Journal of Computational Physics 227 (2008) 9572–9595

Contents lists available at ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

The multi-element probabilistic collocation method (ME-PCM): Error analysis and applications

Jasmine Foo, Xiaoliang Wan, George Em Karniadakis *

Division of Applied Mathematics, Brown University, 182 George Street, Box F, Providence, RI 02912, USA

ARTICLE INFO

Article history:
Received 18 February 2008
Received in revised form 30 June 2008
Accepted 6 July 2008
Available online 24 July 2008

Keywords:
Domain decomposition
Stochastic partial differential equations
Sparse grids

ABSTRACT

Stochastic spectral methods are numerical techniques for approximating solutions to partial differential equations with random parameters. In this work, we present and examine the multi-element probabilistic collocation method (ME-PCM), which is a generalized form of the probabilistic collocation method. In the ME-PCM, the parametric space is discretized and a collocation/cubature grid is prescribed on each element. Both full and sparse tensor product grids based on Gauss and Clenshaw-Curtis quadrature rules are considered. We prove analytically and observe in numerical tests that as the parameter space mesh is refined, the convergence rate of the solution depends on the quadrature rule of each element only through its *degree of exactness*. In addition, the L^2 error of the tensor product interpolant is examined and an adaptivity algorithm is provided. Numerical examples demonstrating adaptive ME-PCM are shown, including low-regularity problems and long-time integration. We test the ME-PCM on two-dimensional Navier-Stokes examples and a stochastic diffusion problem with various random input distributions and up to 50 dimensions. While the convergence rate of ME-PCM deteriorates in 50 dimensions, the error in the mean and variance is two orders of magnitude lower than the error obtained with the Monte Carlo method using only a small number of samples (e.g., 100). The computational cost of ME-PCM is found to be favorable when compared to the cost of other methods including stochastic Galerkin, Monte Carlo and quasi-random sequence methods.

1. Introduction

Problems with parametric uncertainty arise in various applications from engineering, biology, and many other fields. This uncertainty may be due to either the random nature of the quantity being modeled or a lack of information about the true value of the parameter. In this paper we concentrate on a class of methods designed to calculate the *moment statistics* of solutions to PDE/ODE systems with parametric uncertainty. Our work builds heavily upon previous contributions to the field of stochastic numerical methods, which include the generalized polynomial chaos method (gPC), multi-element generalized polynomial chaos method (ME-gPC), probabilistic collocation method (PCM), and many other variants (see, e.g., [1–13] and references therein).

Calculation of moment statistics can essentially be described as a high-dimensional integration problem, where the dimensionality refers to the cardinality of random dimensions. For this reason, problems with large random dimension suffer from the same computational challenges in this field as in the field of high-dimensional numerical integration. Advances such as the use of sparse grid techniques ([2,14,15]) for numerical integration have greatly alleviated this problem, but such

* Corresponding author. Tel.: +1 401 863 1217; fax: +1 401 863 3369.
E-mail address: gk@dam.brown.edu (G.E. Karniadakis).

0021-9991/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jcp.2008.07.009

techniques rely upon the *essential regularity* of the solution in parameter space. Thus, systems with discontinuous dependence on random parameters cause difficulties for the convergence of these methods. As a result, these problems still require prohibitively large computational resources. It is worthwhile noting that currently for problems with very large dimension ($\geqslant 100$), traditional Monte Carlo (MC) methods are more attractive than any method in this class because of their favorable scaling with increasing dimension. In this work, however, we restrict ourselves to problems of *moderate* dimension, where large improvements in cost can be made relative to the cost of MC.

As mentioned before, many variations have been introduced to improve the efficiency of stochastic numerical methods including sparse grid collocation, anisotropic sparse grid collocation [16], as well as sparse polynomial bases [9] and wavelet expansions [17] for the Galerkin formulation. A multi-element formulation for the stochastic Galerkin method was proposed in [1]. This method was then generalized in [18] to deal with arbitrary probability distributions with the numerical construction of generalized polynomial chaos bases on the fly. This approach, called the multi-element generalized polynomial chaos (ME-gPC) method was found to effectively deal with problems exhibiting low regularity in parametric space as well as for long-time integration [18]. However, as with most Galerkin methods, high random dimension often necessitates a prohibitively large number of basis functions in nonlinear problems. In addition, the ME-gPC method requires derivation of a new numerical scheme and solver to deal with nonlinear problems.

In this paper we introduce a multi-element probabilistic collocation method (ME-PCM) which is an extension of the probabilistic collocation method, which was first introduced in [11], and later explored in [3,12]. The method we propose offers the advantages of *domain decomposition* in parametric space, similar to the ME-gPC method, and also the computational ease of *sampling-based* methods. In particular, we note that using the ME-PCM requires only a wrapper around a deterministic solver of the ODE/PDE for each sample calculation. Thus, nonlinear problems are significantly easier to compute using the ME-PCM instead of the ME-gPC method. We are interested in setting the theoretical foundations of the method and in answering some practical questions about the usage of the ME-PCM. Specifically, we would like to study how the choice of an integration/approximation rule in each element affects the h-convergence rate of the solution (here, h-convergence refers to the refinement of elements in the parametric space). We are also interested in analyzing how this method compares to pre−existing methods in terms of computational cost for difficult problems with low regularity in parametric space, and how well it fares in long-time integration problems.

In Section 2 the general framework and model problem are discussed. As the reader will see, most of the error analysis is confined within the context of the stochastic diffusion problem which is presented as an example in this section. However, many non-elliptic problems are addressed in the examples section, so a fairly general model problem formulation is maintained here to discuss the method. Also in this section we will introduce assumptions on the random input that are essential for ME-PCM as well as for the other methods we have discussed so far. In Section 3 ME-PCM is introduced, with two options presented for the choice of an integration/approximation rule in each element. Section 4 contains error analysis of ME-PCM and is divided into two main parts. The first of these addresses the error in moments of the solution or numerical integration error. Here, we show that the choice of integration rule affects the convergence rate only through its degree of polynomial exactness. In the second part we investigate the L^2 error of the ME-PCM approximant by building upon the previously published error analysis for the PCM method in [12] and [15]. In Section 5, numerical examples using ME-PCM for various problems are shown. We numerically verify the findings in Section 4 regarding the h-convergence rate of moment errors using simple numerical integration examples, ODEs and some two dimensional Navier-Stokes problems. We investigate the effectiveness of the adaptive ME-PCM on a problem with low regularity in parametric space space and make some comparison studies of computational costs with other methods. In the last example we solve a 50-dimensional stochastic diffusion problem with the ME-PCM and study the h-convergence of the method using an *a priori* adapted mesh. Finally, in the appendices we include details of the proofs of theorems presented in the main text as well as a short review of material on sparse grids and on adaptivity criteria.

2. Formulation

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space, where Ω is the space of events, $\mathcal{F} \subset 2^{\Omega}$ is the σ-algebra of sets in Ω, and P is the probability measure. Also, define D to be a subset of $\mathbb{R}^d (d \in \{1,2,3\})$ with boundary ∂D. Let $\mathcal{L}$ and $\mathcal{B}$ be operators on D and ∂D, respectively, where $\mathcal{L}$ may depend upon $\omega \in \Omega$. In this work we consider the following problem: find $u : \Omega \times \overline{D} \to \mathbb{R}$ such that P-almost everywhere (a.e.) in Ω the following equation holds:

$$\begin{cases} \mathcal{L}(\boldsymbol{x}, \omega; u) = f(\boldsymbol{x}, \omega), & \boldsymbol{x} \in D, \\ \mathcal{B}(\boldsymbol{x}; u) = g(\boldsymbol{x}), & \boldsymbol{x} \in \partial D. \end{cases} \tag{1}$$

We assume that the boundary has sufficient regularity and that f and g are imposed so that the problem is well-posed P-a.e. We also assume that for P-a.e. $\omega \in \Omega$, the solution $u(\cdot, \omega)$ takes values in a Banach space, $\mathcal{W}(D)$, of functions over the physical domain taking values in $\mathbb{R}$. The main goal of the ME-PCM, like other methods of its class, is to approximate the moment statistics of the solution of this problem.

In order to apply the methods that will be discussed later, the random dependence of operators $\mathcal{L}$ and f must satisfy a few important properties. The first requirement, commonly known as a "finite dimensional noise assumption" [3,12], is that the random input can be represented with a finite-dimensional probability space. More specifically, the random input can be

represented by a finite set of random variables $\{Y_1(\omega), Y_2(\omega), \ldots, Y_N(\omega)\}$, with a known joint density function ρ. With this assumption on the random input, the problem (1) can be restated as follows. Find $u : \Omega \times \overline{D} \to \mathbb{R}$ such that

$$\mathcal{L}(\boldsymbol{x}, Y_1(\omega), Y_2(\omega), \ldots, Y_N(\omega); u) = f(\boldsymbol{x}, Y_1(\omega), Y_2(\omega), \ldots, Y_N(\omega))$$

holds $\forall \boldsymbol{x} \in D$ and for P-a.e. $\omega \in \Omega$, with corresponding boundary conditions. Using the Doob-Dynkin lemma [19] we can assert that the solution $u(\boldsymbol{x}, \omega)$ can be written as $u(\boldsymbol{x}, \boldsymbol{Y}(\omega))$ with $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_N)$. Then, the problem may be recast from the space Ω into the target space of the N random variables. Let $\boldsymbol{y} = (y_1, y_2, \ldots, y_N) \in \Gamma \equiv \prod_{j=1}^{N} \Gamma_j$, where Γ_j is the image of $Y_j(\Omega)$ for $j = 1, \ldots, N$. Let $\rho(\boldsymbol{y})$ be the probability density function (PDF) of $\boldsymbol{Y}$. The problem can be restated: find $u : \Gamma \times \overline{D} \to \mathbb{R}$ such that ρ-almost everywhere for $\boldsymbol{y} \in \Gamma$ the following equation holds:

$$\begin{cases} \mathcal{L}(\boldsymbol{x}, \boldsymbol{y}; u) = f(\boldsymbol{x}, \boldsymbol{y}), & \boldsymbol{x} \in D, \\ \mathcal{B}(\boldsymbol{x}, u) = g(\boldsymbol{x}), & \boldsymbol{x} \in \partial D. \end{cases} \tag{2}$$

Thus, the original problem (1) is recast as a fully deterministic problem in Eq. (2). It is sometimes useful to think of the solution u as a function on Γ, taking values in a proper Banach space $\mathcal{W}(D)$. In this case we would denote $u(\boldsymbol{y})$ to be the Banach-valued solution to the problem for a particular $\boldsymbol{y} \in \Gamma$. Although numerical examples will be shown later in Section 5 for various differential operators $\mathcal{L}$, most of the error analysis presented in Section 4 will be confined to the following prototype elliptic problem.

2.1. A model problem: elliptic problem with stochastic input

We consider the following stochastic linear boundary value problem: find a stochastic function, $u : \Omega \times \overline{D} \to \mathbb{R}$, such that the following equation holds P-a.e:

$$\begin{cases} -\nabla \cdot (a(\boldsymbol{x}; \omega) \nabla u(\boldsymbol{x}; \omega)) = f(\boldsymbol{x}) \quad \text{in } D, \\ u(\boldsymbol{x}; \omega) = 0 \quad \text{on } \partial D, \end{cases} \tag{3}$$

where $f(\boldsymbol{x})$ is assumed to be deterministic for simplicity and $a(\boldsymbol{x}; \omega)$ is a second-order random process satisfying the following assumption:

Assumption 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,

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

Under this assumption, the problem has a unique solution u such that $u(\cdot, \omega)$ takes realizations in the space $\mathcal{W}(D) = H_0^1(D)$, P-a.e. in Ω. We can approximate $a(\boldsymbol{x}, \omega)$ using a truncated Karhunen–Loève (K–L) expansion:

$$a_N(\boldsymbol{x}; \omega) = \mathbb{E}[a](\boldsymbol{x}) + \sum_{j=1}^{N} \sqrt{\lambda_j} \phi_j(\boldsymbol{x}) Y_j(\omega), \tag{4}$$

where $\{Y_j\}_{j=1}^N$ are mutually *uncorrelated* random variables with zero mean and unit variance [7,8,20]. The eigenpairs $\{\lambda_j, \phi_j\}_{j=1}^N$ satisfy

$$\int_D R_{aa}(\boldsymbol{x}, \boldsymbol{y}) \phi_i(\boldsymbol{x}) d\boldsymbol{x} = \lambda_i \phi_i(\boldsymbol{y}), \tag{5}$$

where R_{aa} is the covariance kernel of a. The use of this truncated approximation for $a(\boldsymbol{x}; \omega)$ guarantees that the finite dimensional noise assumption is satisfied. For simplicity we make the additional assumption that Y_i are *independent* so that the density function $\rho(\boldsymbol{y}) = \prod_{j=1}^{N} \rho_j(y_j)$, where ρ_j is the density function of each Y_j.

3. ME-PCM method

In this section we describe the multi-element probabilistic collocation method (ME-PCM). The main idea of ME-PCM is to discretize the space Γ (which is assumed to be bounded) into non-overlapping elements and perform the standard probabilistic collocation method on each element. This yields approximate local moment statistics in each element, which can then be assembled to obtain global statistics. Key considerations arising in practice include finding suitable mesh discretizations, choice of collocation points, and adaptive mesh refinement. In the interest of uniformity, we adopt many of the same notations used in the works of [12] and [15] on the probabilistic collocation method in our description of the ME-PCM extension.

3.1. Spatial discretization

The method involves first discretizing the problem (2) in the physical space D using a standard finite or spectral element solver to obtain the *deterministic semidiscrete solution*. Let us define $\mathcal{W}_k(D)$ to be a standard solution subspace of $\mathcal{W}(D)$

(e.g., finite/spectral element space), containing piecewise polynomials defined on regular mesh $\mathcal{T}_{k,D}$ of D with maximum mesh spacing parameter k. Define the deterministic semidiscrete solution $u_k(\boldsymbol{y}) \equiv \pi_k u(\boldsymbol{y}) \in \mathcal{W}_k(D)$ to be the finite/spectral element ([21,22]) approximate solution of the deterministic problem (2) for each $\boldsymbol{y} \in \Gamma$, where π_k is the finite/spectral element projection operator. Thus, $u_k : \Gamma \to \mathcal{W}_k(D)$. We assume that the discretization in physical space satisfies the usual convergence property:

$$\|u(\boldsymbol{y}) - u_k(\boldsymbol{y})\|_{\mathcal{W}(D)} \leqslant C(u(\boldsymbol{y}), l)k^l, \quad \forall \boldsymbol{y} \in \Gamma, \tag{6}$$

where l is a positive number dependent on the regularity of $u(\boldsymbol{y})$ in D. In the case of the example problem (3), we have:

$$\|u(\boldsymbol{y}) - u_k(\boldsymbol{y})\|_{H_0^1(D)} \leqslant Ck^l \|u(\boldsymbol{y})\|_{H^{l+1}(D)}, \quad \forall \boldsymbol{y} \in \Gamma. \tag{7}$$

3.2. Stochastic discretization

The next step is to discretize the parametric space Γ into a nonoverlapping mesh of open hypercubes. We begin by defining $\{B^i\}_{i=1}^{N_e}$ to be a finite collection of open subsets of Γ such that $\bigcup_{i=1}^{N_e} \overline{B^i} = \Gamma$ and $B^i \cap B^j = \emptyset$ whenever $i \neq j$. We will assume for simplicity that the B^i are rectangular (i.e., $B^i = \prod_{j=1}^N B_j^i$, where $B_j^i \subset \Gamma_j$). These sets will be referred to as 'elements' of a mesh on the parametric space Γ, and thus N_e refers to the number of elements in a particular mesh discretization. Let us denote a particular mesh discretization of Γ as $\mathcal{T}_{h,\Gamma}$, where h refers to the maximum mesh spacing parameter. Once a mesh is prescribed, a set of collocation points $\{\boldsymbol{q}_j^i\}_{j=1}^r$ is prescribed in each element B^i, where r refers to the number of points used. These points are usually chosen to *coincide* with the points of an cubature rule on B^i with integration weights $\{w_j^i\}_{j=1}^r$. In this work, we consider full tensor products of Gauss quadrature points and sparse grids (see Sections A.1 and A.1.1 the appendix for details).

