Proceedings of the ASME 2012 11th Biennial Conference on Engineering Systems Design and Analysis
ESDA 2012
July 2-4, 2012, Nantes, France

ESDA2012-82301

A REGRESSION BASED NON-INTRUSIVE METHOD USING SEPARATED REPRESENTATION FOR UNCERTAINTY QUANTIFICATION

Prashant Rai
Mathilde Chevreuil
Anthony Nouy*
LUNAM Université
Ecole Centrale Nantes, Université de Nantes
GeM, UMR CNRS 6183
Nantes, France

Jayant Sen Gupta
EADS Innovation Works- Applied Mathematics
EADS France Campus Engineering
Blagnac, France

ABSTRACT

This paper aims at handling high dimensional uncertainty propagation problems by proposing a tensor product approximation method based on regression techniques. The underlying assumption is that the model output functional can be well represented in a separated form, as a sum of elementary tensors in the stochastic tensor product space. The proposed method consists in constructing a tensor basis with a greedy algorithm and then in computing an approximation in the generated approximation space using regression with sparse regularization. Using appropriate regularization techniques, the regression problems are well posed for only few sample evaluations and they provide accurate approximations of model outputs.

1 INTRODUCTION

Uncertainty quantification has emerged as a crucial field of investigation for various branches of science and engineering. Over the last decade, considerable efforts have been made in the development of new methodologies based on a functional point of view in probability, where random outputs of simulation codes are approximated with suitable functional expansions. Typically, when considering a function $u(\xi)$ of input random parameters $\xi = (\xi_1 \dots \xi_d)$, an approximation is searched under the form $u(\xi) \approx \sum_{i=1}^{N} u_i \phi_i(\xi)$ where the $\phi_i(\xi)$ constitute a suitable basis of multiparametric functions (e.g. polynomial chaos basis).

Several methods have been proposed for the evaluation of functional expansions [1, 2]. Non intrusive techniques as L^2 projection or regression methods allow the estimation of expansion coefficients by using evaluations of the numerical model at certain sample points, thus allowing the simple use of existing deterministic simulation codes. However the dimension N of classical approximation spaces has an exponential (or factorial) increase with dimension d and hence the computational cost becomes prohibitively high as one needs to evaluate the model for a large number of samples $Q \approx N$. The question is: can we construct a representation of the high dimensional object u, given the fact that we have only limited information on it ? We are particularly interested in the case where the dimension d is large but the "effective dimensionality" of the underlying function is fairly small.

In order to handle high-dimensional models, we here propose a regression-based tensor approximation method, which exploits the tensor structure of the stochastic function spaces. The underlying assumption is that the model output functional can be well represented in a low dimensional basis composed of elementary tensors (rank-one functions). Tensor approximation methods have recently been applied to many areas of scientific computing for representing elements in high dimensional tensor

*Address all correspondence to this author.

product spaces [3]. In the context of uncertainty quantification, for problems involving very high stochastic dimension, instead of evaluating the coefficients of an expansion in a given approximation basis (e.g. polynomial chaos), function u is approximated in suitable low-dimensional tensor subsets (e.g. rank m tensors) which are low-dimensional manifolds of the underlying tensor space. The dimensionality of these manifolds typically grows linearly with dimension d and therefore, it addresses the curse of dimensionality. Note that a regression-based method has already been proposed in [4] for the construction of tensor approximations of multivariate functionals. Here, we propose an alternative construction of tensor approximations using greedy algorithms and sparse regularization techniques.

The proposed method consists in approximating the model with a m-term representation $u_m(\xi) = \sum_{i=1}^m \alpha_i w_i(\xi)$ where the w_i are selected in a suitable low-dimensional tensor subset $\mathscr{M}$ (typically rank-one elements) and where the α_i are real coefficients. Sparse regularization techniques are used in order to retain only the most significant basis functions, which results in an improvement of robustness of the regression-based tensor approximation method when dealing with a limited number of samples. As a result, the proposed technique allows to approximate the response of models with a large number of random inputs even with a limited number of model evaluations.

The outline of the paper is as follows. In section 2, we introduce some basic concepts about functional approaches in uncertainty propagation. We also detail several methods based on regression for the computation of approximate functional expansions. In section 3, we introduce the proposed tensor approximation method based on regularized regression. Finally the ability of the proposed method to handle high dimensional uncertainty propagation problems is illustrated on numerical applications in section 4.

