COMMUNICATIONS IN COMPUTATIONAL PHYSICS
Vol. **5**, No. 2-4, pp. 793-820

Commun. Comput. Phys.
February 2009

Error Control in Multi-Element Generalized Polynomial Chaos Method for Elliptic Problems with Random Coefficients

Xiaoliang Wan*and George Em Karniadakis

Division of Applied Mathematics, Brown University, Providence, RI 02912, USA.

Received 23 August 2007; Accepted (in revised version) 25 January 2008

Available online 1 August 2008

Abstract. We develop the theory for a robust and efficient adaptive multi-element generalized polynomial chaos (ME-gPC) method for elliptic equations with random coefficients for a moderate number ($\mathcal{O}(10)$) of random dimensions. We employ low-order ($p \leq 3$) polynomial chaos and refine the solution using adaptivity in the parametric space. We first study the approximation error of ME-gPC and prove its hp-convergence. We subsequently generate local and global *a posteriori* error estimators. In order to resolve the error equations efficiently, we construct a reduced space using much smaller number of terms in the enhanced polynomial chaos space to capture the errors of ME-gPC approximation. Based on the *a posteriori* estimators, we propose and implement an adaptive ME-gPC algorithm for elliptic problems with random coefficients. Numerical results for convergence and efficiency are also presented.

AMS subject classifications: 65C20, 65C30

Key words: Stochastic PDE, a posteriori error estimate, elliptic problems, adaptive numerical methods, uncertainty quantification.

1 Introduction

Error control in large-scale simulations is based primarily on a combination of heuristic algorithms and physical considerations, often ignoring the mathematical properties of the governing equations. Progress has been made, however, especially for finite element discretizations, where techniques such as adaptive mesh refinement based on *a posteriori* error estimation (see [1, 2] and references therein), and adaptive modeling refinement [3, 4] have been developed and applied to physical applications to reduce the simulation

*Corresponding author. *Email addresses:* xlwan@dam.brown.edu (X. Wan), gk@dam.brown.edu (G. E. Karniadakis)

errors. These techniques are used for deterministic numerical simulations mostly in two- and less often in three-dimensions.

Our interest in this work is to develop error control methods to high-dimensional stochastic problems, where the input data, e.g., transport coefficients, boundary conditions or forcing and interaction terms, are modeled as random processes. A general framework in modeling such stochastic problems has been developed in [5–9] where Galerkin type expansions were employed in conjunction with a random trial basis to obtain a deterministic set of equations that can subsequently be solved with standard numerical methods. In particular, in [8] we developed an adaptive multi-element generalized polynomial chaos (ME-gPC) method.

In this paper we consider elliptic partial differential equations with stochastic coefficients which have many physical applications, e.g., random vibrations, composite materials, etc.; see [5, 10–12] and the references therein. Our objective is to set the theoretical foundations for ME-gPC and derive rigorous algorithms for error control in solving such equations.

Let $(\Omega,\mathcal{F},P)$ be a complete probability space, where Ω is the sample space, $\mathcal{F}$ is the σ-algebra of subsets of Ω, and P is a probability measure. Let D be a bounded, connected, open subset of $\mathbb{R}^d(d=1,2,3)$ with a Lipchitz continuous boundary ∂D. We consider the following stochastic linear boundary value problem: find a stochastic function, $u:\Omega\times\overline{D}\to\mathbb{R}$, such that almost surely (a.s.) the following equation holds:

$$\begin{aligned} &-\nabla\cdot(a(\boldsymbol{x};\omega)\nabla u(\boldsymbol{x};\omega))=f(\boldsymbol{x}) \quad \text{in } D,\\ &u(\boldsymbol{x};\omega)=0 \quad \text{on } \partial D, \end{aligned} \tag{1.1}$$

where $a(\boldsymbol{x};\omega)$ is a second-order random process satisfying the following assumption:

Assumption 1.1. Let $a(\boldsymbol{x};\omega)\in L_\infty(D;\Omega)$ be strictly positive with lower and upper bounds a_{min} and a_{max}, respectively,

$$0<a_{min}<a_{max} \text{ and } \Pr(a(\boldsymbol{x};\omega)\in[a_{min},a_{max}],\forall\boldsymbol{x}\in\overline{D})=1. \tag{1.2}$$

To obtain reliable simulation results for this problem in physical applications, we need to quantify the uncertainty associated with the random inputs and control the approximation errors. The traditional approach to deal with uncertainty is the Monte Carlo method and its variants, which rely on a relative large amount of realizations of the random solution field. These methods are not sensitive to the number of random dimensions but suffer from a relative low convergence rate. To this end, some non-sampling methods, such as perturbation methods [11] and second-moment analysis [13, 14] have been developed. These methods are usually restricted to systems with relatively small number of random inputs and outputs. Recently, another alternative of the non-sampling methods, the polynomial chaos method, has received considerable attention. Polynomial chaos is based on a set of basis functionals and a Galerkin projection, which is also called stochastic Galerkin method in the literature. The polynomial chaos bases can be classified based

on the global and local bases. The global bases include the Wiener-Hermite polynomial chaos [5,15–17] and the Wiener-Askey or generalized polynomial chaos (gPC) [6]; the local bases include piece-wise polynomial chaos [18], wavelet polynomial chaos [19,20] and multi-element generalized polynomial chaos (ME-gPC) [8,9]. Given enough regularity in the parametric space, the global polynomial chaos expansion converges (exponentially) fast, i.e., p-convergence, in the L_2 sense and provides a means to control the error by increasing the polynomial order p. The local polynomial chaos expansion provides another path for convergence, i.e., h-convergence, where h denotes the largest size length of random elements. Also, the local polynomial chaos expansion can control the error by h-adaptivity.

In [7] the convergence of stochastic Galerkin method for the elliptic problem (1.1) was proved and some *a priori* error estimates were derived. The convergence rate was further studied in [21] based on sparse global and local bases, where the basis modes were carefully chosen based on the decay rate of eigenvalues of the covariance kernel of $a(\boldsymbol{x};\omega)$ and some optimal convergence rates were proved.

In this paper, we focus on the error control of ME-gPC by using the strategy of *a posteriori* error estimate and h-adaptivity. We briefly present the Karhunen-Loeve decomposition of the random process $a(\boldsymbol{x};\omega)$ in the next section. We prove the hp-convergence of ME-gPC in Section 3. The *a posteriori* error estimate is developed in Section 4, and it is based on the solution of the error equation in an enhanced polynomial chaos space. The main problem of such an *a posteriori* error estimate is that the cost of solving the error equation is much larger than the cost of solving the original equation. To this end, we propose a reduced space whose cardinality can be much smaller than that of the normal space for the error equation. Subsequently, the underestimated error estimate from the reduced space is corrected by a factor to obtain the final one. The computational savings of such a procedure are about 90%. Based on the *a posteriori* error estimate we construct an h-type adaptive ME-gPC method for the model problem in Section 5. We then present some numerical results to show the efficiency of the proposed adaptive ME-gPC algorithm. We conclude our work in Section 7 with a brief discussion.

2 Karhunen-Loève (K-L) decomposition

Karhunen-Loève decomposition is an optimal representation of a *second-order* random process $a(\boldsymbol{x};\omega)$ [22]. We assume that the covariance kernel $K(\boldsymbol{x}_1,\boldsymbol{x}_2)$ is known and defined as

$$K(\boldsymbol{x}_1,\boldsymbol{x}_2)=\int_\Omega (a(\boldsymbol{x}_1;\omega)-\mathbb{E}[a](\boldsymbol{x}_1))(a(\boldsymbol{x}_2;\omega)-\mathbb{E}[a](\boldsymbol{x}_2))P(\mathrm{d}\omega),\quad \forall(\boldsymbol{x}_1,\boldsymbol{x}_2)\in D\times D. \tag{2.1}$$

For such a covariance kernel $K(\boldsymbol{x}_1,\boldsymbol{x}_2)$, there exists eigenpairs $\{(\lambda_i,h_i)\}_{i=1}^{\infty}$, where $\lambda_i\geq\lambda_j>0$ for $i<j$, satisfying

$$\int_D K(\boldsymbol{x}_1,\boldsymbol{x}_2)h_i(\boldsymbol{x}_1)\mathrm{d}\boldsymbol{x}_1=\lambda_i h_i(\boldsymbol{x}_2)\quad\text{and}\quad\int_D h_i(\boldsymbol{x})h_j(\boldsymbol{x})\mathrm{d}\boldsymbol{x}=\delta_{ij}. \tag{2.2}$$

Based on the eigenpairs $\{(\lambda_i,h_i)\}_{i=1}^{\infty}$, the truncated K-L decomposition of $a(\boldsymbol{x};\omega)$ is

$$a_M(\boldsymbol{x};\omega)=\mathbb{E}[a](\boldsymbol{x})+\sum_{i=1}^{M}\sqrt{\lambda_i}h_i(\boldsymbol{x})y_i(\omega), \tag{2.3}$$

where $\{y_i\}_{i=1}^{\infty}$ are mutually *uncorrelated* random variables with zero mean and unit variance, determined by $a(\boldsymbol{x};\omega)$ as

$$y_i(\omega)=\frac{1}{\sqrt{\lambda_i}}\int_D(a(\boldsymbol{x};\omega)-\mathbb{E}[a](\boldsymbol{x}))h_i(\boldsymbol{x})\mathrm{d}\boldsymbol{x}. \tag{2.4}$$

For a Gaussian process, $\{y_i\}_{i=1}^{\infty}$ are independent identically-distributed (i.i.d.) Gaussian random variables; otherwise, the probability density functions (PDFs) of $\{y_i\}_{i=1}^{\infty}$ depend on $a(\boldsymbol{x};\omega)$. $a_M(\boldsymbol{x};\omega)$ converges to $a(\boldsymbol{x};\omega)$ in the L_2 sense

$$\sup_{\boldsymbol{x}\in D}\mathbb{E}[(a-a_M)^2](\boldsymbol{x})=\sup_{\boldsymbol{x}\in D}\sum_{i=M+1}^{\infty}\lambda_i h_i^2(\boldsymbol{x})\to 0,\quad \text{as } M\to\infty. \tag{2.5}$$

For a fixed correlation length, the decay rate of the eigenvalues is determined by the regularity of the covariance kernel $K(\boldsymbol{x}_1,\boldsymbol{x}_2)$ [23]. For example, the eigenvalues of the exponential kernel $\exp(-|\boldsymbol{x}_1-\boldsymbol{x}_2|/A)$, which has finite Sobolev regularity, have an algebraic decay rate; however, the eigenvalues of the Gaussian kernel $\exp(-|\boldsymbol{x}_1-\boldsymbol{x}_2|^2/A)$, which is analytic, have an exponential decay rate. The decay rate of eigenvalues influences significantly the number M for a fixed truncation error, and is representative of the computational complexity of the method.

