ICES REPORT 13-08

April 2013

A posteriori error control for partial differential equations with random data

by

Corey M. Bryant, Serge Prudhomme, and Timothy Wildey

The Institute for Computational Engineering and Sciences
The University of Texas at Austin
Austin, Texas 78712

Reference: Corey M. Bryant, Serge Prudhomme, and Timothy Wildey, A posteriori error control for partial differential equations with random data, ICES REPORT 13-08, The Institute for Computational Engineering and Sciences, The University of Texas at Austin, April 2013.

A POSTERIORI ERROR CONTROL FOR PARTIAL DIFFERENTIAL EQUATIONS WITH RANDOM DATA

C. M. BRYANT*, S. PRUDHOMME†, AND T. WILDEY‡

Abstract.
In this work, we investigate adaptive approaches to control errors in numerical approximations of differential equations with uncertain or random data and coefficients. The adaptivity is based on *a posteriori* error estimation and the approach relies on the ability to decompose the *a posteriori* error estimate into contributions from the physical discretization and the approximation in stochastic space. Errors are evaluated in terms of linear quantities of interest using adjoint-based methodologies. We demonstrate that a significant reduction in the computational cost required to reach a given error tolerance can be achieved by refining the dominant error contributions rather than uniformly refining both the physical and stochastic discretization. Error decomposition is demonstrated for a two-dimensional flow problem and adaptive procedures are tested on a convection-diffusion problem with discontinuous parameter dependence and a diffusion problem where the diffusion coefficient is characterized by a 10-dimensional parameter space.

1. Introduction. Virtually all mathematical models of physical and biological phenomena present uncertainties resulting from a lack of knowledge, or miss-specification, of the governing physics of the system, such as material properties or external forcing. Uncertainty quantification involves the process of characterizing these uncertainties using probabilistic models based on observations of the system response. Assuming these uncertainty models are correct, accurate propagation of the uncertainty through the model is still necessary to verify the probabilistic models and is thus vital in increasing confidence in predictions obtained from mathematical and computational models.

A thorough uncertainty propagation requires a large number of additional simulations to reproduce the range of behavior of the system; one often resorts to approximate representations of the uncertainty, or surrogate models, that can be constructed from a limited number of simulations of the fill system. Yet, for complex simulations, with many uncertain parameters, the number of full system simulations may still be computationally prohibitive. Adaptive solution techniques represent an attractive enhancement to surrogate methods in order to accelerate convergence of the approximations to the true system response. These techniques include adaptive sparse grid collocation approaches [1,2,25,26] and methods based on the use of piecewise polynomial or multi-wavelet representations, which allow for local refinement of the surrogate representation [20,21,31]. A significant improvement is seen over uniform refinement in terms of computational efficiency [20], but the indicators used to drive adaptive refinement are predominately heuristic and are often tailored to specific problems.

Methodologies based on *a posteriori* error estimation of the solution are more effective at reducing approximation errors in computed solutions [3, 4, 13, 14, 16, 22]. One may consider estimating errors in terms of global norms but it is often more useful to define errors with respect to specific features of the solution, the so-called

*Institute for Computational Engineering and Sciences (ICES), The University of Texas at Austin, Austin, Texas.

†Département de Mathématiques et de Génie Industriel, Ecole Polytechnique de Montréal, Montréal, Québec, Canada.

‡Optimization and Uncertainty Quantification Department, Sandia National Labs, Albuquerque, NM. Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

quantities of interest. Methods for error estimation in quantities of interest, referred to as dual-weighted residual methods or goal-oriented error estimation methods, are based on the solution of an adjoint problem related to the particular quantity of interest (see, e.g. [3,4,15,19,27] and the references therein). While these goal-oriented error estimation methods do require solving an additional linear problem they do have a strong theoretical foundation and have been demonstrated to be quite efficient in guiding adaptive mesh refinement.

Recent works have extended these approaches to problems with uncertain data [1,5,6,23,24]. Mathelin and Le Maître [24] for example, present a goal-oriented error estimation approach based on a straightforward extension of existing a posteriori error estimation techniques to the stochastic finite element framework. They propose error indicators to perform adaptive mesh refinement in both the physical and parameter spaces [23,24]. Butler et al. [6] also propose error estimates in a stochastic Galerkin framework using polynomial chaos expansions of the solution and demonstrate application to non-linear problems. However, their work focused on the estimation of errors due to physical discretization only and neglected errors due to approximation in parameter space. Furthermore, the authors did not address adaptivity. Error estimation in non intrusive approaches for uncertainty quantification has also been examined; these approaches are particularly attractive for the ability to use existing software in a sampling fashion. Almeida and Oden [1] solved a convective-diffusion problem with uncertain coefficients by the collocation method and proposed a procedure to estimate the error in the quantity of interest at each collocation point. These error estimates were used to guide adaptivity in both physical space and parameter space until the estimate of the variance converged. In [5], Butler et al. considered parameterized linear systems, such as those that arise from discretized partial differential equations, but did not investigate errors due to the physical discretization. *A posteriori* error estimates were established for the error due to approximation in parameter space, but were not used for adaptivity.

If the objective in an *a posteriori* error analysis is to define local error indicators to guide adaptive refinement, then it is critical for the error estimate to discern whether discretization in physical space or parameter space contributes more to the total approximation error. In fact, Le Maître and Knio [23, Ch. 9], stated in a section on refinement strategies that:

> "Specifically, it is not possible to decide (a) between h or p refinement and (b) whether one should enrich the approximation space $\mathcal{V}^h$ [physical] or $\mathcal{S}^h$ [stochastic]."

The authors emphasize in particular that, while trial based approaches exist, "[...] better approaches, yet to be conceived, are consequently needed." Development of novel methods that address this issue is the objective of the current work. By combining polynomial chaos expansions, spectral projection, and goal-oriented error estimation methods similar to [5,6] we are able to identity the sources of error that are most significant and hence the approximation space that needs to be refined. This allows for accurate assessment of the error in parameter space even with coarse discretization in physical space (and vice versa). In addition to the choice of which space to refine, knowledge of the error that arises solely from discretization in parameter space allows one to perform adaptive refinement, h or p, in parameter space without the interference of physical discretization error. We explore the use of the stochastic space error estimate for adaptive h refinement and finally demonstrate the use of these error estimates to guide anisotropic p refinement in a higher dimension parameter space.

This paper is organized as follows. We introduce in Section 2 an abstract model problem and precisely define the notion of a quantity of interest. In Section 3 we describe the discretization in physical space and recall existing a posteriori error estimation techniques for quantities of interest applied to the semi-discrete equations. We detail in Section 4 the parametric discretization of the problem. Section 5 describes the changes to the error representation formula under uncertainty and Section 6 gives the decomposition of the error estimate into physical and parametric contributions. Some adaptive refinement strategies are presented in Section 7. Numerical examples and results are given in Section 8. Section 8.1 shows an example of using the error estimate to perform adaptivity in both physical and parameter space on a simple fluid flow problem. Sections 8.2 and 8.3 demonstrate the use of the error estimate as an error indicator for h refinement and anisotropic p refinement, respectively. Finally, we provide concluding remarks in Section 9.

2. Abstract model problem and quantity of interest. We consider here an abstract model problem consisting of a boundary-value problem defined on an open bounded domain $D \subset \mathbb{R}^d$, with boundary ∂D, where d is the dimension of the physical space. In particular, we focus on second-order partial differential equations with uncertain parameters. The parameters shall be denoted by the random vector, or set of random variables, $\boldsymbol{\xi} = \{\xi_i(\theta)\}_{i=1}^n$. We thus introduce a probability space $\{\Theta, \Sigma, P\}$, where Θ is the sample space, Σ a σ-algebra on Θ, and P the probability measure. The abstract model problem can be cast in general form as,

$$\mathcal{A}(\boldsymbol{x}, \boldsymbol{\xi}; u) = f(\boldsymbol{x}, \boldsymbol{\xi}), \quad \forall \boldsymbol{x} \in D, \tag{2.1}$$

with appropriate boundary conditions on ∂D. Here $\mathcal{A}$ is a non-linear, second-order differential operator, $u = u(\boldsymbol{x}, \boldsymbol{\xi})$ is the solution of the differential equation (2.1), and $f = f(\boldsymbol{x}, \boldsymbol{\xi})$ represents the forcing term. Uncertainty is accounted for through the random parameters $\boldsymbol{\xi}$ and may appear in various coefficients in the definition of the operator $\mathcal{A}$, source terms, or boundary conditions. We consider a number of forms for $\mathcal{A}$ in the numerical experiments that follow. We assume that u is a solution in an appropriate functional space $\mathcal{V} = V_D \otimes V_\Theta$, where V_D is, for example, a Hilbert space on D and V_Θ a Hilbert space on Θ.

Although uncertainties in (2.1) lead u to be a random process and will require discretization in the space V_Θ, we interpret (2.1) for the moment as a collection of differential equations parameterized by $\boldsymbol{\xi}$. For any such equation we seek a variational, or weak, solution,

$$\text{find } u(\cdot, \boldsymbol{\xi}) \in V_D \text{ s.t.} \quad B_{\boldsymbol{\xi}}(u; v) = F_{\boldsymbol{\xi}}(v) \quad \forall v \in V_D, \tag{2.2}$$

where $B_{\boldsymbol{\xi}}(\cdot;\cdot)$ is a differentiable, semi-linear integral formulation of $\mathcal{A}$.

