Sparse polynomial chaos expansions based on an adaptive Least Angle Regression algorithm

G. BLATMAN[a,b], **B. SUDRET**[b,c]

a. IFMA-LaMI, Campus des Cézeaux, BP 265, 63175 Aubière cedex
b. EDF R&D, Site des Renardières, 77250 Moret-sur-Loing cedex
c. Phimeca Engineering S.A., Centre d'affaires du Zénith, 34 rue de Sarlième, 63800 Cournon

Résumé :

Dans cette communication, on propose un algorithme basé sur la technique Least Angle Regression *(LAR) pour construire une représentation par* chaos polynomial creux *de la réponse d'un modèle mécanique dont les paramètres d'entrée sont aléatoires. Le plan d'expériences est automatiquement enrichi de sorte à éviter les problèmes se surapprentissage. On obtient au final une représentation ne comportant qu'un faible nombre de termes non nuls, qui peuvent être estimés au moyen d'un nombre réduit d'évaluations du modèle. L'algorithme est appliqué au calcul des moments statistiques du tassement d'une fondation sur un sol dont le module d'Young est modélisé par un champ aléatoire.*

Abstract :

A method is proposed to build up a sparse polynomial chaos (PC) expansion of a mechanical model whose input parameters are random. In this respect, an adaptive algorithm based on Least Angle Regression *(LAR) is described for automatically detecting the significant coefficients of the PC expansion. The experimental design is automatically enriched in order to avoid overfitting problems. Eventually the coefficients of the resulting sparse PC approximation may be computed by means of a relatively small number of possibly costly model evaluations, using a non intrusive regression scheme. The method is illustrated by the moment analysis of the settlement of a foundation on a soil layer whose Young's modulus is a random field.*

Mots clefs : Sparsity-of-effects principle, sparse polynomial chaos expansion, adaptive Least Angle Regression

1 Introduction

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 [1] or regression [2]. 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 (*i.e.* retain all the multivariate polynomials of total degree not greater than a prescribed p). To circumvent this problem, a truncation strategy based on the use of q-norms with $0 < q < 1$ is proposed. It is motivated by the so-called *sparsity-of-effects principle* [3], which states that most models are principally governed by main effects and low-order interactions. The related truncated PC expansions contain a low number of likely important terms compared to the full representation.
Using these truncation strategies, an adaptive algorithm based on *Least Angle Regression* (LAR) [4] is proposed in order to retain progressively a small number of significant PC coefficients, leading to a *sparse* PC representation. Beside the adaptivity in terms of PC basis, the experimental design is systematically complemented such that the overfitting phenomenon is avoided.

2 Polynomial chaos expansion of the model response

2.1 Mathematical framework

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 $f_{\boldsymbol{X}}(\boldsymbol{x})$. 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 that the random variable Y has a finite variance, it may be recast as follows [5] :

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

This expansion is referred to as the *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 $f_{\boldsymbol{X}}$ of the input random vector $\boldsymbol{X}$, *i.e.* $\mathbb{E}\left[\psi_{\boldsymbol{\alpha}}(\boldsymbol{X})\psi_{\boldsymbol{\beta}}(\boldsymbol{X})\right] = 1$ 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 Estimation of the polynomial chaos coefficients

The PC coefficients can be estimated using a non intrusive regression scheme [2, 6]. This method requires the choice of a truncation of the PC *ab initio, i.e.* 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 :

$$\boldsymbol{a} = \{a_{\boldsymbol{\alpha}_0}, \ldots, a_{\boldsymbol{\alpha}_{P-1}}\}^{\mathsf{T}} \tag{2}$$

$$\boldsymbol{\psi}(\boldsymbol{X}) = \{\psi_{\boldsymbol{\alpha}_0}(\boldsymbol{X}),\; \ldots\; , \psi_{\boldsymbol{\alpha}_{P-1}}(\boldsymbol{X})\}^{\mathsf{T}} \tag{3}$$