3 Formulation

Let

$$\boldsymbol{\zeta}=(\zeta_1,\cdots,\zeta_n):(\Omega,\mathcal{F})\to(\mathbb{R}^n,\mathcal{B}^n) \tag{3.1}$$

be an $\mathbb{R}^n$-valued continuous random variable, where $n\in\mathbb{N}$ and $\mathcal{B}^n$ is the σ-algebra of Borel subsets of $\mathbb{R}^n$. For a multi-index $\alpha=(\alpha_1,\cdots,\alpha_n)\in\mathbb{N}_0^n$ we define the operations

$$\begin{aligned}&|\alpha|=\alpha_1+\alpha_2+\cdots+\alpha_n,\quad \alpha!=\alpha_1!\alpha_2!\cdots\alpha_n!,\quad \boldsymbol{x}^{\alpha}=x_1^{\alpha_1}x_2^{\alpha_2}\cdots x_n^{\alpha_n},\\ &\alpha=\beta,\quad \text{if and only if } \alpha_i=\beta_i,\ \forall\alpha,\beta\in\mathbb{N}_0^n.\end{aligned}$$

A general second-order random process $R(\omega)\in L_2(\Omega,\mathcal{F},P)$ can be expressed by generalized polynomial chaos (gPC) [6] as

$$R(\omega)=\sum_{|\alpha|=0}^{\infty}a_{\alpha}\phi_{\alpha}(\boldsymbol{\zeta}(\omega)), \tag{3.2}$$

where $\phi_\alpha(\boldsymbol{\zeta}(\omega))$ denotes the gPC basis of degree $|\alpha|=p$ in terms of the random variable $\boldsymbol{\zeta}$. ϕ_α is, in general, a tensor product of one-dimensional basis $\{\phi_{\alpha_i}(\zeta_i)\}$. The family $\{\phi_\alpha\}$ is an orthogonal basis in $L_2(\Omega,\mathcal{F},P)$ with the orthogonality relation

$$\mathbb{E}[\phi_\alpha\phi_\beta]=\mathbb{E}[\phi_\alpha^2]\delta_{\alpha\beta}, \tag{3.3}$$

where $\delta_{\alpha\beta}$ is the Kronecker delta and $\mathbb{E}$ denotes the expectation with respect to the probability measure

$$P(\mathrm{d}\omega)=w(\boldsymbol{\zeta}(\omega))\mathrm{d}\omega.$$

The index in Eq. (3.2) and $n\in\mathbb{N}$ are, in general, infinite. In practice, the expansion (3.2) is usually truncated at a certain level according to the convergence behavior.

3.1 Multi-element generalized polynomial chaos (ME-gPC)

3.1.1 Background

Based on gPC and the decomposition of parametric space, a multi-element polynomial chaos method was proposed in [8, 9]. We assume that ζ_i are i.i.d. random variables and the support of $\boldsymbol{\zeta}$ is $B=\prod_{i=1}^n(a_i,b_i)$, where a_i and b_i are finite or infinite in $\mathbb{R}$. A decomposition $\mathbf{D}$ of B is defined as

$$\mathbf{D}=\begin{cases} B_k=(a_{k,1},b_{k,1})\times(a_{k,2},b_{k,2})\times\cdots\times(a_{k,n},b_{k,n}),\\ \bar{B}=\bigcup_{k=1}^N\bar{B}_k,\\ B_{k_1}\bigcap B_{k_2}=\emptyset,\ \text{if } k_1\neq k_2, \end{cases} \tag{3.4}$$

where $k,k_1,k_2=1,2,\cdots,N$. Let I_{B_k} be the indicator random variable related to element B_k. It is seen that $\Omega=\cup_{k=1}^N I_{B_k}^{-1}(1)$ is a decomposition of the sample space Ω with

$$I_{B_i}^{-1}(1)\cap I_{B_j}^{-1}(1)=\emptyset,\quad \text{for } i\neq j.$$

Subsequently, a local random variable $\boldsymbol{\zeta}_k:I_{B_k}^{-1}(1)\mapsto B_k$ on the probability space $(I_{B_k}^{-1}(1),\mathcal{F}\cap I_{B_k}^{-1}(1),P(\cdot|I_{B_k}=1))$ is defined subject to a conditional PDF. We assume that $\Pr(I_{B_k}=1)>0$.

Proposition 3.1 ([9]). Let $\mathcal{P}_p u(\boldsymbol{\zeta})$ denote the Galerkin projection of $u(\boldsymbol{\zeta})$ onto the polynomial chaos basis $\{\phi_\alpha(\boldsymbol{\zeta})\}$ with $\alpha_i\leq p$. If $\mathcal{P}_p u(\boldsymbol{\zeta})$ converges to $u(\boldsymbol{\zeta})$ in the L_2 sense with respect to the PDF $w(\boldsymbol{\zeta})$, then $\mathcal{P}_p u(\boldsymbol{\zeta}_k)$ converges to $u(\boldsymbol{\zeta}_k)$ in the L_2 sense with respect to the conditional PDF $w_k(\boldsymbol{\zeta}_k|I_{B_k}=1)$, $k=1,2,\cdots,N$.

However, the gPC basis, which is orthogonal on the entire parametric space, will lose orthogonality locally in element k. For efficiency, we employ local orthogonal basis $\{\phi_{k,\alpha}(\boldsymbol{\zeta}_k)\}$, which are orthogonal polynomials constructed numerically since the conditional PDF $w_k(\boldsymbol{\zeta}_k|I_{B_k}=1)$, in general, does not correspond to any weight function of classical orthogonal polynomials. We note that the multi-dimensional basis $\{\phi_{k,\alpha}(\boldsymbol{\zeta}_k)\}$

are tensor products of the one-dimensional basis $\{\phi_{k,\alpha_i}(\zeta_{k,i})\}$, which can be generated efficiently [24] using the three-term recurrence relation,

$$\begin{aligned}&\pi_{i+1}(\tau)=(\tau-c_{i,0})\pi_i(\tau)-c_{i,1}\pi_{i-1}(\tau),\quad i=0,1,\cdots,\\&\pi_0(\tau)=1,\quad \pi_{-1}(\tau)=0,\end{aligned}\tag{3.5}$$

where $\{\pi_i(\tau)\}$ is a set of (monic) orthogonal polynomials,

$$\pi_i(\tau)=\tau^i+\text{lower-degree terms},\quad i=0,1,\cdots\tag{3.6}$$

and the coefficients $c_{i,0}$ and $c_{i,1}$ are uniquely determined by a positive measure corresponding to the conditional PDF $w_k(\boldsymbol{\zeta}_k|I_{B_k}=1)$.

We note that $\|\phi_{k,\alpha_i}(\zeta_{k,i})\|_{L_2}$ decreases fast as $\alpha_i\to\infty$ if the support $(a_{k,i},b_{k,i})$ is small, which may introduce the underflow problem [9] in numerical computation. In practice, we consider the following linear mapping

$$\zeta_{k,i}=\frac{b_{k,i}-a_{k,i}}{2}y_{k,i}+\frac{b_{k,i}+a_{k,i}}{2},\tag{3.7}$$

where $\boldsymbol{\zeta}_k$ is mapped to a new n-dimensional random variable $\boldsymbol{y}_k$. The support of $\boldsymbol{y}_k$ is a n-dimensional hypercube $\text{supp}(\boldsymbol{y}_k):=(-1,1)^n$ if $a_{k,i}$ and $b_{k,i}$ are finite. The PDF of $\boldsymbol{y}_k$ is

$$w_k(\boldsymbol{y}_k|I_{B_k}=1)=\frac{w(\boldsymbol{\zeta}(\boldsymbol{y}_k))}{\Pr(\boldsymbol{\zeta}\in B_k)}\prod_{i=1}^{n}\frac{b_{k,i}-a_{k,i}}{2}.\tag{3.8}$$

Then, we construct orthogonal polynomials $\{\phi_{k,\alpha}(\boldsymbol{y}_k)\}$ satisfying

$$\mathbb{E}_k[\phi_{k,\alpha}\phi_{k,\beta}]=\mathbb{E}_k[\phi_{k,\alpha}^2]\delta_{\alpha\beta},$$

where $\mathbb{E}_k[\cdot]$ indicate the expectation with respect to the PDF $w_k(\boldsymbol{y}_k|I_{B_k}=1)$. It was shown in [9] that $\{\phi_{k,\alpha_i}(y_{k,i})\}$ can be obtained on-the-fly, which are usually accurate and well scaled for $\alpha_i\le 20$.

Let $u_k^p(\boldsymbol{y}_k)$ be the local polynomial chaos expansion corresponding to element B_k. Then, the global approximation for $\boldsymbol{\zeta}\in B$ can be expressed as

$$u^p(\boldsymbol{\zeta})=\sum_{k=1}^{N}u_k^p(\boldsymbol{y}_k(\boldsymbol{\zeta}))I_{B_k}=\sum_{k=1}^{N}\sum_{\alpha_i\le p}u_{k,\alpha}\phi_{k,\alpha}(\boldsymbol{y}_k(\boldsymbol{\zeta}))I_{B_k},\tag{3.9}$$

which converges to $u(\boldsymbol{\zeta})$ in the L_2 sense, in other words,

$$\int_B(u^p(\boldsymbol{\zeta})-u(\boldsymbol{\zeta}))^2f(\boldsymbol{\zeta})\mathrm{d}\boldsymbol{\zeta}\to 0,\quad \text{as } p\to\infty.\tag{3.10}$$

Remark 3.1. The following C^0 continuity

$$u_i^p(\boldsymbol{y}_i(\boldsymbol{\zeta}))=u_j^p(\boldsymbol{y}_j(\boldsymbol{\zeta})),\quad \boldsymbol{\zeta}\in\bar{B}_i\cap\bar{B}_j,\tag{3.11}$$

where $\bar{B}_i$ and $\bar{B}_j$ indicate the closure of two adjacent random elements, respectively, is not required since the Lebesgue measure of the interface is zero.

By Bayes' rule and the law of total probability, any statistics can be obtained as

$$\int_B g(u(\boldsymbol{\zeta}))w(\boldsymbol{\zeta})\mathrm{d}\boldsymbol{\zeta} \approx \sum_{k=1}^{N} \Pr(I_{B_k}=1)\int_{B_k} g\left(u_k^p(\boldsymbol{y}_k)\right) w_k(\boldsymbol{y}_k|I_{B_k}=1)\mathrm{d}\boldsymbol{y}_k, \tag{3.12}$$

where $g(\cdot)\in L_1(\Omega,\mathcal{F},P)$ is a function of the random field $u(\boldsymbol{\zeta})$.

3.1.2 *hp*-convergence of ME-gPC

In this section we study the *hp*-convergence of ME-gPC, where h denotes the side length of random element and p the polynomial order. We consider the Legendre-chaos expansion, which corresponds to a uniform distribution. Let $I^n=(-1,1)^n$ be the n-dimensional hypercube. Let $P_p(I^n)$ denote the collection of tensor products of one-dimensional Legender polynomials of (separate) degree $\leq p$ defined on I^n.

We have the following theorem:

Theorem 3.1. *If $u(\boldsymbol{y})\in H^m(I^n)$, $m\geq 0$, and $\boldsymbol{y}$ is a uniform random variable on I^n, then we have*

$$\|u-\mathcal{P}_p u\|_{L_2(I^n)} = 2^n \mathbb{E}[(u-\mathcal{P}_p u)^2]^{1/2} \leq Cp^{-m}\|u\|_{H^m(I^n)}, \tag{3.13}$$

where C is a constant independent of p.

Proof. The theorem can be proved by considering the Legendre expansion of u and also from the approximation results of p-th order polynomials; see [25] for more details. □

We now quote the following approximation results.

Lemma 3.1. *Let A be an open subset of $\mathbb{R}^n$. There exist a constant $C(A)$ such that*

$$\forall v\in H^{m+1}, \quad \inf_{\hat{v}\in P_m(A)} \|v-\hat{v}\|_{H^{m+1}} \leq C(A)|v|_{H^{m+1}},$$

where the semi-norm $|v|_{H^{m+1}}$ is defined as

$$|v|_{H^{m+1}} = \left(\sum_{|\alpha|=m+1}\int_A |\partial_{\boldsymbol{y}} v|^2 \mathrm{d}\boldsymbol{y}\right)^{1/2}.$$

Proof. See Theorem 3.1.1 in [26]. □

We also need the scaling argument.