2 FUNCTIONAL REPRESENTATION AND REGRESSION METHODS

2.1 Stochastic Function Spaces and Their Tensor Structure

We here introduce the definitions of stochastic functions spaces and their approximations. Let $(\Xi_k, \mathscr{B}_k, P_{\xi_k})$ denote the probability space associated with a random variable ξ_k, with $\Xi_k \subset \mathbb{R}$ and P_{ξ_k} the probability measure of ξ_k. We suppose that these random variables are mutually independent. Therefore, the probability space $(\Xi, \mathscr{B}, P_\xi)$ associated with $\xi = (\xi_1, \ldots, \xi_d)$ has the following product structure: $\Xi = \times_{k=1}^d \Xi_k$, $\mathscr{B} = \otimes_{k=1}^d \mathscr{B}_k$, $P_\xi = \otimes_{k=1}^d P_{\xi_k}$. We denote by $L^2_{P_\xi}(\Xi)$ the Hilbert space of second order random variables defined on $(\Xi, \mathscr{B}, P_\xi)$, which is a tensor Hilbert space with the following tensor structure:

$$L^2_{P_\xi}(\Xi) = L^2_{P_{\xi_1}}(\Xi_1) \otimes \ldots \otimes L^2_{P_{\xi_d}}(\Xi_d)$$

We now introduce approximation spaces $\mathscr{S}^k_{n_k} \subset L^2_{P_{\xi_k}}(\Xi_k)$, such that

$$\mathscr{S}^k_{n_k} = \text{span}\left\{\phi^{(k)}_j\right\}_{j=1}^{n_k} = \left\{ v^{(k)}(y_k) = \sum_{j=1}^{n_k} v^k_j \phi^{(k)}_j(y_k); v^k_j \in \mathbb{R} \right\}$$

with $\{\phi^{(k)}_j\}_{j=1}^{n_k}$ forming an orthonormal basis. An approximation space $\mathscr{S}_n \subset L^2_{P_\xi}(\Xi)$ is then obtained by tensorization

$$\begin{aligned}\mathscr{S}_n &= \mathscr{S}^1_{n_1} \otimes \ldots \otimes \mathscr{S}^d_{n_d} \\ &= \left\{ v = \sum_{i_1=1}^{n_1} \ldots \sum_{i_d=1}^{n_d} v_{i_1,\ldots,i_d} \phi^{(1)}_{i_1} \otimes \ldots \otimes \phi^{(d)}_{i_d} \; ; \; v_{i_1,\ldots,i_d} \in \mathbb{R} \right\}\end{aligned}$$

where $\left(\phi^{(1)}_{i_1} \otimes \ldots \otimes \phi^{(d)}_{i_d}\right)(y_1, \ldots, y_d) = \phi^{(1)}_{i_1}(y_1) \ldots \phi^{(d)}_{i_d}(y_d)$. An element $v \in \mathscr{S}_n$ can be identified with the algebraic tensor $\mathbf{v} = (v_{i_1,\ldots,i_d})$, thus yielding the identification $\mathscr{S}_n \simeq \mathbb{R}^{n_1} \otimes \ldots \otimes \mathbb{R}^{n_d}$. Approximation space $\mathscr{S}_n$ has a dimension $\prod_{k=1}^d n_k$ which grows exponentially with the dimension d, thus making impossible the numerical representation of an element $v \in \mathscr{S}_n$ for high dimensional applications. Approximation subspaces $\mathscr{S}_N \subset \mathscr{S}_n$ are typically constructed by suitable tensorization rules

$$\mathscr{S}_N = \left\{ v = \sum_{i \in I_N} v_{i_1,\ldots,i_d} \phi^{(1)}_{i_1} \otimes \ldots \otimes \phi^{(d)}_{i_d}; v_{i_1,\ldots,i_d} \in \mathbb{R} \right\}$$

where $I_N \subset I_n = \times_{k=1}^d \{1, \ldots, n_k\}$ is an index set which can be chosen a priori. A typical construction consists in taking for $\mathscr{S}^k_{n_k}$ the space of degree p polynomials $\mathbb{P}_p(\Xi_k)$, and for $I_N = \{i \in I_n; \sum_{k=1}^d (i_k - 1) \leq p\}$. Thus, $\mathscr{S}_N$ appears to be the so called polynomial chaos composed of multidimensional polynomials with total degree less than p.

We here suppose that approximation space $\mathscr{S}_n$ is given and sufficiently rich to allow accurate representations of a large class of functions (e.g. by choosing polynomial spaces with very high degree). Then, the aim of the present strategy will be to approximate these representations for high dimensional applications.

2.2 Regression Methods

We here consider the case of a real-valued model output $u : \Xi \to \mathbb{R}$. We denote by $\{y^q\}_{q=1}^Q \subset \Xi$ a set of Q samples of ξ, and by $\{u(y^q)\}_{q=1}^Q \subset \mathbb{R}$ the corresponding model evaluations. We suppose that an approximation space $\mathscr{S}_N = span\{\phi_i\}_{i=1}^N$ is given. Classical least-square regression for the construction of

an approximation $u_N \in \mathscr{S}_N$ then consists in solving the following problem

$$\|u-u_N\|_Q^2 = \min_{v\in\mathscr{S}_N} \|u-v\|_Q^2 \quad \text{with} \quad \|u\|_Q^2 = \sum_{q=1}^{Q} u(y^q)^2 \qquad (1)$$