In general, one is often interested in a particular feature of the response, referred here to as a quantity of interest. A quantity of interest can usually be represented as a functional of the solution, $Q(u(\cdot, \boldsymbol{\xi}))$. We restrict ourselves, for the sake of simplicity in the exposition, to bounded linear functionals of the form:

$$Q(u(\cdot, \boldsymbol{\xi})) = \int_D q \; u \; d\boldsymbol{x} \tag{2.3}$$

where q is the kernel associated with the quantity of interest Q. Furthermore, we assume that the quantity of interest depends on the parameters $\boldsymbol{\xi}$ only through the solution itself, i.e. $q = q(\boldsymbol{x})$.

3. Discretization in physical space. We use a standard continuous finite element method to discretize on the physical domain. Let $V_D^h \subset V_D$ represent the finite element subspace consisting of piecewise continuous polynomials of order p, on a suitable partition of D with maximal element diameter h. Then, the semi-discrete weak formulation of (2.2) is,

$$\text{find } u^h(\cdot, \boldsymbol{\xi}) \in V_D^h \text{ s.t.} \quad B_{\boldsymbol{\xi}}(u^h; v) = F_{\boldsymbol{\xi}}(v) \quad \forall v \in V_D^h. \tag{3.1}$$

Equivalently, we seek the expansion in terms of finite element basis functions $\{\phi_j^h\}_{j=1}^l$, i.e. $V_D^h = \text{span}\{\phi_j^h\}$, resulting in a representation for u^h,

$$u^h(\boldsymbol{x}, \boldsymbol{\xi}) = \sum_{j=1}^{l} a_j(\boldsymbol{\xi}) \phi_j^h(\boldsymbol{x}), \tag{3.2}$$

with parameterized coefficients $a_j(\boldsymbol{\xi})$. Solving for the vector of coefficients $\boldsymbol{a}(\boldsymbol{\xi})$ requires the solution of a parameterized non-linear algebraic system,

$$\boldsymbol{B}(\boldsymbol{\xi}; \boldsymbol{a}(\boldsymbol{\xi})) = \boldsymbol{F}(\boldsymbol{\xi}), \tag{3.3}$$

where $\boldsymbol{B}(\boldsymbol{\xi}; \boldsymbol{a}(\boldsymbol{\xi}))$ and $\boldsymbol{F}(\boldsymbol{\xi})$ are defined as:

$$\left[\boldsymbol{B}(\boldsymbol{\xi}; \boldsymbol{a}(\boldsymbol{\xi}))\right]_i = B_{\boldsymbol{\xi}} \left(\sum_j a_j(\boldsymbol{\xi}) \phi_j^h; \phi_i^h \right), \tag{3.4}$$

$$\left[\boldsymbol{F}(\boldsymbol{\xi})\right]_i = F_{\boldsymbol{\xi}} \left(\phi_i^h \right). \tag{3.5}$$

We briefly recall here some existing *a posteriori* error estimation techniques to investigate the error in the semi-discrete solution for a given, fixed $\boldsymbol{\xi}$, i.e. $Q\left(u(\cdot, \boldsymbol{\xi})\right) - Q\left(u^h(\cdot, \boldsymbol{\xi})\right)$. It requires the definition of the so-called adjoint problem. For quantities of interest in the form of (2.3) we seek a generalized Green's function associated with the particular quantity of interest Q [3,4,15,27]. The definition of the adjoint problem requires a linear (or linearized) operator. For a general nonlinear PDE one approach to define the linear operator is to assume $B_{\boldsymbol{\xi}}(\cdot;\cdot)$ is convex and use the Integral Mean Value Theorem yielding

$$B'_{\boldsymbol{\xi}}(\hat{u}; e, v) = B_{\boldsymbol{\xi}}(u; z) - B_{\boldsymbol{\xi}}(u^h; z)$$

where $e = u - u^h$, and $\hat{u}$ lies on the segment connecting u and u^h. In practice, $\hat{u}$ is unknown so we linearize around u^h giving,

$$B'_{\boldsymbol{\xi}}(u^h; e, v) = B_{\boldsymbol{\xi}}(u; z) - B_{\boldsymbol{\xi}}(u^h; z) + \Delta_B,$$

where the remainder, Δ_B, is a higher-order perturbation term and is, in general, neglected [3, 4, 15]. The adjoint problem is then defined as,

$$\text{find } z(\cdot, \boldsymbol{\xi}) \in V_D \text{ s.t.} \quad B'_{\boldsymbol{\xi}}(u^h; v, z) = Q(v) \quad \forall v \in V_D. \tag{3.6}$$

Solution of the adjoint equation provides the following error representation formula. Let u and z be the solution to forward and adjoint problems, respectively, and

$u^h \in V_D^h$ the semi-discrete approximation of u, then

$$\begin{aligned} Q(u(\cdot,\boldsymbol{\xi})) - Q(u^h(\cdot,\boldsymbol{\xi})) &= B'_{\boldsymbol{\xi}}(u^h; u-u^h, z) && (3.7)\\ &= B_{\boldsymbol{\xi}}(u;z) - B_{\boldsymbol{\xi}}(u^h;z) + \Delta_B && (3.8)\\ &= F_{\boldsymbol{\xi}}(z) - B_{\boldsymbol{\xi}}(u^h;z) + \Delta_B && (3.9)\\ &:= \mathcal{R}_{\boldsymbol{\xi}}(u^h;z) + \Delta_B. && (3.10) \end{aligned}$$

If we solve (3.6) for a given choice of $\boldsymbol{\xi}$, then we can also utilize Galerkin orthogonality and write the error estimate as

$$Q(u(\cdot,\boldsymbol{\xi})) - Q(u^h(\cdot,\boldsymbol{\xi})) = \mathcal{R}_{\boldsymbol{\xi}}(u^h; z-\pi_{h,p}z), \tag{3.11}$$

where $\pi_{h,p} : V_D \to V_D^h$ is a projection operator into the semi-discrete space. Of course, one cannot, in general, solve (3.6) exactly and must resort to an approximation of the adjoint solution z. However, due to orthogonality, the adjoint approximation must be made in an enriched function space V_D^+, where $V_D^h \subset V_D^+ \subset V_D$. A common choice for the enriched space is to increase the polynomial order of the finite element basis $p \leftarrow p+1$, yielding the discrete function space V_D^+. The semi-discrete adjoint problem is thus,

$$\text{find } z^+(\cdot,\boldsymbol{\xi}) \in V_D^+ \text{ s.t.} \quad B'_{\boldsymbol{\xi}}(u^h; v, z^+) = Q(v) \quad \forall v \in V_D^+. \tag{3.12}$$

Similarly to the primal problem, the discrete solution to the adjoint problem can be expanded in terms of finite element basis functions $\{\phi_i^+\}$, where $V_D^+ = \text{span}\{\phi_i^+\}$,

$$z^+(\boldsymbol{x},\boldsymbol{\xi}) = \sum_i b_i(\boldsymbol{\xi})\phi_i^+(\boldsymbol{x}). \tag{3.13}$$

In contrast to the primal problem, the coefficients $\boldsymbol{b}(\boldsymbol{\xi})$ are computed by solving the parameterized *linear* system of equations

$$\boldsymbol{B}^*(\boldsymbol{\xi})\boldsymbol{b}(\boldsymbol{\xi}) = \boldsymbol{Q}, \tag{3.14}$$

where $\boldsymbol{B}^*$ and $\boldsymbol{Q}$ are defined as:

$$[\boldsymbol{B}^*(\boldsymbol{\xi})]_{ji} = B'_{\boldsymbol{\xi}}\left(u^h; \phi_j^+, \phi_i^+\right), \tag{3.15}$$

$$[\boldsymbol{Q}]_j = Q\left(\phi_j^+\right). \tag{3.16}$$

A computable error estimate can then be obtained using the semi-discrete primal and adjoint solutions,

$$\begin{aligned} Q\Big(u(\cdot,\boldsymbol{\xi}) - u^h(\cdot,\boldsymbol{\xi})\Big) &= \mathcal{R}_{\boldsymbol{\xi}}(u^h;z) + \Delta_B && (3.17)\\ &= \mathcal{R}_{\boldsymbol{\xi}}(u^h;z^+) + \mathcal{R}_{\boldsymbol{\xi}}(u^h; z-z^+) + \Delta_B && (3.18)\\ &\approx \mathcal{R}_{\boldsymbol{\xi}}(u^h;z^+) && (3.19) \end{aligned}$$

The second term in (3.18) can be neglected since it involves the product of the forward residual and the higher order adjoint error $z - z^+$.

4. Parameter space discretization. We now turn our attention to the discretization in parameter or stochastic space and the effect it has on the error representation formula established in the previous section. For the derivation that follows we will use generalized polynomial chaos expansions to represent the random variables and the so-called pseudospectral projection method to compute the coefficients. However, the error estimates and decomposition strategy that are described below are not restricted to this approach and can be applied to most non-intrusive methods for uncertainty quantification that only require point-wise solutions in parameter space.

Wiener [32] initially proposed a series expansion in terms of Hermite polynomials, which are orthogonal with respect to standard Gaussian random variables. Cameron and Martin [8] proved that the Fourier-Hermite series of any second-order random process $X(\theta) \in L_P^2(\Theta)$ converge in the $L_P^2(\Theta)$ sense. Extensive work has been done to generalize this approach to expansions in terms of non-Gaussian random variables, resulting in methods referred to as Wiener-Askey polynomial chaos, or generalized polynomial chaos [12, 20, 29, 31, 33].

