Proper Generalized Decompositions and separated representations for the numerical solution of high dimensional stochastic problems

Anthony Nouy

Archives of Computational Methods in Engineering, 2010 (In Press)

Proper Generalized Decompositions and separated representations for the numerical solution of high dimensional stochastic problems

Anthony Nouy

Received: date / Accepted: date

Abstract Uncertainty quantification and propagation in physical systems appear as a critical path for the improvement of the prediction of their response. Galerkin-type spectral stochastic methods provide a general framework for the numerical simulation of physical models driven by stochastic partial differential equations. The response is searched in a tensor product space, which is the product of deterministic and stochastic approximation spaces. The computation of the approximate solution requires the solution of a very high dimensional problem, whose calculation costs are generally prohibitive. Recently, a model reduction technique, named Generalized Spectral Decomposition method, has been proposed in order to reduce these costs. This method belongs to the family of Proper Generalized Decomposition methods. It takes part of the tensor product structure of the solution function space and allows the *a priori* construction of a quasi optimal separated representation of the solution, which has quite the same convergence properties as *a posteriori* Hilbert Karhunen-Loève decompositions. The associated algorithms only require the solution of a few deterministic problems and a few stochastic problems on deterministic reduced basis (algebraic stochastic equations), these problems being uncoupled. However, this method does not circumvent the "curse of dimensionality" which is associated with the dramatic increase in the dimension of stochastic approximation spaces, when dealing with high stochastic dimension. In this paper, we propose a mariage between the Generalized Spectral Decomposition algorithms and a separated representation methodology, which exploits the tensor product structure of stochastic functions spaces. An efficient algorithm is proposed for the *a priori* construction of separated representations of square integrable vector-valued functions defined on a high-dimensional probability space, which are the solutions of systems of stochastic algebraic equations.

Keywords Stochastic partial differential equation · Spectral stochastic methods · Curse of dimensionality · Model reduction · Generalized Spectral Decomposi-

This work is supported by the French National Research Agency (grant ANR-06-JCJC-0064) and by GdR MoMaS with partners ANDRA, BRGM, CEA, CNRS, EDF, IRSN.

A. Nouy
GeM - Research Institute in Civil Engineering and Mechanics, UMR CNRS 6183, Ecole Centrale Nantes, University of Nantes
E-mail: anthony.nouy@ec-nantes.fr

tion · Proper Generalized Decomposition · Separation of variables · Finite sums decomposition

1 Introduction

The numerical prediction of the impact of uncertainties on the response of physical models appears as a crucial issue in many branches of science and engineering. These last two decades, spectral stochastic methods have been extensively investigated for the propagation of uncertainties through physical models driven by finite dimensional noise (see *e.g.* [18,50,38,34] and the references therein). These methods rely on a representation of the response as a function of basic random variables modeling the input uncertainties. An approximation of the response is sought on suitable approximation basis. Several methods have been proposed for the definition and computation of the approximate solution: L^2 projection [20,22], interpolation [2,15,52,51,49], regression [6] or Galerkin projections [19,3,35,17].

Galerkin spectral stochastic methods inherit from nice mathematical results in functional analysis. They lead to accurate predictions and allow for a better control on numerical simulations through *a posteriori* error estimation and adaptive approximation [23,46,45,33,48]. However, the computation of the approximate solution requires the solution of a very high dimensional problem, which is generally prohibitive with traditional techniques. Moreover, it requires a good knowledge of the mathematical structure of the physical model in order to extend classical deterministic solvers to the stochastic framework (preconditioners, non linear solvers, ...).

In order to circumvent the above mentioned drawbacks of Galerkin spectral stochastic methods, an *a priori* model reduction technique, named *Generalized Spectral Decomposition* (GSD) method, has been recently proposed for solving stochastic partial differential equations (SPDEs) [36–38,42]. This method, which takes part of the tensor product structure of the solution function space, allows the *a priori* computation of a quasi optimal separated representation of the solution, which has quite the same convergence properties as classical spectral decompositions (*i.e.* Hilbert Karhunen-Loève decompositions). A decomposition of the solution is sought in the form

$$u(\boldsymbol{x},\boldsymbol{\xi}) \approx \sum_{i=1}^{M} w_i(\boldsymbol{x})\lambda_i(\boldsymbol{\xi}), \tag{1}$$

where the $w_i(\boldsymbol{x})$ are deterministic functions of the physical variables $\boldsymbol{x}$ (*e.g.* space and/or time) and where the $\lambda_i(\boldsymbol{\xi})$ are functions of the basic random variables $\boldsymbol{\xi}$. The basic principle of the GSD method consists in defining optimal reduced basis from a double orthogonality criterium. Reduced basis functions then appear as the solutions of a pseudo eigenproblem whose dominant eigenspace is associated with the desired optimal reduced basis. Dedicated algorithms, inspired from classical algorithms for solving eigenproblems, have been proposed for the approximation of the optimal decomposition [37]. The main advantage of these algorithms is that they only ask for the solution of a few uncoupled deterministic problems for computing functions w_i and stochastic algebraic equations for computing stochastic functions λ_i. Stochastic algebraic equations can be solved with classical spectral stochastic methods, leading to an approximation of random variables $\lambda_i(\boldsymbol{\xi}) \approx \sum_{\alpha=1}^{P} \lambda_{i,\alpha} H_\alpha(\boldsymbol{\xi})$, where the $H_\alpha(\boldsymbol{\xi})$ form a basis of classical stochastic approximation spaces, such as polynomial or piecewise

polynomial spaces [11,53,43,31,47]. Deterministic problems being uncoupled, classical deterministic solution techniques can be used. It then makes the GSD method a partially non-intrusive Galerkin spectral stochastic approach.

The separation of deterministic problems and stochastic algebraic equations leads to drastic computational savings, especially for large scale applications. However, this deterministic/stochastic separation does not circumvent the "*curse of dimensionality*" which is associated with the dramatic increase in the dimension P of stochastic approximation spaces, when dealing with a high stochastic dimension, *i.e.* with a large number of random variables $\boldsymbol{\xi} = (\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_r)$. In this paper, we propose a mariage between GSD algorithms and a separated variables representation technique which exploits the tensor product structure of stochastic functions space. The separation of variables is used for the approximate representation of square-integrable vector-valued functions $\Lambda(\boldsymbol{\xi})$ (or second order random vectors) defined on a high-dimensional probability space

$$\Lambda(\boldsymbol{\xi}) = \Lambda(\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_r) \approx \sum_{i=1}^{Z} \phi_i^0 \phi_i^1(\boldsymbol{\xi}_1) \ldots \phi_i^r(\boldsymbol{\xi}_r) \tag{2}$$

where the $\phi_i^j(\boldsymbol{\xi}_j)$ are real valued functions of basic random variables $\boldsymbol{\xi}_j$. A representation (2) of order Z appears as a classical spectral stochastic expansion of a random variable $\Lambda(\boldsymbol{\xi})$ on an Z-dimensional approximation basis $\{\Psi_i(\boldsymbol{\xi})\}_{i=1}^{Z}$, with $\Psi_i(\boldsymbol{\xi}) = \prod_{i=1}^{r} \phi_i^r(\boldsymbol{\xi}_r)$, which is not selected *a priori* but chosen such that it gives a quasi optimal approximation for a given dimension Z. A natural extension of the GSD method is proposed for the *a priori* construction of separated representation (2). The algorithm proposed in this paper, which can be applied to many problems defined in tensor product spaces, yield rather good convergence properties with respect to the order Z of the decomposition.

The overall methodology proposed in this paper allows computing an approximate solution of the model in very high dimensional approximation spaces (10^{20}, 10^{50}, ...), with algorithms having a complexity which is (quasi)linear with the stochastic dimension r. It then allows to deal with problems which are unaffordable with conventional spectral stochastic approaches and usually require the use of classical Monte-Carlo simulations.

Let us note that the overall methodology and algorithms could be naturally applied to the solution of other types of problems defined in tensor product spaces. Some variants of this methodology have been proposed for the *a priori* construction of such separated representations of functions in tensor product spaces [27,28,41,5,1,21,29, 39]. In the context of spectral stochastic methods, a basic methodology has already been proposed in [14,13]. This kind of methodologies is receiving a growing interest in many applications where numerical simulations suffer from the curse of dimensionality. The obtained decompositions have been recently called *Proper Generalized Decompositions* (PGD). PGD methods can be seen as a family of methods for the *a priori* construction of separated representations of functions which are solutions of problems defined in tensor product spaces (GSD method belongs to this family). For some variants of algorithms and some very particular frameworks, some mathematical results are available [7,16]. However, the mathematical bases of these methods are still badly mastered. Further mathematical investigations will be necessary in order to better understand this type of decomposition in a general framework and to propose more efficient algorithms. Nevertheless, as it will be illustrated in this paper, these types of

algorithms are already of great practical interest.

The outline of the paper is as follows. In section 2, we briefly recall the principle of classical stochastic spectral approaches for solving stochastic partial differential equations. In section 3, we recall the basics of the GSD method and related algorithms for the construction of decomposition (1). In section 4, we introduce a methodology for the solution of stochastic algebraic equations defined on high dimensional product probability spaces, which is based on the *a priori* construction of decomposition (2). The proposed method belongs to the family of *Proper Generalized Decomposition* (PGD) methods. Sections 5 and 6 will illustrate the overall methodology (coupling GSD algorithms and PGD in high dimension) for model stochastic partial differential equations, namely stochastic advection diffusion reaction equations.

2 Stochastic partial differential equations and Galerkin spectral stochastic methods

2.1 Weak formulation of stochastic partial differential equations

We consider a stochastic partial differential equation (SPDE) defined on a physical domain (*e.g.* space or space-time domain) whose operator and right-hand side depend on a finite set of m real valued random variables $\boldsymbol{\xi} = (\xi_1, \dots, \xi_m)$. We introduce the associated finite-dimensional probability space $(\boldsymbol{\Xi}, \mathcal{B}, P_{\boldsymbol{\xi}})$, where $\boldsymbol{\Xi} \subset \mathbb{R}^m$ is the set of elementary events, $\mathcal{B}$ is a σ-algebra on $\boldsymbol{\Xi}$ and $P_{\boldsymbol{\xi}}$ is the probability measure. We consider that the solution u of the SPDE is a random variable with values in a Hilbert space $\mathcal{V}$ of functions defined on the physical domain. A strong-stochastic formulation of the SPDE writes: find $u : \boldsymbol{\Xi} \to \mathcal{V}$ such that we have $P_{\boldsymbol{\xi}}$ almost surely

$$u(\boldsymbol{\xi}) \in \mathcal{V}, \quad a(u(\boldsymbol{\xi}), v; \boldsymbol{\xi}) = b(v; \boldsymbol{\xi}) \quad \forall v \in \mathcal{V}, \tag{3}$$

where a and b and bilinear[1] and linear forms on $\mathcal{V}$. We consider the particular class of SPDEs whose solution u is a second order random variable with values in $\mathcal{V}$, which is supposed to be independent on the random event $\boldsymbol{\xi}$[2]. The solution then belongs to Hilbert space $L^2(\boldsymbol{\Xi}, \mathcal{B}, P_{\boldsymbol{\xi}}; \mathcal{V})$, which can be identified with the tensor product space $\mathcal{V} \otimes \mathcal{S}$, where $\mathcal{S} := L^2(\boldsymbol{\Xi}, \mathcal{B}, P_{\boldsymbol{\xi}})$ denotes the space of real valued second order random variables defined on $(\boldsymbol{\Xi}, \mathcal{B}, P_{\boldsymbol{\xi}})$ (or equivalently the space of real-valued functions defined on $\boldsymbol{\Xi}$ which are $\mathcal{B}$-measurable and square integrable). A weak-stochastic formulation of (3) writes:

$$u \in \mathcal{V} \otimes \mathcal{S}, \quad A(u, v) = B(v) \quad \forall v \in \mathcal{V} \otimes \mathcal{S}, \tag{4}$$

where bilinear form A and linear form B are defined by

$$A(u, v) := E\big(a(u(\boldsymbol{\xi}), v(\boldsymbol{\xi}); \boldsymbol{\xi})\big), \tag{5}$$

$$B(v) := E\big(b(v(\boldsymbol{\xi}); \boldsymbol{\xi})\big), \tag{6}$$

[1] In this article, we only consider the case of linear SPDEs. Problem (3) can be associated with a linear physical model but also with one step of a nonlinear iterative strategy for solving a nonlinear SPDE.

[2] For SPDEs defined on random domains, a suitable reformulation of the problem on a deterministic domain allows to work in a deterministic function space $\mathcal{V}$ [54, 8, 40].

where E is the mathematical expectation defined by

$$E(f(\boldsymbol{\xi})) = \int_{\boldsymbol{\Xi}} f(\mathbf{y}) dP_{\boldsymbol{\xi}}(\mathbf{y}). \tag{7}$$

2.2 Product structure of stochastic function space

We suppose that the set of m random variables $\boldsymbol{\xi}$ can be split into r mutually independent sets of random variables $\{\boldsymbol{\xi}_i\}_{i=1}^r$, *i.e.* $\boldsymbol{\xi} = \{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_r\}$. Let $(\boldsymbol{\Xi}_i, \boldsymbol{\mathcal{B}}_i, P_{\boldsymbol{\xi}_i})$, with $\boldsymbol{\Xi}_i \subset \mathbb{R}^{m_i}$, denote the probability space associated with the set of random variables $\boldsymbol{\xi}_i$, with $m = \sum_{i=1}^r m_i$. The probability space $(\boldsymbol{\Xi}, \boldsymbol{\mathcal{B}}, P_{\boldsymbol{\xi}})$ have a product structure:

$$\boldsymbol{\Xi} = \times_{i=1}^r \boldsymbol{\Xi}_i, \quad \boldsymbol{\mathcal{B}} = \otimes_{i=1}^r \boldsymbol{\mathcal{B}}_i, \quad P_{\boldsymbol{\xi}} = \otimes_{i=1}^r P_{\boldsymbol{\xi}_i} \tag{8}$$

Hilbert space $\mathcal{S} = L^2(\boldsymbol{\Xi}, \boldsymbol{\mathcal{B}}, P_{\boldsymbol{\xi}})$ then have the following tensor product structure:

$$\mathcal{S} \simeq \mathcal{S}^1 \otimes \ldots \otimes \mathcal{S}^r, \quad \mathcal{S}^i := L^2(\boldsymbol{\Xi}_i, \boldsymbol{\mathcal{B}}_i, P_{\boldsymbol{\xi}_i}) \tag{9}$$

If the m_i random variables $\boldsymbol{\xi}_i = (\xi_{i,1}, \ldots, \xi_{i,m_i})$ are mutually independent, probability space $(\boldsymbol{\Xi}_i, \boldsymbol{\mathcal{B}}_i, P_{\boldsymbol{\xi}_i})$ has itself a product structure: $\boldsymbol{\Xi}_i = \times_{j=1}^{m_i} \Xi_{i,j}$, $\boldsymbol{\mathcal{B}}_i = \otimes_{j=1}^{m_i} \mathcal{B}_{i,j}$, $P_{\boldsymbol{\xi}_i} = \otimes_{j=1}^{m_i} P_{\xi_{i,j}}$. Therefore, Hilbert space $\mathcal{S}^i$ has the following tensor product structure: $\mathcal{S}^i = \mathcal{S}^{i,1} \otimes \ldots \otimes \mathcal{S}^{i,m_i}$, with $\mathcal{S}^{i,j} = L^2(\Xi_{i,j}, \mathcal{B}_{i,j}, P_{\xi_{i,j}})$.

2.3 Stochastic approximation spaces

Approximation spaces in Hilbert space $\mathcal{S} = L^2(\boldsymbol{\Xi}, \boldsymbol{\mathcal{B}}, P_{\boldsymbol{\xi}})$ can naturally be built by tensorization of approximation spaces in $\mathcal{S}^i = L^2(\boldsymbol{\Xi}_i, \boldsymbol{\mathcal{B}}_i, P_{\boldsymbol{\xi}_i})$. Let $\mathcal{S}^i_{P_i}$ denote a P_i-dimensional approximation space in $\mathcal{S}^i$. A full tensorization leads to a P-dimensional approximation space $\mathcal{S}_P \subset \mathcal{S}$ defined by

$$\mathcal{S}_P = \mathcal{S}^1_{P_1} \otimes \ldots \otimes \mathcal{S}^r_{P_r}, \quad P = \prod_{i=1}^r P_i \tag{10}$$

Let $\{h^i_{\alpha_i}(\boldsymbol{\xi}_i)\}_{\alpha_i=1}^{P_i}$ denote a basis of $\mathcal{S}^i_{P_i}$ and let $\mathcal{I}_P = \{\alpha = (\alpha_j)_{j=1}^r; \alpha_j \in \{1, \ldots, P_j\}\}$ denote a set of multi-indices. A basis $\{H_\alpha(\boldsymbol{\xi})\}_{\alpha \in \mathcal{I}_P}$ of $\mathcal{S}_P$ is then simply obtained by letting $H_\alpha(\boldsymbol{\xi}) = \prod_{i=1}^r h^i_{\alpha_i}(\boldsymbol{\xi}_i)$. For simplicity, we introduce a one-to-one mapping between the set of multi-indices $\mathcal{I}_P$ and $\{1, \ldots, P\}$ and equivalently denote $\{H_\alpha\}_{\alpha=1}^P$ the basis of $\mathcal{S}_P$.

The reader can refer to [43] for a general methodology for the construction of approximation spaces $\mathcal{S}^i_{P_i}$ in the case of arbitrary probability measures $P_{\boldsymbol{\xi}_i}$. For the case where $\boldsymbol{\xi}_i$ is composed by m_i independent random variables, classical choices consist in introducing orthogonal complete polynomial basis [19, 53] (classical polynomial chaos basis), or piecewise polynomial basis [11, 46, 30]. These constructions are classical and will not be detailed in this paper (see *e.g.* [38]).

2.4 Galerkin spectral stochastic approximation

Galerkin stochastic approaches consist in defining an approximate solution of problem (4) by

$$u \in \mathcal{V} \otimes \mathcal{S}_P, \quad A(u,v) = B(v) \quad \forall v \in \mathcal{V} \otimes \mathcal{S}_P, \tag{11}$$

where $\mathcal{S}_P \subset \mathcal{S}$ is a P-dimensional approximation space. Let $\{H_\alpha\}_{\alpha=1}^P$ denote a basis of $\mathcal{S}_P$. Equation (11) can be interpreted as a system of P coupled SPDEs: find $\{u_\alpha\}_{\alpha=1}^P \in (\mathcal{V})^P$ such that $\forall \beta \in \{1,\dots,P\}$, $\forall v_\beta \in \mathcal{V}$,

$$\sum_{\alpha=1}^{P} E\big(a\left(u_\alpha, v_\beta; \boldsymbol{\xi}\right) H_\alpha(\boldsymbol{\xi}) H_\beta(\boldsymbol{\xi})\big) = E\big(b(v_\beta;\boldsymbol{\xi}) H_\beta(\boldsymbol{\xi})\big)$$

3 Generalized spectral decomposition method

In this section, we recall the basics of the Generalized Spectral Decomposition method (GSD) [36,37,42], which is a method for the *a priori* construction of a separated representation of the solution u of (4):

$$u \approx u_M = \sum_{i=1}^{M} w_i \lambda_i, \quad w_i \in \mathcal{V}, \quad \lambda_i \in \mathcal{S} \tag{12}$$

where neither the functions w_i nor the functions λ_i are fixed *a priori*. Decomposition (12) is called a separated representation of order M. Functions w_i and λ_i are said to be optimal reduced basis functions with respect to a given metric if the order M is minimal for a given accuracy, measured with this particular metric. The GSD method provides a methodology and dedicated algorithms for the *a priori* definition and construction of a decomposition of type (12). In the context of spectral stochastic methods, it can be seen as a method for the *a priori* construction of a very low dimensional stochastic approximation space $\mathcal{S}_M := span(\{\lambda_i\})_{i=1}^M \subset \mathcal{S}$.

Remark 1 - Here, we use a terminology associated with stochastic problems although the method could be applied to the approximate solution of a large class of problems (4) *defined in a tensor product space* $\mathcal{V} \otimes \mathcal{S}$.

3.1 A posteriori separated representation: classical spectral decomposition

When the solution u is known, an optimal separated representation u_M can be naturally defined by introducing an inner product $\ll \cdot,\cdot \gg_{\mathcal{V}\otimes\mathcal{S}}$ on tensor product space $\mathcal{V} \otimes \mathcal{S}$, this inner product being built from inner products $< \cdot,\cdot >_{\mathcal{V}}$ and $< \cdot,\cdot >_{\mathcal{S}}$ on Hilbert spaces $\mathcal{V}$ and $\mathcal{S}$, *i.e.* such that $\forall \lambda, \lambda^* \in \mathcal{S}$ and $\forall w, w^* \in \mathcal{V}$

$$\ll \lambda w, \lambda^* w^* \gg_{\mathcal{V}\otimes\mathcal{S}} = < w, w^* >_{\mathcal{V}} < \lambda, \lambda^* >_{\mathcal{S}}$$

The optimal order M separated representation u_M is then defined as the one which minimizes $\|u - u_M\|_{\mathcal{V}\otimes\mathcal{S}}$, where $\|\cdot\|_{\mathcal{V}\otimes\mathcal{S}}$ is the norm associated with $\ll \cdot,\cdot \gg_{\mathcal{V}\otimes\mathcal{S}}$. It turns out that this optimal decomposition corresponds to the Hilbert Karhunen-Loève

decomposition, where functions $\{w_i\}_{i=1}^M$ span the M-dimensional dominant eigenspace of the following eigenproblem:

$$T_u(w) = \sigma_u(w)w \tag{13}$$

where operator $T_u : \mathcal{V} \to \mathcal{V}$ and $\sigma_u : \mathcal{V} \to \mathbb{R}^+$ are defined by

$$T_u(w) =< u, < u, w >_{\mathcal{V}} >_{\mathcal{S}} \tag{14}$$

$$\sigma_u(w) = \frac{< T_u(w), w >_{\mathcal{V}}}{< w, w >_{\mathcal{V}}} \tag{15}$$

Under regularity assumptions on u, T_u is a symmetric compact operator on $\mathcal{V}$, such that classical spectral theory applies. When selecting an orthogonal basis $\{w_i\}_{i=1}^M$ of the dominant eigenspace of T_u, *i.e.* such that $< w_i, w_j >_{\mathcal{V}} = 0$ for $i \neq j$, stochastic functions are defined by $\lambda_i =< w_i, w_i >_{\mathcal{V}}^{-1} < u, w_i >_{\mathcal{V}}$. For many problems, the *a posteriori* computation of such a separated representation reveals that a good accuracy can be obtained with a low order M. In other words, there often exists a very low-dimensional reduced basis of deterministic and stochastic functions allowing to accurately represent the solution.

3.2 *A priori* separated representation: Generalized Spectral Decomposition

When the solution u is not known, the above classical Hilbert Karhunen-Loève decomposition can not be obtained. The Generalized Spectral Decomposition method (GSD) provides a methodology for the *a priori* construction (*i.e.* without knowing u) of a separated representation which has quite the same convergence properties as classical Hilbert Karhunen-Loève decompositions. This method belongs to the so called family of *Proper Generalized Decomposition* methods (PGD).