Lemma 3.2. *Let A and $\hat{A}$ be two open subsets of $\mathbb{R}^n$ such that there exists an affine mapping $F(x)=B(x)+b$ of $\hat{A}$ onto A and $F(\hat{A})=A$. Let $diam(A)=1$, $\rho_A=K$, $diam(\hat{A})=h$, $\rho_{\hat{A}}=\bar{K}h$. If a function $\hat{v}$ belongs to the Sobolev space $W^{m,q}(A)$ for some integer $m\geq 0$ and some number $q\in[1,\infty]$, the function $v=\hat{v}\circ F\in W^{m,q}(\hat{A})$ then*

$$|\hat{v}|_{H^m(A)} \leq Ch^{m-\frac{n}{2}}|v|_{H^m(\hat{A})}, \tag{3.14a}$$

$$|\hat{v}|_{H^m(\hat{A})} \leq Ch^{\frac{n}{2}-m}|v|_{H^m(A)}. \tag{3.14b}$$

Proof. See Theorem 3.1.2 in [26]. □

We are now ready to prove the *hp*-convergence of ME-gPC for Legendre-chaos.

Theorem 3.2. *Let $u(\boldsymbol{y})\in H^m(I^n)$ and $\boldsymbol{D}$ be a uniform decomposition of I^n with N_1 element along each y_i. Then the ME-gPC approximation $u^p(\boldsymbol{y})\in P_p(\boldsymbol{y})$ converges to $u(\boldsymbol{y})$ with an error*

$$\|u^p-u\|_{L_2(I^n)}\leq Cp^{-m}N_1^{-m}|u|_{H^m(I^n)}, \tag{3.15}$$

where we assume that $m\leq p+1$.

Proof. We consider element B_k, where we define $\boldsymbol{y}_k(\boldsymbol{y}):B_k\mapsto I^n$ by an affine mapping. Let $v(\boldsymbol{y}_k)\in P_p(I^n)$. Using Theorem 3.1 and Lemma 3.2, we obtain

$$\begin{aligned}&\|u(\boldsymbol{y}_k)-\mathcal{P}_p(\boldsymbol{y}_k)\|_{L_2(I^n)}=\|u-v-\mathcal{P}_p(u-v)\|_{L_2(I^n)}\\ \leq& Cp^{-m}\inf_{v\in P_p(I^n)}\|u-v\|_{H^m(I^n)}\leq Cp^{-m}N_1^{-m+\frac{n}{2}}|u|_{H^m(B_k)}.\end{aligned}$$

Using the formula (3.12), we obtain

$$\begin{aligned}&\|u(\boldsymbol{y})-u^p(\boldsymbol{y})\|^2_{L_2(I^n)}=2^n\mathbb{E}[(u(\boldsymbol{y})-u^p(\boldsymbol{y}))^2]\\ =&2^n\sum_{k=1}^N\Pr(I_{B_k}=1)\mathbb{E}_k[(u(\boldsymbol{y}_k)-\mathcal{P}_pu(\boldsymbol{y}_k))^2]\\ =&\sum_{k=1}^N\text{volume}(B_k)\|u(\boldsymbol{y}_k)-\mathcal{P}_pu(\boldsymbol{y}_k)\|^2_{L_2(I^n)}\leq\sum_{k=1}^N C|u(y)|^2_{H^m(B_k)}p^{-2m}N_1^{-2m}\\ \leq& Cp^{-2m}N_1^{-2m}\sum_{k=1}^N|u(y)|^2_{H^m(B_k)}=Cp^{-2m}N_1^{-2m}|u(y)|^2_{H^m(I^n)}.\end{aligned} \tag{3.16}$$

This completes the proof of this theorem. □

Remark 3.2. If we take $m=p+1$ and consider the error of the second-order moment we recover the *h*-convergence rate

$$|\mathbb{E}[u^2]-\mathbb{E}[\mathcal{P}_p^2u]|\leq C(p)N_1^{-2(p+1)}$$

shown in [8,18].

Remark 3.3. The convergence rate given in Theorem 3.2 is similar with that of the deterministic *hp* polynomial approximation, which usually takes the form [27]

$$\|u-u^p\|_{L_2}\leq Cp^{-m}h^\mu\|u\|_{H^m},$$

where $\mu=\min(p+1,m)$. Since we focus on the *h*-type adaptivity in this paper, we do not pay much attention on the regularity parameter m, so we have assumed that $m\leq p+1$.

Remark 3.4. In Theorem 3.2, we use a full tensor-product basis. Such an approximation will suffer from the "curse of dimensionality". In practice, the importance of each random dimension is usually different, sparse bases can be employed adaptively to reduce the cost while maintaining good accuracy. Convergence rates of certain sparse approximations based on analyticity were given in [21]. We will discuss the adaptivity of ME-gPC based on an *a posteriori* error estimators later.

In applications of gPC we usually consider a relative small perturbation, i.e., we associate a random variable with a small number (standard deviation). Using the same idea as in the proof of Theorem 3.2, we obtain the following corollary.

Corollary 3.1. Let $\mathcal{P}_p u(\boldsymbol{y})$ be the Legendre-chaos expansion of $u(\delta_c \boldsymbol{y})$, where $\boldsymbol{y}$ is a uniform random variable on I^n, and $0<\delta_c<1$ is a constant. If $u(\boldsymbol{y})\in H^m$ with $m\leq p+1$, then the approximation errors can be expressed as

$$\|u(\delta_c \boldsymbol{y})-\mathcal{P}_p u(\delta_c \boldsymbol{y})\|_{L_2(I^n)}\leq C p^{-m}\delta_c^m, \tag{3.17a}$$

$$|\mathbb{E}[u^2]-\mathbb{E}[\mathcal{P}_p^2 u]|\leq C p^{-2m}\delta_c^{2m}. \tag{3.17b}$$

It is seen that if the degree of perturbation decreases the error will decrease algebraically for a fixed polynomial order p.

3.2 Galerkin projection

For each random element B_k, an independent problem of the same type is defined as

$$\begin{aligned} &-\nabla\cdot(a(\boldsymbol{x};\boldsymbol{y}_k)\nabla u(\boldsymbol{x};\boldsymbol{y}_k))=f(\boldsymbol{x}) \quad \text{in } D,\\ &u(\boldsymbol{x};\boldsymbol{y}_k)=0 \quad \text{on } \partial D. \end{aligned} \tag{3.18}$$

All previous assumptions on $a(\boldsymbol{x};\cdot)$ are satisfied with respect to the conditional PDF $w_k(\boldsymbol{y}_k|I_{B_k}=1)$. We note that such a system is complete.

We define a bilinear form $\mathscr{B}_k(\cdot,\cdot)$ as

$$\mathscr{B}_k(u,v)=\mathbb{E}_k[\int_D a\nabla u\cdot\nabla v \mathrm{d}\boldsymbol{x}] \tag{3.19}$$

and a linear form $\mathscr{L}_k(\cdot)$ as

$$\mathscr{L}_k(v)=\mathbb{E}_k[\int_D fv\mathrm{d}\boldsymbol{x}]. \tag{3.20}$$

We consider the following Hilbert space

$$W_k=\left\{v(\boldsymbol{x};\boldsymbol{y}_k(\omega))\,\middle|\,\|v\|_{W_k}^2=\mathbb{E}_k[\int_D \nabla v\cdot\nabla v\mathrm{d}\boldsymbol{x}]<\infty,\quad \forall v\in H_0^1(D)\otimes L_2(I_{B_k}^{-1}(1))\right\} \tag{3.21}$$

with an inner product $(u,v)_{W_k}=\mathbb{E}_k[\int_D \nabla u\cdot\nabla v\mathrm{d}\boldsymbol{x}]$. For the entire parametric space, a similar Hilbert space W can be defined with an inner product $(u,v)_W=\mathbb{E}[\int_D \nabla u\cdot\nabla v\mathrm{d}\boldsymbol{x}]$.

Due to Assumption 1.1, the following conditions are naturally satisfied

$$\mathscr{B}_k(u,v) \leq a_{max}\|u\|_{W_k}\|v\|_{W_k}, \quad \forall u,v \in W_k \text{ (Continuity)}; \tag{3.22a}$$
$$a_{min}\|v\|_{W_k} \leq \mathscr{B}_k(v,v), \quad \forall v \in W_k \text{ (Coercivity)}. \tag{3.22b}$$

Thus, we can claim by the Lax-Milgram theorem that an unique solution exists for the following weak form

$$\mathscr{B}_k(u,v) = \mathscr{L}_k(v), \quad \forall v \in W_k. \tag{3.23}$$

4 *A posteriori* error estimation of gPC

In the deterministic finite element method (FEM), the local error estimator can be obtained either explicitly from the errors of a proper interpolation [2], or implicitly from a local problem [1]. In ME-gPC we use similar ideas to obtain a local error estimator. We focus on how to estimate the error of gPC efficiently with a relative low cost.

For simplicity we drop the subscript k, since the following discussion is valid for every ME-gPC element. In practice, the space W needs to be truncated for an approximation of $u(\boldsymbol{x};\boldsymbol{y})$. To focus on the error of polynomial chaos, we use the following assumption:

Assumption 4.1. The error given by spatial discretization can be neglected in comparison to the error introduced by the truncated polynomial chaos.

Let $\|v\|_E = \mathbb{E}[\int_D a|\nabla v|^2]^{1/2}$ denote the energy norm. We define the approximation error as

$$e = \|u - u^p\|_E. \tag{4.1}$$

4.1 Local error e

Since the dimension of W is infinite, we need to truncate it to a certain level for an approximate solution of Eq. (3.23). Let

$$Q^p = \text{span}\{\phi_\alpha(\boldsymbol{y}), \quad |\alpha| \leq p\}$$

denote the truncated polynomial chaos space up to polynomial order p. We define a truncated version of W as $W^p = \{v(\boldsymbol{x};\boldsymbol{y}) | v \in H_0^1(D) \otimes Q^p\}$. Due to Assumption 4.1, we here keep $H_0^1(D)$ unchanged although in numerical computation it will also be truncated. Let u^p satisfy the following equation

$$\mathscr{B}(u^p,v) = \mathscr{L}(v), \quad \forall v \in W^p. \tag{4.2}$$

Since Q^p is an hierarchical orthogonal space we introduce the following *saturation assumption*:

Assumption 4.2. There exists a constant $\beta\in[0,1)$ such that

$$\|u-u^{p+q}\|_E\leq\beta\|u-u^p\|_E,\quad\forall q\in\mathbb{N}.$$

Let $e^*=u^{p+q}-u^p$, which satisfies that

$$\mathscr{B}(e^*,v)=\mathscr{L}(v)-\mathscr{B}(u^p,v),\quad\forall v\in W^{p+q}. \tag{4.3}$$

Under Assumption 4.2, the true error e in the energy norm can be bounded from both ends by e^* as [1]

$$\|e^*\|_E\leq\|e\|_E\leq\frac{1}{\sqrt{1-\beta^2}}\|e^*\|_E. \tag{4.4}$$

However, it will be more expensive to solve Eq. (4.3) than the original problem. Thus, we decompose e^* as

$$e^*=e_1+e_2, \tag{4.5}$$

where $e_1\in W^p$ and $e_2\in Y_{p,q}:=W^{p+q}-W^p$ satisfy

$$\mathscr{B}(e_1,v)+\mathscr{B}(e_2,v)=L(v)-\mathscr{B}(u^p,v)=0,\quad\forall v\in W^p, \tag{4.6a}$$

$$\mathscr{B}(e_1,w)+\mathscr{B}(e_2,w)=L(w)-\mathscr{B}(u^p,w),\quad\forall w\in Y_{p,q}. \tag{4.6b}$$

We note that e_1 and e_2 are coupled together. To obtain a reasonable approximation $\bar{e}$ to e^*, the coupling terms can be ignored [1] such that

$$\mathscr{B}(\bar{e}_1,v)=0,\quad\forall v\in W^p, \tag{4.7a}$$

$$\mathscr{B}(\bar{e}_2,w)=L(w)-\mathscr{B}(u^p,w),\quad\forall w\in Y_{p,q}, \tag{4.7b}$$

where $\bar{e}=\bar{e}_1+\bar{e}_2=\bar{e}_2$. By assuming that a *strengthened Cauchy-Schwarz inequality* [1] holds, the error estimator e^* in energy norm can be bounded from both ends as

$$\|\bar{e}\|_E\leq\|e^*\|_E\leq\frac{1}{\sqrt{1-\gamma^2}}\|\bar{e}\|_E, \tag{4.8}$$

where $\gamma\in[0,1)$ is a constant.

