13th AIAA/ISSMO Multidisciplinary Analysis Optimization Conference
13 - 15 September 2010, Fort Worth, Texas

AIAA 2010-9125

Towards Goal-Oriented Stochastic Design Employing Adaptive Collocation Methods

M. S. Eldred* L. P. Swiler†
Sandia National Laboratories‡, Albuquerque, NM 87185

Non-intrusive polynomial chaos expansion (NIPCE) methods based on orthogonal polynomials and stochastic collocation (SC) methods based on Lagrange interpolation polynomials are attractive techniques for uncertainty quantification (UQ) due to their strong mathematical basis and ability to produce functional representations of stochastic dependence. Both techniques reside in the collocation family, in that they sample the response metrics of interest at selected locations within the random domain without intrusion to simulation software. In this work, we explore the use of polynomial order refinement (p-refinement) approaches, both uniform and adaptive, in order to automate the assessment of UQ convergence and improve computational efficiency. In the first class of p-refinement approaches, we employ a general-purpose metric of response covariance to control the uniform and adaptive refinement processes. For the adaptive case, we detect anisotropy in the importance of the random variables as determined through variance-based decomposition and exploit this decomposition through anisotropic tensor-product and anisotropic sparse grid constructions. In the second class of p-refinement approaches, we move from anisotropic sparse grids to generalized sparse grids and employ a goal-oriented refinement process using statistical quantities of interest. Since these refinement goals can frequently involve metrics that are not analytic functions of the expansions (i.e., beyond low order response moments), we additionally explore efficient mechanisms for accurately and efficiently estimated tail probabilities from the expansions based on importance sampling.

I. Introduction

Uncertainty quantification (UQ) is the process of determining the effect of input uncertainties on response metrics of interest. These input uncertainties may be characterized as either aleatory uncertainties, which are irreducible variabilities inherent in nature, or epistemic uncertainties, which are reducible uncertainties resulting from a lack of knowledge. Since sufficient data is available for characterizing aleatory uncertainties, probabilistic methods are commonly used for computing response distribution statistics based on input probability distribution specifications. Conversely, for epistemic uncertainties, data is generally too sparse to support objective probabilistic input descriptions, leading either to subjective probabilistic descriptions (e.g., assumed priors in Bayesian analysis) or nonprobabilistic methods based on interval specifications.

One technique for the analysis of aleatory uncertainties using probabilistic methods is the polynomial chaos expansion (PCE) approach to UQ. In this work, we start from a foundation of generalized polynomial chaos using the Wiener-Askey scheme,[1] in which Hermite, Legendre, Laguerre, Jacobi, and generalized Laguerre orthogonal polynomials are used for modeling the effect of uncertain variables described by normal, uniform, exponential, beta, and gamma probability distributions, respectively. These polynomial selections are optimal for these distribution types since they are orthogonal with respect to an inner product weighting function that corresponds (identical support range, weight differs by at most a constant factor) to the probability density functions for these continuous distributions. Orthogonal polynomials can be computed for any positive weight function, so these five classical orthogonal polynomials may be augmented with

*Distinguished Member of Technical Staff, Optimization and Uncertainty Estimation Department, MS-1318, Associate Fellow AIAA.

†Principal Member of Technical Staff, Optimization and Uncertainty Estimation Department, MS-1318.

‡Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under Contract DE-AC04-94AL85000.

numerically-generated polynomials for other probability distributions (e.g., for lognormal, extreme value, and histogram distributions). When independent standard random variables are used (or computed through transformation), the variable expansions are uncoupled, allowing the polynomial orthogonality properties to be applied on a per-dimension basis. This allows one to mix and match the polynomial basis used for each variable without interference with the spectral projection scheme for the response. With usage of the optimal basis corresponding to each the random variable types, exponential convergence rates can be obtained in integrated statistical quantities of interest (i.e., mean, variance, probability), given finite input variance and adequate response smoothness.

In non-intrusive PCE, simulations are used as black boxes and the calculation of chaos expansion coefficients for response metrics of interest is based on a set of simulation response evaluations. To calculate these response PCE coefficients, two primary classes of approaches have been proposed: spectral projection and linear regression. The spectral projection approach projects the response against each basis function using inner products and employs the polynomial orthogonality properties to extract each coefficient. Each inner product involves a multidimensional integral over the support range of the weighting function, which can be evaluated numerically using sampling, tensor-product quadrature, Smolyak sparse grid, or cubature approaches. The linear regression approach (also known as point collocation or stochastic response surfaces) uses a single linear least squares solution to solve for the PCE coefficients which best match a set of response values obtained from a design of computer experiments.

Stochastic collocation (SC) is a second stochastic expansion approach that is closely related to PCE. Whereas PCE estimates coefficients for known orthogonal polynomial basis functions, SC forms Lagrange interpolation functions for known coefficients. Since the i^{th} interpolation function is 1 at collocation point i and 0 for all other collocation points, it is easy to see that the expansion coefficients are just the response values at each of the collocation points. The formation of multidimensional interpolants with this property requires the use of structured collocation point sets derived from tensor products or sparse grids. The key to the approach is performing collocation using the Gauss points and weights from the same optimal orthogonal polynomials used in PCE, which results in the same exponential convergence rates.

Once PCE or SC representations have been obtained for a response metric of interest, analytic expressions can be derived for the moments of the expansion (from integration over the aleatory/probabilistic random variables) as well as for various sensitivity measures. Local sensitivities (i.e., derivatives) and global sensitivities[2] (i.e., ANOVA, variance-based decomposition) of the response metrics may be computed with respect to the expansion variables, and local sensitivities of response moments may be computed with respect to other nonprobabilistic variables[3] (i.e., design, epistemic uncertain). The ability to analytically compute response sensitivities, moments, and moment derivatives enables bi-level, sequential, and multifidelity formulations to design under uncertainty[3] as well as second-order probability and Dempster-Shafer evidence approaches to mixed aleatory-epistemic UQ.[4–6]

In this work, we build on this algorithmic foundation to develop uniform and adaptive refinement schemes to expansion polynomial order, applied on a per-dimension basis. In other words, we pursue *dimension p-adaptive* stochastic expansions, with an eye towards future *h-adaptive* and *hp-adaptive* methods. Guidance and assessment of these refinement schemes is based initially on response covariance, which is a convenient analytic metric for stochastic expansion methods. However, to support *goal-oriented adaptivity* for use within an optimization process[a], we extend our guidance/assessment metrics to include other design possibilities, notably cumulative and complementary cumulative distribution function (CDF/CCDF) probabilities of failure for supporting reliability-based design. As probabilities of failure must be computed numerically from a stochastic expansion and tail probabilities are of foremost interest, we require techniques that can be more efficient and precise than basic random sampling; in particular, we explore importance sampling estimation performed on the stochastic expansions to compute these measures.

Section II describes the orthogonal polynomial and interpolation polynomial basis functions, Section III describes the generalized polynomial chaos and stochastic collocation methods, Section IV describes non-intrusive approaches for calculating the polynomial chaos coefficients or forming the set of stochastic collocation points, Section V describes procedures for computing statistical metrics from stochastic expansions, Section VI presents approaches for uniform and adaptive p-refinement of stochastic expansions, Section VII describes a set of computational experiments, and Section VIII provides concluding remarks.

[a]This progression can be viewed as similar to the progression from Zienkiewicz-Zhu global error estimation and adaptivity[7] towards goal-oriented error estimation and adaptivity that has been occurring in the field of deterministic finite elements.

II. Polynomial Basis

A. Orthogonal polynomials in the Askey scheme

Table 1 shows the set of orthogonal polynomials which provide an optimal basis for different continuous probability distribution types. It is derived from the family of hypergeometric polynomials known as the Askey scheme,[8] for which the Hermite polynomials originally employed by Wiener[9] are a subset. The optimality of these basis selections derives from their orthogonality with respect to weighting functions that correspond to the probability density functions (PDFs) of the continuous distributions when placed in a standard form. The density and weighting functions differ by a constant factor due to the requirement that the integral of the PDF over the support range is one.

Table 1. Linkage between standard forms of continuous probability distributions and Askey scheme of continuous hyper-geometric polynomials.