We here introduce a definition of the separated representation (12) based on two Galerkin orthogonality criteria. Let us denote $u_M = \sum_{i=1}^M w_i\lambda_i := W_M \cdot \Lambda_M$, where $W_M = (w_i)_{i=1}^M \in (\mathcal{V})^M$ and $\Lambda_M = (\lambda_i)_{i=1}^M \in (\mathcal{S})^M$. The set of deterministic functions W_M and stochastic functions Λ_M are then defined by:

$$A(W_M \cdot \Lambda_M, W_M \cdot \Lambda_M^*) = B(W_M \cdot \Lambda_M^*)\ \forall \Lambda_M^* \in (\mathcal{S})^M \tag{16}$$

$$A(W_M \cdot \Lambda_M, W_M^* \cdot \Lambda_M) = B(W_M^* \cdot \Lambda_M)\ \forall W_M^* \in (\mathcal{V})^M \tag{17}$$

Let $f : W_M \in (\mathcal{V})^M \mapsto f(W_M) \in (\mathcal{S})^M$ denote the mapping such that for a given W_M, $\Lambda_M = f(W_M)$ is the unique solution of (16). Let $F : \Lambda_M \in (\mathcal{S})^M \mapsto F(\Lambda_M) \in (\mathcal{V})^M$ denote the mapping such that for a given Λ_M, $W_M = F(\Lambda_M)$ is the unique solution of (17). Equations (16) and (17) are then respectively equivalent to $\Lambda_M = f(W_M)$ and $W_M = F(\Lambda_M)$. These two equations can be rescasted as follows:

$$T(W_M) = W_M, \text{ with } T(W_M) := (F \circ f)(W_M) \tag{18}$$

$$\Lambda_M = f(W_M) \tag{19}$$

Equation (18) can be interpreted as a pseudo eigenproblem where the linear subspace spanned by W_M is interpreted as a M-dimensional generalized eigenspace of operator T (see [37]).

Remark 2 - Denoting by $\mathcal{V}_M = span(W_M)$ *and* $\mathcal{S}_M = span(\Lambda_M)$ *the linear subspaces spanned by* $(w_i)_{i=1}^M$ *and* $(\lambda_i)_{i=1}^M$ *respectively, the proposed definition of the decomposition can be interpreted as follows: find optimal* M*-dimensional subspaces* $\mathcal{V}_M$ *and* $\mathcal{S}_M$ *such that* $u_M \in \mathcal{V}_M \otimes \mathcal{S}_M$ *verifies simultaneously the two following Galerkin orthogonality criteria:*

$$A(u_M, v) = B(v) \quad \forall v \in \mathcal{V}_M \otimes \mathcal{S} \tag{20}$$

$$A(u_M, v) = B(v) \quad \forall v \in \mathcal{V} \otimes \mathcal{S}_M \tag{21}$$

Equation (20) *(resp.* (21)*) defines* u_M *as the Galerkin approximation of* u *in the approximation space* $\mathcal{V}_M \otimes \mathcal{S}$ *(resp.* $\mathcal{V} \otimes \mathcal{S}_M$*). The proposed GSD definition can then be interpreted as an a priori Galerkin model reduction technique, where none of the reduced approximation spaces* $\mathcal{V}_M$ *and* $\mathcal{S}_M$ *are selected a priori (see [38] for the connection with other model reduction techniques).*

3.3 Interpretation of GSD

Definition (18) appears as a generalization of Hilbert-Karhunen-Loève decomposition where optimality is defined with respect to the bilinear form A of the problem. For the particular case where bilinear form A defines an inner product $\ll \cdot, \cdot \gg_A := A(\cdot, \cdot)$ on $\mathcal{V} \otimes \mathcal{S}$ with the following separation property:

$$\ll w\lambda, w^*\lambda^* \gg_A = < w, w^* >_{A,\mathcal{V}} < \lambda, \lambda^* >_{A,\mathcal{S}}, \tag{22}$$

the proposed definition exactly coincides with a Hilbert Karhunen-Loève decomposition. Indeed, in this case, $T(w) = \sigma_u(w)^{-1}\widetilde{T}_u(w)$, with

$$\widetilde{T}_u(w) = < u, < u, w >_{A,\mathcal{V}} >_{A,\mathcal{S}} \tag{23}$$

$$\widetilde{\sigma}_u(w) = \frac{< \widetilde{T}_u(w), w >_{A,\mathcal{V}}}{< w, w >_{A,\mathcal{V}}} \tag{24}$$

and equation (18) is equivalent to an eigenproblem on operator $\widetilde{T}_u$, which is the correlation operator of u based on inner products $< \cdot, \cdot >_{A,\mathcal{V}}$ and $< \cdot, \cdot >_{A,\mathcal{S}}$. Choosing W_M as a basis of the dominant eigenspace of $\widetilde{T}_u$ and choosing $\Lambda_M = f(W_M)$ leads to a decomposition u_M of order M which is optimal with respect to the norm $\| \cdot \|_A$ associated with $\ll \cdot, \cdot \gg_A$.

In the general case, (18) can not be interpreted as a classical eigenproblem. For problems where (4) are the Euler-Lagrange of a quadratic optimization problem on $\mathcal{V} \otimes \mathcal{S}$ (*i.e.* if A is a symmetric and coercive bilinear form), the concept of optimal decomposition associated with a dominant eigenspace can still be derived (see [37]). However, since it is not a classical eigenproblem, dedicated algorithms must be introduced in order to construct this optimal decomposition. For more general problems, although optimality properties are no longer available, algorithms inspired from classical algorithms for the solution of eigenproblems lead in practise to the construction of separated representations which have good convergence properties with M.

Remark 3 - For non symmetric problems, in order to rigourously define an optimality criterium and to obtain a rigorous definition of the dominance of generalized eigenspaces, the problem could be reformulated as an optimization problem, e.g. by

introducing a minimal residual formulation. This type of reformulation can be easily introduced in a finite dimensional (discretized) framework. However, in the continuous framework, it requires to manipulate non classical formulations of partial differential equations and induces many computational issues since non standard computation codes have to be implemented. In section 4.5.3, this type of reformulation will be discussed in a more general framework.

3.4 GSD algorithms

We here briefly recall different algorithms that have been proposed for the capture of quasi optimal decompositions. For a detailed description and in depth study of these algorithms, see [36,37].

3.4.1 Subspace iterations

A first algorithm for capturing the dominant eigenspace of operator T consists in building the series $W_M^{(k+1)} = T(W_M^{(k)})$, starting from an arbitrary set of functions $W_M^{(0)}$. This algorithm can be interpreted as a subspace iteration method for capturing the dominant eigenspace of operator T. In practise, $span(W_M^{(k)})$ often rapidly converges towards a subspace $span(W_M)$, which defines a generalized spectral decomposition $u_M = W_M \cdot f(W_M)$ which verifies the two Galerkin orthogonality criteria (16) and (17). In the context of the solution of an SPDE, one iteration of this algorithm can be interpreted as follows: first, for a given set of M deterministic functions W_M, we compute $\Lambda_M = f(W_M)$ by solving a system of M stochastic algebraic equations corresponding to a Galerkin approximation of the SPDE on the subspace $\mathcal{V}_M \otimes \mathcal{S}$, with $\mathcal{V}_M = span(W_M)$. In a second time, we compute $W_M = F(\Lambda_M)$ by solving a system of M coupled PDEs corresponding to a Galerkin approximation of the SPDE on the subspace $\mathcal{V} \otimes \mathcal{S}_M$, with $\mathcal{S}_M = span(\Lambda_M)$. From a computational point of view, this algorithm has two main drawbacks. First, such as classical stochastic Galerkin methods, it still requires the solution of a coupled system of deterministic PDEs. Secondly, since we do not know *a priori* the order M required for a given accuracy, this algorithm has to be repeated for increasing orders M until reaching the desired accuracy, thus leading to unnecessary intermediate computations. Other algorithms have been proposed in order to minimize the computational efforts and in order to only require the solution of uncoupled deterministic PDEs.

3.4.2 Power algorithm

Power algorithm consists in performing subspace iterations on a one-dimensional subspace in order to capture the dominant eigenfunctions w_i of successive operators $T^i = F^i \circ f^i$, where mappings $f^i : \mathcal{V} \to \mathcal{S}$ and $F^i : \mathcal{S} \to \mathcal{V}$ are defined such that $\lambda = f^i(w)$ and $w = F^i(\lambda)$ are respectively the unique solutions of the two following problems:

$$A(w\lambda, w\lambda^*) = B(w\lambda^*) - A(u_i, w\lambda^*) \quad \forall \lambda^* \in \mathcal{S} \tag{25}$$

$$A(w\lambda, w^*\lambda) = B(w^*\lambda) - A(u_i, w^*\lambda) \quad \forall w^* \in \mathcal{V} \tag{26}$$

where u_i is the previously computed order i decomposition. This algorithm allows a progressive construction of the set of deterministic functions W_M. The separated decomposition u_M of order M can be defined by letting the $\lambda_i = f^i(w_i)$, for $i \in \{1, \ldots, M\}$. In the case where the generalized spectral decomposition corresponds to a classical eigenproblem, this construction leads to the optimal decomposition. However, for the general case, it only leads to a sub-optimal decomposition. An update of stochastic functions often significantly improves the accuracy of the decomposition. This update consists in defining the stochastic functions associated with W_M by $\Lambda_M = f(W_M)$, which requires the solution of a system of M stochastic algebraic equations.

3.4.3 Arnoldi algorithm

Another algorithm, inspired from Arnoldi algorithm, has been proposed in [37] in order to further minimize the computational efforts. This algorithm leads to a decomposition which for a given order M is less accurate than with subspace iteration (and sometimes than power method with update). However, it only requires the solution of M uncoupled PDEs in order to build the set of functions W_M. An Arnoldi procedure for the construction of W_M is as follows: starting from a function $\lambda \in \mathcal{S}$, we compute an initial function $w_1 = F(\lambda)$ by solving a simple deterministic PDE. Then, we compute the generalized Krylov subspace $\mathcal{K}_M(T, w_1) = span\{w_i\}_{i=1}^M$, defined by $w_{i+1} = \Pi_{\mathcal{K}_i^\perp} T(w_i)$, where $\Pi_{\mathcal{K}_i^\perp}$ is a projector onto the orthogonal of the i-dimensional Krylov subspace. The computation of w_{i+1} from w_i can be decomposed into three steps: in a first time, we compute $\lambda = f(w_i)$ by solving a simple stochastic algebraic equation, which is equivalent to a Galerkin projection of the initial SPDE on a 1-dimensional deterministic reduced basis $span\{w_i\} \subset \mathcal{V}$. In a second time, we compute $w_{i+1} = F(\lambda)$ by solving a simple deterministic PDE, which is equivalent to a stochastic Galerkin projection on a 1-dimensional stochastic reduced basis $span\{\lambda\} \subset \mathcal{S}$. In a third time, we orthogonalize w_{i+1} with respect to $\mathcal{K}_i = span\{w_j\}_{j=1}^i$ (orthogonalization with respect to a chosen inner product on $\mathcal{V}$). A basis W_M being obtained, the associated stochastic functions $\Lambda_M = f(W_M)$ are obtained by solving a system of M stochastic algebraic equations. This procedure is summarized in the following algorithm.

Algorithm 1 *Arnoldi algorithm for GSD*

Inititialize $\lambda \in \mathcal{S}$
for $i = 1 \ldots M$ ***do***
  Compute $w_i = F(\lambda)$ *{Deterministic PDE}*
  Orthogonalize w *with respect to* $span(W_{i-1})$
  Compute $\lambda = f(w_i)$ *{Stochastic algebraic equation}*
end for
Compute $\Lambda_M = f(W_M)$ *{System of stochastic algebraic equations}*

Remark 4 - In practise, the Arnoldi procedure may break at a given iteration i. If the associated decomposition $u_i = W_i \cdot f(\Lambda_i)$ has not reached the desired accuracy, the algorithm is then restarted on the "deflated" operator T^i, defined in section 3.4.2. For a detailed description and in depth study of the above algorithms, see [37].

3.5 Computational aspects of GSD algorithms

GSD algorithms have been introduced in a quite abstract setting. Here, we detail the computational aspects of the algorithms by simply specifying how to apply the mappings F, f, F^i and f^i.

3.5.1 Separated representation of bilinear and linear forms

We consider that bilinear form a and linear form b in equation (3) admit the following separated representations: $\forall w, w^* \in \mathcal{V}$,

$$a(w, w^*; \boldsymbol{\xi}) = \sum_{k=1}^{K_A} a_k(w, w^*) A_k(\boldsymbol{\xi}), \tag{27}$$

$$b(w^*; \boldsymbol{\xi}) = \sum_{k=1}^{K_B} b_k(w^*) B_k(\boldsymbol{\xi}), \tag{28}$$

where the a_k are deterministic bilinear forms on $\mathcal{V}$, where the b_k are deterministic linear forms on $\mathcal{V}$, and where the A_k and B_k are real-valued random variables defined on $(\boldsymbol{\Xi}, \mathcal{B}, P_{\boldsymbol{\xi}})$.

3.5.2 Application of mappings F and F^i

Mapping $F : \mathcal{S} \to \mathcal{V}$ is defined such that $w = F(\lambda)$ is the solution of the following problem:

$$a_\lambda(w, w^*) = b_\lambda(w^*) \quad \forall w^* \in \mathcal{V} \tag{29}$$

where a_λ and b_λ are deterministic bilinear and linear forms on $\mathcal{V}$ defined by

$$a_\lambda(w, w^*) = \sum_{k=1}^{K_A} E(A_k \lambda \lambda) a_k(w, w^*) \tag{30}$$

$$b_\lambda(w^*) = \sum_{k=1}^{K_B} E(B_k \lambda) b_k(w^*) \tag{31}$$

Equation (29) is then a classical deterministic PDE.

Mapping $F^i : \mathcal{S} \to \mathcal{V}$ is defined such that $w = F^i(\lambda)$ is the solution of (29) with the following modified right-hand side:

$$\begin{aligned} b^i_\lambda(w^*) = &\sum_{k=1}^{K_B} E(B_k \lambda) b_k(w^*) \\ &- \sum_{j=1}^{i} \sum_{k=1}^{K_A} E(A_k \lambda \lambda_j) a_k(w_j, w^*) \end{aligned} \tag{32}$$

In practise, problem (29) is solved using classical discretization techniques.

3.5.3 Application of mappings f and f^i

Mapping $f : \mathcal{V} \to \mathcal{S}$ is defined such that $\lambda = f(w)$ is the solution of the following problem:

$$\alpha_w(\lambda, \lambda^*) = \beta_w(\lambda^*) \quad \forall \lambda^* \in \mathcal{S} \tag{33}$$

where α_w and β_w are bilinear and linear forms on $\mathcal{S}$ defined by

$$\alpha_w(\lambda, \lambda^*) = E(\lambda^*(\boldsymbol{\xi}) A(\boldsymbol{\xi}) \lambda(\boldsymbol{\xi})), \tag{34}$$

$$A(\boldsymbol{\xi}) = \sum_{k=1}^{K_A} a_k(w, w) A_k(\boldsymbol{\xi}) \tag{35}$$

$$\beta_w(\lambda^*) = E(\lambda^*(\boldsymbol{\xi}) B(\boldsymbol{\xi})), \tag{36}$$

$$B(\boldsymbol{\xi}) = \sum_{k=1}^{K_B} b_k(w) B_k(\boldsymbol{\xi}) \tag{37}$$

Equation (33) corresponds to a weak formulation of the simple stochastic algebraic equation $A(\boldsymbol{\xi})\lambda(\boldsymbol{\xi}) = B(\boldsymbol{\xi})$.

Mapping $f^i : \mathcal{V} \to \mathcal{S}$ is defined such that $\lambda = f^i(w)$ is the solution of (33) with the following modified right-hand side:

$$\beta_w^i(\lambda^*) = E(\lambda^*(\boldsymbol{\xi}) B^i(\boldsymbol{\xi})), \tag{38}$$

where

$$\begin{aligned} B^i(\boldsymbol{\xi}) = & \sum_{k=1}^{K_B} b_k(w) B_k(\boldsymbol{\xi}) \\ & - \sum_{j=1}^{i} \sum_{k=1}^{K_A} A_k(\boldsymbol{\xi}) \lambda_j(\boldsymbol{\xi}) a_k(w_j, w) \end{aligned} \tag{39}$$

Mapping $f : (\mathcal{V})^M \to (\mathcal{S})^M$ is defined such that $\Lambda_M = f(W_M)$ is the solution of the following problem:

$$\alpha_W(\Lambda_M, \Lambda_M^*) = \beta_W(\Lambda_M^*) \quad \forall \Lambda_M^* \in (\mathcal{S})^M \tag{40}$$

where α_W and β_W are bilinear and linear forms on $(\mathcal{S})^M$ defined by

$$\alpha_W(\Lambda_M, \Lambda_M^*) = E(\boldsymbol{\Lambda}^{*T}(\boldsymbol{\xi}) \mathbf{A}(\boldsymbol{\xi}) \boldsymbol{\Lambda}(\boldsymbol{\xi})), \tag{41}$$

$$\beta_W(\Lambda_M^*) = E(\boldsymbol{\Lambda}^{*T}(\boldsymbol{\xi}) \mathbf{B}(\boldsymbol{\xi})) \tag{42}$$

where $\Lambda_M \in (\mathcal{S})^M$ has been assimilated with a random vector $\boldsymbol{\Lambda} \in L^2(\boldsymbol{\Xi}, \mathcal{B}, P_{\boldsymbol{\xi}}; \mathbb{R}^M) \simeq \mathbb{R}^M \otimes \mathcal{S}$, and where random matrix $\mathbf{A}$ and random vector $\mathbf{B}$ are defined by

$$(\mathbf{A}(\boldsymbol{\xi}))_{ij} = \sum_{k=1}^{K_A} a_k(w_j, w_i) A_k(\boldsymbol{\xi}), \tag{43}$$

$$(\mathbf{B}(\boldsymbol{\xi}))_i = \sum_{k=1}^{K_B} b_k(w_i) B_k(\boldsymbol{\xi}) \tag{44}$$

3.5.4 How to solve stochastic algebraic equations ?

Stochastic algebraic equations (33) and (40) can be classically solved using a Galerkin spectral stochastic method. After the introduction of an approximation space $\mathcal{S}_P$, computing the Galerkin projection $\lambda \in \mathcal{S}_P$ (resp. $\Lambda_M \in (\mathcal{S}_P)^M$) requires the solution of a system of P (resp. PM) equations (see appendix A for details on this classical solution technique).

For high-dimensional stochastic problems (requiring a very large P), the solution of these stochastic algebraic equations may be computationally costly or even unaffordable. In the following section, we introduce a methodology based on separation of variables in order to solve these stochastic algebraic equations in the case of high-dimensional probability spaces.

4 Proper generalized decomposition for solving equations defined on tensor product spaces

In this section, we introduce a methodology for the *a priori* construction of a separated representation of the solution of the following problem defined on a multi-dimensional tensor product space:

$$\begin{aligned} &u \in \mathcal{S}^0 \otimes \mathcal{S}^1 \otimes \ldots \otimes \mathcal{S}^r, \\ &\alpha(u,v) = \beta(v) \quad \forall v \in \mathcal{S}^0 \otimes \mathcal{S}^1 \otimes \ldots \otimes \mathcal{S}^r \end{aligned} \tag{45}$$

where α and β are bilinear and linear forms. This problem can be associated with the initial SPDE (4), by letting $\alpha := A$, $\beta := B$ and $\mathcal{S}^0 := \mathcal{V}$. Letting $\mathcal{S}^0 := \mathbb{R}^n$, equation (45) can be interpreted as a system of stochastic algebraic equations. For example, such a system is obtained after a discretization of the SPDE at the deterministic level (*e.g.* after introducing a finite dimensional approximation space $\mathcal{V}_n \subset \mathcal{V}$). It is also associated with stochastic algebraic equations (33) and (40) whose solution is required by GSD algorithms introduced in section (3) (see section 3.5.4). The proposed methodology can be seen as an extension of GSD method to the case $r \geqslant 2$ and it belongs to the family of Proper Generalized Decomposition (PGD) methods.

4.1 Separated representation of the solution

An order Z separated representation of the solution of (45) is defined by

$$u(\boldsymbol{\xi}) \approx u_Z(\boldsymbol{\xi}) = \sum_{i=1}^{Z} \phi_i^0 \phi_i^1(\boldsymbol{\xi}_1) \ldots \phi_i^r(\boldsymbol{\xi}_r) \tag{46}$$

where $\phi^i \in \mathcal{S}^i$. The optimality of such a decomposition is clearly related to the metric which is used for estimating the distance between u and u_Z. An optimal separated representation (46) could be naturally defined *a posteriori* by introducing a classical norm $\|\cdot\|$ on $\otimes_{j=0}^r \mathcal{S}^j$ and by letting

$$\|u - u_Z\| = \min_{\{\phi_1^j\}_{j=0}^r, \ldots, \{\phi_Z^j\}_{j=0}^r} \|u - \sum_{i=1}^{Z} \phi_i^0 \ldots \phi_i^r\| \tag{47}$$

In the case $r = 1$, this definition corresponds to a classical order M singular value decomposition, also named Karhunen-Loève decomposition or Proper Orthogonal Decomposition. In the general case $r > 1$, this appears as a multi-dimensional generalization of singular value decomposition which has been extensively studied in the literature in the finite dimensional case (see e.g. [9,26,25] and the references therein) and in the infinite dimensional case [32]. In this general case, the *a posteriori* construction of an optimal decomposition, *i.e.* leading to the minimal order Z for a given accuracy, is a non trivial and sometimes ill-posed problem [24,10]. Various algorithms have been proposed which lead to quasi optimal but not necessarily optimal decompositions.

In this section, we focus on the more complicated problem of the *a priori* construction of the separated representation u_Z, without knowing the solution u *a priori*. A basic algorithm is proposed that leads to quite good convergence properties of the decomposition in many situations.

4.2 Circumvent the curse of dimensionality for spectral stochastic methods

Decomposition (46) can be equivalently rewritten

$$u_Z(\boldsymbol{\xi}) = \sum_{i=1}^{Z} \phi_i^0 \Psi_i(\boldsymbol{\xi}), \quad \Psi_i(\boldsymbol{\xi}) := \phi_i^1(\boldsymbol{\xi}_1) \dots \phi_i^r(\boldsymbol{\xi}_r) \tag{48}$$

with $\Psi_i(\boldsymbol{\xi}) \in \otimes_{j=1}^r \mathcal{S}^j \simeq \mathcal{S} = L^2(\boldsymbol{\Xi}, \mathcal{B}, P_{\boldsymbol{\xi}})$. It then appears as a spectral stochastic expansion of a second order random variable u with values in $\mathcal{S}^0$ on a basis $\{\Psi_i\}_{i=1}^Z$, defining a Z-dimensional approximation space $\mathcal{S}_Z \subset \mathcal{S}$. Here, the difference with a classical spectral stochastic approach is that the stochastic approximation basis is not selected *a priori* but is selected in order to accurately approximate the solution with a very low dimension Z. The following algorithms aim at capturing *a priori* such an optimal representation. We will see in the numerical examples that for a given accuracy of the approximation, several orders of magnitude (10, 10^{10}, 10^{100}, ...) may exist between the optimal Z and the dimension P of classical stochastic approximation spaces $\mathcal{S}_P$ defined in section 2.3. For high-dimensional stochastic problems, this methodology can be seen as a way to circumvent the curse of dimensionality associated with the dramatic increase in the dimension of stochastic approximation spaces, when increasing the dimension of the underlying probability space.