Let us recall the notation for a probability space $\{\Theta, \Sigma, P\}$ and consider a finite set of independent random variables $\boldsymbol{\xi} = \{\xi_i(\theta)\}_{i=1}^n$. Assume that $\boldsymbol{\xi}$ maps the space of random events, Θ, into the n-dimensional space Ω, $\boldsymbol{\xi} : \Theta \to \Omega$ with a known probability distribution $\rho(\boldsymbol{\xi})$ on Ω. Let $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ be a multi-index, and $\Psi_\alpha(\boldsymbol{\xi}(\theta))$ (multivariate) polynomials

$$\Psi_\alpha(\boldsymbol{\xi}(\theta)) = \psi_{\alpha_1}(\xi_1) \cdots \psi_{\alpha_n}(\xi_n), \tag{4.1}$$

where the univariate polynomial $\psi_{\alpha_i}(\xi_i)$ is the orthonormal polynomial of degree α_i, orthogonal with respect to the probability distribution of ξ_i [11, 17, 18]. That is, for $\alpha, \beta \in \mathbb{N}^n$

$$\int_\Theta \Psi_\alpha(\boldsymbol{\xi}(\theta))\Psi_\beta(\boldsymbol{\xi}(\theta))P(d\theta) = \int_\Omega \Psi_\alpha(\boldsymbol{\xi})\Psi_\beta(\boldsymbol{\xi})\rho(\boldsymbol{\xi})d\boldsymbol{\xi} = \delta_{\alpha\beta}, \tag{4.2}$$

where

$$\delta_{\alpha\beta} = \begin{cases} 1, & \alpha_i = \beta_i, i = 1, \ldots, n \\ 0, & \text{otherwise.} \end{cases} \tag{4.3}$$

Generalization of the Cameron-Martin theorem guarantees convergence of the generalized polynomial chaos expansion,

$$X(\boldsymbol{\xi}(\theta)) = \sum_{\alpha \in \mathbb{N}^n} c_\alpha \Psi_\alpha(\boldsymbol{\xi}(\theta)), \tag{4.4}$$

of any second-order process X, with coefficients

$$c_\alpha = \langle X, \Psi_\alpha \rangle_\Omega := \int_\Omega X(\boldsymbol{\xi})\Psi_\alpha(\boldsymbol{\xi})\rho(\boldsymbol{\xi})d\boldsymbol{\xi} \tag{4.5}$$

for the so-called Wiener-Askey polynomials which form a complete basis in the Hilbert space on Ω [12, 33]. A more thorough investigation of the Wiener-Askey polynomials and associated weighting functions can be found in [29]. Without loss of generality, we restrict ourselves to uniform random variables. That is, we consider expansions in terms of Legendre polynomials which are orthogonal with respect to a uniform distribution on $[-1, 1]$. For uniform random variables defined over a different domain, the

Legendre polynomials can easily be shifted and rescaled to maintain orthonormality. More specically we will express the coefficients $a_i(\boldsymbol{\xi})$ in (3.2) as generalized polynomial chaos expansions,

$$a_i(\boldsymbol{\xi}) = \sum_{\alpha \in \mathbb{N}^n} a_{i,\alpha} \Psi_\alpha(\boldsymbol{\xi}),\ i = 1, \ldots, l \tag{4.6}$$

where $a_{i,\alpha} = \langle a_i, \Psi_\alpha \rangle_\Omega$. In vector notation, it reads:

$$\boldsymbol{a}(\boldsymbol{\xi}) = \sum_{\alpha \in \mathbb{N}^n} \boldsymbol{a}_\alpha \Psi_\alpha(\boldsymbol{\xi}). \tag{4.7}$$

In practice the above expansion needs to be truncated to a finite expansion order. Let $\mathcal{I}_N$ be a set of multi-indices,

$$\mathcal{I}_N = \{\alpha \in \mathbb{N}^n : \alpha_1 + \cdots + \alpha_n \leq N\}, \tag{4.8}$$

which corresponds to the tensor product of polynomials with maximum total degree N. The coefficients $\boldsymbol{a}(\boldsymbol{\xi})$ can be approximated as

$$\boldsymbol{a}(\boldsymbol{\xi}) \approx \sum_{\alpha \in \mathcal{I}_N} \boldsymbol{a}_\alpha \Psi_\alpha(\boldsymbol{\xi}). \tag{4.9}$$

Unfortunately, the definition of $\boldsymbol{a}_\alpha$ still involves the unknown process $\boldsymbol{a}(\boldsymbol{\xi})$, therefore we can not compute the coefficients exactly. The pseudospectral method approximates the coefficients using quadrature,

$$\boldsymbol{a}_\alpha = \langle \boldsymbol{a}, \Psi_\alpha \rangle_\Omega \approx \sum_{k=1}^{m} \boldsymbol{a}(\boldsymbol{\xi}^k) \Psi_\alpha(\boldsymbol{\xi}^k) w_k := \boldsymbol{a}_{\alpha,m}, \tag{4.10}$$

where $\{\boldsymbol{\xi}^k\}$ and $\{w_k\}$ are the integration points and weights, respectively. It has been shown in [10] that using a Gaussian quadrature rule yields a geometric convergence rate. It does however require a number of quadrature points $m(N, n)$ that scales exponentially with the dimension of Ω. One can partially mitigate this curse of dimensionality through the use of sparse quadrature rules without any significant change to the theoretical results that follow. In any case, the use of quadrature amounts to solving a series of *uncoupled* deterministic problems for $\boldsymbol{a}(\boldsymbol{\xi}^k)$, namely

$$\boldsymbol{B}(\boldsymbol{\xi}^k; \boldsymbol{a}(\boldsymbol{\xi}^k)) = \boldsymbol{F}(\boldsymbol{\xi}^k),\ k = 1, \ldots, m(N, n), \tag{4.11}$$

where $\boldsymbol{B}(\boldsymbol{\xi}^k, \boldsymbol{a}(\boldsymbol{\xi}^k))$ and $\boldsymbol{F}(\boldsymbol{\xi}^k)$ are as defined in (3.4) and (3.5), respectively. Note that the solution of (4.11) requires a non-linear solve for each $\boldsymbol{\xi}^k$. Finally, we are able to construct the fully discrete approximation for $u(\boldsymbol{x}, \boldsymbol{\xi})$,

$$u^{h,N}(\boldsymbol{x}, \boldsymbol{\xi}) = \sum_{j=1}^{l} \sum_{\alpha \in \mathcal{I}_N} a_{j_{\alpha,m}} \Psi_\alpha(\boldsymbol{\xi}) \phi_j^h(\boldsymbol{x}). \tag{4.12}$$

5. Error estimate for fully discretized solution. We now examine the changes to the error representation formula (3.10), needed in order to account for the fully discrete solution. First a new adjoint problem must be defined, this time based on a linearization about the fully discrete solution,

$$\text{find } \hat{z}(\cdot, \boldsymbol{\xi}) \in V_D \text{ s.t.} \quad B'_{\boldsymbol{\xi}}(u^{h,N}; v, \hat{z}) = Q(v) \quad \forall v \in V_D. \tag{5.1}$$

As a result of the change in the linearization point, the solution to the above equation represents a distinct solution from that of (3.6); note the change in notation to $\hat{z}$. This is a subtle issue that only appears in the non-linear case. The impact of the change is only in the adjoint solution itself, the derivation can proceed in exactly the same way as before. A discrete version of the equation is obtained, again on an enriched space V_D^+,

$$\text{find } \hat{z}^+(\cdot,\boldsymbol{\xi}) \in V_D^+ \text{ s.t.} \quad B'_{\boldsymbol{\xi}}(u^{h,N}; v, \hat{z}^+) = Q(v) \quad \forall v \in V_D^+. \tag{5.2}$$

The error representation formula for the fully discrete solution follows analogously to (3.10),

$$\begin{aligned} Q\Big(u(\cdot,\boldsymbol{\xi}) - u^{h,N}(\cdot,\boldsymbol{\xi})\Big) &= B'_{\boldsymbol{\xi}}(u^{h,N}; u - u^{h,N}, \hat{z}) && (5.3)\\ &= B_{\boldsymbol{\xi}}(u;\hat{z}) - B_{\boldsymbol{\xi}}(u^{h,N};\hat{z}) + \Delta_{B,N} && (5.4)\\ &= \mathcal{R}_{\boldsymbol{\xi}}(u^{h,N};\hat{z}) + \Delta_{B,N} && (5.5)\\ &= \mathcal{R}_{\boldsymbol{\xi}}(u^{h,N};\hat{z}^{+,N}) + \mathcal{R}_{\boldsymbol{\xi}}(u^{h,N};\hat{z}-\hat{z}^{+,N}) + \Delta_{B,N} && (5.6)\\ &\approx \mathcal{R}_{\boldsymbol{\xi}}(u^{h,N};\hat{z}^{+,N}), && (5.7) \end{aligned}$$

where $\hat{z}^{+,N}$ is a fully discrete approximation of the adjoint solution,

$$\hat{z}^{+,N}(\boldsymbol{x},\boldsymbol{\xi}) = \sum_{i=1}^{l} \sum_{\alpha\in\mathcal{I}_N} b_{i,\alpha,m}\Psi_\alpha(\boldsymbol{\xi})\phi_i^+(\boldsymbol{x}). \tag{5.8}$$

Here, the coefficients $b_{i,\alpha,m}$ are determined by solving an equation similar to (3.14) at the quadrature points,

$$\boldsymbol{B}^{*,N}(\boldsymbol{\xi}^k)\boldsymbol{b}(\boldsymbol{\xi}^k) = \boldsymbol{Q},\ k = 1,\ldots,m(N,n), \tag{5.9}$$

the difference being,

$$\left[\boldsymbol{B}^{*,N}(\boldsymbol{\xi})\right]_{ji} = B'_{\boldsymbol{\xi}}\left(u^{h,N};\phi_j^+,\phi_i^+\right). \tag{5.10}$$

We note that it is not necessary to obtain a higher order approximation in parameter space since the error representation is evaluated only at specific points $\boldsymbol{\xi} \in \Omega$, and not integrals on Ω. In other words, issues with orthogonality are not encountered in our case since we are not using the Galerkin method to approximate the solutions of the primal and adjoint problems. In fact, one could even use a lower order expansion of the adjoint solution and still obtain a error estimate, although its accuracy may be degraded significantly [5]. The term $\mathcal{R}_{\boldsymbol{\xi}}(u^{h,N};\hat{z}-\hat{z}^{+,N})$ in (5.6) involves the product of $e_u = u - u^{h,N}$ and $e_z = z - z^{+,N}$, and was shown in [5, 7] to be higher order if the forward and adjoint approximations converge point-wise in parameter space (see [28, 30] for conditions on this form of convergence).