Distribution	Density function	Polynomial	Weight function	Support range
Normal	$\frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$	Hermite $He_n(x)$	$e^{\frac{-x^2}{2}}$	$[-\infty, \infty]$
Uniform	$\frac{1}{2}$	Legendre $P_n(x)$	1	$[-1, 1]$
Beta	$\frac{(1-x)^\alpha (1+x)^\beta}{2^{\alpha+\beta+1} B(\alpha+1, \beta+1)}$	Jacobi $P_n^{(\alpha,\beta)}(x)$	$(1-x)^\alpha (1+x)^\beta$	$[-1, 1]$
Exponential	e^{-x}	Laguerre $L_n(x)$	e^{-x}	$[0, \infty]$
Gamma	$\frac{x^\alpha e^{-x}}{\Gamma(\alpha+1)}$	Gen. Laguerre $L_n^{(\alpha)}(x)$	$x^\alpha e^{-x}$	$[0, \infty]$

Note that Legendre is a special case of Jacobi for $\alpha = \beta = 0$, Laguerre is a special case of generalized Laguerre for $\alpha = 0$, $\Gamma(a)$ is the Gamma function which extends the factorial function to continuous values, and $B(a, b)$ is the Beta function defined as $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$. Some care is necessary when specifying the α and β parameters for the Jacobi and generalized Laguerre polynomials since the orthogonal polynomial conventions[10] differ from the common statistical PDF conventions. The former conventions are used in Table 1.

B. Numerically generated orthogonal polynomials

If all random inputs can be described using independent normal, uniform, exponential, beta, and gamma distributions, then generalized PCE can be directly applied. If correlation or other distribution types are present, then additional techniques are required. One solution is to employ nonlinear variable transformations as described in Section III.C.1 such that an Askey basis can be applied in the transformed space. This can be effective as shown in Ref. 11, but convergence rates are typically degraded. In addition, correlation coefficients are warped by the nonlinear transformation,[12] and transformed correlation values are not always readily available. An alternative is to numerically generate the orthogonal polynomials (using Gauss-Wigert,[13] discretized Stieltjes,[14] Chebyshev,[14] or Gramm-Schmidt[15] approaches) and then compute their Gauss points and weights (using the Golub-Welsch[16] tridiagonal eigensolution). These solutions are optimal for given random variable sets having arbitrary probability density functions and preserve the exponential convergence rates for general UQ applications, but performing this process for general joint density functions with correlation is a topic of ongoing research.

C. Interpolation polynomials

Lagrange polynomials interpolate a set of points in a single dimension using the functional form

$$L_j = \prod_{\substack{k=1 \\ k \neq j}}^{m} \frac{\xi - \xi_k}{\xi_j - \xi_k} \tag{1}$$

where it is evident that L_j is 1 at $\xi = \xi_j$, is 0 for each of the points $\xi = \xi_k$, and has order $m - 1$.

For interpolation of a response function R in one dimension over m points, the expression

$$R(\xi) \cong \sum_{j=1}^{m} r(\xi_j)\, L_j(\xi) \tag{2}$$

reproduces the response values $r(\xi_j)$ at the interpolation points and smoothly interpolates between these values at other points. For interpolation in multiple dimensions, a tensor-product approach can be used wherein

$$R(\boldsymbol{\xi}) \cong \sum_{j_1=1}^{m_{i_1}} \cdots \sum_{j_n=1}^{m_{i_n}} r\left(\xi_{j_1}^{i_1}, \ldots, \xi_{j_n}^{i_n}\right) \left(L_{j_1}^{i_1} \otimes \cdots \otimes L_{j_n}^{i_n}\right) \tag{3}$$

where $\boldsymbol{i} = (m_1, m_2, \cdots, m_n)$ are the number of nodes used in the n-dimensional interpolation and ξ_j^i is the j-th point in the i-th direction. As will be seen later (Section IV.A.2), interpolation on sparse grids involves a summation of these tensor products with varying $\boldsymbol{i}$ levels.

III. Stochastic Expansion Methods

A. Generalized Polynomial Chaos

The set of polynomials from Sections II.A and II.B are used as an orthogonal basis to approximate the functional form between the stochastic response output and each of its random inputs. The chaos expansion for a response R takes the form

$$R = a_0 B_0 + \sum_{i_1=1}^{\infty} a_{i_1} B_1(\xi_{i_1}) + \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{i_1} a_{i_1 i_2} B_2(\xi_{i_1}, \xi_{i_2}) + \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{i_1} \sum_{i_3=1}^{i_2} a_{i_1 i_2 i_3} B_3(\xi_{i_1}, \xi_{i_2}, \xi_{i_3}) + \ldots \tag{4}$$

where the random vector dimension is unbounded and each additional set of nested summations indicates an additional order of polynomials in the expansion. This expression can be simplified by replacing the order-based indexing with a term-based indexing

$$R = \sum_{j=0}^{\infty} \alpha_j \Psi_j(\boldsymbol{\xi}) \tag{5}$$

where there is a one-to-one correspondence between $a_{i_1 i_2 \ldots i_n}$ and α_j and between $B_n(\xi_{i_1}, \xi_{i_2}, ..., \xi_{i_n})$ and $\Psi_j(\boldsymbol{\xi})$. Each of the $\Psi_j(\boldsymbol{\xi})$ are multivariate polynomials which involve products of the one-dimensional polynomials. For example, a multivariate Hermite polynomial $B(\boldsymbol{\xi})$ of order n is defined from

$$B_n(\xi_{i_1}, ..., \xi_{i_n}) = e^{\frac{1}{2}\boldsymbol{\xi}^T \boldsymbol{\xi}} (-1)^n \frac{\partial^n}{\partial \xi_{i_1} \ldots \partial \xi_{i_n}} e^{-\frac{1}{2}\boldsymbol{\xi}^T \boldsymbol{\xi}} \tag{6}$$

which can be shown to be a product of one-dimensional Hermite polynomials involving a multi-index m_i^j:

$$B_n(\xi_{i_1}, ..., \xi_{i_n}) = \Psi_j(\boldsymbol{\xi}) = \prod_{i=1}^{n} \psi_{m_i^j}(\xi_i) \tag{7}$$

In the case of a mixed basis, the same multi-index is employed although the one-dimensional polynomials $\psi_{m_i^j}$ are heterogeneous.

In practice, one truncates the infinite expansion at a finite number of random variables and a finite expansion order

$$R \cong \sum_{j=0}^{P} \alpha_j \Psi_j(\boldsymbol{\xi}) \tag{8}$$

Traditionally, the polynomial chaos expansion includes a complete basis of polynomials up to a fixed total-order specification. The total number of terms N_t in an expansion of total order p involving n random variables is given by

$$N_t \;=\; 1 + P \;=\; 1 + \sum_{s=1}^{p} \frac{1}{s!} \prod_{r=0}^{s-1} (n + r) \;=\; \frac{(n+p)!}{n!p!} \tag{9}$$

This traditional approach will be referred to as a "total-order expansion."

An important alternative approach is to employ a "tensor-product expansion," in which polynomial order bounds are applied on a per-dimension basis (no total-order bound is enforced) and all combinations of the one-dimensional polynomials are included. In this case, the total number of terms N_t is

$$N_t \;=\; 1 + P \;=\; \prod_{i=1}^{n} (p_i + 1) \tag{10}$$

where p_i is the polynomial order bound for the i-th dimension.

It is apparent from Eq. 10 that the tensor-product expansion readily supports anisotropy in polynomial order for each dimension, since the polynomial order bounds for each dimension can be specified independently. It is also feasible to support anisotropy with total-order expansions, although this involves pruning polynomials that satisfy the total-order bound (potentially defined from the maximum of the per-dimension bounds) but which violate individual per-dimension bounds. In this case, the number of pruned polynomials must be subtracted from Eq. 9.

B. Stochastic Collocation

The SC expansion is formed as a sum of a set of multidimensional Lagrange interpolation polynomials, one polynomial per collocation point. Since these polynomials have the feature of being equal to 1 at their particular collocation point and 0 at all other points, the coefficients of the expansion are just the response values at each of the collocation points. This can be written as:

$$R \cong \sum_{j=1}^{N_p} r_j \boldsymbol{L}_j(\boldsymbol{\xi}) \tag{11}$$

where the set of N_p collocation points involves a structured multidimensional grid (a tensor-product grid as in Eq. 3 or a Smolyak sparse grid). There is no need for tailoring of the expansion form as there is for PCE (i.e., to synchronize the expansion polynomials with the set of integrable monomials) since the polynomials that appear in the expansion are determined by the Lagrange construction (Eq. 1). That is, any tailoring or refinement of the expansion occurs through the selection of points in the interpolation grid and the polynomial orders of the basis adapt automatically.