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*. Let us denote by $\mathcal{Y}$ the associated set of model response quantities, 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, *i.e.* by minimizing the mean-square truncation error $1/N \sum_{i=1}^{N} \left(\mathcal{M}(\boldsymbol{x}^{(i)}) - \mathcal{M}_{\mathcal{A}}(\boldsymbol{x}^{(i)})\right)^2$. Using the above notation the solution reads :

$$\hat{\boldsymbol{a}} = (\boldsymbol{\Psi}^{\mathsf{T}}\boldsymbol{\Psi})^{-1}\boldsymbol{\Psi}^{\mathsf{T}}\mathcal{Y} \tag{4}$$

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.

2.3 Moment analysis by post-processing the polynomial chaos expansion

The second moments of the PC approximation of the model response may be derived *analytically* from the coefficients. In particular, the mean and the standard deviation are respectively given by :

$$\mu_Y^{(P)} \;=\; a_{\boldsymbol{0}} \qquad , \qquad \sigma_Y^{(P)} \;=\; \sum_{0<|\boldsymbol{\alpha}|\leq p} a_{\boldsymbol{\alpha}}^2 \tag{5}$$

where P denotes the number of terms in the PC truncation. The skewness and the kurtosis coefficients may be also computed by algebraic combinations of the coefficients [7]. However the corresponding computational cost blows up in case of a large number M of input variables. To circumvent this problem, one may generate a large random sample and evaluate the PC approximation at all the sample points. Thus the higher-order statistics of the PC approximation may be obtained by an elementary statistical analysis of the PC-based sample. It is also possible to compute the so-called *global sensitivity indices* directly from the PC coefficients, see [8].

2.4 The issue of truncating the polynomial chaos expansion

For computational purpose it is necessary to truncate the polynomial chaos expansion in Eq.(1) as mentioned in Section 2.2. In most papers in the literature the truncation sets $\mathcal{A}^p$ correspond to those multivariate basis polynomials $\psi_{\boldsymbol{\alpha}}$ whose total degree is not greater than p, that is :

$$\mathcal{A}^{M,p} \;\equiv\; \mathcal{A}_1^{M,p} \;=\; \left\{\boldsymbol{\alpha} \in \mathbb{N}^M \;:\; \|\boldsymbol{\alpha}\|_1 \leq p\right\} \tag{6}$$

where $\|\cdot\|_1$ denotes the 1-norm on $\mathbb{R}^M$ defined by :

$$\|\boldsymbol{z}\|_1 \;=\; \sum_{i=1}^{M} |z_i| \quad \forall\, \boldsymbol{z} = \{z_1, \ldots, z_M\}^{\mathsf{T}} \in \mathbb{R}^M \tag{7}$$

Note that the subscript 1 in $\mathcal{A}_1^{M,p}$ corresponds to the choice of the 1-norm. One limitation of this truncation strategy lies in the strong increase of unknown PC coefficients with M and p. Indeed the number of multi-indices in $\mathcal{A}_1^{M,p}$ is given by :

$$\text{card}\left(\mathcal{A}_1^{M,p}\right) = \binom{M+p}{p} \tag{8}$$

As a consequence, the minimal size N of the experimental design (*i.e.* the number of model evaluations) that ensures the well-posedness of the regression problem increases itself considerably with M and p. The above truncation strategy may thus lead to intractable calculations in high dimensions for a computationally demanding model $\mathcal{M}$.

3 Sparse polynomial chaos approximation

3.1 PC expansions based on the sparsity-of-effects principle

An alternative strategy is proposed in the present paper for truncating the PC expansion of the model response. It is motivated by the so-called *sparsity-of-effects principle* [3], which states that most models are principally governed by main effects and low order interactions. One proposes the use of the following truncation sets based on q-norms, $0 < q < 1$:

$$\mathcal{A}_q^{M,p} = \left\{ \boldsymbol{\alpha} \in \mathbb{N}^M \, : \, \|\boldsymbol{\alpha}\|_q \equiv \left(\sum_{i=1}^{M} \alpha_i^q \right)^{1/q} \leq p \right\} \tag{9}$$

Such norms penalize the higher order interaction terms all the more since q is low. Note that setting q equal to 1 corresponds to the usual truncation scheme mentioned in Section 2. The proposed truncation strategy thus leads to PC expansions with a reduced number of unknown coefficients, which may be computed using a moderate number N of model evaluations.
The computational cost may be further reduced by taking into account the fact that the PC expansion of the model response contains only a small number of significant terms (*sparse* PC expansion). This is the scope of the next section.

