C. R. Mecanique 338 (2010) 698–703

Contents lists available at ScienceDirect

Comptes Rendus Mecanique

www.sciencedirect.com

Identification of multi-modal random variables through mixtures of polynomial chaos expansions

Identification de variables aléatoires multi-modales par mélange de décompositions sur la chaos polynômial

Anthony Nouy

GeM - UMR CNRS 6183, École centrale Nantes, Université de Nantes, 1, rue de la Noë, 44321 Nantes, France

ARTICLE INFO

Article history:
Received 18 March 2010
Accepted after revision 21 September 2010

Keywords:
Statistics
Uncertainty quantification
Identification
Multi-modal density
Polynomial chaos
Finite mixture model
Spectral stochastic methods

Mots-clés :
Statistique
Quantification d'incertitudes
Identification
Densité multi-modale
Chaos Polynômial
Modèle de mélange fini
Méthodes spectrales stochastiques

ABSTRACT

A methodology is introduced for the identification of a multi-modal real-valued random variable from a collection of samples. The random variable is seen as a finite mixture of uni-modal random variables. A functional representation of the random variable is used, which can be interpreted as a mixture of polynomial chaos expansions. After a suitable separation of samples into sets of uni-modal samples, the coefficients of the expansion are identified by using an empirical projection technique. This identification procedure allows for a generic representation of a large class of multi-modal random variables with low-order generalized polynomial chaos representations.

RÉSUMÉ

Une méthodologie est proposée pour l'identification d'une variable aléatoire multi-modale à partir d'échantillons. La variable aléatoire est vue comme un mélange fini de variables aléatoires uni-modales. Une représentation fonctionnelle de la variable aléatoire est utilisée. Elle peut être interprétée comme un mélange de décompositions sur le chaos polynômial. Après une séparation adaptée des échantillons en sous-ensembles d'échantillons uni-modaux, les coefficients de la décomposition sont identifiés en utilisant une technique de projection empirique. Cette procédure d'identification permet une représentation générique d'une large classe de variables aléatoires multi-modales avec une décomposition sur chaos polynômial généralisé de faible degré.

1. Introduction

Uncertainty quantification and propagation in physical systems appear as a critical path for the improvement of the prediction of their response. For the numerical estimation of outputs of stochastic systems driven by finite-dimensional noise, the so-called spectral stochastic methods [1–4] have received a growing attention in the last two decades. These methods rely on a functional representation of random outputs, considered as second-order random variables, by using truncated expansions on suitable Hilbertian basis. Classical basis consist in polynomial functions (finite-dimensional Polynomial Chaos [5,6,1]), piecewise polynomial functions [7–9] or more general orthogonal basis [10]. Of course, the accuracy of predictions

E-mail address: anthony.nouy@ec-nantes.fr.

1631-0721/$ – see front matter
doi:10.1016/j.crme.2010.09.003

depends on the quality of the input probabilistic model. Some works have been recently devoted to the identification of random variables (or processes), from a collection of samples, using Polynomial Chaos (PC) representations. Classical inference techniques have been used to identified the coefficients of functional expansions, such as maximum likelihood estimation [11,12] or Bayesian inference [13,14]. Polynomial Chaos *a priori* allows for the representation of second-order random variables with arbitrary probability laws. However, for some classes of random variables, classical PC expansions may exhibit very slow convergence rates, thus requiring very high-order expansions for an accurate representation. When introducing such representations for random input parameters of a physical model, very high-order expansions are also required for an accurate approximation of random outputs. Classical spectral stochastic methods, such as Galerkin-type methods, then require to deal with high-dimensional approximation spaces, which leads to prohibitive computational costs. Although the use of efficient solvers or model reduction techniques based on separated representations [15–17] may help to reduce computational costs, a convenient alternative consists in identifying more suitable representations of random inputs.