The semidiscrete solution u_k is then collocated on the set of points $\bigcup_{i=1}^{N_e}\bigcup_{j=1}^r \{\boldsymbol{q}_j^i\}$. In other words, at each of these points $\boldsymbol{q}_j^i$ we find the finite/spectral element solution of the deterministic problem:

$$\begin{cases} \mathcal{L}(\boldsymbol{x}, \boldsymbol{q}_j^i; u) = f(\boldsymbol{x}, \boldsymbol{q}_j^i), & \boldsymbol{x} \in D \\ \mathcal{B}(\boldsymbol{x}, u) = g(\boldsymbol{x}), & \boldsymbol{x} \in \partial D. \end{cases} \tag{8}$$

This approximate solution is denoted by $u_k(\cdot, \boldsymbol{q}_j^i) : \Gamma \mapsto \mathcal{W}_k(D)$. We are then interested in constructing a fully discrete approximant $\mathcal{I}_{B^i} u_k(\boldsymbol{x}, \boldsymbol{y})$ using the set of solutions $\{u_k(\cdot, \boldsymbol{q}_j^i)\}_{j=1}^r$ over each element B^i. For example, the operator $\mathcal{I}_{B^i}$ can be chosen to be the tensor product Lagrangian interpolant, i.e.,

$$\mathcal{I}_{B^i} u_k(\boldsymbol{x}, \boldsymbol{y}) \equiv L_{B^i}^{\mathbf{p}} u_k(\boldsymbol{x}, \boldsymbol{y}) = \sum_{j=1}^r u_k(\boldsymbol{x}, \boldsymbol{q}_j^i) l_j^i(\boldsymbol{y}), \tag{9}$$

where $l_j^i(\boldsymbol{y})$ is the Lagrange polynomial corresponding to the point $\boldsymbol{q}_j^i$ and $\mathbf{p}$ determines the degree of the interpolant in each dimension. The operator $L_{B^i}^{\mathbf{p}}$ is defined and described in more depth in the appendix, Section A.1.

Another choice for the operator $\mathcal{I}_{B^i}$ is the isotropic Smolyak sparse grid operator $\mathcal{S}_{B^i}(s)$, which was introduced in [23] by Smolyak. Here, the *sparseness parameter* s controls the order of the approximant. The construction and details of this operator are also given in the appendix, Section A.1.1.

We now define the global approximant:

$$\tilde{u}(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{N_e} \mathcal{I}_{B^i} u_k(\boldsymbol{x}, \boldsymbol{y}) \mathbb{I}_{\{\boldsymbol{y} \in B^i\}} \quad \forall \boldsymbol{x} \in \overline{D}, \quad \forall \boldsymbol{y} \in \Gamma$$

where $\mathbb{I}_{\boldsymbol{y} \in A}$ denotes the characteristic function of set A.

We subsequently consider the computation of statistics and define the *conditional probability density function* in each element:

$$\eta^i(\boldsymbol{y}) = \frac{\rho(\boldsymbol{y})}{\int_{B^i} \rho(\boldsymbol{y})\, \mathrm{d}\boldsymbol{y}}. \tag{10}$$

We assume that the density function is in tensor product form $\rho(\boldsymbol{y}) = \prod_{j=1}^N \rho_j(y_j)$, so this property is inherited by the local density $\eta^i(\boldsymbol{y}) = \prod_{j=1}^N \eta_j(y_j)$. The local mean of a function $v : \overline{D} \times B^i \to \mathbb{R}$ in an element i is given by:

$$\mathbb{E}^i[v(\boldsymbol{x}, \cdot)] = \mathbb{E}[v(\boldsymbol{x}, \cdot) | \boldsymbol{Y}(\omega) \in B^i] = \int_{B^i} v(\boldsymbol{x}, \boldsymbol{y}) \eta^i(\boldsymbol{y})\, \mathrm{d}\boldsymbol{y}.$$

Using the cubature rule over each element, we can easily compute the *approximate* local mean of $\tilde{u}$ as

$$\mathbb{E}_a^i[\tilde{u}](\boldsymbol{x}) = \sum_{j=1}^r u_k(\boldsymbol{x}, \boldsymbol{q}_j^i) w_j^i \approx \mathbb{E}^i[\tilde{u}](\boldsymbol{x}).$$

Here we use the notation $\mathbb{E}_a^i$ to denote the expected value approximation operator using numerical quadrature. Note that $\mathbb{E}_a^i$ is defined through the particular choice of mesh and collocation grids for any given ME-PCM procedure. Finally, the approximate global mean can be assembled from the local means via Bayes' formula

$$\mathbb{E}_a[\tilde{u}](\boldsymbol{x}) \equiv \sum_{i=1}^{N_e} \mathbb{E}_a^i[\tilde{u}](\boldsymbol{x}) P(\mathbf{Y}(\omega) \in B^i) \approx \mathbb{E}[\tilde{u}](\boldsymbol{x}). \tag{11}$$

Other statistics can be computed by the same procedure. For example, we can compute the energy norm in physical space of the MEPCM solution $\tilde{u}$ at each collocation point. Then, using quadrature we can compute the mean of the energy norm of the solution:

$$\mathbb{E}_a[\|\tilde{u}\|_{\mathcal{W}(D)}] \equiv \sum_{i=1}^{N_e} \left(\sum_{j=1}^{r} \|u_k(\cdot, \boldsymbol{q}_j^i)\|_{\mathcal{W}(D)} w_j^i \right) P(\mathbf{Y}(\omega) \in B^i) \approx \mathbb{E}[\|\tilde{u}\|_{\mathcal{W}(D)}]. \tag{12}$$

Recall that r denotes the number of collocation points in each element.

Remark 2. If the PDF $\rho(\boldsymbol{y})$ is uniform, grids in each element can be obtained by an affine mapping from a reference element. Otherwise, grids are, in general, element-dependent, since $\eta^i(\boldsymbol{y})$ are different in each element when the parametric space Γ is decomposed. We refer the reader to [4] for more details on constructing a local GPC bases orthogonal with respect to the conditional PDF.

Remark 3 Adaptivity. Adaptive mesh refinement is necessary when the solution has low regularity in the parametric space Γ. The adaptive procedure developed in [4,18] for the ME-gPC method can be employed directly for ME-PCM. The key idea of adaptive criterion in [4,18] is to refine an element in the parametric space when the decay of the coefficients in the chaos expansion is relatively slow, see Section A.4 in the appendix. We note that $\mathcal{I}_{B^i} u_k(\boldsymbol{x}, \boldsymbol{y})$ corresponds to a unique polynomial chaos expansion and all the local statistics used in the adaptive criterion can be computed easily by the collocation solution. It is only necessary to project the collocation solution onto the highest modes of the basis to evaluate the adaptivity criterion. General adaptive formulas for sparse grids have been developed in [24].

4. Error analysis

In this section, we present come convergence results of ME-PCM, in particular, we focus on the model problem (3).

4.1. Moments error

We first examine the relation between h-convergence rate of ME-PCM and the degree of exactness of cubature rules. We begin by defining two standard norms. For a function $f : A \to \mathbb{R}$, we consider the Sobolev space $W^{m,\infty}(A)$ with the norm:

$$\|f\|_{m,\infty,A} = \max_{|\boldsymbol{\alpha}| \leqslant m} \operatorname{ess\,sup}_{\boldsymbol{x} \in A} |D^{\alpha} f(\boldsymbol{x})|$$

and the seminorm:

$$|f|_{m,\infty,A} = \max_{|\boldsymbol{\alpha}| = m} \operatorname{ess\,sup}_{\boldsymbol{x} \in A} |D^{\alpha} f(\boldsymbol{x})|,$$

where $A \subset \mathbb{R}^N$, $\alpha \in \mathbb{N}_0^N$, $|\alpha| = \alpha_1 + \cdots + \alpha_N$ and $m \in \mathbb{N}_0$.

In this section we assume that the density function $\rho(\boldsymbol{y})$ is *uniform* for convenience in analysis. Then, a general h-convergence rate of ME-PCM is given by the following theorem:

Theorem 4. *Suppose $f \in W^{m+1,\infty}(\Gamma)$ with $\Gamma = (0,1)^N$, and $\{B^i\}_{i=1}^{N_e}$ is a nonoverlapping mesh of Γ. Let h indicate the maximum side length of each element and $\mathcal{Q}_m^{\Gamma}$ a quadrature rule with degree of exactness m in domain Γ. (In other words $\mathcal{Q}_m$ exactly integrates polynomials up to order m). Let $\mathcal{Q}_m^A$ be the quadrature rule in subset $A \subset \Gamma$, corresponding to $\mathcal{Q}_m^{\Gamma}$ through an affine linear mapping. We define a linear functional on $W^{m+1,\infty}(A)$:*

$$E_A(g) \equiv \int_A g(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} - \mathcal{Q}_m^A(g) \tag{13}$$

whose norm is defined as

$$\|E_A\|_{k,\infty,A} = \sup_{\|g\|_{k,\infty,A} \leqslant 1} |E_A(g)|. \tag{14}$$

Then the following error estimate holds:

$$\left| \int_{\Gamma} f(\boldsymbol{x}) d\boldsymbol{x} - \sum_{i=1}^{N_e} \mathcal{Q}_m^{B^i} f \right| \leqslant C h^{m+1} \|E_{\Gamma}\|_{m+1,\infty,\Gamma} |f|_{m+1,\infty,\Gamma} \tag{15}$$

where C is a constant and $\|E_{\Gamma}\|_{m+1,\infty,\Gamma}$ refers to the norm in the dual space of $W^{m+1,\infty}(\Gamma)$.

Proof. See Section A.2 in the appendix. □

Remark 5. In Theorem 4, the only information used for the h-convergence rate of ME-PCM is the degree of exactness of cubature rules. The norm of the error functional $\|E_\Gamma\|_{m+1,\infty,\Gamma}$ usually exhibits p-type convergence for polynomial interpolation. Roughly speaking, Theorem 4 shows hp-convergence of the moments error of the ME-PCM.

We subsequently present several examples based on different choices of interpolation rule $\mathcal{Q}_m^\Gamma$ and an application to the stochastic elliptic problem.

Example 1. Tensor-product Gauss grid – Let $\mathcal{Q}_m^\Gamma$ signify tensor product integration rule Γ, based upon one-dimensional Gauss formulas $\mathcal{U}_j^k$, described in appendix (Section A.1). If we choose $n+1$ points in each dimension, the degree of exactness is $m = 2n+1$.

We now consider the degree of exactness associated with sparse grids, as described in the appendix (Section A.1.1). In the sparse quadrature procedure, the sparseness parameter 's' controls the number of points in the grid. A sparseness level equal to the number of dimensions ($s = N$) corresponds to the trivial grid consisting of only one point. As s increases, the number of points and hence the accuracy of the quadrature increases. In the following we state some results from [25] relating the degree of exactness m of sparse quadrature rules to the sparseness parameter s.

Example 2. Smolyak Clenshaw-Curtis grid – Let $\mathcal{Q}_m^\Gamma$ be the Clenshaw Curtis Smolyak operator in Γ of dimension N with sparseness parameter s, which is described in Section A.1.1. There we choose the number of points for each order to be $n_1 = 1$ and $n_k = 2^{k-1}+1$ for $k > 1$. With this choice, the degree of exactness of $\mathcal{Q}_m^\Gamma$ can be stated as follows [25]: Let $\sigma = floor(s/N)$ and let $\tau = s \bmod N$. Then $\mathcal{Q}_m^\Gamma$ has degree of exactness:

$$m(s,N) = \begin{cases} 2(s-N)+1, & \text{if } s < 4N \\ 2^{\sigma-1}(N+1+\tau)+2N-1, & \text{otherwise.} \end{cases} \tag{16}$$

From [26] we also obtain a bound on the operator norm of the error functional:

$$\|E_\Gamma\|_{m+1,\infty,\Gamma} \leqslant r^{-m(s,N)}(\log r)^{(N-1)(m(s,N)+1)}. \tag{17}$$

where r denotes the total number points used in the quadrature rule.

Example 3. Smolyak Gauss grid – Let $\mathcal{Q}_m^\Gamma$ be the Smolyak-Gauss operator in Γ of dimension N with sparseness parameter s. Again we choose $n_1 = 1$ and $n_k = 2^{k-1}+1$ for $k > 1$. Then, the degree of exactness of $\mathcal{Q}_m^\Gamma$ can be stated as follows [25]: Let $\sigma = floor(s/N)$ and let $\tau = s \bmod N$. Then $\mathcal{Q}_m^\Gamma$ has degree of exactness:

$$m(s,N) = \begin{cases} 2(s-N)+1, & \text{if } s \leqslant 3N \\ 2^{\sigma-1}(N+1+\tau)-1, & \text{otherwise.} \end{cases} \tag{18}$$

An application to stochastic elliptic problem – Let u be the exact solution of the stochastic problem (3) given in Section 2.1 satisfying $u \in H_0^1(D) \cap H^2(D)$ almost surely. We assume that the coefficient $a(\boldsymbol{x};\omega)$ is represented by independent uniform random variables $(Y_1,\ldots,Y_N) \in \Gamma$ and that Γ is compact in $\mathbb{R}^N$ (see Eq. (4)). We obtain the following result:

Lemma 6. *Using a nonoverlapping mesh of Γ with maximum side length h and a collocation grid associated with a quadrature rule with degree of exactness m in each element, define $\tilde{u}$ to be the approximate solution given by ME-PCM with a spatial discretization projection π_k satisfying (7) for $l = 1$. Then the error of the energy norm $\mathbb{E}_a\|\tilde{u}\|_{H_0^1(D)}$ (defined in Eq. (12)) can be bounded as:*

$$\left|\mathbb{E}\|u\|_{H_0^1(D)} - \mathbb{E}_a\|\tilde{u}\|_{H_0^1(D)}\right| \leqslant C_1 k\|u\|_{H^2(D)} + C_2 h^{m+1}, \tag{19}$$

where the constants C_1 and C_2 are independent of k and h.

