Available online at www.sciencedirect.com
ScienceDirect
C. R. Mecanique 336 (2008) 518–523
COMPTES RENDUS MECANIQUE
http://france.elsevier.com/direct/CRAS2B/

Sparse polynomial chaos expansions and adaptive stochastic finite elements using a regression approach

Géraud Blatman [a,b,*], Bruno Sudret [b]

[a] *IFMA-LaMI, Campus des Cézeaux, BP 265, 63175 Aubière cedex, France*
[b] *EDF R&D, Département matériaux et mécanique des composants, site des Renardières, 77250 Moret-sur-Loing cedex, France*

Received 18 January 2008; accepted after revision 25 February 2008

Available online 28 April 2008

Presented by Jean-Baptiste Leblond

Abstract

A method is proposed to build a sparse polynomial chaos (PC) expansion of a mechanical model whose input parameters are random. In this respect, an adaptive algorithm is described for automatically detecting the significant coefficients of the PC expansion. The latter can thus be computed by means of a relatively small number of possibly costly model evaluations, using a non-intrusive regression scheme (also known as stochastic collocation). The method is illustrated by a simple polynomial model, as well as the example of the deflection of a truss structure. ***To cite this article: G. Blatman, B. Sudret, C. R. Mecanique 336 (2008).***

Résumé

Chaos polynomial creux et éléments finis stochastiques adaptatifs : une approche par régression. Dans cette communication, on propose un algorithme permettant de construire une représentation par *chaos polynomial creux* de la réponse d'un modèle mécanique dont les paramètres d'entrée sont aléatoires. L'algorithme construit de façon adaptative une représentation creuse en détectant automatiquement les termes importants et en supprimant ceux qui sont négligeables. A chaque étape, le calcul des coefficients s'effectue par minimisation au sens des moindres carrés (méthode non-intrusive dite de *régression*). L'algorithme est déroulé pas à pas sur un modèle polynomial, puis appliqué à l'étude de la fiabilité d'un treillis élastique. ***Pour citer cet article : G. Blatman, B. Sudret, C. R. Mecanique 336 (2008).***

Keywords: Solids and structures; Adaptive stochastic finite elements; Sparse polynomial chaos; Stochastic collocation; Regression; Structural reliability

Mots-clés : Solides et structures ; Eléments finis stochastiques adaptatifs ; Chaos polynomial creux ; Collocation stochastique ; Régression ; Fiabilité

* Corresponding author.
E-mail addresses: geraud.blatman@edf.fr (G. Blatman), bruno.sudret@edf.fr (B. Sudret).

1631-0721/$ – see front matter
doi:10.1016/j.crme.2008.02.013

Version française abrégée

La propagation des incertitudes sur les paramètres dans les modèles mécaniques a fait l'objet de nombreux travaux de recherche depuis une vingtaine d'années. Les approches spectrales s'appuyant sur un développement de la réponse aléatoire dans une base de chaos polynomial [1,2] ont montré leur efficacité pour traiter des problèmes de mécanique, diffusion, transfert thermique et plus récemment, fiabilité des structures [6,7,9].

Les méthodes non-intrusives de calcul des coefficients du chaos (projection ou régression) permettent d'obtenir ces derniers à partir d'un certain nombre d'évaluations du modèle déterministe (e.g. modèle éléments finis) sans avoir à intervenir dans le code lui-même. Cependant, le nombre de termes du développement, et donc le nombre de calculs déterministes à effectuer pour les estimer, augmente très rapidement avec le nombre de variables aléatoires d'entrée du modèle. En pratique cependant, beaucoup de ces coefficients s'avèrent négligeables a posteriori.

L'objectif de cette communication est de proposer une approximation par *chaos polynomial creux*, dans laquelle seuls les termes importants du développement sont calculés de façon adaptative par un algorithme itératif. A chaque étape, des termes d'ordre supérieur sont ajoutés un par un au développement courant, et seuls ceux qui conduisent à une augmentation significative du coefficient de détemination R^2 de la régression sont retenus. Puis les coefficients retenus aux étapes précédentes sont supprimés un par un tant que cette suppression ne conduit pas à dégrader R^2.