Alternatively, it is possible to neglect the change from z to $\hat{z}$ using the following observation,

$$\mathcal{R}_{\boldsymbol{\xi}}(u^{h,N};\hat{z}^{+,N}) = \mathcal{R}_{\boldsymbol{\xi}}(u^{h,N};z^{+,N}) + \mathcal{R}_{\boldsymbol{\xi}}(u^{h,N};\hat{z}^{+,N} - z^{+,N}). \tag{5.11}$$

The first term on the right-hand side of (5.11) is the usual dual-weighted residual. The second term can be attributed to the discrepancy in the definition of the adjoint

problems, (3.6) and (5.2), as a result of linearizing at different solution states, u^h and $u^{h,N}$, respectively. In general, this is a higher-order linearization error if $u^{h,N}$ is a reasonable approximation for u^h. This is expected to be the case for higher order stochastic approximations. Indeed, if the surrogate is constructed using interpolation, as it is in stochastic collocation approaches, the forward solution would be the same at the collocation points and thus the adjoint solutions would agree and the linearization error could be neglected. It has even been shown that for particular quadrature rules the interpolating polynomial and pseudospectral approximation agree at the quadrature points [9].

Given $\boldsymbol{\xi}$, the evaluation of (5.7), simply requires an inner product calculation. However, assuming one wants to estimate the error at many points in parameter space, or estimate the L^2_Ω-norm of the error, it may be computationally more efficient to form a higher order spectral approximation of the error estimates,

$$\mathcal{E}(\boldsymbol{\xi}) = \sum_{\alpha \in \mathcal{I}_M} \mathcal{E}_{\alpha,m} \Psi_\alpha(\boldsymbol{\xi}), \quad M > N, \tag{5.12}$$

rather than carry out many inner product calculations. Forming the spectral expansion will only require inner product calculations at the quadrature points,

$$\mathcal{E}_{\alpha,m} = \sum_{j=1}^{m(M,n)} \mathcal{R}_{\boldsymbol{\xi}^j}\left(u^{h,N}; \hat{z}^{+,N}\right) \Psi_\alpha(\boldsymbol{\xi}^j). \tag{5.13}$$

Obviously, for a higher order expansion a greater number of quadrature points will be needed, $m(M,n) > m(N,n)$ for $M > N$. In the examples below we use $M = 2N$. For all of the computational results presented in section 8, the time required to form the error surrogate (5.12) is much smaller than the time required to solve the forward and adjoint problems.

6. Error Decomposition. The error estimate (5.7) contains finite element discretization error as well as the error due to the spectral approximation in parameter space. To efficiently control approximation error through adaptivity we seek to separate the errors due to the physical discretization and the errors due to the discretization in stochastic space, and then to use those indicators to drive adaptive refinement. As mentioned previously most methods to date [1, 5, 6, 23, 24] do not provide such an indicator. In this section we explain the decomposition of the resulting error estimate into separate error indicators for each approximation space.

In order to obtain separate error estimates for each space, consider an alternative representation of the total error. Since, we assume that u^h has been computed at specific points $\boldsymbol{\xi}$, it can be used to separate the error into contributions from each space,

$$\begin{aligned} Q\Big(u(\cdot,\boldsymbol{\xi}) - u^{h,N}(\cdot,\boldsymbol{\xi})\Big) = & \underbrace{Q\Big(u(\cdot,\boldsymbol{\xi}) - u^{h}(\cdot,\boldsymbol{\xi})\Big)}_{\text{error due to physical discretization}} \\ + & \underbrace{Q\Big(u^{h}(\cdot,\boldsymbol{\xi}) - u^{h,N}(\cdot,\boldsymbol{\xi})\Big)}_{\text{error due to stochastic discretization}} \end{aligned} \tag{6.1}$$

One could instead write the decomposition in terms of u^N; however, this semi-discrete solution is usually not computable since it requires an exact solution on the physical

domain. The first term on the right-hand side of (6.1) corresponds to that introduced in (3.17). It only involves the contributions due to the approximation in the physical domain and can be approximated using (3.19). The second term is much more difficult to estimate directly but represents the error due to discretization in the stochastic domain. Using (5.6) and (3.18) we can however obtain an expression for the error due to discretization in Ω only,

$$
\begin{aligned}
Q\Big(u^h(\cdot,\boldsymbol{\xi}) - u^{h,N}(\cdot,\boldsymbol{\xi})\Big) &= Q\Big(u(\cdot,\boldsymbol{\xi}) - u^{h,N}(\cdot,\boldsymbol{\xi})\Big) \\
&\quad - \Big[Q\Big(u(\cdot,\boldsymbol{\xi}) - u^h(\cdot,\boldsymbol{\xi})\Big)\Big] && (6.2)\\
&= \mathcal{R}_{\boldsymbol{\xi}}(u^{h,N};\hat{z}^{+,N}) + \mathcal{R}_{\boldsymbol{\xi}}(u^{h,N};\hat{z}-\hat{z}^{+,N}) + \Delta_{B,N} \\
&\quad - \Big[\mathcal{R}_{\boldsymbol{\xi}}(u^h;z^+) + \mathcal{R}_{\boldsymbol{\xi}}(u^h;z-z^+) + \Delta_B\Big] && (6.3)\\
&= \mathcal{R}_{\boldsymbol{\xi}}(u^{h,N};\hat{z}^{+,N}) - \mathcal{R}_{\boldsymbol{\xi}}(u^h;z^+) \\
&\quad + \mathcal{R}_{\boldsymbol{\xi}}(u^{h,N};\hat{z}-\hat{z}^{+,N}) + \Delta_{B,N} \\
&\quad - \mathcal{R}_{\boldsymbol{\xi}}(u^h;z-z^+) - \Delta_B && (6.4)\\
&\approx \mathcal{R}_{\boldsymbol{\xi}}(u^{h,N};\hat{z}^{+,N}) + \mathcal{R}_{\boldsymbol{\xi}}(u^h;z^+) && (6.5)
\end{aligned}
$$

Equation (6.5) provides a computable error estimate for the error due to discretization in parameter space while (6.4) reflects the accuracy of such an estimate. It is interesting to see here how the two adjoint solution z and $\hat{z}$ are used to compute the contributions from each approximation space.

An expansion for the total error estimate has already been defined in (5.12); for computational efficiency a pseudospectral expansion can be obtained for error due to discretization in physical space as well,

$$
\mathcal{E}^D(\boldsymbol{\xi}) = \sum_{\alpha\in\mathcal{I}_N} \mathcal{E}^D_{\alpha,m}\Psi_\alpha(\boldsymbol{\xi}), \tag{6.6}
$$

where

$$
\mathcal{E}^D_{\alpha,m} = \sum_{j=1}^{m} \mathcal{R}_{\boldsymbol{\xi}^j}\left(u^h;z^+\right)\Psi_\alpha(\boldsymbol{\xi}^j). \tag{6.7}
$$

If the order of this expansion is chosen to be N then, the solutions $\{u^h\}_{k=1}^m$ have already been computed for the construction of $u^{h,N}$. If we further assume that the adjoint expansion is also order N, then $\mathcal{E}^D$ is easily constructed using the standard deterministic adjoint-based error estimates evaluated at each of the quadrature points. With the representations $\mathcal{E}$, and $\mathcal{E}^D$, the error due to discretization in parameter space can be easily computed from (6.5),

$$
\mathcal{E}^\Omega(\boldsymbol{\xi}) = \mathcal{E}(\boldsymbol{\xi}) - \mathcal{E}^D(\boldsymbol{\xi}). \tag{6.8}
$$

We note that $\mathcal{E}^\Omega$ is only as accurate to the minimum of the expansion orders of $\mathcal{E}$ and $\mathcal{E}^D$. However, the total error estimate, $\mathcal{E}$, is higher-order and is used initially as an error indicator to determine any refinement is required. If this is the case, then the relative magnitude of $\mathcal{E}^D$ and $\mathcal{E}^\Omega$ is more important than the overall accuracy of either component since they are only used to determine which approximation space should be refined.

7. Refinement strategies. In this section we outline a few adaptive strategies that will be used for the numerical results that follow. They have been simplified for presentation purposes; more sophisticated strategies can be devised to further take advantage of the error decomposition. Algorithm 1 represents the simplest form of an adaptive strategy that takes advantage of the error estimators for each approximation space separately. This algorithm is used to adapt the solution of a simple Navier-Stokes flow problem in Section 8.1.