C. Transformations to independent standard variables

Polynomial chaos and stochastic collocation are expanded using polynomials that are functions of independent standard random variables $\boldsymbol{\xi}$. Thus, a key component of either approach is performing a transformation of variables from the original random variables $\boldsymbol{x}$ to independent standard random variables $\boldsymbol{\xi}$ and then applying the stochastic expansion in the transformed space. This notion of independent standard space is extended over the notion of "u-space" used in reliability methods[17, 18] in that in includes not just independent standard normals, but also independent standardized uniforms, exponentials, betas and gammas. For problems directly involving independent input distributions of these five types, conversion to standard form involves a simple linear scaling transformation (to the form of the density functions in Table 1) and then the corresponding chaos/collocation points can be employed. For other independent distributions, one has a choice of two different approaches:

1. Numerically generate an optimal polynomial basis for each independent distribution as described in Section II.B.
2. Perform a nonlinear variable transformation from a given input distribution to the most similar Askey basis and employ the Askey orthogonal polynomials and associated Gauss points/weights. For example, lognormal distributions might employ a Hermite basis in a transformed standard normal space and loguniform, triangular, and histogram distributions might employ a Legendre basis in a transformed standard uniform space.

For correlated non-normal distributions, a third approach is currently the only rigorous option due to the independence of decorrelated Gaussians:

3. Perform a nonlinear variable transformation from all given input distributions to uncorrelated standard normal distributions and employ Hermite bases and associated Gauss points/weights. This involves the Nataf transformation, described in Section III.C.1 below.

1. Nataf transformation

The transformation from correlated non-normal distributions to uncorrelated standard normal distributions is denoted as $\boldsymbol{\xi} = T(\mathbf{x})$ with the reverse transformation denoted as $\mathbf{x} = T^{-1}(\boldsymbol{\xi})$. These transformations are nonlinear in general, and possible approaches include the Rosenblatt,[19] Nataf,[12] and Box-Cox[20] transformations. The results in this paper employ the Nataf transformation, which is suitable for the common case when marginal distributions and a correlation matrix are provided, but full joint distributions are not known[b]. The Nataf transformation occurs in the following two steps. To transform between the original correlated x-space variables and correlated standard normals ("z-space"), a CDF matching condition is applied for each of the marginal distributions:

$$\Phi(z_i) = F(x_i) \tag{12}$$

where $\Phi()$ is the standard normal cumulative distribution function and $F()$ is the cumulative distribution function of the original probability distribution. Then, to transform between correlated z-space variables and uncorrelated ξ-space variables, the Cholesky factor $\mathbf{L}$ of a modified correlation matrix is used:

$$\mathbf{z} = \mathbf{L}\boldsymbol{\xi} \tag{13}$$

where the original correlation matrix for non-normals in x-space has been modified to represent the corresponding "warped" correlation in z-space.[12]

IV. Non-intrusive methods for expansion formation

The major practical difference between PCE and SC is that, in PCE, one must estimate the coefficients for known basis functions, whereas in SC, one must form the interpolants for known coefficients. PCE estimates its coefficients using any of the following approaches: random sampling, tensor-product quadrature, Smolyak sparse grids, cubature, or linear regression. In SC, the multidimensional interpolants need to be formed over structured data sets, such as point sets from quadrature or sparse grids; approaches based on random sampling may not be used. Refer to Ref. 5 for additional information on random sampling and linear regression approaches; below we focus our attention on our preferred spectral projection approaches employing tensor-product quadrature and Smolyak sparse grids.

A. Spectral projection

The spectral projection approach projects the response against each basis function using inner products and employs the polynomial orthogonality properties to extract each coefficient. Similar to a Galerkin projection, the residual error from the approximation is rendered orthogonal to the selected basis. From Eq. 8, it is evident that

$$\alpha_j \;=\; \frac{\langle R, \Psi_j \rangle}{\langle \Psi_j^2 \rangle} \;=\; \frac{1}{\langle \Psi_j^2 \rangle} \int_\Omega R\, \Psi_j \, \varrho(\boldsymbol{\xi})\, d\boldsymbol{\xi}, \tag{14}$$

where each inner product involves a multidimensional integral over the support range of the weighting function. In particular, $\Omega = \Omega_1 \otimes \cdots \otimes \Omega_n$, with possibly unbounded intervals $\Omega_j \subset \mathbb{R}$ and the tensor product form $\varrho(\boldsymbol{\xi}) = \prod_{i=1}^{n} \varrho_i(\xi_i)$ of the joint probability density (weight) function. The denominator in Eq. 14 is the norm squared of the multivariate orthogonal polynomial, which can be computed analytically using the product of univariate norms squared

$$\langle \Psi_j^2 \rangle \;=\; \prod_{i=1}^{n} \langle \psi_{m_i^j}^2 \rangle \tag{15}$$

where the univariate inner products have simple closed form expressions for each polynomial in the Askey scheme[10] and are readily computed as part of the numerically-generated solution procedures described in Section II.B. Thus, the primary computational effort resides in evaluating the numerator, which is evaluated numerically using sampling, quadrature, cubature, or sparse grid approaches (and this numerical approximation leads to use of the term "pseudo-spectral" by some investigators).

[b] If joint distributions are known, then the Rosenblatt transformation is preferred.

1. Tensor product quadrature

In quadrature-based approaches, the simplest general technique for approximating multidimensional integrals, as in Eq. 14, is to employ a tensor product of one-dimensional quadrature rules. Since there is little benefit to the use of nested quadrature rules in the tensor-product case, we choose Gaussian abscissas, i.e. the zeros of polynomials that are orthogonal with respect to a density function weighting, e.g. Gauss-Hermite, Gauss-Legendre, Gauss-Laguerre, generalized Gauss-Laguerre, Gauss-Jacobi, or numerically-generated Gauss rules.

We first introduce an index $i \in \mathbb{N}_+$, $i \geq 1$. Then, for each value of i, let $\{\xi_1^i, \ldots, \xi_{m_i}^i\} \subset \Omega_i$ be a sequence of abscissas for quadrature on Ω_i. For $f \in C^0(\Omega_i)$ and $n = 1$ we introduce a sequence of one-dimensional quadrature operators

$$\mathscr{U}^i(f)(\xi) = \sum_{j=1}^{m_i} f(\xi_j^i)\, w_j^i, \tag{16}$$

with $m_i \in \mathbb{N}$ given. When utilizing Gaussian quadrature, Eq. 16 integrates exactly all polynomials of degree less than or equal to $2m_i - 1$, for each $i = 1, \ldots, n$. Given an expansion order p, the highest order coefficient evaluations (Eq. 14) can be assumed to involve integrands of at least polynomial order $2p$ (Ψ of order p and R modeled to order p) in each dimension such that a minimal Gaussian quadrature order of $p + 1$ will be required to obtain good accuracy in these coefficients.

Now, in the multivariate case $n > 1$, for each $f \in C^0(\Omega)$ and the multi-index $\mathbf{i} = (i_1, \ldots, i_n) \in \mathbb{N}_+^n$ we define the full tensor product quadrature formulas

$$\mathcal{Q}_{\mathbf{i}}^n f(\xi) = \left(\mathscr{U}^{i_1} \otimes \cdots \otimes \mathscr{U}^{i_n}\right)(f)(\boldsymbol{\xi}) = \sum_{j_1=1}^{m_{i_1}} \cdots \sum_{j_n=1}^{m_{i_n}} f\left(\xi_{j_1}^{i_1}, \ldots, \xi_{j_n}^{i_n}\right) \left(w_{j_1}^{i_1} \otimes \cdots \otimes w_{j_n}^{i_n}\right). \tag{17}$$

Clearly, the above product needs $\prod_{j=1}^n m_{i_j}$ function evaluations. Therefore, when the number of input random variables is small, full tensor-product quadrature is a very effective numerical tool. On the other hand, approximations based on tensor-product grids suffer from the *curse of dimensionality* since the number of collocation points in a tensor grid grows exponentially fast in the number of input random variables. For example, if Eq. 17 employs the same order for all random dimensions, $m_{i_j} = m$, then Eq. 17 requires m^n function evaluations.

