \dots, \mathfrak{h}_m)$  from  $\mathbb{R}^n$  into  $\mathbb{R}^m$  such that  $\Xi = \mathfrak{h}(\mathbf{H})$ , and for all  $\ell = 1, \dots, \nu_s$  and for all  $j = 1, \dots, m$ , we have  $\xi_j^{\text{sim},\ell} = \mathfrak{h}_j(\boldsymbol{\eta}^{\text{sim},\ell})$  that can globally be rewritten as

$$\xi^{\text{sim},\ell} = \mathfrak{h}(\boldsymbol{\eta}^{\text{sim},\ell}), \quad \ell = 1, \dots, \nu_s. \quad (3.10)$$

**3.2. Polynomial chaos expansion with respect to the arbitrary measure defined by the multimodal probability measure of  $\mathbf{H}$ .** In order to define the algebraic notations for the polynomial chaos expansions in a finite approximation, the following parameters and multi-indexes are introduced. Let  $\mathbb{N} = \{0, 1, 2, \dots\}$  be the set of all the null and positive integers. For  $N_d \geq 1$  ( $N_d$  will be the maximum degree of the polynomials), the following set  $\mathcal{A}_{N_d}$  of multi-indexes are introduced,

$$\mathcal{A}_{N_d} = \{\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n \mid 0 \leq \alpha_1 + \dots + \alpha_n \leq N_d\}. \quad (3.11)$$

The  $1 + K$  elements of set  $\mathcal{A}_{N_d}$  are denoted by  $\boldsymbol{\alpha}^{(0)}, \dots, \boldsymbol{\alpha}^{(K)}$  in which  $\boldsymbol{\alpha}^{(0)}$  is the multi-index  $(0, \dots, 0)$ , and where the integer  $K$  is such that

$$K = \frac{(n + N_d)!}{n! N_d!} - 1. \quad (3.12)$$

It is assumed that  $K < \nu_s$ . Let  $\{\Psi_{\boldsymbol{\alpha}}(\boldsymbol{\eta}), \boldsymbol{\alpha} \in (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n\}$  be the family of multivariate orthonormal polynomials with respect to the multimodal probability measure  $p_{\mathbf{H}}(\boldsymbol{\eta}) d\boldsymbol{\eta}$  on  $\mathbb{R}^n$  defined by Eqs. (2.7) to (2.9), such that, for all  $\boldsymbol{\alpha}$  and  $\boldsymbol{\beta}$  in  $\mathbb{N}^n$ ,

$$\int_{\mathbb{R}^n} \Psi_{\boldsymbol{\alpha}}(\boldsymbol{\eta}) \Psi_{\boldsymbol{\beta}}(\boldsymbol{\eta}) p_{\mathbf{H}}(\boldsymbol{\eta}) d\boldsymbol{\eta} = E\{\Psi_{\boldsymbol{\alpha}}(\mathbf{H}) \Psi_{\boldsymbol{\beta}}(\mathbf{H})\} = \delta_{\boldsymbol{\alpha}\boldsymbol{\beta}}. \quad (3.13)$$

In Eq. (3.13),  $\delta_{\boldsymbol{\alpha}\boldsymbol{\beta}}$  is the kronecker symbol and by convention, for  $\boldsymbol{\alpha} = \mathbf{0}$ ,  $\Psi_{\mathbf{0}}(\boldsymbol{\eta}) = 1$  is the constant normalized polynomial. It should be noted that the probability measure  $p_{\mathbf{H}}(\boldsymbol{\eta}) d\boldsymbol{\eta}$  is non separable (the components of random vector  $\mathbf{H}$  are not statistically independent). Consequently, the integral on  $\mathbb{R}^n$  defined by Eq. (3.13) cannot be written as a product of  $n$  integrals over  $\mathbb{R}$ , and each multivariate orthonormal polynomial  $\Psi_{\boldsymbol{\alpha}}(\boldsymbol{\eta})$  in  $\boldsymbol{\eta} = (\eta_1, \dots, \eta_n)$  cannot be written as  $\Psi_{\boldsymbol{\alpha}}(\boldsymbol{\eta}) = \psi_{\alpha_1}(\eta_1) \times \dots \times \psi_{\alpha_n}(\eta_n)$ , but this type of tensorial product of univariate polynomials for the multivariate polynomials (tensorial product that does not hold in the present case) is not used. From a theoretical point of view, the multivariate orthonormal polynomials can be viewed as the result of a Gram-Schmidt orthonormalization algorithm of the multivariate monomials defined, for  $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n)$  belonging to  $\mathcal{A}_{N_d}$  (defined by Eq. (3.11)), and for  $\boldsymbol{\eta} = (\eta_1, \dots, \eta_n)$  in  $\mathbb{R}^n$ , by

$$\mathcal{M}_{\boldsymbol{\alpha}}(\boldsymbol{\eta}) = \eta_1^{\alpha_1} \times \dots \times \eta_n^{\alpha_n}. \quad (3.14)$$

It should be noted that an explicit algebraic formula (constructed, for instance, with a symbolic toolbox) for each multivariate orthonormal polynomial  $\Psi_{\alpha}(\boldsymbol{\eta})$ , cannot be obtained due to the non separable expression of  $p_{\mathbf{H}}(\boldsymbol{\eta})$  defined by Eqs. (2.7) to (2.9). Nevertheless, even if such an explicit expression could be obtained, this expression would not be used for the reasons detailed in [41] and summarized in Section 3.4, and is effectively not used in the methodology proposed for constructing the polynomial chaos expansion of the random vector  $\Xi$  (see Sections 3.3 and 3.4). Finally, it can be seen that the degree of multivariate polynomial  $\Psi_{\alpha}(\boldsymbol{\eta})$  is  $\alpha_1 + \dots + \alpha_n$ , and that for all  $\alpha$  in  $\mathcal{A}_{N_d}$ , the maximum degree of the polynomials is  $N_d$ .