Algorithm 1: General adaptive refinement in both D and Ω

while $\|\mathcal{E}\|_{L^2_\Omega} > TOL$ **do**
- Compute $\{\boldsymbol{a}(\xi^k)\}_{k=1}^{m(N,n)}$, $\{\boldsymbol{b}(\xi^k)\}_{k=1}^{m(N,n)}$ according to (4.11) and (5.9) respectively;
- Compute $\{\mathcal{R}_{\boldsymbol{\xi}^k}(u^h; z^+)\}_{k=1}^{m(N,n)}$;
- Construct $u^{h,N}$, $z^{+,N}$, and $\mathcal{E}^D$;
- Compute $\{\mathcal{R}_{\boldsymbol{\xi}^k}(u^{h,N}; z^{+,N}\}_{k=1}^{m(M,n)}$ and construct $\mathcal{E}$;
- Compute $\|\mathcal{E}\|_{L^2_\Omega}$, $\left\|\mathcal{E}^D\right\|_{L^2_\Omega}$ and $\left\|\mathcal{E}^\Omega\right\|_{L^2_\Omega} = \left\|\mathcal{E} - \mathcal{E}^D\right\|_{L^2_\Omega}$;
- **if** $\left\|\mathcal{E}^D\right\|_{L^2_\Omega} > \left\|\mathcal{E}^\Omega\right\|_{L^2_\Omega}$ **then**
  - Refine finite element mesh ;
- **else**
  - Refine discretization on Ω $(N \leftarrow N+1)$;
- **end**

end

Next, in Algorithm 2, we present a procedure to perform adaptive h refinement in Ω. As in the example in Section 8.2, the physical discretization is assumed fixed. This procedure could easily be used in combination with Algorithm 1 to accommodate refinement in D as well. Additionally, one could devise an adaptive procedure that considers h and p refinement in Ω.

Partitioning of parameter space will be driven by the magnitude of the stochastic discretization error over each element. Let $\Omega = \cup_i \Omega_i$ represent the current partition of the domain, and $\mathcal{E}^{\Omega_i}$ be the local spectral approximation

Algorithm 2: Adaptive h refinement in Ω

while $\|\mathcal{E}\|_{L^2_\Omega} > TOL$ **do**
- Construct $\mathcal{E}^{\Omega_i}$;
- Compute $\left\|\mathcal{E}^{\Omega_i}\right\|_{L^2_{\Omega_i}}$;
- **if** $\left\|\mathcal{E}^{\Omega_i}\right\|_{L^2_{\Omega_i}} > TOL_2$ **then**
  - Split Ω_i into 2^n new elements by bisection in each stochastic direction;
- **end**

end

Since the h-adaptive procedure described in Algorithm 2 is really only practical on low-dimensional stochastic spaces, the final adaptive strategy targets the curse of dimensionality. As in Algorithm 2 we focus only on adaptivity in Ω. First a generalization of the definition of a set of multi-indices is needed to allow for anisotropy.

Let

$$\mathcal{I}_{\mathcal{N}} = \{\alpha \in \mathbb{N}^n : \alpha_i \leq N_i, i = 1, \ldots, n\}, \tag{7.1}$$

where $\mathcal{N} = (N_1, \ldots, N_n)$ represents the maximum polynomial order in each direction and higher-order expansions of the error will now be obtained using $\mathcal{M} = \mathcal{N} + \mathbf{1} = \{N_1 + 1, \ldots, N_n + 1\}$.

Since the expansion of the error estimate is represented by a higher-order expansion, refinement can be driven by the higher-order terms. Specifically the relative magnitude of the higher-order coefficients, those that are not included in the expansion of the solution, will be used to determine in which direction the polynomial order of approximation should be increased. Adding polynomials associated with the largest contributions to the error accounts for the most important directions first, hence reducing the overall approximation error in the most efficient way. Section 8.3 demonstrates the use of the adaptive algorithm presented in Algorithm 3.

Algorithm 3: Anisotropic p refinement in Ω

```
while ‖E‖_{L²_Ω} > TOL do
    Construct E(ξ) = Σ_{α∈I_M} E_{α,m} Ψ_α(ξ) ;
    Let α* ∈ I_M \ I_N such that, E_{α*,m} = max_{α∈I_M\I_N} E_{α,m}. ;
    for i = 1 → n do
        if α*_i > N_i then
            Increase polynomial order of approximation in component i,
            N_i ← N_i + 1 ;
        end
    end
end
```

8. Numerical Results. We present in this section several numerical examples in order to demonstrate the performance of the algorithms described in the previous section for adaptive refinement.

8.1. Navier-stokes flow past a cylinder. We begin with the solution of steady state Navier-stokes flow around a cylinder in a channel. In this example, the inlet velocity profile and fluid viscosity are treated as uncertain parameters and the goal is to efficiently compute an accurate solution of the x-velocity behind the cylinder. We will rely on the error estimates detailed above to guide adaptivity, choosing to refine physical space or parameter space based on the error estimates, $\mathcal{E}^D$ and $\mathcal{E}^\Omega$. Our decomposition provides error indicators to prevent over-refining in either space when the total approximation error is predominantly due to the discretization in one space.

We define the non-linear boundary-value problem

$$\mathcal{A}(\boldsymbol{x}, \boldsymbol{\xi}; \boldsymbol{u}) = \begin{cases} -\nu(\boldsymbol{\xi})\Delta\boldsymbol{u} + \boldsymbol{u} \cdot \nabla\boldsymbol{u} + \nabla p &= \mathbf{0}, \\ \nabla \cdot \boldsymbol{u} &= 0, \end{cases} \quad \text{in } D, \tag{8.1}$$

with boundary conditions

$$\boldsymbol{u}(\boldsymbol{x}) = \boldsymbol{h}(\boldsymbol{x}, \boldsymbol{\xi}) \qquad \boldsymbol{x} \in \Gamma_i, \tag{8.2}$$

$$\boldsymbol{u} = 0, \qquad \boldsymbol{x} \in \Gamma_{\mathrm{w}} \cup \Gamma_{\mathrm{cyl}}, \tag{8.3}$$

$$(\nu(\boldsymbol{\xi})\nabla \boldsymbol{u} - p\boldsymbol{I}) \cdot \boldsymbol{n} = 0, \qquad \boldsymbol{x} \in \Gamma_{\mathrm{o}}, \tag{8.4}$$

where the domain D and boundary ∂D are shown in Figure 8.1 along with the initial mesh in physical space. Suppose that the kinematic viscosity and inlet velocity field are parameterized by uniformly distributed random variables such that, $\nu(\boldsymbol{\xi}) = \nu(\xi_1) = \xi_1 \sim U(0.01, 0.1)$ and

$$\boldsymbol{h}(\boldsymbol{x}, \boldsymbol{\xi}) = \boldsymbol{h}(\boldsymbol{x}, \xi_2) = \xi_2 \begin{bmatrix} \frac{3}{32}(4 - x_2^2) \\ 0 \end{bmatrix}, \tag{8.5}$$

where $\xi_2 \sim U(1, 3)$. The parameter distributions combined with the geometry of the problem correspond to the range of Reynolds number, $[1.25, 37.5]$. For these values of Re, the flow remains laminar the steady state solution is stable.

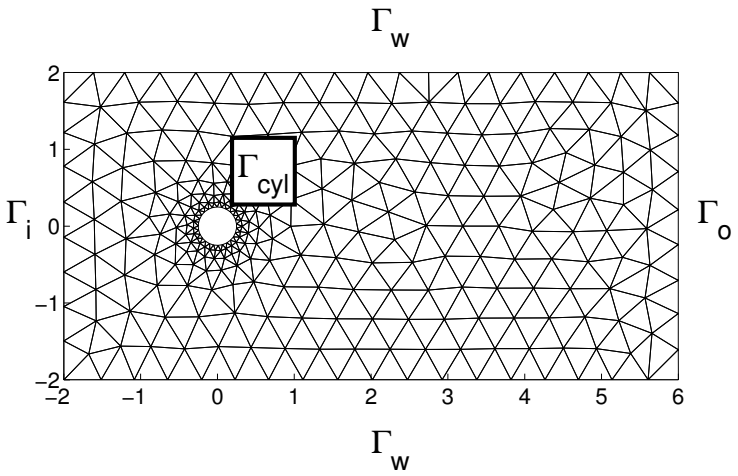


FIG. 8.1. *Physical domain*

The quantity of interest is chosen to be the horizontal component of the velocity u_x evaluated at a point directly behind the cylinder, $\boldsymbol{x}_0 = (1, 0)^T$,

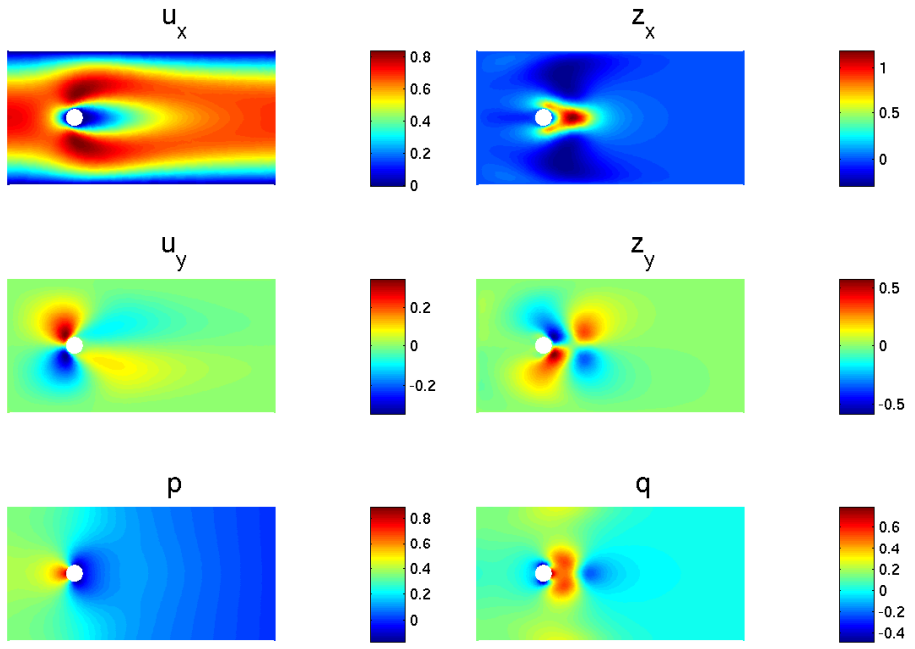
$$Q\left(\boldsymbol{u}(\cdot, \boldsymbol{\xi})\right) = u_x\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \boldsymbol{\xi}\right) = \int_D \boldsymbol{u}(\boldsymbol{x}, \boldsymbol{\xi}) \cdot \boldsymbol{e}_x \; \delta(\boldsymbol{x}, \boldsymbol{x}_0) \; d\boldsymbol{x}. \tag{8.6}$$