Based on the above discussion, we know that $\|\bar{e}\|_E$ can be used as a local error estimator. Due to the equivalence between the norm $\|\cdot\|_W$ and $\|\cdot\|_E$ we use

$$\eta=\|e\|_W\approx\|\bar{e}\|_W=\|\bar{e}\|_{L_2(B,H_0^1(D))}$$

as a local error estimator.

The cost of obtaining the error estimator $\eta\approx\|\bar{e}\|_{L_2(B,H_0^1(D))}$ is determined by the number $\dim(Y_{p,q})$. We set $q=1$. In Table 1 we show some typical values of $\dim(Y_{p,q})$ for $p=2$, where we also give the dimensions of the solution space W^p in parenthesis for comparison. It is seen that $\dim(Y_{p,q})$ can be much larger than the dimension of W^p, which implies that the cost to obtain η will be much more expensive than that to solve the original problem. Clearly, to make the numerical method practical we need to reduce the cost of obtaining the local error estimators.

Table 1: Typical values of $\dim(Y_{p,q})$ with respect to the number n of random dimensions. $p=2$ and $q=1$. The numbers in parenthesis denote the dimensions of the space W^p.

n	2	4	6	8	10
$\dim(Y_{p,q})$	4(6)	20(15)	56(28)	120(45)	220(66)

4.2 Stochastic regularity

4.2.1 Taylor expansion

We first consider the Taylor expansion of $u(\cdot,\boldsymbol{y})$ to examine the behavior of error contribution of each term $\boldsymbol{y}^\alpha$. Let

$$\rho_i = r_i\sqrt{\lambda_i}\|h_i\|_{L_\infty(D)}, \tag{4.9}$$

where r_i are constants from the linear mapping (3.7) in the ME-gPC decomposition.

In [21], the following proposition about the regularity of $u(\cdot,\boldsymbol{y})$ is given.

Proposition 4.1. If $u(\cdot,\boldsymbol{y})$ is a solution of the model problem, then

$$\|\partial_{\boldsymbol{y}}^\alpha u(\cdot,\boldsymbol{y})\|_{L_\infty(B,H_0^1(D))} \le c_{a,|\alpha|}|\alpha|!\rho^\alpha\|u(\cdot,\boldsymbol{y})\|_{H_0^1(D)}, \forall\alpha\in\mathbb{N}^n, \tag{4.10}$$

where $\rho^\alpha=\prod_{i=1}^n\rho_i^{\alpha_i}$ and $c_{a,|\alpha|}$ is a constant depending on the random process $a(\boldsymbol{x};\omega)$ and $|\alpha|$.

We now examine the Taylor expansion of $u(\cdot,\boldsymbol{y})$ around $\boldsymbol{y}=0$, which takes the form

$$u(\cdot,\boldsymbol{y}) = \sum_{|\alpha|\le p}\frac{\partial_{\boldsymbol{y}}^\alpha u(\cdot,0)}{\alpha!}\boldsymbol{y}^\alpha + \sum_{|\alpha|>p}\frac{\partial_{\boldsymbol{y}}^\alpha u(\cdot,0)}{\alpha!}\boldsymbol{y}^\alpha, \tag{4.11}$$

where $\boldsymbol{y}^\alpha=\prod_{i=1}^n y_i^{\alpha_i}$. In particular, we consider the terms satisfying $|\alpha|=p+1$, which contribute most to the error estimate. Using proposition 4.1, the terms with $|\alpha|=p+1$ can be written as

$$\|\sum_{\alpha=p+1}\frac{\partial_{\boldsymbol{y}}^\alpha u(\cdot,0)}{\alpha!}\boldsymbol{y}^\alpha\|_{L^\infty(B,H_0^1(D))} \approx \sum_{|\alpha|=p+1} c_{a,\alpha}T(\alpha)\rho^\alpha, \tag{4.12}$$

where $T(\alpha)=|\alpha|!/\alpha!$.

4.2.2 Spectral expansion

We consider the spectral expansion of $u(\cdot,\boldsymbol{y})$ as

$$u=\sum_{|\alpha|=0}^{\infty}u_\alpha\phi_\alpha(\boldsymbol{y}),$$

where $\{\phi_\alpha\}$ is the normalized orthogonal basis. By comparing the spectral expansion and the Taylor expansion, we obtain

$$u_\alpha = \sum_{|\beta|=0}^{\infty} \frac{\partial_{\boldsymbol{y}}^\beta u(\cdot,0)}{\beta!} \mathbb{E}[\boldsymbol{y}^\beta \phi_\alpha(\boldsymbol{y})]. \tag{4.13}$$

By noting the fact that

$$\boldsymbol{y}^\alpha = \sum_{|\beta|=0}^{|\alpha|} a_\beta \phi_\beta(\boldsymbol{y}),$$

where the coefficients a_β can be uniquely determined, and the orthogonality of $\{\phi_\alpha\}$, we obtain

$$u_\alpha = \sum_{\beta_i \geq \alpha_i} \frac{\partial_{\boldsymbol{y}}^\beta u(\cdot,0)}{\beta!} \mathbb{E}[\boldsymbol{y}^\beta \phi_\alpha(\boldsymbol{y})]. \tag{4.14}$$

For the error estimate we are interested in the terms with polynomial order $|\alpha| \geq p+1$. Specifically, we consider the dominant terms

$$\|u-u^p\|^2_{L_2(B,H_0^1(D))} \approx \sum_{|\alpha|=p+1} \int_D (\nabla u_\alpha)^2 \mathrm{d}\boldsymbol{x},$$

where we assume that the polynomial chaos basis $\{\phi_\alpha\}$ has been normalized. If in the approximation of u_α the term satisfying $\beta=\alpha$ is dominant, we can rewrite the error approximation as

$$\|u-u^p\|^2_{L_2(B,H_0^1(D))} \approx \sum_{|\alpha|=p+1} c_{a,\alpha}^2 T^2(\alpha) \mathbb{E}[\boldsymbol{y}^\alpha \phi_\alpha]^2 \rho^{2\alpha}, \tag{4.15}$$

where we separate the factor ρ^α from $\|\partial_{\boldsymbol{y}}^\alpha u(\cdot,0)\|_{H_0^1(D)}$ motivated by the proposition 4.1, and $c_{a,\alpha}$ is constant satisfying

$$c_{a,\alpha} = \frac{\|\partial_{\boldsymbol{y}}^\alpha u(\cdot,0)\|_{H_0^1(D)}}{|\alpha|!\rho^\alpha}. \tag{4.16}$$

Based on Eq. (4.15) we aim to reduce the cost to obtain the *a posteriori* error estimate η.

4.3 Estimation of η

4.3.1 Patterns of α

We first classify the patterns of α for a given $|\alpha|=p+1$.

Definition 4.1. For $\alpha \in \mathbb{N}^n$, the pattern of α is defined as

$$s_\alpha = [\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_n}], \quad \alpha_{i_j} \geq \alpha_{i_k}, \quad \forall j<k,$$

where $i_j = 1,2,\cdots,n$. We let $S_{|\alpha|}$ indicate the set $\{s_\alpha^j \,||\alpha| \text{ is a constant}\}, j=1,2,\cdots,n_S$, where n_S is the cardinality of $S_{|\alpha|}$.

The cardinality n_S of $S_{|\alpha|}$ is a function of n and $|\alpha|$. For example, let $|\alpha|=3$ and $n\geq 3$. s_α has at most the following three possibilities

$$s_\alpha=[3,0,0,\cdots,0],[2,1,0,\cdots,0],[1,1,1,\cdots,0].$$

For $|\alpha|=2,3,4,5$, it can be easily verified that

$$n_S=\begin{cases} 2,2,3,3, & \text{if } n=2,\\ 2,3,4,5, & \text{if } n\geq 3. \end{cases} \tag{4.17}$$

In other words, $n_S\leq|\alpha|$ for $|\alpha|\leq 5$.

Based on $s_{|\alpha|}$ we group the index α satisfying $|\alpha|=p+1$. Let $Z_{|\alpha|=p+1}=\{\alpha||\alpha|=p+1\}$. We define

$$Z^i_{|\alpha|=p+1}=\{\alpha|s_\alpha=s^i_\alpha,s^i_\alpha\in S_{|\alpha|=p+1}\},\quad i=1,\cdots,n_S. \tag{4.18}$$

Thus

$$Z_{|\alpha|=p+1}=\bigcup_{i=1}^{n_S} Z^i_{|\alpha|=p+1},$$

where $Z^i_{|\alpha|=p+1}\bigcap Z^j_{|\alpha|=p+1}=\emptyset,\forall i\neq j$.

4.3.2 A reduced space V^p

Let $q=1$. We now take a closer look at the subspace

$$Y_{p,1}=\{\phi_\alpha(\boldsymbol{y})|\phi_\alpha=\prod_{i=1}^{n}\phi_{\alpha_i}(y_i),|\alpha|=p+1\}. \tag{4.19}$$

We divide the space $Y_{p,1}$ according to the patterns of α as

$$Y_{p,1}=\bigcup_{i=1}^{n_S} A_i, A_i=\{\phi_\alpha|\alpha\in Z^i_{|\alpha|=p+1}\}. \tag{4.20}$$

For each A_i, we define a subset $A'_i\subset A_i$ with

$$A'_i=\{\phi_\alpha|\text{the corresponding } c_{a,\alpha}\mathbb{E}[\boldsymbol{y}^\alpha\phi_\alpha]\rho^{2\alpha} \text{ are the largest } \theta_s n_{A_i} \text{ ones for } \phi_\alpha\in A_i\},$$

where $0<\theta_s<1$ is a prescribed constant and n_{A_i} is the cardinality of A_i. We now construct the reduced space V^p as

$$V^p=\bigcup_{i=1}^{n_S} A'_i. \tag{4.21}$$

It is obvious that we keep the most important modes ϕ_α for each $s^i_\alpha\in S_{|\alpha|=p+1}$ based on the values $c_{a,\alpha}\mathbb{E}[\boldsymbol{y}^\alpha\phi_\alpha]^2\rho^{2\alpha}$, and

$$\frac{\dim(V^p)}{\dim(Y_{p,1})}=\theta_s<1.$$

Instead of using the space $Y_{p,1}$, we want to estimate η on the reduced space V^p. We note that the interaction between V^p and the rest modes in $Y_{p,1}$ is ignored. Let $\tilde{\eta}$ indicate the error estimate given by V^p. Since the number of modes in V^p can be significantly reduced compared to $Y_{p,1}$, the error η will be underestimated if the eigenvalues ρ_i decrease slowly. The simplest way to improve $\tilde{\eta}$ is to obtain a constant $\tilde{c}$ such that $\eta \sim \tilde{c}\tilde{\eta}$.