The second-order random variable  $\Xi$  can then be expanded in polynomial chaos  $\Psi_{\alpha}$  as,

$$\Xi = \lim_{N_d \rightarrow +\infty} \Xi^{(N_d)}, \quad (3.15)$$

with convergence in the space  $L^2(\Theta, \mathbb{R}^m)$  of all the  $\mathbb{R}^m$ -valued second-order random variables defined on  $(\Theta, \mathcal{T}, \mathcal{P})$ , and where

$$\Xi^{(N_d)} = \sum_{k=0}^K \mathbf{y}^k \Psi_{\alpha^{(k)}}(\mathbf{H}) \quad , \quad \mathbf{y}^0, \dots, \mathbf{y}^K \in \mathbb{R}^m. \quad (3.16)$$

Taking into account Eq. (3.9), we choose to identify the coefficients  $\mathbf{y}^0, \dots, \mathbf{y}^K$  in  $\mathbb{R}^m$ , such that for each fixed value of  $N_d$ ,

$$E\{\Xi^{(N_d)}\} = \mathbf{0} \quad , \quad E\{\Xi^{(N_d)} (\Xi^{(N_d)})^T\} = [I_m]. \quad (3.17)$$

From Eqs. (3.13), (3.16), and (3.17), it can be deduced that coefficients  $\mathbf{y}^0, \dots, \mathbf{y}^K$  must satisfy the constraints  $\mathbf{y}^0 = \mathbf{0}$  and  $\sum_{k=1}^K \mathbf{y}^k (\mathbf{y}^k)^T = [I_m]$ . Introducing the rectangular matrix

$$[y] = [\mathbf{y}^1 \dots \mathbf{y}^K] \in \mathbb{M}_{m,K}(\mathbb{R}), \quad (3.18)$$

and the  $\mathbb{R}^K$ -valued random variable  $\Psi(\mathbf{H}) = (\Psi_{\alpha^{(1)}}(\mathbf{H}), \dots, \Psi_{\alpha^{(K)}}(\mathbf{H}))$ , then Eq. (3.16) can be rewritten as

$$\Xi^{(N_d)} = [y] \Psi(\mathbf{H}), \quad (3.19)$$

in which the coefficients  $[y]$  verifies  $[y] [y]^T = [I_m]$ . Finally, due to the orthonormality of the polynomial chaos (see the second equation in Eq. (3.17)), it can easily be deduced the classical equation,

$$[y] = E\{\Xi \Psi(\mathbf{H})^T\} = E\{\mathfrak{h}(\mathbf{H}) \Psi(\mathbf{H})^T\}, \quad (3.20)$$

in which  $\Xi = \mathfrak{h}(\mathbf{H})$  is the  $\mathbb{R}^m$ -random variable whose  $\nu_s$  independent realizations  $\boldsymbol{\xi}^{\text{sim},1}, \dots, \boldsymbol{\xi}^{\text{sim},\nu_s}$  have been computed with Eq. (3.7) (see also Eq. (3.10)).

**3.3. Computing the coefficients of the PCE.** For  $N_d$  fixed, we propose to compute the coefficients  $[y]$  of the PCE in using the Monte Carlo method. The use of the ergodic method would lead us to numerically evaluate function  $\mathfrak{h}$  for the  $M$  values  $\boldsymbol{\eta}^k = \mathbf{U}^k(\theta)$  for  $k = M_0, \dots, M$  of  $\boldsymbol{\eta}$ , while the Monte Carlo method only requires  $\nu_s$  evaluations, and we have seen that  $M \gg \nu_s$ . In addition, the  $\nu_s$  evaluations of function  $\mathfrak{h}$  have already be carried out for constructing the statistical reduced-order model of random vector  $\mathbf{Q}$  (see Eqs. (3.2), (3.3) and (3.7)). Therefore, there is a weak

interest in developing an approach based on the ergodic method for computing the coefficients of the PCE of random vector  $\Xi$ . We then use the  $\nu_s$  independent realizations  $\boldsymbol{\eta}^{\text{sim},1}, \dots, \boldsymbol{\eta}^{\text{sim},\nu_s}$  of random vector  $\mathbf{H}$ , defined by Eq. (3.3), and the  $\nu_s$  independent realizations  $\boldsymbol{\xi}^{\text{sim},1}, \dots, \boldsymbol{\xi}^{\text{sim},\nu_s}$  of random vector  $\Xi$  computed in using Eq. (3.7) (i.e. Eq. (3.10)). Taking into account Eqs. (3.8) and (3.9), the estimation  $[y^{N_d}]$  of  $[y]$  given by Eq. (3.20) is written as

$$[y^{N_d}] = \frac{1}{\nu_s - 1} [\boldsymbol{\xi}^{\text{sim}}] [\Psi]^T. \quad (3.21)$$

in which the  $(m \times \nu_s)$  real matrix  $[\boldsymbol{\xi}^{\text{sim}}]$  and the  $(K \times \nu_s)$  real matrix  $[\Psi]$  are defined, for all  $j = 1, \dots, m, k = 1, \dots, K$  and  $\ell = 1, \dots, \nu_s$ , as

$$[\boldsymbol{\xi}^{\text{sim}}]_{j\ell} = \xi_j^{\text{sim},\ell}, \quad [\Psi]_{k\ell} = \Psi_{\alpha^{(k)}}(\boldsymbol{\eta}^{\text{sim},\ell}). \quad (3.22)$$