Proof. Using the triangle inequality,

$$\left|\mathbb{E}\left[\|u\|_{H_0^1(D)}\right] - \mathbb{E}_a\left[\|\tilde{u}\|_{H_0^1(D)}\right]\right| \leqslant \left|\mathbb{E}\left[\|u\|_{H_0^1(D)}\right] - \mathbb{E}\left[\|u_k\|_{H_0^1(D)}\right]\right| + \left|\mathbb{E}\left[\|u_k\|_{H_0^1(D)}\right] - \mathbb{E}_a\left[\|\tilde{u}\|_{H_0^1(D)}\right]\right| \tag{20}$$

Note that $\|u_k(\cdot,y)\|_{H_0^1(D)} \in W^{m+1,\infty}(\Gamma)$ for the problem at hand. Then, it is clear that we can obtain the bound on the second term from Theorem 4 by noting that

$$\mathbb{E}_a\|\tilde{u}\|_{H_0^1(D)} = \frac{1}{\mathrm{vol}(\Gamma)}\sum_{i=1}^{N_e}\mathcal{Q}_m^{B^i}\|u_k(\cdot,\cdot)\|_{H_0^1(D)}$$

(where the norm is taken over the spatial variable and the quadrature acts in the probability space) as well as by noting that

$$\mathbb{E}\|u_k\|_{H_0^1(D)} = \frac{1}{\mathrm{vol}(\Gamma)}\int_\Gamma \|u_k(\cdot,\boldsymbol{y})\|_{H_0^1(D)}\mathrm{d}\boldsymbol{y}.$$

Then,

$$|\mathbb{E}\|u_k\|_{H_0^1(D)} - \mathbb{E}_a\|\tilde{u}\|_{H_0^1(D)}| \leqslant C\|E_\Gamma\|_{m+1,\infty,\Gamma}|u_k|_{W^{m+1,\infty}(\Gamma;H_0^1(D))}h^{m+1}.$$

We note that u_k has the same regularity as u. Moreover, we know that there exists an analytic extension for u in the parametric space [12], which implies that $|u_k|_{W^{m+1,\infty}}$ is finite. The first term in (20) is a spatial discretization error term and

$$\left|\mathbb{E}\|u\|_{H_0^1(D)} - \mathbb{E}\|u_k\|_{H_0^1(D)}\right| \leqslant \mathbb{E}\|u - u_k\|_{H_0^1(D)} \leqslant \mathbb{E}\Big[Ck\|u\|_{H^2(D)}\Big] \leqslant Ck \cdot \mathbb{E}\|u\|_{H^2(D)}. \qquad \square$$

4.2. L^2 error of the global approximant

In this section, the error of the global approximant $\tilde{u}$ is compared with the exact solution u in the $L^2(\Gamma;H_0^1(D))$ norm:

$$\|\tilde{u} - u\|_{L^2(\Gamma;H_0^1(D))} \leqslant \|\tilde{u} - u_k\|_{L^2(\Gamma;H_0^1(D))} + \|u_k - u\|_{L^2(\Gamma;H_0^1(D))}$$

We are interested in the first term, which corresponds to the stochastic discretization error. We assume that the second term, which corresponds to the deterministic discretization error, is zero.

4.2.1. Tensor product interpolation using Gauss abscissas

We consider the global ME-PCM approximant $\tilde{u}$ for a uniform PDF $\rho(\boldsymbol{y})$, where the interpolation operator is based on full tensor product of Gauss–Legendre abscissas (see appendix, Section A.1), i.e., $\mathcal{I}_\Gamma = L_\Gamma^{\boldsymbol{p}}$, where $\boldsymbol{p} = (p_1,\ldots,p_N) \in \mathbb{Z}_+^N$ indicates polynomial order used in each random dimension. We begin by restating a few relevant results from [12], where the same stochastic elliptic problem is addressed in a one-element formulation. It is proven there that the solution satisfies the following regularity property:

Lemma 7. Regularity property [Babuska et al.] *Let $\Gamma_j^* = \prod_{i=1,i\neq j}^N \Gamma_j$, and let y_j^* be an arbitrary element of Γ_j^*. For each $y_j \in \Gamma_j$, there exists $\alpha_j > 0$ such that the solution $u(y_j,y_j^*,\boldsymbol{x})$ as a function of y_j, $u : \Gamma_j \to C^0(\Gamma_j^*,H_0^1(D))$ admits an analytic extension $u_a^j(z,y_j^*,\boldsymbol{x})$, $z \in \mathbb{C}$ of u, in the region of the complex plane:*

$$\Sigma(\Gamma_j;\alpha_j) \equiv \{z \in \mathbb{C}, dist(z,\Gamma_j) \leqslant \alpha_j\}. \tag{21}$$

Moreover, $\forall z \in \Sigma(\Gamma_z;\alpha_j)$, $\|u_a(z)\|_{C^0(\Gamma_j^*;H_0^1(D))} \leqslant \lambda$, where λ is independent of j.

We now state a main result of [12], which provides the error estimate of $L_\Gamma^{\boldsymbol{p}}u$.

Lemma 8. [Babuska et al.] *Let u be the exact solution of (3). Let $\boldsymbol{p} = (p_1,\ldots,p_N)$ be a vector in $\mathbb{Z}_+^N$ and $L_\Gamma^{\boldsymbol{p}}$ represent the Lagrangian tensor product interpolation operator using Gauss–Legendre abscissas on Γ, as defined in Section A.1. Then,*

$$\|L_\Gamma^{\boldsymbol{p}}u - u\|_{L^2(\Gamma;H_0^1(D))} \leqslant C\sum_{j=1}^N \exp\{-r_jp_j\} \tag{22}$$

where

$$r_j = \log\left[\frac{\alpha_j}{2}\left(1 + \sqrt{1 + \frac{1}{\alpha_j^2}}\right)\right], \tag{23}$$

and the α_j are the parameters related to the size of the analyticity domain in Lemma 7.

We note that in Lemmas 7 and 8, u_k is replaced by u because u_k possesses the same regularity as u. We next consider the error estimate of ME-PCM. Recall that $\tilde{u}(\boldsymbol{x},\boldsymbol{y}) = \sum_{i=1}^{N_e}(L_{B^i}^{\boldsymbol{p}}u_k)(\boldsymbol{x},\boldsymbol{y})\mathbb{I}_{\{\boldsymbol{y}\in B^i\}}$. For convenience, we instead consider the error between quantities $\sum_{i=1}^{N_e}(L_{B^i}^{\boldsymbol{p}}u)(\boldsymbol{x},\boldsymbol{y})\mathbb{I}_{\{\boldsymbol{y}\in B^i\}}$ and u. However, we will abuse notation a bit and still refer to the former quantity as $\tilde{u}$. We now state the result:

Theorem 9. *Let u be the exact solution of (3). Let $\mathcal{T}_{h,\Gamma} = \{B^i\}_{i=1}^{N_e}$ represent a uniform mesh on Γ where each element has side length h ($N_e = (\frac{1}{h})^N$). Let $\boldsymbol{p} = (p_1,\ldots,p_N)$ be a vector in $\mathbb{Z}_+^N$ and let $L_\Gamma^{\boldsymbol{p}}$ represent the Lagrangian tensor product interpolation operator using Gauss–Legendre abscissas on Γ. Let $L_{B^i}^{\boldsymbol{p}}$ represent $L_\Gamma^{\boldsymbol{p}}$ affinely mapped to B^i. Then,*

$$\|\tilde{u} - u\|_{L^2(\Gamma;H_0^1(D))} \leqslant C\sum_{j=1}^N \exp\{-r_j(h)p_j\} \tag{24}$$

where

$$r_j(h) = \log\left[\frac{\alpha_j}{2h}\left(1 + \sqrt{1 + \frac{h^2}{\alpha_j^2}}\right)\right]. \tag{25}$$

Proof. See Section A.3 in the appendix. $\square$

Remark 10. It is easy to see that for fixed **p**, if we rewrite $r_j(h)$ as

$$r_j(h) = \log\frac{1}{h} + \log\left[\frac{\alpha_j}{2}\left(1+\sqrt{1+\frac{h^2}{\alpha_j^2}}\right)\right],$$

we obtain a factor in the error estimate as

$$\exp(-p_j \log\frac{1}{h}) = h^{p_j},$$

which is the desired h-type convergence of ME-PCM for this problem. Alternatively for fixed h, $r_j(h)$ is a constant and we obtain p-type convergence through the term

$$\exp(-r_j(h)p_j).$$

Remark 11. It is important to note that Lemma 8 and Theorem 9 are not restricted to uniform distribution and can be easily generalized to other $\rho(\boldsymbol{y})$ by considering the norm equivalence [12]. Also error estimates of ME-PCM based on other interpolation operators [15], e.g., sparse grids, can be obtained following a similar procedure as in the proof of Theorem 9.

5. Numerical examples

5.1. Approximation of GENZ test functions

In this section we will numerically verify the h-convergence rate of ME-PCM stated in the previous section by approximating the integrals of the following functions defined on $[0,1]^2$ from the GENZ test suite [27]. Sparse grids based on one-dimensional Gauss and Clenshaw-Curtis quadrature rules are examined. We use the following functions:

$$\text{OSCILLATORY}: \quad f_1(x_1,x_2) = \cos(2\pi w_1 + c_1x_1 + c_2x_2)$$

$$\text{GAUSSIAN}: \quad f_4(x_1,x_2) = \exp(-c_1^2(x_1-w_1)^2 - c_2^2(x_2-w_2)^2)$$

$$\text{CONTINUOUS}: \quad f_5(x_1,x_2) = \exp(-c_1|x_1-w_1|^2 - c_2|x_2-w_2|^2)$$

$$\text{DISCONTINUOUS}: \quad f_6(x_1,x_2) = \begin{cases} 0, & \text{if } x_1 > w_1 \text{ or } x_2 > w_2, \\ \exp(c_1x_1 + c_2x_2), & \text{otherwise.} \end{cases}$$

where w_1, w_2, c_1, c_2 are constants.

Fig. 1 shows the approximation error of the OSCILLATORY and GAUSSIAN integrals by sparse grid cubature on uniform meshes. The number of elements per side N_{es} corresponds to $\frac{1}{h}$ where h is the size of each element. The sparseness parameter is $s = 3$ in both cases, leading to a degree of exactness of $m = 3$ for both types of grids. Both of these functions f_1, f_4 lie in $\mathcal{W}^{4,\infty}([0,1]^2)$ so the conditions of Theorem 4 hold. From the figure, we see that the convergence rate is indeed $O(h^4) = O(N_{es}^{-4})$ as predicted, since $m+1 = 4$.

We also consider the CONTINUOUS and DISCONTINUOUS functions (f_5, f_6), which do not satisfy the regularity requirements needed in the assumptions of Theorem 4. Fig. 2 shows the approximation error for these functions using the same

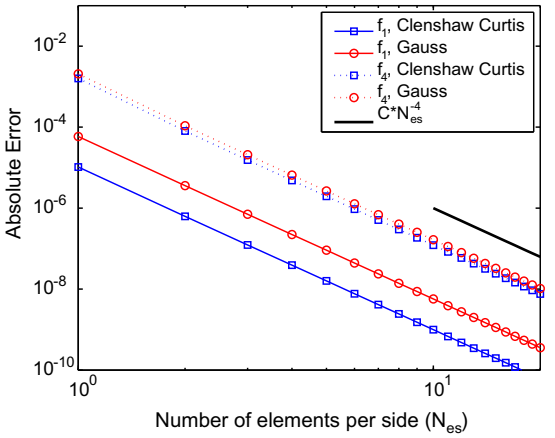

Fig. 1. h-Convergence of Gauss and Clenshaw-Curtis sparse grid integration of smooth functions f_1 and f_4 from the GENZ test suite. Solid lines indicate f_1 errors while dotted lines indicate f_4 errors. Square markers: Clenshaw-Curtis points, circle markers: Gauss points.

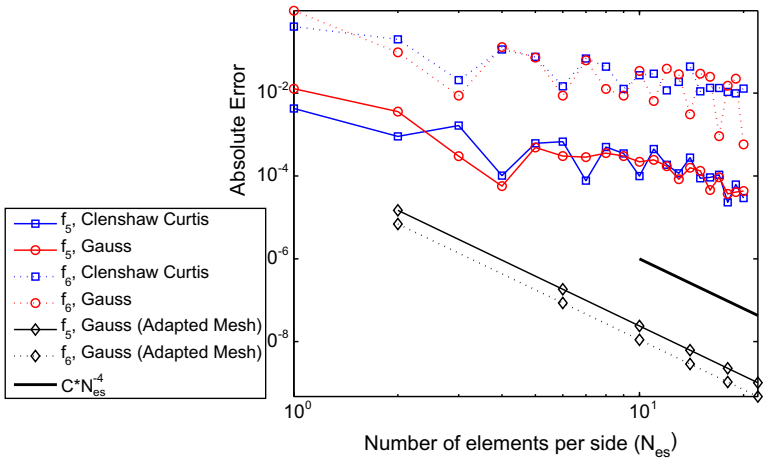

Fig. 2. *h*-Convergence of sparse grid integration error for nonsmooth functions f_5 and f_6 from the GENZ test suite. Solid lines indicate f_1 errors while dotted lines indicate f_4 errors. Marker key – square: Clenshaw-Curtis points on uniform mesh, circle: Gauss points on uniform mesh, diamond: Gauss points on adapted mesh.

sparse grids as in the previous example. Without the proper regularity, it can be seen that the convergence rate is reduced to approximately *order one*. Tests were also performed to verify that *adapting* the mesh to the discontinuities of the function recovers the optimal convergence rate. Specifically, the mesh is chosen such that the discontinuities fall on the borders of elements. From the figure it can be seen that under these conditions, the convergence rate indeed returns to $O(N_{es}^{-4})$.

To investigate the constant in the error bound in (15), the integral of functions f_1 and f_4 are approximated using a Clenshaw-Curtis sparse grid with varying sparseness parameter s. The errors are shown in tables in Fig. 3. In the last column, the absolute error ϵ is normalized by the term

$$R = r^{-m}(\log r)^{(N-1)(m+1)}h^{m+1}|\cdot|_{m+1,\infty,[0,1]^2}, \tag{26}$$

where r is the total number of points in the quadrature rule. This gives an estimate of the constants in the error bound of (15).

It can be seen that by normalizing the error by the factor R, we obtain a constant which depends on the function being integrated as well as the sparseness parameter.

5.2. ODE examples

Next we examine errors in time-dependent solutions using, first, simple ordinary differential equations (ODEs) and, second, a nonlinear system.