4.3.3 Estimation of $\tilde{c}$

Both values of η and $\tilde{\eta}$ can be divided into n_S parts according to the patterns s_α^i in $S_{|\alpha|=p+1}$, and η can be rewritten as

$$\eta^2 = \sum_{i=1}^{n_S} \eta_{s_\alpha^i}^2 \sim \sum_{i=1}^{n_S} \sum_{\phi_\alpha \in A_i} c_{a,\alpha} T(\alpha) \mathbb{E}[\boldsymbol{y}^\alpha \phi_\alpha]^2 \rho^{2\alpha}. \tag{4.22}$$

Correspondingly, $\tilde{\eta}$ takes the form

$$\tilde{\eta}^2 = \sum_{i=1}^{n_S} \tilde{\eta}_{s_\alpha^i}^2. \tag{4.23}$$

Thus, using ρ we can estimate the following ratios

$$\frac{\eta_{s_\alpha^i}^2}{\tilde{\eta}_{s_\alpha^i}^2} = \frac{\sum_{\phi_\alpha \in A_i} c_{a,\alpha} \mathbb{E}[\boldsymbol{y}^\alpha \phi_\alpha]^2 \rho^{2\alpha}}{\sum_{\phi_\alpha \in A_i'} c_{a,\alpha} \mathbb{E}[\boldsymbol{y}^\alpha \phi_\alpha]^2 \rho^{2\alpha}}. \tag{4.24}$$

We now take a closer look at the factor $T(\alpha)$ for each s_α^i. We note that $T(\alpha) = |\alpha|!/\alpha!$, which depends on s_α only. If $T(\alpha)$ is the same for all s_α^i, it is easy to verify that

$$\min_{s_\alpha^i} \frac{\eta_{s_\alpha^i}^2}{\tilde{\eta}_{s_\alpha^i}^2} \leq \frac{\eta^2}{\tilde{\eta}^2} \leq \max_{s_\alpha^i} \frac{\eta_{s_\alpha^i}^2}{\tilde{\eta}_{s_\alpha^i}^2}, \tag{4.25}$$

which implies that we can choose $\tilde{c}$ as

$$\tilde{c}^2 = \max_{s_\alpha^i} \frac{\eta_{s_\alpha^i}^2}{\tilde{\eta}_{s_\alpha^i}^2}. \tag{4.26}$$

We note that the terms $\mathbb{E}[\boldsymbol{y}^\alpha \phi_\alpha]$ are the same for a particular pattern s_α if the components of $\boldsymbol{y}$ are i.i.d random variables. In ME-gPC for arbitrary probability measures the original PDFs of y_i are decomposed simultaneously with the parametric space. $\mathbb{E}[\boldsymbol{y}^\alpha \phi_\alpha]$ will change correspondingly and can be regarded as a measure of such a change.

5 Algorithm for *h*-type adaptive ME-gPC

We define a global error as

$$\eta_g = \|u-u^p\|_{L_2(B,H_0^1(D))} \tag{5.1}$$

and a local error in element k as

$$\eta_k = \|u-u^p\|_{L_2(B_k,H_0^1(D))}. \tag{5.2}$$

We have the following lemma.

Lemma 5.1. *The relation between η_g and η_k is*

$$\eta_g^2 = \sum_{k=1}^{N} \eta_k^2 \Pr(I_{B_k=1}). \tag{5.3}$$

Proof. Observe that

$$\begin{aligned}
\eta_g^2 &= \int_B \int_D (\nabla(u-u^p))^2 f(\boldsymbol{y}) \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{y} \\
&= \int_B \int_D (\nabla(u-\sum_{k=1}^{N} u_k^p(\boldsymbol{y}_k(\boldsymbol{y}))I_{B_k}))^2 f(\boldsymbol{y}) \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{y} \\
&= \sum_{k=1}^{N} \Pr(I_{B_k}=1) \int_{B_k} \int_D (\nabla(u-u_k^p))^2 f_k(\boldsymbol{y}_k|I_{B_k}=1) \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{y}_k \\
&= \sum_{k=1}^{N} \Pr(I_{B_k}=1)\eta_k^2,
\end{aligned}$$

which leads to (5.3). □

5.1 Convergence rate of η_g

If the importance (degree of perturbation) of each random dimension is the same, we expect an *h*-type convergence rate

$$\eta_g \sim \mathcal{O}\left(N^{-(p+1)/n}\right),$$

where N is the *total* number of random elements [7, 9, 21]. We note that the algebraic index decreases by the factor $\frac{1}{n}$. Under an analyticity assumption – the eigenvalues of the K-L expansion of $a(\boldsymbol{x};\boldsymbol{y})$ decrease exponentially, the following convergence rate was proved in [21]

$$\eta_g \sim \mathcal{O}(N_{ace}^{-p+o(1)}), \quad n \to \infty \tag{5.4}$$

for the sparse wavelet approximation, where N_{ace} is the number of deterministic PDEs to be solved. We note here that N_{ace}/N is constant since we use the same polynomial order

in each random element. In this work, we relax the optimal assumption and implement adaptivity based on the *a posteriori* error estimator. In practice, the optimal convergence rate in Eq. (5.4) is difficult to maintain due to the low regularity of covariance kernel, e.g., the exponential kernel. The convergence rate can be weakened greatly by the dimensionality. Thus, we focus on a moderate number $\mathcal{O}(10)$ of random dimensions.

5.2 Criteria of adaptivity

The relation between η_g and η_k provides a natural strategy for the *h*-type adaptivity of ME-gPC. We define another local error indicator as

$$\hat{\eta}_k = \eta_k \Pr(I_{B_k}=1)^{1/2} \approx \tilde{c}_k \tilde{\eta}_k \Pr(I_{B_k=1})^{1/2}, \tag{5.5}$$

where $\tilde{c}_k$ is the correction factor of the error estimator $\tilde{\eta}_k$ from the reduced space V_k^p in element k. We note that $\hat{\eta}_k$ includes an extra factor: the square root of $\Pr(I_{B_k}=1)$. The value of $\tilde{c}_k\tilde{\eta}_k$ indicate the approximation error of the local gPC while $\Pr(I_{B_k}=1)$ is the probability that the random variable $\boldsymbol{y}$ is located in the element B_k. These two factors should be balanced for the *h*-type refinement. For example, it may not be necessary to refine an element where η_k is large while $\Pr(I_{B_k}=1)$ is very small.

When the global error estimator is larger than a threshold, then *h*-type refinement is needed; however, it is not wise to refine all the random elements. We examine the values of $\hat{\eta}_k$ for each element and refine the elements satisfying

$$\text{Criterion I:} \quad \hat{\eta}_k \geq \theta_I \max_{i=1,\cdots,N} \hat{\eta}_i, \quad 0<\theta_I<1, \tag{5.6}$$

where θ_I is a prescribed constant.

Alternatively, we can sort the local error indicators $\hat{\eta}_k$

$$\hat{\eta}_{i_1}, \hat{\eta}_{i_2}, \cdots, \hat{\eta}_{i_N}, \quad \hat{\eta}_{i_m} \geq \hat{\eta}_{i_n}, \ \forall m<n, \tag{5.7}$$

where $(i_1, i_2, \cdots, i_N)$ is a permutation of $(1,2,\cdots,N)$, and refine a certain percent of the total elements. This criterion can be expressed as

$$\text{Criterion II:} \quad \hat{\eta}_{i_j} \ \text{ with } j \leq \theta_{II} N, \quad 0<\theta_{II}<1. \tag{5.8}$$

In other words, Criterion I refines the elements according to a threshold for the local error indicator without any restriction on the number of elements while Criterion II refines a certain percent of the total elements with the largest local error indicators.

5.3 Importance of random dimension

Another question about adaptivity is how to refine the local *multi-dimensional* random element. It is too expensive to refine all the random dimensions due to the "curse of dimensionality", e.g., bisection in each random dimension will result in 2^n child elements.

Thus we need to consider the importance of each random dimension. In this work, we use the value ρ_i to measure the importance of each random dimension and choose just a few random dimensions to refine while leave the rest unchanged. If ρ_i decreases fast, the random dimension corresponding to the largest ρ_i should be refined; otherwise, we can relax the restriction and refine the first two or three most important random dimensions simultaneously. The support of each random dimension is divided into two equidistant elements.

5.4 Revisiting the correction factor $\tilde{c}$

In the construction of V^p and the estimation of $\tilde{c}$, we include the constant $c_{a,\alpha}$. However, such a constant is, in general, unknown. Furthermore, in different random elements the random process $a(\boldsymbol{x},\boldsymbol{y})$ takes different forms $a_k(\boldsymbol{x},\boldsymbol{y}_k)$:

$$a_k(\boldsymbol{x},\boldsymbol{y}_k)=\mathbb{E}[a](\boldsymbol{x})+\sigma\sum_{i=1}^{M}\frac{b_{k,i}+a_{k,i}}{2}\sqrt{\lambda_i}h_i(\boldsymbol{x})+\sigma\sum_{i=1}^{M}\frac{b_{k,i}-a_{k,i}}{2}\sqrt{\lambda_i}h_i(\boldsymbol{x})y_{k,i},$$

where the element $\times_{i=1}^{M}[a_{k,i},b_{k,i}]$ is mapped to a unit hypercube $[-1,1]^M$. Thus, $c_{a_k,\alpha}$ is also element-dependent.

We first present our strategy to deal with $c_{a_k,\alpha}$ and explain it subsequently.

1. Compute $c_{a,\alpha}$ in the global parametric space using the gPC basis. In other words, we need to solve the error equation in the full space $Y_{p,1}$.
2. Construct the reduced space V_k^p in each random element using the coefficients $c_{a,\alpha}$.
3. Take into account the relative fluctuation of $c_{a_k,\alpha}$ with respect to $c_{a,\alpha}$.

We start from the following observation:

Proposition 5.1. If in element $\times_{i=1}^{M}[a_{k,i},b_{k,i}]$ the linear mapping (3.7) is defined, then the following relation holds

$$\|\partial_{\boldsymbol{y}}^{\alpha}u(\boldsymbol{x},\boldsymbol{y}=\boldsymbol{y}_c)\|_{H_0^1(D)}\prod_{i=1}^{M}\left(\frac{b_{k,i}-a_{k,i}}{2}\right)^{\alpha_i}=\|\partial_{\boldsymbol{y}_k}^{\alpha}u_k(\boldsymbol{x},0)\|_{H_0^1(D)} \tag{5.9}$$

where $u(\boldsymbol{x},\boldsymbol{y})$ and $u_k(\boldsymbol{x},\boldsymbol{y}_k)$ are global and local solutions, respectively, and

$$y_{c,i}=\frac{b_{k,i}+a_{k,i}}{2}.$$

Proof. We note that $u(\boldsymbol{y})=u_k(\boldsymbol{y}_k(\boldsymbol{y}))$. The conclusion follows from the chain rule of differentiation. □

Using the Eqs. (4.16) and (5.9) we can obtain the relation between $c_{a_k,\alpha}$ and $c_{a,\alpha}$ as

$$\frac{c_{a_k,\alpha}}{c_{a,\alpha}}=\frac{\|\partial_{\boldsymbol{y}}^{\alpha}u(\boldsymbol{x},\boldsymbol{y}_c)\|_{H_0^1(D)}}{\|\partial_{\boldsymbol{y}}^{\alpha}u(\boldsymbol{x},0)\|_{H_0^1(D)}}. \tag{5.10}$$

We now examine the fluctuation of $\|\partial_{\boldsymbol{y}}^{\alpha}u(\boldsymbol{x},\boldsymbol{y}_c)\|_{H_0^1(D)}$ around $\|\partial_{\boldsymbol{y}}^{\alpha}u(\boldsymbol{x},0)\|_{H_0^1(D)}$. For simplicity and without loss of generality we consider the following one-dimensional model:

$$-\nabla\cdot((b(\boldsymbol{x})+\delta b(\boldsymbol{x})\xi)\nabla u(\boldsymbol{x},\xi))=f(\boldsymbol{x}), \tag{5.11}$$

where $\delta b(\boldsymbol{x})$ is a perturbation function around $b(\boldsymbol{x})$ and ξ is a random variable with zero mean on $[-1,1]$.