From Eq. (3.19), random vector  $\Xi^{(N_d)}$ , for which  $[y]$  is estimated by the Monte Carlo method, is written as

$$\Xi^{(N_d)} = [y^{N_d}] \Psi(\mathbf{H}). \quad (3.23)$$

The  $\nu_s$  independent realizations  $\boldsymbol{\xi}^{(N_d),1}, \dots, \boldsymbol{\xi}^{(N_d),\nu_s}$  of random vector  $\Xi^{(N_d)}$  are thus given, for all  $\ell = 1, \dots, \nu_s$ , by  $\boldsymbol{\xi}^{(N_d),\ell} = [y^{N_d}] \Psi(\boldsymbol{\eta}^{\text{sim},\ell})$ . Let  $[\boldsymbol{\xi}^{N_d}]$  be the  $(m \times \nu_s)$  real matrix defined, for all  $j = 1, \dots, m$  and  $\ell = 1, \dots, \nu_s$ , by

$$[\boldsymbol{\xi}^{N_d}]_{j\ell} = \xi_j^{(N_d),\ell}. \quad (3.24)$$

Consequently, matrix  $[\boldsymbol{\xi}^{N_d}]$  can be computed by the following equation,

$$[\boldsymbol{\xi}^{N_d}] = [y^{N_d}] [\Psi]. \quad (3.25)$$

**3.4. Computing the realizations of the polynomial chaos.** The objective of this section is to compute the  $(K \times \nu_s)$  real matrix  $[\Psi]$  defined in Eq. (3.22) and constituted of the  $\nu_s$  independent realizations  $\Psi_{\alpha^{(k)}}(\boldsymbol{\eta}^{\text{sim},\ell})$  of the polynomial chaos  $\Psi_{\alpha^{(k)}}(\mathbf{H})$  for the arbitrary multimodal probability measure  $p_{\mathbf{H}}(\boldsymbol{\eta}) d\boldsymbol{\eta}$  defined by Eqs. (2.7) to (2.9). It is assumed that  $K < \nu_s$ . Such a calculation must preserve the orthogonality property of the PCE in high dimension which is written as

$$\lim_{\nu_s \rightarrow +\infty} \frac{1}{\nu_s - 1} [\Psi] [\Psi]^T = [I_K]. \quad (3.26)$$

As it has been proved in [41], this problem is not trivial at all. For instance, for classical measures for which the polynomial chaos are explicitly known (Hermite, Legendre, and so on), if the explicit algebraic formulas (constructed with a symbolic Toolbox) or the computational recurrence relations with respect to the degree are used, then important numerical noise is induced, and the orthogonality property is lost. If a global orthogonalization was done to correct this problem, then the independence of the realizations would be lost. In addition, in the present case, we have an arbitrary multimodal measure, and the explicit expression of the polynomial chaos are unknown. We then propose to use the new algorithm proposed in [35], that we summarize below. It consists (i) in constructing the  $\nu_s$  independent realizations of the multimodal multivariate monomials using the generator of the multimodal random vector  $\mathbf{H}$  presented in Section 2.4.4; (ii) in performing an orthogonalization of the realizations of the multimodal multivariate monomials with an algorithm different from the Gram-Schmidt orthogonalization algorithm as the latter is not stable in high dimension. For  $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n)$  belonging

to  $\mathcal{A}_{N_d}$  (defined by Eq. (3.11)), and for  $\boldsymbol{\eta} = (\eta_1, \dots, \eta_n)$  in  $\mathbb{R}^n$ , let  $\mathcal{M}_{\boldsymbol{\alpha}}(\boldsymbol{\eta}) = \eta_1^{\alpha_1} \times \dots \times \eta_n^{\alpha_n}$  be the multivariate monomials. Let  $K_0 = 1 + K$ . Let  $[\mathbb{M}]$  be the  $(K_0 \times \nu_s)$  real matrix of independent realizations of the monomials such that

$$[\mathbb{M}]_{k\ell} = \mathcal{M}_{\boldsymbol{\alpha}^{(k)}}(\boldsymbol{\eta}^{\text{sim},\ell}) \quad , \quad k = 0, 1, \dots, K, \quad (3.27)$$

in which  $\boldsymbol{\alpha}^{(0)} = (0, \dots, 0)$ , and where  $\boldsymbol{\eta}^{\text{sim},1}, \dots, \boldsymbol{\eta}^{\text{sim},\nu_s}$  are the  $\nu_s$  independent realizations of multimodal random vector  $\mathbf{H}$ , computed with Eq. (3.3). Let  $[\Psi_0]$  be the  $(K_0 \times \nu_s)$  real matrix such that  $[\Psi_0]_{k\ell} = \Psi_{\boldsymbol{\alpha}^{(k)}}(\boldsymbol{\eta}^{\text{sim},\ell})$  for  $k = 0, 1, \dots, K$  and  $\ell = 1, \dots, \nu_s$ . Matrix  $[\Psi_0]$  can be written as  $[\Psi_0] = [\mathbb{A}] [\mathbb{M}]$  in which  $[\mathbb{A}]$  is a  $(K_0 \times K_0)$  real matrix, and  $[\mathbb{M}]$  is a  $(K_0 \times \nu_s)$  real matrix. Let  $[\mathbb{F}]$  be the positive-definite symmetric  $(K_0 \times K_0)$  real matrix defined by  $[\mathbb{F}] = \frac{1}{\nu_s - 1} [\mathbb{M}] [\mathbb{M}]^T = [\mathbb{A}]^{-1} [\mathbb{A}]^{-T}$ . The algorithm is then the following:

- Compute matrix  $[\mathbb{M}]$  defined by Eq. (3.27), and then compute the  $(K_0 \times K_0)$  real matrix  $[\mathbb{F}] = \frac{1}{\nu_s - 1} [\mathbb{M}] [\mathbb{M}]^T$ .
- Compute the lower triangular  $(K_0 \times K_0)$  real matrix  $[\mathbb{L}]$  from the Cholesky decomposition  $[\mathbb{L}] [\mathbb{L}]^T$  of positive-definite symmetric matrix  $[\mathbb{F}]$ .
- Compute the lower triangular  $(K_0 \times K_0)$  real matrix  $[\mathbb{A}] = [\mathbb{L}]^{-1}$ .
- Compute the  $(K_0 \times \nu_s)$  real matrix  $[\Psi_0] = [\mathbb{A}] [\mathbb{M}]$ .
- Deduce the  $(K \times \nu_s)$  real matrix  $[\Psi]$  obtained in deleting the first row of matrix  $[\Psi_0]$ .