4.3 Progressive definition of the decomposition based on Galerkin orthogonality criteria

We first consider a progressive definition of the decomposition (46). We suppose that an approximate order Z decomposition u_Z has been determined. The aim is then to define a new set of functions $(\phi^0, \phi^1, \dots, \phi^r) \in \mathcal{S}^0 \times \mathcal{S}^1 \times \dots \times \mathcal{S}^r$, leading to the following $Z + 1$ decomposition:

$$u_{Z+1} = u_Z + \phi^0 \phi^1 \dots \phi^r \tag{49}$$

We here propose to define the new set of functions by the following $r+1$ Galerkin orthogonality criteria:

$$\begin{array}{l} \forall(\widetilde{\phi^0},\ldots,\widetilde{\phi^r}) \in \mathcal{S}^0 \times \ldots \times \mathcal{S}^r, \\ \alpha(u_Z + \phi^0\phi^1 \ldots \phi^r, \widetilde{\phi^0}\phi^1 \ldots \phi^r) = \beta(\widetilde{\phi^0}\phi^1 \ldots \phi^r) \\ \alpha(u_Z + \phi^0\phi^1 \ldots \phi^r, \phi^0\widetilde{\phi^1} \ldots \phi^r) = \beta(\phi^0\widetilde{\phi^1} \ldots \phi^r) \\ \ldots \\ \alpha(u_Z + \phi^0\phi^1 \ldots \phi^r, \phi^0\phi^1 \ldots \widetilde{\phi^r}) = \beta(\phi^0\phi^1 \ldots \widetilde{\phi^r}) \end{array} \tag{50}$$

We introduce the following mappings

$$\begin{array}{l} F_0^Z : \mathcal{S}^1 \times \mathcal{S}^2 \times \ldots \times \mathcal{S}^r \to \mathcal{S}^0 \\ F_1^Z : \mathcal{S}^0 \times \mathcal{S}^2 \times \ldots \times \mathcal{S}^r \to \mathcal{S}^1 \\ \ldots \\ F_r^Z : \mathcal{S}^0 \times \mathcal{S}^1 \times \ldots \times \mathcal{S}^{r-1} \to \mathcal{S}^r \end{array} \tag{51}$$

such that the set of equations (50) can be equivalently written:

$$\begin{array}{l} \phi^0 = F_0^Z(\phi^1, \phi^2, \ldots, \phi^r) \\ \phi^1 = F_1^Z(\phi^0, \phi^2, \ldots, \phi^r) \\ \ldots \\ \phi^r = F_r^Z(\phi^0, \phi^1, \ldots, \phi^{r-1}) \end{array} \tag{52}$$

Let us note that the product $\prod_{j=0}^r \phi^j$ is unchanged by the following rescaling of functions:

$$\prod_{j=0}^{r} \phi^j = \prod_{j=0}^{r} \gamma^j \phi^j, \quad \prod_{j=0}^{r} \gamma^j = 1, \tag{53}$$

This defines an equivalence class of separated functions. Selecting for the rescaling factor $\gamma^j = \|\phi^j\|_{\mathcal{S}^j}^{-1}$, for $j \in \{1,\ldots,r\}$, and $\gamma^0 = \prod_{j=1}^r 1/\gamma^j$, yields normalized functions $\{\gamma^j\phi^j\}_{j=1}^r$. We now introduce the following iterative algorithm 2 for the construction of the set of functions $(\phi^0, \phi^1, \ldots, \phi^r)$ having the above normalization property.

Algorithm 2 *Power-type iterations*

Require: u_Z

1: *Initialize* $(\phi^0, \ldots, \phi^r)$
2: ***loop***
3: ***for*** $j = 1 \ldots r$ ***do***
4: $\phi^j = F_j^Z(\{\phi^l\}_{l=0, l\neq j}^r)$
5: $\phi^j = \phi^j / \|\phi^j\|_{\mathcal{S}^j}$
6: ***end for***
7: $\phi^0 = F_0^Z(\{\phi^l\}_{l=1}^r)$
8: *Check convergence of* $\phi^0 \ldots \phi^r$ *{tolerance* ε_{tol}*}*
9: ***end loop***

In practise, a simple stagnation criterium is used for checking convergence in step 8. The initialization is usually generated randomly. For many types of problems, we observe that this initialization has only a slight influence on the convergence of the algorithm. The tolerance ε_{tol} in algorithm 2 can be relatively coarse (in practise, we take $\varepsilon_{tol} \approx 10^{-2}$). Also, the maximum number of iterations in the loop is usually taken relatively small (≈ 4). These choices will be justified in the numerical examples.

4.4 Global update of functions

In many situations, the above progressive construction of the decomposition may have a very slow convergence with Z, far slower than the ideal *a posteriori* separated representation defined in equation (47). We here propose to perform a global update of functions, which in practise significantly improves the convergence properties of the decomposition. Let $\Phi_Z^j := \{\phi_1^j, \dots, \phi_Z^j\} \in (\mathcal{S}^j)^Z$. The whole set of functions $\{\Phi_Z^j\}_{j=0}^r$ can be defined by the following $r+1$ Galerkin orthogonality criteria:

$$\begin{array}{l} \alpha(\sum_{i=1}^Z \phi_i^0\phi_i^1\dots\phi_i^r, \sum_{i=1}^Z \widetilde{\phi_i^0}\phi_i^1\dots\phi_i^r) = \\ \quad \beta(\sum_{i=1}^Z \widetilde{\phi_i^0}\phi_i^1\dots\phi_i^r) \quad \forall \{\widetilde{\phi_i^0}\}_{i=1}^Z \in (\mathcal{S}^0)^Z \\ \dots \\ \alpha(\sum_{i=1}^Z \phi_i^0\phi_i^1\dots\phi_i^r, \sum_{i=1}^Z \phi_i^0\phi_i^1\dots\widetilde{\phi_i^r}) = \\ \quad \beta(\sum_{i=1}^Z \phi_i^0\phi_i^1\dots\widetilde{\phi_i^r}) \quad \forall \{\widetilde{\phi_i^r}\}_{i=1}^Z \in (\mathcal{S}^r)^Z \end{array} \tag{54}$$

We introduce the following mappings:

$$\begin{array}{l} F_0 : (\mathcal{S}^1)^Z \times \dots \times (\mathcal{S}^r)^Z \to (\mathcal{S}^0)^Z \\ \dots \\ F_r : (\mathcal{S}^0)^Z \times \dots \times (\mathcal{S}^{r-1})^Z \to (\mathcal{S}^r)^Z \end{array} \tag{55}$$

such that the set of equations (54) can be equivalently written:

$$\begin{array}{l} \Phi_Z^0 = F_0(\Phi_Z^1, \dots, \Phi_Z^r) \\ \dots \\ \Phi_Z^r = F_r(\Phi_Z^0, \dots, \Phi_Z^{r-1}) \end{array} \tag{56}$$

We now propose the following algorithm for the *a priori* construction of a separated representation of the solution of problem (45).

Algorithm 3 *Progressive construction with update (multidimensional PGD)*

Set $u_0 := 0$
for $Z = 1 \dots Z_{max}$ ***do***
  Compute a new set $(\phi_Z^0, \dots, \phi_Z^r)$ *with algo. 2*
  for $n = 1$ *to* N_{update} ***do***
    for all $j \in J_{update}$ ***do***
      $\Phi_Z^j = F_j(\{\Phi_Z^l\}_{l=0, l\neq j}^r)$
    end for
  end for
  Check convergence of u_Z
end for

The set $J_{update} \subset \{0, \dots, r\}$ is composed by the dimensions j for which the sets of functions Φ_Z^j are updated. One usually observes that the accuracy of the decomposition is improved when increasing the set J_{update}. In practice, when the updating along a dimension j is achievable from a computational point of view, this dimension should be added to the set J_{update}. Repeating the updating step several times (*i.e.* taking $N_{update} > 1$) may improve the quality of the obtained decomposition. However, since the computational cost of this updating step increases (non linearly) with the order Z, unnecessary updates should be avoided. There is no general theoretical results about the efficiency of this updating step, which is clearly problem dependent. Numerical

experiences may help deriving guidelines for a specific class of problems. From the experiences of the author, one observes that $N_{update} = 1$ is sufficient in many situations, especially for the case of SPDEs dealt with in this article. For the practical implementation of this algorithm, see appendix B.

4.5 Interpretation of algorithm and comments

4.5.1 The case $r = 1$: Generalized Spectral Decomposition

The case $r = 1$ (*i.e.* when function space $\mathcal{S}^0 \otimes \mathcal{S}^1$ is a tensor product of two spaces) corresponds to the case of the generalized spectral decomposition described in section 3, which appears as a generalization of Karhunen-Loève decomposition. We show in this case that optimal functions $\phi_i^0 \in \mathcal{S}^0$ (resp. $\phi_i^1 \in \mathcal{S}^1$) are associated with the dominant eigenspace of a pseudo eigenproblem on operator $F_0 \circ F_1$ (resp. $F_1 \circ F_0$). Several algorithms have been proposed and studied for the capture of an approximation of the dominant eigenspace (see section 3.4). Here, algorithm 2 corresponds to power-type iterations for finding the dominant eigenfunction of the deflated operator $(F_0^Z \circ F_1^Z)$. Algorithm 3 then corresponds to a power-type method with deflation and update for capturing an approximate generalized spectral decomposition (see section 3.4.2 and [36, 37]). For classical eigenproblems (*i.e.* for classical spectral decomposition), it can be proved that updating has no effect [37]. However, in general (for the pseudo eigenproblem), it has been observed that updating can significantly improve the approximation of dominant eigenspaces and can lead to a better convergence with Z of the generalized spectral decomposition [36, 37, 42].

Further mathematical investigations are still necessary for a better understanding of this pseudo eigenproblem, for which – to the knowledge of the author – there is no mathematical framework available (see [37] for discussions on this pseudo eigenproblem). However, the proposed power-type algorithm with update seems to lead to a rather good approximation of the optimal decomposition in many situations.

4.5.2 The case $r > 1$

In the case $r > 1$, there is no straightforward interpretation in terms of an pseudo eigenproblem. Further investigations will be necessary in order to correctly interpret the decomposition and propose more efficient algorithms, possibly still inspired from algorithms for solving classical eigenproblems, or from other algorithms for the *a posteriori* construction of separated representations.

For the particular case where $\alpha(\cdot, \cdot)$ is a symmetric continuous coercive bilinear form on $\otimes_{j=0}^r \mathcal{S}^j$, the proposed construction can also be interpreted as a nonlinear approximation algorithm. Indeed, for this particular case, problem (45) can be reformulated as the following minimization problem

$$u = \arg \min_{v \in \mathcal{S}^0 \otimes \ldots \otimes \mathcal{S}^r} \frac{1}{2}\alpha(v, v) - b(v) \tag{57}$$

$$= \arg \min_{v \in \mathcal{S}^0 \otimes \ldots \otimes \mathcal{S}^r} \|u - v\|_\alpha^2, \tag{58}$$

where $\|u\|_\alpha^2 = \alpha(u, u)$ denotes the norm induced by α. Equations (50) are then associated with stationarity conditions (or Euler-Lagrange equations) of the following

optimization problem:

$$\min_{\phi^0,\ldots,\phi^r} \|u - u_Z - \phi^0 \ldots \phi^r\|_\alpha^2 \tag{59}$$

while equations (54) are associated with stationarity conditions of the following optimization problem:

$$\min_{\{\phi_i^0\}_{i=0}^Z,\ldots,\{\phi_i^r\}_{i=1}^Z} \|u - \sum_{i=1}^{Z} \phi_i^0 \ldots \phi_i^r\|_\alpha^2 \tag{60}$$

The construction of the decomposition can then be interpreted as a nonlinear approximation problem, where the optimal separated representation is defined as the one which minimizes the distance to u with respect to the metric induced by the bilinear form α. A proof of the convergence of the progressive decomposition u_Z, defined by

$$\|u - u_{Z+1}\|_\alpha^2 = \min_{\phi^0,\ldots,\phi^r} \|u - u_Z - \phi^0 \ldots \phi^r\|_\alpha^2 \tag{61}$$

can be found in [16] in an abstract setting, for problems defined in tensor product spaces. In [7], the progressive construction (without update) has been interpreted as a Greedy algorithm in nonlinear approximation [12, 4], where the "dictionary" is composed by separated functions of type $\prod_{j=0}^r \phi^j$, $\phi^j \in \mathcal{S}^j$.

Algorithm 2 then corresponds to an alternated minimization procedure, where minimization is performed on a function $\phi^j \in \mathcal{S}^j$ while letting fixed the other functions $\phi^{j'}$, $j' \neq j$. In algorithm 3, the updating step corresponds to the minimization problem (60), where successive minimizations are performed along dimensions $j \in J_{update}$. It is easy to prove that iterative algorithm 2 has a monotonic convergence. It is also straightforward to prove that algorithm 3 leads to a monotone convergence of the decomposition u_Z with Z. Performing several updates in algorithm 3 ($N_{update} > 1$) corresponds to performing several iterations of an alternated minimization procedure for solving (60). In practise, one observes that performing only one iteration (*i.e.* only one update per updated dimension, $N_{update} = 1$) is often sufficient. Additional iterations do not significantly improve the accuracy. This has been observed on several numerical examples but since only a few mathematical results are available, it should be confirmed on a larger set of examples.

In the opinion of the author, the interpretation as a pseudo eigenproblem seems more pertinent than an interpretation as a nonlinear approximation problem, and could lead to the development of more efficient algorithms to capture an optimal decomposition or an approximation of it (as it is done in the case $r = 1$ with the GSD algorithms).

4.5.3 Reformulation as an optimization problem: necessary or not ?

If problem (45) corresponds to stationarity conditions of a quadratic optimization problem, monotone convergence of algorithm 3 can be proved. It is a property of robustness of the algorithm and of the proposed construction. In order to recover this robustness for more general problems (*e.g.* for non-symmetric bilinear form α), a reformulation of problem (45) as an optimization problem can be introduced. Let $R(u) \in \mathcal{S}^0 \otimes \ldots \otimes \mathcal{S}^r$ denote the residual of equation (45), defined by

$$< v, R(u) >_{\mathcal{S}^0 \otimes \ldots \otimes \mathcal{S}^r} := < v, \beta - \alpha(u) >_{\mathcal{S}^0 \otimes \ldots \otimes \mathcal{S}^r} \tag{62}$$

$$:= \beta(v) - \alpha(u, v) \tag{63}$$

where $< \cdot, \cdot >_{\mathcal{S}^0 \otimes \ldots \otimes \mathcal{S}^r}$ denotes an inner product on Hilbert space $\otimes_{j=0}^{r} \mathcal{S}^j$ and where $\beta \in \otimes_{j=0}^{r} \mathcal{S}^j$ and $\alpha(u) \in \otimes_{j=0}^{r} \mathcal{S}^j$ are associated with linear forms $\beta(\cdot)$ and $a(u, \cdot)$ by Riez representation. Then, denoting by $\| \cdot \|$ the associated norm, the separated decomposition can be progressively defined as follows

$$\min_{\phi^0, \ldots, \phi^r} \|R(u_Z + \phi^0 \ldots \phi^r)\|^2 \tag{64}$$

which can be rewritten as (59) by replacing bilinear form $\alpha(u, v)$ and linear form $\beta(v)$ by bilinear form $< \alpha(v), \alpha(u) >$ and linear form $< \alpha(v), \beta >$ respectively. Equations (50) then have to be interpreted as the stationarity conditions associated with optimization problem (64). The obtained decomposition u_Z then satisfies an optimality criterium with respect to the residual norm. Under suitable assumptions, the convergence of the progressive decomposition u_Z defined by (64) can be proved [16].

However, one observes in practise that it leads to poor convergence properties of u_Z with respect to natural norms in tensor product Hilbert spaces (*e.g.* L^2 norm). Although monotone convergence is not guaranteed for non-variational problems (non symmetric problems), in many cases, a construction based on Galerkin orthogonality criteria appears to yield better convergence properties with respect to usual norms and should be preferred when one tries to obtain the lowest order of decomposition for a given precision with respect to a usual norm.
The minimal residual formulation also presents another drawback from the computational point of view. Indeed, algorithms based on separation of variables take part of the separated representation of the operator and right-hand side (see appendix B on computational aspects). In this minimal residual formulation, the initial operator and right-hand side are multiplied by the adjoint operator, which drastically increase the separation order of the operator and right-hand side of the new formulation.

Remark 5 - This minimal residual formulation (or least-square formulation) has been proposed in [5] for the solution of algebraic equations in finite dimensional tensor product spaces and applied to the solution of stochastic algebraic equations in [13]. For each order Z, the authors proposed an algorithm based on an alternated minimization procedure for solving

$$\min_{\{\phi_i^0\}_{i=0}^{Z}, \ldots, \{\phi_i^r\}_{i=1}^{Z}} \|R(\sum_{i=1}^{Z} \phi_i^0 \ldots \phi_i^r)\|^2 \tag{65}$$

For each order Z, iterations are performed until convergence or stagnation. If the residual does not satisfy a desired accuracy, the algorithm is restarted with order $Z+1$. In the case $r = 1$, this corresponds to the subspace iterations for solving the pseudo eigenproblem (see section 3.4.1). For $r > 1$, this alternated minimization technique corresponds to the steps 4 to 8 of algorithm 3 (so called updating steps), with $J_{update} = \{0, \ldots, r\}$ (all dimensions). Since the required order Z for a given accuracy is not known a priori, this type of algorithm can lead to high computational costs. In this article, a progressive construction with updates is then preferred.

5 Example 1: advection diffusion reaction equation

5.1 Formulation of the problem and discretization

Formulation of the problem. We consider an advection diffusion reaction equation defined on a spatial domain $\mathit{\Omega} = (0,1) \times (0,1)$ and a time interval $I = (0,T)$, with $T = 0.03$. We denote by $\boldsymbol{\xi} \in \boldsymbol{\Xi}$ the random input parameters. The solution field $u(\boldsymbol{x}, t, \boldsymbol{\xi})$, defined on $\mathit{\Omega} \times I \times \boldsymbol{\Xi}$ verifies

$$\dot{u} - \mu(\boldsymbol{\xi})\Delta u + c(\boldsymbol{\xi}) \cdot \nabla u + \kappa(\boldsymbol{\xi})u = f(\boldsymbol{\xi}) \quad on \quad \mathit{\Omega} \times I \tag{66a}$$

$$u = 0 \quad on \quad \partial\mathit{\Omega} \times I \tag{66b}$$

$$u = 0 \quad on \quad \mathit{\Omega} \times \{0\} \tag{66c}$$

where $\dot{u} \equiv \partial_t u$, where μ and κ are random diffusion and reaction parameters, where c is a random advection velocity, and where f is a random source term. We take

$$\begin{aligned}
&\mu(\boldsymbol{\xi}) = 1 + 0.2\xi_1, \\
&c(\boldsymbol{\xi}) = 250(1 + 0.2\xi_2)(x - \frac{1}{2}, \frac{1}{2} - y), \\
&\kappa(\boldsymbol{\xi}) = 10(1 + 0.2\xi_3) \\
&f(\boldsymbol{\xi}) = 100(1 + 0.2\xi_4) I_{\mathit{\Omega}_1}
\end{aligned}$$

where $(x, y) = \boldsymbol{x} \in \mathit{\Omega}$, $I_{\mathit{\Omega}_1}$ is the indicator function of a subdomain $\mathit{\Omega}_1 = (0.7, 0.8) \times (0.7, 0.8) \subset \mathit{\Omega}$ (see figure 1) and where $\boldsymbol{\xi} = (\xi_i)_{i=1}^4$ is a set of 4 mutually independent uniform random variables $\xi_i \in U(-1,1)$. The set of elementary events is then $\boldsymbol{\Xi} = \times_{i=1}^4 \Xi_i$, with $\Xi_i = (-1,1)$, and is endowed with the uniform probability measure $P_{\boldsymbol{\xi}}$.

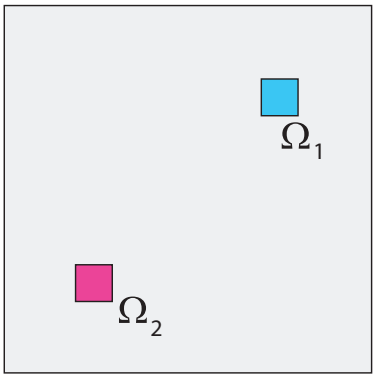
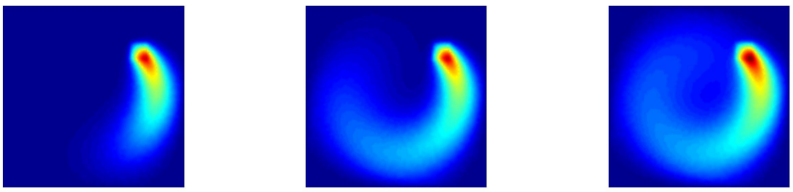


Fig. 1 Example 1.

On Figure 2, plotted is the solution corresponding to outcome $\boldsymbol{\xi} = 0$ (mean value of parameters).

Weak formulation. We introduce the weak formulation (4) of problem (66) with the following definition of function spaces

$$\begin{aligned}
&\mathcal{V} = \mathcal{V}^x \otimes \mathcal{V}^t, \quad \mathcal{V}^x = H_0^1(\mathit{\Omega}), \quad \mathcal{V}^t = L^2(I), \\
&\mathcal{S} = L^2(\boldsymbol{\Xi}, \mathcal{B}, P_{\boldsymbol{\xi}})
\end{aligned}$$



(a) $t = \frac{20}{80}T$ (b) $t = \frac{50}{80}T$ (c) $t = T$

Fig. 2 Example 1. Solution $u(\boldsymbol{\xi})$ for $\boldsymbol{\xi} = 0$ (mean values of parameters) at different time steps

and the following definitions of bilinear and linear forms:

$$a(u, v; \boldsymbol{\xi}) = \int_I \int_\Omega \dot{u} v \, dx \, dt + \int_\Omega u(0^+) v(0^+) \, dx$$
$$+ \int_I \int_\Omega \mu(\boldsymbol{\xi}) \nabla u \cdot \nabla v \, dx \, dt + \int_I \int_\Omega c(\boldsymbol{\xi}) \cdot \nabla u \, v \, dx \, dt$$
$$+ \int_I \int_\Omega \kappa(\boldsymbol{\xi}) u \, v \, dx \, dt \tag{67}$$
$$l(v; \boldsymbol{\xi}) = \int_I \int_\Omega v f(\boldsymbol{\xi}) \, dx \, dt \tag{68}$$

where $u(0^+) \equiv \lim_{t \downarrow 0} u(\boldsymbol{x}, t, \boldsymbol{\xi})$. Let us note that with this weak formulation, the initial condition is verified in a weak sense.