3.2 Sparse PC approximation using an adaptive LAR algorithm

3.2.1 Estimation of the PC coefficients using LAR

Regularization is a technique that allows one to perform a least-square regression when the number of model evaluations N is less than the number P of basis functions. It relies upon a penalization of some norm (or more generally functional) of the regression coefficients. In particular, $\mathcal{L}^1$-regularized regression consists in fitting a metamodel $\mathcal{M}_{\mathcal{A}_q^{M,p}}$ by solving :

$$\text{Minimize} \quad \sum_{i=1}^{N} \left(y^{(i)} - \sum_{\boldsymbol{\alpha} \in \mathcal{A}_q^{M,p}} a_{\boldsymbol{\alpha}} \psi(\boldsymbol{x}^{(i)}) \right)^2 \qquad \text{subject to} \quad \sum_{\boldsymbol{\alpha} \in \mathcal{A}_q^{M,p}} |a_{\boldsymbol{\alpha}}| \leq s \tag{10}$$

where $s \geq 0$ is a tuning parameter. The $\mathcal{L}^1$-type constraint of this optimization problem yields a *sparse* solution (*i.e.* with many components equal to zero). In other words, a selection of a small number of significant terms in the basis $\mathcal{A}_q^{M,p}$ is performed. The solution is all the sparser since the value of s is low.
Least Angle Regression (LAR) [4] is an efficient algorithm for solving the problem in Eq.(10). It provides in one shot the entire paths of solution coefficients as s is increased from 0 up to a maximum value. The LAR procedure is described below :

1. Standardize the vectors $\{\boldsymbol{\Psi}_{\boldsymbol{\alpha}_i}, i = 1, \ldots, P-1\}$ to have empirical mean zero and empirical variance one, where $\boldsymbol{\Psi}_{\boldsymbol{\alpha}_i} \equiv \{\psi_{\alpha_i}(\boldsymbol{x}^{(1)}), \ldots, \psi_{\alpha_i}(\boldsymbol{x}^{(N)})\}^\mathsf{T}$. Initialize the approximated response vector $\hat{\mathcal{Y}} = \boldsymbol{0}$, which corresponds to the initial coefficients $a_{\boldsymbol{\alpha}_0}, \ldots, a_{\boldsymbol{\alpha}_{P-1}} = 0$. Define the residual $\boldsymbol{R} = \mathcal{Y} - \hat{\mathcal{Y}}$.
2. Find the vector $\boldsymbol{\Psi}_{\boldsymbol{\alpha}_j}$ which is most correlated with $\boldsymbol{R}$.
3. Move $a_{\boldsymbol{\alpha}_j}$ from 0 towards the value $\boldsymbol{\Psi}_{\boldsymbol{\alpha}_j}^\mathsf{T} \boldsymbol{R}$, until some other vector $\boldsymbol{\psi}_{\boldsymbol{\alpha}_k}$ has as much correlation with the current residual as does $\boldsymbol{\psi}_{\boldsymbol{\alpha}_j}$.
4. Move jointly $\{a_{\boldsymbol{\alpha}_j}, a_{\boldsymbol{\alpha}_k}\}^\mathsf{T}$ in the direction defined by their least-square coefficients of the current residual on $\{\boldsymbol{\psi}_{\boldsymbol{\alpha}_j}, \boldsymbol{\psi}_{\boldsymbol{\alpha}_k}\} \equiv \boldsymbol{\Phi}$, until some other vector $\boldsymbol{\psi}_{\boldsymbol{\alpha}_l}$ has as much correlation with the current residual. This direction is defined explicitly by the vector $\{\hat{a}_{\boldsymbol{\alpha}_j}, \hat{a}_{\boldsymbol{\alpha}_k}\}^\mathsf{T} = (\boldsymbol{\Phi}^\mathsf{T}\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^\mathsf{T}\boldsymbol{R}$.