The aim of the present article is to propose a PC-based numerical methodology for the identification of real-valued multi-modal random variables. In Section 2, we briefly recall the basics of PC expansions of uni-variate random variables and introduce an empirical projection technique in order to identify these expansions from samples. This projection technique is an efficient alternative to classical inference techniques. We then illustrate the limitations of classical PC expansions when trying to represent multi-modal random variables. In Section 3, we introduce a methodology for representing multi-modal real-valued random variables. From a theoretical point of view, it consists in introducing a complete set of events allowing a separation of modes. The probability density function of the random variable to be identified appears as a mixture of probability density functions of random variables (finite mixture model [18]). We then propose a natural representation of the random variable on a generalized Polynomial Chaos, which can be interpreted as a "mixture of chaos expansions", estimated from samples using an efficient empirical projection. Section 4 will illustrate the efficiency of the proposed methodology.

2. Polynomial chaos decomposition of a second-order random variable

2.1. Polynomial chaos representation

Let X denote a real-valued random variable defined on an abstract probability space $(\Theta, \mathcal{B}_\Theta, P)$. Let F_X denote its cumulative density function (CDF) and p_X its probability density function (PDF). We introduce a random variable ξ defined on $(\Theta, \mathcal{B}_\Theta, P)$, with known support $\Xi \subset \mathbb{R}$ and probability law P_ξ, thus defining a new probability space $(\Xi, \mathcal{B}_\Xi, P_\xi)$. The random variable $g(\xi) := F_X^{-1} \circ F_\xi(\xi)$ have the same probability law as X. We then make the hypothesis that g is a P_ξ-square integrable function from Ξ to $\mathbb{R}$, i.e. $g \in L^2(\Xi, dP_\xi)$. Introducing an Hilbertian basis $\{h_i\}_{i\in\mathbb{N}}$ of $L^2(\Xi, dP_\xi)$, the random variable X then admits the following representation: $X = \sum_{i\in\mathbb{N}} X_i h_i(\xi)$, with coefficients X_i being defined by $X_i = \langle g, h_i\rangle_{L^2(\Xi,dP_\xi)} := E(g(\xi)h_i(\xi))$, where E denotes the mathematical expectation. An approximate representation of X can be obtained by truncating the decomposition: $X \approx \sum_{i=0}^{p} X_i h_i(\xi)$. A classical choice for the h_i consists in polynomial functions orthonormal with respect to scalar product $\langle\cdot,\cdot\rangle_{L^2(\Xi,dP_\xi)}$, thus leading to the so-called uni-dimensional Polynomial Chaos (PC) expansion of degree p of X [1,19,10].

2.2. Identification of the decomposition

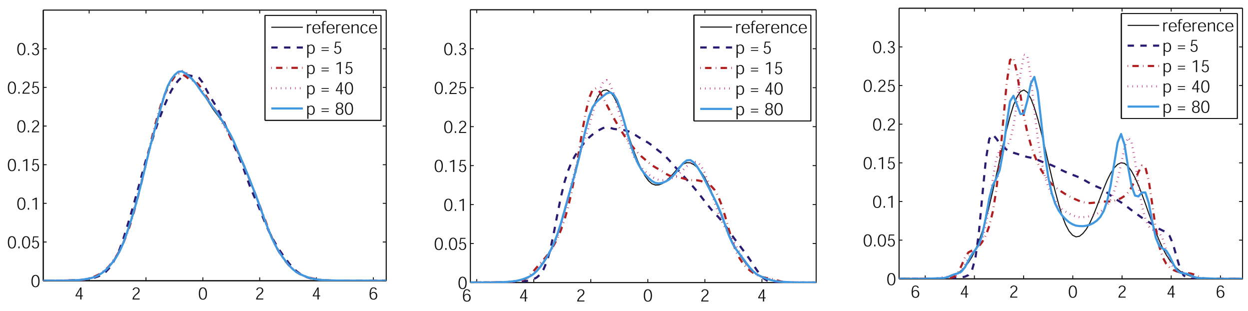
Classical inference techniques have been applied for the identification of coefficients X_i from a collection of independent samples $\{X^{(k)}\}_{k=1}^{Q}$ of X: maximum likelihood estimation [11], Bayesian inference [13]. Here, in order to estimate the coefficients, we use a simple and efficient numerical technique, named "empirical projection". It is based on the estimation of mapping g from samples and the introduction of a quadrature scheme to compute its projection on the PC basis. We denote by $\widetilde{F}_X$ the empirical CDF of X, estimated from samples: $\widetilde{F}_X(x) = \frac{1}{Q}\sum_{k=1}^{Q} I(X^{(k)} \leqslant x)$, where $I(A)$ is the indicator function of event A. We then introduce the following approximation $g(\xi) \approx \widetilde{F}_X^{-1} \circ F_\xi(\xi)$, where $\widetilde{F}_X^{-1} : [0,1] \to \mathbb{R}$ is uniquely defined as $\widetilde{F}_X^{-1}(y) = \min\{x \in \{X^{(k)}\}_{k=1}^{Q};\ \widetilde{F}_X(x) \geqslant y\}$. Then, the coefficients of the chaos expansion can be approximated using a numerical integration:

$$X_i = \int_\Xi F_X^{-1}\big(F_\xi(y)\big)h_i(y)\,\mathrm{d}P_\xi(y) \approx \sum_{k=1}^{N} \omega_k \widetilde{F}_X^{-1}\big(F_\xi(y_k)\big)h_i(y_k) \tag{1}$$

where the $\{\omega_k, y_k\}_{k=1}^{N}$ are integration weights and points. In practice, an accurate Gauss-quadrature associated with measure P_ξ can be used.

2.3. Limitations of classical polynomial chaos representations

Classical polynomial chaos decompositions allow for an accurate representation of a wide class of probability laws. The accuracy can be simply improved by choosing a suitable germ ξ (Gaussian, Uniform, etc.), associated with classical orthogonal polynomial basis (Hermite, Legendre, etc.). However, these classical polynomial decompositions may not be adapted



(a) $\delta = 1$ (b) $\delta = 1.5$ (c) $\delta = 2$

Fig. 1. PC expansion of a bi-modal random variable: convergence with the expansion's degree p.

for some classes of random variables, particularly for multi-modal random variables. Fig. 1 illustrates the convergence of a Hermite polynomial chaos expansion[1] of a bi-modal random variable $X(\theta)$, defined by

$$X(\theta) = \begin{cases} a(\theta) - \delta & \text{if } b(\theta) < 1/3 \\ a(\theta) + \delta & \text{if } b(\theta) > 1/3 \end{cases}$$

where a and b are independent standard Gaussian random variables and where δ is a parameter controlling the separation of modes. We observe that when increasing δ, the convergence of a classical PC expansion drastically deteriorates, thus needing for a high polynomial degree for an accurate representation of the PDF.

3. Identification of a multi-modal random variable

3.1. Mixture of probability laws

Let us denote by m the number of modes of the random variable X (which can be defined as the number of local maxima of p_X). We introduce a complete set[2] of m events $\{\Theta_i\}_{i=1}^m$ of $\mathcal{B}_\Theta$ and we define associated real-valued random variables Y_i with probability law defined for all $B \in \mathcal{B}_\mathbb{R}$ by

$$P_{Y_i}(B) = P(X \in B|\Theta_i) = P\big(X^{-1}(B)|\Theta_i\big) = P\big(X^{-1}(B) \cap \Theta_i\big)/P(\Theta_i)$$

We admit that the events Θ_i are such that the Y_i are uni-modal random variables. The probability law of X can then be defined by: $\forall B \in \mathcal{B}_\mathbb{R}$,

$$P_X(B) = \sum_{i=1}^m P\big(X^{-1}(B) \cap \Theta_i\big) = \sum_{i=1}^m P_{Y_i}(B)P(\Theta_i)$$

Its probability density function then appears as a mixture [18] (i.e. convex combination) of probability density functions of uni-modal random variables:

$$p_X(x) = \sum_{i=1}^m p_{Y_i}(x)P(\Theta_i) \tag{2}$$

The identification of X is then replaced by the identification of random variables Y_i, which are expected to admit accurate low-order chaos representations. The questions are now: how to define the partition $\{\Theta_i\}_{i=1}^m$, what kind of chaos representation can be used for X and how to identify this representation from samples?

3.2. Definition of Θ_i by separation of samples

We introduce an artificial separation of samples $\{X^{(k)}\}_{k=1}^Q$ into m sets of uni-modal samples, which allows for the construction of the desired partition of Θ. We suppose that the empirical PDF allows estimating a set of points $\{x_i\}_{i=1}^{m-1}$ that separate samples into m sets of uni-modal samples, defined as follows:

$$\mathcal{X}_i = \big\{X^{(k)}, k \in \{1, \dots, Q\}; X^{(k)} \in [x_{i-1}, x_i)\big\}, \quad i \in \{1, \dots, m\} \tag{3}$$

[1] Expansions are identified with the empirical projection technique, using a highly accurate (and converged) Gauss-Hermite quadrature.