Remark 1. *For the case of a function $u : \Xi \to \mathscr{V}$ with $\mathscr{V} = \mathbb{R}^n$, we could introduce $\|u\|_Q^2 = \sum_{q=1}^{Q} \|u(y^q)\|_{\mathscr{V}}^2$, with $\|\cdot\|_{\mathscr{V}}$ a norm on $\mathscr{V}$.*

Note that $\|\cdot\|_Q$ only defines a semi-norm on $L^2_{P_\xi}(\Xi)$ but it may define a norm on the finite dimensional subspace $\mathscr{S}_N$ if we have a sufficient number Q of model evaluations. A necessary condition is $Q \geq N$. However, this condition may be unreachable in practice for high dimensional stochastic problems and usual a priori (non adapted) construction of approximation spaces $\mathscr{S}_N$. Moreover, classical regression may yield bad results because of ill-conditioning (solution very sensitive to sampling points). A way to circumvent these issues is to introduce a regularized regression functional

$$\mathscr{J}^\lambda(v) = \|u-v\|_Q^2 + \lambda \mathscr{R}(v) \qquad (2)$$

where λ is a regularization parameter and $\mathscr{R}$ a regularization functional. The regularized regression problem then consists in solving

$$\mathscr{J}^\lambda(u_N^\lambda) = \min_{v\in\mathscr{S}_N} \mathscr{J}^\lambda(v) \qquad (3)$$

Denoting by $\mathbf{v} = (v_1,\ldots,v_N)^T \in \mathbb{R}^N$ the coefficients of an element $v = \sum_{i=1}^N v_i\phi_i \in \mathscr{S}_N$, we can write

$$\|u-v\|_Q^2 = \|\mathbf{z}-\Phi\mathbf{v}\|_2^2 \qquad (4)$$

with $\mathbf{z} = (u(y^1),\ldots,u(y^q))^T \in \mathbb{R}^Q$ the vector of random evaluations of $u(\xi)$ and $\Phi \in \mathbb{R}^{Q\times N}$ the matrix with components $(\Phi)_{q,i} = \phi_i(y^q)$. We can then introduce a function $R : \mathbb{R}^N \to \mathbb{R}$ such that $\mathscr{R}(\sum_i v_i\phi_i) = R(\mathbf{v})$, and a function $J^\lambda : \mathbb{R}^N \to \mathbb{R}$ such that $\mathscr{J}^\lambda(\sum_i v_i\phi_i) = J^\lambda(\mathbf{v}) = \|\mathbf{z}-\Phi\mathbf{v}\|_2^2 + \lambda R(\mathbf{v})$. An algebraic version of regression problem (3) can then be written as follows:

$$\min_{\mathbf{v}\in\mathbb{R}^N} \|\mathbf{z}-\Phi\mathbf{v}\|_2^2 + \lambda R(\mathbf{v}) \qquad (5)$$

Regularization introduces additional information such as smoothness, bounds on norms, sparsity... Under some assumptions on the regularization functional $\mathscr{R}$, problem (3) may have a unique solution. However, the choice of regularization strongly influences the quality of the obtained approximation.

2.3 Sparse Regularization

Over the last decade, methods based on sparse regularization have been rediscovered under the umbrella of compressed sensing that aims at recovering sparse signals from a few linear measurements [5–7]. A sparse function is one that can be represented using few terms when expanded on a suitable basis. In the context of uncertainty quantification, if a stochastic function is known to be sparse on a particular function basis, e.g. polynomial chaos (or tensor basis), sparse regularization methods can be used for quasi optimal recovery with only a few sample evaluations. In general, a successful reconstruction of sparse solution vector depends on *sufficient* sparsity of the coefficient vector and on additional technical properties (e.g. incoherence). This strategy has been found to be effective for non-adapted sparse approximation of PDEs [8,9].

An approximation $u_N(\xi) = \sum_{i=1}^N u_i\phi_i(\xi)$ of a function $u(\xi)$ is considered as sparse on a particular basis $\{\phi_i(\xi)\}_{i=1}^N$ if only a small fraction of coefficients $\mathbf{u} = (u_1,\ldots,u_N)^T$ are significant. Under certain conditions, the significant coefficients can be computed accurately using only $Q \ll N$ random samples of $u(\xi)$ via sparse regularization. Given the random samples $\mathbf{z} \in \mathbb{R}^Q$ of the model output $u(\xi)$, sparse regularization aims at finding the nearly sparsest coefficient $\mathbf{u}$ by solving an optimization problem of the form:

$$\min_{\mathbf{v}\in\mathbb{R}^N} \|\Phi\mathbf{v}-\mathbf{z}\|_2^2 + \lambda\|\mathbf{v}\|_s$$

where $\|\mathbf{v}\|_s$ is a measure of the sparsity of $\mathbf{v}$. Let us briefly explain the construction of such a regression problem. We assume that the solution u is approximately sparse, such that for a given precision δ, the set $\{\mathbf{v}\in\mathbb{R}^N; \|\Phi\mathbf{v}-\mathbf{z}\|_2 \leq \delta\}$ is non empty. The sparsest approximation in this set can be ideally obtained by solving a problem of type