Discretization. At the space level, we introduce a finite element approximation space $\mathcal{V}^x_{N_x} \subset \mathcal{V}^x$ with dimension $N_x = 4435$. The finite element mesh composed of 3-nodes triangles is shown on figure 3. At the time level, we introduce a piecewise constant

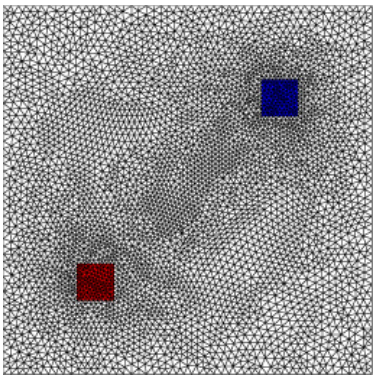


Fig. 3 Example 1. Finite element mesh

approximation space $\mathcal{V}^t_{N_t} \subset \mathcal{V}^t$ associated with a partition $\{I_i = (t_{i-1}, t_i)\}_{i=1}^{N_t}$ of the time interval I. A time discontinuous Galerkin framework is used by introducing the

following definition of time derivatives:

$$\int_I \int_\Omega \dot{u} v \, dx \, dt := \sum_{i=1}^{N_t - 1} \int_\Omega \left(u(t_i^+) - u(t_i^-) \right) v(t_i^+) \, dx$$

where $u(t_i^\pm) \equiv \lim_{\epsilon \downarrow 0} u(t_i \pm \epsilon)$. We here introduce a uniform partition with $N_t = 80$. Finally, at the stochastic level, we first introduce a classical polynomial approximation space $\mathcal{S}_P = \otimes_{i=1}^4 \mathcal{S}_{P_i}^i \subset \mathcal{S}$, where the $\mathcal{S}_{P_i}^i = \mathbb{P}_p(\Xi_i)$ are unidimensional polynomial spaces of degree $p = 5$ ($P_i = 6$). The dimension of $\mathcal{S}_P$ is then $P = 1296$. The classical Galerkin approximation is defined by

$$\begin{aligned} & u \in \mathcal{V}_{N_x}^x \otimes \mathcal{V}_{N_t}^t \otimes \mathcal{S}_P, \\ & A(u, v) = B(v) \quad \forall v \in \mathcal{V}_{N_x}^x \otimes \mathcal{V}_{N_t}^t \otimes \mathcal{S}_P \end{aligned} \tag{69}$$

Remark 6 Let us note that approximation space $\mathcal{S}_P$ is here defined as the full tensorization of unidimensional polynomial spaces (polynomial space with partial degree p). It does not correspond to the classical polynomial chaos approximation space (polynomial space with total degree p).

5.2 Generalized spectral decomposition

In this section, we apply the GSD algorithm 1 (Arnoldi-type algorithm) for the *a priori* construction of a decomposition of the solution

$$u(\boldsymbol{x}, t, \boldsymbol{\xi}) \approx u_M(\boldsymbol{x}, t, \boldsymbol{\xi}) = \sum_{i=1}^{M} w_i(\boldsymbol{x}, t) \lambda_i(\boldsymbol{\xi}) := W_M \cdot \Lambda_M$$

where the $w_i(\boldsymbol{x}, t) \in \mathcal{V}_{N_x}^x \otimes \mathcal{V}_{N_t}^t$ are deterministic modes (space-time modes) and the $\lambda_i \in \mathcal{S}_P$ are stochastic modes. In this section, we only focus on the properties of the GSD method introduced in section 3. We do not focus on the solution of stochastic algebraic equations and we consider that these equations are solved with a very good accuracy (error less than the error associated with the truncation order M of the GSD). The solution of these stochastic algebraic equations with the algorithm proposed in section 4 will be analyzed in the following section 5.3.

5.2.1 Algorithm and computational aspects of GSD

We recall that for building a decomposition of order M, the Arnoldi-type algorithm 1 requires the solution of M classical deterministic problems (problems $w_i = F(\lambda)$), M stochastic algebraic equations (problems $\lambda = f(w_i)$) and a system of stochastic algebraic equations (problem $\Lambda_M = f(W_M)$) for the update of stochastic functions. The set of M deterministic modes w_i are computed by solving only M uncoupled deterministic problems $w_i = F(\lambda)$ for different $\lambda \in \mathcal{S}_P$ (equation (29)). These problems correspond to classical advection diffusion reaction problems associated with different deterministic parameters $\mu_\lambda = E(\mu\lambda\lambda)$, $c_\lambda = E(c\lambda\lambda)$ and $\kappa_\lambda = E(\kappa\lambda\lambda)$ (respectively

for the diffusion, advection and reaction terms) and with a deterministic source term $f_\lambda = E(f\lambda)$. Bilinear and linear forms in equation (29) write

$$a_\lambda(w, w^*) = \int_I \int_\Omega E(\lambda\lambda)\dot{w}w^* \, dx \, dt + \int_\Omega E(\lambda\lambda)w(0^+)w^*(0^+) \, dx + \int_I \int_\Omega \kappa_\lambda w \, w^* \, dx \, dt + \int_I \int_\Omega \mu_\lambda \nabla w \cdot \nabla w^* \, dx \, dt + \int_I \int_\Omega c_\lambda \cdot \nabla w \, w^* \, dx \, dt \tag{70}$$

$$l_\lambda(v) = \int_I \int_\Omega w^* f_\lambda \, dx \, dt \tag{71}$$

5.2.2 Illustration of the obtained decomposition

We here illustrate the decomposition $u_9 = W_9 \cdot \Lambda_9$ of order $M = 9$ obtained by the Arnoldi-type algorithm. Figure 4 shows the first 4 deterministic modes $\{w_i\}_{i=1}^4$. These modes are orthonormalized with respect to the natural inner product in $L^2(\Omega) \otimes L^2(I)$. Figure 5 shows the probability density functions of stochastic modes Λ_9. In Table 1, we indicate the mean $m_1(\lambda_i) := E(\lambda_i)$ and second moment $m_2(\lambda_i) := E(\lambda_i^2)$ of each stochastic mode λ_i.

Table 1 First and second moments of random variables $\{\lambda_i\}_{i=1}^9$

i	$m_1(\lambda_i)$	$m_2(\lambda_i)$
1	12.458	157.7
2	0.603	0.8521
3	0.139	0.5362
4	−0.084	0.0467
5	−0.055	0.0073
6	0.035	0.0029
7	0.123	0.0387
8	0.008	0.0002
9	0.050	0.0065

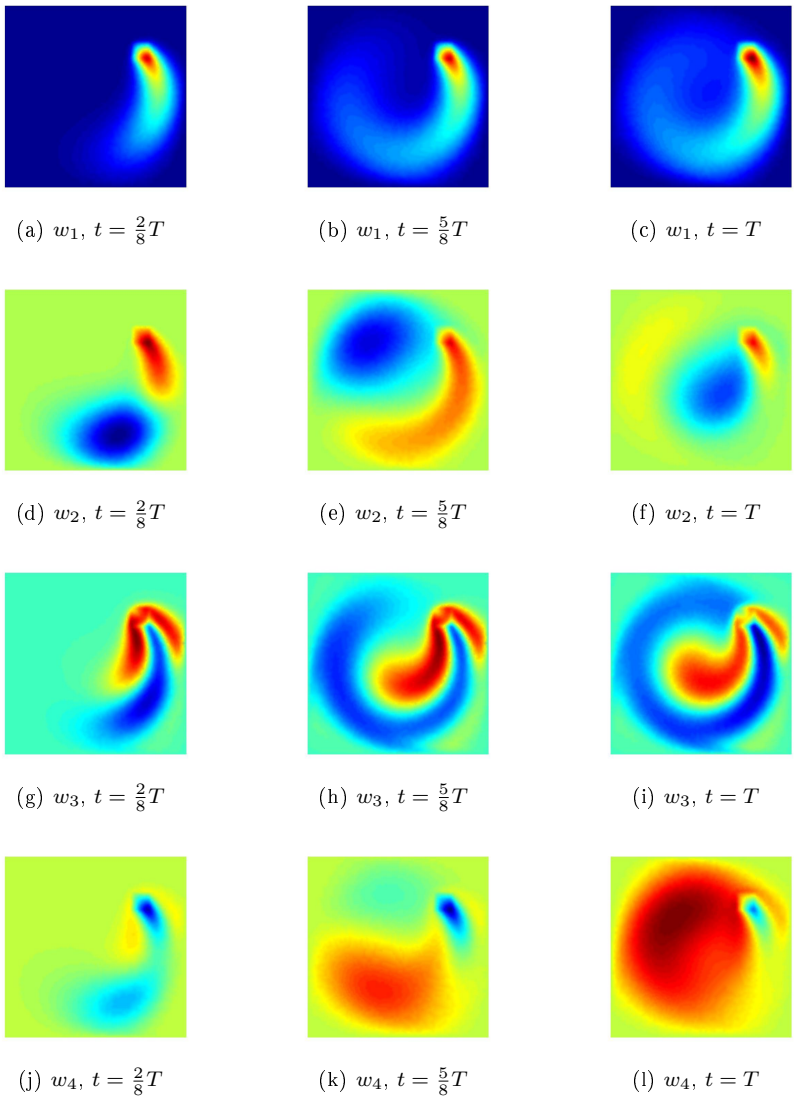
Since the deterministic modes are orthonormalized with respect to the inner product in $L^2(\Omega) \otimes L^2(I)$, the values $m_2(\lambda_i)$ reflect the contribution of the different modes to the L^2 norm of the solution:

$$\|u_M\|^2_{L^2(\Omega \times I \times \Xi)} = E(< u_M, u_M >_{L^2(\Omega \times I)}) = \sum_{i=1}^{M} m_2(\lambda_i)$$

We observe a global decrease in the contribution of the modes to the norm of the decomposition u_M. However, we notice that the convergence is not monotonic.

5.2.3 Convergence of the generalized spectral decomposition

We here study the convergence of the GSD decomposition with respect to the order M of the decomposition.



(a) w_1, $t = \frac{2}{8}T$ (b) w_1, $t = \frac{5}{8}T$ (c) w_1, $t = T$

(d) w_2, $t = \frac{2}{8}T$ (e) w_2, $t = \frac{5}{8}T$ (f) w_2, $t = T$

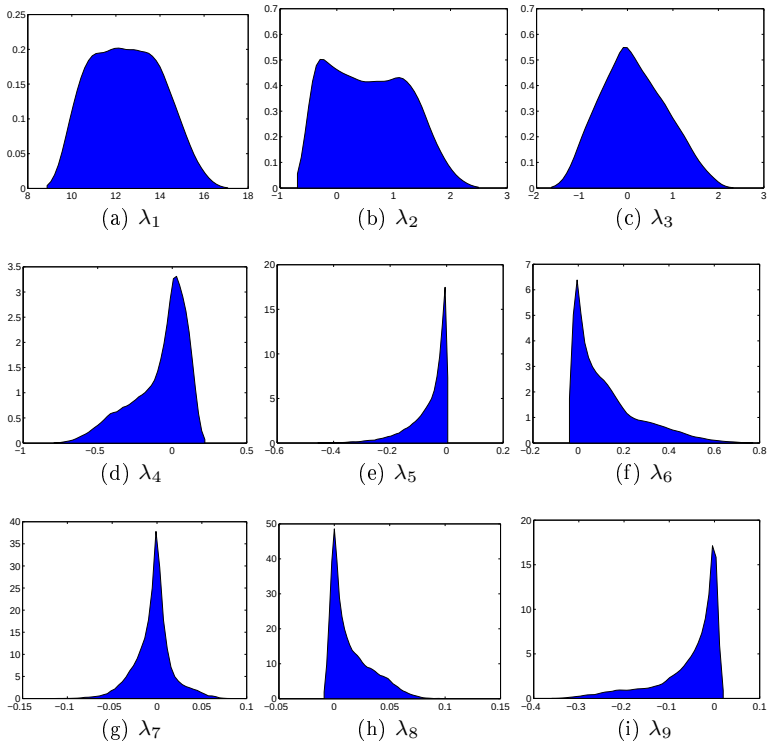
(g) w_3, $t = \frac{2}{8}T$ (h) w_3, $t = \frac{5}{8}T$ (i) w_3, $t = T$

(j) w_4, $t = \frac{2}{8}T$ (k) w_4, $t = \frac{5}{8}T$ (l) w_4, $t = T$

Fig. 4 Example 1. First 4 deterministic modes $\{w_i(\boldsymbol{x}, t)\}_{i=1}^{4}$ of the GSD decomposition built by algorithm 1 (shown at three different time steps)

Error in solution. We estimate the relative error between u_M and the semi-discretized solution $u \in \mathcal{V}_{N_x}^x \otimes \mathcal{V}_{N_t}^t \otimes \mathcal{S}$:

$$\epsilon_\gamma^M = \frac{\|u - u_M\|_\gamma}{\|u\|_\gamma} \tag{72}$$



(a) λ_1 (b) λ_2 (c) λ_3

(d) λ_4 (e) λ_5 (f) λ_6

(g) λ_7 (h) λ_8 (i) λ_9

Fig. 5 Example 1. Probability density functions of stochastic modes $\Lambda_9 = \{\lambda_i\}_{i=1}^9 = f(W_9)$ of GSD decomposition u_9

We introduce two different norms $\|\cdot\|_\gamma$ defined as follows

$$\|u\|_{L^2(\boldsymbol{\Xi};L^2(\Omega\times I))} = E\left(\|u(\boldsymbol{\xi})\|^2_{L^2(\Omega\times I)}\right)^{1/2} \tag{73}$$

$$\|u\|_{L^\infty(\boldsymbol{\Xi};L^2(\Omega\times I))} = \sup_{\boldsymbol{\xi}\in\boldsymbol{\Xi}} \|u(\boldsymbol{\xi})\|_{L^2(\Omega\times I)} \tag{74}$$

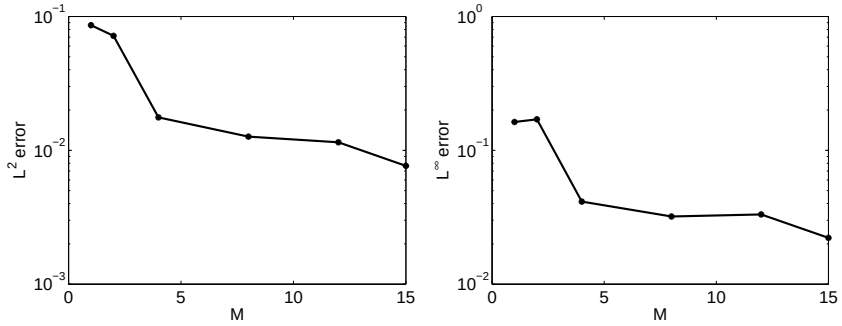
and we denote the corresponding relative errors (72) by ϵ_2^M and ϵ_∞^M respectively. These two norms are estimated by Monte-Carlo simulations:

$$\|v\|^2_{L^2(\boldsymbol{\Xi};L^2(\Omega\times I))} \approx \frac{1}{Q}\sum_{q=1}^{Q} \|v(\boldsymbol{\xi}^{(q)})\|^2_{L^2(\Omega\times I)} \tag{75}$$

$$\|v\|_{L^\infty(\boldsymbol{\Xi};L^2(\Omega\times I))} \approx \sup_{q\in\{1,\ldots,Q\}} \|v(\boldsymbol{\xi}^{(q)})\|_{L^2(\Omega\times I)} \tag{76}$$

where the $\{\boldsymbol{\xi}^{(q)}\}_{q=1}^Q$ are Q samplings of random variables $\boldsymbol{\xi}$. The reference values $u(\boldsymbol{\xi}^{(q)})$ are obtained by solving the corresponding deterministic problems with a classical deterministic numerical solution technique. Here, we take $Q = 100$, which leads to a good estimation of error indicators. Figure 6 shows the convergence with M of error indicators ϵ_γ^M. We observe a good convergence with M in the L^2-norm (error less

than 10^{-2} for $M = 15$) and also in the L^∞-norm (error 2.10^{-2} for $M = 15$). The good convergence in the L^∞-norm indicates that with a low order M, the approximation $u_M(\boldsymbol{\xi})$ is relatively good for almost every elementary events $\boldsymbol{\xi} \in \boldsymbol{\Xi}$.



(a) $L^2(\boldsymbol{\Xi}; L^2(\Omega \times I))$-norm (b) $L^\infty(\boldsymbol{\Xi}; L^2(\Omega \times I))$-norm

Fig. 6 Example 1. Convergence with M of u_M. Relative errors ϵ_2^M and ϵ_∞^M estimated with Monte-Carlo simulations.

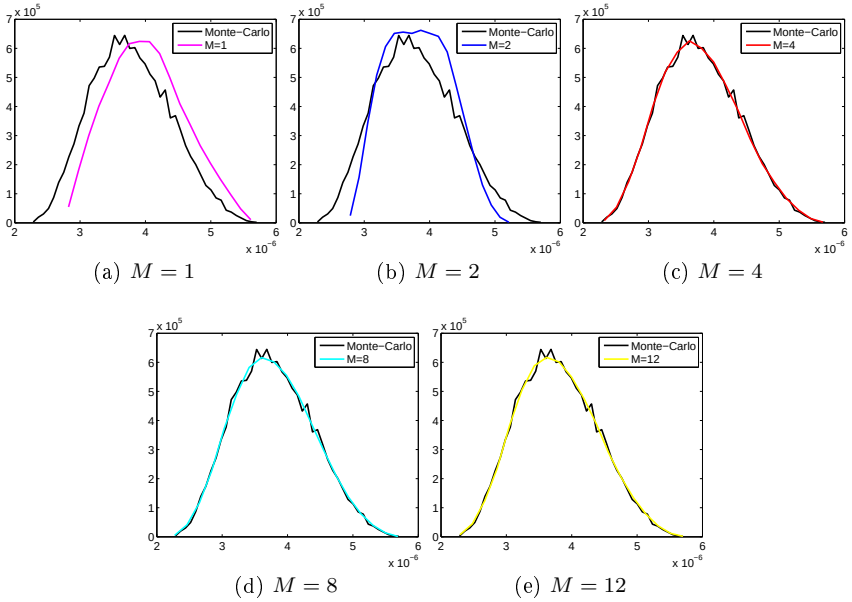
Error on quantities of interest. In order to further analyze the convergence, we focus on two quantities of interest:

$$Q_1(u)(t, \boldsymbol{\xi}) = \int_{\Omega_2} u(\boldsymbol{x}, t, \boldsymbol{\xi})\, dx$$
$$Q_2(u)(\boldsymbol{\xi}) = \int_I \int_{\Omega_2} u(\boldsymbol{x}, t, \boldsymbol{\xi})\, dx\, dt = \int_I Q_1(u)(t, \boldsymbol{\xi})\, dt$$

where $\Omega_2 = (0.2, 0.3) \times (0.2, 0.3) \subset \Omega$ is a subdomain shown on Figure 1. Let us note that Q_2 is a random variable and that Q_1 is a stochastic process in time. Figure 7 shows the convergence with M of the probability density function (pdf) of $Q_2(u_M)$. The reference pdf is computed with a classical Monte-Carlo method with $30,000$ samples (resolution of $30,000$ advection-diffusion-reaction deterministic problems). On Figure 8, we observe the convergence with M of the mean $\mu_{Q_2}^M$ and standard deviation $\sigma_{Q_2}^M$ of $Q_2(u_M)$. The plots indicate the relative error of these statistical quantities with respect to reference values obtained with the Monte-Carlo method. We observe a very quick convergence with M (although non monotonic) of the quantity of interest Q_2.

On Figure 9, we observe the convergence with M of the mean $\mu_{Q_1}^M(t)$ and standard deviation $\sigma_{Q_1}^M(t)$ of $Q_1(u_M)(t, \boldsymbol{\xi})$, which are time functions. The plots indicate the relative error with respect to reference values obtained with the Monte-Carlo method, the error being computed in the $L^2(I)$-norm. We observe a very quick convergence with M of these statistical quantities (relative error less than 10^{-2} with $M = 10$). On Figure 10, we observe the convergence with M of the 99.9% quantiles of $Q_1(u_M)(t, \boldsymbol{\xi})$. These quantiles (which are time functions) represent the envelope such that the probability of $Q_1(u_M)(t, \boldsymbol{\xi})$ being inside this envelope is 99.9%. We also observe a very good approximation of these quantiles with a low order decomposition ($M \approx 12$).

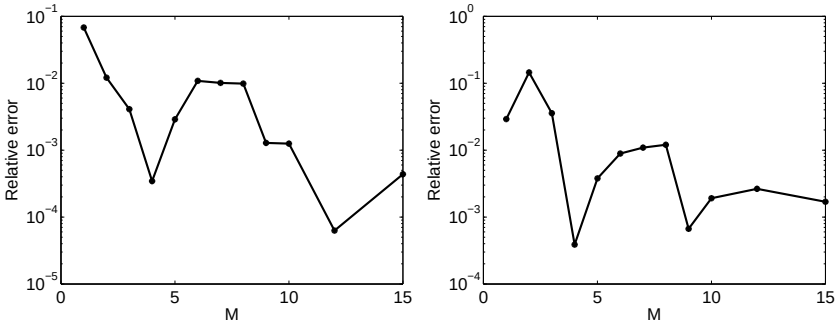
Let us recall that only M classical deterministic problems have to be solved in order to compute an order M generalized spectral decomposition. This low number



(a) $M = 1$ (b) $M = 2$ (c) $M = 4$

(d) $M = 8$ (e) $M = 12$

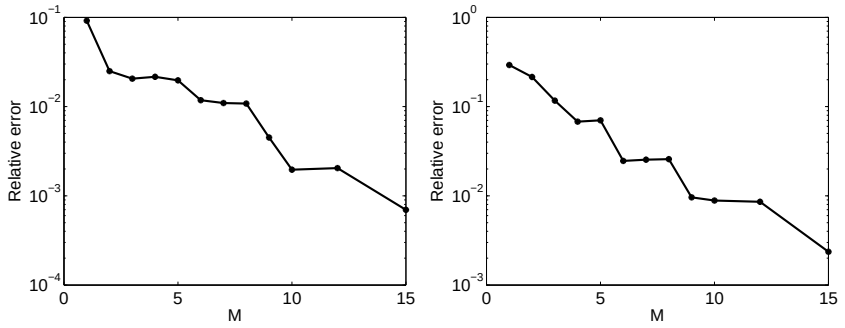
Fig. 7 Example 1. Convergence with M of the probability density function of the quantity of interest $Q_2(u_M)(\boldsymbol{\xi})$. Reference computed with Monte-Carlo.



(a) $\mu^M_{Q_2}$ (b) $\sigma^M_{Q_2}$

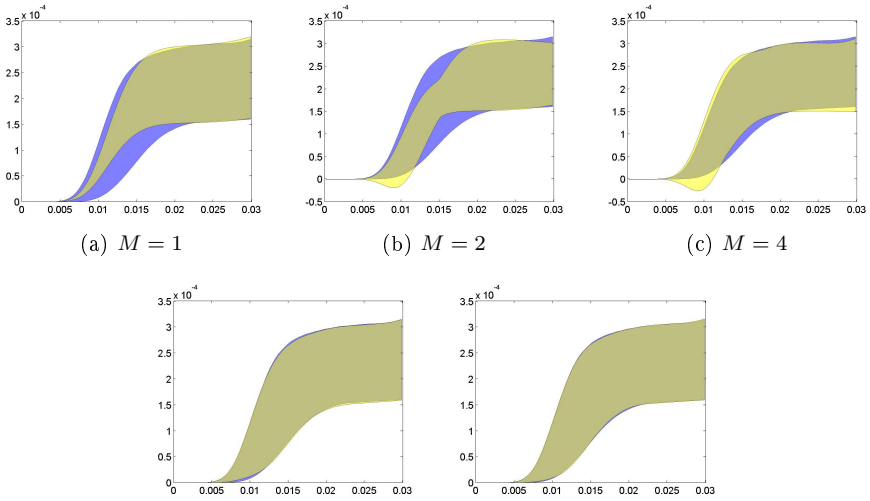
Fig. 8 Example 1. Convergence with M of the mean $\mu^M_{Q_2}$ (a) and standard deviation $\sigma^M_{Q_2}$ (b) of the quantity of interest $Q_2(u_M)(\boldsymbol{\xi})$. Relative error with respect to the reference Monte-Carlo simulations.

of deterministic problems to be solved must be compared with the huge number of deterministic simulations required by classical sampling techniques such as Monte-Carlo.