In Ref. 21, it is demonstrated that close synchronization of expansion form with the monomial resolution of a particular numerical integration technique can result in significant performance improvements. In particular, the traditional approach of exploying a total-order PCE neglects a significant portion of the monomial coverage for a tensor-product quadrature approach, and one should rather employ a tensor-product PCE (Eq. 10) to provide improved synchronization and more effective usage of the Gauss point evaluations. When the quadrature points are standard Gauss rules (i.e., no Clenshaw-Curtis or Gauss-Patterson), it has been shown that PCE and SC result in identical polynomial forms, completely eliminating a performance gap that exists between total-order PCE and SC.

2. Smolyak sparse grids

If the number of random variables is moderately large, one should rather consider sparse tensor product spaces as first proposed by Smolyak[22] and further investigated by Refs. 23–27 that reduce dramatically the number of collocation points, while preserving a high level of accuracy.

Here we follow the notation in Ref. 27 to describe the Smolyak *isotropic* formulas $\mathscr{A}(\mathrm{w}, n)$, where w is a level that is independent of dimension[c]. The Smolyak formulas are just linear combinations of the product formulas in Eq. 17 with the following key property: only products with a relatively small number of points are used. With $\mathscr{U}^0 = 0$ and for $i \geq 1$ define

$$\Delta^i = \mathscr{U}^i - \mathscr{U}^{i-1}. \tag{18}$$

and we set $|\mathbf{i}| = i_1 + \cdots + i_n$. Then the isotropic Smolyak quadrature formula is given by

$$\mathscr{A}(\mathrm{w}, n) = \sum_{|\mathbf{i}| \leq \mathrm{w}+n} \left(\Delta^{i_1} \otimes \cdots \otimes \Delta^{i_n}\right). \tag{19}$$

[c]Other common formulations use a level q where $q \geq n$. We use $\mathrm{w} = q - n$, where $\mathrm{w} \geq 0$ for all n.

Equivalently, formula Eq. 19 can be written as[28]

$$\mathscr{A}(\mathrm{w}, n) = \sum_{\mathrm{w}+1 \le |\mathbf{i}| \le \mathrm{w}+n} (-1)^{\mathrm{w}+n-|\mathbf{i}|} \binom{n-1}{\mathrm{w}+n-|\mathbf{i}|} \cdot \left(\mathscr{U}^{i_1} \otimes \cdots \otimes \mathscr{U}^{i_n}\right). \tag{20}$$

For each index set $\mathbf{i}$ of levels, linear or nonlinear growth rules are used to define the corresponding one-dimensional quadrature orders. The following growth rules are employed for indices $i \geq 1$:

$$\mathrm{Clenshaw-Curtis}: \quad m = \begin{cases} 1 & i = 1 \\ 2^{i-1} + 1 & i > 1 \end{cases} \tag{21}$$

$$\mathrm{Gauss-Patterson}: \quad m = 2^i - 1 \tag{22}$$

$$\mathrm{Gaussian}: \quad m = 2i - 1 \tag{23}$$

where the nonlinear growth rules for Clenshaw-Curtis and Gauss-Patterson take full advantage of the point nesting in these rules, and the linear growth rules for Gaussian quadrature take advantage of, at most, "weak" nesting (e.g., reuse of the center point).

Examples of isotropic sparse grids, constructed from the fully nested Clenshaw-Curtis abscissas and the weakly-nested Gauss-Legendre abscissas are shown in Figure 1, where $\Omega = [-1, 1]^2$. There, we consider a two-dimensional parameter space and a maximum level $\mathrm{w} = 5$ (sparse grid $\mathscr{A}(5, 2)$). To see the reduction in function evaluations with respect to full tensor product grids, we also include a plot of the corresponding Clenshaw-Curtis isotropic full tensor grid having the same maximum number of points in each direction, namely $2^5 + 1 = 33$.

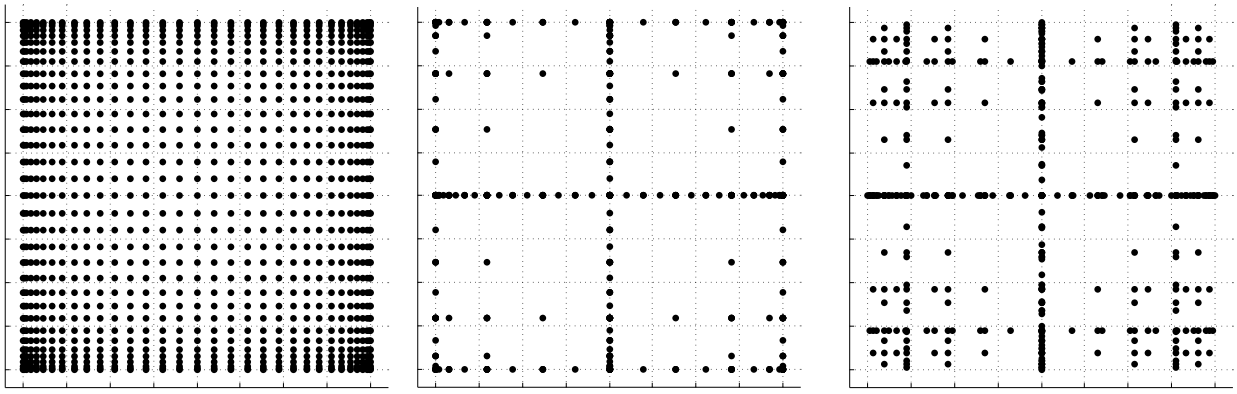


Figure 1. For a two-dimensional parameter space ($n = 2$) and maximum level $\mathrm{w} = 5$, we plot the full tensor product grid using the Clenshaw-Curtis abscissas (left) and isotropic Smolyak sparse grids $\mathscr{A}(5, 2)$, utilizing the Clenshaw-Curtis abscissas (middle) and the Gauss-Legendre abscissas (right).

In Ref. 21, it is demonstrated that the synchronization of total-order PCE with the monomial resolution of a sparse grid is imperfect, and that sparse grid SC consistently outperforms sparse grid PCE when employing the sparse grid to directly evaluate the integrals in Eq. 14. In this paper, we depart from the use of sparse integration of total-order expansions, and instead employ a linear combination of tensor expansions. That is, we compute separate tensor polynomial chaos expansions for each of the underlying tensor quadrature rules (for which there is no synchronization issue) and then sum them using the Smolyak combinatorial coefficient (from Eq. 20 in the isotropic case). This improves accuracy, preserves the PCE/SC consistency property mentioned in Section IV.A.1, and also simplifies PCE for the case of anisotropic sparse grids described next.

For anisotropic Smolyak sparse grids, a dimension preference vector is used to emphasize important stochastic dimensions. Given a mechanism for defining anisotropy, we can extend the definition of the sparse grid from that of Eq. 20 to weight the contributions of different index set components. First, the sparse grid index set constraint becomes

$$\mathrm{w}\underline{\gamma} < \mathbf{i} \cdot \gamma \leq \mathrm{w}\underline{\gamma} + |\gamma| \tag{24}$$

where $\underline{\gamma}$ is the minimum of the dimension weights γ_k, $k = 1$ to n. The dimension weighting vector γ amplifies the contribution of a particular dimension index within the constraint, and is therefore inversely related to

the dimension preference (higher weighting produces lower index set levels). For the isotropic case of all $\gamma_k = 1$, it is evident that you reproduce the isotropic index constraint $\mathrm{w}+1 \leq |\mathbf{i}| \leq \mathrm{w}+n$ (note the change from $<$ to $\leq$). Second, the combinatorial coefficient for adding the contribution from each of these index sets is modified as described in Ref. 29.

V. Statistics from Stochastic Expansions

Stochastic expansion methods have a number of convenient analytic features that make them attractive for use within higher level analyses, such as local and global sensitivity analysis, mixed aleatory/epistemic uncertainty quantification, and design under uncertainty algorithms. First, moments of the response expansion are available analytically. Second, the response expansions are readily differentiated with respect to the underlying expansion variables, and terms may be reorganized to provide Sobol sensivities from a variance-based decomposition.[2,30] Finally, response moment expressions may be differentiated with respect to auxilliary nonprobabilistic variables.[3] In the sections that follow, we focus on moments and global response sensitivities as they provide the initial foundation for adaptivity.