5. Continue this way until all P predictors have been entered. After P steps, one gets the full least-square solution.
6. Compute accuracy estimates of the metamodels associated with all the solution coefficients, and select the truncation set $\mathcal{A}^{p,*}$ that corresponds to the largest accuracy estimate.

LAR may be thus regarded as a way of selecting an optimal sparse PC basis. One chooses to eventually apply ordinary least-square regression to recompute the corresponding coefficients, as it is considered to provide more accurate estimates. The accuracy estimates that are used in Step 6 are based on *leave-one-out cross validation* [9] and are denoted by $\hat{Q}^2$.

3.2.2 Sparse PC expansion using an adaptive LAR scheme

A limitation of LAR lies on the requirement of an *a priori* truncation set $\mathcal{A}_q^{M,p}$. To circumvent this difficulty, one proposes a procedure for progressively enriching the truncation set of the PC approximation, *i.e.* the set of active basis functions. One first selects the type q of the norm that is used to truncate the PC expansions as well as a target accuracy Q^2_{tgt}. The proposed adaptive procedure is outlined below :

1. Select an ED $\mathcal{X}$ and collect the corresponding model evaluations in $\mathcal{Y}$ once and for all.
2. Initialize the PC degree $p = 0$ and the accuracy estimate $\hat{Q}^{2,p^*} = 0$.
3. Set $\mathcal{A}^p = \mathcal{A}_q^{M,p}$. Apply the LAR algorithm to fit the metamodel $\mathcal{M}_{\mathcal{A}^p}$. Store the accuracy estimate $\hat{Q}^{2,p}$.
4. If $p \geq 2$: if $\hat{Q}^{2,p} \leq \hat{Q}^{2,p-1} \leq \hat{Q}^{2,p-2}$ (overfitting), then enrich the ED and go back to Step 3.
5. Set $p^* = p$ if $\hat{Q}^{2,p} \geq \hat{Q}^{2,p^*}$.
6. If $\hat{Q}^{2,p^*} \geq Q^2_{tgt}$: stop and return the optimal truncation set $\mathcal{A}^{p^*}$. Otherwise set $p = p + 1$ and go back to Step 3.

The condition in Step 4 is a heuristic criterion to avoid overfitting. One eventually applies ordinary least-square regression to recompute the coefficients corresponding to $\mathcal{A}^{p^*}$.

4 Statistical moment analysis of a foundation

4.1 Problem statement

Let us study the problem of the settlement of a foundation, already adressed in [10]. An elastic soil layer of thickness t and lying on a rigid substratum is considered. A structure to be founded on this soil mass is idealized as a uniform pressure P applied over a length $2B$ of the free surface (see Figure 1). The soil is modelled as an elastic linear isotropic material. A plane strain analysis is carried out.

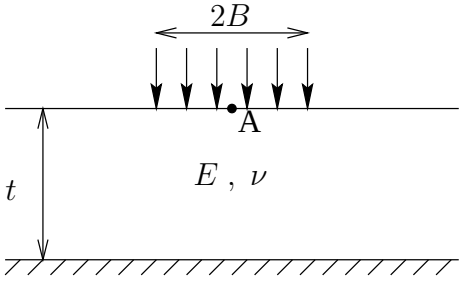

FIG. 1 – Settlement of a foundation - problem definition

Due to the symmetry, half of the structure is modelled by finite elements. Strictly speaking, there is no symmetry in the system when random fields of material properties are introduced. However, it is believed that this simplification does not significantly influence the results. The foundation (resp. mesh) width is equal to $10\ \mathrm{m}$ (resp. $60\ \mathrm{m}$). The soil layer thickness is equal to $30\ \mathrm{m}$ and its Poisson's ratio is equal to 0.3. The applied pressure is equal to $0.2\ \mathrm{MPa}$. The finite element mesh displayed in Figure 2-a was chosen.