(a) $\mu^M_{Q_1}(t)$ (b) $\sigma^M_{Q_1}(t)$

Fig. 9 Example 1. Convergence with M of the mean $\mu^M_{Q_1}(t)$ (a) and standard deviation $\sigma^M_{Q_1}(t)$ (b) of the quantity of interest $Q_1(u_M)(t,\boldsymbol{\xi})$. Relative error (in $L^2(I)$-norm) with respect to the reference Monte-Carlo simulations.



(a) $M = 1$ (b) $M = 2$ (c) $M = 4$

(d) $M = 8$ (e) $M = 12$

Fig. 10 Example 1. Convergence with M of the 99.9% quantiles of the quantity of interest $Q_1(u_M)(t,\boldsymbol{\xi})$. Reference (in blue) computed by Monte-Carlo simulation.

5.3 Proper generalized decomposition for the solution of stochastic algebraic equations: separated representation at the stochastic level

We now illustrate the behavior of the Proper Generalized Decomposition (PGD) technique introduced in section 4 for the solution of the stochastic algebraic equations required in the construction of the generalized spectral decomposition $u_M = W_M \cdot \Lambda_M$ (steps 5 and 7 of algorithm 1). These stochastic algebraic equations correspond to problems of type $\lambda = f(w_i)$ and $\Lambda_M = f(W_M)$ (see section 3.5.3 for the definition of these problems). We use the following tensor product structure of the probability space: $\boldsymbol{\Xi} = \Xi_1 \times \ldots \times \Xi_4$, $P_{\boldsymbol{\xi}} = P_{\xi_1} \otimes \ldots \otimes P_{\xi_4}$, where P_{ξ_i} is the uniform probability measure on Ξ_i. The stochastic function space $\mathcal{S}$ has the following tensor product struc-

ture: $\mathcal{S} = \mathcal{S}_1 \otimes \ldots \otimes \mathcal{S}_4$, with $\mathcal{S}^i = L^2(\Xi_i, dP_{\xi_i})$, and we introduce an approximation space $\mathcal{S}_P = \otimes_{i=1}^4 \mathcal{S}^i_{P_i}$ as detailed in section 5.1.

5.3.1 Solution of problems $\lambda = f(w)$

We first analyze the solution of problems $\lambda = f(w_i)$ for the different modes i, corresponding to step 5 of algorithm 1). These problems correspond to the solution of equation (33) which can be seen as a Galerkin projection of the initial stochastic problem on the 1-dimensional deterministic basis spanned by w_i. Let us denote by $A(\boldsymbol{\xi})\lambda(\boldsymbol{\xi}) = B(\boldsymbol{\xi})$ the strong-stochastic form of these problems. We use the algorithm 3 for the approximate solution of these problems. For the updating step (steps 5 to 7), we use an updating along each stochastic dimension, *i.e.* $J_{update} = \{1, \ldots, r\}$, and a number of updates N_{update} which will be indicated later. This algorithm leads to the construction of the following order Z decomposition of stochastic function $\lambda \in \mathcal{S} \simeq \mathbb{R} \otimes \mathcal{S}^1 \otimes \ldots \otimes \mathcal{S}^4$:

$$\lambda(\boldsymbol{\xi}) \approx \lambda_Z(\boldsymbol{\xi}) = \sum_{i=1}^{Z} \phi_i^0 \phi_i^1(\xi_1) \ldots \phi_i^4(\xi_4)$$

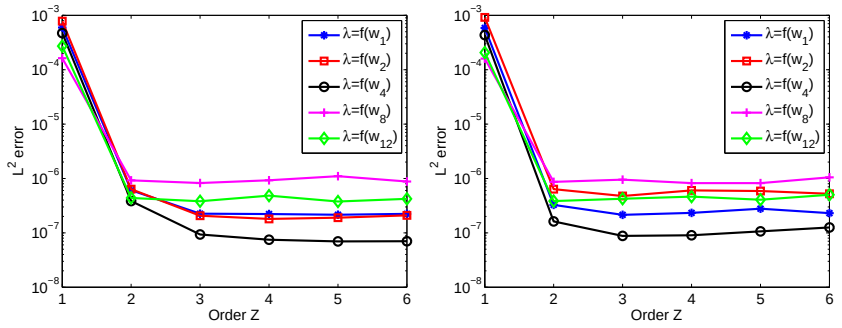
with $\phi_i^0 \in \mathbb{R}$ and $\phi_i^j \in \mathcal{S}^j_{P_j}$. In order to analyze the convergence of the decomposition, we introduce the following error indicator in L^2-norm:

$$\epsilon^Z = \frac{\|\lambda - \lambda_Z\|_{L^2(\boldsymbol{\Xi})}}{\|\lambda\|_{L^2(\boldsymbol{\Xi})}}, \tag{77}$$

with $\|\lambda\|_{L^2(\boldsymbol{\Xi})} = E(\lambda(\boldsymbol{\xi})^2)^{1/2}$. The L^2-norm is estimated with Monte-Carlo simulations:

$$\|\lambda\|^2_{L^2(\boldsymbol{\Xi})} \approx \frac{1}{Q} \sum_{q=1}^{Q} \lambda(\boldsymbol{\xi}^{(q)})^2, \tag{78}$$

where the $\{\boldsymbol{\xi}^{(q)}\}_{q=1}^Q$ are Q samplings of random variables $\boldsymbol{\xi}$. The reference values are defined by $\lambda(\boldsymbol{\xi}^{(q)}) = A(\boldsymbol{\xi}^{(q)})^{-1} B(\boldsymbol{\xi}^{(q)})$. Here, we take $Q = 100$. Let us note that error indicator ϵ^Z evaluates the distance between the approximate solution $\lambda_Z \in \mathcal{S}_P$ and the strong stochastic solution $\lambda \in \mathcal{S}$. It then takes into account two contributions of errors: the approximation error (introduction of $\mathcal{S}_P \subset \mathcal{S}$) and the error due to the separated representation technique (truncation error). In this example, the approximation error is negligible compared to the truncation error (sufficiently high polynomial degree used for $\mathcal{S}_P$). On Figure 11, we illustrate the convergence with Z of λ_Z for different problems $\lambda_Z \approx f(w_i)$. We plot the convergence for a parameter $N_{update} = 0$ or 1 in the algorithm 3. We observe a very fast convergence in $Z = 2$ or 3 modes for each mode and we do not observe any significant influence of parameter N_{update}. The error value which is reached after $Z = 2$ or 3 corresponds to the lowest numerical precision which can be reached with separated representation technique (corresponding to an error about 10^{-8} in algebraic norms). These results indicate that for problems $\lambda = f(w_i)$, a very good accuracy is obtained with Z only equal to 1 or 2 (*i.e.* the λ admits a very low order separated representation).



(a) $N_{update} = 0$ (b) $N_{update} = 1$

Fig. 11 Example 1. Approximate solution of stochastic algebraic equations $\lambda = f(w_i)$ with algorithm 3. Convergence with Z of λ_Z, in L^2-norm, for $N_{update} = 0$ (a) and $N_{update} = 1$ (b) ($J_{update} = \{1, \ldots, r\}$).

5.3.2 Solution of problems $\Lambda_M = f(W_M)$

We now focus on the solution of the system of stochastic algebraic equations $\Lambda_M = f(W_M)$, corresponding to step 7 of algorithm 1 (update of stochastic functions). This problem is solved with algorithm 3. For the updating step (steps 5 to 7 of algorithm 3), we use an updating along each dimension, *i.e.* $J_{update} = \{0, \ldots, r\}$, and a number of updates N_{update} which will be indicated later. This problem corresponds to the solution of equation (40) which can be seen as a Galerkin projection of the initial SPDE on the M-dimensional deterministic basis spanned by $W_M = \{w_i\}_{i=1}^M$ (reduced basis of space-time functions).

Remark 7 We will test the algorithm 3 for different orders M. However, let us recall that in practise, when using the Arnoldi-type algorithm 1, problem $\Lambda_M = f(W_M)$ is solved only one time, after the construction of a set of deterministic functions $\{w_i\}_{i=1}^M$. More precisely, if the Arnoldi procedure is restarted, it is solved one time after the construction of each Krylov subspace.

We assimilate $\Lambda_M \in (\mathcal{S})^M$ with a random vector $\boldsymbol{\Lambda} \in \mathbb{R}^M \otimes \mathcal{S}$ and we denote by $\mathbf{A}(\boldsymbol{\xi})\boldsymbol{\Lambda}(\boldsymbol{\xi}) = \mathbf{B}(\boldsymbol{\xi})$ the strong-stochastic form of problem $\Lambda_M = f(W_M)$. We use the algorithm 3 for the approximate solution of this problem. It leads to the construction of the following order Z decomposition of stochastic functions $\Lambda_M \in (\mathcal{S})^M \equiv \boldsymbol{\Lambda} \in \mathbb{R}^M \otimes \mathcal{S}^1 \otimes \ldots \otimes \mathcal{S}^4$:

$$\boldsymbol{\Lambda}(\boldsymbol{\xi}) \approx \boldsymbol{\Lambda}_Z(\boldsymbol{\xi}) = \sum_{i=1}^{Z} \boldsymbol{\phi}_i^0 \phi_i^1(\xi_1) \ldots \phi_i^4(\xi_4),$$

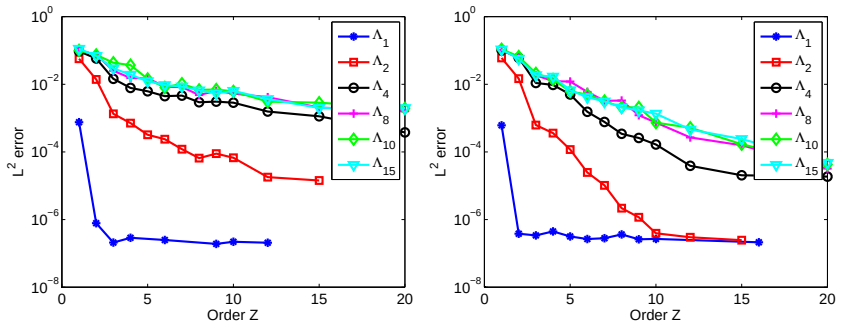
with $\boldsymbol{\phi}_i^0 \in \mathbb{R}^M$ and $\phi_i^j \in \mathcal{S}_{P_j}^j$. In order to analyze the convergence of the decomposition, we introduce the following error indicator in L^2-norm:

$$\epsilon_M^Z = \frac{\|\boldsymbol{\Lambda} - \boldsymbol{\Lambda}_Z\|_{\mathbb{R}^M \otimes L^2(\boldsymbol{\Xi})}}{\|\boldsymbol{\Lambda}\|_{\mathbb{R}^M \otimes L^2(\boldsymbol{\Xi})}} \tag{79}$$

with $\|\boldsymbol{\Lambda}\|_{\mathbb{R}^M \otimes L^2(\boldsymbol{\Xi})} = E(\|\boldsymbol{\Lambda}(\boldsymbol{\xi})\|^2_{\mathbb{R}^M})^{1/2}$. The L^2-norm is estimated with Monte-Carlo simulations

$$\|\boldsymbol{\Lambda}\|^2_{\mathbb{R}^M \otimes L^2(\boldsymbol{\Xi})} \approx \frac{1}{Q} \sum_{q=1}^{Q} \|\boldsymbol{\Lambda}(\boldsymbol{\xi}^{(q)})\|^2_{\mathbb{R}^M} \tag{80}$$

where the $\{\boldsymbol{\xi}^{(q)}\}_{q=1}^{Q}$ are Q samplings of random variables $\boldsymbol{\xi}$. Reference values $\boldsymbol{\Lambda}(\boldsymbol{\xi}^{(q)}) = \mathbf{A}(\boldsymbol{\xi}^{(q)})^{-1}\mathbf{B}(\boldsymbol{\xi}^{(q)})$ are obtained by solving a simple system of deterministic equations. Here, we take $Q = 100$. As mentioned in the previous section, the approximation error, due to the introduction of $\mathcal{S}_P \subset \mathcal{S}$, is here negligible. Then, ϵ_M^Z quantifies the truncation error (for a truncation order Z). On Figure 12, we illustrate the convergence with Z of $\boldsymbol{\Lambda}_Z$ for different problems $\Lambda_{M,Z} \approx f(W_M)$. We plot the convergence for a parameter $N_{update} = 0$ or 1 in the algorithm 3. We here notice for $M > 1$ a significant influence of the updating step in algorithm 3. Indeed, for a given order Z, the accuracy of the decomposition $\boldsymbol{\Lambda}_Z$ obtained with $N_{update} = 1$ is better than the one obtained without update ($N_{update} = 0$).



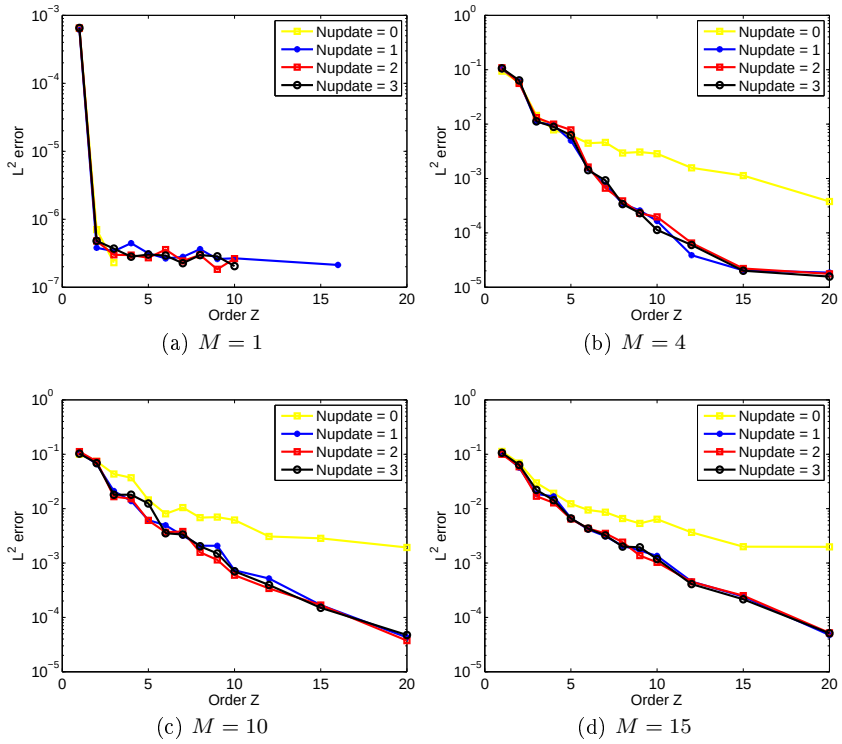
(a) $N_{update} = 0$ (b) $N_{update} = 1$

Fig. 12 Example 1. Approximate solution of the system of stochastic algebraic equations $\Lambda_M = f(W_M)$ with algorithm 3. Convergence with Z of $\Lambda_{M,Z}$, in L^2-norm, for $N_{update} = 0$ (a) and $N_{update} = 1$ (b).

On figure 13, we test the influence of the number of updates N_{update}. As mentioned in section 4.4, we observe in this example that performing more than 1 update ($N_{update} > 1$) does not improve the accuracy of the decomposition obtained with $N_{update} = 1$. We observe that when increasing M, a higher order Z is required for reaching a given accuracy. However, this order Z is always very small compared to the dimension of the stochastic approximation space $P = 1296$. A L^2 error less than 10^{-2} is obtained with only $Z = 5$ whatever the order M.

The overall methodology can be seen as a technique for constructing automatically a very low dimensional stochastic approximation space $\mathcal{S}_Z = span\{\Psi_i\}_{i=1}^{Z} \subset \mathcal{S}_P$, with $\Psi_i(\boldsymbol{\xi}) = \prod_{j=1}^{4} \phi_i^j(\xi_j)$, which is well adapted to the representation of the solution u of the present stochastic problem. Here, $Z \approx 5$ only is sufficient to reach a good approximation.

Let us finally note that computational costs associated with the overall numerical strategy are very low. For example, for the construction of a GSD decomposition u_M of



(a) $M = 1$ (b) $M = 4$

(c) $M = 10$ (d) $M = 15$

Fig. 13 Example 1. Approximate solution of the system of stochastic algebraic equations $\Lambda_M = f(W_M)$ with algorithm 3 for different orders M. Convergence with Z of $\Lambda_{M,Z}$, in L^2-norm. Influence of parameter N_{update} of the algorithm.

order $M = 15$, it took a few seconds on a simple laptop. As illustrated in this example, u_{15} provides a very good approximation of u and of the quantities of interest.

6 Example 2: stationary advection diffusion reaction equation

6.1 Formulation of the problem and discretization

Formulation of the problem. We consider a stationary advection diffusion reaction equation defined on a spatial domain $\Omega = (0,1) \times (0,1)$ (see figure 1). It is a stationary version of example 1 where the only source of uncertainty comes from the diffusion coefficient which is chosen as a random field, depending on a set of random variables $\boldsymbol{\xi} \in \boldsymbol{\Xi}$. The solution field $u(\boldsymbol{x}, \boldsymbol{\xi})$, defined on $\Omega \times \boldsymbol{\Xi}$ verifies

$$-\nabla \cdot (\mu(\boldsymbol{x}, \boldsymbol{\xi}) \nabla u) + c \cdot \nabla u + \kappa u = f \quad on \quad \Omega \tag{81a}$$

$$u = 0 \quad on \quad \partial\Omega \tag{81b}$$

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

$f = 100 I_{\Omega_1}$ (see figure 1), where $\Omega_1 = (0.7, 0.8) \times (0.7, 0.8) \subset \Omega$, with I_{Ω_1} the indicator function of Ω_1. $\mu(\boldsymbol{x}, \boldsymbol{\xi})$ is a random field defined by

$$\mu(\boldsymbol{x}, \boldsymbol{\xi}) = \mu_0 + \sum_{i=1}^{40} \sqrt{\sigma_i} \mu_i(\boldsymbol{x}) \xi_i \tag{82}$$

where $\mu_0 = 1$ is the mean value of μ, where the $\xi_i \in U(-1, 1)$ are mutually independent uniform random variables and where the $\mu_i(\boldsymbol{\xi})$ are a set of $L^2(\Omega)$-orthonormal spatial functions. These spatial functions are plotted in figure 14. The associated amplitudes $\sqrt{\sigma_i}$ are plotted on figure 15. The $m = 40$ random parameters $\boldsymbol{\xi} = (\xi_i)_{i=1}^m$ define a probability space $(\boldsymbol{\Xi}, \mathcal{B}, P_{\boldsymbol{\xi}})$, with $\boldsymbol{\Xi} = (-1, 1)^m$ and $P_{\boldsymbol{\xi}}$ the uniform probability measure on Borel σ-algebra $\mathcal{B}$.

Remark 8 The couples $(\mu_i, \sigma_i) \in L^2(\Omega) \times \mathbb{R}^+$ are chosen as the 40 dominant eigenpairs of eigenproblem $T(\mu_i) = \sigma_i \mu_i$, where T is the kernel operator

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

with $\alpha(\boldsymbol{x}, \boldsymbol{y}) = 0.2^2 \exp(-\frac{\|\boldsymbol{x}-\boldsymbol{y}\|^2}{0.3^2})$. The expression (82) for $\mu(\boldsymbol{x}, \boldsymbol{\xi})$ then corresponds to a truncated version of a homogeneous random field with mean 1, standard deviation $0.2/\sqrt{3}$ and exponential square covariance function with correlation length 0.3.

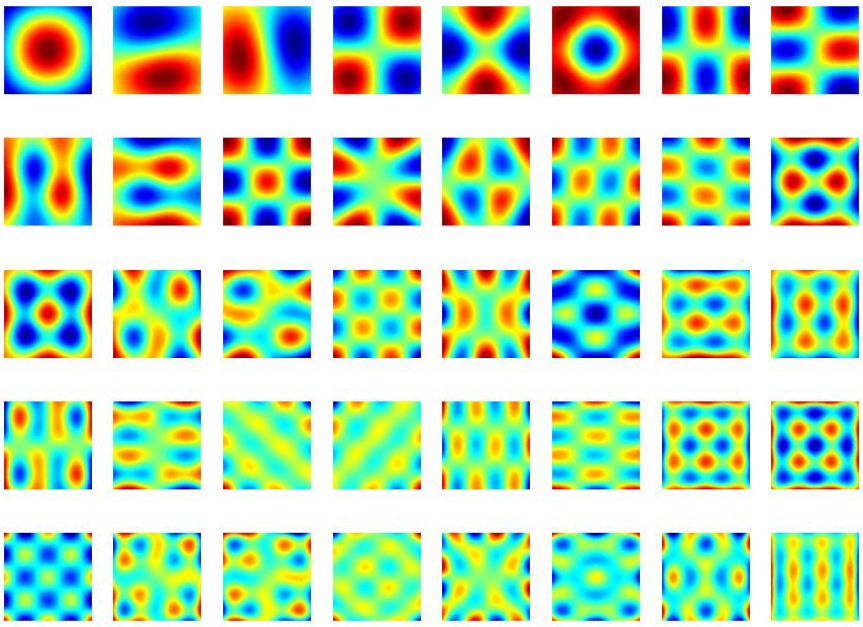


Fig. 14 Example 2. Spatial modes $\{\mu_i(\boldsymbol{x})\}_{i=1}^{40}$ of the decomposition (82) of random field $\mu(\boldsymbol{x}, \boldsymbol{\xi})$ (modes sorted from left to right and top to bottom)

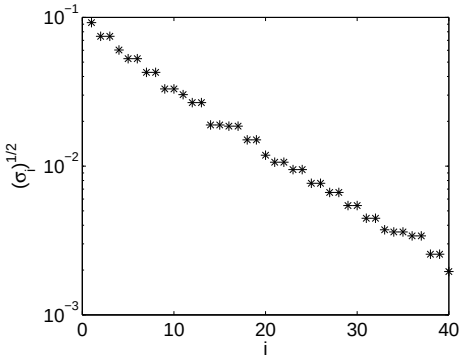


Fig. 15 Example 2. Amplitudes $\sqrt{\sigma_i}$ of the modes of the decomposition (82) of random field $\mu(\boldsymbol{x}, \boldsymbol{\xi})$

Weak formulation. We introduce the weak formulation (4) of problem (81) with the following definition of function spaces

$$\mathcal{V} = H_0^1(\Omega), \quad \mathcal{S} = L^2(\boldsymbol{\Xi}, \mathcal{B}, P_{\boldsymbol{\xi}}) \tag{83}$$

and the following definitions of bilinear and linear forms:

$$\begin{aligned} a(u, v; \boldsymbol{\xi}) = & \int_\Omega \mu(\boldsymbol{x}, \boldsymbol{\xi}) \nabla u \cdot \nabla v \, dx \\ & + \int_\Omega c \cdot \nabla u \; v \, dx + \int_\Omega \kappa u \; v \, dx \end{aligned} \tag{84}$$

$$l(v) = \int_\Omega v f \, dx \tag{85}$$

Discretization. At the space level, we introduce a finite element approximation space $\mathcal{V}_N \subset \mathcal{V}$ with dimension $N = 4435$. The finite element mesh composed of 3-nodes triangles is shown on figure 3. At the stochastic level, we introduce different approximation strategies, associated with different separations of function space $\mathcal{S} \simeq \mathcal{S}^1 \otimes \ldots \otimes \mathcal{S}^r$, where $\mathcal{S}^j = L^2(\boldsymbol{\Xi}_j, \mathcal{B}_j, P_{\boldsymbol{\xi}_j})$ and $\boldsymbol{\Xi}_j = (-1, 1)^{m^*}$, with $r \times m^* = m = 40$. We introduce complete polynomial approximation spaces $\mathcal{S}^j_{P^*} = \mathbb{P}_p(\boldsymbol{\Xi}_j)$ of degree $p = 4$, with $P^* = \frac{(p+m^*)!}{p!}$ and define

$$\mathcal{S} \supset \mathcal{S}_P \simeq \mathcal{S}^1_{P^*} \otimes \ldots \otimes \mathcal{S}^r_{P^*}$$

We will take for the reference computation $(r, m^*) = (8, 5)$. The associated dimension of $\mathcal{S}_P$ is then $P = (P^*)^r \approx 6.10^{16}$. Let us note that with such a dimension, a direct computation of the stochastic Galerkin projection is unaffordable in this example. The overall methodology proposed in this article (sections 3 and 4) allows obtaining an approximation of this Galerkin projection.