5.2.1. Simple ODEs

In this section we investigate the *h*-convergence rate of ME-PCM for ODEs with uncertain parameters. The collocation grid in each element is kept fixed as the mesh is uniformly refined. Only 'smooth' problems are considered, i.e., problems where the solution exhibits regularity in the parameter space. The error from the time integration solver (fourth-order Runge–Kutta) is negligible, and thus dominated by the more interesting stochastic semidiscrete error term. We are interested in verifying the convergence rates for ODEs in both the mean and variance error, while using tensor product grids instead of the sparse grids investigated in the previous examples. The following equations are investigated:

$$\text{Case I}: \frac{dy}{dt} = -(\xi_1 + M)y$$
$$\text{Case II}: \frac{dy}{dt} = -(\xi_1 + M)y^2$$
$$\text{Case III}: \frac{dy}{dt} = -(\xi_1 + M)y^3$$
$$\text{Case IV}: \frac{dy}{dt} = -(\xi_1 + \xi_2 + M)y^2$$

where ξ_1, ξ_2 are i.i.d. random variables distributed uniformly on $[-1, 1]$. M is chosen to avoid singularities in the solution and $y(0) = 1$. Exact solutions can be found for all four cases. The error in mean and variance between the numerical solution using ME-PCM and the exact solution is taken at time $T = 5$.

s, m, r	N_{es}	$\lvert\epsilon\rvert$	$\frac{\lvert\epsilon\rvert}{R}$	s, m, r	N_{es}	$\lvert\epsilon\rvert$	$\frac{\lvert\epsilon\rvert}{R}$
3,3,5	2	6.3e-7	0.000405	3,3,5	1	8.0e-5	0.000319
	4	3.9e-8	0.000402		2	8.1e-5	0.000319
	10	1.0e-9	0.000402		4	4.8e-6	0.000305
	20	6.2e-11	0.000402		8	3.0e-6	0.000302
	100	1.0e-13	0.000404		10	1.2e-7	0.000301
4,5,13	1	2.2e-7	0.000774		50	1.9e-10	0.000300
	2	3.6e-9	0.000760	4,5,13	3	1.8e-8	0.000900
	4	5.2e-11	0.000756		10	1.3e-11	0.000888
	5	1.4e-11	0.000755		15	1.1e-12	0.000887
	8	8.1e-13	0.000755		20	2.1e-13	0.000892
	15	1.9e-14	0.000752	5,7,29	2	3.7.0e-10	0.000283
5,7,29	1	3.1e-10	0.001109		3	2.6e-11	0.000271
	2	1.2e-12	0.001085		5	4.3e-13	0.000266
	3	4.7e-15	0.001080		8	9.9e-15	0.000265

Fig. 3. Sparse integration errors ϵ and numerical approximation of constants for the functions f_1 (left) and f_4 (right) using varying sparseness parameter s. Recall that s is the sparseness parameter of the grid in each element, m is the degree of exactness, and r is the total number of points used in each element. Also, N_{es} is the number of elements per direction.

In this example a tensor product Gauss–Legendre collocation grid is used over each element. An asymptotic index of algebraic convergence, κ, is numerically calculated as h decreases. This is shown in Tables 1 and 2 for the mean and variance, respectively. The results are close to the expected value of κ, which is $m+1$ where m is now the degree of exactness for the Gauss tensor product grids. Furthermore, the same result holds for the variance as well as the mean.

5.2.2. Kraichnan–Orszag problem

The deterministic Kraichnan–Orszag (K–O) three-mode problem is as follows [28]:

$$\begin{aligned}\frac{dy_1}{dt} &= y_1 y_3\\ \frac{dy_2}{dt} &= -y_2 y_3\\ \frac{dy_3}{dt} &= -y_1^2 + y_2^2\end{aligned} \tag{27}$$

Table 1
Numerically calculated index of algebraic convergence κ of the *mean* solution at time $T = 5$ for ODE examples I–IV

degree of exactness m	Case I	Case II	Case III	Case IV
3	−3.95	−3.99	−3.99	−3.93
5	−5.94	−5.98	−5.98	−5.91
7	−7.96	−7.91	−7.91	−7.92

ME-PCM is used with a tensor product Gauss–Legendre grid in each element. These results show good agreement with the expected values of $\kappa \approx m+1$.

Table 2
Numerically calculated index of algebraic convergence κ of the *variance* of the solution at time $T = 5$ for ODE examples I–IV

degree of exactness m	Case I	Case II	Case III	Case IV
3	−3.95	−3.99	−3.99	−3.96
5	−5.94	−5.97	−5.97	−5.90
7	−7.92	−7.97	−7.89	−7.91

ME-PCM is used with a tensor product Gauss–Legendre grid in each element. These results show good agreement with the expected values of $\kappa \approx m+1$.

subject to random initial conditions:

$$y_1(0) = y_1(0;\omega), \quad y_2(0) = y_2(0;\omega), \quad y_3(0) = y_3(0;\omega)$$

The solution exhibits *low regularity* with respect to the parameter $y_2(0)$. Fig. 4 shows phase plots of solutions to the deterministic problem with varying initial conditions for y_2. If $y_2(0)$ is negative, the solution is periodic and travels in the region where y_2 is negative. If it is positive, the solution is also periodic but travels in the region where y_2 is positive. If it is zero, the solution is a constant. It is clear from this plot that the frequency of the solution is also dependent on this initial condition. The temporal discretization here and throughout this section is performed by fourth-order Runge–Kutta integration.

To further explore this, suppose that the initial condition $y_2(0;\omega)$ in the stochastic K–O problem is ξ, a *uniform* random variable on $[-0.1, 0.1]$ and that $y_1(0) = 1$ and $y_3(0) = 0$. The solution $\boldsymbol{y} = (y_1, y_2, y_3)$ is only C^0 continuous with respect to the random parameter ξ. Fig. 5 shows the solution of the deterministic K–O problem as a function of the parameter $y_2(0)$ ranging from -0.1 to 0.1 at varying times. We can see that the solution is very oscillatory; thus, it requires a large amount of computational time to be solved by existing methods such as ME-gPC, PCM, and gPC. In the following sections, we will investigate the performance of both regular and *h*-adaptive ME-PCM on this problem for short- and long-time integration. Computational costs of ME-PCM will be compared to those of other existing methods mentioned above. In addition, a sparse grid ME-PCM example will be demonstrated for a case where $N = 3$.

In the following, ξ_1, ξ_2 are i.i.d. random variables distributed *uniformly* on $[-1, 1]$. First, we consider a two-dimensional $(N = 2)$ case of problem in Eq. (27), where $y_1(0) = \xi_2$, $y_2(0) = 0.1\xi_1$, and $y_3(0) = 0$. To illustrate *p*-convergence of ME-PCM, we hold N_e constant using a uniform grid of 4 elements on $\Gamma = [-1, 1] \times [-1, 1]$. A tensor product of Gauss–Legendre quadrature points is used and the number of points is increased between each run. Fig. 6 shows the errors in mean and variance of y_1. The exact solution is taken to be the ME-PCM solution with $N_e = 100$ elements and $r = 100$ points in each element.

Next, *h*-convergence is demonstrated by holding r constant with just one collocation point per element and refining the mesh uniformly. Here the one-dimensional input used: $y_2(0) = \xi_1, y_1(0) = 1.0, y_3(0) = 0.0$. Fig. 7 shows the errors in mean and variance of y_1 at time $t = 5$ as a function of the number of elements used. Note that here at time $t = 5$ the solution is still smooth, and we expect the convergence rate defined in the previous section. Hence, we expect the error to decrease by $O(N_e^{-1})$ as the mesh is refined.

For convergence in long-term integration we demonstrate the *h*-adaptive version of ME-PCM using the one-dimensional random input as in the previous example. In Fig. 8 the ME-PCM solution is shown for varying tolerance level θ (see Section A.4 in the appendix). A Gauss–Legendre grid of 3 collocation points in each element is used. As θ decreases, the elements split more frequently; thus the solution does indeed converge to the reference solution which is obtained using a quasi random Sobol (MC-SOBOL) sequence with 10^6 iterations. Also plotted on the same graph is an example PCM solution where 30 Gauss–Legendre points are used. The PCM solution and the ME-PCM solution are both shown only until they diverge from the correct solution.

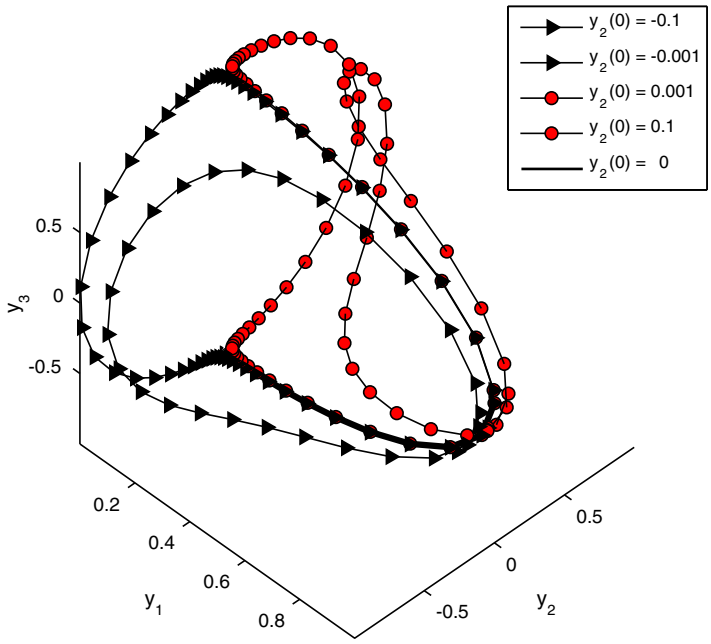

Fig. 4. y_1–y_2–y_3 Phase plot of solutions to the Kraichnan–Orszag problem (27) with initial conditions $y_1(0) = 1, y_3(0) = 0$ and varying $y_2(0)$. Triangle markers indicate solutions when $y_2(0) > 0$. Circle markers indicate solutions when $y_2(0) < 0$. The wide black line (situated on the lower half of the $y_2 = 0$ plane separating the two sets of trajectories for $y_2(0) < 0$ and $y_2(0) > 0$) indicates the solution when $y_2(0) = 0$.

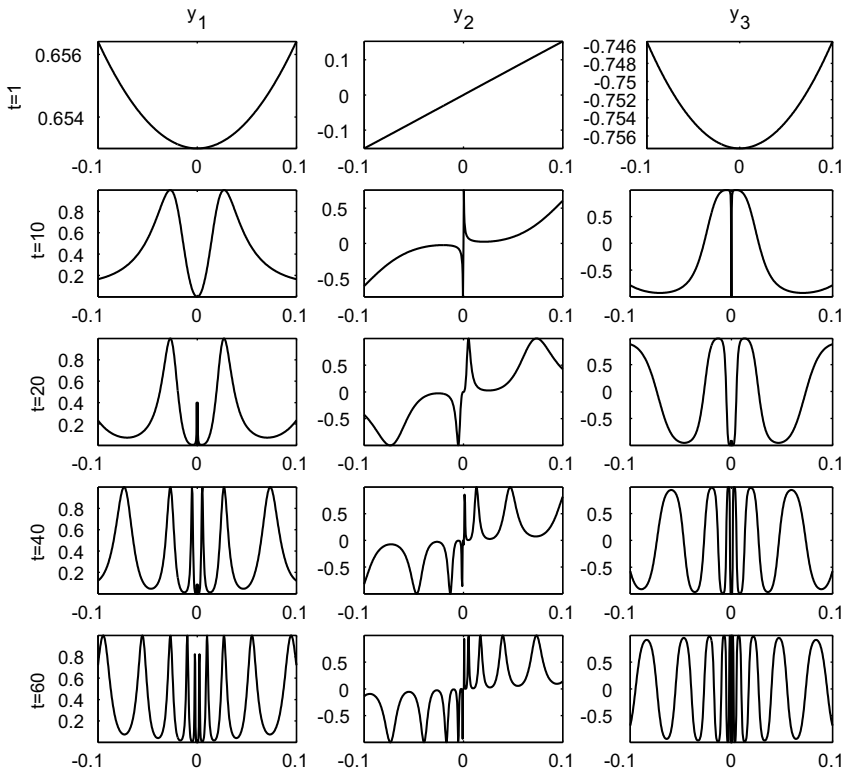

Fig. 5. Plots of the solution to (27) y_1, y_2, y_3 as a function of $y_2(0)$ at various times. $(y_1(0) = 1, y_3(0) = 0)$.

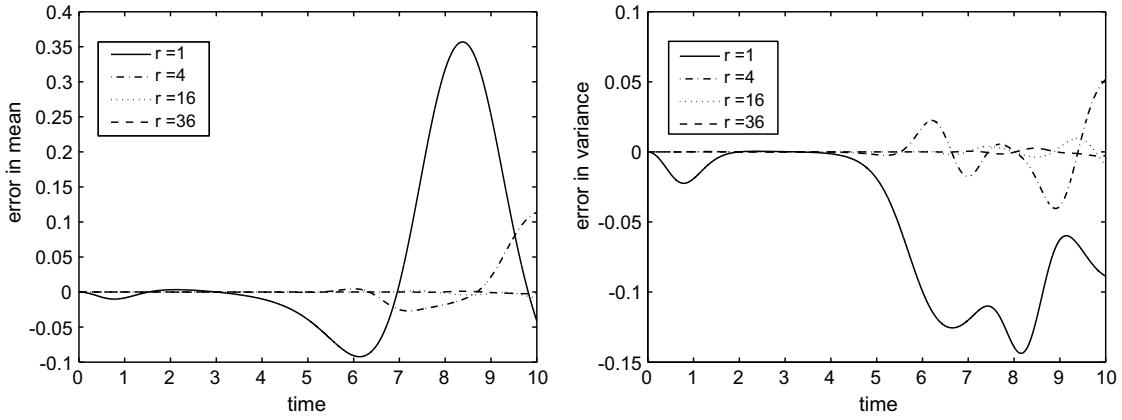

Fig. 6. Error in mean (left) and variance (right) of y_1 for 2D K–O problem with varying r.

Computational cost: Here the computational costs of ME-PCM for the K–O problem are compared to those of existing methods. Since the solution to this problem has low regularity in the parametric space it represents a 'worst case scenario' for all of these methods. First, we consider the two-dimensional K–O problem from the previous section. The error in variance of y_1 is considered at final time $t = 10$. In this section we define the error between two numerically integrated functions $g(t_j)$ and $f(t_j), j = 1, \ldots, n_t$ as:

$$\epsilon_{L^2} = \frac{\frac{1}{n_t}\sqrt{\sum_{j=1}^{n_t}(g(t_j) - f(t_j))^2}}{\frac{1}{n_t}\sqrt{\sum_{j=1}^{n_t} f(t_j)^2}}, \tag{28}$$

where f is considered to be the reference solution.

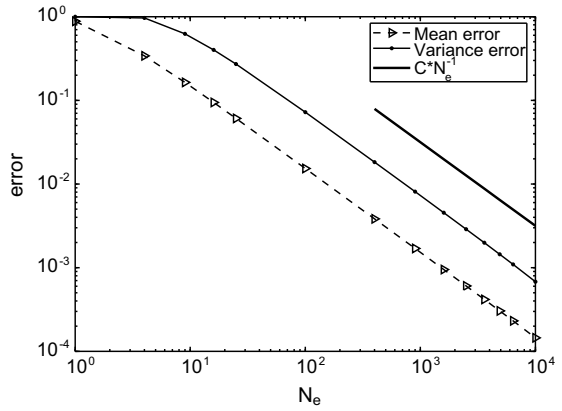

Fig. 7. *h*-Convergence of ME-PCM for the 1D K–O problem.