[2] $\bigcup_{i=1}^m \Theta_i = \Theta$, $\Theta_i \cap \Theta_j = \emptyset$ for $i \neq j$.

where by convention $x_0 = -\infty$ and $x_m = +\infty$. Then, Θ_i is defined as the abstract event associated with samples in $\mathcal{X}_i$ (i.e. $\mathcal{X}_i \subset X(\Theta_i)$). The probability of event Θ_i is then defined by $P(\Theta_i) = Card(\mathcal{X}_i)/Q$. In order to define the set of events $\{\Theta_i\}_{i=1}^m$, we introduce a partition of $[0,1)$, defined by intervals $B_i = [z_{i-1}, z_i)$, $i = 1,\dots,m$, where $0 = z_0 < z_1 < \dots < z_m = 1$. Then, introducing a uniform random variable $\xi_1 \in U(0,1)$, we define $\Theta_i = \xi_1^{-1}(B_i)$, for $i = 1,\dots,m$. It completely characterizes the partition of $[0,1) := \Xi_1$, with $z_i = \sum_{j=1}^i P(\Theta_j)$, $i = 1,\dots,m$.

3.3. Mixture of polynomial chaos expansions

Let us now denote ξ_2 another random variable, independent on ξ_1. Letting $\boldsymbol{\xi} = (\xi_1, \xi_2)$, we define a 2-dimensional probability space $(\boldsymbol{\Xi}, \mathcal{B}_{\boldsymbol{\Xi}}, P_{\boldsymbol{\xi}})$, with $\boldsymbol{\Xi} = \Xi_1 \times \Xi_2$ and $P_{\boldsymbol{\xi}} = P_{\xi_1} \otimes P_{\xi_2}$. Random variable X is then seen as a function of $\boldsymbol{\xi}$, defined by $X(\boldsymbol{\xi}) = \sum_{i=1}^m I_{B_i}(\xi_1) Y_i(\xi_2)$, where I_{B_i} denotes the indicator function of B_i. We next introduce a chaos representation of each random variable $Y_i = \sum_{j\in\mathbb{N}} X_{i,j} h_j(\xi_2)$, where the h_j are classical orthonormal polynomials in $L^2(\Xi_2, dP_{\xi_2})$. A generalized chaos representation of X is then sought as:

$$X(\boldsymbol{\xi}) = \sum_{i=1}^m I_{B_i}(\xi_1)\left(\sum_{j\in\mathbb{N}} X_{i,j} h_j(\xi_2)\right) = \sum_{i=1}^m \sum_{j\in\mathbb{N}} X_{i,j} I_{B_i}(\xi_1) h_j(\xi_2) \tag{4}$$

which can be interpreted as a mixture of polynomial chaos expansions. Functions $\{I_{B_i} h_j\}$ form an orthogonal set of functions in $L^2(\boldsymbol{\Xi}, dP_{\boldsymbol{\xi}}) = L^2(\Xi_1, dP_{\xi_1}) \otimes L^2(\Xi_2, dP_{\xi_2})$, composed by piecewise polynomial functions (polynomial with respect to ξ_2 and piecewise constant with respect to ξ_1). The L^2-norm of a basis function is $E((I_{B_i}(\xi_1) h_j(\xi_2))^2)^{1/2} = P(\Theta_i)^{1/2}$. Coefficients $\{X_{i,j}\}$ of the decomposition of X are defined as the orthogonal projections of X on these basis functions:

$$X_{i,j} = P(\Theta_i)^{-1} \int_{\Xi_1 \times \Xi_2} X(y_1, y_2) I_{B_i}(y_1) h_j(y_2) p_{\xi_2}(y_2)\, \mathrm{d}y_1\, \mathrm{d}y_2 = \int_{\Xi_2} Y_i(y_2) h_j(y_2) p_{\xi_2}(y_2)\, \mathrm{d}y_2 \tag{5}$$