L'algorithme est déroulé pas à pas sur un modèle polynomial, montrant ainsi sa capacité à retrouver la structure creuse. On applique enfin la méthode à un calcul de fiabilité d'un treillis élastique (10 variables aléatoires). La probabilité de défaillance est obtenue avec une grande précision, pour un nombre de termes dans le chaos à peu près 7 fois plus petit que dans une représentation classique. Le coût de calcul (nombre d'appels au code éléments finis) est de façon corrolaire divisé par un facteur compris entre 3 et 5.

1. Introduction

Uncertainty propagation methods, which aim at studying the influence of the randomness of input parameters of a model onto the model response, have received much attention in the past twenty years. Among others, *spectral stochastic methods* based on polynomial chaos (PC) expansions [1,2] have shown their great potential in various applications ranging from diffusion and thermal problems [3], computational fluid dynamics [4,5] and structural reliability [6,7].

Polynomial chaos (PC) expansions allow one to represent explicitly the random response of a mechanical system whose input parameters are modelled by random variables. The PC coefficients may be efficiently computed using *non-intrusive* techniques such as projection [8] or regression [9]. However, the required number of model evaluations (i.e. the computational cost) increases with the PC size, which itself dramatically increases with the number of input variables when the common truncation scheme of the PC expansion is applied. To circumvent this problem, an adaptive algorithm is proposed in order to retain only the significant PC coefficients, leading to a sparse PC representation (see also the *stochastic reduced basis* technique proposed by [10]).

The basics of polynomial chaos expansion is first recalled in Section 2. Then the sparse PC representation and the associated adaptive algorithm is detailed in Section 3. The method is finally applied to the study of the reliability of a truss structure having uncertain mechanical properties.

2. Polynomial chaos representation

2.1. Polynomial chaos representation of a numerical model with uncertain input

Consider a mechanical system described by a numerical model $\mathcal{M}$ which can be analytical or more generally algorithmic (e.g. a finite element model). Suppose that this model has M uncertain input parameters which are represented by *independent* random variables $(X_1, \ldots, X_M)$ gathered in a random vector $\boldsymbol{X}$ of prescribed joint probability density function $p_{\boldsymbol{X}}(\boldsymbol{x})$. Hence the model response denoted by $Y = \mathcal{M}(\boldsymbol{X})$ is also random. For the sake of simplicity, Y is assumed to be scalar throughout the paper (in case of a vector response $\boldsymbol{Y}$, the following derivations hold componentwise). Provided the random variable Y has finite variance, it can be expressed as follows [2]:

$$Y = \mathcal{M}(\boldsymbol{X}) = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^M} a_{\boldsymbol{\alpha}} \psi_{\boldsymbol{\alpha}}(\boldsymbol{X}) \tag{1}$$

This expansion is referred to as the *finite-dimensional polynomial chaos* (PC) *representation* of Y. The $a_{\boldsymbol{\alpha}}$'s are unknown deterministic coefficients and the $\psi_{\boldsymbol{\alpha}}$'s are multivariate polynomials which are orthonormal with respect to the joint PDF $p_{\boldsymbol{X}}$ of the input random vector $\boldsymbol{X}$, i.e. $\mathrm{E}[\psi_{\boldsymbol{\alpha}}(\boldsymbol{X})\psi_{\boldsymbol{\beta}}(\boldsymbol{X})] = 1$ if $\boldsymbol{\alpha} = \boldsymbol{\beta}$ and 0 otherwise. For instance, if $\boldsymbol{X}$ is a standard normal random vector, the $\psi_{\boldsymbol{\alpha}}$ are normalized multivariate Hermite polynomials.