6.2 Generalized spectral decomposition

In this section, we apply the GSD algorithm 1 (Arnoldi-type algorithm) for the *a priori* construction of a decomposition of the solution

$$u(\boldsymbol{x},\boldsymbol{\xi}) \approx u_M(\boldsymbol{x},\boldsymbol{\xi}) = \sum_{i=1}^{M} w_i(\boldsymbol{x})\lambda_i(\boldsymbol{\xi}) := W_M \cdot \Lambda_M$$

where the $w_i \in \mathcal{V}_N$ are spatial modes and the $\lambda_i \in \mathcal{S}_P$ are stochastic modes. In this section, we only focus on the properties of the GSD method introduced in section 3 (for deterministic/stochastic separation). We do not focus on the solution of stochastic algebraic equations and we consider that these equations are solved with a good accuracy (error less than the error associated with the truncation order M of the GSD). The solution of these stochastic algebraic equations with the algorithm proposed in section 4 will be analyzed in the following section 6.3.

6.2.1 Algorithm and computational aspects

We recall that for building a decomposition of order M, the Arnoldi-type algorithm 1 requires the solution of M classical deterministic PDEs (problems $w_i = F(\lambda)$), M stochastic algebraic equations (problems $\lambda = f(w_i)$) and a system of stochastic algebraic equations (problem $\Lambda_M = f(W_M)$) for the update of stochastic functions. The set of M deterministic modes w_i are computed by solving only M uncoupled deterministic problems $w_i = F(\lambda)$ for different $\lambda \in \mathcal{S}_P$ (equation (29)). These problems correspond to classical stationary advection diffusion reaction problems associated with different deterministic parameters $\mu_\lambda(\boldsymbol{x}) = E(\mu(\boldsymbol{x},\boldsymbol{\xi})\lambda(\boldsymbol{\xi})^2)$, $c_\lambda = E(c\lambda^2) = cE(\lambda^2)$ and $\kappa_\lambda = E(\kappa\lambda^2) = \kappa E(\lambda^2)$ (respectively for the diffusion, advection and reaction terms) and with a deterministic source term $f_\lambda = E(f\lambda) = fE(\lambda)$. Bilinear and linear forms in equation (29) write

$$a_\lambda(w,w^*) = \int_\Omega \mu_\lambda(\boldsymbol{x})\nabla w \cdot \nabla w^* \, dx \tag{86}$$

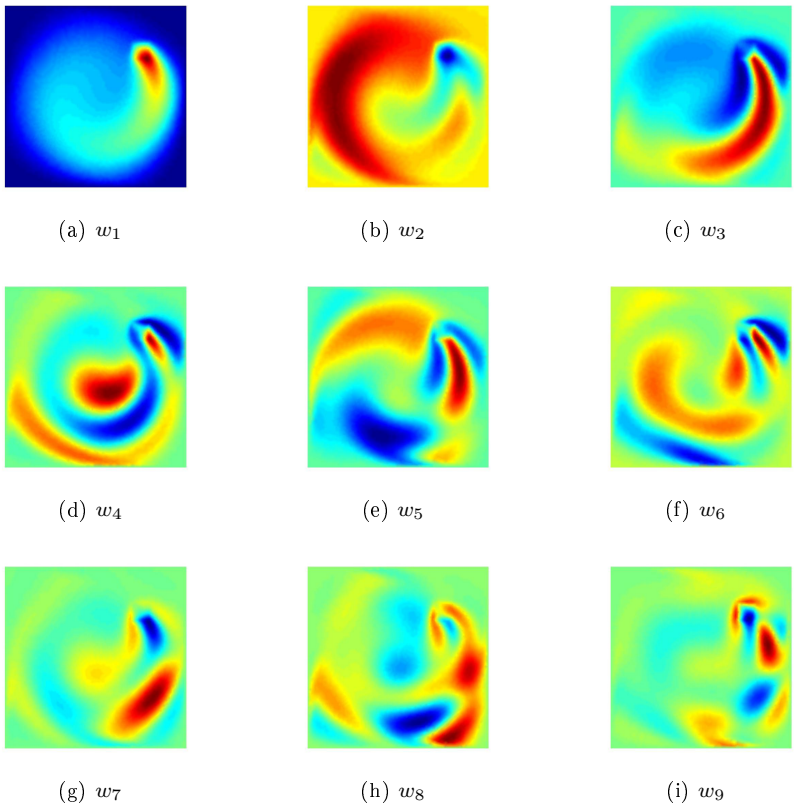
$$+ E(\lambda^2)\int_\Omega c(\boldsymbol{x}) \cdot \nabla w \; w^* \, dx + E(\lambda^2)\int_\Omega \kappa w \; w^* \, dx \tag{87}$$

$$l_\lambda(v) = E(\lambda)\int_\Omega w^* f \, dx \tag{88}$$

6.2.2 Illustration of the obtained decomposition

We here illustrate the decomposition $u_9 = W_9 \cdot \Lambda_9$ of order 9 obtained by the Arnoldi-type algorithm. Figure 16 shows the first 9 deterministic modes $\{w_i\}_{i=1}^9$. These modes are orthonormalized with respect to the natural inner product in $L^2(\Omega)$ (in the construction of generalized Krylov subspace). Figure 17 shows the stochastic modes Λ_9. In Table 2, we indicate the mean $m_1(\lambda_i) := E(\lambda_i)$ and second moment $m_2(\lambda_i) := E(\lambda_i^2)$ of each stochastic mode λ_i.

Since the deterministic modes are orthonormalized with respect to the inner product in $L^2(\Omega)$, the values $m_2(\lambda_i)$ reflect the contribution of the different modes to the



(a) w_1 (b) w_2 (c) w_3

(d) w_4 (e) w_5 (f) w_6

(g) w_7 (h) w_8 (i) w_9

Fig. 16 Example 2. Deterministic modes $\{w_i(\boldsymbol{x})\}_{i=1}^{9}$ of the GSD decomposition u_9

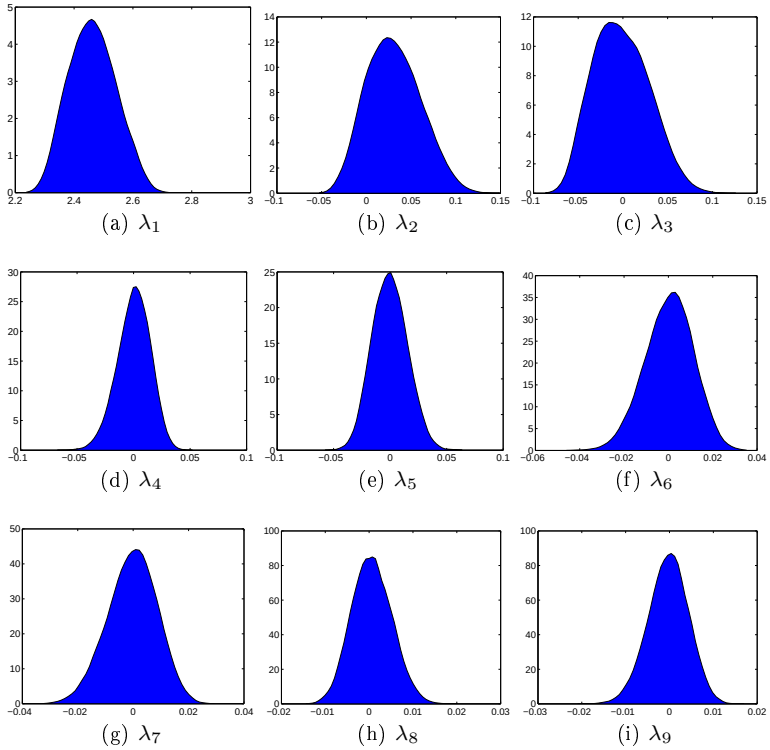
Table 2 First and second moments of random variables $\{\lambda_i\}_{i=1}^{9}$

i	$m_1(\lambda_i)$	$m_2(\lambda_i)$
1	2.4628	6.1
2	0.0307	$1.9\,10^{-3}$
3	−0.0017	$9.8\,10^{-4}$
4	0.0002	$2.1\,10^{-4}$
5	−0.0002	$2.4\,10^{-4}$
6	0.0003	$1.2\,10^{-4}$
7	−0.0003	$7.8\,10^{-5}$
8	0.0006	$2.1\,10^{-5}$
9	−0.0003	$2.1\,10^{-5}$

L^2 norm of the solution:

$$\|u_M\|^2_{L^2(\Omega\times\boldsymbol{\Xi})} = E((u_M, u_M)_{L^2(\Omega)}) = \sum_{i=1}^{M} m_2(\lambda_i)$$

We observe a global decrease in the contribution of the modes to the norm of the decomposition u_M.



(a) λ_1 (b) λ_2 (c) λ_3

(d) λ_4 (e) λ_5 (f) λ_6

(g) λ_7 (h) λ_8 (i) λ_9

Fig. 17 Example 2. Probability density functions of stochastic modes $\Lambda_9 = \{\lambda_i\}_{i=1}^9 = f(W_9)$ of GSD decomposition u_9

6.2.3 Convergence of the generalized spectral decomposition

We here study the convergence of the GSD decomposition with respect to the order M of the decomposition.

Error in solution. We estimate the relative error between u_M and the semi-discretized solution $u \in \mathcal{V}_N \otimes \mathcal{S}$:

$$\epsilon_\gamma^M = \frac{\|u - u_M\|_\gamma}{\|u\|_\gamma} \tag{89}$$

We introduce two different norms $\|\cdot\|_\gamma$ defined as follows

$$\|u\|_{L^2(\boldsymbol{\Xi};L^2(\Omega)} = E(\|u(\boldsymbol{\xi})\|^2_{L^2(\Omega)})^{1/2} \tag{90}$$

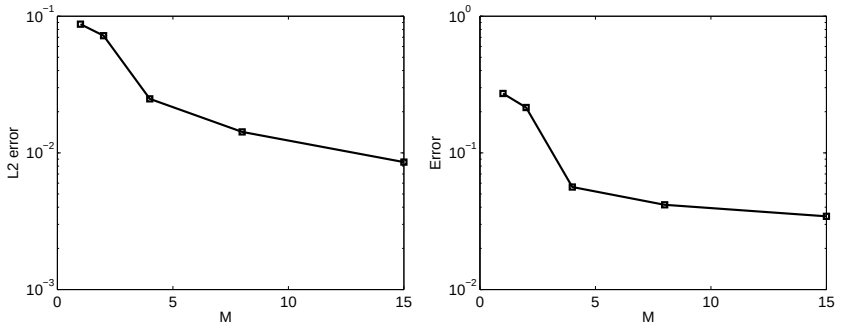
$$\|u\|_{L^\infty(\boldsymbol{\Xi};L^2(\Omega))} = \sup_{\boldsymbol{\xi}\in\boldsymbol{\Xi}} \|u(\boldsymbol{\xi})\|_{L^2(\Omega)} \tag{91}$$

and we denote the corresponding relative errors (89) by ϵ_2^M and ϵ_∞^M respectively. These two norms are estimated by Monte-Carlo simulations:

$$\|v\|^2_{L^2(\boldsymbol{\Xi};L^2(\Omega))} \approx \frac{1}{Q}\sum_{q=1}^{Q} \|v(\boldsymbol{\xi}^{(q)})\|^2_{L^2(\Omega)} \tag{92}$$

$$\|v\|_{L^\infty(\boldsymbol{\Xi};L^2(\Omega))} \approx \sup_{q\in\{1,\dots,Q\}} \|v(\boldsymbol{\xi}^{(q)})\|_{L^2(\Omega)} \tag{93}$$

where the $\{\boldsymbol{\xi}^{(q)}\}_{q=1}^Q$ are Q samplings of random variables $\boldsymbol{\xi}$. The reference values $u(\boldsymbol{\xi}^{(q)})$ are obtained by solving the corresponding deterministic problems. Here, we take $Q = 100$, which leads to a good estimation of error indicators. Figure 18 shows the convergence with M of error indicators ϵ_γ^M. We observe a good convergence with M of the L^2-norm (error less than 10^{-2} for $M = 15$) and also in the L^∞-norm (error 3.10^{-2} for $M = 15$). The good convergence in the L^∞-norm indicates that with a low order M, the approximation $u_M(\boldsymbol{\xi})$ is relatively good for almost every elementary events $\boldsymbol{\xi} \in \boldsymbol{\Xi}$ (see figure 19 for the illustration of this fact).



(a) $L^2(\boldsymbol{\Xi};L^2(\Omega))$-norm (b) $L^\infty(\boldsymbol{\Xi};L^2(\Omega))$-norm

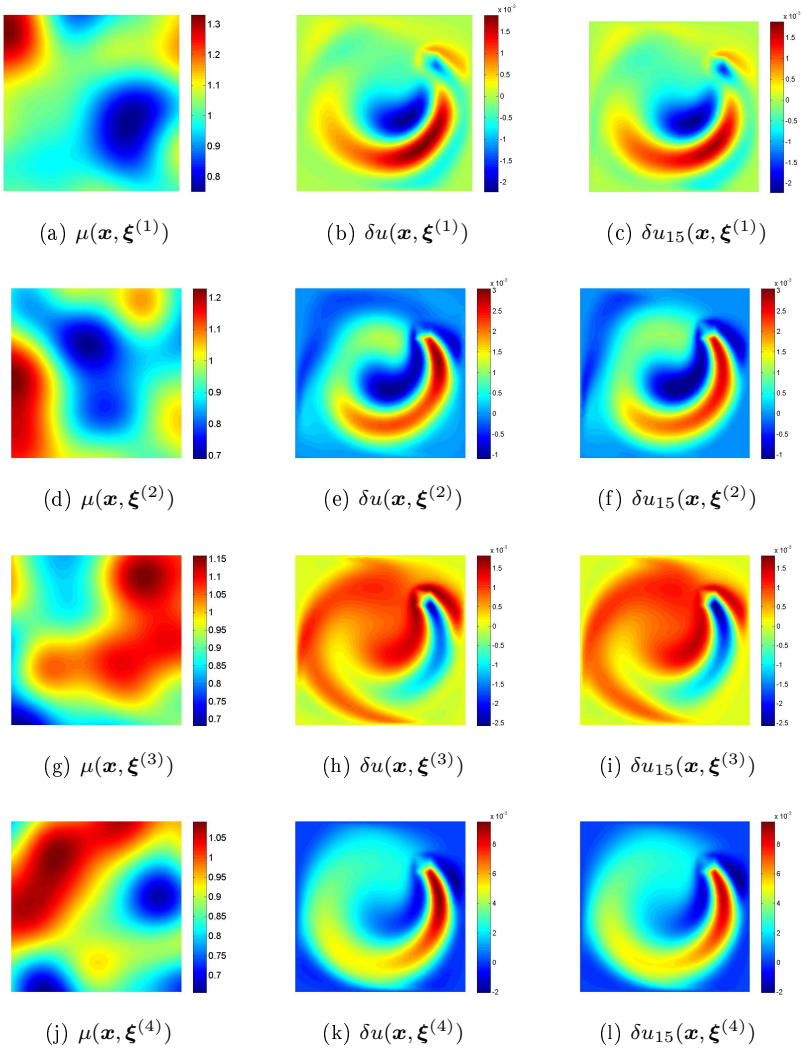
Fig. 18 Example 2. Convergence with M of u_M. Relative errors ϵ_2^M and ϵ_∞^M estimated with Monte-Carlo simulations.

Convergence of quantities of interest. In order to further analyze the convergence, we focus on a quantity of interest:

$$Q(u)(\boldsymbol{\xi}) = \int_{\Omega_2} u(\boldsymbol{x},\boldsymbol{\xi})\, dx \tag{94}$$

where $\Omega_2 = (0.2, 0.3) \times (0.2, 0.3) \subset \Omega$ is a subdomain shown on Figure 1. Let us note that Q_2 is a random variable. Figure 20 shows the convergence with M of the probability density function (pdf) of $Q(u_M)$. The reference pdf is computed with a classical Monte-Carlo method with 36,000 samples (resolution of 36,000 advection diffusion reaction deterministic problems). We observe a very good convergence with M of the quantity of interest Q_2.

On figure 21, we observe the convergence with M of the probability of the event $\{Q(u_M)(\boldsymbol{\xi}) > q\}$, *i.e.* $P_{\boldsymbol{\xi}}\{Q(u_M) > q\}$ for different values of q. We observe that the



(a) $\mu(\boldsymbol{x}, \boldsymbol{\xi}^{(1)})$ (b) $\delta u(\boldsymbol{x}, \boldsymbol{\xi}^{(1)})$ (c) $\delta u_{15}(\boldsymbol{x}, \boldsymbol{\xi}^{(1)})$

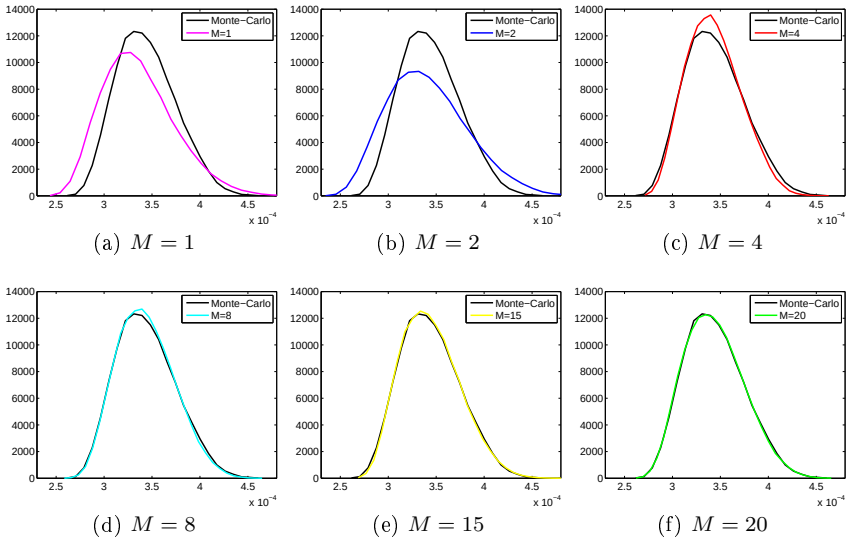
(d) $\mu(\boldsymbol{x}, \boldsymbol{\xi}^{(2)})$ (e) $\delta u(\boldsymbol{x}, \boldsymbol{\xi}^{(2)})$ (f) $\delta u_{15}(\boldsymbol{x}, \boldsymbol{\xi}^{(2)})$

(g) $\mu(\boldsymbol{x}, \boldsymbol{\xi}^{(3)})$ (h) $\delta u(\boldsymbol{x}, \boldsymbol{\xi}^{(3)})$ (i) $\delta u_{15}(\boldsymbol{x}, \boldsymbol{\xi}^{(3)})$

(j) $\mu(\boldsymbol{x}, \boldsymbol{\xi}^{(4)})$ (k) $\delta u(\boldsymbol{x}, \boldsymbol{\xi}^{(4)})$ (l) $\delta u_{15}(\boldsymbol{x}, \boldsymbol{\xi}^{(4)})$

Fig. 19 Example 2. Comparison between GSD approximation u_{15} and direct computations u for different outcomes $\boldsymbol{\xi}^{(q)}$ of random variables. Associated outcomes of diffusion coefficient μ (first column), direct simulation δu (second column), and GSD approximation δu_{15} (third column). $\delta u(\boldsymbol{x}, \boldsymbol{\xi}) = u(\boldsymbol{x}, \boldsymbol{\xi}) - u(\boldsymbol{x}, 0)$, where $u(\boldsymbol{x}, 0)$ is the solution with a mean random field $\mu = \mu_0$.

number of modes M must be increased in order to accurately predict events with lower and lower probabilities. However, we observe that a relatively low order decomposition ($M = 20$) allows to accurately predict the probability of rare events (events with a probability lower than of 10^{-3}).



(a) $M = 1$ (b) $M = 2$ (c) $M = 4$

(d) $M = 8$ (e) $M = 15$ (f) $M = 20$

Fig. 20 Example 2. Convergence with M of the probability density function of the quantity of interest $Q(u_M)(\boldsymbol{\xi})$. Reference computed with Monte-Carlo.

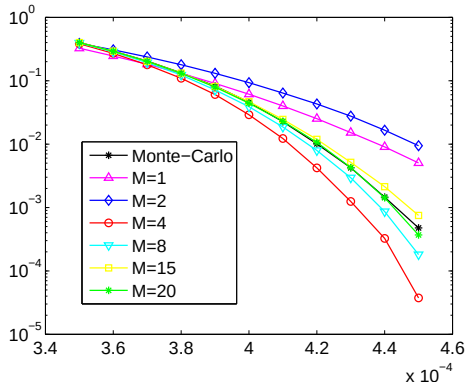


Fig. 21 Example 2. Convergence with M of the probability $P_{\boldsymbol{\xi}}\{Q(u_M) > q\}$ for different values q. Reference computed with Monte-Carlo.

6.3 Proper Generalized Decomposition for the solution of stochastic algebraic equations: separated representation at the stochastic level

We now illustrate the behavior of the Proper Generalized Decomposition (PDG) technique introduced in section 4 for the solution of the stochastic algebraic equations required in the construction of the generalized spectral decomposition $u_M = W_M \cdot \Lambda_M$.