It contains 768 elements and 825 nodes. The maximal displacement under the foundation (point A in Figure 1) computed with this mesh is equal to 5.49 cm.

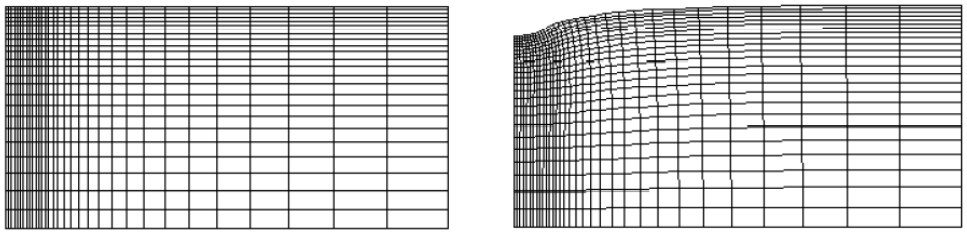

FIG. 2 – Settlement of a foundation - finite element mesh and deformed shape for mean values of the parameters by a deterministic analysis

4.2 Probabilistic model

The Young's modulus of the soil is considered to vary both in the vertical and the horizontal directions. It is modelled by a two-dimensional homogeneous lognormal random field. Its mean value is set equal to $\mu_E = 50\,\text{MPa}$ and its coefficient of variation is $\delta_E = \sigma_E/\mu_E = 0.3$. The autocorrelation coefficient function of the underlying Gaussian field $N(\boldsymbol{x}, \omega)$ is :

$$\rho_N(\boldsymbol{x}, \boldsymbol{x}') \;=\; \exp\left[-\frac{\|\boldsymbol{x} - \boldsymbol{x}'\|^2}{\ell^2}\right] \tag{11}$$

where $\ell = 10$ m. $N(\boldsymbol{x}, \omega)$ is discretized using the *Karhunen-Loève expansion* [11] :

$$N(\boldsymbol{x}, \omega) \;\simeq\; \sum_{i=1}^{M} \sqrt{\lambda_i}\xi_i(\omega)\varphi_i(\boldsymbol{x}) \tag{12}$$

where the $\sqrt{\lambda_i}$'s (resp. the $\varphi_i(\boldsymbol{x})$'s) are the eigenvalues (resp. eigenfunctions) of the random field covariance kernel, and the $\xi_i(\omega)$'s form a set of *independent* standard Gaussian random variables. A relative accuracy in the variance less than 1% is obtained when using $M = 40$ terms in the discretization of $N(\boldsymbol{x}, \omega)$.

4.3 Moment analysis

Of interest are the four first statistical moments of the maximum vertical displacement. Reference results are obtained by direct Monte Carlo simulation (50,000 model evaluations are performed) and bootstrap resampling (1,000 resamples are used). On the other hand, estimates of the moments are provided using the proposed adaptive LAR methodology. In this purpose, one considers Hermite PC approximations of the model response that are truncated using a $q = 0.6$-norm. The target accuracy Q^2_{tgt} is successively set equal to 0.95 and 0.995. The various moments estimates are gathered in Table 1.

TAB. 1 – Estimates of the four first statistical moments of the maximum vertical displacement

Moments	Reference	Sparse PC approximations	
		$Q^2_{tgt} = 0.95$	$Q^2_{tgt} = 0.995$
Mean (cm)	5.91	5.90	5.91
StD (cm)	1.13	1.07	1.13
Skewness	0.6	0.2	0.6
Kurtosis	3.7	3.1	3.4
# FE runs	50,000	250	1,000
Final degree	-	4	4
$IS\ (\%)$ †	-	52	53