$$\min_{\mathbf{v}\in\mathbb{R}^N} \|\mathbf{v}\|_0 \quad \text{subject to} \quad \|\Phi\mathbf{v}-\mathbf{z}\|_2 \leq \delta \qquad (P_0^\delta)$$

where $\|\mathbf{v}\|_0 = \#\{i \in \{1,\ldots,N\} \,:\, v_i \neq 0\}$ is the number of non zero components of $\mathbf{v}$. In general, this problem is not computationally tractable as it is NP hard to compute. Under certain assumptions, problem (P_0^δ) can be reasonably well approximated by the following minimization problem:

$$\min_{\mathbf{v}\in\mathbb{R}^N} \|\mathbf{v}\|_1 \quad \text{subject to} \quad \|\Phi\mathbf{v}-\mathbf{z}\|_2 \leq \delta \qquad (P_1^\delta)$$

where $\|\mathbf{v}\|_1 = \sum_{i=1}^N |v_i|$ is the 1-norm of $\mathbf{v}$. Since the 1-norm $\|\mathbf{v}\|_1$ is strictly convex, the optimization problem (P_1^δ) admits a unique

solution. An equivalent form of (P_1^δ) can be written

$$\min_{\mathbf{v}\in\mathbb{R}^N} \|\Phi\mathbf{v}-\mathbf{z}\|_2^2 + \lambda\|\mathbf{v}\|_1 \qquad (P_1^\lambda)$$

where $\lambda > 0$ is some regularization parameter which is related to tolerance δ. Several optimization algorithms have been proposed for solving (P_1^λ), such as interior point methods [10].

Of course, in order to successfully recover the coefficients $\mathbf{u}$ of function $u_N \in \mathscr{S}_N$ with sparse regularization, the function u_N must be sparse relatively to the basis which is used. The question is now: how to select a basis in which the approximation u_N is sparse ?

3 TENSOR APPROXIMATIONS BASED ON NON INTRUSIVE REGRESSION

3.1 Canonical Tensor Subsets

We now introduce some basic definitions of tensor subsets of the finite dimensional tensor space $\mathscr{S}_n = \mathscr{S}_{n_1}^1 \otimes \ldots \otimes \mathscr{S}_{n_d}^d$ in which approximations of stochastic functions will be searched. We first introduce the set of elementary tensors $\mathscr{R}_1 \subset \mathscr{S}_n$ (or rank-one tensors) defined by

$$\mathscr{R}_1 = \left\{ v(y) = \left(\otimes_{k=1}^d v^{(k)}\right)(y) = \prod_{k=1}^d v^{(k)}(y_k)\ ;\ v^{(k)} \in \mathscr{S}_{n_k}^k \right\}$$

$\mathscr{R}_1$ is a (nonlinear) submanifold of $\mathscr{S}_n$ with a dimension $(\sum_{k=1}^d n_k - 1)$ which grows only linearly with dimension d.

Remark 2. *Denoting by $\mathbf{v}^{(k)} = (v_j^k)_{j=1}^{n_k} \in \mathbb{R}^{n_k}$ the vector of coefficients of an element $v^{(k)} \in \mathscr{S}_{n_k}^k$, an element $v \in \mathscr{R}_1$ can be identified with an algebraic rank-one element $\mathbf{v} = \otimes_{k=1}^d \mathbf{v}^{(k)}$ in $\mathbb{R}^{n_1} \otimes \ldots \otimes \mathbb{R}^{n_d}$. In the following, we will omit this identification for clarity, although it is crucial for practical implementation.*

We also introduce the set of rank-m (canonical) tensors $\mathscr{R}_m$ defined by

$$\mathscr{R}_m = \left\{ v = \sum_{i=1}^m v_i\ ;\ v_i = \otimes_{k=1}^d v_i^{(k)} \in \mathscr{R}_1 \right\} = \mathscr{R}_{m-1} + \mathscr{R}_1$$

Note that we have the property that $\mathscr{S}_n = span(\mathscr{R}_1)$, such that each element in $\mathscr{S}_n$ can be represented as a sum of elementary tensors. In the following, we will propose algorithms for the construction of approximations in tensor subsets $\mathscr{R}_1$ and $\mathscr{R}_m$, which are low-dimensional subsets of the approximation space $\mathscr{S}_n$, but which are not linear spaces nor convex sets, thus making more difficult the analysis and practical resolution of optimization problems in these sets.

Remark 3. *Other tensor subsets have been introduced which have better approximation properties, such as Tucker tensor sets or Hierarchical tensor sets (see [11]). These tensor formats are not considered here.*

3.2 Updated Greedy Construction of a Canonical Tensor Decomposition

We here present an algorithm for the construction of a rank-m approximation $u_m \in \mathscr{R}_m$ of u of the form

$$u_m = \sum_{i=1}^m \alpha_i w_i, \quad w_i = \otimes_{k=1}^d w_i^{(k)} \in \mathscr{R}_1 \qquad (6)$$

We use an updated greedy procedure which is as follows. We start by setting $u_0 = 0$. Then, knowing an approximation u_{m-1} of u, we proceed as follows.