A. Analytic moments

Mean and covariance of the polynomial chaos expansion are available in simple closed form:

$$\mu_i = \langle R_i \rangle \cong \sum_{k=0}^{P} \alpha_{ik} \langle \Psi_k(\boldsymbol{\xi}) \rangle = \alpha_{i0} \tag{25}$$

$$\Sigma_{ij} = \langle (R_i - \mu_i)(R_j - \mu_j) \rangle \cong \sum_{k=1}^{P} \sum_{l=1}^{P} \alpha_{ik}\alpha_{jl} \langle \Psi_k(\boldsymbol{\xi}) \Psi_l(\boldsymbol{\xi}) \rangle = \sum_{k=1}^{P} \alpha_{ik}\alpha_{jk} \langle \Psi_k^2 \rangle \tag{26}$$

where the norm squared of each multivariate polynomial is computed from Eq. 15. The moments μ_R and σ_R are exact moments of the expansion, which converge under refinement to moments of the true response function.

Similar expressions can be derived for stochastic collocation:

$$\mu_i = \langle R_i \rangle \cong \sum_{k=1}^{N_p} r_{ik} \langle \boldsymbol{L}_k(\boldsymbol{\xi}) \rangle = \sum_{k=1}^{N_p} r_{ik} w_k \tag{27}$$

$$\Sigma_{ij} = \langle R_i R_j \rangle - \mu_i \mu_j \cong \sum_{k=1}^{N_p} \sum_{l=1}^{N_p} r_{ik} r_{jl} \langle \boldsymbol{L}_k(\boldsymbol{\xi}) \boldsymbol{L}_l(\boldsymbol{\xi}) \rangle - \mu_i \mu_j = \sum_{k=1}^{N_p} r_{ik} r_{jk} w_k - \mu_i \mu_j \tag{28}$$

where the expectation of a particular Lagrange polynomial constructed at Gauss points and then integrated at these same Gauss points leaves only the weight corresponding to the point for which the interpolation value is one.

B. Analytic global sensitivities: variance-based decomposition

In addition to obtaining derivatives of stochastic expansions with respect to the random variables, it is possible to obtain variance-based sensitivity indices from the stochastic expansions. Variance-based decomposition is a global sensitivity method that summarizes how the uncertainty in model output can be apportioned to uncertainty in individual input variables. VBD uses two primary measures, the main effect sensitivity index S_i and the total effect index T_i. These indices are also called the Sobol' indices. The main effect sensitivity index corresponds to the fraction of the uncertainty in the output, Y, that can be attributed to input x_i alone. The total effects index corresponds to the fraction of the uncertainty in the output, Y, that can be attributed to input x_i and its interactions with other variables. The main effect sensitivity index compares the variance of the conditional expectation $Var_{x_i}[E(Y|x_i)]$ against the total variance $Var(Y)$. Formulas for the indices are:

$$S_i = \frac{Var_{x_i}[E(Y|x_i)]}{Var(Y)} \tag{29}$$

and

$$T_i = \frac{E(Var(Y|x_{-i}))}{Var(Y)} = \frac{Var(Y) - Var(E[Y|x_{-i}])}{Var(Y)} \tag{30}$$

where $Y = f(\mathbf{x})$ and $x_{-i} = (x_1, ..., x_{i-1}, x_{i+1}, ..., x_m)$.

The calculation of S_i and T_i requires the evaluation of m-dimensional integrals which are typically approximated by Monte-Carlo sampling. However, in stochastic expansion methods, it is possible to obtain the sensitivity indices as analytic functions of the coefficients in the stochastic expansion. The derivation of these results is presented in Ref. 2. The sensitivity indices are printed as a default when running either polynomial chaos or stochastic collocation in DAKOTA. Note that in addition to the first-order main effects, S_i, we are able to calculate the sensitivity indices for higher order interactions such as the two-way interaction $S_{i,j}$.

C. Numerical probabilities: importance sampling

In Refs. 3 and 5, forward and inverse reliability design metrics were approximated from reliability indices based on simple moment projections (assumes linear Gaussian response):

$$\beta_{cdf} = \frac{\mu_R - \bar{z}}{\sigma_R} \tag{31}$$

$$z = \mu_R - \sigma_R \bar{\beta}_{cdf} \tag{32}$$

such that the analytic moment expressions could be employed and design sensitivities of β and z could be readily formed from the analytic moment sensitivities. In this work, we seek to eliminate this approximation and replace it with numerical evaluation of probability metrics based on importance sampling.

Importance sampling is a biased sampling method used to sample random variables from different densities than originally defined. These importance sampling densities are constructed to pick important values of input random variables to improve the estimation of a statistical response of interest, such as a mean or probability of failure of a response from a simulation model. The use of biased input sampling densities will result in biased estimators if they are applied directly to the simulation results. However, the simulation results are weighted to correct for the use of the importance sampling densities, and this ensures that the importance sampling estimators are unbiased. Conceptually, importance sampling is very attractive: for example one wants to generate more samples in a failure region when estimating failure probabilities. In practice, however, importance sampling can be challenging to implement efficiently, especially in a general framework that will allow solutions for many classes of problems.

We have developed an importance sampler that has three variations: plain importance sampling (IS), adaptive importance sampling (AIS), and multimodal adaptive importance sampling (MMAIS). These three variations were developed to refine probability of failure estimates in analytic reliability methods. In reliability methods, the random inputs are converted to standard normal space (i.e. independent, identically distributed normal distributions with mean zero and variance one). This conversion makes the calculations associated with importance sampling easier, because the importance densities can also be independent normal random variables. The importance densities are simply taken to be standard normals "shifted" so that the means are no longer zero: the means are (for example) the most probable point of failure (MPP) associated with a reliability method.

In IS, which is typically called after the convergence of a local reliability method, the importance sampling density is centered around the MPP. The method then creates the requested number of samples from the importance density (using Latin Hypercube sampling), evaluates these, weights them by their probability of occurrence given the original density, and calculates the probability. In AIS, which is also typically called after the convergence of a local reliability method, the process is similar to IS but is performed iteratively until the probability of failure converges. Note "adaptive" refers to the updating of the probability of failure estimate: there is no adaptive change in importance sampling density, which is still the product of independent normals centered around probable failure points. Finally, MMAIS was designed to be used with a global reliability method such as EGRA. MMAIS uses all of the samples located in the failure region (or near the limit state) to build a multimodal sampling density. First, it generates a small number of LHS samples around each of the likely failure points. These samples are alloted to the different points based on their relative probabilities of occurrence - more probable points get more samples. The goal is to search for "representative" points. Once these are located, the multimodal sampling density is set and then proceeds

similarly to AIS (sample until convergence of the failure probability). Note that IS and AIS were designed to be started with one likely failure point and MMAIS with several.

The goal in this work was to utilize these importance sampling methods to improve the probability estimation generated by sampling on stochastic expansions. For example, if one wants to use a polynomial chaos expansion to estimate the probability that the (uncertain) response is greater than 78, one could generate 1000 uncertain input variable vectors, substitute them into the PCE expression, and calculate the percentage that are greater than 78. If 630 out of 1000 samples have a value greater than 78, we would say the failure probability estimate is 0.63. This approach works well for probability estimates that are large. However, it may not work for very tiny failure estimates or tail probabilities. For example, if we are trying to ascertain a response level which has a 1.e-6 failure probability, if we only take 1000 samples on the expansion, we might get a probability estimate of zero because we have not sufficiently resolved the probability. Often, it is not trivial to generate millions of random numbers, and thus we want to examine importance sampling as a means of obtaining accurate failure probabilities for rare event, tail probability situations. In this work, we took both the "build points" (the points used to create the stochastic expansion, such as sparse grid points or quadrature points) and the points we generated upon the expansion as initial points with which to seed the importance sampling.