2.2. Non-intrusive computation of the polynomial chaos coefficients

The PC coefficients can be estimated using a non-intrusive regression scheme [11,9,12]. This method requires the choice of a truncation of the PC *ab initio*, hence of a non-empty finite set $\mathcal{A} = \{\boldsymbol{\alpha}_0, \ldots, \boldsymbol{\alpha}_{P-1}\} \subset \mathbb{N}^M$ which contains the multi-indices of the retained basis polynomials $\psi_{\boldsymbol{\alpha}_0}, \ldots, \psi_{\boldsymbol{\alpha}_{P-1}}$. $\mathcal{A}$ is referred to as the *truncation set* in the sequel. The corresponding PC approximation is denoted by $Y_{\mathcal{A}} \equiv \mathcal{M}_{\mathcal{A}}(\boldsymbol{X}) = \sum_{\boldsymbol{\alpha}\in\mathcal{A}} a_{\boldsymbol{\alpha}}\psi_{\boldsymbol{\alpha}}(\boldsymbol{X})$ which rewrites $Y_{\mathcal{A}} = \boldsymbol{a}^{\mathsf{T}}\boldsymbol{\psi}(\boldsymbol{X})$, by introducing the vector notation $\boldsymbol{a} = \{a_{\boldsymbol{\alpha}_0}, \ldots, a_{\boldsymbol{\alpha}_{P-1}}\}^{\mathsf{T}}$ and $\boldsymbol{\psi}(\boldsymbol{X}) = \{\psi_{\boldsymbol{\alpha}_0}(\boldsymbol{X}), \ldots, \psi_{\boldsymbol{\alpha}_{P-1}}(\boldsymbol{X})\}^{\mathsf{T}}$.

Let us consider a set of realizations of $\boldsymbol{X}$ denoted by $\mathcal{X} = \{\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(N)}\}$ and referred to as the *experimental design* (ED). Let us denote by $\mathcal{Y}$ the associated set of model response, say $\mathcal{Y} = \{\mathcal{M}(\boldsymbol{x}^{(1)}), \ldots, \mathcal{M}(\boldsymbol{x}^{(N)})\}$. The unknown coefficients $\boldsymbol{a}$ may be computed by performing a least-square minimization [9], i.e. by minimizing the mean-square truncation error $1/N\sum_{i=1}^{N}(\mathcal{M}(\boldsymbol{x}^{(i)}) - \mathcal{M}_{\mathcal{A}}(\boldsymbol{x}^{(i)}))^2$. Using the above notation the solution reads:

$$\hat{\boldsymbol{a}} = (\boldsymbol{\Psi}^{\mathsf{T}}\boldsymbol{\Psi})^{-1}\boldsymbol{\Psi}^{\mathsf{T}}\mathcal{Y} \tag{2}$$

where $\boldsymbol{\Psi}$ is a $N \times P$ matrix such that $\boldsymbol{\Psi}_{ij} = \psi_{\boldsymbol{\alpha}_j}(\boldsymbol{x}^{(i)})$, $i = 1, \ldots, N$, $j = 0, \ldots, P-1$.

The size N of the ED must be greater than P to make this problem well posed. For practical implementation, the series in Eq. (1) is commonly truncated by retaining those polynomials $\psi_{\boldsymbol{\alpha}}$ whose total degree $|\boldsymbol{\alpha}|$ is less than p. This leads to the truncation set $\mathcal{A}^{M,p} = \{\boldsymbol{\alpha} \in \mathbb{N}^M\colon \sum_{i=1}^{M}\alpha_i \leqslant p\}$. Accordingly, the number of PC terms is given by $P = \binom{M+p}{p}$. Hence it dramatically increases with both p and M. Consequently, the minimal size of the ED that is required for an accurate solution of the regression problem [13] blows up. Thus increasing the accuracy of the PC expansion may lead to intractable calculations in high dimensions. Nevertheless, as all the input variables do not have the same influence on the response and as only low order interactions are physically meaningful in practice, both P and N might be reduced by only retaining a small number of important coefficients, i.e. by an appropriate choice of a *sparse* truncation set $\mathcal{A} \subset \mathbb{N}^M$ such that $\mathrm{card}\,\mathcal{A} \ll \mathrm{card}\,\mathcal{A}^{M,p}$ for a given accuracy.