**3.5. Convergence analysis with respect to  $N_d$ .** The numerical convergence with respect to parameter  $N_d$  is classically analyzed. For  $n$  fixed, let  $K = K(N_d)$  be defined by Eq. (3.12). Taking into account the convergence in  $L^2(\Theta, \mathbb{R}^m)$  of the sequence  $\{\boldsymbol{\Xi}^{(N_d)}\}_{N_d}$  of random vectors (defined by Eq. (3.19)) towards random vector  $\boldsymbol{\Xi}$ , the  $L^2$ -convergence of the polynomial chaos expansion can be performed in analyzing the graph

$$N_d \mapsto \|\boldsymbol{\Xi} - \boldsymbol{\Xi}^{(N_d)}\|_{L^2} = (E\{\|\boldsymbol{\Xi} - \boldsymbol{\Xi}^{(N_d)}\|^2\})^{1/2}, \quad (3.28)$$

in which  $\|\boldsymbol{\xi}\|^2 = \xi_1^2 + \dots + \xi_n^2$ . Introducing the estimation of the second-order moment of real valued random variable  $\|\boldsymbol{\Xi} - \boldsymbol{\Xi}^{(N_d)}\|$  in using the  $\nu_s$  independent realizations  $\boldsymbol{\eta}^{\text{sim},1}, \dots, \boldsymbol{\eta}^{\text{sim},\nu_s}$  of  $\mathbf{H}$ , yields  $\|\boldsymbol{\Xi} - \boldsymbol{\Xi}^{(N_d)}\|_{L^2}^2 \simeq \frac{1}{\nu_s - 1} \|[\boldsymbol{\xi}^{\text{sim}}] - [\boldsymbol{\xi}^{N_d}]\|_F^2$  in which  $\|\cdot\|_F$  is the Frobenius matrix norm, and where matrices  $[\boldsymbol{\xi}^{\text{sim}}]$  and  $[\boldsymbol{\xi}^{N_d}]$  are defined by Eqs. (3.22) and (3.25), respectively. For the numerical convergence analysis, we therefore introduce the relative error function  $N_d \mapsto \text{err}_{\boldsymbol{\Xi}}(N_d)$  such that

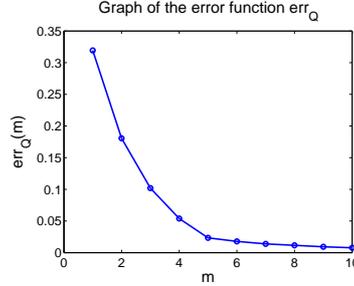
$$\text{err}_{\boldsymbol{\Xi}}(N_d) = \frac{\|[\boldsymbol{\xi}^{\text{sim}}] - [\boldsymbol{\xi}^{N_d}]\|_F}{\|[\boldsymbol{\xi}^{\text{sim}}]\|_F}. \quad (3.29)$$

### 3.6. Numerical application.

#### 3.6.1. Definition of the data transformation.

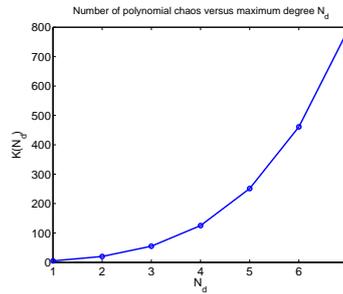
We reuse the numerical application presented in Section 2.6 for which the values of the parameters are  $N = 100$ ,  $\nu = 500$ ,  $n = 5$ ,  $f_0 = 5.0565$ ,  $\varepsilon_0 = 1/200$ ,  $m_{\text{overs}} = 10$  that yields  $M_0 = 260$ ,  $\Delta r = 0.01638$ ,  $r_0 = 4.26$ ,  $i_0 = 1$ . The transformation  $\mathbf{h}$  from  $\mathbb{R}^N$  in  $\mathbb{R}^M$ , introduced in Section 3 (see Eq. (3.1)), is described in Appendix B. The reader can then simulate the "data transformation" used in this numerical application. Taking into account the convergence analysis presented in Section 2.6.3-(i), and since the number of evaluation of  $\mathbf{h}$  must not be too large for the practical applications, we choose  $\nu_s = 1,000$ .

**3.6.2. Statistical reduced-order model of  $\mathbf{Q}$ .** The estimations  $\widehat{\mathbf{m}}_{\mathbf{Q}}$  and  $[\widehat{C}_{\mathbf{Q}}]$  are computed using Eq. (3.4). Concerning the construction of the reduced-order statistical model of  $\mathbf{Q}$  defined by Eq. (3.5), Fig. 3.1 displays the graph of the error function  $m \mapsto \text{err}_{\mathbf{Q}}(m)$  defined by Eq. (3.6). We choose  $m = 4$  corresponding to an error of 0.054.



**Figure 3.1.** Reduced-order statistical model: graph of the error function  $m \mapsto \text{err}_{\mathbf{Q}}(m)$

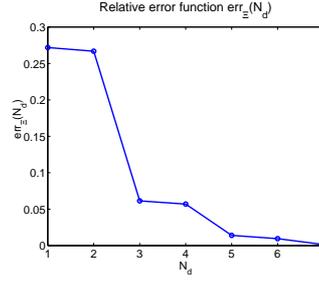
**3.6.3. Computing the coefficient and convergence analysis.** The coefficients  $[y^{N_d}]$  are computed using Eq. (3.21) for  $N_d = 1, \dots, 7$  with  $\nu_s = 1,000$ . Fig. 3.2 displays the graph  $N_d \mapsto K(N_d)$  that gives the number  $K$  of polynomial chaos calculated with Eq. (3.12) as a function of the maximum degree  $N_d$  of the polynomial chaos. The result of the convergence analysis has been



**Figure 3.2.** Graph  $N_d \mapsto K(N_d)$  of the number  $K$  of polynomial chaos versus their maximum degree  $N_d$ .

studied as a function of the maximum degree  $N_d$  of the polynomial chaos. Fig. 3.3 displays the graph  $N_d \mapsto \text{err}_{\Xi}(N_d)$  calculated with Eq. (3.29). For  $N_d = 2$ , the error is 0.27 while for  $N_d = 3$  and 4, the error is 0.061 and 0.057, and for  $N_d = 7$ , the error is 0.0014.