6.3.1 Solution of problems $\lambda = f(w)$

We first analyze the solution of problems $\lambda = f(w_i)$ for the different modes i, corresponding to step 5 of algorithm 1. These problems correspond to the solution of equa-

tion (33) which can be seen as a Galerkin projection of the initial stochastic problem on the 1-dimensional deterministic basis spanned by w_i. Let us denote by $A(\boldsymbol{\xi})\lambda(\boldsymbol{\xi}) = B(\boldsymbol{\xi})$ the strong-stochastic form of these problems. We use the algorithm 3 for the approximate solution of these problems. For the updating step (steps 5 to 7), we use an updating along each stochactic dimension, *i.e.* $J_{update} = \{1, \ldots, r\}$, and a number of updates N_{update} which will be indicated later. This algorithm leads to the construction of the following order Z decomposition of stochastic function $\lambda \in \mathcal{S} \simeq \mathbb{R} \otimes \mathcal{S}^1 \otimes \ldots \otimes \mathcal{S}^r$:

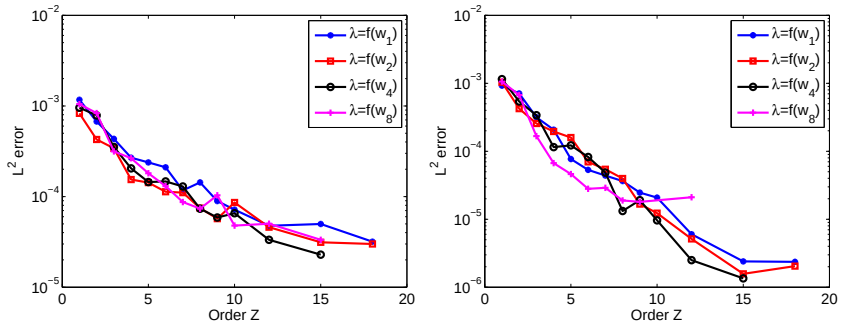
$$\lambda(\boldsymbol{\xi}) \approx \lambda_Z(\boldsymbol{\xi}) = \sum_{i=1}^{Z} \phi_i^0 \phi_i^1(\boldsymbol{\xi}_1) \ldots \phi_i^r(\boldsymbol{\xi}_r),$$

with $\phi_i^0 \in \mathbb{R}$ and $\phi_i^j \in \mathcal{S}_{P_j}^j$. In order to analyze the convergence with Z, we use the error indicator ϵ^Z defined in (77). The L^2-norm is estimated with equation (78) (Monte-Carlo simulations), where the $\{\boldsymbol{\xi}^{(q)}\}_{q=1}^{Q}$ are Q samplings of random variables $\boldsymbol{\xi}$. The reference values are defined by $\lambda(\boldsymbol{\xi}^{(q)}) = A(\boldsymbol{\xi}^{(q)})^{-1} B(\boldsymbol{\xi}^{(q)})$. Here, we take $Q = 100$. Let us note that error indicator ϵ^Z evaluates the distance between the approximate solution $\lambda_Z \in \mathcal{S}_P$ and the strong stochastic solution $\lambda \in \mathcal{S}$. It then takes into account two contributions of errors: the approximation error (introduction of $\mathcal{S}_P \subset \mathcal{S}$) and the error due to the separated representation technique. In this example, the approximation error is still negligible compared to the truncation error (sufficiently high polynomial degree used for $\mathcal{S}_P$). On Figure 22, we illustrate the convergence with Z of λ_Z for different problems $\lambda_Z \approx f(w_i)$. We plot the convergence for a parameter $N_{update} = 0$, 1 or 2 in algorithm 3. For each problem, we observe very low error values for small orders Z and a relatively good convergence rate with Z. We notice that the convergence rate with Z is increased when increasing the number N_{update} of updates (for a given order Z, better approximation when increasing N_{update}). However, performing more than 2 updates ($N_{update} > 2$) is not necessary. That means that for a given order Z and when updating the decomposition, the updating procedure converges very fast with N_{update} towards the optimal decomposition of order Z. Figure 23 illustrates this influence of N_{update}.

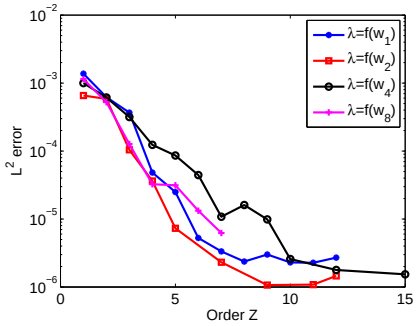
For each problem $\lambda = f(w_i)$, the algorithm allows the capture of a very low dimensional stochastic approximation space $\mathcal{S}_Z = span\{\Psi_i\}_{i=1}^{Z} \subset \mathcal{S}_P$, with $\Psi_i(\boldsymbol{\xi}) = \prod_{j=1}^{r} \phi_i^j(\boldsymbol{\xi}_j)$, which is well adapted to the representation of the solution λ of each stochastic algebraic equation. This order Z must be compared to the dimension of the underlying approximation space $P = 6.10^{16}$. In fact, for these problems, an order $Z = 1$ seems sufficient (error about 10^{-3}). These results indicate that the λ is well approximated by an order one (rank-one) separated representation and this representation is well captured by the proposed algorithm.

6.3.2 Solution of problems $\Lambda_M = f(W_M)$

We now focus on the solution of the system of stochastic algebraic equations $\Lambda_M = f(W_M)$, corresponding to step 7 of algorithm 1 (update of stochastic functions). This problem is solved with algorithm 3. For the updating step (steps 5 to 7 of algorithm 3), we use an updating along each dimension, *i.e.* $J_{update} = \{0, \ldots, r\}$, and a number of updates N_{update} which will be indicated later. This problem corresponds to the solution of equation (40) which can be seen as a Galerkin projection of the initial stochastic problem on the M-dimensional deterministic basis spanned by $W_M = \{w_i\}_{i=1}^{M}$.



(a) $N_{update} = 0$ (b) $N_{update} = 1$



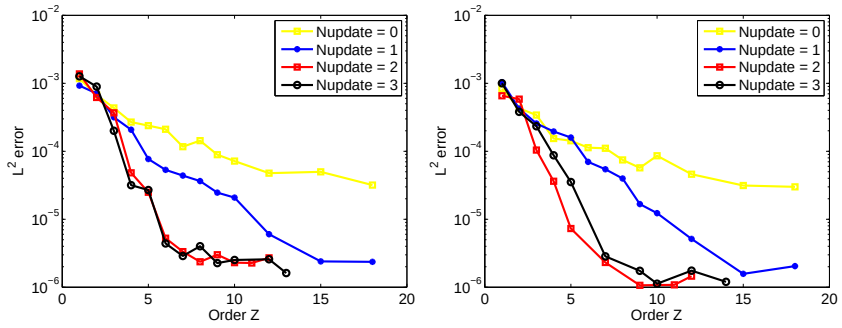
(c) $N_{update} = 2$

Fig. 22 Example 2. Approximate solution of stochastic algebraic equations $\lambda = f(w_i)$ with algorithm 3. Convergence with Z of λ_Z, in L^2-norm, for $N_{update} = 0$ (a), $N_{update} = 1$ (b) and $N_{update} = 2$ (c) ($J_{update} = \{1, \ldots, r\}$).

We assimilate $\Lambda_M \in (\mathcal{S})^M$ with a random vector $\boldsymbol{\Lambda} \in \mathbb{R}^M \otimes \mathcal{S}$ and we denote by $\mathbf{A}(\boldsymbol{\xi})\boldsymbol{\Lambda}(\boldsymbol{\xi}) = \mathbf{B}(\boldsymbol{\xi})$ the strong-stochastic form of problem $\Lambda_M = f(W_M)$. Algorithm 3 leads to the construction of the following order Z decomposition of stochastic functions $\Lambda_M \in (\mathcal{S})^M \equiv \boldsymbol{\Lambda} \in \mathbb{R}^M \otimes \mathcal{S}^1 \otimes \ldots \otimes \mathcal{S}^r$:

$$\boldsymbol{\Lambda}(\boldsymbol{\xi}) \approx \boldsymbol{\Lambda}_Z(\boldsymbol{\xi}) = \sum_{i=1}^{Z} \boldsymbol{\phi}_i^0 \phi_i^1(\boldsymbol{\xi}_1) \ldots \phi_i^r(\boldsymbol{\xi}_r),$$

with $\boldsymbol{\phi}_i^0 \in \mathbb{R}^M$, $\phi_i^j \in \mathcal{S}_{P^*}^j$. In order to analyze the convergence of the decomposition, we introduce the error indicator ϵ_M^Z, defined in (79). The L^2-norm is estimated with equation (80) (Monte-Carlo integration), where the $\{\boldsymbol{\xi}^{(q)}\}_{q=1}^Q$ are Q samplings of random variables $\boldsymbol{\xi}$. Reference values $\boldsymbol{\Lambda}(\boldsymbol{\xi}^{(q)}) = \mathbf{A}(\boldsymbol{\xi}^{(q)})^{-1}\mathbf{B}(\boldsymbol{\xi}^{(q)})$ are obtained by solving a simple system of equations. Here, we take $Q = 100$. As mentioned in the previous section, the approximation error, due to the introduction of $\mathcal{S}_P \subset \mathcal{S}$, is here negligible. Then, ϵ_M^Z quantifies the truncation error (for truncation order Z).
Figure 24 illustrates the convergence with Z of $\boldsymbol{\Lambda}_Z$ for different problems $\Lambda_{M,Z} \approx f(W_M)$. We plot the convergence for a parameter $N_{update} = 1$ in the algorithm 3. We observe that when increasing M, a higher order Z is required for reaching a given



(a) $\lambda_Z \approx f(w_1)$ (b) $\lambda_Z \approx f(w_2)$

Fig. 23 Example 2. Approximate solution of stochastic algebraic equations $\lambda = f(w_i)$ with algorithm 3, for $i = 1$ (a) and $i = 2$ (b). Convergence with Z of λ_Z, in L^2-norm. Influence of the number of updates N_{update} ($J_{update} = \{1, \ldots, r\}$).

accuracy. However, the required order seems to stabilize for $M > 10$. We obtain a good accuracy with a low order Z (error less than 10^{-2} for $Z = 7$).

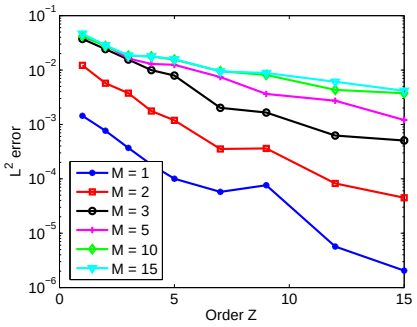


Fig. 24 Example 2. Approximate solution of the system of stochastic algebraic equations $\Lambda_M = f(W_M)$ with algorithm 3. Convergence with Z of $\Lambda_{M,Z}$ in L^2-norm for different orders M ($N_{update} = 1$).

From now on, we only focus on the problem $\Lambda_M = f(W_M)$ for $M = 15$. In figure 25, we test the influence of the number of updates N_{update}. As mentioned in section 4.4, we observe in this example that performing more than 1 update ($N_{update} > 1$) does not improve the accuracy of the decomposition for a given order Z.

6.3.3 Influence of the way to separate function space $\mathcal{S}$

We finally test the influence of the way to separate function space $\mathcal{S} = \otimes_{i=1}^{r} \mathcal{S}^i$, with $\mathcal{S}^i = L^2(\boldsymbol{\Xi}_i, \boldsymbol{\mathcal{B}}_i, P_{\boldsymbol{\xi}_i})$. The corresponding approximation space is $\mathcal{S}_P = \otimes_{i=1}^{r} \mathcal{S}^i_{P^*}$, with $\mathcal{S}^i_{P^*} = \mathbb{P}_p((-1,1)^{m^*})$. In the above reference computation, we selected $(r, m^*) = (8, 5)$. We now consider the alternatives indicated in the following table (for each couple

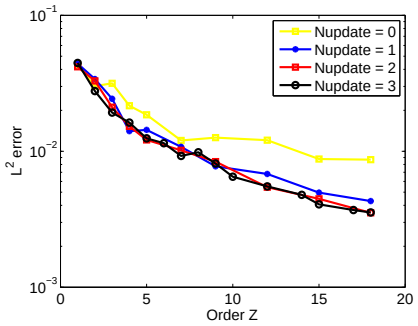


Fig. 25 Example 2. Approximate solution of the system of stochastic algebraic equations $\Lambda_M = f(W_M)$, for $M = 15$, with algorithm 3. Convergence with Z of $\Lambda_{M,Z}$, in L^2-norm. Influence of parameter N_{update} of the algorithm.

(r, m^*), the dimension of P^* and the total dimension P are indicated).

r	40	20	10	8	5
m^*	1	2	4	5	8
P^*	5	15	70	126	495
P	$\approx 9.10^{27}$	$\approx 3.10^{23}$	$\approx 3.10^{18}$	$\approx 6.10^{16}$	$\approx 3.10^{13}$

Let us remark that the change in P comes from the fact that function spaces $\mathcal{S}^i_{P^*}$ are polynomial spaces with total degree p (and not partial degree) in m^* dimensions. On Figure 26, we plot for these different alternatives, the convergence with Z for problem $\Lambda_{M,Z} \approx f(W_M)$, with $M = 15$. We observe that in this example, the way to separate the function space $\mathcal{S}$ does not have a significant influence on the convergence with Z. For all alternatives, an order $Z \approx 7$ allows to obtain an error 10^{-2}.

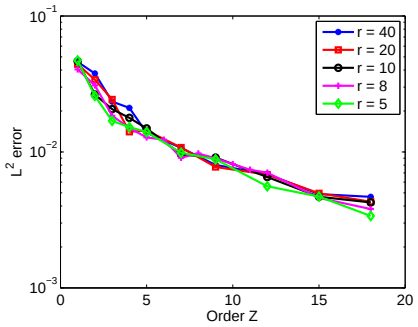


Fig. 26 Example 2. Approximate solution of the system of stochastic algebraic equations $\Lambda_M = f(W_M)$, for $M = 15$, with algorithm 3 ($N_{update} = 1$). Convergence with Z of $\Lambda_{M,Z}$, in L^2-norm. Influence of the separation of function space $\mathcal{S}$.

For the case $(r, m^*) = (40, 1)$, corresponding to a complete separation of function space, it turns out that the algorithm allows to construct a very low dimensional subspace $\mathcal{S}_Z \subset \mathcal{S}_P$, which is adapted to the solution of the problem. The solution

appears to be well represented with $Z \approx 7$, to be compared with $P = 9.10^{27}$. The proposed methodology can be seen as a method for constructing an adapted "highly sparse" representation of a solution in tensor product spaces.

6.3.4 Sensitivity analysis

Finally, we perform a sensitivity analysis of the quantity of interest with respect to random variables ξ_i. We use first order Sobol sensitivity indices defined by

$$S_i = Var(E(Q|\widetilde{\mathcal{B}}_i))/Var(Q) \tag{95}$$

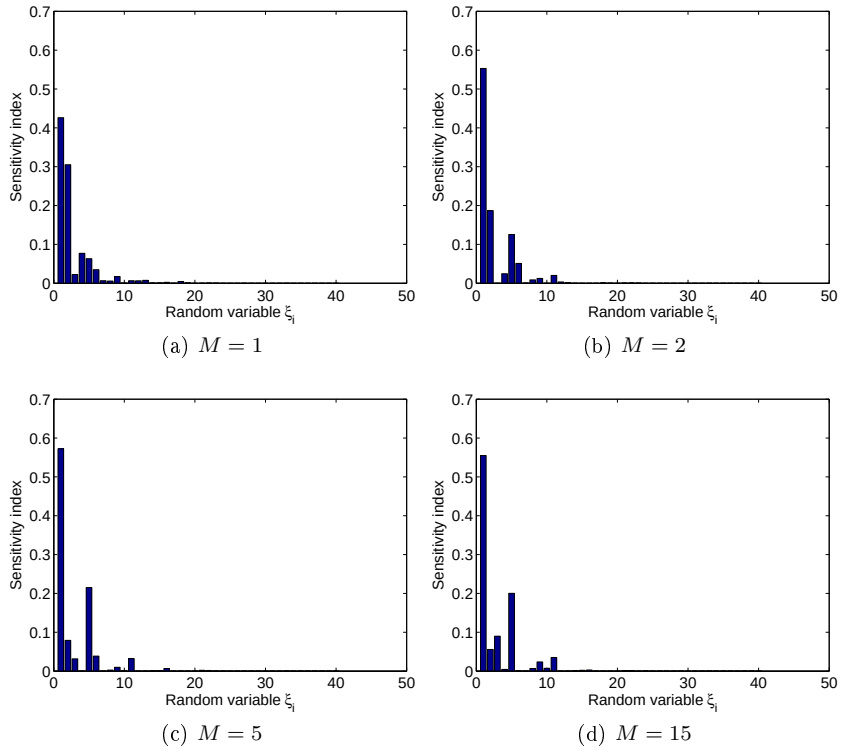
where $Var(A) = E(A^2) - E(A)^2$ denotes the variance of a random variable A and where $E(Q|\widetilde{\mathcal{B}}_i)$ is the random variable obtained by the projection of $Q \in L^2(\boldsymbol{\Xi}, \mathcal{B}, P_{\boldsymbol{\xi}})$ onto the subspace $L^2(\boldsymbol{\Xi}, \widetilde{\mathcal{B}}_i, P_{\boldsymbol{\xi}})$, where $\widetilde{\mathcal{B}}_i := \sigma^{-1}(\xi_i) := \ldots \otimes \{\Xi_{i-1}\} \otimes \mathcal{B}_i \otimes \{\Xi_{i+1}\} \otimes \ldots \subset \mathcal{B}$ is the σ-algebra generated by random variable ξ_i. This projection is the conditional expectation $E(\cdot|\widetilde{\mathcal{B}}_i)$. The reader can refer to [44] for an introduction to sensitivity analysis in the context of spectral stochastic methods. The computation of the conditional expectation operation is very simple when we have a separated representation of the quantity of interest Q under the form $Q = \sum_{k=1}^{Z} \phi_k^0 \prod_{i=1}^{m} \phi_k^i(\xi_i)$. Indeed, we have

$$E(Q|\widetilde{\mathcal{B}}_i) = \sum_{k=1}^{Z} \phi_k^j(\xi_j)\alpha_k^j, \; \alpha_k^j = \phi_k^0 \prod_{i=1, i\neq j}^{m} E(\phi_k^i(\xi_i))$$

where the expectations are simply obtained since the expansion of functions ϕ_k^i on polynomial basis is known (simple operations in the context of spectral stochastic methods). On figure 27, we plot the sensitivity index of each random variable for different values of decomposition order M. We observe a fast convergence with M of sensitivity indices (good estimation with $M = 5$). This analysis illustrates that many random variables, and then many modes in the decomposition of the diffusion parameter, are not important in the prediction of this quantity of interest. The proposed method allows to characterize accurately the significant random variables among a large number of random variables. Let us note that in this example, the sensitivity indices of random variables ξ_i do not monotically decrease with i, although the random variables were sorted by decreasing contribution in the representation of the random field $\mu(\boldsymbol{x}, \boldsymbol{\xi})$. Then, the selection of the most significant random variables was not trivial in this example.

7 Conclusion

A model reduction technique, based on *a priori* separated representations, has been proposed for solving high-dimensional stochastic partial differential equations with spectral stochastic approaches. It combines Generalized Spectral Decomposition algorithms, for a quasi optimal deterministic/stochastic separation, and a new Proper Generalized Decomposition (PGD) algorithm for the solution of systems of stochastic algebraic equations. This PGD algorithm exploits the tensor product structure of stochastic functions space and allows the *a priori* construction of a separated representation of a random solution defined on a very high-dimensional product probability



(a) $M = 1$ (b) $M = 2$

(c) $M = 5$ (d) $M = 15$

Fig. 27 Example 2. First order Sobol sensitivity indices of Q with respect to random variables ξ_i, $i = 1 \dots 40$. Convergence with M.

space. The method can handle with problems with such a dimension that their solution is unfeasible with standard spectral stochastic techniques. In that sense, the overall methodology appears as a way to circumvent the curse of dimensionality.

The ability of the proposed algorithms to solve high-dimensional stochastic problems has been illustrated on numerical examples. Further works will be devoted to the validation of these algorithms for a larger class of stochastic problems and to other types of problems formulated in tensor product spaces.

A Computational aspects of Generalized Spectral Decomposition

We here consider the computational aspects associated with the solution of problem:

$$u \in \mathcal{V} \otimes \mathcal{S}, \quad A(u, v) = B(v) \quad \forall v \in \mathcal{V} \otimes \mathcal{S} \tag{96}$$

with Generalized Spectral Decomposition algorithms introduced in section 3.

A.1 Separated representation of bilinear and linear forms

We consider that bilinear form a and linear form b in equation (3) admit the following separated representation:

$$a(w,\widetilde{w};\boldsymbol{\xi}) = \sum_{k=1}^{K_A} a_k(w,\widetilde{w}) A_k(\boldsymbol{\xi}), \quad b(\widetilde{w};\boldsymbol{\xi}) = \sum_{k=1}^{K_B} b_k(\widetilde{w}) B_k(\boldsymbol{\xi}) \tag{97}$$

where the a_k are deterministic bilinear forms on $\mathcal{V}$, where the b_k are deterministic linear forms on $\mathcal{V}$, and where the A_k and B_k are real-valued random variables defined on probability space $(\boldsymbol{\Xi}, \mathcal{B}, P_{\boldsymbol{\xi}})$. An approximation space $\mathcal{V}_N = span\{\varphi_i\}_{i=1}^N \subset \mathcal{V}$ is introduced. A function $w \in \mathcal{V}_N$ is identified with a vector $\mathbf{w} \in \mathbb{R}^N$, such that $w = \sum_{i=1}^N w_i \varphi_i$. Let $\mathbf{A} : \boldsymbol{\Xi} \to \mathbb{R}^{N\times N}$ and $\mathbf{b} : \boldsymbol{\Xi} \to \mathbb{R}^N$ denote the random matrix and random vector such that $\forall w, \widetilde{w} \in \mathcal{V}_N$

$$a(w,\widetilde{w};\boldsymbol{\xi}) := \widetilde{\mathbf{w}}^T \mathbf{A}(\boldsymbol{\xi}) \mathbf{w}, \quad b(\widetilde{w};\boldsymbol{\xi}) := \widetilde{\mathbf{w}}^T \mathbf{b}(\boldsymbol{\xi}) \tag{98}$$

Random matrix $\mathbf{A}$ and random vector $\mathbf{b}$ can be decomposed as follows:

$$\mathbf{A}(\boldsymbol{\xi}) = \sum_{k=1}^{K_A} \mathbf{A}_k^0 A_k(\boldsymbol{\xi}), \quad \mathbf{b}(\boldsymbol{\xi}) = \sum_{k=1}^{K_B} \mathbf{b}_k^0 B_k(\boldsymbol{\xi}) \tag{99}$$

where the $\mathbf{A}_k^0 \in \mathbb{R}^{N\times N}$ and $\mathbf{b}_k^0 \in \mathbb{R}^N$ are matrices and vectors associated with bilinear forms a_k and linear forms b_k on $\mathcal{V}_N$.