3. Adaptive polynomial chaos approximation

3.1. Assessment of the goodness-of-fit of the PC approximation

Let $\mathcal{A}$ be a non-empty finite subset of $\mathbb{N}^M$ and $Y_{\mathcal{A}}$ be the associated truncated PC expansion. Let us denote by $\hat{Y}_{\mathcal{A}}$ its regression-based approximation, whose coefficients are computed by regression from a given ED. The latter can be selected randomly [11] or based on the roots of orthogonal polynomials [9]. The error of approximation can be quantified by the empirical *lack of fit* defined by:

$$\widehat{LOF}_{\mathcal{A}} = \frac{\sum_{i=1}^{N}(\mathcal{M}(\boldsymbol{x}^{(i)}) - \boldsymbol{a}^{\mathsf{T}}\boldsymbol{\psi}(\boldsymbol{x}^{(i)}))^2}{\sum_{i=1}^{N}(\mathcal{M}(\boldsymbol{x}^{(i)}) - \bar{f})^2} \tag{3}$$

where $\bar{f} = 1/N\sum_{i=1}^{N}\mathcal{M}(\boldsymbol{x}^{(i)})$. The accuracy of the response approximation $\hat{Y}_{\mathcal{A}}$ can thus be assessed by means of the *determination coefficient* defined by $R^2_{\mathcal{A}} = 1 - \widehat{LOF}_{\mathcal{A}}$. The value $R^2_{\mathcal{A}} = 1$ indicates a perfect fit of the true model response Y, whereas $R^2_{\mathcal{A}} = 0$ indicates no linear relationship between Y and the multivariate polynomials $\{\psi_{\boldsymbol{\alpha}}, \boldsymbol{\alpha} \in \mathcal{A}\}$.

3.2. Adaptive algorithm to construct a sparse PC expansion

Let us respectively define the *interaction order* j and the *total degree* p of any multi-index $\boldsymbol{\alpha} \in \mathbb{N}^M$ by:

$$j = \sum_{i=1}^{M}\mathbf{1}_{\alpha_i>0}(\boldsymbol{\alpha}), \qquad p = |\boldsymbol{\alpha}| = \sum_{i=1}^{M}\alpha_i \tag{4}$$

where $\mathbf{1}_{\alpha_i>0}(\boldsymbol{\alpha})$ equals 1 if $\alpha_i > 0$ and 0 otherwise. Moreover, for any truncation set $\mathcal{A} \subset \mathbb{N}^M$, the quantity $\max_{\boldsymbol{\alpha}\in\mathcal{A}} |\boldsymbol{\alpha}|$ is referred to as the *maximal degree* of the PC expansion. Lastly, let us denote by $\mathcal{I}_{j,p}$ the set of multi-indices of interaction order j and total degree p. An iterative adaptive procedure is now presented for constructing a sparse PC approximation of the system response:

(i) Choose an ED $\mathcal{X}$ and perform the model evaluations $\mathcal{Y}$ once and for all.
(ii) Select the values of the algorithm parameters, i.e. the target accuracy R^2_{tgt}, the maximal PC degree $p_{\max}$ and maximal interaction order $j_{\max}$ and the cut-off values ε_1, ε_2.
(iii) Initialize the algorithm: $p = 0$, truncation set $\mathcal{A}^0 = \{\mathbf{0}\}$, where $\mathbf{0}$ is the null element of $\mathbb{N}^M$.
(iv) For any degree $p \in \{1, \ldots, p_{\max}\}$:
- **Forward step**: for any interaction order $j \in \{1, \ldots, j_{\max}\}$, gather the candidate terms in a set $\mathcal{I}_{j,p}$. Add each candidate term to $\mathcal{A}^{p-1}$ *one-by-one* and compute the PC expansion coefficients by regression (Eq. (2)) and the associated determination coefficient R^2 in each case. Retain eventually those candidate terms that lead to a significant increase in R^2, i.e. greater than ε_1, and discard the other candidate terms. Let $\mathcal{A}^{p,+}$ be the final truncation set at this stage.
- **Backward step**: remove in turn each term in $\mathcal{A}^{p,+}$ of degree strictly not greater than p. In each case, compute the PC expansion coefficients and the associated determination coefficient R^2. Eventually discard from $\mathcal{A}^{p,+}$ those terms that lead to an insignificant decrease in R^2, i.e. less than ε_2. Let $\mathcal{A}^p$ be the final truncation set.
- If $R^2_{\mathcal{A}^p} \geqslant R^2_{\text{tgt}}$, stop.