3.4. Identification of the decomposition from samples

Classical inference techniques [20,21] could be used in order to identify from samples the set of $m(p+1)$ coefficients[3] of the mixture of chaos expansions (4). However, the number of parameters is such that these classical techniques lead to high computational costs. With a maximum likelihood estimation, the identification requires the resolution of a hard optimization problem (high dimension, objective function with many local maxima, possibly non-smooth function) for which classical algorithms may lack robustness. Here, we propose to apply the empirical projection technique introduced in Section 2.2. The random variable Y_i is written in terms of ξ_2 in the following way: $Y_i(\xi_2) = F_{Y_i}^{-1} \circ F_{\xi_2}(\xi_2)$, where $F_{Y_i} = F_{X|\Theta_i}$ is the conditional CDF of X knowing Θ_i. The subset of samples $\mathcal{X}_i$ corresponds to independent samples of Y_i. Therefore, an approximation $\widetilde{F}_{Y_i}$ of F_{Y_i} can be simply estimated by

$$\widetilde{F}_{Y_i}(y) = \frac{1}{\mathrm{Card}(\mathcal{X}_i)} \sum_{z\in\mathcal{X}_i} I(z \leqslant y)$$

Random variable X is then approximated by truncating polynomial chaos expansions to a degree p, the coefficients being estimated from samples in the following way: $\forall i \in \{1,\dots,m\}$, $\forall j \in \{0,\dots,p\}$,

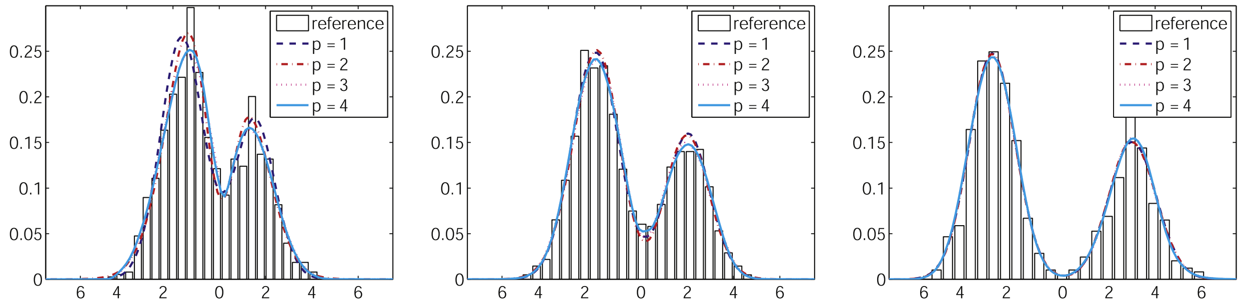
$$X_{i,j} \approx \sum_{k=1}^N \omega_k \widetilde{F}_{Y_i}^{-1}\big(F_{\xi_2}(y_k)\big) h_j(y_k) \tag{6}$$

where the $\{\omega_k, y_k\}_{k=1}^N$ are integration weights and points of a classical quadrature rule (e.g. Gauss-quadrature) associated with measure P_{ξ_2}.

4. Numerical illustration

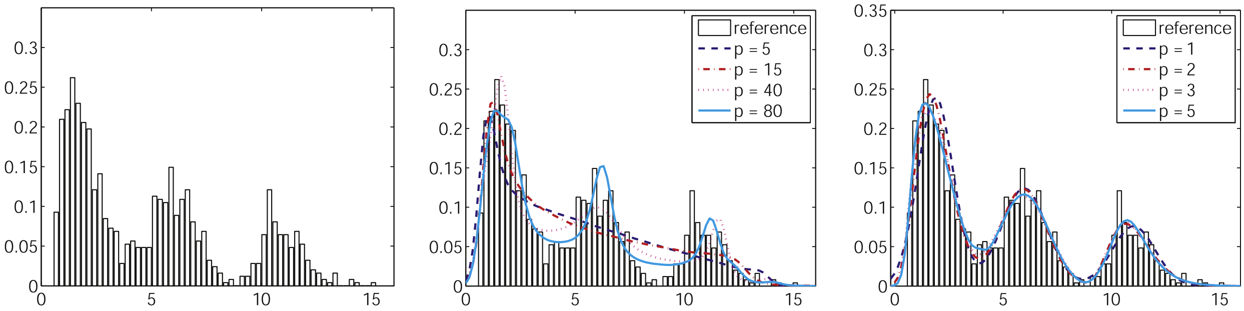
Example 1. We generate an artificial collection of $Q = 1000$ samples from the bi-modal random variable defined in Section 2.3. We consider three cases corresponding to the following three values of the mode-separation parameter: $\delta = 1.5$, $\delta = 2$ and $\delta = 3$. The corresponding empirical PDFs of samples are shown in Figs. 2(a)–(c). On the same figures, also plotted are the PDFs associated with a mixture of Hermite polynomial chaos expansions (ξ_2 is a Gaussian random variable). The coefficients of the expansion have been obtained with the empirical projection technique (see Section 3.4), using a 15-points Gauss-Hermite quadrature for the numerical integration. For the three cases, samples have been separated into two sets of uni-modal samples by choosing separation values $x_1 = 0$, 0.2 and 0 respectively. Theses values have been determined by