Correction step. We first compute a correction $w_m \in \mathscr{R}_1$ of u_{m-1} which is based on the following regression problem:

$$w_m \in \arg\min_{w\in\mathscr{R}_1} \|u - u_{m-1} - w\|_Q^2 \qquad (7)$$

Note that the nonlinearity of this regression problem comes from the fact that the set $\mathscr{R}_1$ is not a linear vector space. In practice, minimization problem (7) is solved using an alternating minimization algorithm. Denoting $w = \otimes_{k=1}^d w^{(k)}$, it consists in successively solving regression problems $\min_{w^{(j)}\in\mathscr{S}_{n_j}^j} \|u - u_{m-1} - \otimes_{k=1}^d w^{(k)}\|_Q^2$ for fixed values of functions $\{w^{(k)}\}_{k\neq j}$.

Updating step. The correction step provides an approximation $u_{m-1} + w_m$. Then, the next approximation $u_m = \sum_{i=1}^m \alpha_i w_i$ is computed using regularized regression:

$$u_m = \arg\min_{v\in\mathscr{S}_m} \|u - v\|_Q^2 + \lambda\mathscr{R}(v) \qquad (8)$$

where $\mathscr{S}_m = span\{w_i\}_{i=1}^m$ is the linear space generated by the previously computed elementary tensors. This algorithm can be interpreted as an updated greedy procedure for the construction of a small dimensional linear approximation space $\mathscr{S}_m$ in which an approximation u_m is computed using a suitable regularized regression technique. Regularizations which are adapted to the present context are discussed in the following section.

Remark 4. *A pure greedy construction would consist in letting $u_m = u_{m-1} + w_m$. This pure greedy construction yields to a rather bad behavior of the sequence of approximations $\{u_m\}_{m\geq 1}$. A*

remedy could be to introduce a regularized version of the correction

$$w_m \in \arg\min_{w \in \mathscr{R}_1} \|u - u_{m-1} - w\|_Q^2 + \lambda \mathscr{R}(u_{m-1} + w)$$

If the functional $\mathscr{R}$ is chosen as a strictly convex function, then it can be proved (following [12]), that the sequence $u_m = u_{m-1} + w_m$ converges towards the unique solution u_n^λ of the regularized regression problem $\min_{v \in \mathscr{S}_n} \mathscr{J}^\lambda(v)$, with $\mathscr{J}^\lambda$ defined in (2)*. However, the difficulty is to find a pertinent regularization functional $\mathscr{R}$ yielding a good approximation u^λ and such that it is compatible with the tensor framework.*

3.3 Regularized Regression in Reduced Bases of Elementary Tensors

Here, we detail the definition of the updating step (8) for computing the approximation $u_m = \sum_{i=1}^m \alpha_i w_i$. Denoting by $\mathbf{z} \in \mathbb{R}^Q$ the vector of samples of $u(\xi)$, by $\alpha = (\alpha_1, \cdots, \alpha_m)^T$ the vector of coefficients and by $W \in \mathbb{R}^{Q \times m}$ the matrix whose components ar $(W)_{q,i} = w_i(y^q)$, problem (8) can be reformulated as the following optimization problem:

$$\min_{\alpha \in \mathbb{R}^m} \|W\alpha - \mathbf{z}\|_2^2 + \lambda R(\alpha);$$

In fact, the set of functions $\{w_i\}_{i=1}^m$ of u_m can be interpreted as the first m elements of a stochastic basis in which the solution u is approximately sparse.

Choice of the Regularization. A first natural choice consists in taking

$$R(\alpha) = \|\alpha\|_2^2$$

As illustrated in numerical examples, we will observe that this choice of ℓ_2 regularization yields a deterioration with m of the approximation u_m.
Depending on the way of generating the elements, the vector of coefficients $\alpha \in \mathbb{R}^m$ may also be searched as a sparse vector. It is usually the case when we use sub-optimal greedy constructions, yielding to high values of rank m for reaching a given precision, much higher than optimal rank representation of the solution. Therefore, we propose to introduce sparse ℓ_1-regularization in the definition of the updating step, by choosing

$$R(\alpha) = \|\alpha\|_1$$

In practice, it is observed that this sparse regularization allows a pertinent selection of significant terms in the canonical decomposition and allows to avoid a deterioration with m of the sequence u_m. Note that when some α_i are found to be negligible, it yields an approximation $u_m = \sum_{i=1}^m \alpha_i w_i \in \mathscr{R}_m$ with a lower effective rank representation.
The influence of the choice of regression functional in the updating step will be analyzed on numerical examples.

Selection of Regularization Parameter. An optimal parameter λ can be selected using suitable error criteria, e.g. based on cross validation.