However, the above functional is unbounded in V_D. We thus choose the following q,

$$q(\boldsymbol{x}) = \begin{bmatrix} \frac{10}{\pi} \exp\left(-10(x_{0,1} - x_1)^2 - 10(x_{0,2} - x_2)^2\right) \\ 0 \end{bmatrix}, \tag{8.7}$$

so that

$$Q\left(\boldsymbol{u}(\cdot, \boldsymbol{\xi})\right) \approx \int_D q(\boldsymbol{x}) \; u(\boldsymbol{x}, \boldsymbol{\xi}) \; dx, \tag{8.8}$$



(a) Average solution (b) Average adjoint

FIG. 8.2. *Average solution fields*

the modified quantity of interest, is a bounded functional. Figure 8.2 shows the average forward and adjoint solution fields over the range of parameter values for the initial mesh.

Beginning with the physical mesh in figure 8.1 and polynomial order $N = 1$, four simulation studies are performed. First we use the proposed error decomposition and Algorithm 1 to adaptively refine both physical and parameter spaces. Next, we perform the refinement in each space individually; uniform h-refinement in D for one study and successively increasing polynomial order of the spectral expansion for the other. Finally, we simultaneously adapt both spaces, performing uniform h-refinement on D *and* increasing the polynomial order of the stochastic expansion at each adaptive step. Figure 8.3 shows the convergence of the total error estimate $\|\mathcal{E}\|_{L^2_\Omega}$ versus the total number of degrees of freedom used for each approach.

As one might expect, improving the approximation in physical space leads to little or no reduction in the overall error. Clearly the most significant source of error is the discretization in parameter space. Similarly, only refining the approximation in parameter space leads to a plateau in the error estimate after six or seven iterations. More precisely, an initial reduction is seen as a result of improving the approximation in parameter space, but once the error $\left\|\mathcal{E}^\Omega\right\|_{L^2_\Omega}$ is reduced below the level of the physical discretization error $\left\|\mathcal{E}^D\right\|_{L^2_\Omega}$, no further reduction in the total error can be expected.

In contrast, the adaptive procedure we propose continues to achieve a significant reduction in the total approximation error with each refinement. This is because the greater source is identified and the appropriate approximation is refined accordingly.

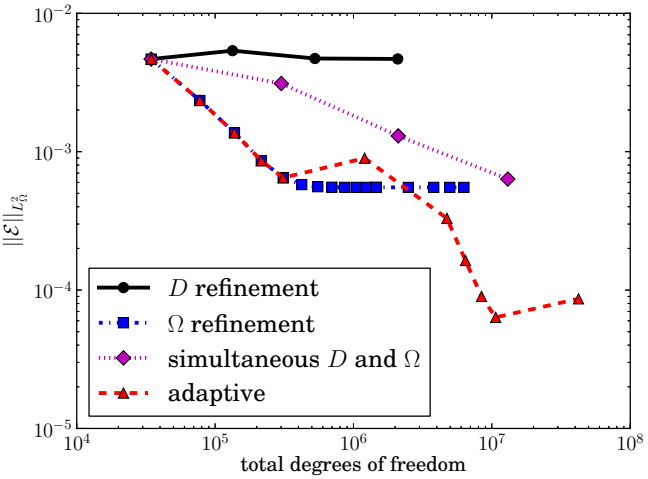


FIG. 8.3. *Convergence of error estimate, $\mathcal{E}$*

Figure 8.4 shows this more clearly by comparing the magnitude of $\mathcal{E}^D$ and $\mathcal{E}^\Omega$ for each stage of the adaptive procedure as well as the total error estimate. It is clear that the correct choices for refinement are made based on the error estimate. We observe that the total error estimate actually increases after refinement in D which indicates that the initial mesh may be too coarse for the adjoint-based error to behave as one expects asymptotically. However, the fact that the adaptive approach reduces the error beyond the point where stochastic refinement stagnates indicates that the physical adaptivity was effective.

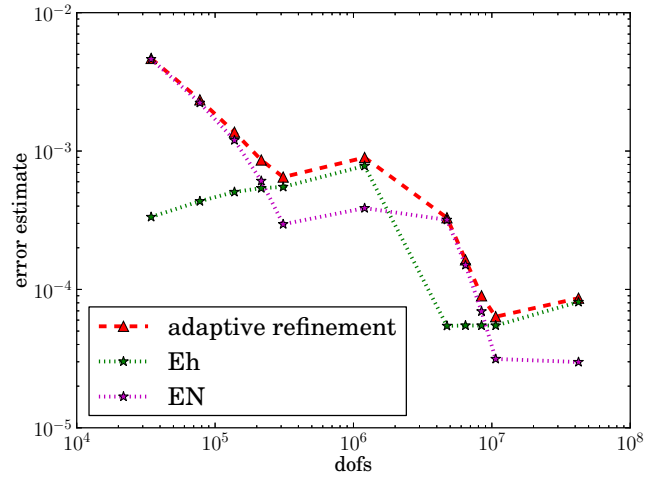


FIG. 8.4. *Convergence of error estimate, $\mathcal{E}$*

8.2. Adaptive h-refinement in parameter space. To demonstrate the advantages of h-refinement in parameter space, we consider an example where the quantity of interest depends discontinuously on the parameter. Attempting to capture responses with steep gradients or discontinuities using global spectral approximations result in the well known Gibb's phenomenon, producing spurious oscillations. As a result, many authors have proposed changing the expansion basis to reduce the effect of the oscillatory behavior. Among the bases used are piecewise polynomials and multiwavelet basis [20, 21, 24, 31]. Wan and Karniadakis [31] use a multi-element generalized polynomial chaos approach, where the decision to add elements is based on the contribution of higher-order terms to the solution variance. In addition to this type of h-refinement they adapt the order of the local expansions, i.e. p-refinement, and achieve similar convergence results as spectral hp methods used for spatial discretization. Multiwavelet approaches of Le Maître et al. [20, 21] rely on a similar type of heuristic refinement indicator to perform h-refinement in stochastic space. Convergence is shown to be superior to Monte Carlo sampling approaches in estimating the mean and variance of the solution. Mathelin and Le Maître [24] later proposed a more rigorous refinement criterion based on goal-oriented *a posteriori* error estimates. They propose adaptive refinement strategies for both physical and stochastic spaces based on these error estimates and show promising results for flow problems with uncertainties. However, to guide adaptive refinement they rely on an estimate of the physical discretization error using the expansion of the fully discretized solution, which may be corrupted by approximation errors from the stochastic space discretization. Furthermore, while the total error estimate is in terms of the quantity of interest, the estimate of the physical discretization error used as an indicator for refinement is not.

In contrast, the error estimates proposed in this work are not only given with respect to quantities of interest but also indicate the contributions from each space. While non-uniform distributions would require the use of a different expansion basis, it would not alter the adaptive procedure we present. For this reason we will only consider the case of uniform random variables and some of the technicalities of using multi-element spectral approaches, such as properly weighting the local expansion variables, can be omitted. A detailed discussion of multi-element polynomial chaos expansions using non-uniform distributions can be found in [31]. To further simplify the presentation we will consider only linear polynomials in the stochastic expansions. Lower order expansions additionally limit the effects of Gibb's phenomena and are well suited to tackle problems with discontinuities.

We consider here the convection-diffusion problem,

$$\mathcal{A}(\boldsymbol{x}, \boldsymbol{\xi}; u) = -2\Delta u + \boldsymbol{b}(\boldsymbol{\xi}) \cdot \nabla u = f(\boldsymbol{x}, \boldsymbol{\xi}), \quad \forall \boldsymbol{x} \in D = (0,1)^2, \tag{8.9}$$

where uncertainty is present in the convection and source terms. The discontinuity arises from the definition of the velocity term,

$$\boldsymbol{b}(\boldsymbol{\xi}) = \begin{bmatrix} \sin\left(\frac{3\pi}{2}\xi_1\right) \\ 4\lfloor \xi_2 - \xi_1 \rfloor \end{bmatrix} \tag{8.10}$$

where

$$\lfloor \xi_2 - \xi_1 \rfloor = \begin{cases} 0 & \xi_1 \leq \xi_2 \\ -1 & \xi_1 > \xi_2 \end{cases}. \tag{8.11}$$

It is advantageous if element boundaries coincide with discontinuities of the solution. To make the response more difficult to model and to avoid partitioning directly along the discontinuity we have chosen the discontinuity to lie on the line $\xi_1 = \xi_2$, while the element boundaries are oriented along the axis of ξ_i. The source term is chosen so that the exact solution is,

$$u(\boldsymbol{x}, \boldsymbol{\xi}) = 10 \sin\left(\frac{3\pi}{2}\xi_1\right)(4\lfloor \xi_2 - \xi_1 \rfloor)(x_1 - x_1^2)(x_2 - x_2^2). \tag{8.12}$$

We choose as a quantity of interest the solution at the point $\boldsymbol{x}_0 = (0.33, 0.33)^T$, which will again be approximated using

$$q(\boldsymbol{x}) = \frac{100}{\pi} \exp\left(-100(x_1 - x_{0,1})^2 - 100(x_2 - x_{0,2})^2\right), \tag{8.13}$$

that is

$$Q(u(\cdot, \boldsymbol{\xi})) = u\left(\begin{bmatrix}0.33\\0.33\end{bmatrix}, \boldsymbol{\xi}\right) \approx\approx \int_D q(\boldsymbol{x})\ u(\boldsymbol{x}, \boldsymbol{\xi})\ d\boldsymbol{x} \tag{8.14}$$

As stated above we will assume the two parameters are uniformly distributed, with $\xi_1 \sim U(0,1)$ and $\xi_2 \sim U(0,1)$. Figure 8.5 shows the true response for the quantity of interest over parameter space. Notice that on one side of the discontinuity the response is actually zero while the other region shows a smooth non-linear dependence. As a result of this feature, and the fact that only linear polynomials are used in the expansion, local refinement will be needed away from the discontinuity to obtain an accurate solution.