**3.6.4. Probability density functions relative to  $\Xi$  and  $\Xi^{(N_d)}$ .** In this section, we compare the graphs of the probability density functions relative to random vector  $\Xi$  (for which the  $\nu_s$  independent realizations are  $\xi^{\text{sim},1}, \dots, \xi^{\text{sim},\nu_s}$ ), with the graphs of the probability density functions relative to random vector  $\Xi^{(N_d)}$  (for which the  $\nu_s$  independent realizations are  $\xi^{(N_d),1}, \dots, \xi^{(N_d),\nu_s}$ ). The computations are performed for  $\nu_s = 1,000$ ,  $N_d = 3$  that corresponds to  $K = 55$  with an error  $\text{err}_{\Xi}(N_d) = 0.061$ . For  $j$  in  $\{1, \dots, m\}$ , Fig. 3.4 displays the graphs of the probability density



**Figure 3.3.** Graph  $N_d \mapsto err_{\Xi}(N_d)$  of the relative error of the polynomial-chaos coefficients versus the maximum degree  $N_d$  of the polynomial chaos.

functions  $p_{\Xi_j}$  and  $p_{\Xi_j^{(N_d)}}$  calculated by

$$p_{\Xi_j}(\xi_j) = \frac{1}{\nu_s} \sum_{\ell=1}^{\nu_s} \mu_{1, \hat{\sigma}_m^{MC}} \left( \frac{\hat{\sigma}_m^{MC}}{\sigma_m^{MC}} \xi_j^{\text{sim}, \ell} - \xi_j \right), \quad (3.30)$$

$$p_{\Xi_j^{(N_d)}}(\xi_j) = \frac{1}{\nu_s} \sum_{\ell=1}^{\nu_s} \mu_{1, \hat{\sigma}_m^{MC}} \left( \frac{\hat{\sigma}_m^{MC}}{\sigma_m^{MC}} \xi_j^{(N_d), \ell} - \xi_j \right), \quad (3.31)$$

in which  $\sigma_m^{MC}$  and  $\hat{\sigma}_m^{MC}$  are defined by Eqs. (2.46), and where  $\mu_{1, \hat{\sigma}_m^{MC}}$  is defined by Eq. (2.40). Fig. 3.4 shows that the densities are well captured. For  $i$  and  $j$  fixed in  $\{1, \dots, n\}$ , Fig. 3.5 displays the graphs of the joint probability density functions  $p_{\Xi_i \Xi_j}$  and  $p_{\Xi_i^{(N_d)} \Xi_j^{(N_d)}}$  calculated by

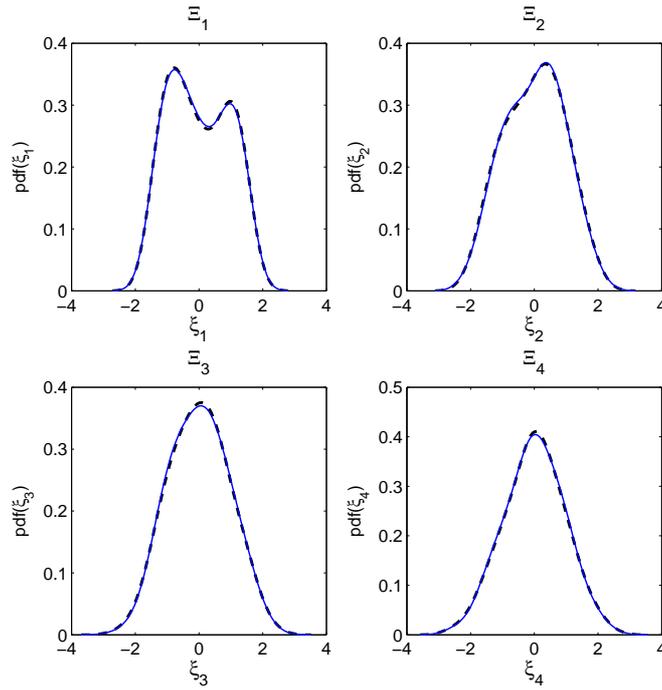
$$p_{\Xi_i \Xi_j}(\xi_i, \xi_j) = \frac{1}{\nu_s} \sum_{\ell=1}^{\nu_s} \mu_{2, \hat{\sigma}_m^{MC}} \left( \frac{\hat{\sigma}_m^{MC}}{\sigma_m^{MC}} \boldsymbol{\xi}^{\text{sim}, \ell, ij} - \boldsymbol{\xi}^{ij} \right), \quad (3.32)$$

$$p_{\Xi_i^{(N_d)} \Xi_j^{(N_d)}}(\xi_i, \xi_j) = \frac{1}{\nu_s} \sum_{\ell=1}^{\nu_s} \mu_{2, \hat{\sigma}_m^{MC}} \left( \frac{\hat{\sigma}_m^{MC}}{\sigma_m^{MC}} \boldsymbol{\xi}^{(N_d), \ell, ij} - \boldsymbol{\xi}^{ij} \right), \quad (3.33)$$

in which  $\boldsymbol{\xi}^{\text{sim}, \ell, ij} = (\xi_i^{\text{sim}, \ell}, \xi_j^{\text{sim}, \ell})$ ,  $\boldsymbol{\xi}^{(N_d), \ell, ij} = (\xi_i^{(N_d), \ell}, \xi_j^{(N_d), \ell})$  and  $\boldsymbol{\xi}^{ij} = (\xi_i, \xi_j)$  belong to  $\mathbb{R}^2$ . Fig. 3.5 shows that there is good agreement between the joint densities.