A.2 Classical stochastic approximation and tensor product notation

We now introduce an approximation space $\mathcal{S}_P = span\{H_\alpha\}_{\alpha=1}^P \subset \mathcal{S}$ and introduce matrices $\mathbf{A}_k^1 \in \mathbb{R}^{P\times P}$ and vectors $\mathbf{b}_k^1 \in \mathbb{R}^P$ such that

$$(\mathbf{A}_k^1)_{\alpha\beta} = E(A_k(\boldsymbol{\xi}) H_\alpha(\boldsymbol{\xi}) H_\beta(\boldsymbol{\xi})), \quad (\mathbf{b}_k^1)_\alpha = E(B_k(\boldsymbol{\xi}) H_\alpha(\boldsymbol{\xi})) \tag{100}$$

A function $u \in \mathcal{V}_N \otimes \mathcal{S}_P$ is identified with $\mathbf{u} = \sum_{\alpha=1}^P \mathbf{u}_\alpha \otimes \mathbf{e}_\alpha \in \mathbb{R}^N \otimes \mathbb{R}^P$, where $\mathbf{e}_\alpha \in \mathbb{R}^P$ is identified with $H_\alpha \in \mathcal{S}_P$. Bilinear form A and linear form B on $\mathcal{V}_N \otimes \mathcal{S}_P$ are identified with $\mathbf{A} \in \mathbb{R}^{N\times N} \otimes \mathbb{R}^{P\times P}$ and $\mathbf{b} \in \mathbb{R}^N \otimes \mathbb{R}^P$ defined by

$$\mathbf{A} = \sum_{k=1}^{K_A} \mathbf{A}_k^0 \otimes \mathbf{A}_k^1, \quad \mathbf{b} = \sum_{k=1}^{K_B} \mathbf{b}_k^0 \otimes \mathbf{b}_k^1 \tag{101}$$

and such that

$$A(u,v) := \mathbf{v} \cdot \mathbf{A} \cdot \mathbf{u}, \quad B(v) := \mathbf{v} \cdot \mathbf{b} \tag{102}$$

where operations between tensor products must be interpreted as follows: denoting $\mathbf{A}^0 \in \mathbb{R}^{N\times N}$, $\mathbf{A}^1 \in \mathbb{R}^{P\times P}$, $\mathbf{w} \in \mathbb{R}^N$, $\boldsymbol{\lambda} \in \mathbb{R}^P$

$$(\mathbf{A}^0 \otimes \mathbf{A}^1) \cdot (\mathbf{w} \otimes \boldsymbol{\lambda}) := (\mathbf{A}^0 \mathbf{w}) \otimes (\mathbf{A}^1 \boldsymbol{\lambda}) \tag{103}$$

$$(\mathbf{w} \otimes \boldsymbol{\lambda}) \cdot (\mathbf{w} \otimes \boldsymbol{\lambda}) := (\mathbf{w}^T \mathbf{w})(\boldsymbol{\lambda}^T \boldsymbol{\lambda}) \tag{104}$$

A separated representation u_M of order M is equivalently denoted

$$u_M \equiv \mathbf{u}_M = \sum_{i=1}^M \mathbf{w}_i \otimes \boldsymbol{\lambda}_i, \quad \mathbf{w}_i \in \mathbb{R}^N, \ \boldsymbol{\lambda}_i \in \mathbb{R}^P$$

A.3 Discretized versions of mappings

The residual associated with $\mathbf{u}_M$ is defined by

$$\mathbf{b}_M = \mathbf{b} - \mathbf{A}\cdot\mathbf{u}_M := \sum_{k=1}^{K_{B_M}} \widetilde{\mathbf{b}}_k^0 \otimes \widetilde{\mathbf{b}}_k^1 \tag{105}$$

Mappings $f^M : \mathcal{V}_N \to \mathcal{S}_P$, $F^M : \mathcal{S}_P \to \mathcal{V}_N$, $f : (\mathcal{V}_N)^M \to (\mathcal{S}_P)^M$, $F : (\mathcal{S}_P)^M \to (\mathcal{V}_N)^M$ are identified with mappings $\mathbf{f}^M : \mathbb{R}^N \to \mathbb{R}^P$, $\mathbf{F}^M : \mathbb{R}^P \to \mathbb{R}^N$, $\mathbf{f} : \mathbb{R}^{N\times M} \to \mathbb{R}^{P\times M}$, $\mathbf{F} : \mathbb{R}^{P\times M} \to \mathbb{R}^{N\times M}$, defined by

$$\boldsymbol{\lambda} = \mathbf{f}^M(\mathbf{w}) = \left(\sum_{k=1}^{K_A} (\mathbf{w}^T\mathbf{A}_k^0\mathbf{w})\mathbf{A}_k^1\right)^{-1}\left(\sum_{k=1}^{K_{B_M}} (\mathbf{w}^T\widetilde{\mathbf{b}}_k^0)\widetilde{\mathbf{b}}_k^1\right)$$

$$\mathbf{w} = \mathbf{F}^M(\boldsymbol{\lambda}) = \left(\sum_{k=1}^{K_A} \mathbf{A}_k^0(\boldsymbol{\lambda}^T\mathbf{A}_k^1\boldsymbol{\lambda})\right)^{-1}\left(\sum_{k=1}^{K_{B_M}} \widetilde{\mathbf{b}}_k^0(\boldsymbol{\lambda}^T\widetilde{\mathbf{b}}_k^1)\right)$$

$$\boldsymbol{\Lambda} = \mathbf{f}(\mathbf{W}) = \left(\sum_{k=1}^{K_A} (\mathbf{W}^T\mathbf{A}_k^0\mathbf{W})\otimes\mathbf{A}_k^1\right)^{-1}\left(\sum_{k=1}^{K_b} (\mathbf{W}^T\mathbf{b}_k^0)\otimes\mathbf{b}_k^1\right)$$

$$\mathbf{W} = \mathbf{F}(\boldsymbol{\Lambda}) = \left(\sum_{k=1}^{K_A} \mathbf{A}_k^0\otimes(\boldsymbol{\Lambda}^T\mathbf{A}_k^1\boldsymbol{\Lambda})\right)^{-1}\left(\sum_{k=1}^{K_b} \mathbf{b}_k^0\otimes(\boldsymbol{\Lambda}^T\mathbf{b}_k^1)\right)$$

B Computational aspects of multi-dimensional Proper Generalized Decomposition

We here consider the computational aspects associated with the solution of problem

$$u \in \mathcal{S}^0\otimes\ldots\otimes\mathcal{S}^r, \quad \alpha(u,v) = \beta(v) \quad \forall v \in \mathcal{S}^0\otimes\ldots\otimes\mathcal{S}^r \tag{106}$$

with the Proper Generalized Decomposition algorithm introduced in section 4.

B.1 Separated representation of bilinear and linear forms

We consider that $\mathcal{S}^0 \simeq \mathbb{R}^n$ and assimilate $u \in \mathcal{S}^0\otimes\ldots\otimes\mathcal{S}^r$ with a random vector $\mathbf{u}(\boldsymbol{\xi})$. We consider that bilinear form α and linear form β in equation (106) write:

$$\alpha(u,v) = E(\mathbf{v}^T\mathbf{A}\mathbf{u}), \quad \beta(v) = E(\mathbf{v}^T\mathbf{b}) \tag{107}$$

where random matrix $\mathbf{A}(\boldsymbol{\xi}) \in \mathbb{R}^{n\times n}$ and random vector $\mathbf{b}(\boldsymbol{\xi}) \in \mathbb{R}^n$ admit the following separated representation:

$$\mathbf{A}(\boldsymbol{\xi}) = \sum_{k=1}^{K_A} \mathbf{A}_k^0 A_k^1(\boldsymbol{\xi}_1)\ldots A_k^r(\boldsymbol{\xi}_r) \tag{108}$$

$$\mathbf{b}(\boldsymbol{\xi}) = \sum_{k=1}^{K_B} \mathbf{b}_k^0 B_k^1(\boldsymbol{\xi}_1)\ldots B_k^r(\boldsymbol{\xi}_r) \tag{109}$$

where $\mathbf{A}_k^0 \in \mathbb{R}^{n\times n}$, $\mathbf{b}_k^0 \in \mathbb{R}^n$, and where $A_k^j, B_k^j : \boldsymbol{\Xi}_j \to \mathbb{R}$ are random variables defined on probability space $(\boldsymbol{\Xi}_j, \mathcal{B}_j, P_{\boldsymbol{\xi}_j})$.

B.2 Stochastic approximation and tensor product notation

For each $j \in \{1, \ldots, r\}$, we introduce an approximation space $\mathcal{S}^j_{P_j} = span\{h^j_\alpha\}_{\alpha=1}^{P_j} \subset \mathcal{S}^j$ and introduce matrices $\mathbf{A}^j_k \in \mathbb{R}^{P_j \times P_j}$ and vectors $\mathbf{b}^j_k \in \mathbb{R}^{P_j}$ such that

$$(\mathbf{A}^j_k)_{\alpha\beta} = E(A^j_k(\boldsymbol{\xi}_j) h^j_\alpha(\boldsymbol{\xi}_j) h^j_\beta(\boldsymbol{\xi}_j)), \tag{110}$$

$$(\mathbf{b}^j_k)_\alpha = E(B^j_k(\boldsymbol{\xi}_j) h^j_\alpha(\boldsymbol{\xi}_j)) \tag{111}$$

A function $u \in \mathcal{S}^0 \otimes \mathcal{S}^1_{P_1} \otimes \ldots \mathcal{S}^r_{P_r}$ is identified with $\mathbf{u} \in \mathbb{R}^n \otimes \mathbb{R}^{P_1} \otimes \ldots \otimes \mathbb{R}^{P_r}$. For simplicity, let $n := P_0$. Bilinear form α and linear form β are then identified with $\mathbf{A} \in \mathbb{R}^{P_0 \times P_0} \otimes \mathbb{R}^{P_1 \times P_1} \otimes \ldots \otimes \mathbb{R}^{P_r \times P_r}$ and $\mathbf{b} \in \mathbb{R}^{P_0} \otimes \mathbb{R}^{P_1} \otimes \ldots \otimes \mathbb{R}^{P_r}$ defined by

$$\alpha(u, v) := \mathbf{v} \cdot \mathbf{A} \cdot \mathbf{u}, \quad B(v) := \mathbf{v} \cdot \mathbf{b} \tag{112}$$

with

$$\mathbf{A} = \sum_{k=1}^{K_A} \mathbf{A}^0_k \otimes \mathbf{A}^1_k \otimes \ldots \otimes \mathbf{A}^r_k \tag{113}$$

$$\mathbf{b} = \sum_{k=1}^{K_B} \mathbf{b}^0_k \otimes \mathbf{b}^1_k \otimes \ldots \otimes \mathbf{b}^r_k \tag{114}$$

and where operations between multi-dimensional tensors must be interpreted as follows: $\forall \mathbf{A}^j \in \mathbb{R}^{P_j \times P_j}$ and $\forall \boldsymbol{\phi}^j \in \mathbb{R}^{P_j}$,

$$(\mathbf{A}^0 \otimes \ldots \otimes \mathbf{A}^r) \cdot (\boldsymbol{\phi}^0 \otimes \ldots \otimes \boldsymbol{\phi}^r) := (\mathbf{A}^0 \boldsymbol{\phi}^0) \otimes \ldots \otimes (\mathbf{A}^r \boldsymbol{\phi}^r) \tag{115}$$

$$(\boldsymbol{\phi}^0 \otimes \ldots \otimes \boldsymbol{\phi}^r) \cdot (\boldsymbol{\phi}^0 \otimes \ldots \otimes \boldsymbol{\phi}^r) := \prod_{j=0}^{r} ({\boldsymbol{\phi}^j}^T \boldsymbol{\phi}^j) \tag{116}$$

A separated representation $u_Z \in \mathcal{S}^0 \otimes \ldots \otimes \mathcal{S}^r$ of order Z is equivalently denoted

$$u_Z \equiv \mathbf{u}_Z = \sum_{i=1}^{Z} \boldsymbol{\phi}^0_i \otimes \ldots \otimes \boldsymbol{\phi}^r_i, \quad \boldsymbol{\phi}^j_i \in \mathbb{R}^{P_j}$$

B.3 Discretized versions of mappings

B.3.1 Mappings F^Z_j

The residual associated with $\mathbf{u}_Z$ is defined by

$$\mathbf{b}_Z = \mathbf{b} - \mathbf{A} \cdot \mathbf{u}_Z := \sum_{k=1}^{K_{B_Z}} \widetilde{\mathbf{b}}^0_k \otimes \ldots \otimes \widetilde{\mathbf{b}}^r_k \tag{117}$$

Mappings

$$F^Z_j : \ldots \otimes \mathcal{S}^{j-1} \otimes \mathcal{S}^{j+1} \otimes \ldots \to \mathcal{S}^j \tag{118}$$

are identified with mappings

$$\mathbf{F}^Z_j : \ldots \otimes \mathbb{R}^{P_{j-1}} \otimes \mathbb{R}^{P_{j+1}} \otimes \ldots \to \mathbb{R}^{P_j} \tag{119}$$

defined by

$$\begin{aligned}\boldsymbol{\phi}^j &= \mathbf{F}_j^Z(\ldots,\boldsymbol{\phi}^{j-1},\boldsymbol{\phi}^{j+1},\ldots)\\ &= \left(\sum_{k=1}^{K_A} \Delta_k^j \mathbf{A}_k^j\right)^{-1}\left(\sum_{k=1}^{K_{B_Z}} \delta_k^j \widetilde{\mathbf{b}}_k^j\right)\end{aligned} \tag{120}$$

with

$$\Delta_k^j = \prod_{\substack{l=0\\ l\neq j}}^{r} \boldsymbol{\phi}^{l^T}\mathbf{A}_k^l\boldsymbol{\phi}^l, \quad \delta_k^j = \prod_{\substack{l=0\\ l\neq j}}^{r} \boldsymbol{\phi}^{l^T}\widetilde{\mathbf{b}}_k^l \tag{121}$$

B.3.2 Mappings F_j

Mappings

$$F_j : \ldots \otimes (\mathcal{S}^{j-1})^Z \otimes (\mathcal{S}^{j+1})^Z \otimes \ldots \to (\mathcal{S}^j)^Z \tag{122}$$

are identified with mappings

$$\mathbf{F}_j : \ldots \otimes \mathbb{R}^{P_{j-1}\times Z} \otimes \mathbb{R}^{P_{j+1}\times Z} \otimes \ldots \to \mathbb{R}^{P_j\times Z} \tag{123}$$

Denoting $\boldsymbol{\Phi}^j = (\boldsymbol{\phi}_1^j,\ldots,\boldsymbol{\phi}_Z^j) \in \mathbb{R}^{P_j\times Z}$, mapping $\mathbf{F}_j$ is defined by

$$\begin{aligned}\boldsymbol{\Phi}^j &= \mathbf{F}_j(\ldots,\boldsymbol{\Phi}^{j-1},\boldsymbol{\Phi}^{j+1},\ldots)\\ &:= \left(\sum_{k=1}^{K_A} \boldsymbol{\Delta}_k^j \otimes \mathbf{A}_k^j\right)^{-1}\left(\sum_{k=1}^{K_{B_Z}} \boldsymbol{\delta}_k^j \otimes \mathbf{b}_k^j\right)\end{aligned} \tag{124}$$

where $\boldsymbol{\Delta}_k^j \in \mathbb{R}^{Z\times Z}$ and $\boldsymbol{\delta}_k^j \in \mathbb{R}^Z$ are defined by

$$(\boldsymbol{\Delta}_k^j)_{pq} = \prod_{\substack{l=0\\ l\neq j}}^{r} \boldsymbol{\phi}_p^{l^T}\mathbf{A}_k^l\boldsymbol{\phi}_q^l, \quad (\boldsymbol{\delta}_k^j)_p = \prod_{\substack{l=0\\ l\neq j}}^{r} \boldsymbol{\phi}_p^{l^T}\mathbf{b}_k^l \tag{125}$$

References

1. A. Ammar, B. Mokdad, F. Chinesta, and R. Keunings. A new family of solvers for some classes of multidimensional partial differential equations encountered in kinetic theory modelling of complex fluids. *Journal of Non-Newtonian Fluid Mechanics*, 139(3):153–176, 2006.
2. I. Babuška, F. Nobile, and R. Tempone. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM J. Num. Anal.*, 45(3):1005–1034, 2007.
3. I. Babuška, R. Tempone, and G. E. Zouraris. Solving elliptic boundary value problems with uncertain coefficients by the finite element method: the stochastic formulation. *Computer Methods in Applied Mechanics and Engineering*, 194:1251–1294, 2005.
4. A.R. Barron, A. Cohen, W. Dahmen, and R.A. DeVore. Approximation and learning by greedy algorithms. *Ann. Statist.*, 36(1):64–94, 2008.
5. G. Beylkin and M. J. Mohlenkamp. Algorithms for numerical analysis in high dimensions. *SIAM J. Sci. Comput.*, 26(6):2133–2159, 2005.
6. G. Blatman and B. Sudret. Sparse polynomial chaos expansions and adaptive stochastic finite elements using a regression approach. *Comptes Rendus Mécanique*, 336(6):518–523, 2007.

7. C. Le Bris, T. Lelievre, and Y. Maday. Results and questions on a nonlinear approximation approach for solving high-dimensional partial differential equations. *Constructive Approximation*, 30(3):621–651, 2009.
8. C. Canuto and T. Kozubek. A fictitious domain approach to the numerical solution of pdes in stochastic domains. *Numerische Mathematik*, 107(2):257–293, 2007.
9. L. De Lathauwer, B. De Moor, and J. Vandewalle. A multilinear singular value decomposition. *SIAM J. Matrix Anal. Appl.*, 21(4):1253–1278, 2000.
10. V. de Silva and L.-H. Lim. Tensor rank and ill-posedness of the best low-rank approximation problem. *SIAM Journal of Matrix Analysis & Appl.*, 30(3):1084–1127, 2008.
11. M. Deb, I. Babuška, and J. T. Oden. Solution of stochastic partial differential equations using galerkin finite element techniques. *Computer Methods in Applied Mechanics and Engineering*, 190:6359–6372, 2001.
12. R.A. DeVore and V.N. Temlyakov. Some remarks on greedy algorithms. *Adv. Comput. Math.*, 5:173–187, 1996.
13. A. Doostan and G. Iaccarino. A least-squares approximation of partial differential equations with high-dimensional random inputs. *Journal of Computational Physics*, 228(12):4332–4345, 2009.
14. A. Doostan, G. Iaccarino, and N. Etemadi. A least-squares approximation of high-dimensional uncertain systems. Technical report, Center for Turbulence Research, Annual Research Briefs, 2007.
15. C.G. Webster F. Nobile, R. Tempone. A sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 46(5):2309–2345, 2007.
16. A. Falco and A. Nouy. 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*, submitted, 2010.
17. P. Frauenfelder, C. Schwab, and R. A. Todor. Finite elements for elliptic problems with stochastic coefficients. *Computer Methods in Applied Mechanics and Engineering*, 194(2-5):205–228, 2005.
18. R. Ghanem. Ingredients for a general purpose stochastic finite elements implementation. *Computer Methods in Applied Mechanics and Engineering*, 168:19–34, 1999.
19. R. Ghanem and P. Spanos. *Stochastic finite elements: a spectral approach*. Springer, Berlin, 1991.
20. D. Ghiocel and R. Ghanem. Stochastic finite-element analysis of seismic soil-structure interaction. *ASCE Journal Engrg. Mech.*, 128(1):66–77, 2002.
21. D. Gonzalez, A. Ammar, F. Chinesta, and E. Cueto. Recent advances on the use of separated representations. *Int. J. for Numerical Methods in Engineering*, 2009. doi:10.1002/nme.2710.
22. A. Keese. A review of recent developments in the numerical solution of stochastic pdes (stochastic finite elements). Technical Report 2003-6, Technical report, Institute of Scientific Computing, Tech. Univ. Braunschweig, Germany, 2003. http://opus.tu-bs.de/opus/volltexte/2003/504/.
23. A. Keese and H. G. Mathhies. Adaptivity and sensitivity for stochastic problems. In P. D. Spanos and G. Deodatis, editors, *Computational Stochastic Mechanics*, volume 4, pages 311–316. Millpress, Rotterdam, 2004.
24. T. G. Kolda. A counterexample to the possibility of an extension of the Eckart-Young low-rank approximation theorem for the orthogonal rank tensor decomposition. *Siam Journal On Matrix Analysis And Applications*, 24(3):762–767, 2003.
25. T. G. Kolda and B. W. Bader. Tensor decompositions and applications. *SIAM Review*, 51(3):455–500, September 2009.
26. T.G. Kolda. Orthogonal tensor decompositions. *SIAM J. Matrix Analysis & Applications*, 23(1):243–255, 2001.
27. P. Ladevèze. *Nonlinear Computational Structural Mechanics - New Approaches and Non-Incremental Methods of Calculation*. Springer Verlag, 1999.
28. P. Ladevèze and A. Nouy. On a multiscale computational strategy with time and space homogenization for structural mechanics. *Computer Methods in Applied Mechanics and Engineering*, 192:3061–3087, 2003.
29. P. Ladevèze, J.C. Passieux, and D. Néron. The LATIN multiscale computational method and the Proper Generalized Decomposition. *Computer Methods in Applied Mechanics and Engineering*, In press. doi:10.1016/j.cma.2009.06.023.

30. O.P. Le Maître, O. M. Knio, H. N. Najm, and R. G. Ghanem. Uncertainty propagation using Wiener-Haar expansions. *Journal of Computational Physics*, 197(1):28–57, 2004.
31. O.P. Le Maître, H.N. Najm, R.G. Ghanem, and O.M. Knio. Multi-resolution analysis of Wiener-type uncertainty propagation schemes. *Journal of Computational Physics*, 197(2):502–531, 2004.
32. D. Leibovici and H. El Maâche. A singular value decomposition of an element belonging to a tensor product of k separable hilbert spaces. *C. R. Acad. Sci. Paris, Série I - Math.*, 325(7):779–782, 1997.
33. L. Mathelin and O. Le Maître. Dual-based a posteriori error estimate for stochastic finite element methods. *Communications in Applied Mathematics and Computational Science*, 2(1):83–116, 2007.
34. H. G. Matthies. Stochastic finite elements: Computational approaches to stochastic partial differential equations. *Zamm-Zeitschrift Fur Angewandte Mathematik Und Mechanik*, 88(11):849–873, 2008.
35. H. G. Matthies and A. Keese. Galerkin methods for linear and nonlinear elliptic stochastic partial differential equations. *Computer Methods in Applied Mechanics and Engineering*, 194(12-16):1295–1331, 2005.
36. 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):4521–4537, 2007.
37. 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*, 197:4718–4736, 2008.
38. 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):251–285, 2009.
39. A. Nouy. A priori model reduction through Proper Generalized Decomposition for solving time-dependent partial differential equations. *Computer Methods in Applied Mechanics and Engineering*, doi:10.1016/j.cma.2010.01.009, 2010.
40. A. Nouy, A. Clément, F. Schoefs, and N. Moës. An extended stochastic finite element method for solving stochastic partial differential equations on random domains. *Computer Methods in Applied Mechanics and Engineering*, 197:4663–4682, 2008.
41. A. Nouy and P. Ladevèze. Multiscale computational strategy with time and space homogenization: a radial-type approximation technique for solving micro problems. *International Journal for Multiscale Computational Engineering*, 170(2):557–574, 2004.
42. A. Nouy and O.P. Le Maître. Generalized spectral decomposition method for stochastic non linear problems. *Journal of Computational Physics*, 228(1):202–235, 2009.
43. 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.
44. B. Sudret. Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering & System Safety*, 93(7):964–979, 2008.
45. R.A. Todor and C. Schwab. Convergence rates for sparse chaos approximations of elliptic problems with stochastic coefficients. *IMA J. Numer. Anal.*, 27(2):232Ű261, 2007.
46. X. Wan and G.E. Karniadakis. An adaptive multi-element generalized polynomial chaos method for stochastic diffential equations. *J. Comp. Phys.*, 209:617–642, 2005.
47. X. Wan and G.E. Karniadakis. Multi-element generalized polynomial chaos for arbitrary propability measures. *SIAM J. Sci. Comp.*, 28(3):901–928, 2006.
48. X. Wan and G.E. Karniadakis. Error control in multi-element generalized polynomial chaos method for elliptic problems with random coefficients. *Comm. Comput. Phys.*, 5(2-4):793–820, 2009.
49. D. Xiu. Efficient collocational approach for parametric uncertainty analysis. *Comm. Comput. Phys.*, 2(2):293–309, 2007.
50. D. Xiu. Fast numerical methods for stochastic computations: a review. *Comm. Comput. Phys.*, 5:242–272, 2009.
51. D. Xiu, B. Ganapathysubramanian, and N. Zabaras. Sparse grid collocation schemes for stochastic natural convection problems. *Journal of Computational Physics*, 225(1):652–685, 2007.
52. D. Xiu and J.S. Hesthaven. High-order collocation methods for differential equations with random inputs. *SIAM J. Sci. Comput.*, 27(3):1118Ű1139, 2005.
53. D. Xiu and G. E. Karniadakis. The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.*, 24(2):619–644, 2002.
54. D. Xiu and D. M. Tartakovsky. Numerical methods for differential equations in random domains. *SIAM J. Sci. Comput.*, 28(3):1167–1185, 2006.