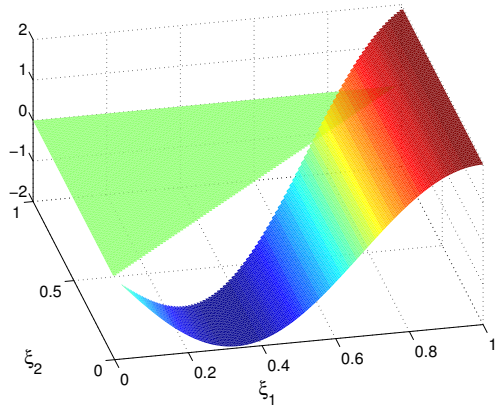


FIG. 8.5. *Exact response surface*

To start, one element will be used in parameter space and the spectral expansions will be formed with $N = 1$. In order to focus on the refinement procedure in parameter space only, the physical mesh is kept fixed in all studies. The relative error on that particular mesh is $\left\|\mathcal{E}^D\right\|_{L^2_\Omega} \sim 10^{-3}$. Three simulation studies are performed, the first being our proposed h-refinement strategy (Algorithm 2) with $\text{TOL}_2 = 0.75\ \max_j\{\left\|\mathcal{E}_{\Omega_j}\right\|_{L^2_\Omega}\}$. To evaluate the efficiency of the adaptive approach standard uniform h-refinement in parameter space is also performed. Finally, p-refinement of a global spectral basis is compared. While fairer comparison can

be made against uniform h-refinement, we include the example of p-refinement to demonstrate the shortcomings of such an approach and highlight the advantage of the multi-element approach for this particular problem. Convergence results for each algorithm are shown in Figure 8.6.

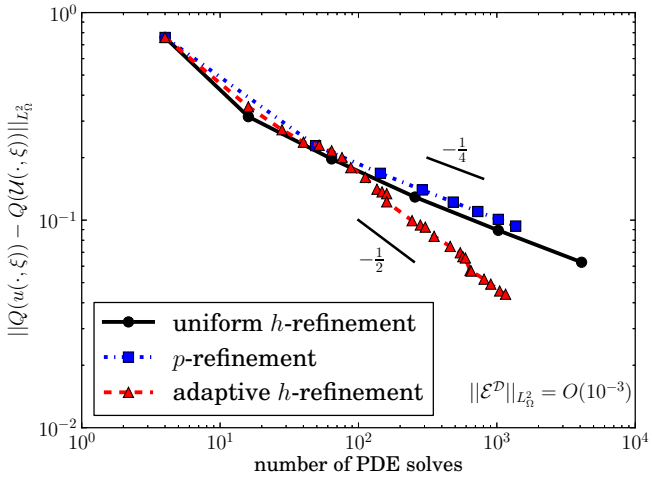


FIG. 8.6. *Convergence of true error*

We have included a comparison of the total error plotted as a function of $\boldsymbol{\xi}$ after the final refinement step using p-refinement in figure 8.7(a), and our proposed h-refinement strategy in figure 8.7(b). Notice the oscillatory behavior in the case of the global approach. Clearly the multi-element approach does a better job of identifying the discontinuity, and localizes it to smaller elements along the discontinuity. This can be seen more clearly in the final parameter space mesh, shown in figure 8.8. Recall that all elements have linear approximations; as a result the region below $\xi_1 = \xi_2$ does require some refinement to capture the nonlinear response. Uniform h-refinement in parameter space produces a similar result to 8.7(b) but requires many more degrees of freedom due to the fact that the region above $\xi_1 = \xi_2$ would also be refined.

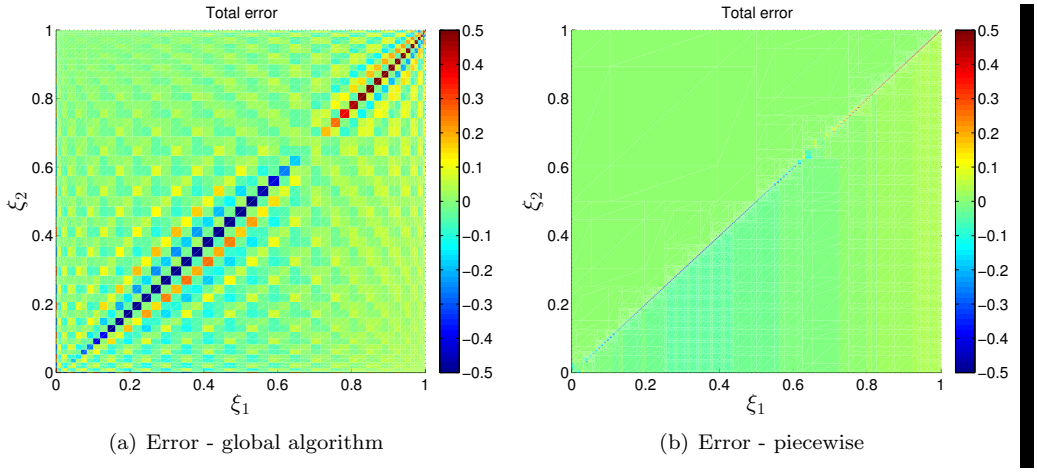
8.3. Higher parameter dimension. Finally, we demonstrate the use of the error estimate to perform anisotropic p-refinement in parameter space. To make the example more informative a problem is chosen where the influence of different directions in parameter space vary by orders of magnitude. Consider the following diffusion problem where the log of the diffusion coefficient is expanded in a Karhunen-Loève type of expansion,

$$\mathcal{A}(\boldsymbol{x}, \boldsymbol{\xi}; u) = -\nabla \cdot (K(\boldsymbol{x}, \boldsymbol{\xi}) \nabla u) = 10, \qquad \forall \boldsymbol{x} \in D = (0,1)^2, \tag{8.15}$$

with homogeneous Dirichlet boundary conditions. Let

$$K(\boldsymbol{x}, \boldsymbol{\xi}) = \exp\left[\sum_{k=1}^{10} c_k \sin(\lambda_k \pi x_1) \cos(\lambda_k \pi x_2)\ \xi_k\right], \tag{8.16}$$

where c_k, λ_k are the coefficients and periods associated with the expansion of K, respectively; Table 8.1 lists the values for c_k and λ_k. Clearly the first dimension is the



(a) Error - global algorithm (b) Error - piecewise

FIG. 8.7. *Distribution of error*

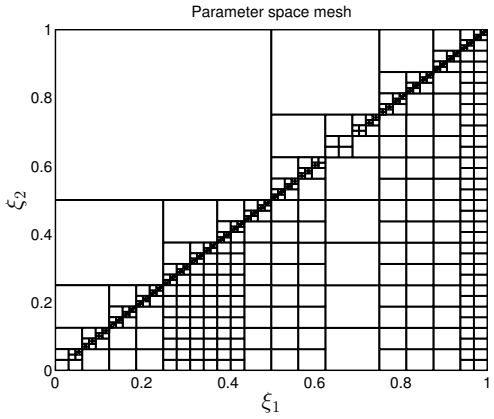


FIG. 8.8. *Parameter space mesh*

most important and will require a higher polynomial order to capture the predominant behavior of the response; influence of subsequent terms in the expansion decreases by an order of magnitude for each term. We assume that all random variables are uniformly distributed, $\xi_i \sim U(0,1)$. Again, for the quantity of interest we will take the solution at the point $\boldsymbol{x}_0 = (0.5, 0.5)^T$

$$Q(u(\cdot,\boldsymbol{\xi})) = u\left(\begin{bmatrix}0.5\\0.5\end{bmatrix},\boldsymbol{\xi}\right) \approx \langle q,u\rangle_{L^2_D}\,, \tag{8.17}$$

using

$$q(\boldsymbol{x}) = \frac{10}{\pi}\exp\left(-10(x_1-x_{0,1})^2 - 10(x_2-x_{0,2})^2\right). \tag{8.18}$$

A comparison of the adaptive refinement strategy (Algorithm 3) to isotropic refinement is shown in figure 8.9. As in the previous example we keep the physical discretization fixed at a level that makes the error $\mathcal{E}^D$ negligible. The sequence of polynomial orders corresponding to the iterations shown in figure 8.9 are provided

TABLE 8.1
Values for the expansion of a

k	c_k	λ_k
1	10	0.239
2	1	0.313
3	1e-1	0.406
4	1e-2	0.643
5	1e-3	0.740
6	1e-4	2.049
7	1e-5	3.105
8	1e-6	3.396
9	1e-7	8.731
10	1e-8	9.979