Applying the operator ∂_{ξ}^{m} to both sides of Eq. (5.11), we obtain

$$-\nabla\cdot((b+\delta b\xi)\nabla\partial_{\xi}^{m}u(\boldsymbol{x},\xi))=m\nabla\cdot(\delta b\nabla\partial_{\xi}^{m-1}u(\boldsymbol{x},\xi)). \tag{5.12}$$

When $\xi=0$, Eq. (5.12) becomes

$$-\nabla\cdot(b\nabla\partial_{\xi}^{m}u(\boldsymbol{x},0))=m\nabla\cdot(\delta b\nabla\partial_{\xi}^{m-1}u(\boldsymbol{x},0)).$$

For a particular value $\xi=\xi_i$, we consider the change of $\partial_{\xi}^{m}u(\boldsymbol{x},\xi_i)$ with respect to $\partial_{\xi}^{m}u(\boldsymbol{x},0)$. Let $u_{\xi}^{m}(\boldsymbol{x},\xi_i)=\partial_{\xi}^{m}u(\boldsymbol{x},\xi_i)$. We write $\partial_{\xi}^{m}u(\boldsymbol{x},\xi_i)$ as

$$u_m^{\xi}(\boldsymbol{x},\xi_i)=u_{\xi}^{m}(\boldsymbol{x},0)+\delta u_{\xi}^{m}(\boldsymbol{x},0),$$

which satisfies the following equation

$$-\nabla\cdot((b+\delta b\xi_i)\nabla(u_{\xi}^{m}(\boldsymbol{x},0)+\delta u_{\xi}^{m}(\boldsymbol{x},0)))=m\nabla\cdot(\delta b\nabla(u_{\xi}^{m-1}(\boldsymbol{x},0)+\delta u_{\xi}^{m-1}(\boldsymbol{x},0))).$$

The above equation can be simplified as

$$-\nabla\cdot((b+\delta b\xi_i)\nabla\delta u_{\xi}^{m}(\boldsymbol{x},0))=\nabla\cdot(\delta b\xi_i\nabla u_{\xi}^{m}(\boldsymbol{x},0))+m\nabla\cdot(\delta b\nabla\delta u_{\xi}^{m-1}(\boldsymbol{x},0)).$$

For a low-order derivative and small perturbation, we neglect the second-order terms and obtain

$$-\nabla\cdot(b\nabla\delta u_{\xi}^{m}(\boldsymbol{x},0))=\nabla\cdot(\delta b\xi_i\nabla u_{\xi}^{m}(\boldsymbol{x},0)), \tag{5.13}$$

which yields

$$(b\nabla\delta u_{\xi}^{m}(\boldsymbol{x},0),\nabla\delta u_{\xi}^{m}(\boldsymbol{x},0))=-(\delta b\xi_i\nabla u_{\xi}^{m}(\boldsymbol{x},0),\nabla\delta u_{\xi}^{m}(\boldsymbol{x},0)) \tag{5.14}$$

by Green's formula. The above equation implies that it is reasonable to use the degree of perturbation of the random inputs to model the fluctuation of $c_{a_k,\alpha}$ with respect to $c_{a,\alpha}$.

Finally, for the elliptic problem we take the value of $\tilde{c}$ as

$$\tilde{c}^2=1+\left(\frac{1}{n_S}\sum_{s_{\alpha}^i}\frac{\eta_{s_{\alpha}^i}^2}{\tilde{\eta}_{s_{\alpha}^i}^2}-1\right)\left(\frac{1+\sigma_c}{1-\sigma_c}\right)^2, \tag{5.15}$$

where σ_c is degree of perturbation of $a(\boldsymbol{x},\boldsymbol{y})$.

Remark 5.1. We note that we take the mean of $\eta_{s_{\alpha}^i}^2/\tilde{\eta}_{s_{\alpha}^i}^2$ instead of the maximum since we have a new factor $(1+\sigma_c)/(1-\sigma_c)$. When the number of terms in the reduced space increases, $\tilde{c}$ goes to 1, in other words, V^p becomes $Y_{p,1}$.

Algorithm 5.1: ***h*-type adaptive ME-gPC**

1. **Choose** a global tolerance ϵ and a steering parameter $0<\theta<1$.
2. **Compute** $c_{a,\alpha}$ by solving the error equation on the space $Y_{p,1}$.
3. **Construct** local polynomial chaos basis, implement gPC element-by-element.
4. **Sort** the values $\rho_{k,i}\approx c_{k,i}\sqrt{\lambda_i}$ in element k to measure the importance of each random dimension.
5. **Construct** the reduced space V_k^p.
6. **Compute** the local error indicator $\hat{\eta}_k$ for new random elements.
7. **if** the global error indicator η_g satisfies

$$\eta_g=\left(\sum_{k=1}^{N}\hat{\eta}_k^2\right)^{1/2}>\epsilon,$$

   then
8. **if** $\hat{\eta}_k$ satisfies Criteria (I) or (II) **then**
9. Refine the first two or three leading random dimensions.
10. **end if**
11. **Go** to step 3.
12. **else**
13. **Stop** and exit.
14. **end if**

5.5 Discussion of cost

Since the PDE system for the polynomial chaos coefficients can be decoupled [7], we can use the number of PDEs that needs to be solved as a measure of the overall cost. We assume that the adaptive ME-gPC simulation leads to N random elements. If the space $Y_{p,1}$ is used, the total cost for the *a posteriori* error estimates is $\mathcal{O}(\dim(Y_{p,1})NT_1)$, where T_1 is the time for resolving one deterministic PDE; if the reduced space V^p is used the cost is $\mathcal{O}(\dim(Y_{p,1})T_1+\dim(V^p)(N-1)T_1)$. When N is large enough, the ratio of the two costs is about θ_s (see Section 4.3.2). We note here that unlike the deterministic adaptive methods, we do not need to resolve the problem in the entire parametric space when splitting occurs; we only need to solve the local problems in the newly added random elements.

6 Numerical results

In this section we present a numerical study on the proposed *a posteriori* error estimators and the adaptive ME-gPC method using stochastic algebraic and elliptic model problems.

6.1 The stochastic algebraic model

To illustrate the effectiveness of the presented concepts, we first consider the following algebraic equation

$$(\beta_0+\sigma\sum_{i=1}^{M}\sqrt{\lambda_i}y_i)u=1, \tag{6.1}$$

where β_0 and σ are constants, λ_i are eigenvalues from a second-order random process, and $y_i\sim U[-1,1]$ are i.i.d. uniform random variables with zero mean and unit variance. To avoid the singularity at zero we assume that

$$\beta_0+\sigma\sum_{i=1}^{M}\sqrt{\lambda_i}y_i>0.$$

We can appreciate that such an algebraic equation shares common properties with the elliptic model problem with random coefficients. For the algebraic model we consider the norm $\|\cdot\|_{L_2(B)}$.

We consider the Taylor expansion of the solution. The term $\partial_{\boldsymbol{y}}^{\alpha}u(0)$ takes the form

$$\partial_{\boldsymbol{y}}^{\alpha}u(0)=\rho^{\alpha}\partial_z^{|\alpha|}u(z=0), \tag{6.2}$$

where $z=\sigma\sum_{i=1}^{M}\sqrt{\lambda_i}y_i$ and $\rho_i=\sigma\sqrt{\lambda_i}$. Thus, the Taylor expansion of the solution can be expressed as

$$u(\boldsymbol{y})=\sum_{\alpha}\frac{\partial_{\boldsymbol{y}}^{\alpha}u(0)}{\alpha!}\boldsymbol{y}^{\alpha}=\sum_{\alpha}\frac{\partial_z^{|\alpha|}u(z=0)}{\alpha!}\rho^{\alpha}\boldsymbol{y}^{\alpha}. \tag{6.3}$$

We note that given a particular point $\boldsymbol{y}_i$,

$$\partial_{\boldsymbol{y}}^{\alpha}u(\boldsymbol{y}_i)/\partial_{\boldsymbol{y}}^{\alpha}u(0)=\partial_z^{|\alpha|}u(z=z(\boldsymbol{y}_i))/\partial_z^{|\alpha|}u(z=0)=\text{constant},$$

where the constant only depends on $|\alpha|$ instead of α.

We consider two typical covariance kernels $\exp(-(x-y)^2/A)$ and $\exp(-|x-y|/A)$, where $x,y\in[0,1]$ and A is the correlation length. We note that the decay rate of eigenvalues is exponential for the Gaussian kernel and asymptotically algebraic for the exponential kernel. In this work we focus on $M=10$ random dimensions and consider two correlation lengthes: $A=1$ and 0.1. When $A=1$, $\lambda_{10}/\lambda_1=3.4\times10^{-3}$,$2.8\times10^{-14}$ for exponential and Gaussian kernels, respectively; when $A=0.1$, $\lambda_{10}/\lambda_1=1.1\times10^{-1}$,$5.7\times10^{-6}$ correspondingly.

6.1.1 Error estimates

In Figs. 1 and 2 we present the error contribution u_{α}^2 of each term in $Y_{p,1}$ for $p=1,2$, respectively, using the exponential kernel. It is observed that for each pattern s_{α}^i, the error contribution shows an overall decreasing trend, which is determined by the factor ρ^{α}

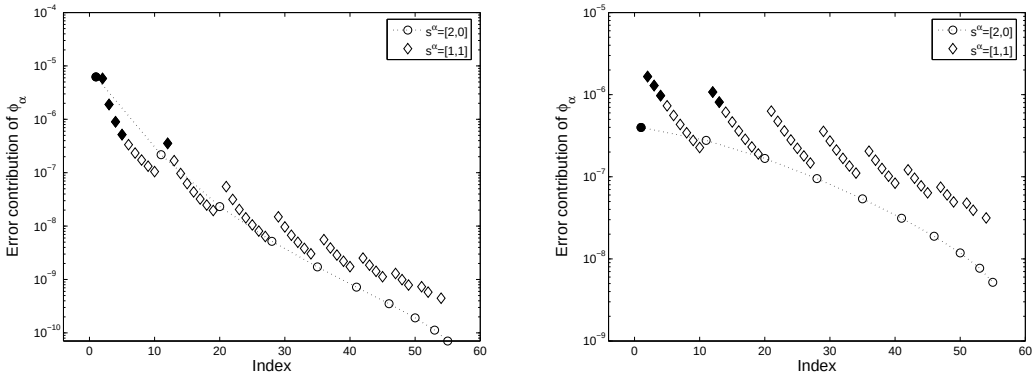

Figure 1: Error contribution of each mode in $Y_{p,1}$ with $p=1$ and $M=10$. Eigenvalues are from kernel $\exp(|x-y|/A)$ in $[0,1]$. The markers of the largest 10% terms for each s^i_α are filled. Left: $A=1$; Right: $A=0.1$.