[3] Note that samples separation values x_i, $i = 1,\dots,m-1$, could also be added to the set of parameters to be identified.



(a) $\delta = 1.5$ (b) $\delta = 2$ (c) $\delta = 3$

Fig. 2. Mixture of PC expansions for a bi-modal random variable: convergence with the expansion's degree p.



(a) (b) (c)

Fig. 3. Probability density functions: samples (a), Hermite PC expansion (b), mixture of Hermite PC expansions (c).

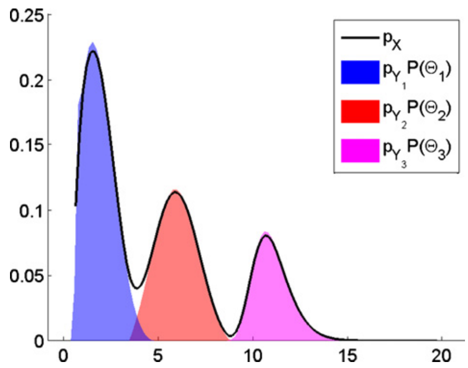


Fig. 4. Weighted probability density functions $\{y \mapsto P(\Theta_i) p_{Y_i}(y)\}$ of random variables Y_i, $i = 1, \ldots, 3$, identified with a Hermite PC expansion of degree $p = 3$.

simply locating local minima of the empirical PDFs. Whatever the separation of modes, we observe a very good agreement between the empirical PDFs of samples and the identified mixture of polynomial chaos expansions (4), even with a low degree of expansion ($p = 2$ or 3).

Example 2. We consider a collection of $Q = 1000$ samples corresponding to a 3-modal distribution, represented on Fig. 3(a). On Fig. 3(b), we illustrate the bad convergence of a classical Hermite PC expansion $X \approx \sum_{i=0}^{p} X_i h_i(\xi)$ (ξ being a standard Gaussian random variable). The coefficients are computed using the empirical projection technique (see Section 2.2) with a high-order Gauss-Hermite quadrature (100 quadrature points). Fig. 3(c) illustrates the PDF obtained with a 3-modal mixture of Hermite PC expansions of degree p: $X(\xi_1, \xi_2) = \sum_{i=1}^{3} \sum_{j=0}^{p} X_{i,j} I_{B_i}(\xi_1) h_j(\xi_2)$. The samples have been separated into three sets of uni-modal samples by choosing separation values $x_1 = 4$ and $x_2 = 9$. These values have been chosen by approximatively locating the two local minima of the empirical PDF. For the computation of expansion coefficients, we have used the empirical projection technique introduced in Section 3.4, with a 15-points Gauss-Hermite quadrature. We observe that a very good representation of the random variable is obtained with a mixture of PC expansions of low degree ($p \approx 3$).

Fig. 4 shows the weighted PDFs $y \mapsto P(\Theta_i) p_{Y_i}(y)$ of random variables Y_i. It also shows the resulting PDF of X, which appears as the mixture of the weighted conditional PDFs.