**4. Conclusions.** A methodology and algorithms have been proposed for constructing the polynomial chaos expansion (PCE) of any  $\mathbb{R}^m$ -valued random variable with respect to an arbitrary multimodal probability measure on  $\mathbb{R}^n$ . The algorithms proposed are efficient and robust, and can be used in high dimension. This new class of algorithms, which is adapted to the multimodal case, can be useful in the context of uncertainty quantification for direct and inverse problems.

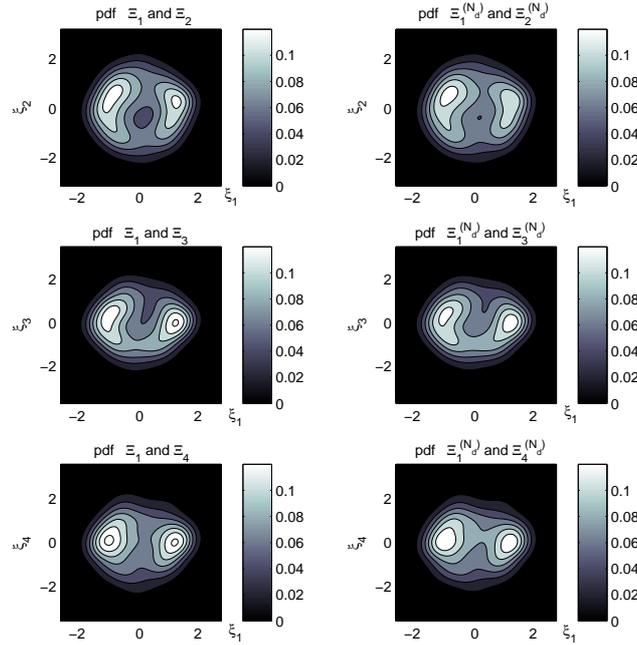
**Appendix A. Data generation for the numerical application.** In this Appendix, the algorithm used for generating the  $\nu$  independent realizations  $\mathbf{x}^{\text{exp}, 1}, \dots, \mathbf{x}^{\text{exp}, \nu}$  of the random vector  $\mathbf{X}$  with values in  $\mathbb{R}^N$ , is described. We consider,  $N = 100$ ,  $\nu = 500$ ,  $n = 5$ ,  $n_r = 15$ ,  $c_{\text{sym}} = 0.2$ , and  $c_{\text{sep}} = 2$ . We introduce a deterministic  $(n_r \times n_r)$  real matrix  $[a]$  which is generated by  $\text{randn}('state', 0)$ , and then,  $[a] = [I_{n_r}] + [\text{randn}(n_r, n_r)]/2$ , in which we have used the Matlab function, "randn". Then



**Figure 3.4.** For  $j$  in  $\{1, \dots, m\}$ , graph of the probability density functions  $\xi_j \mapsto p_{\Xi_j}(\xi_j)$  (dashed line) and  $\xi_j \mapsto p_{\Xi_j^{(N_d)}}(\xi_j)$  (solid line).

$\nu$  independent realizations  $(Y_{1\ell}, \dots, Y_{n_r\ell})$  (for  $\ell = 1, \dots, \nu$ ) of a  $\mathbb{R}^{n_r}$ -valued random vector  $\mathbf{Y} = (Y_1, \dots, Y_{n_r})$  are generated such that,  $\text{randn}('state', 0)$ , and then,  $Y_{i\ell} = \sum_{j=1}^{n_r} [a]_{ij} [\text{randn}(n_r, \nu)]_{j\ell}$ . The  $\nu$  independent realizations  $\boldsymbol{\eta}^{\text{exp},1}, \dots, \boldsymbol{\eta}^{\text{exp},\nu}$  of the  $\mathbb{R}^{n_r}$ -valued random vector  $\boldsymbol{\eta}$  are generated as follows. For  $i = 1, \dots, n_r$  and for  $\ell = 1, \dots, \nu$ ,  $\eta_i^{\text{exp},\ell} = Z_{i\ell} + c_{\text{sep}} \sigma_i (2 \mathbb{1}_{[0,+\infty[}(Y_{i\ell} - c_{\text{sym}}) - 1)$ , in which  $Z_{i\ell} = \underline{z}_i + \sigma_i [\text{randn}(n_r, \nu)]_{i\ell}$  where  $\underline{z}_i = 2i/n_r - 0.5$ , and where  $\sigma_i = 1 + 0.2(2i/n_r - 1)$ . Let  $\mathbf{m} = (m_1, \dots, m_{n_r})$  and  $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_{n_r})$  be the vectors in  $\mathbb{R}^{n_r}$  such that  $m_j = 1.5$  for all  $j$ , and where  $\boldsymbol{\lambda}$  is the vector  $(1.000, 0.2459, 0.1082, 0.0605, 0.0385, 0.0266, 0.0195, 0.0149, 0.01117, 0.0095, 0.0078, 0.0065, 0.0056, 0.0048, 0.0042)$ . A family of orthonormal vectors  $\{\boldsymbol{\phi}^1, \dots, \boldsymbol{\phi}^{n_r}\}$  is constructed as the normalized eigenvectors associated with the  $n_r$  largest eigenvalues of the positive symmetric  $(N \times N)$  real matrix  $[b]'$   $[b]$  in which the deterministic  $(N \times N)$  real matrix  $[b]$  is generated by  $\text{rand}('state', 0)$ , and then,  $[b] = [\text{rand}(N, N)]$ . Finally the  $\nu$  independent realizations  $\mathbf{x}^{\text{exp},1}, \dots, \mathbf{x}^{\text{exp},\nu}$  are generated, for  $\ell = 1, \dots, \nu$ , by  $\mathbf{x}^{\text{exp},\ell} = \mathbf{m} + \sum_{i=1}^{n_r} \sqrt{\lambda_i} \eta_i^{\text{exp},\ell} \boldsymbol{\phi}^i$ .