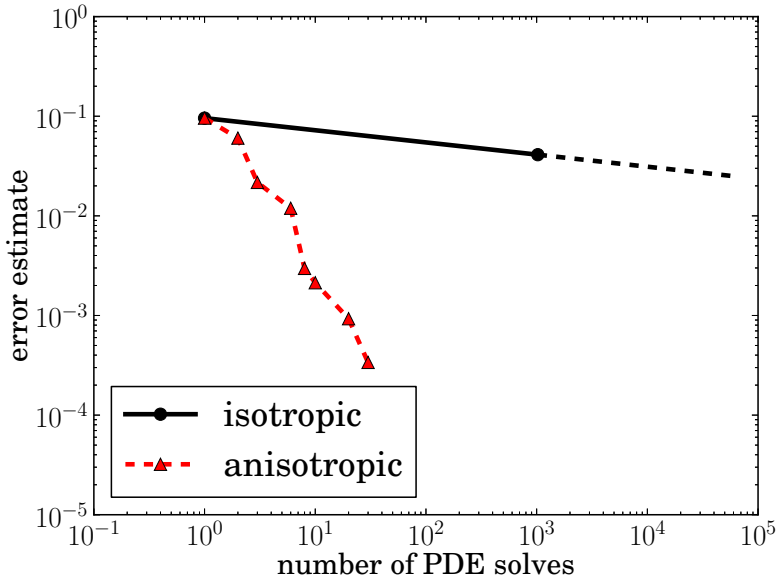


FIG. 8.9. *Convergence of error estimates for isotropic and anisotropic refinement*

in table 8.2. The proposed indicator reasonable identifies the first few terms of the expansion as the most important ones and can be used to adapt the polynomial order accordingly. When compared with isotropic p refinement one sees significant computational savings in terms of the number of solutions of the PDE. Note that the comparison can be made only up to isotropic order $N = 1$. Constructing the expansion for $N = 2$ requires the solution of $59,049$ two-dimensional finite element solutions, which becomes prohibitive due to the amount of memory needed to store all solutions. The dashed line shows the extrapolation of the error from the previous two solutions, assuming the same trend continues.

TABLE 8.2
Expansion orders obtained from adaptation

	isotropic	anisotropic				
iter	N_i	N_1	N_2	N_3	N_4	$N_5 - N_{10}$
1	0	0	0	0	0	0
2	1	1	0	0	0	0
3	(2)	1	1	0	0	0
4		2	1	0	0	0
5		3	1	0	0	0
6		4	1	0	0	0
7		4	1	1	0	0
8		4	2	1	0	0

9. Conclusion. In this work, we have developed *a posteriori* error estimates for second-order differential equations with uncertain coefficients. Error was defined in terms of specific quantities of interest represented as functionals of the solution. A novel decomposition of the error estimate into contributions from physical and stochastic approximation spaces was proposed and shown to properly identify which approximation space should be refined. This was in particular demonstrated on fluid flows around a cylinder within a channel at low to moderate Reynolds numbers. Adaptive refinement strategies were also developed that took advantage of the indicator of the error in the stochastic approximation. Adaptive mesh refinement in stochastic space using the error estimator proved to be superior to uniform refinement. Finally, anisotropic p-refinement was performed to significantly reduce the computational cost needed to obtain an accurate solution for a diffusion problem in a 10-dimensional stochastic space.

Acknowledgments. Serge Prudhomme is a participant of the KAUST SRI Center for Uncertainty Quantification in Computational Science and Engineering. Corey Bryant participated in the Visitors' Program of the KAUST SRI Center for Uncertainty Quantification in Computational Science and Engineering. This material is based upon work supported by the Department of Energy [National Nuclear Security Administration] under Award Number [DE-FC52-08NA28615].

REFERENCES

[1] R. C. Almeida and J. T. Oden. Solution verification, goal-oriented adaptive methods for stochastic advection-diffusion problems. *Comput. Methods Appl. Mech. Engrg.*, 199(37-40):2472–2486, 2010.

[2] I. Babuška, F. Nobile, and R. Tempone. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM J. Numer. Anal.*, 45(3):1005–1034 (electronic), 2007.

[3] W. Bangerth and R. Rannacher. *Adaptive Finite Element Methods for Differential Equations*. Birkhauser Verlag, 2003.

[4] R. Becker and R. Rannacher. An optimal control approach to a posteriori error estimation in finite element methods. *Acta Numer.*, 10:1–102, 2001.

[5] T. Butler, P. Constantine, and T. Wildey. A posteriori error analysis of parameterized linear systems using spectral methods. *SIAM. J. Matrix Anal. Appl.*, 33:195–209, 2012.

[6] T. Butler, C. Dawson, and T. Wildey. A posteriori error analysis of stochastic spectral methods. *SIAM J. Sci. Comput.*, 33:1267–1291, 2011.

[7] T. Butler, C. Dawson, and T. Wildey. Propagation of uncertainties using improved surrogate models. Submitted to SIAM J. Sci. Comp., 2013.

[8] R. Cameron and W. Martin. The orthogonal development of non-linear functionals in series of fourier-hermite functionals. *Annals of Mathematics*, 48:385–392, 1947.

[9] P. G. Constantine, M. S. Eldred, and E. T. Phipps. Sparse pseudospectral approximation method. *Comput. Methods Appl. Mech. Engrg.*, 229232(0):1 – 12, 2012.

[10] P. G. Constantine, D. F. Gleich, and G. Iaccarino. Spectral methods for parameterized matrix equations. *SIAM. J. Matrix Anal. Appl.*, 31:2681–2699, 2010.

[11] C. F. Dunkl and Y. Xu. *Orthogonal Polynomials in Several Variables.* Cambridge University Press, 2001.

[12] O. G. Ernst, A. Mugler, H-J. Starkloff, and E. Ullmann. On the convergence of generalized polynomial chaos expansions. *ESAIM: Mathematical Modelling and Numerical Analysis*, 46(2):317–339, 2011.

[13] D. Estep. A posteriori error bounds and global error control for approximation of ordinary differential equations. *SIAM J. Numer. Anal.*, 32(1):1–48, 1995.

[14] D. Estep and D. French. Global error control for the continuous Galerkin finite element method for ordinary differential equations. *RAIRO Modél. Math. Anal. Numér.*, 28:815–852, 1994.

[15] D. Estep, M. G. Larson, and R. D. Williams. Estimating the error of numerical solutions of systems of reaction-diffusion equations. *Mem. Amer. Math. Soc.*, 146(696):viii+109, 2000.

[16] D. Estep, Målqvist A., , and S. Tavener. Nonparametric density estimation for randomly perturbed elliptic problems I: Computational methods, a posteriori analysis, and adaptive error control. *SIAM J. Sci. Comput.*, 31:2935–2959, 2009.

[17] W. Gautschi. Orthogonal polynomials: applications and computation. *Acta Numerica*, 5:45–119, 1996.

[18] W. Gautschi. *Orthogonal Polynomials: Computation and Approximation.* Clarendon Press, Oxford, 2004.

[19] M. B. Giles and E. Suli. Adjoint methods for PDEs: a posteriori error analysis and post-processing by duality. Acta Numer., pages 105–158. Cambridge Univ. Press, Cambridge, 2002.

[20] O. Le Maître, R. Ghanem, O. Knio, and H. Najm. Multi-resolution analysis of Wiener-type propagation schemes. *J. Comput. Phys.*, 197:502–531, 2004.

[21] O. Le Maître, R. Ghanem, O. Knio, and H. Najm. Uncertainty propagation using Wiener-Haar expansions. *J. Comput. Phys.*, 197(1):28–57, 2004.

[22] X. Ma and N. Zabaras. An adaptive hierarchical sparse grid collocation algorithm for the solution of stochastic differential equations. *J. Comput. Phys.*, 228(8):3084–3113, 2009.

[23] O. Le Maître and O. Knio. *Spectral Methods for Uncertainty Quantification: With Applications to Computational Fluid Dynamics.* Springer, 2010.

[24] L. Mathelin and O. Le Maître. Dual-based a posteriori error estimate for stochastic finite element methods. *Commun. Appl. Math. and Comp. Sci.*, 2(1):83–115, 2007.

[25] F. Nobile, R. Tempone, and C. G. Webster. An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM J. Numer. Anal*, 46(5):2411–2442, 2008.

[26] F. Nobile, R. Tempone, and C. G. Webster. A sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM J. Numer. Anal*, 46(5):2309—-2345, 2008.

[27] S. Prudhomme and J. T. Oden. On goal-oriented error estimation for elliptic problems: application to the control of pointwise errors. *Comput. Methods Appl. Mech. Engrg.*, 176(14):313 – 331, 1999.

[28] R. A. Todor and C. Schwab. Convergence rates for sparse chaos approximations of elliptic problems with stochastic coefficients. *IMA J. Numer. Anal.*, 27(2):232–261, 2006.

[29] X. Wan and G. Karniadakis. Beyond Wiener-Askey expansions: Handling arbitrary pdfs. *Journal of Scientific Computing*, 27:455–464, 2006.

[30] X. Wan and G. Karniadakis. Solving elliptic problems with non-Gaussian spatially-dependent random coefficients. *Comput. Methods Appl. Mech. Engrg.*, 198(21-26):1985–1995, 2009.

[31] X. Wan and G. E. Karniadakis. An adaptive multi-element generalized polynomial chaos method for stochastic differential equations. *J. Comput. Phys.*, 209(2):617–642, 2005.

[32] N. Wiener. The homogeneous chaos. *American Journal of Mathematics*, 60:897–936, 1938.

[33] D. Xiu and G. Karniadakis. The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.*, 24:619–644, 2002.