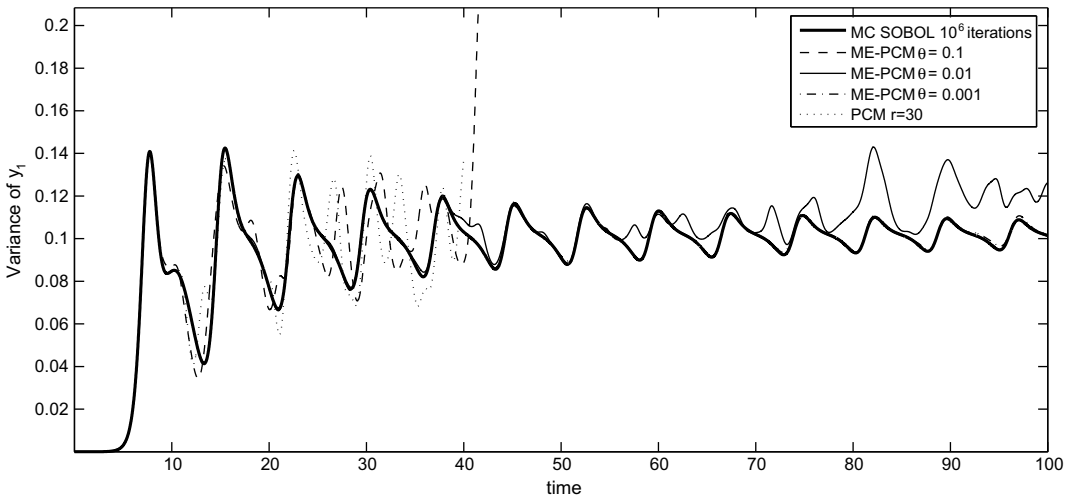

Fig. 8. *h*-Adaptive ME-PCM for the 1D K–O problem. We note that the single-element PCM diverges after $t \approx 20$.

To compare computational costs, an error tolerance is prescribed and the cost for achieving that error level is measured. The methods examined are: ME-PCM with Clenshaw-Curtis sparse grids, PCM with Clenshaw-Curtis sparse grids, and MC-SOBOL. For PCM and ME-PCM, the error levels are achieved by increasing the number of points in each element (*p-refinement*). For ME-PCM we are also able to refine the mesh simultaneously (*h-refinement*). For MC-SOBOL, the error level is achieved by increasing the length of the Sobol sequence. Table 3 shows the results of these comparisons. From the results we can see that the ME-PCM is much faster than either PCM or MC-SOBOL for all error tolerances and that for some error levels PCM is actually more expensive than MC-SOBOL.

Next, we consider the costs of the *h*-adaptive formulations of ME-PCM and ME-gPC for the same problem. We note that for the K–O problem, the solution's spectrum is continuously growing so that elements are continuously splitting and this leads to large computational costs. This property makes it suitable for comparing two such *h*-adaptive methods as it will accentuate the relative difference between the methods and their weaknesses. For both methods, the error level is achieved by decreasing θ, the adaptivity tolerance. Legendre chaos is used for the basis in the ME-gPC method, and a tensor product

Table 3
Comparison of number of samples required for the 2D K–O problem $(y_1(0) = 1, y_2(0) = 0.1\xi_1, y_3(0) = \xi_2)$ for time $t \in [0, 10]$

Error level ϵ_{L^2}	ME-PCM	PCM	MC-SOBOL
10^{-2}	36 ($N_e = 16, s = 2$)	321 ($s = 8$)	100
10^{-3}	320 ($N_e = 64, s = 3$)	3329 ($s = 11$)	950
10^{-4}	3328 ($N_e = 256, s = 4$)	7169 ($s = 12$)	9500

For ME-PCM and PCM the sparseness parameter s is provided, and for ME-PCM the number of elements N_e is also provided.

Gauss–Legendre grid is used in each element for the ME-PCM. The highest polynomial order is 2 for ME-GPC and 3 collocation points per element are used for ME-PCM. Table 4 shows the results of this comparison. We note that even with extra projection steps between physical and modal parametric space, ME-PCM outperforms ME-gPC. It is important to note that the same adaptivity tolerance is required for achieving the prescribed error tolerance for both methods; thus, the difference in computational costs is due to the actual cost of each method and not a discrepancy in the element splitting frequency.

In Fig. 9 we demonstrate the h-convergence of sparse grid ME-PCM for a three-dimensional K–O problem $(y_1(0) = \xi_1, y_2(0) = \xi_2, y_3(0) = \xi_3)$. A sparse Gauss–Legendre grid with 195 points is used in every element and the mesh is uniformly refined as in previous examples. We also plot the solution obtained from the h-adaptive version of ME-PCM, using $\theta = 0.0001$ and a Gauss–Legendre tensor product grid with 7^3 points in each element.

5.3. Kovasznay flow

We consider next the problem of steady, laminar flow behind a two-dimensional grid, solved by Kovasznay in 1948 [29]. The exact solution to the Navier-Stokes equations is given by:

$$u = 1 - e^{\lambda x}\cos(2\pi y), \quad v = \frac{\lambda}{2\pi}e^{\lambda x}\sin(2\pi y),$$

where u, v are velocities in the x and y direction, respectively, and

$$\lambda = \frac{1}{2\nu} - \left(\frac{1}{4\nu^2} + 4\pi^2\right)^{\frac{1}{2}},$$

Here we model the kinematic viscosity, ν, as a random parameter as:

$$\nu = \nu_0(1 + \delta\xi), \quad |\delta| < 1. \tag{29}$$

The random variable ξ has Beta distributions $B(1,1)$ and $B(0,0)$ with support $[-1,1]$ for the two cases we address, and ν_0 represents the mean viscosity. Each deterministic problem is performed by a spectral/hp element method using a 32-element mesh. The accuracy of the spatial solver is of order 10^{-10} in the L_∞ norm.

In the first case $\xi \sim B(1,1)$, $\delta = 0.95$, $\nu_0 = 0.05$ and the collocation points are chosen to be Gauss points from the gPC basis constructed to be orthogonal to the PDF on each element. We use $r = 2$ points in each element with a degree of exactness $m = 2r - 1 = 3$. Since the problem is smooth we expect a convergence rate of $O(N_e^{-(m+1)}) = O(N_e^{-4})$ (note that $N_e = N_{es}$ here). In Fig. 10 the L_∞ errors in mean and variance are plotted as a function of N_e. The numerical results show agreement with the expected convergence rate of the method.

Table 4
Comparison of computational costs (s) for 2D K–O problem $(y_1(0) = 1, y_2(0) = 0.1\xi_1, y_3(0) = \xi_2)$ for time $t \in [0, 10]$

Error level ϵ_{L^2}	h-Adaptive ME-PCM	h-Adaptive ME-gPC
10^{-2}	0.5	11.95
10^{-3}	3.43	29.31
10^{-4}	38.8	337.7

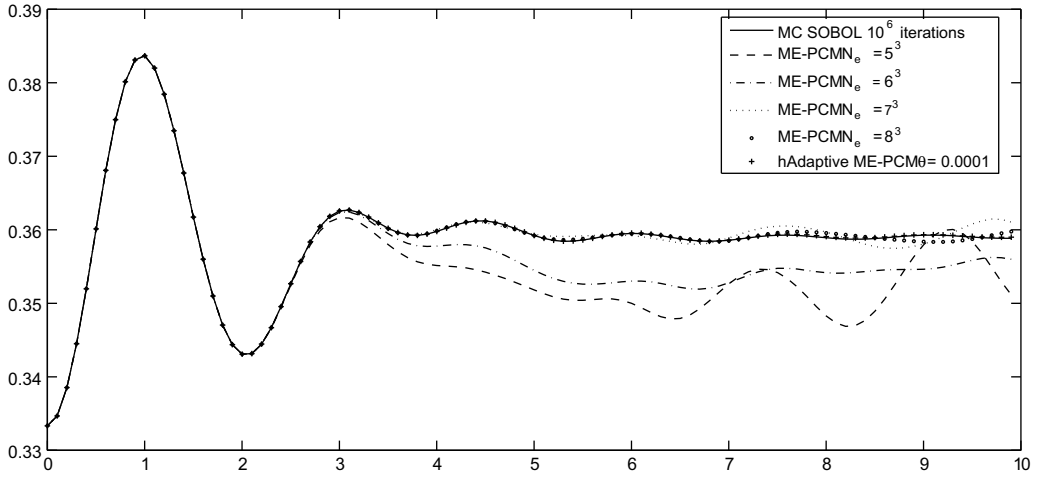

Fig. 9. h-Convergence of sparse grid ME-PCM for the 3D K–O problem $(y_1(0) = \xi_1, y_2(0) = \xi_2, y_3(0) = \xi_3)$.

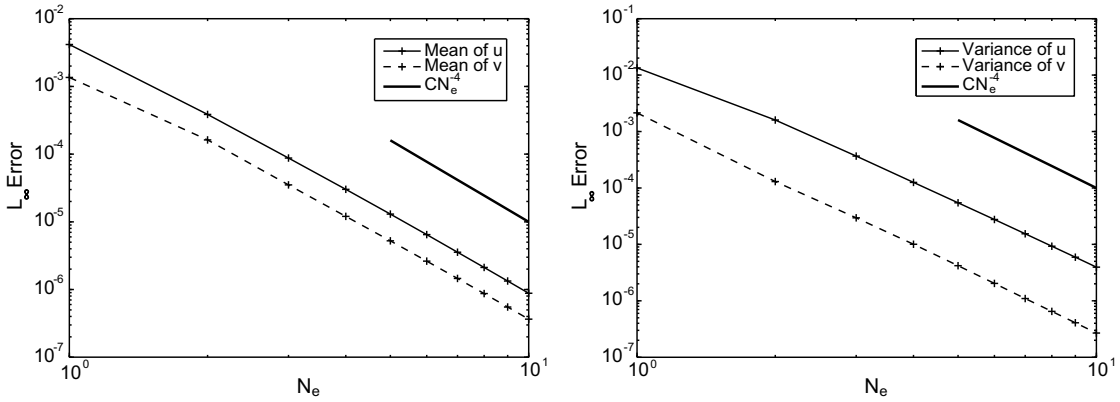

Fig. 10. Kovasznay flow problem: h-convergence rate for $\xi \sim B(-1,1)$. Left: mean; right: variance.

We also consider the case where $\xi \sim B(0,0)$ which is the same as the uniform distribution on $[-1,1]$. In this case $\delta = 0.8$ and $v_0 = 0.05$. Here, $r = 1$ Gauss–Legendre points are used in each element; thus the degree of exactness is $m = 1$. The expected convergence rate is then $O(N_e^{-2})$. Fig. 11 shows the results for this case. For comparison we include ME-gPC results with the same h-convergence rate (highest order of polynomial is zero). It is seen that ME-gPC provides a better accuracy for the mean value due to the Galerkin projection.

5.4. Navier-Stokes: noisy flow past a 2D stationary circular cylinder

We also study noisy flow past a 2D stationary cylinder to demonstrate the long-term behavior of the ME-PCM method. We consider the following inflow boundary conditions:

$$u = 1 + \delta\xi, \quad v = 0,$$

where δ is a constant equal to 0.1 and ξ is a uniform random variable on $[-\sqrt{3}, \sqrt{3}]$. The Reynolds number based on the mean velocity is $Re = 100$.

This problem was studied in [30] using the ME-gPC method. Due to the sensitivity of the vortex shedding frequency to the inflow noise, high-order polynomial chaos is not efficient for this problem. However, the convergence can be improved by using multi-element formulations and h-type refinement. Here, we re-visit this problem using the ME-PCM method.

The error evolution of variance of the lift coefficient is shown in Fig. 12. The errors are normalized based on a reference solution given by ME-PCM with $N_e = 20$, $r = 9$. It is observed that the errors given by single-element PCM increase fast to $O(1)$. Similar results are shown in [30] for the high-order single-element gPC. This is due to the expanding frequency spectrum in the parametric space with respect to ξ. Thus, the polynomial order or number of collocation points must be increased rapidly to maintain low errors. ME-PCM effectively slows down the error increase by h-refinement in agreement with the analysis of [30].

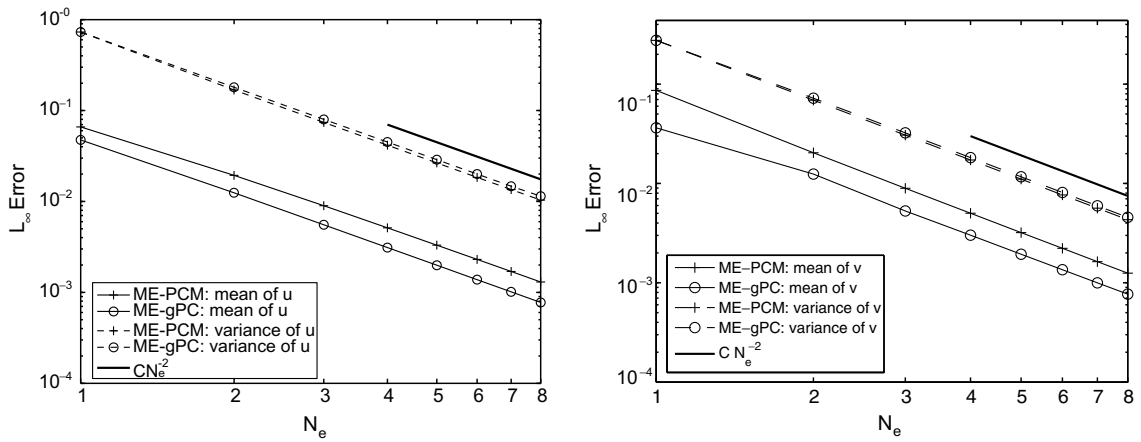

Fig. 11. Kovasznay flow problem: h-convergence rate for $\xi \sim B(0,0)$. Left: x-velocity; right: y-velocity.

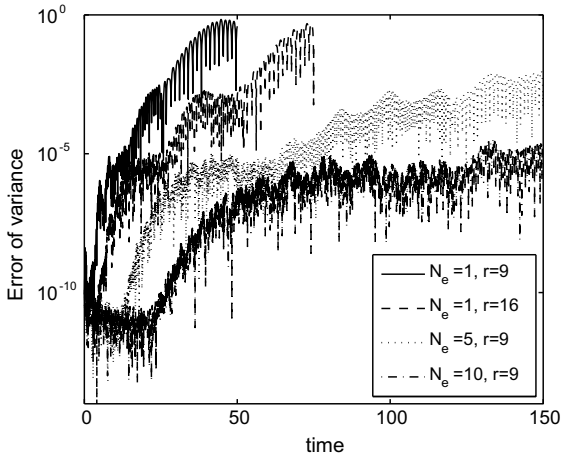

Fig. 12. Error evolution of variance of the lift coefficient given by ME-PCM for the noisy flow past 2D stationary circular cylinder. More than 25 vortex shedding cycles are simulated.