We performed these three importance sampling variations on the multi-modal test problem as defined in Ref. 31. This problem has a highly nonlinear and multimodal response defined by:

$$g(\mathbf{x}) = \frac{\left(x_1^2 + 4\right)(x_2 - 1)}{20} - \sin\frac{5x_1}{2} - 2 \tag{33}$$

The distribution of x_1 is Normal(1.5,1) and x_2 is Normal(2.5,1); the variables are uncorrelated. The response level of interest for this study is $\bar{z} = 0$ with failure defined by $g > \bar{z}$. Figure 2 shows a contour plot in x-space of this response function throughout the ± 5 standard deviation search space. This function clearly has several local optima to the forward-reliability MPP search problem. The importance sampling estimates are given in Table 2. EGRA refers to the Efficient Global Reliability Analysis method explained in Ref. 31. PCE 1000 refers to the failure probability obtained by sampling the polynomial chaos expansion 1000 times. PCE 1000 IS refers to the failure probability obtained by using IS on the PCE, and initializing the IS with 1000 samples on the expansion. PCE1000 AIS and PCE1000 MMAIS refer to the failure probabilities obtained by using AIS and MMAIS on the expansion, again initialized with the 1000 samples on the expansion. IS, AIS, and MMAIS appear to perform better than calculating the failure probability using 1000 samples on the expansion, which is to be expected. However, EGRA appears to be more accurate than the importance sampling methods. This is to be expected, since EGRA focuses on resolving the limit state function whereas PCE is trying to fully resolve the characterization of the response function.

Table 2. Failure probability estimates obtained by importance sampling on polynomial chaos expansions of the multi-modal test function.

Response Threshold	True Failure Probability	EGRA	PCE 1000	PCE 1000 IS	PCE 1000 AIS	PCE 1000 MMAIS
-5	0.0000006	0.0000006	0.000000	0.000013	0.000020	0.000074
-1	0.644443	0.641916	0.669000	0.661667	0.653241	0.653578
0	0.968656	0.968888	0.967000	0.969471	0.964992	0.965110
1	0.997409	0.997367	0.994000	0.995434	0.995695	0.995753
5	0.999999	0.999999	0.999000	0.999898	0.999888	0.999891

VI. Refinement Strategies

Several approaches for refinement of stochastic expansions are presented here: uniform p-refinement, adaptive p-refinement using anisotropic sparse and tensor grids, and goal-oriented adaptive p-refinement using generalized sparse grids.

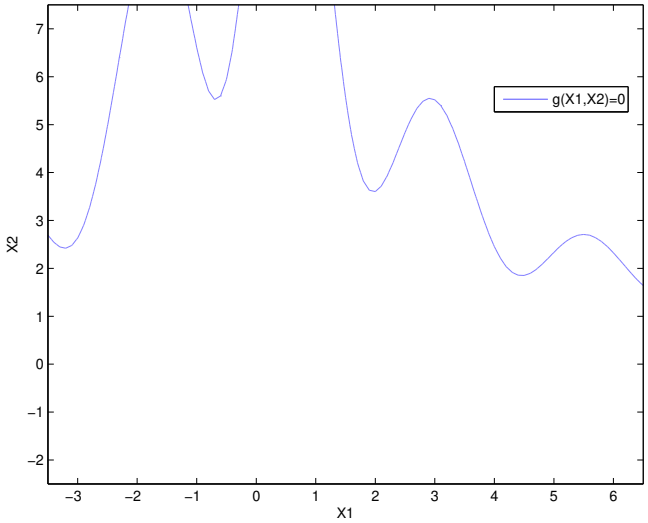


Figure 2. Contour plot of the multimodal function. The solid line is $g = \bar{z} = 0$.

A. Uniform p-refinement

Uniform p-refinement involves ramping the order of a tensor-product quadrature grid or the level of a Smolyak sparse grid isotropically, that is, with no dimension preference. In this case, Sobol' indices are not employed, and the only algorithmic requirements are:

- With the usage of nested integration rules and "restricted growth" schemes that attempt to reduce inefficiency due to the hyperbolic cross phenomenon, one must ensure that a change in level results in an adequate change in the grid. If it does not, continue incrementing the order/level (without grid evaluation) until sufficient change is obtained. An insufficient grid change is likely to cause premature convergence in the refinement process.
- A convergence criterion is required. One useful general-purpose metric is the L^2 norm of the change in the response covariance matrix.

B. Adaptive p-refinement

Adaptive p-refinement involves ramping the order of a tensor-product quadrature grid or the level of a Smolyak sparse grid anisotropically, that is, using a defined dimension preference. This dimension preference may be computed from local sensitivity analysis, global sensitivity analysis, a posteriori error estimation, or decay rate estimation. Herein, we initially focus on global sensitivity analysis from low order isotropic expansions where dimension preference is defined from total Sobol' indices (Eq. 30) and is updated on every iteration. These main effects admit anisotropic sparse grids based on a linear index-set constraint (Eq. 24); with the introduction of interaction effects, nonlinear index-set constraints can also be considered.

C. Goal-oriented p-refinement

If the objectives and constraints of a design under uncertainty problem are focused on variance (e.g., robust design), then the general-purpose formulations described above are sufficiently goal-oriented. However, for other classes of problems (e.g., reliability-based design or stochastic inverse problems), we may prefer to employ refinement approaches guided by assessments of accuracy in other statistical quantities of interest. A particular focus in this effort is to refine adaptively with the goal of accuracy in tail probability estimates. There are two parts to this effort: goal-oriented p-refinement using generalized sparse grids and efficient tail probability estimation.

Since probability levels are not available analytically from stochastic expansions, they must be evaluated numerically using some form of sampling on the expansion. For tail probability estimates, standard sampling approaches (e.g., LHS) can become expensive even for this surrogate-based sampling and we require a more directed probability estimation procedure, in particular the importance sampling procedure described previously in Section V.C.

Given a capability to estimate alternative goals, we must then employ an adaptive refinement procedure that flexibly accommodates these goals. The generalized sparse grid algorithm,[32] originally intended for adaptive numerical integration, is readily adapted to the adaptive refinement of NIPCE and SC expansions. Customizations of our approach from that of Ref. 32 include:

- Rather than hierarchical interpolants, we employ independent polynomial chaos expansions for each active and accepted index set. Pushing and popping index sets then involves increments of tensor chaos expansions (as described in Section IV.A.2) along with corresponding increments to the Smolyak combinatorial coefficients.
- Since we support bases for more than uniform and normal random variables, we exploit rule nesting when possible (i.e., Gauss-Patterson for uniform or transformed uniform variables, and Genz-Keister for normal or transformed normal variables), but we do not require it. This implies a loss of some algorithmic simplifications that occur when grids are strictly hierarchical.
- In the evaluation of the effect of a trial index set, the goal in Ref. 32 is numerical integration and the metric is the size of the increment induced by the trial set on the expectation of the function of interest. It is straightforward to instead measure the effect of a trial index set on response covariance, numerical probability, or other statistical quantities of interest (QOI) computed by post-processing the resulting NIPCE or SC expansion (it is much less straightforward to embed QOI in the calculation of dimension preference for anisotropic tensor/sparse grids).

Given these customizations, the algorithmic steps can be summarized as:

1. *Initialization:* Starting from an initial isotropic or anisotropic reference grid (often the $w = 0$ grid corresponding to a single collocation point), accept the reference index sets as the old set and define active index sets using the admissible forward neighbors of all old index sets.
2. *Trial set evaluation:* Evaluate the tensor grid corresponding to each trial active index set, form the tensor polynomial chaos expansion corresponding to it, update the Smolyak combinatorial coefficients, and combine the trial expansion with the reference expansion. Perform necessary bookkeeping to allow efficient restoration of previously evaluated tensor polynomial chaos expansions.
3. *Trial set selection:* Select the trial index set that induces the largest change in the statistical quantity of interest. In our studies, we employ an L^2 norm of change in probaility/reliability/response level mappings, when level mappings are present, or L^2 norm of change in response covariance, when level mappings are not present.
4. *Update sets:* If the largest change induced by the trial sets exceeds a specified convergence tolerance, then promote the selected trial set from the active set to the old set and update the active sets with new admissible forward neighbors; return to step 2. If the convergence tolerance is satisfied, advance to step 5.
5. *Finalization:* Promote all remaining active sets to the old set, update the Smolyak combinatorial coefficients, and perform a final combination of tensor expansions to arrive at the final result for the statistical quantity of interest.

VII. Computational Results

Capabilities for uncertainty analysis and design based on nonintrusive polynomial chaos and stochastic collocation have been implemented in DAKOTA,[33] an open-source software framework for design and performance analysis of computational models on high performance computers. This section compares NIPCE and SC performance results for several algebraic benchmark test problems. These results build upon NIPCE and SC results for UQ presented in Refs. 11,21, NIPCE-based and SC-based optimization under uncertainty results presented in Ref. 3, and mixed aleatory-epistemic UQ results presented in Refs. 4–6.