Note that the various regression analyses only involve analytical computations (see Eq. (2)). Thus their computational cost is usually small compared to that associated to the model evaluations onto the ED.

3.3. The adaptive algorithm in action

The adaptive procedure detailed above is now tested on the following simple polynomial model: $\mathcal{M}(\xi_1, \xi_2) = 1 + H_1(\xi_1)H_1(\xi_2) + H_3(\xi_1)$, where H_k is the k-th Hermite polynomial and ξ_1, ξ_2 are independent standard normal variables. A random design made of $N = 100$ *Latin Hypercube samples* (LHS) (see e.g. [14]) is used. The various steps of the PC construction are illustrated in Fig. 1. As the model itself is polynomial, the exact solution should be retrieved. Thus the target determination coefficient is set equal to $R^2_{\text{tgt}} = 1$.

The iterations on the PC total degree p and the interaction order j are respectively displayed from left to right and from top to bottom in Fig. 1. The polynomials that are discarded during the procedure are grey-shaded. In this example the polynomials $H_1(\xi_1)$ and $H_2(\xi_2)$ are correctly neglected in the forward steps associated with $p = 1$ and $p = 2$ respectively. All the remaining useless polynomials are removed in the last backward step that is associated with $p = 3$.

Note that the standard "full" truncation set $\mathcal{A}^{2,3}$ has 10 coefficients, whereas only three coefficients are eventually required to represent the model.

4. Application example: reliability of a truss structure

Let us consider the truss structure sketched in Fig. 2. Ten independent input random variables are considered, namely the Young's moduli and the cross-section areas of the horizontal and the oblical bars (respectively denoted by E_1, A_1 and E_2, A_2) and the applied loads (denoted by $P_i, i = 1, \ldots, 6$) [7].

The deflection at midspan V_1 is regarded as the model random response. It is approximated by a PC expansion in normalized multivariate Hermite polynomials (note that the input random variables listed in Fig. 2 are first transformed into standard normal variables). A reliability analysis is carried out with respect to the failure criterion $\{V_1 > 11 \text{ cm}\}$. An ED made of $N = 100$ LHS is used. The maximum degree $p_{\max}$ is set equal to 5, the maximum interaction order $j_{\max}$ to 2 and the cut-off values $\varepsilon_1, \varepsilon_2$ to 5×10^{-6}. A parametric study is performed varying the target accuracy R^2_{tgt}.

Results are compared in terms of the generalized reliability index $\beta = -\Phi^{-1}(P_f)$, where $P_f = \mathbb{P}(V_1 > 11 \text{ cm})$ denotes the probability of failure. The reference value is obtained using crude Monte Carlo simulation of the problem (1,000,000 samples, i.e. 1,000,000 finite element runs are used). A PC-based solution is also computed using a full PC of degree $p = 3$ ($P = \binom{10+3}{3} = 286$), whose coefficients were computed by regression (a Latin Hypercube design