5.5. Stochastic elliptic problem

We consider the two-dimensional ($D = [0,1]^2$) elliptic problem with random coefficients from Section 2.1 using the ME-PCM method. To avoid introducing large errors from physical discretization, we consider a smooth problem in the physical domain. For simplicity, we use a non-zero force term with homogeneous boundary conditions

$$f(x) = \sin(x_1)\cos(x_2) \quad \text{and } \mathbb{E}[a](x) = 1. \tag{30}$$

Assume that the random field $a(x,\omega)$ satisfies the Gaussian correlation function: $K(x_1,x_2) = \delta^2\, e^{\frac{-|x_1-x_2|^2}{A}}$ with A being the correlation length and δ the standard deviation. Due to the analyticity of the Gaussian kernel, the eigenvalues decay exponentially [8]. The decay rate is determined by the value of the correlation length, where a larger A corresponds to a faster decay rate.

Since the Gaussian kernel is analytic, high-order element methods for spatial discretization converge quickly, resulting in highly accurate numerical solutions for the eigenvalue problem in the Karhunen-Loève expansion. Spectral/*hp* discretization with 64 quadrilateral elements is used in the physical space. In each element, a 12th order Jacobi polynomial basis is used to construct the approximation space. The accuracy is close to the machine accuracy for numerical solutions of both deterministic elliptic PDEs and the eigenvalue problem. Therefore, we assume from now on that no substantial errors come from the physical discretization.

We perform the ME-PCM for this problem using Smolyak Gauss sparse grids in each element of a uniform mesh. We first verify the relation between the *h*-convergence rate and the degree of exactness of the quadrature rule. Consider a two-dimensional random input Y_1, Y_2 with uniform and Beta distributions, taking values in $\Gamma = [-1,1]^2$. We note here that for the case where Y_1 and Y_2 have uniform distribution, the local conditional probability density function is still uniform in each

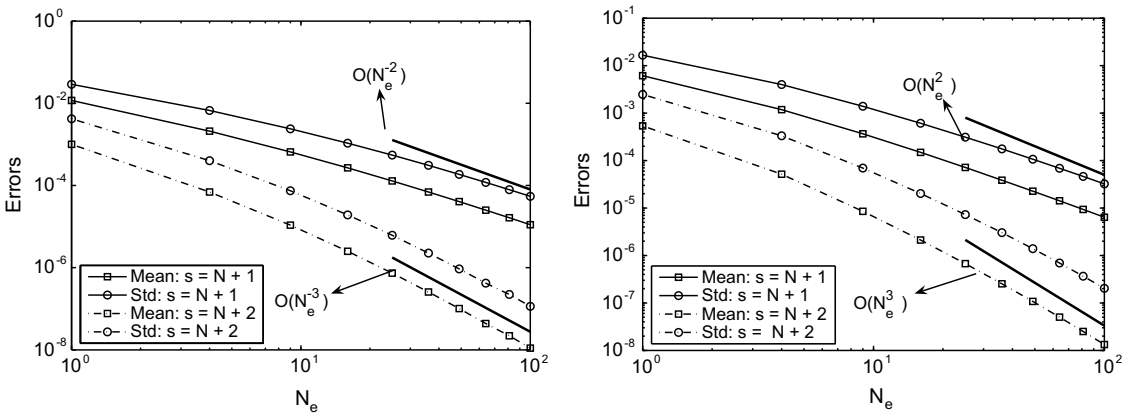

Fig. 13. *h*-Convergence rate of the ME-PCM for the stochastic diffusion problem with $N = 2$ and the degree of exactness of quadrature rules. Left: Uniform distribution, $\delta = 0.66$; right: Beta(1,1) distribution, $\delta = 0.72$.

direction; however, in the case where Y_1 and Y_2 have Beta distribution the conditional density function can differ from dimension to dimension. Due to the symmetry of the tensor-product rule in the Smolyak algorithm, we still expect the h-convergence rate given in theorem for the Beta distribution case. We use $\delta = 0.72$ for Beta(1,1) distribution and $\delta = 0.66$ for uniform distribution so that the variance of the input for both cases is equal. The h-convergence behavior is shown in Fig. 13. It is seen that the h-convergence rate asymptotically approaches $O\left(N_e^{\frac{-(m+1)}{N}}\right) = O\left(N_{es}^{-(m+1)}\right)$ for both uniform and Beta distributions. Note that $m(3,2) = 3$ and $m(4,2) = 5$ for these non-nested sparse grids.

We next study some higher dimensional cases:

(i) $N = 4$, $A = 2.8367936716$,
(ii) $N = 10$, $A = 0.4898834872$,
(iii) $N = 25$, $A = 0.1121059863$,
(iv) $N = 50$, $A = 0.04890758154$.

For all these cases the smallest eigenvalue is less than 0.314% of the largest eigenvalue, see Fig. 14. We compute the error by comparing the solution with a reference solution computed using a highly refined grid in Γ. The error is computed in the norm $\|\cdot\|_{L^2(\Gamma;H_0^1(D))}$, where $\Gamma = [-1,1]^N$ is the parametric domain.

According to the regularity study in [9], the importance of each random dimension can be roughly estimated by the value $\sqrt{\lambda_i}\|\phi_i(x)\|_{L_\infty}$, where λ_i and ϕ_i are the eigenpairs from the Karhunen-Loève expansion satisfying (5). In [18] it is shown that refining the random dimension with the largest value of $\sqrt{\lambda_i}\|\phi_i(x)\|_{L_\infty}$ is an efficient adaptivity method. For collocation-type methods, one straightforward approach to approximate the local errors is to compare the results given by sparse grids at two

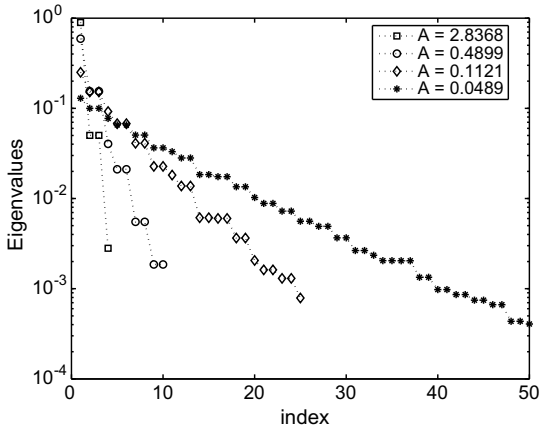

Fig. 14. Eigenvalues given by different correlation lengths.

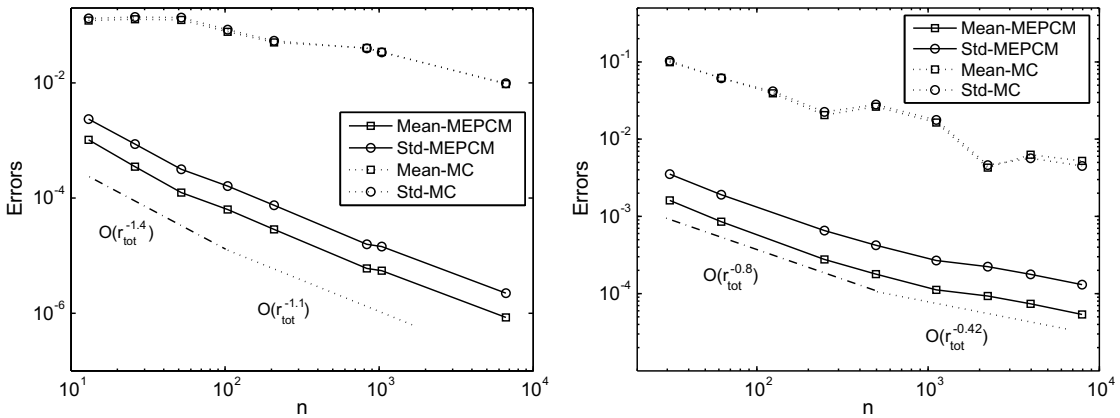

Fig. 15. h-Convergence of the ME-PCM method. Error in mean and standard deviation vs. number of sampling points are plotted for both ME-PCM and Monte Carlo methods. The local observed convergence rates of ME-PCM are denoted with dashed lines adjacent to the graphs. Left: $N = 4$, $s = N + 1$; right: $N = 10$, $s = N + 1$.

different levels. However, due to the big jump between numbers of points at two consecutive sparseness levels the cost for error estimation could be much larger than solving the original problem. For example, if $N = 10$ and $s = N + 1$, the cost for error estimation is about 15 times as large as the cost for solving the equation. Thus, here we investigate the h-convergence of ME-PCM method by building *a priori* adaptive meshes in the parametric space. More specifically, we use a pre-constructed mesh where only the dimensions with the largest values of $\sqrt{\lambda_i}\|\phi_i(x)\|_{L_\infty}$ are decomposed.

In Fig. 15 we plot the normalized errors of mean and standard deviation versus the total number $r_{\text{tot}} = N_e r$ of collocation points in Γ for $N = 4, 10$. For both cases the three most important random dimensions have been decomposed in every parameter space mesh. We refine the mesh by further decomposing elements in only the three most important dimensions. According to Theorem 4 we expect the theoretical convergence rate $O\left(r_{\text{tot}}^{-(m+1)/N}\right) = O\left(N_e^{-(m+1)/N}\right)$ as the mesh is refined uniformly in all directions. For the cases shown in Fig. 15, these expected theoretical convergence rates are $O(r_{\text{tot}}^{-1})$ for $N = 4$ and $O(r_{\text{tot}}^{-0.4})$. However, since we refine in only the most important random dimensions, at low discretization levels we observe that the h-convergence rate is actually faster than the expected rate. As the mesh is refined further in these three dimensions, we observe that the h-convergence rate asymptotically approaches the expected theoretical values. This is due to the fact that the values of $\sqrt{\lambda_i}\|\phi_i(x)\|_{L_\infty}$ for each dimension become more comparable after decomposition of the parametric space. We note that if the number of random dimensions is of the order $O(10)$, we can easily obtain an h-convergence rate better than Monte Carlo methods using sparse grids of relatively low level. For example, the expected asymptotic h-convergence rate is $O(r_{\text{tot}}^{-0.6})$ and $O(r_{\text{tot}}^{-0.8})$, for $s = N + 2$ and $s = N + 3$, respectively, with $N = 10$.

Note that if the approximated random function is sufficiently regular the h-convergence rate is determined by m and N. The ME-PCM method will become less efficient for a fixed m as N increases. In Tables 5 and 6, we show the h-convergence for $N = 25, 50$, where only the *six* most important dimensions are refined. We note that for moderate error levels, the efficiency of ME-PCM is far better than that of Monte Carlo since using approximately 100 samples results in a very small error two orders of magnitude less than the corresponding Monte Carlo error. The overall observed h-convergence rate is $O(r_{\text{tot}}^{-0.26})$ and $O(r_{\text{tot}}^{-0.13})$ for $N = 25$ and $N = 50$, respectively, with $s = N + 1$, which represents the sparsest nontrivial level of sampling. It is seen that adaptivity improves the h-convergence rate, since if no adapted meshes are used the expected h-convergence rate would be $O(r_{\text{tot}}^{-0.16})$ and $O(r_{\text{tot}}^{-0.08})$ for $N = 25$ and $N = 50$, respectively, for $s = N + 1$. If coupled with *a posteriori* error estimation [18], the adaptive meshes can be improved further by refining the elements with largest local errors.

In summary, the relation between the degree of exactness and h-refinement is verified for the elliptic problem with random coefficients. Using *a priori* adaptive meshes, the ME-PCM method based on sparse grids of relative low level can be an

Table 5
Errors in mean and standard deviation for elliptic problem with $N = 25$; comparison of ME-PCM and MC errors for the same amount of work done (number of points sampled)

# of Sample points	ME-PCM mean	ME-PCM std	MC mean	MC std
76	1.72e−4	3.41e−4	3.56e−2	3.62e−2
152	1.29e−4	2.57e−4	2.44e−2	2.44e−2
608	7.55e−5	1.52e−4	1.05e−2	1.08e−2
1216	6.94e−5	1.40e−4	9.90e−3	1.03e−2
4104	5.86e−5	1.19e−4	4.14e−3	4.27e−3
9728	4.67e−5	9.50e−5	1.51e−3	1.56e−3
38912	2.82e−5	5.76e−5	1.11e−3	1.13e−3
58368	2.76e−5	5.64e−5	6.27e−4	6.42e−4

Here, ME-PCM with a sparse grid with parameter $s = N + 1$ is used in every element. Mesh refinement is performed in only the six most important dimensions.

Table 6
Errors in mean and standard deviation for elliptic problem with $N = 50$; comparison of ME-PCM and MC errors for the same amount of work done (number of points sampled)

# of Sample points	ME-PCM mean	ME-PCM std	MC mean	MC std
151	4.31e−5	7.92e−5	9.23e−3	9.37e−3
302	3.70e−5	6.92e−5	8.49e−3	8.63e−3
604	3.27e−5	6.16e−5	5.78e−3	5.75e−3
2416	2.65e−5	5.08e−5	3.12e−3	3.11e−3
4832	2.55e−5	4.93e−5	2.48e−3	2.47e−3
19328	2.40e−5	4.66e−5	1.25e−3	1.23e−3
77312	1.83e−5	3.55e−5	5.88e−4	5.86e−4
173952	1.53e−5	3.01e−5	5.11e−4	5.12e−4
309248	1.50e−5	2.96e−5	3.72e−4	3.78e−4

Here, ME-PCM with a sparse grid with parameter $s = N + 1$ is used in every element. Mesh refinement is performed in only the six most important dimensions.

efficient numerical approach for a moderate number $O(10)$ of random dimensions. A surprising result for high dimensions is that although the convergence rate of ME-PCM degrades the method produces very good results even for a very small number of samples, of the order of 100.

6. Summary

In this work we introduced an *h*-adaptive multi-element formulation of the probabilistic collocation method. Two choices for collocation point sets were addressed: tensor product and Smolyak sparse grids. It was proven in Section 4 that the *h*-convergence rate of ME-PCM moment errors is dependent on the choice of cubature rule only through its degree of exactness, in the case of uniform inputs. We also presented there an L^2 error bound for the ME-PCM solution to the stochastic diffusion problem in (4.2).

The *h*-convergence rate result was verified in the numerical examples section, for both uniform and nonuniform inputs. The choice of grid points should be made in a problem-dependent manner; one needs to consider the expected regularity of the solution, PDF of the inputs, and boundary point issues in order to choose a sufficiently robust and efficient cubature rule for the problem. In addition, if *h*-adaptivity is required one must consider the accuracy of projection onto the orthogonal basis when prescribing the rule.

In the numerical examples section we also demonstrated ME-PCM convergence for the discontinuous stochastic Kraichnan-Orzag problem. It was shown that for the two-dimensional K–O problem the ME-PCM outperforms PCM and the quasi Monte Carlo method. We also demonstrated the convergence of the *h*-adaptive ME-PCM on the one- and three-dimensional K–O problems, and showed that the computational cost of the *h*-adaptive ME-PCM is much less than that of the *h*-adaptive ME-gPC method. The effectiveness of the ME-PCM in long-term integration was addressed using the problem of noisy flow past a 2D stationary cylinder. Lastly, convergence was demonstrated for the stochastic diffusion problem with higher-dimensional inputs and the efficiency of the method was shown to be more favorable than that of the Monte Carlo method even for 50 dimensions. A surprising result, which we plan to investigate further in future studies, is the good performance of the adaptive ME-PCM for high-dimensional problems in cases of *very coarse* sampling.