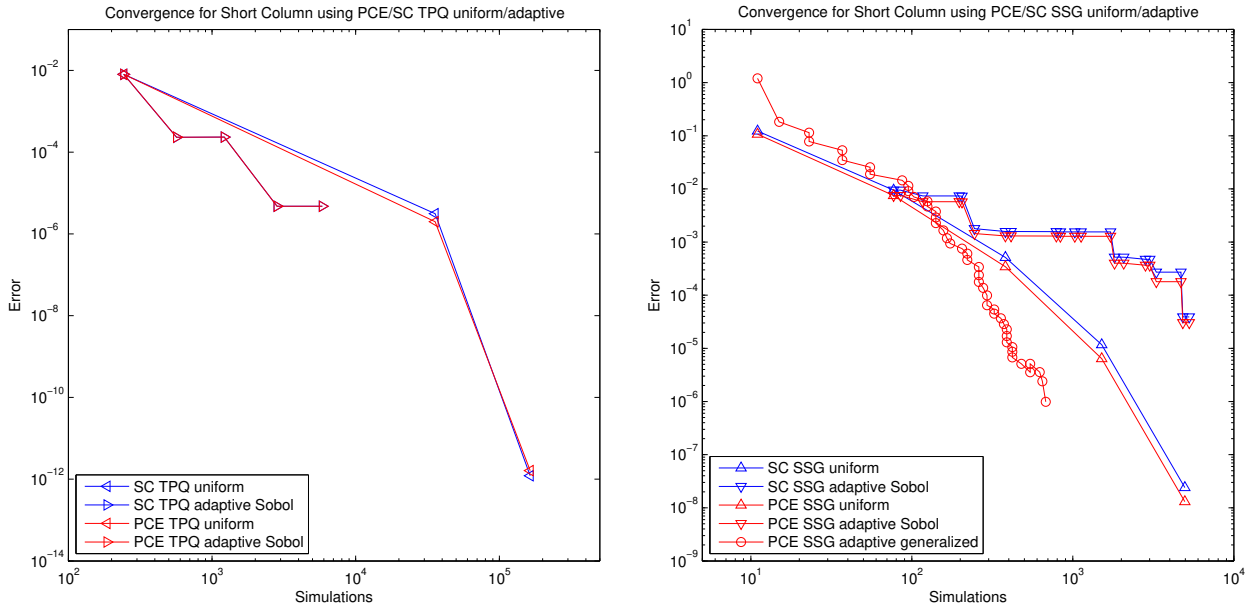
A. Short Column

This test problem involves the plastic analysis of a short column with rectangular cross section (width b and depth h) having uncertain material properties (yield stress Y) and subject to uncertain loads (bending moment M and axial force P).[34] The limit state function is defined as:

$$g(\mathbf{x}) = 1 - \frac{4M}{bh^2Y} - \frac{P^2}{b^2h^2Y^2} \tag{34}$$

The distributions for P, M, and Y are $\mathcal{N}(500, 100)$, $\mathcal{N}(2000, 400)$, and Lognormal with $(\mu, \sigma) = (5,\ 0.5)$[d], respectively, with a correlation coefficient of 0.5 between P and M (uncorrelated otherwise). For P and M, a linear variable transformation is applied, and for Y, a nonlinear variable transformation is applied; in both cases, Hermite orthogonal polynomials and Genz-Keister nested integration rules are employed in the transformed standard normal space. When b and h are included in combined expansions, linear scaling, Legendre polynomials, and Gauss-Patterson nested integration rules are employed.

Figure 3 shows convergence of NIPCE and SC expansion under uniform and adaptive refinement, using the metric of CDF reliability index for failure region $g(x) \leq 0$. For purposes of UQ, we treat all five variables as random variables (b and h are uniform on [5,15] and [15,25], respectively). For tensor-product quadrature (TPQ) in Figure 3(a), the general-purpose adaptive p-refinement approach using Sobol' indices shows a modest improvement in convergence rate over uniform p-refinement. For Smolyak sparse grids (SSG) in Figure 3(b), the Sobol'-based adaptive p-refinement approach is outperformed by simple uniform refinement. However, the goal-oriented adaptive p-refinement approach using generalized sparse grids converges the most rapidly of all approaches, achieving 10^{-6} error tolerance in fewer than 1000 total function evaluations.



(a) TPQ (b) SSG

Figure 3. Convergence of Short Column under uniform/adaptive refinement.

B. Cantilever Beam

The next test problem involves the simple uniform cantilever beam[35,36] shown in Figure 4. Random variables in the problem include the yield stress R and Youngs modulus E of the beam material and the horizontal and vertical loads, X and Y, which are modeled with independent normal distributions using $\mathcal{N}(40000, 2000)$, $\mathcal{N}(2.9E7, 1.45E6)$, $\mathcal{N}(500, 100)$, and $\mathcal{N}(1000, 100)$, respectively. Problem constants include $L = 100$ in. and

[d] $\log(Y) \sim \mathcal{N}(1.6045, .099751)$.

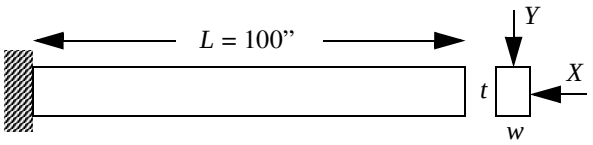


Figure 4. Cantilever beam test problem.

$D_0 = 2.2535$ in. The beam response metrics have the following analytic form for stress S and displacement D:

$$S = \frac{600}{wt^2}Y + \frac{600}{w^2t}X \leq R \quad (35)$$

$$D = \frac{4L^3}{Ewt}\sqrt{(\frac{Y}{t^2})^2 + (\frac{X}{w^2})^2} \leq D_0 \quad (36)$$

where we place them in standard form using

$$g_S = S - R \leq 0 \quad (37)$$

$$g_D = D - D_0 \leq 0 \quad (38)$$

such that negative g values indicate safe regions of the parameter space. For polynomial approximation of g_S and g_D, a linear variable transformation is used and Hermite orthogonal polynomials and Genz-Keister nested integration rules are employed in the transformed standard normal space. When w and t are included in combined expansions, linear scaling, Legendre polynomials, and Gauss-Patterson nested integration rules are employed for these variables. It is worth noting that: (1) g_S is linear and g_D is only mildly nonlinear in the random variables, but both are highly nonlinear in the design variables w and t, and (2) g_D contains a singularity due to the infinite tails of E and therefore does not have finite variance; however, this occurs at 20 standard deviations and, given that Youngs modulus cannot physically be negative, its presence is not of practical or numerical interest.

Figure 5 shows convergence of NIPCE and SC expansion under uniform and adaptive refinement, using metrics of stress and displacement CCDF reliability indices for failure regions $g_S \geq 0$ and $g_D \geq 0$. For purposes of UQ, we again treat all variables as random variables (w and t are uniform on [1,10]). For tensor-product quadrature (TPQ) in Figure 5(a,c) and for Smolyak sparse grids (SSG) in Figure 5(b,d), the general-purpose adaptive p-refinement approach using Sobol' indices show improvements in convergence rate over uniform p-refinement. The goal-oriented adaptive p-refinement approach using generalized sparse grids again converges the most rapidly of all approaches, achieving 10^{-6} error tolerances in just under 1000 total function evaluations, although the improvement relative to Sobol'-based adaptive p-refinement is less pronounced than that shown previously in Figure 3.