5. Conclusion

In this Note, we have introduced an efficient numerical technique for the identification of real-valued multi-modal random variables. A mixture of chaos representations is used, which can be interpreted as a 2-dimensional generalized polynomial chaos expansion. The expansion basis is defined by the product of polynomial functions of a first random variable and piecewise constant functions of a second random variable. The coefficients of the expansion are estimated from samples by using an efficient empirical projection technique. Classical inference techniques such as maximum likelihood or Bayesian inference could also be used for estimating the coefficients of the expansion, although leading to much higher computational costs. The proposed mixture of polynomial chaos expansions and the empirical projection technique can be extended to vector-valued random variables. The empirical projection technique however requires more and more samples as the dimension increases.

References

[1] R. Ghanem, P. Spanos, Stochastic Finite Elements: A Spectral Approach, Springer, Berlin, 1991.
[2] A. Nouy, Recent developments in spectral stochastic methods for the numerical solution of stochastic partial differential equations, Archives of Computational Methods in Engineering 16 (3) (2009) 251–285.
[3] D. Xiu, Fast numerical methods for stochastic computations: A review, Communications in Computational Physics 5 (2009) 242–272.
[4] H.G. Matthies, Stochastic finite elements: Computational approaches to stochastic partial differential equations, Zamm-Zeitschrift fur Angewandte Mathematik und Mechanik 88 (11) (2008) 849–873.
[5] N. Wiener, The homogeneous chaos, Am. J. Math. 60 (1938) 897–936.
[6] R.H. Cameron, 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) 385–392.
[7] M. Deb, I. Babuška, J.T. Oden, Solution of stochastic partial differential equations using Galerkin finite element techniques, Computer Methods in Applied Mechanics and Engineering 190 (2001) 6359–6372.
[8] O.P. Le Maître, O.M. Knio, H.N. Najm, R.G. Ghanem, Uncertainty propagation using Wiener–Haar expansions, Journal of Computational Physics 197 (1) (2004) 28–57.
[9] X. Wan, G.E. Karniadakis, Multi-element generalized polynomial chaos for arbitrary propability measures, SIAM Journal on Scientific Computing 28 (3) (2006) 901–928.
[10] C. Soize, R. Ghanem, Physical systems with random uncertainties: Chaos representations with arbitrary probability measure, SIAM Journal on Scientific Computing 26 (2) (2004) 395–410.
[11] C. Desceliers, R. Ghanem, C. Soize, Maximum likelihood estimation of stochastic chaos representations from experimental data, International Journal for Numerical Methods in Engineering 66 (6) (2006) 978–1001.
[12] G. Stefanou, A. Nouy, A. Clément, Identification of random shapes from images through polynomial chaos expansion of random level-set functions, International Journal for Numerical Methods in Engineering 79 (2) (2009) 127–155.
[13] R. Ghanem, A. Doostan, On the construction and analysis of stochastic models: Characterization and propagation of the errors associated with limited data, Journal of Computational Physics 217 (1) (2006) 63–81.
[14] M. Arnst, R. Ghanem, C. Soize, Identification of bayesian posteriors for coefficients of chaos expansions, Journal of Computational Physics 229 (9) (2010) 3134–3154.
[15] A. Nouy, A generalized spectral decomposition technique to solve a class of linear stochastic partial differential equations, Computer Methods in Applied Mechanics and Engineering 196 (45-48) (2007) 4521–4537.
[16] A. Nouy, Generalized spectral decomposition method for solving stochastic finite element equations: Invariant subspace problem and dedicated algorithms, Computer Methods in Applied Mechanics and Engineering 197 (2008) 4718–4736.
[17] A. Nouy. Proper Generalized Decompositions and separated representations for the numerical solution of high dimensional stochastic problems. Archives of Computational Methods in Engineering (2010), doi:10.1007/s11831-010-9054-1, in press.
[18] G.J. McLachlan, D. Peel, Finite Mixture Models, Wiley, New York, 2000.
[19] D. Xiu, G.E. Karniadakis, The Wiener–Askey polynomial chaos for stochastic differential equations, SIAM Journal on Scientific Computing 24 (2) (2002) 619–644.
[20] C.P. Robert, The Bayesian Choice, Springer Verlag, New York, 1994.
[21] A. Gelman, J.B. Carlin, H.S. Stern, D.B. Rubin, Bayesian Data, Analysis, Chapman and Hall, Boca Raton, FL, 2003.