Acknowledgment

This work was partially supported by the computational mathematics programs of DOE, AFOSR and NSF/AMC-SS program. J. Foo would like to acknowledge the support of DOE Computational Science Graduate Fellowship under Grant number DE-FG02-97ER25308 and the Krell Institute and travel support by NSF Grant OISE-0456114. In addition, we would like to thank Prof. Christoph Schwab and Marcel Bieri at ETH Zurich for their helpful suggestions.

Appendix A

A.1. Tensor product Lagrangian interpolation

The description in this section closely mirrors the description in the work of [12] since it addresses the procedure over out chosen reference element, Γ, only. However, we include it here to explicitly define the procedure and notation for the reader. First define the polynomial space $\mathbb{P}_{p_j}(\Gamma_j)$ as the span of polynomials of degree at most p_j in Γ_j, for $j = 1, \ldots, N$. Then define $\mathbb{P}_{\mathbf{p}}(\Gamma)$ to be the span of tensor product polynomials $\mathbb{P}_{p_j}(\Gamma_j)$, where $\mathbf{p} = (p_1, p_2, \ldots, p_N)$:

$$\mathbb{P}_{\mathbf{p}}(\Gamma) = \operatorname{span}\{\mathbb{P}_{p_1}(\Gamma_1) \otimes \cdots \mathbb{P}_{p_N}(\Gamma_N)\}$$

In this section we seek an interpolation operator $\mathcal{I}_\Gamma$ which will take $u_k(\cdot,\cdot)$ to the space $\mathbb{P}_{\mathbf{p}}(\Gamma) \otimes \mathcal{W}_k(D)$. In each dimension $j = 1, \ldots, N$ consider the set of polynomials on Γ_j which are orthogonal with respect to some density function ρ_j, and let $\rho = \prod_{j=1}^{N} \rho_j$. Define $\{q_{j,i}\}_{i=1}^{p_j+1}$ to be the $p_j + 1$ roots of the p_jth order polynomial of this family. For any coordinate N-tuple of integer indices $[m_1, m_2, \ldots, m_N]$ where each index $m_j \in [1, p_j + 1]$, a global index may be associated:

$$m = m_1 + p_1(m_2 - 1) + p_1 p_2(m_3 - 1) + \cdots$$

and the associated N-tuple $(q_{1,m_1}, q_{2,m_2}, \ldots, q_{N,m_N})$ is called q_m. Considering all possible vectors $[m_1, \ldots, m_N]$ we obtain a set of points $\{q_m\}_{m=1}^{r}$ where the total number of points $r = \prod_{j=1}^{N}(p_j + 1)$.

Now define $\{l_i^j\}_{i=1}^{p_j+1}$ to be the one-dimensional Lagrange polynomial basis for $\mathbb{P}_{p_j}(\Gamma_j)$ on the abscissas $\{q_{j,i}\}_{i=1}^{p_j+1}$. The N-dimensional tensor product Lagrange basis on Γ is then defined to be the set of polynomials $l_m(y) = \prod_{j=1}^{N} l_{m_j}^j(y_j)$. The N-dimensional Lagrangian interpolant of u_k in Γ is then given by the following:

$$L_\Gamma^{\mathbf{p}}(u)(y) = \sum_{m=1}^{r} u_k(x, q_m) \cdot l_m(y)$$

Hence, the operator $L_\Gamma^{\mathbf{p}}$ takes $u_k \in C^0(\Gamma; \mathcal{W}_k(D))$ to the space $\mathbb{P}_{\mathbf{p}}(\Gamma) \otimes \mathcal{W}_k(D)$.

It is also important that the collocation points coincide with the points of a cubature rule in each element with respect to the weight ρ. With the tensor product Gaussian abscissas we have that for any continuous $v : \Gamma \to \mathbb{R}$,

$$\sum_{m=1}^{r} v(q_m) w_m \approx \int_{\Gamma} v(y)\rho(y)\,dy,$$

where the $w_m = \prod_{j=1}^{N} w_{m_j}$, and $w_{m_j} = \int_{\Gamma_j} l^2_{m_j}(y)\rho_j(y)\,dy$. In Section 3, $L^{\mathbf{p}}_{B^i}$ refers to the Lagrangian interpolant $L^{\mathbf{p}}_{\Gamma}$ defined in element B^i of the mesh $\mathcal{T}_{h,\Gamma}$ constructed using roots of polynomials orthogonal with respect to the conditional probability density function $\eta^i(y) = \prod_{j=1}^{N} \eta_j(y_j)$ defined in (10). For more details on the construction of polynomials orthogonal with respect to the conditional density function, please see [4].

A.1.1. Smolyak sparse grid interpolation

Another choice for the operator $\mathcal{I}_\Gamma$ is the isotropic Smolyak sparse grid operator, which was introduced in [23] by Smolyak. This algorithm provides an alternative to the more costly tensor product rule described above. It has previously been used in other works for stochastic collocation ([2,14,15]). Recently an anisotropic formulation of the Smolyak sparse grid has been introduced in [16] for the probabilistic collocation method. While we do not use it in this work, this anisotropic formulation could also be implemented in the ME-PCM. The following closely follows the description of the Smolyak method in [26,31], and [25] where the interpolation and cubature errors and costs of this method are investigated.

In this section we will assume that $\Gamma = [-1,1]^N$ without loss of generality, since the N-dimensional element can always be mapped to this standard element. We begin by choosing a one-dimensional interpolation formula $\mathcal{V}^i_j$ suited to the setting in which we are interested. For functions $v : [-1,1] \to \mathcal{W}_k(D)$, define

$$\mathcal{V}^i_j(v) = \sum_{m=1}^{n_i} v(y^i_m) \cdot a^i_m,$$

where $i \in \mathbb{N}$ specifies the degree of the interpolation, n_i is the number of points used, $a^i_m \in C([-1,1])$, and the $\{y^i_m\}_{m=1}^{n_i}$ are interpolation abscissas in $[-1,1]$. The index j indicates that this interpolation formula is used in the jth dimension. In practice, we will always use the same formulas in every dimension, but for now we will retain the subscript in order to better elucidate the Smolyak construction.

This one-dimensional formula could be chosen to be the Lagrangian interpolant on Gaussian abscissas as described above. In that case,

$$\mathcal{V}^i_j(v) = \sum_{m=1}^{n_i} v(q_{j,m}) \cdot l^j_m,$$

where the $q_{j,m}$ are the roots of the $(n_i - 1)$th degree orthogonal polynomial in the jth dimension as described above and $\{l^j_m\}_{m=1}^{n_i}$ are the Lagrange interpolating polynomials through these abscissas. Recall that the orthogonality of the polynomials generating the abscissas is with respect to the weight ρ_j.

The Clenshaw-Curtis interpolant is another choice for the one-dimensional formula $\mathcal{V}^i_j$. In this case, the abscissas $\{y^i_m\}_{m=1}^{n_i}$ would be Clenshaw-Curtis points, which can be found in [25] and the interpolating polynomials a^i_m are chosen such that $\mathcal{V}^i_j$ reproduces exactly all polynomials of degree less than n_i. Please see the references [25,26,31] and others for more detail on the Clenshaw-Curtis interpolant.

The one-dimensional interpolant serves as a building block for the Smolyak formula, as we will see soon. In this work we choose $n_1 = 1$ and $n_i = 2^{i-1} + 1$ as recommended in [31]. With this choice the Clenshaw-Curtis point sets are nested, which reduces the number of points used in total.

Define $\mathcal{V}^0_j = 0$ and $\Delta^i_j = \mathcal{V}^i_j - \mathcal{V}^{i-1}_j$. The Smolyak algorithm is:

$$\mathcal{S}_\Gamma(s) = \sum_{|\mathbf{i}| \leqslant s} (\Delta^{i_1}_1 \otimes \cdots \otimes \Delta^{i_N}_N) \tag{31}$$

where the summation is over N-dimensional vectors $\mathbf{i}$ with components $i_1, \ldots, i_N \in \mathbb{N}$. The parameter s controls the 'sparseness' of the grid; larger s results in more points.

The operator can also be rewritten as:

$$\mathcal{S}_\Gamma(s) = \sum_{s-N+1 \leqslant |\mathbf{i}| \leqslant s} (-1)^{s-|\mathbf{i}|} \binom{N-1}{s-|\mathbf{i}|} \cdot (\mathcal{V}^{i_1}_1 \otimes \cdots \otimes \mathcal{V}^{i_N}_N). \tag{32}$$

From now on we assume that the interpolation rule $\mathcal{V}^i_j$ is the same for all dimensions j, so that we can drop the subscript. Let χ^i denote the one dimensional point set used in $\mathcal{V}^i$. The total set of points used in $\mathcal{S}_\Gamma(s)$ is:

$$\mathcal{H}_\Gamma(s) = \bigcup_{s-N+1 \leqslant |\mathbf{i}| \leqslant s} (\chi^{i_1} \times \cdots \times \chi^{i_N})$$

Then, in the general notation we have used above, the collocation points are given by $\{q_j\}_{j=1}^{r}$ where each $q_j \in \mathcal{H}_\Gamma(s)$ and the total number of points $r = \mathrm{Card}(\mathcal{H}_\Gamma(s))$. When Clenshaw-Curtis one-dimensional rules are used with this choice of n_i, the

point sets are nested (i.e. $\chi^{i-1} \subset \chi^i$). When using nested one-dimensional rules the Smolyak formula is actually interpolatory (see [31] for details).

The Smolyak formula can also be used as a cubature formula over Γ with respect to the weight $\rho = \prod_{j=1}^{N} \rho_j$ as in the previous section. To do this, we simply replace $\mathcal{V}^i$ with the corresponding one-dimensional quadrature formula $\mathcal{U}^i$ in Eqs. (31) and (32) of the Smolyak construction. In other words, let

$$\mathcal{U}_j^i(v) = \sum_{m=1}^{n_i} v(y_m^i) \cdot w_m^i$$

approximate the integral

$$\int_{[-1,1]} v(y)\rho_j(y)\,\mathrm{d}y$$

for $i \in \mathbb{N}$ and $j = 1, \ldots, N$. Then, the Smolyak $N-$dimensional cubature operates

$$\sum_{s-N+1 \leqslant |\mathbf{i}| \leqslant s} (-1)^{s-|\mathbf{i}|} \binom{N-1}{s-|\mathbf{i}|} \cdot (\mathcal{U}_1^{i_1} \otimes \cdots \otimes \mathcal{U}_N^{i_N}).$$

approximates the N-dimensional integral

$$\int_{\Gamma} v(y)\rho(y)\,\mathrm{d}y.$$

for smooth functions $v : \Gamma \to \mathbb{R}$. Thus, the weights for the N-dimensional cubature are combinations of products of the one-dimensional weights; see [25] for more details. In the Section 3, $\mathcal{S}_{B^i}(s)$ refers to the Smolyak operator $\mathcal{S}_\Gamma(s)$ constructed in the element B^i of the mesh $\mathcal{T}_{h,\Gamma}$ using conditional probability density function $\eta^i(y) = \prod_{j=1}^{N} \eta_j(y_j)$ defined in (10).

We make a choice for the approximating operator $\mathcal{I}_{B^i}$ in each element based on the needs of the problem. We may choose $\mathcal{I}_{B^i} \equiv L_{B^i}^{\mathbf{p}}$ where $\mathbf{p}$ determines the degree of the interpolant and thus the number of points used. Alternatively, we can choose $\mathcal{I}_{B^i} \equiv \mathcal{S}_{B^i}(s)$, where s controls the order of the approximant and also the number of points used. In this work we consider both tensor product and sparse grids, built from both Gaussian and Clenshaw-Curtis formulations for one-dimensional bases. In practice, the choice of what type of approximation to use should be problem-dependent, and factors to take into consideration include: the suitability of a rule to the density function ρ, preference for points on or off boundaries, regularity of the integrand and robustness of the rule, and of course the number of points required to achieve a particular degree of exactness. This last consideration is addressed in detail in [26]. Since nested quadrature rules require less overall points in the Smolyak algorithm, the Clenshaw-Curtis rule is often an attractive choice.

A.2. Proof of Theorem 4

Proof. Recall that we have defined a linear functional on $W^{m+1,\infty}(A)$:

$$E_A(g) \equiv \int_A g(x)\,\mathrm{d}x - \mathcal{Q}_m^A(g) \tag{33}$$

whose norm is defined as

$$\|E_A\|_{k,\infty,A} = \sup_{\|g\|_{k,\infty,A} \leqslant 1} |E_A(g)|. \tag{34}$$

Let $P_m(A)$ denote the space of all polynomials of degree $\leqslant m$ in the variable $x \in A \subset \Gamma$. Since the quadrature rule $\mathcal{Q}_m^A$ has a degree of exactness of m, i.e.,

$$E_A(g) = 0, \quad \forall g \in P_m(A)$$

we know from the Bramble-Hilbert lemma that there exists a constant $C(A)$ such that

$$|E_A(f)| \leqslant C(A)\|E_A\|_{m+1,\infty,A}|f|_{m+1,\infty,A}, \tag{35}$$

where the constant $C(A)$ is determined by A.

For element $B^i = \prod_{k=1}^{N}(a_k, b_k)$ we define an invertible affine mapping

$$F_i : x \in \Gamma \to F_i(x) = q^{\mathrm{T}}x + c \in B^i,$$

where the constant vectors $q, c \in \mathbb{R}^n$. Since $\Gamma = (0,1)^N$, it is easy to see that the component $q_k = b_k - a_k \leqslant h$, $k = 1, \ldots, N$. We now examine the relation between $|f|_{m+1,\infty,F_i^{-1}(B^i)}$ and $|f|_{m+1,\infty,B^i}$.

$$|f|_{m+1,\infty,F_i^{-1}(B^i)} = \max_{|\alpha|=m+1} \operatorname{ess\,sup}_{x \in F_i^{-1}(B^i)} |D^\alpha f| = \max_{|\alpha|=m+1} \operatorname{ess\,sup}_{x \in B^i} q^\alpha |D^\alpha f| \leqslant h^{m+1}|f|_{m+1,\infty,B^i}, \tag{36}$$

where $q^\alpha = \prod_{k=1}^{N} q_k^{\alpha_k}$.