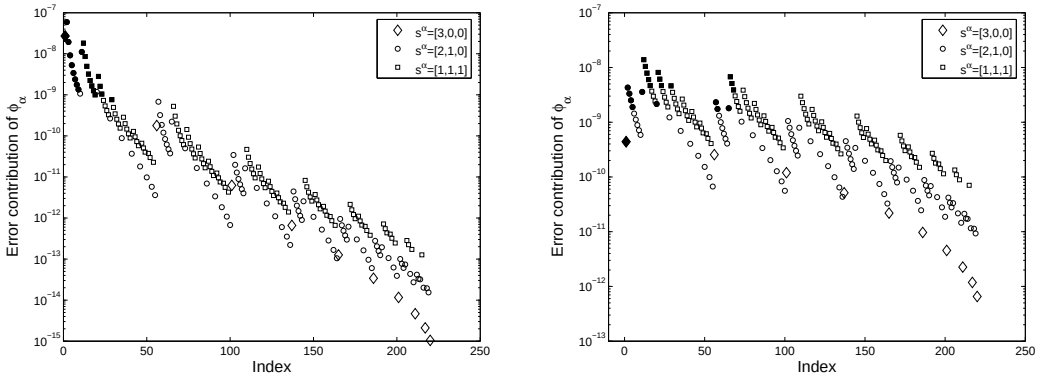

Figure 2: Same as Fig. 1, except with $p=2$ and $M=10$.

since the factors $T(\alpha)$ are comparable for each pattern s^i_α, i.e., the decay rate is determined by the decay rate of eigenvalues. We note that only the first two or three components of s^i_α are kept in the legends since the rest are zero. In Fig. 2, similar behaviors are observed. In Figs. 3 and 4, we plot the values $u^2_\alpha/\rho^{2\alpha}$, which are almost constant for each s^i_α. This implies that it is reasonable to separate the factor ρ^α in the error estimate. Furthermore, it is observed that for a fixed polynomial order the profiles of $u^2_\alpha/\rho^{2\alpha}$ are almost the same for $A=0.1,1$. Such observations can be easily explained using Eqs. (6.3) and (4.15). Also, this is the motivation in constructing the reduced space based on the pattern s^i_α since we do not need to consider the constants related to each pattern.

We now examine the effectiveness of the reduced space V^p. Let η_W denote the error from the full space $Y_{p,1}$ and η_V the error from the reduced space V^p. We use the largest 10% terms for each s^i_α in space $Y_{p,1}$, which are indicated in Figs. 1 and 2 by filled markers. In Table 2 we show the values of η and $\tilde{c}\tilde{\eta}$ for different cases. It is observed the *a posteriori* error estimator $\tilde{c}\tilde{\eta}$ is effective for the algebraic model with about 90% savings.

Table 2: *A posteriori* error estimates given by $Y_{p,1}$ and V^p. $\dim(V^p)/\dim(Y_{p,1})=0.1$. Eigenvalues are from the exponential kernel.

	η_V	η_W	η_V/η_W
$A=1, p=1$	4.30e-3	4.20e-3	1.02
$A=0.1, p=1$	4.08e-3	4.06e-3	1.00
$A=1, p=2$	4.65e-4	4.47e-4	1.04
$A=0.1, p=2$	5.12e-4	5.01e-4	1.02
$A=0.1, \lambda_i=\lambda_1, p=2$	1.93e-3	1.93e-3	1.00

6.1.2 Adaptive ME-gPC

We now examine the behavior of the adaptive ME-gPC method for the algebraic problem. We let the polynomial order be $p=2$. If a uniform discretization is considered in the parametric space, the h-convergence rate of ME-gPC for the second-order moment should be $\mathcal{O}(N^{-2(2+1)/10})=\mathcal{O}(N^{-0.6})$ [7, 9, 18], which is close to the convergence rate of the standard Monte Carlo method.

It is known that if the eigenvalues decrease fast, the error contribution mainly comes from the first several random dimensions. For a limit case, if the correlation length goes to infinity, the random behavior can be described by one random variable. This implies that the most effective way to do h-type refinement is to refine the random dimension related to the largest ρ_i. Let n_r denote the number of random dimensions to be refined. We present the evolution of errors for two different refinement strategies: $n_r=1$ or 2, in Fig. 5. Both exponential and Gaussian kernels are examined. The local error estimators are obtained from spaces V^p and $Y_{p,1}$, where the correction factor $\tilde{c}$ is chosen using Eq. (4.26) since $\partial^\alpha_{\boldsymbol{y}} u(\boldsymbol{y}_i)/\partial^\alpha_{\boldsymbol{y}} u(0)=$ constant for a given $|\alpha|$. We let $\theta_{II}=0.1$ for Criterion II (see Eq. (5.8)), in other words, we refine 10% of the total number of elements when necessary. It is observed that it is more effective to refine only one random dimension. Numerical experiments show that for a certain n_r the error evolution is almost the same for $\theta_{II}=0.1$ and 0.2. (We do not include here the results for $\theta_{II}=0.2$.) All the curves can be roughly decomposed to two parts (indicated by the dotted lines). The h-convergence rate is much better in the first part where the difference between ρ_i is large. As the h-refinement goes on, ρ_i approach each other and the h-convergence rate asymptotes $\mathcal{O}(N^{-2(p+1)/n})$. Since the eigenvalues of Gaussian kernel decrease very fast when $A=1$, only the first two random dimensions contribute to the errors $\geq 10^{-10}$, and thus the h-convergence rate is $\approx\mathcal{O}(N^{-2(2+1)/2=-3})$. For the errors $\leq 10^{-10}$, the third random dimension must be included resulting in an h-convergence rate $\approx\mathcal{O}(N^{-2(2+1)/3=-2})$. In other words, the eigenvalues indicate properly the importance of each random dimension.

In Fig. 6, we present the convergence of h-type adaptive ME-gPC for different random distributions. The eigenvalues are from the exponential kernel. Firstly, we note that the reduced space V^p provides a comparable error estimator with the space $Y_{p,1}$. Secondly,

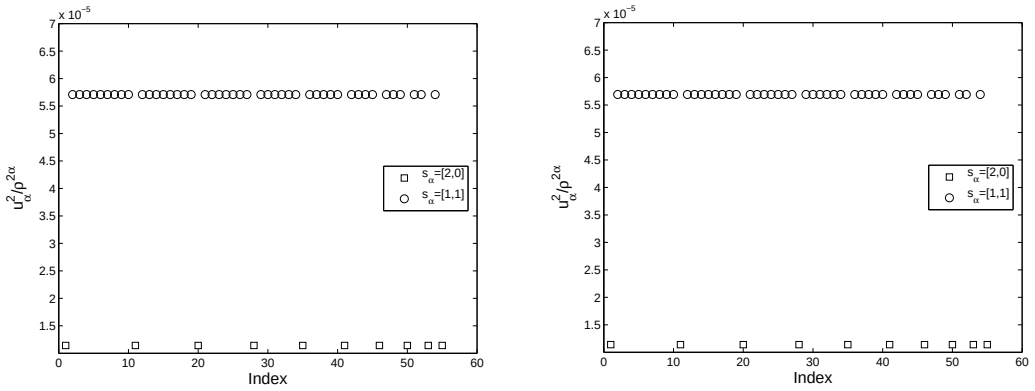

Figure 3: The factor $u_\alpha^2/\rho^{2\alpha}$ of each term in $Y_{p,1}$ with $p=1$ and $M=10$. Eigenvalues are from kernel $\exp(|x-y|/A)$ in $[0,1]$. Left: $A=1$; Right: $A=0.1$.

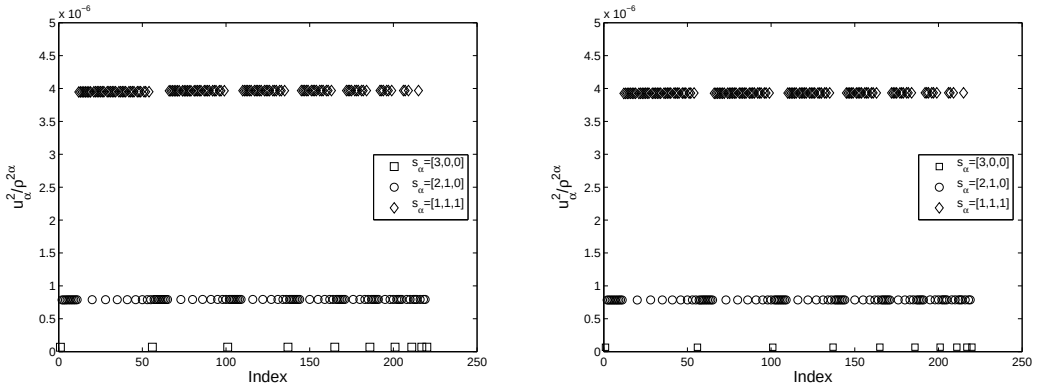

Figure 4: Same as Fig. 3, except with $p=2$ and $M=10$.

the error estimator given by V^p works well not only for the uniform distribution but also for other random distributions, although the PDFs for each random dimension in a certain random element may be different. This observation implies that the factor $\mathbb{E}[y^\alpha\phi_\alpha]\rho^\alpha$ provides a good prediction.

6.2 The stochastic elliptic model

We next consider a one-dimensional elliptic problem

$$-\frac{\mathrm{d}}{\mathrm{d}x}(a(x;\omega)\frac{\mathrm{d}u}{\mathrm{d}x})=1, \quad x\in[0,1] \tag{6.4}$$

subject to homogeneous Dirichlet boundary conditions. Using the exponential covariance kernel, we approximate $a(x;\omega)$ by the K-L decomposition as

$$a_M(x;\omega)=\mathbb{E}[a](x)+\sigma\sum_{i=1}^{M}\sqrt{\lambda_i}h_i(x)y_i,$$

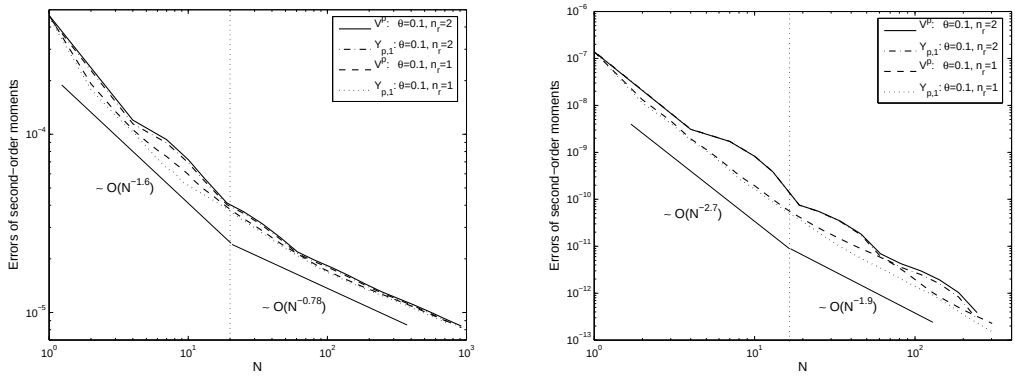

Figure 5: Adaptive errors of the second-order moments for the stochastic algebraic model. N is the number of random elements. $p=2$, $M=10$ and $A=1$. Uniform distribution is considered. Left: exponential kernel; Right: Gaussian kernel.