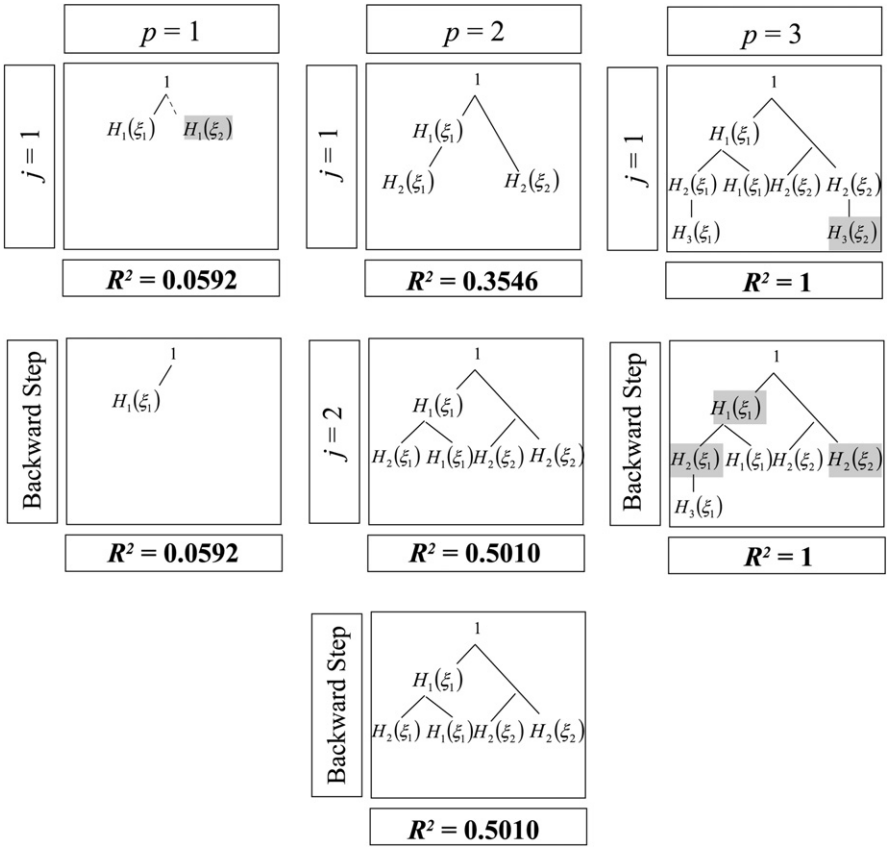
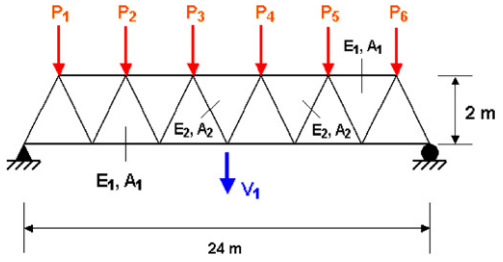


Fig. 1. Polynomial model – adaptive construction of the PC approximation.



Variable	Distribution	Mean	Standard deviation
E_1, E_2 (Pa)	Lognormal	2.10×10^{11}	2.10×10^{10}
A_1 (m^2)	Lognormal	2.0×10^{-3}	2.0×10^{-4}
A_2 (m^2)	Lognormal	1.0×10^{-3}	1.0×10^{-4}
P_1–P_6 (N)	Gumbel	5.0×10^{4}	7.5×10^{3}

Fig. 2. Truss example – Statement of the problem.

of size $N = 500$ was used). The PC-based reliability results are obtained by sampling the PC expansion. Results are reported in Table 1 together with the size P_{final} and total degree p_{final} of the resulting PC approximation as well as the maximum size $P_{\max}$ attained within the iterations of the adaptive algorithm.

It appears that the PC approximations provide estimates of β which are all the more accurate since the target accuracy R^2_{tgt} is high. In particular, a relative error on β less than 5% is obtained when setting $R^2_{\text{tgt}} = 0.9900$, by using only $N = 100$ samples whereas the full PC model of order $p = 3$ would require more than 286 samples. As a whole, the number of terms $P = 43$ required to evaluate accurately the probability of failure is 7 times less than that of a full PC representation. A computational gain factor ranging from 3 (when N is set equal to the minimum required number of samples, i.e. $N = P = 286$) to 5 (for the current choice $N = 2P$) is thus obtained. It is also observed that $P_{\max}$ increases with the target accuracy, and hence choosing values too close to 1 for R^2_{tgt} might lead to overfitting as $N \simeq P$ in this case, misleadingly yielding a value of R^2 very close to 1.