Remark 5. *Let us note that the proposed algorithm not only generates an approximation of the model output but a sequence of model approximations $\{u_m\}_{m \geq 1}$. Therefore, error criteria based on cross validation can also be used in order to select the best model among these generated models.*

4 APPLICATION EXAMPLES

In this section, we report the performance of the proposed tensor-based regression method on two high dimensional stochastic partial differential equations.

4.1 Example 1: Diffusion equation with multiple inclusions

We consider a stationary diffusion problem defined on a two dimensional domain $\Omega = (0.1) \times (0.1)$ (see Fig. 1):

$$\begin{aligned} -\nabla \cdot (\kappa \nabla u) &= I_D(x) \text{ on } \Omega \\ u &= 0 \text{ on } \partial\Omega \end{aligned} \tag{9}$$

where $D \subset \Omega$ is a square domain in the middle and I_D is the indicator function of D. The diffusion coefficient is defined by

$$\kappa = \begin{cases} \xi_k & \text{on } C_k,\ 1 \leq k \leq 8 \\ 1 & \text{on } \Omega \backslash (\cup_{k=1}^8 C_k) \end{cases}$$

where the C_k, $1 \leq k \leq 8$, are circular domains (see Fig. 1) and where the $\xi_k \in U(0.9, 1.1)$ are independent uniform random variables. We define the quantity of interest

$$I(u)(\xi) \ = \ \int_D u(x, \xi) dx$$

We introduce approximation spaces $\mathscr{S}_{p+1}^k = \mathbb{P}_p(0.9, 1.1)$, with polynomial degree $p = 10$. An accurate approximation of the reference solution is evaluated by the Proper Generalized Decomposition method(see [13]).

We obtain the tensor basis by greedy procedure and update the coefficients α after each successive rank one correction. In

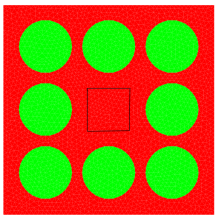

FIGURE 1. Diffusion problem with multiple inclusions.

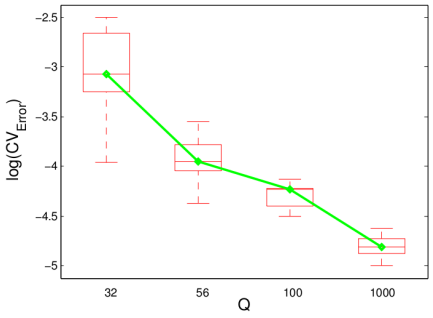

FIGURE 2. Cross validation obtained by ℓ_1 regularized update for different sample sizes.

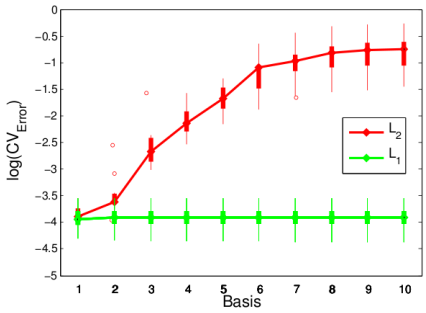

FIGURE 3. Cross validation error of ℓ_1 and ℓ_2 regularized update for sample size $Q = 56$

order to derive reliable conclusions, we compare the performance of ℓ_2 and ℓ_1 regularization by performing a sample independence study. We take 11 sample sets of size $Q = \{32, 56, 100, 1000\}$ and plot median value of the quantities together with quartiles and outliers, if any. As shown in Fig. 2, when using ℓ_1 regularization, the cross validation error reduces with sample size Q, as long as the influence of the sample. In other words, when using ℓ_1 regularization, the obtained tensor approximations seems to converge with Q towards a deterministic approximation. Fig. 3 and Fig. 4 show cross validation error v/s number of tensor basis functions for $Q = 56$ and 1000 samples respectively. We note that for few samples i.e. $Q = 56$, ℓ_2 regularization deteriorates for high rank approximation whereas ℓ_1 regularization yields a

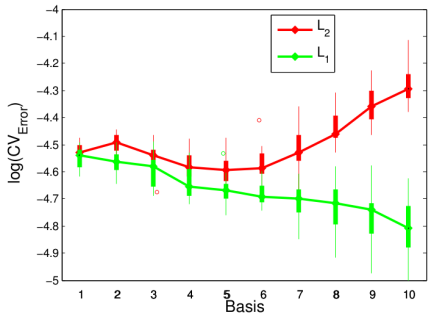

FIGURE 4. Cross validation error of ℓ_1 and ℓ_2 regularized update for sample size $Q = 1000$

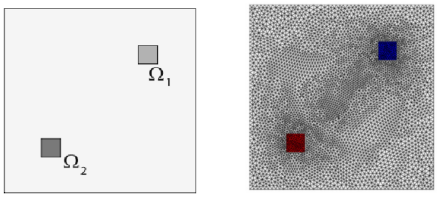

FIGURE 5. Domain and finite element mesh.

stabilization of the cross validation error. For $Q = 1000$, the approximation obtained using ℓ_1 regularization is convergent for high dimensional tensor basis and gives better approximation of the solution. It can also be noted that for Q as low as 56, we obtain second order accurate solutions for almost all sample sets. We therefore draw the following conclusions:

- The number of model evaluations sufficient to obtain a very accurate approximation is very small compared to the dimension of the underlying approximation space $N = dim(\mathscr{S}_n) = (p+1)^8 = 11^8$ but also to the dimension of the full polynomial chaos with total degree less than or equal to p (i.e $N = \frac{(8+p)!}{p!8!} = 43758$).
- ℓ_1 regularization is able to recover sparse solution vector α on stochastic tensor basis and hence is an effective update strategy.
- Tensor approximations with update strategy based on ℓ_1 regularization appears to converge with respect to the number of samples towards a deterministic approximation. In other words, beyond a sample size threshold, the obtained tensor approximations are nearly sample independent.

4.2 Stationary Advection Diffusion Reaction Equation with Random Field

In this example, we consider a stationary advection diffusion reaction equation on a spatial domain $\Omega = (0,1)^2$ (Fig. 5) where the source of uncertainty comes from the diffusion coefficient

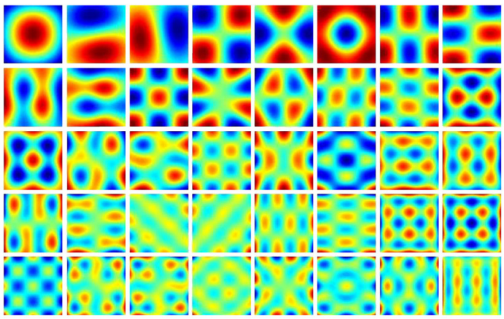

FIGURE 6. Spatial modes $\{\mu_k(x)\}_{k=1}^{40}$ of the decomposition of random field $\mu(x,\xi)$.

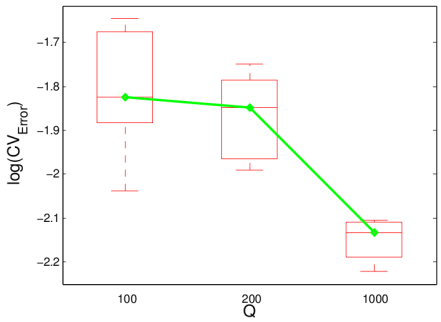

FIGURE 7. Cross validation obtained by ℓ_1 regularized update for different sample sizes.

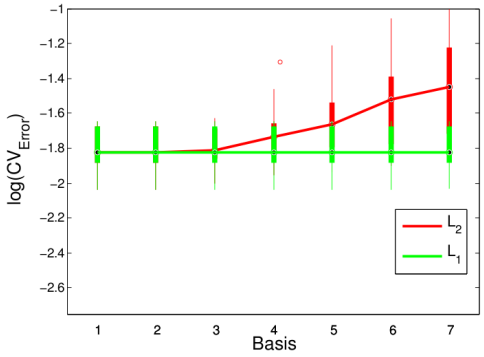

FIGURE 8. Cross validation error of ℓ_1 and ℓ_2 regularized update for sample size $Q = 100$

which is a random field. The problem is:

$$-\nabla\cdot(\mu(x,\xi)\nabla u)+c\cdot\nabla u+\kappa u = f \text{ on } \Omega$$
$$u = 0 \text{ on } \partial\Omega$$

where $\kappa = 10$ is a deterministic reaction coefficient and $c = 250(x-\frac{1}{2},\frac{1}{2}-y)$ is a deterministic advection velocity. The source term is deterministic and is defined by $f = 100I_{\Omega_1}$, where $\Omega_1 = (0.7,0.8)\times(0.7,0.8)\subset\Omega$ and where I_{Ω_1} is the indicator function of Ω_1. $\mu(x,\xi)$ is a random field defined by

$$\mu(x,\xi) = \mu_0 + \sum_{k=1}^{100}\sqrt{\sigma_k}\mu_k(x)\xi_k \tag{10}$$

where $\mu_0 = 1$ is the mean value of μ, where the $\xi_k \in U(-1,1)$ are mutually independent uniform random variables and where the μ_k are a set of $L^2(\Omega)$-orthonormal spatial functions. The couples $(\mu_k,\sigma_k)\in L^2(\Omega)\times\mathbb{R}^+$ are chosen as the 100 dominant eigenpairs of eigenproblem $T(\mu_k) = \sigma_k\mu_k$, where T is the kernel operator

$$T : v\in L^2(\Omega)\mapsto\int_\Omega \alpha(x,y)v(y)dy \in L^2(\Omega)$$

with $\alpha(x,y) = 0.2^2 exp(-\frac{\|x-y\|^2}{l_c^2})$ with l_c the correlation length. The equation (10) then corresponds to a truncated version of a homogeneous random field with mean 1, standard deviation $\frac{0.2}{\sqrt{3}}$ and exponential square covariance function with correlation length l_c. The first 40 spatial functions are plotted in Fig. 6. The $d = 100$ random parameters $\xi = (\xi_k)_{k=1}^d$ define a probability space $(\Xi,\mathscr{B},P_\xi)$, with $\Xi = (-1,1)^d$ and P_ξ the uniform probability measure. We introduce approximation spaces $\mathscr{S}_{p+1}^k = \mathbb{P}_p(-1,1)$ which are spaces of polynomials with degree $p = 3$.