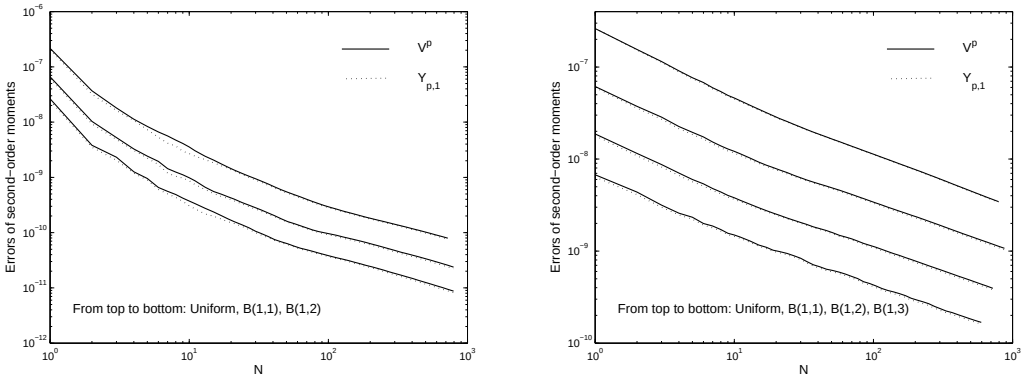

Figure 6: Adaptive errors of the second-order moments for the stochastic algebraic model. N is the number of random elements. $p=2$ and $M=10$. Uniform and Beta distributions are considered. Eigenvalues are from the exponential kernel. Left: $A=1$; Right: $A=0.1$.

where σ is standard deviation and y_i are i.i.d. random variables on $[-1,1]$. We let $\sigma=0.35$ and $M=10$. The random field is approximated by gPC while the physical space is discretized by the spectral/hp element method [27].

6.2.1 Error estimates

We now test the effectiveness of the reduced space V^p, where the norm $\|e\|_{L_2(B,H_0^1(D))}$ is used for the error estimates. We define the effectivity index as

$$C_U=\frac{\tilde{c}\tilde{\eta}}{\eta_{W^{p+1}}}, \tag{6.5}$$

where $\eta_{W^{p+1}}$ is the error estimate from the space W^{p+1} instead of $Y_{p,1}$ and $\tilde{c}\tilde{\eta}$ is the error estimate from the reduced space V^p. We compute $\tilde{c}$ using the Eq. (5.15). Since the number

Table 3: Statistics of local effectivity indices for a mesh with 1060 random elements. $p=2$ and $A=1$.

		Mean of C_U	Std of C_U
$\frac{\dim(V^p)}{\dim(Y_{p,1})}=0.1$	Uniform	1.24	0.0254
	Beta(1,1)	1.18	0.0205
	Beta(4,4)	1.11	0.0309
$\frac{\dim(V^p)}{\dim(Y_{p,1})}=0.2$	Uniform	1.17	0.0203
	Beta(1,1)	1.12	0.0189
	Beta(4,4)	1.08	0.0209

of random dimension is 10, it is hard to generate a standard mesh in the parametric space. We solve the elliptic problem adaptively up to a 1060-element mesh of the parametric space, where we let $\theta_{II}=0.2$ and $n_r=2$. Since the final meshes can be different for different PDFs of y_i, we collect the information in a statistical sense, in other words, we compute the mean and standard deviation of the local effectivity indices, which are shown in Table 3. We can see that the local effectivity indices are distributed basically around the mean value since the standard deviation is very small. In other words, the reduced space V^p can recover the information efficiently although only 10% or 20% terms in $Y_{p,1}$ are used. We note that the reduced space works well for nonlinear PDFs due to the factor $\mathbb{E}[\boldsymbol{y}^\alpha\phi_\alpha]$ (see Eq. (4.15)). The larger V^p yields better local effectivity indices as expected.

6.2.2 Adaptive ME-gPC

We next examine the adaptive behavior of ME-gPC for the stochastic elliptic model problem. In Fig. 7 we plot the global adaptive errors of the second-order moments ($H_0^1(D)$ norm in the physical space) from the space W^{p+1} and the reduced space V^p. We let $\theta_{II}=0.1$ and $n_r=1$. It appears that the *a posteriori* error estimator from V^p provides a good prediction of the true errors with a much smaller overall cost. The curves show a similar two-part structure as we observed in the stochastic algebraic model, which implies that the splitting strategy is effective. Although the PDF is decomposed by ME-gPC and the local polynomial bases may be different between two random dimensions, the proposed *a posteriori* error estimator works well for all tested PDFs.

7 Summary

In this paper we present an adaptive ME-gPC algorithm for elliptic problems with random coefficients. We first prove the hp-convergence of ME-gPC. Based on the hierarchy of polynomial chaos basis and the properties of Galerkin projection, we consider an implicit *a posteriori* error estimate. However, the problem is the cost for the error equation can be much larger than that for the original problem even for a moderate number of random dimensions ($\mathcal{O}(10)$). Based on the properties of the elliptic problems, we extract a factor ρ^α from the error contribution of each polynomial chaos mode (see Eq. (4.15)).

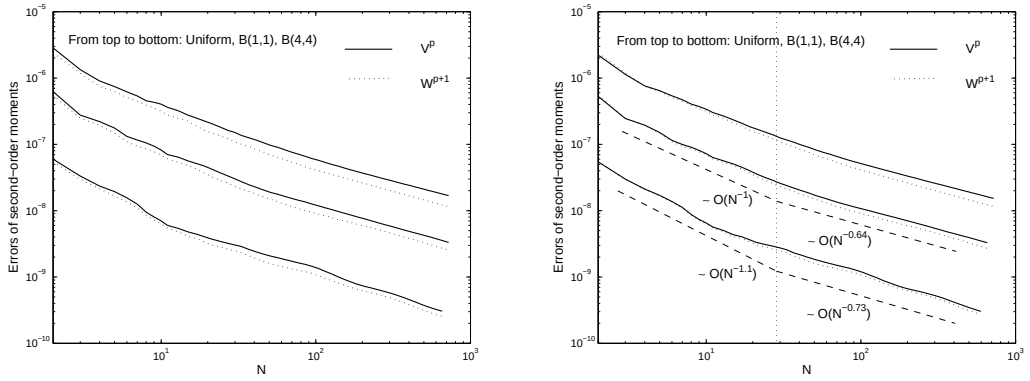

Figure 7: Adaptive errors of the second-order moments for the stochastic elliptic model. N is the number of random elements. $p=2$, $A=1$ and $M=10$. Uniform and Beta distributions are considered. Eigenvalues are from the exponential kernel. Left: $\dim(V^p)/\dim(Y_{p,1})=0.1$; Right: $\dim(V^p)/\dim(Y_{p,1})=0.2$.

The examination of the Taylor expansion implies that we can use the factor $\mathbb{E}[\boldsymbol{y}^\alpha\phi_\alpha]$ to deal with the decomposition of arbitrary PDFs, where the local polynomial chaos bases may be different between two random dimensions. We grouped the terms in the space $Y_{p,1}$ according to the pattern of index α, which helps us to get rid of the pattern related constants. Based on these observations, we constructed a reduced space V^p, which consists of a small number of terms in $Y_{p,1}$. Numerical studies on stochastic algebraic and elliptic models show that V^p can provide a good prediction of the true errors with a large saving in computation.

We do not consider the interaction between the errors of spatial discretization and polynomial chaos expansion, which should be an important issue for the adaptive methods both in physical space and in parametric space. The study of such a problem will be reported later.

Acknowledgements

This work was supported by NSF AM-SS program and also by the Computational Mathematics program of AFOSR. We would like to thank Professor Christoph Schwab for his helpful suggestions.

References

[1] M. Ainsworth and J. T. Oden. *A Posteriori Error Estimation in Finite Element Analysis*. John Wiley & Sons, 2000.

[2] R. Verfurth. *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. John Wiley & Sons Ltd and B. G. Teubner, 1996.

[3] J. T. Oden and S. Prudhomme. Estimation of modeling error in computational mechanics. *J. Comput. Phys.*, 182:496–515, 2002.

[4] J. T. Oden and K. S. Vemaganti. Estimation of local modeling error and goal-oriented adaptive modeling of heterogeneous materials: I. error estimates and adaptive algorithms. *J. Comput. Phys.*, 164:22–47, 2000.
[5] R. G. Ghanem and P. Spanos. *Stochastic Finite Elements: A Spectral Approach*. Springer-Verlag, New York, 1991.
[6] D. Xiu and G. E. Karniadakis. The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.*, 24(2):619–644, 2002.
[7] I. Babuška, R. Tempone, and G. E. Zouraris. Galerkin finite element approximations of stochastic elliptic differential equations. *SIAM J. Numer. Anal.*, 42(2):800–825, 2004.
[8] X. Wan and G. E. Karniadakis. An adaptive multi-element generalized polynomial chaos method for stochastic differential equations. *J. Comput. Phys.*, 209(2):617–642, 2005.
[9] X. Wan and G. E. Karniadakis. Multi-element generalized polynomial chaos for arbitrary probability measures. *SIAM J. Sci. Comput.*, 28(3):901–928, 2006.
[10] H. Holden, B. Oksendal, J. Uboe, and T. Zhang. *Stochastic Partial Differential Equations: A Modeling, White Noise Functional Approach*. Birkhauser, Boston, 1996.
[11] M. Kleiber and T. D. Hien. *The Stochastic Finite Element Method*. Wiley, Chichester, 1992.
[12] J. B. Roberts and P. D. Spanos. *Random Vibration and Statistical Linearization*. Wiley, Chichester, 1990.
[13] W. Liu, T. Belytschko, and A. Mani. Random field finite elements. *J. Num. Meth. Engrg.*, 23:1831–1845, 1986.
[14] W. Liu, T. Belytschko, and A. Mani. Probabilistic finite elements for nonlinear structual dynamics. *Comput. Methods Appl. Mech. Engrg.*, 56:61–81, 1986.
[15] N. Wiener. The homogeneous chaos. *Amer. J. Math.*, 60:897–936, 1938.
[16] R. H. Cameron and W. T. Martin. The orthogonal development of nonlinear functionals in series of Fourier-Hermite functionals. *Ann. Math.*, 48:385, 1947.
[17] H. G. Matthies and C. G. Bucker. Finite element for stochastic media problems. *Comput. Methods Appl. Mech. Engrg.*, 168:3–17, 1999.
[18] M. K. Deb, I. Babuška, and J. T. Oden. Solution of stochastic partial differential equations using Galerkin finite element techniques. *Comput. Methods Appl. Mech. Engrg.*, 190:6359–6372, 2001.
[19] O. P. Le Maitre, H. N. Najm, R. G. Ghanem, and O. M. Knio. Uncertainty propagation using Wiener-Haar expansions. *J. Comput. Phys.*, 197:28–57, 2004.
[20] O. P. Le Maitre, H. N. Najm, R. G. Ghanem, and O. M. Knio. Multi-resolution analysis of Wiener-type uncertainty propagation schemes. *J. Comput. Phys.*, 197:502–531, 2004.
[21] R. A. Todor and C. Schwab. Convergence rates for sparse chaos approximations of elliptic problems with stochastic coefficients. *IMA J. Numer. Anal.*, 27(2):232–261, 2007.
[22] M. Loeve. *Probability Theory, fourth ed.* Springer-Verlag, New York, 1977.
[23] P. Frauenfelder, C. Schwab, and R. A. Todor. Finite elements for elliptic problems with stochastic coefficients. *Comput. Methods Appl. Mech. Engrg.*, 194:205–228, 2005.
[24] W. Gautschi. On generating orthogonal polynomials. *SIAM J. Sci. Stat. Comput.*, 3(3):289–317, 1982.
[25] I. Babuška and B. Guo. Optimal estimates for lower and upper bounds of approximation errors in the p-version of the finite element method in two dimensions. *Numer. Math.*, 85:219–255, 2000.
[26] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems*. SIAM, Philadelphia, 2002.
[27] G. E. Karniadakis and S. J. Sherwin. *Spectral/*hp *Element Methods for CFD*. Oxford University Press, 2005.