**Appendix B. Data transformation for the numerical application.** In this Appendix, the function  $\mathbf{x} \mapsto \mathbf{q} = \mathbf{h}(\mathbf{x})$  from  $\mathbb{R}^N$  into  $\mathbb{R}^M$ , introduced in Section 3, is defined. It is assumed that  $M = N$ . Then, for all  $\alpha$  in  $\{1, \dots, N\}$ ,  $q_\alpha = z_\alpha + \{\widehat{\mathbf{m}}_X\}_\alpha$  in which  $z_\alpha = y_\alpha / (1 + 6y_\alpha^2)$  with  $y_\alpha = x_\alpha - \{\widehat{\mathbf{m}}_X\}_\alpha$ .



**Figure 3.5.** For  $(i, j) = (1, 2), (1, 3)$  and  $(1, 4)$ , graphs of the probability density functions  $(\xi_i, \xi_j) \mapsto p_{\Xi_i \Xi_j}(\xi_i, \xi_j)$  and  $(\xi_i, \xi_j) \mapsto p_{\Xi_i^{(N_d)} \Xi_j^{(N_d)}}(\xi_i, \xi_j)$ .

## REFERENCES

- [1] M. ARNST, R. GHANEM, AND C. SOIZE, *Identification of Bayesian posteriors for coefficients of chaos expansion*, Journal of Computational Physics, 229(9) (2010), pp. 3134-3154.
- [2] M. ARNST, R. GHANEM, E. PHIPPS, AND J. RED-HORSE, *Dimension reduction in stochastic modeling of coupled problems*, International Journal for Numerical Methods in Engineering, 92(11) (2012), pp. 940-968.
- [3] M. ARNST, C. SOIZE, AND R. GHANEM, *Hybrid sampling/spectral method for solving stochastic coupled problems*, SIAM/ASA Journal on Uncertainty Quantification, 1(1) (2013), pp. 218243.
- [4] M. ARNST, R. GHANEM, E. PHIPPS, AND J. RED-HORSE, *Reduced chaos expansions with random coefficients in reduced-dimensional stochastic modeling of coupled problems*, International Journal for Numerical Methods in Engineering, 29(5) (2014), pp. 352-376.
- [5] A.W. BOWMAN AND A. AZZALINI, *Applied Smoothing Techniques for Data Analysis*, Oxford University Press, 1997.
- [6] A. BATOU AND C. SOIZE, *Calculation of Lagrange multipliers in the construction of maximum entropy distributions in high stochastic dimension*, SIAM/ASA Journal on Uncertainty Quantification, 1(1) (2013), pp. 431-451.
- [7] K. BURRAGE, I. LENANE, AND G. LYTHER, *Numerical methods for second-order stochastic differential equations*, SIAM Journal on Scientific Computing, 29 (2007), pp. 245264.
- [8] R.H. CAMERON AND W.T. MARTIN, *The orthogonal development of non-linear functionals in series of Fourier-Hermite functionals*, The Annals of Mathematics, Second Series, 48(2) (1947), pp. 385-392.
- [9] S. DAS, R. GHANEM, AND J. SPALL, *Asymptotic sampling distribution for polynomial chaos representation of data : A maximum-entropy and fisher information approach*, SIAM Journal on Scientific Computing, 30(5) (2008), pp. 2207-2234.
- [10] B.J. DEBUSSCHERE, H.N. NAJM, P.P. PEBAY, O.M. KNIO, R. GHANEM, AND O.P. LE MAITRE, *Numerical challenges in the use of polynomial chaos representations for stochastic processes*, SIAM Journal on Scientific Comput-