We compare the performance of ℓ_2 and ℓ_1 regularizations by performing a sample independence study. We took 11 sample sets of size $Q = \{100,200,1000\}$ and plot median value of the quantities together with quartiles and outliers, if any. Fig. 7 shows cross validation error v/s tensor basis dimension for $Q = 100$ samples. We note that for this sample size ℓ_1 regularization keeps the tensor basis with minimum cross validation error. However, in this example, rank one approximation is very accurate and we observe no improvement in solution even for $Q = 1000$.

From this example, several conclusions can be drawn:

- The proposed regression technique is very effective in very high dimensional stochastic problems. The number of model evaluations Q required to obtain very accurate solution is very small. This is orders of magnitude less than a classical Polynomial Chaos approximation (which would require $N = (p+d)!/(p!d!) = 176851$ model evaluations, when $p = 3$). Note that the efficiency of the proposed tensor approximation methods on this particular example is due to the effective low rank of the solution.
- The cross validation error reduces with the sample size Q.

Tensor approximations seem to converge with Q towards a deterministic tensor approximation.

- ℓ_1 regularization keeps the tensor approximation with minimum cross validation error.

5 CONCLUSION

A non-intrusive regression technique based on tensor product approximation has been proposed for propagation of uncertainty in high dimensional stochastic problems. It involves formulating a minimization problem in stochastic tensor product space and using tensor product approximation strategies to build a sequence of approximations with increasing rank. The rank-one tensors obtained by successive corrections can then be chosen as reduced bases on which coefficients can be updated by ℓ_1 regularization such that a few significant terms are retained in the final solution. Cross validation model selection technique has been used to evaluate the best approximation of the quantity of interest among the different generated approximations. The ability of the proposed method to handle high dimensional uncertainty quantification problem was illustrated on two stochastic partial differential equations and first results are quite promising. Future work will be dedicated to evaluate the capabilities of this method in approximating stochastic functions with discontinuities. Other updating strategies based on sparse regularization and hybrid methods will also be studied for better exploiting the generated information.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the support of EADS Innovation Works and of the French National Research Agency (Grant ANR-2010-BLAN-0904).

REFERENCES

[1] Nouy, A., 2009. "Recent developments in spectral stochastic methods for the numerical solution of stochastic partial differential equations". *Archives of Computational Methods in Engineering,* ***16***(3), pp. 251–285.

[2] Xiu, D., 2009. "Fast numerical methods for stochastic computations: a review". *Comm. Comput. Phys.,* ***5***, pp. 242–272.

[3] Kolda, T. G., and Bader, B. W., 2009. "Tensor decompositions and applications". *SIAM Review,* ***51***(3), September, pp. 455–500.

[4] Beylkin, G., Garcke, B., and Mohlenkamp, M., 2011. "Multivariate regression and machine learning with sums of separable functions". *Journal of Computational Physics,* ***230***, pp. 2345–2367.

[5] Candes, E., Romberg, J., and Tao, T., 2006. "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information". *IEEE Trans. Info. Theory,* ***52(2)***(2), pp. 489–509.

[6] Candes, EJ. Romberg, J. T. T., 2006. "Near optimal signal recovery from random projections: Universal encoding strategies?". *IEEE Transactions on information theory,* ***52(12)***, pp. 5406–5425.

[7] Donoho, D., 2006. "Compressed Sensing". *IEEE Transactions on information theory,* ***52(4)***, pp. 1289–1306.

[8] Blatman, G., and Sudret, B., 2011. "Adaptive sparse polynomial chaos expansion based least angle regression". *Journal of Computational Physics,* ***230***, pp. 2345–2367.

[9] Doostan, A., and Owhadi, H., 2011. "A non-adapted sparse approximation of pdes with stochastic inputs". *Journal of Computational Physics,* ***230***(8), pp. 3015–3034.

[10] Koh, K., Kim, S.-J., and Boyd, S., 2007. "An interior-point method for large-scale l1-regularized logistic regression". *J. Mach. Learn. Res.,* ***8***, Dec., pp. 1519–1555.

[11] Hackbusch, W., 2012. *Tensor Spaces and Numerical Tensor Calculus*, Vol. 42 of *Series in Computational Mathematics*. Springer.

[12] Falco, A., and Nouy, A., 2011. "A Proper Generalized Decomposition for the solution of elliptic problems in abstract form by using a functional Eckart-Young approach". *Journal of Mathematical Analysis and Applications,* ***376***(2), pp. 469–480.

[13] Nouy, A., 2010. "Proper Generalized Decompositions and separated representations for the numerical solution of high dimensional stochastic problems". *Archives of Computational Methods in Engineering,* ***17***(4), pp. 403–434.