Table 1
Truss example – Reliability index obtained from various sparse PC representations

Target accuracy R^2_{tgt}	Number of FE runs	Prob. of failure	Reliability index β	$P_{\max}$	P_{final}	p_{final}
0.9000	100	3.24×10^{-3}	2.7225	11	11	1
0.9900	100	6.73×10^{-3}	2.4716	21	21	2
0.9990	100	8.43×10^{-3}	2.3899	64	41	2
0.9999	100	8.76×10^{-3}	2.3757	70	43	3
Full PC ($p = 3$)	500	8.69×10^{-3}	2.3784	286	286	
Reference solution	1,000,000	8.70×10^{-3}	2.3781			

5. Conclusion

A method is proposed to build iteratively a PC expansion of the random response of a model with random input parameters. It is based on an adaptive algorithm which automatically detects the significant PC terms, leading to a sparse PC representation. The retained PC coefficients can thus be computed efficiently by regression using a rather low number N of model evaluations compared to what would be required to compute a "full" PC approximation. The step-by-step application of the algorithm to a polynomial model $\mathcal{M}$ shows that it satisfactorily yields the exact solution in this case. The truss example shows that the algorithm may be used for solving structural reliability problems. Accurate results may be obtained in reliability analysis using few runs of the (FE) model when the values of the threshold parameters are reasonably low. An adaptive scheme aimed at optimizing adaptively the number N of model evaluations, i.e. the size of the experimental design in order to avoid overfitting problems, is currently being investigated. Furthermore, error estimates based on resampling techniques are being studied to improve the robustness of the method.

References

[1] R. Ghanem, P. Spanos, Stochastic Finite Elements: A Spectral Approach, Courier Dover Publications, 2003.
[2] C. Soize, R. Ghanem, Physical systems with random uncertainties: chaos representations with arbitrary probability measure, SIAM J. Sci. Comput. 26 (2) (2004) 395–410.
[3] D. Xiu, G.E. Karniadakis, A new stochastic approach to transient heat conduction modeling with uncertainty, Int. J. Heat Mass Transfer 46 (2003) 4681–4693.
[4] O.P. Le Maître, O.M. Knio, H.N. Najm, R.G. Ghanem, A stochastic projection method for fluid flow, J. Comput. Phys. 173 (2001) 481–511.
[5] O.M. Knio, O.P. Le Maître, Uncertainty propagation in CFD using polynomial chaos decomposition, Fluid Dyn. Res. 38 (9) (2006) 616–640.
[6] B. Sudret, A. Der Kiureghian, Comparison of finite element reliability methods, Prob. Eng. Mech. 17 (2002) 337–348.
[7] G. Blatman, B. Sudret, M. Berveiller, Quasi-random numbers in stochastic finite element analysis, Mécanique & Industries 8 (2007) 289–297.
[8] D.M. Ghiocel, R.G. Ghanem, Stochastic finite element analysis of seismic soil–structure interaction, J. Eng. Mech. (ASCE) 128 (2002) 66–77.
[9] M. Berveiller, B. Sudret, M. Lemaire, Stochastic finite element: a non-intrusive approach by regression, Rev. Européenne Mécanique Numérique 15 (1–3) (2006) 81–92.
[10] A. Nouy, A generalized spectral decomposition technique to solve stochastic partial differential equations, Comput. Methods Appl. Mech. Engrg. 196 (45–48) (2007) 4521–4537.
[11] S.K. Choi, R.V. Grandhi, R.A. Canfield, Structural reliability under non-Gaussian stochastic behavior, Computers & Structures 82 (2004) 1113–1121.
[12] B. Sudret, Uncertainty propagation and sensitivity analysis in mechanical models – Contributions to structural reliability and stochastic spectral methods, Habilitation à diriger des recherches, Université Blaise-Pascal, Clermont-Ferrand, France, 2007.
[13] B. Sudret, Global sensitivity analysis using polynomial chaos expansions, Reliab. Eng. Sys. Safety 93 (2008) 964–979.
[14] M.D. McKay, R.J. Beckman, W.J. Conover, A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, Technometrics 2 (1979) 239–245.