- ing, 26(2) (2004), pp. 698-719.
- [11] C. DESCIELIERS, R. GHANEM, AND C. SOIZE, *Maximum likelihood estimation of stochastic chaos representations from experimental data*, International Journal for Numerical Methods in Engineering, 66(6) (2006), pp. 978-1001 .
  - [12] J. L. DOOB, *Stochastic Processes*, John Wiley and Sons, New York, 1990.
  - [13] A. DOOSTAN, R. GHANEM, AND J. RED-HORSE, *Stochastic model reduction for chaos representations*, Computer Methods in Applied Mechanics and Engineering, 196(37-40) (2007), pp. 3951-3966.
  - [14] O.G. ERNST, A. MUGLER, H.J. STARKLOFF, AND E. ULLMANN, *On the convergence of generalized polynomial chaos expansions*, ESAIM: Mathematical Modelling and Numerical Analysis, 46(2) (2012), pp. 317-339.
  - [15] S. GEMAN AND D. GEMAN, *Stochastic relaxation, Gibbs distribution and the Bayesian distribution of images*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol PAM I-6 (1984), pp. 721-741.
  - [16] R. GHANEM AND P.D. SPANOS, *Stochastic Finite Elements: A spectral Approach*, Springer-verlag, New-York, 1991 (revised edition, Dover Publications, New York, 2003).
  - [17] R. GHANEM AND R.M. KRUGER, *Numerical solution of spectral stochastic finite element systems*, Computer Methods in Applied Mechanics and Engineering, 129(3) (1996), pp. 289-303.
  - [18] R. GHANEM, R. DOOSTAN, AND J. RED-HORSE, *A probability construction of model validation*, Computer Methods in Applied Mechanics and Engineering, 197(29-32) (2008), pp. 2585-2595.
  - [19] D. GHOSH AND R. GHANEM, *Stochastic convergence acceleration through basis enrichment of polynomial chaos expansions*, International Journal for Numerical Methods in Engineering, 73(2) (2008), pp. 162-184.
  - [20] J. GUILLEMINOT AND C. SOIZE, *Stochastic model and generator for random fields with symmetry properties: application to the mesoscopic modeling of elastic random media*, Multiscale Modeling and Simulation (A SIAM Interdisciplinary Journal), 11(3) (2013), pp. 840-870.
  - [21] E. HAIRER, C. LUBICH, AND G. WANNER, *Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations*, Springer-Verlag, Heidelberg, 2002.
  - [22] W. K. HASTINGS, *Monte Carlo sampling methods using Markov chains and their applications*, Biometrika, 109 (1970), pp. 57-97.
  - [23] J. KAIPIO AND E. SOMERSALO, *Statistical and Computational Inverse Problems*, Springer-Verlag, New York, 2005.
  - [24] V. KESHAVARZZADEH, R. GHANEM, S.F. MASRI, AND O.J. ALDRAIHEM, *Convergence acceleration of polynomial chaos solutions via sequence transformation*, Computer Methods in Applied Mechanics and Engineering, 271 (2014), pp. 167-184.
  - [25] R. KHASHMINSKII, *Stochastic Stability of Differential Equations*, Series: Stochastic Modelling and Applied Probability, Vol. 66, 2nd edition, Springer, Heidelberg, 2012. Originally published in Russian, by Nauka, Moskow, 1969. First English edition published in 1980 under R.Z. Has' minski in the series Mechanics: Analysis by Sijthoff & Noordhoff.
  - [26] P. E. KLOEDEN AND E. PLATEN, *Numerical Solution of Stochastic Differentials Equations*, Springer-Verlag, Heidelberg, 1992.
  - [27] O.P. LE MAITRE AND O.M. KNIO, *Spectral Methods for Uncertainty Quantification with Applications to Computational Fluid Dynamics*, Springer, Heidelberg, 2010.
  - [28] D. LUCOR, C.H. SU, AND G.E. KARNIADAKIS, *Generalized polynomial chaos and random oscillators*, International Journal for Numerical Methods in Engineering, 60(3) (2004), pp. 571-596.
  - [29] Y.M. MARZOUK AND H.N. NAJM, *Dimensionality reduction and polynomial chaos acceleration of Bayesian inference in inverse problems*, Journal of Computational Physics, 228(6) (2009), pp. 1862-1902.
  - [30] N. METROPOLIS AND S. ULAM, *The Monte Carlo method*, Journal of the American Statistical Association, 44 (1949), pp. 335-341.
  - [31] H.N. NAJM, *Uncertainty quantification and polynomial chaos techniques in computational fluid dynamics*, Journal Review of Fluid Mechanics, 41 (2009), pp. 35-52.
  - [32] R. M. NEAL, *Slice sampling*, Annals of Statistics, 31 (2003), pp. 705-767.
  - [33] A. NOUY, *Identification of multi-modal random variables through mixtures of polynomial chaos expansions*, Comptes Rendus Mecanique, 338(12) (2010), pp. 698-703.
  - [34] A. NOUY AND C. SOIZE, *Random fields representations for stochastic elliptic boundary value problems and statistical inverse problems*, European Journal of Applied Mathematics, 25(3) (2014), pp. 339-373.
  - [35] G. PERRIN, C. SOIZE, D. DUHAMEL, AND C. FUNFSCHILLING, *Identification of polynomial chaos representations in high dimension from a set of realizations*, SIAM Journal on Scientific Computing, 34(6) (2012), pp. A2917-A2945.
  - [36] C. P. ROBERT AND G. CASELLA, *Monte Carlo Statistical Methods*, Springer-Verlag, New York, 2005.
  - [37] C. SOIZE, *The Fokker-Planck Equation for Stochastic Dynamical Systems and its Explicit Steady State Solutions*,

- World Scientific, Singapore, 1994.
- [38] C. SOIZE AND R. GHANEM, *Physical systems with random uncertainties : Chaos representation with arbitrary probability measure*, SIAM Journal on Scientific Computing, 26(2) (2004), pp. 395-410.
  - [39] C. 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, 76(10) (2008), pp. 1583-1611.
  - [40] C. SOIZE AND R. GHANEM, *Reduced chaos decomposition with random coefficients of vector-valued random variables and random fields*, Computer Methods in Applied Mechanics and Engineering, 198(21-26) (2009), pp. 1926-1934.
  - [41] C. SOIZE AND C. DESCIELIERS, *Computational aspects for constructing realizations of polynomial chaos in high dimension*, SIAM Journal on Scientific Computing, 32(5) (2010), pp. 2820-2831.
  - [42] C. SOIZE, *Identification of high-dimension polynomial chaos expansions with random coefficients for non-Gaussian tensor-valued random fields using partial and limited experimental data*, Computer Methods in Applied Mechanics and Engineering, 199(33-36) (2010), pp. 2150-2164.
  - [43] J. C. SPALL, *Introduction to Stochastic Search and Optimization*, John Wiley and Sons, Hoboken, New Jersey, 2003.
  - [44] B. SUDRET, *Global sensitivity analysis using polynomial chaos expansions*, Reliability Engineering & System Safety, 93(7) (2008), pp. 964-979.
  - [45] D. TALAY AND L. TUBARO, *Expansion of the global error for numerical schemes solving stochastic differential equation*, Stochastic Analysis and Applications, 8(4) (1990), pp. 94-120.
  - [46] D. TALAY, *Simulation and numerical analysis of stochastic differential systems*, pp. 5496, in *Probabilistic Methods in Applied Physics*, Lecture Notes in Physics, 451, P. Kree and W. Wedig, eds., Springer-Verlag, Heidelberg, 1995.
  - [47] D. TALAY, *Stochastic Hamiltonian system: exponential convergence to the invariant measure and discretization by the implicit Euler scheme*, Markov Processes and Related Fields, 8 (2002), pp. 163-198.
  - [48] R. TIPIREDDY AND R. GHANEM, *Basis adaptation in homogeneous chaos spaces*, Journal of Computational Physics, 259 (2014), pp. 304-317.
  - [49] X.L. WAN AND G.E. KARNIADAKIS, *Multi-element generalized polynomial chaos for arbitrary probability measures*, SIAM Journal on Scientific Computing, 28(3) (2006), pp. 901-928.
  - [50] D.B. XIU AND G.E. KARNIADAKIS, *Wiener-Askey polynomial chaos for stochastic differential equations*, SIAM Journal on Scientific Computing, 24(2) (2002), pp. 619-644.