† $IS \equiv$ Number of non zero terms/Total number of terms $= (\sum_{\boldsymbol{\alpha} \in \mathcal{A}_q^{M,p}} \mathbf{1}_{\{a_{\boldsymbol{\alpha}} \neq 0\}})\ /\ (|\mathcal{A}_q^{M,p}|)$

It appears that the sparse PC approximation corresponding to $Q^2_{tgt} = 0.95$ yields good estimates of the mean and the standard deviation, with respective relative errors of 0.1% and 5.3% (with respect to the reference values). Note that the associated computational cost (*i.e.* $N = 250$ model evaluations) would only allow one to compute an usual (*i.e.* with $q = 1$) full PC metamodel of degree $p = 1$ (with a number of terms given by $P = 41$). In contrast, a full PC representation of degree $p = 2$ would require more than $P = 861$ model evaluations. More accurate estimates of the moments are provided when setting Q^2_{tgt} equal to 0.995. Indeed, a two-digit accuracy is obtained on the mean and the standard deviation as well as a one-digit accuracy on the skewness coefficient. One also gets a relative error of 8.1% on the kurtosis coefficient. Lastly, both PC approximations reveal relatively sparse, with a *index of sparsity* IS close to 50%. Other applications show a significant gain in sparse chaos representation, with an index of sparsity $IS =$ close to 5%, see *e.g.* [12].

5 Conclusion

An adaptive LAR procedure is proposed to build up a sparse PC representation of the random response of a model with random input parameters. In order to reduce the number of unknown PC coefficients to identify and hence the required number of computer experiments (*i.e.* the computational cost), an adaptive algorithm is proposed for automatically detecting the significant PC terms (sparse PC representation). The experimental design is automatically enriched such that the overfitting phenomenon is avoided. The example of the foundation shows that the algorithm may be used to efficiently estimate the statistical moments of the model response, leading to to a considerable reduction of the number of model evaluations compared to crude Monte Carlo simulation.

Références

[1] D.M. Ghiocel and R.G. Ghanem. Stochastic finite element analysis of seismic soil-structure interaction. *J. Eng. Mech.*, 128 :66–77, 2002.

[2] M. Berveiller, B. Sudret, and M. Lemaire. Stochastic finite elements : a non intrusive approach by regression. *Eur. J. Comput. Mech.*, 15(1-3) :81–92, 2006.

[3] D.C. Montgomery. *Design and analysis of experiments*. 2004.

[4] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani. Least angle regression. *Ann. Statist.*, 32 :407–499, 2004.

[5] C. Soize and R. Ghanem. Physical systems with random uncertainties : chaos representations with arbitrary probability measure. *SIAM J. Sci. Comput.*, 26(2) :395–410, 2004.

[6] B. Sudret. *Uncertainty propagation and sensitivity analysis in mechanical models – Contributions to structural reliability and stochastic spectral methods*, 2007. Habilitation à diriger des recherches, Université Blaise Pascal, Clermont-Ferrand, France.

[7] M. Berveiller. *Eléments finis stochastiques : approches intrusive et non intrusive pour des analyses de fiabilité*. PhD thesis, Université Blaise Pascal, Clermont-Ferrand, 2005.

[8] B. Sudret. Global sensitivity analysis using polynomial chaos expansions. *Reliab. Eng. Sys. Safety*, 93 :964–979, 2008.

[9] A. Elisseff and M. Pontil. Leave-one-out error and stability of learning algorithms with applications. In G. Horvath J. Suykens and S. Basu et. Al, editors, *Learning Theory and Practice*. NATO ASI Series, IOS Press, Amsterdam, 2002.

[10] B. Sudret and A. Der Kiureghian. Stochastic finite elements and reliability : a state-of-the-art report. Technical Report UCB/SEMM-2000/08, University of California, Berkeley, 2000. 173 pages.

[11] M. Loève. *Probability theory*. Springer Verlag, New-York, 1977.

[12] G. Blatman and B. Sudret. Use of sparse polynomial chaos expansions in adaptive stochastic finite element analysis. *Prob. Eng. Mech.*, 2008. accepted for publication.