Using inequalities (35) and (36), we have

$$\left|\int_\Gamma f(x)\,dx - \sum_{i=1}^{N_e} \mathcal{Q}_m^{B_i} f(x)\right| = \left|\sum_{i=1}^{N_e}\int_{B^i} f(x)\,dx - \mathcal{Q}_m^{B_i} f(x)\right| \leqslant \sum_{i=1}^{N_e} \mathrm{vol}(B^i)\left|\int_{F_i^{-1}(B^i)=\Gamma} f(y)\,dy - \mathcal{Q}_m^{\Gamma}(f(y))\right|$$
$$\leqslant \sum_{i=1}^{N_e} \mathrm{vol}(B^i) C(\Gamma)\|E_\Gamma\|_{m+1,\infty,\Gamma}|f|_{m+1,\infty,F_i^{-1}(B^i)} \leqslant C(\Gamma)h^{m+1}\sum_{i=1}^{N_e}\mathrm{vol}(B^i)\|E_\Gamma\|_{m+1,\infty,\Gamma}|f|_{m+1,\infty,B^i}$$
$$\leqslant C(\Gamma)h^{m+1}\|E_\Gamma\|_{m+1,\infty,\Gamma}|f|_{m+1,\infty,\Gamma},$$

which concludes the proof. □

A.3. Proof of Theorem 9

Proof. Recall that $L_{B^i}^{\mathbf{p}}$ is defined to be an affine mapping of the interpolation operator $L_\Gamma^{\mathbf{p}}$ from Γ into B^i. The first step is to bound the error in each element B^i. To do this, we first map from the element to the reference element $\Gamma = (0,1)^N$. Let B^i be an element in the mesh. Then, define $\mathbf{c} = (c_1, c_2, \ldots, c_N)$ to be the vector in $\mathbb{R}^N$ such that $B^i = (c_1, c_1+h)\times\cdots\times(c_N, c_N+h)$.

Now let y be any point in $\mathbb{R}^N$ and define the mapping $F_i(y) = yh + \mathbf{c}$ and let $\bar{u} \equiv u \circ F_i$. Then, $F_i : \Gamma \to B^i$, and

$$\|L_{B^i}^{\mathbf{p}} u - u\|_{L^2(B^i;H_0^1(D))} = h^{N/2}\|L_\Gamma^{\mathbf{p}}\bar{u} - \bar{u}\|_{L^2(\Gamma;H_0^1(D))}.$$

We can bound this last quantity using Lemma 1, as long as we ensure that Regularity Property 1 holds for $\bar{u}$ and determine the size of the domain of the analytic extension.

Recall that $u : \Gamma \to H_0^1(D)$ has an analytic extension u_a^j in each dimension j, which we can think of as a function taking $\mathbb{C}\times\Gamma_j^*$ to the space $H_0^1(D)$. Note that u_a can also be thought of as a function of (z_2, y_j, y_j^*), where $z_2 \in \mathbb{R}$, taking values again in $H_0^1(D)$. To see this, consider the element $z \in \mathbb{C}$ associated with the coordinate pair (z_1, z_2) through the relation $z = z_1 + iz_2$. In the context of $u_a^j : \mathbb{C}\times\Gamma_j^* \to H_0^1(D)$, the first space $\mathbb{C}$ is actually the complex plane in which the real axis corresponds to y_j. Thus, dependence on $z \in \mathbb{C}$ can be written as dependence on a coordinate pair (y_j, z_2) where $z_1 = y_j$. We now define a function $\bar{u}_a^j$ by making the following mapping: Let $z_2 \in \mathbb{R}$, $y_j \in \mathbb{R}$, and and $y_j^* \in \mathbb{R}^{N-1}$, and define the new mapping

$$\bar{u}_a^j(z_2, y_j, y_j^*) \equiv u_a^j\left(hz_2, hy_j + c_j, hy_j^* + c_j^*\right)$$

where c_j^* is the vector $\mathbf{c}$ with the jth dimension deleted. This function is defined for all (z_2, y_j, y_j^*) such that the analytic extension u_a is defined. It is clear then that $\bar{u}_a^j$ coincides with $\bar{u}$ for all (z_2, y_j, y_j^*) such that $z_2 = 0$ and $y_j \in \Gamma_j$. Thus, it is an extension of $\bar{u}|_{\Gamma_j}$. In addition, since this mapping does not affect the analyticity properties of a function, $\bar{u}_a^j$ as a function of the pair, $(z_2, y_j) \in \mathbb{C}$, is analytic for all (z_2, y_j) such that $(hz_2, hy_j + c_j) \in \Sigma$. In particular, we can say that $\bar{u}_a^j$ is an analytic extension of $\bar{u}|_{\Gamma_j}$ in the region:

$$\overline{\Sigma}(\Gamma_j; \frac{1}{h}\alpha_j) = \left\{x \in \mathbb{C}, dist(x, \Gamma_j) \leqslant \frac{1}{h}\alpha_j\right\}. \tag{37}$$

Furthermore, $\forall z \in \overline{\Sigma}(\Gamma_j; \frac{1}{h}\alpha_j), \|\bar{u}_a^j(z)\|_{C^0(\Gamma_j^*;H_0^1(D))} \leqslant \lambda$. The analyticity region can be much larger than this for certain elements, but this region is a minimum that is valid for all elements. Note that the bound λ is also independent of the element choice. Thus, using Lemma 1, we can bound

$$\|L_\Gamma^{\mathbf{p}}\bar{u} - \bar{u}\|_{L^2(\Gamma;H_0^1(D))} \leqslant C\sum_{j=1}^{N} p_j \exp\{-r_j(h)p_j\},$$

with $r_j(h)$ defined as

$$r_j(h) = \log\left[\frac{\alpha_j}{2h}\left(1 + \sqrt{1 + \frac{h^2}{\alpha_j^2}}\right)\right].$$

The constant C depends on the function $\bar{u}$ only through the quantity that bounds $\|\bar{u}_a^j(z)\|_{C^0(\Gamma_j^*;H_0^1(D))}$. Since this quantity is bounded by λ, and this bound holds for any j and any element i, C can be chosen to be independent of the element choice.

This bound is independent of the element choice B^i so it is a uniform bound for the L^2 interpolation error over every element. Then,

$$\|\tilde{u} - u\|^2_{L^2(\Gamma;H_0^1(D))} = \|\sum_{i=1}^{N_e}(L_{B^i}^{\mathbf{p}}u)\mathbb{I}_{\{y\in B^i\}} - u\|^2_{L^2(\Gamma;H_0^1(D))} = \sum_{i=1}^{N_e}\|L_{B^i}^{\mathbf{p}}u - u\|^2_{L^2(B^i;H_0^1(D))} = \sum_{i=1}^{N_e} h^N\|L_\Gamma^{\mathbf{p}}\bar{u} - \bar{u}\|^2_{L^2(\Gamma;H_0^1(D))}$$
$$\leqslant \left(C\sum_{j=1}^{N}\exp\{-p_j r_j(h)\}\right)^2. \qquad \square$$

A.4. An adaptive procedure

We denote the gPC expansion of a random field in element B^k as:

$$\hat{u}(y) = \sum_{j=0}^{N_p} \hat{u}_j \Phi_{k,j}(y),$$

where p is the highest order of polynomial chaos, N_p denotes the total number of basis modes for a gPC expansion of maximum order p in N dimensions, and $\{\Phi_{k,j}\}_{j=1}^{\infty}$ is the local orthogonal polynomial chaos basis in element B^k.

By noting that there exists a unique correspondence between the gPC basis and the Lagrange basis defined by grid points used in ME-PCM, we can employ the adaptive criterion developed in [4] in the following manner. We must first obtain the gPC coefficients $\hat{u}_j$ of the solution in each element since the adaptivity criterion is evaluated in terms of these coefficients. To do this, we project the collocation solution onto each basis function $\Phi_{k,j}$ to obtain the coefficient $\hat{u}_j$ using the numerical quadrature rule associated with the collocation points. From the orthogonality of gPC we can easily obtain the local variance given by polynomial chaos with order p:

$$\varrho_{k,p}^2 = \sum_{j=1}^{N_p} \hat{u}_j^2 \mathbb{E}\left[\Phi_{k,j}^2\right].$$

We define the decay rate of relative error of polynomial chaos in each element as follows:

$$\vartheta_k = \frac{\sum_{i=N_{p-1}+1}^{N_p} \hat{u}_i^2 \mathbb{E}\left[\Phi_{k,i}^2\right]}{\varrho_{k,p}^2}.$$

Based on ϑ_k and the scaled parameter $\Pr(Y(\omega) \in B^k)$, we implement *h-type refinement*, in other words, we decompose the current random element into smaller ones, if the following criterion

$$\vartheta_k^{\gamma} Pr\left(Y(\omega) \in B^k\right) \geqslant \theta, \quad 0 < \gamma < 1$$

is satisfied, where γ and θ are prescribed constants. The sensitivity of each random dimension is defined as

$$r_i = \frac{\left(\hat{u}_{i,p}\right)^2 \mathbb{E}[\Phi_{i,p}^2]}{\sum_{j=N_{p-1}+1}^{N_p} \hat{u}_j^2 \mathbb{E}[\Phi_j^2]}, \quad i = 1, 2, \ldots, N \tag{38}$$

where we drop the subscript k for clarity, and the subscript $*_{i,p}$ denotes the mode varying in only the ith random dimension with polynomial order p. All random dimensions which satisfy

$$r_i \geqslant \alpha \cdot \max_{j=1,\ldots,d} r_j, \quad 0 < \alpha < 1, \quad i = 1, 2, \ldots, N \tag{39}$$

will be split into two *equal* random elements in the next time step while all other random dimensions will remain unchanged. To split each element, a new collocation grid on each daughter element is constructed. Then, the ME-PCM interpolant at each new collocation point is evaluated to provide a current set of solutions at every collocation point in the domain.

We note that not all gPC coefficients $\hat{u}_i$, $i = 0, \ldots, N_p$ are utilized in the adaptivity criterion. In fact, from above we can see that it is only necessary to project the collocation solution onto the highest modes of the basis in order to evaluate this criterion. The order p of the polynomial chaos basis used can be approximately determined by the following proposition.

Proposition 12. *To maintain an accurate transformation between collocation solutions and the gPC spectral expansion, the polynomial order of the gPC basis can be taken up to $\lfloor m/2 \rfloor$, where $\lfloor * \rfloor$ indicates the integer not larger than $*$ and m indicated the degree of exactness of the quadrature rule.*

Proof. The conclusion can be obtained directly from the definition of Galerkin projection. □

In the examples we used $\gamma = 0.5$, $\alpha = 0.01$ and we varied θ. More details on the adaptivity criteria can be found in [4,18].

References

[1] I. Babuska, R. Tempone, G.E. Zouraris, Galerkin finite element approximations of stochastic elliptic differential equations, SIAM J. Numer. Anal. 42 (2) (2004) 800–825.
[2] D. Xiu, G.E. Karniadakis, The Wiener–Askey polynomial chaos for stochastic differential equations, SIAM J. Sci. Comput. 24 (2) (2002) 619–644.
[3] D. Xiu, J. Hesthaven, High-order collocation methods for differential equations with random inputs, SIAM J. Sci. Comput. 27 (3) (2005) 1118–1139.
[4] X. Wan, G.E. Karniadakis, Multi-element generalized polynomial chaos for arbitrary probability measures, SIAM J. Sci. Comput. 28 (3) (2006) 901–928.
[5] L. Mathelin, M. Hussaini, A stochastic collocation algorithm for uncertainty analysis, NASA/CR-2003-212153.
[6] M.K. Deb, I. Babuska, J.T. Oden, Solution of stochastic partial differential equations using Galerkin finite element techniques, Comput. Methods Appl. Mech. Eng. 190 (2001) 6359–6372.
[7] R.G. Ghanem, P. Spanos, Stochastic Finite Elements: A Spectral Approach, Springer-Verlag, New York, 1991.

[8] P. Frauenfelder, C. Schwab, R.A. Todor, Finite elements for elliptic problems with stochastic coefficients, Comput. Methods Appl. Mech. Eng. 194 (2005) 205–228.
[9] R.A. Todor, C. Schwab, Convergence rates for sparse chaos approximations of elliptic problems with stochastic coefficients, IMA J. Numer. Anal. 27 (2) (2007) 232–261.
[10] H.G. Matthies, A. Keese, Galerkin methods for linear and nonlinear stochastic partial differential equations, Comput. Methods Appl. Mech. Eng. 194 (12-16) (2005) 1295–1331.
[11] M. Tatang, G. McRae, Direct treatment of uncertainty in models of reaction and transport, MIT Technical Reports.
[12] I. Babuska, F. Nobile, R. Tempone, A stochastic collocation method for elliptic partial differential equations with random input data, SIAM J. Numer. Anal. 45 (3) (2007) 1005–1034.
[13] B. Ganapathysubramanian, N. Zabaras, Sparse grid collocation for stochastic natural convection problems, J. Comput. Phys. 225 (1) (2007) 652–685.
[14] A. Keese, H.G. Matthies, Numerical methods and smolyak quadrature for nonlinear stochastic partial differential equations, Informatikbericht 2003-5, Technische Universitat Braunschweig.
[15] F. Nobile, R. Tempone, C. Webster, A sparse grid collocation method for elliptic partial differential equations with random input data, SIAM J. Numer. Anal. 46 (5) (2008) 2309–2345.
[16] F. Nobile, R. Tempone, C. Webster, An anisotroopic sparse grid collocation method for elliptic partial differential equations with random input data, SIAM J. Numer. Anal. 46 (5) (2008) 2411–2442.
[17] O.P.L. Maitre, H.N. Najm, R.G. Ghanem, O.M. Knio, Multi-resolution analysis of Wiener-type uncertainty propagation schemes, J. Comput. Phys. 197 (2004) 502–531.
[18] X. Wan, G.E. Karniadakis, An adaptive multi-element generalized polynomial chaos method for stochastic differential equations, J. Comput. Phys. 209 (2) (2005) 617–642.
[19] B. Oksendal, Stochastic Differential Equations, Springer-Verlag, 1998.
[20] M. Loéve, Probability Theory, fourth ed., Springer-Verlag, New York, 1977.
[21] S. Brenner, L. Scott, The Mathematical Theory of Finite Element Methods, Springer, 2002.
[22] G.E. Karniadakis, S.J. Sherwin, Spectral/*hp* Element Methods for CFD, Oxford University Press, 2005.
[23] S. Smolyak, Quadrature and interpolation formulas for tensor products of certain classes of functions, Soviet Math. Dokl. 4 (1963) 240–243.
[24] T. Gerstner, M. Griebel, Dimension-adaptive tensor-product quadrature, Computing 71 (1) (2003) 65–87.
[25] E. Novak, K. Ritter, Simple cubature formulas with high polynomial exactness, Constr. Approx. 15 (1999) 499–522.
[26] E. Novak, K. Ritter, High dimensional integration of smooth functions over cubes, Numer. Math. 75 (1996) 79–97.
[27] A. Genz, A package for testing multiple integration subroutines, in: Numerical Integration: Recent Developments, Software and Applications, 1987, pp. 337–340.
[28] S.A. Orszag, L.R. Bissonnette, Dynamical properties of truncated Wiener–Hermite expansions, Phys. Fluids 10 (12) (1967) 2603–2613.
[29] L. Kovasznay, Laminar flow behind a two-dimensional grid, in: Proceedings of the Cambridge Philosophical Society, 1948, pp. 44.
[30] X. Wan, G.E. Karniadakis, Long-term behavior of polynomial chaos in stochastic flow simulations, Comput. Methods Appl. Math. Eng. 195 (41–43) (2006) 5582–5596.
[31] V. Barthelmann, E. Novak, K. Ritter, High dimensional polynomial interpolation on sparse grids, Adv. Comput. Math. 12 (2000) 273–288.