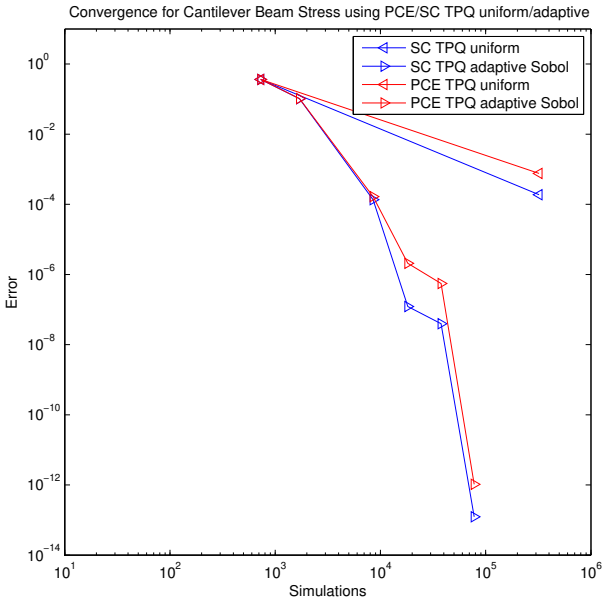
C. Ishigami

The previous two test functions are rational functions with limited regularity, which can degrade convergence rates in polynomial approximations. The Ishigami test problem[37] is a smooth C^∞ function:

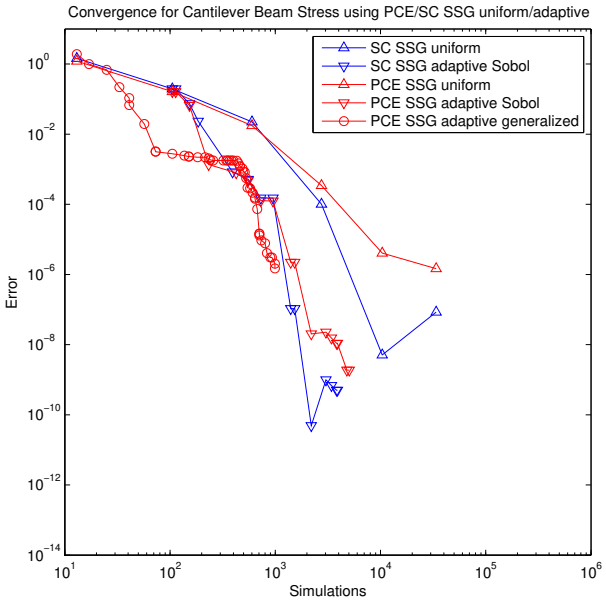
$$f(\mathbf{x}) = \sin(2\pi x_1 - \pi) + 7\sin^2(2\pi x_2 - \pi) + 0.1(2\pi x_3 - \pi)^4 \sin(2\pi x_1 - \pi) \quad (39)$$

where the distributions for x_1, x_2, and x_3 are *iid* uniform on [0,1]. This function was created as a test for global sensitivity analysis methods, but is challenging for any method based on low-order structured grids (due to term cancellations at midpoints and bounds). For approximation by stochastic expansion, a linear scaling transformation is applied and Legendre orthogonal polynomials and nested Gauss-Patterson integration rules are employed in the transformed space.

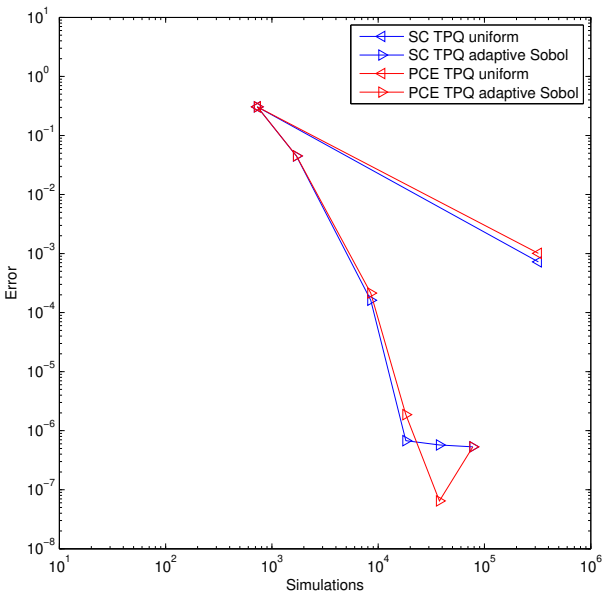
Figure 6 shows convergence of NIPCE and SC under uniform and adaptive refinement. Adaptive p-refinement approaches based on Sobol' indices are initially less effective than uniform refinement due to term cancellations on low order grids. However, with the eventual introduction of higher-order grids, the adaptive approaches jump ahead. Once again, the goal-oriented adaptive p-refinement approach using generalized



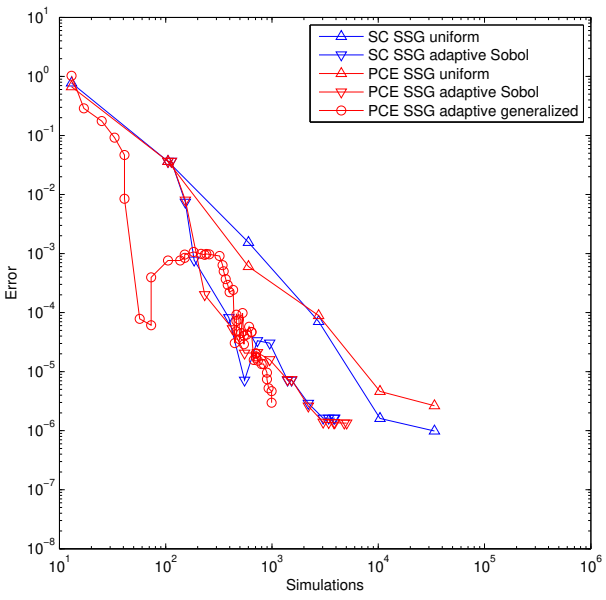
(a) Stress, TPQ



(b) Stress, SSG



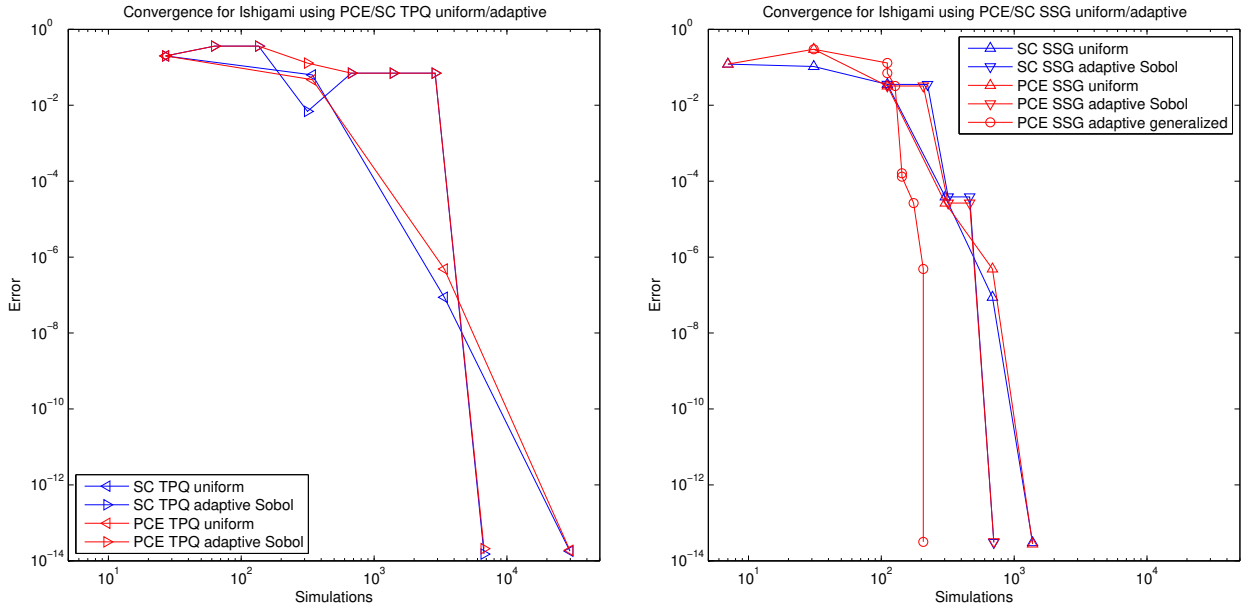
(c) Displacement, TPQ



(d) Displacement, SSG

Figure 5. Convergence of Cantilever Beam stress/displacement under uniform/adaptive refinement.

sparse grids is the most efficient approach, provided that it is started from an initial sparse grid of sufficient order to avoid premature convergence caused by low order term cancellations. This approach achieves 10^{-14} error tolerance in just over 200 total function evaluations.



(a) TPQ (b) SSG

Figure 6. Convergence of Ishigami under uniform/adaptive refinement.

VIII. Conclusions

In this paper, we have explored uniform and adaptive p-refinement approaches for nonintrusive polynomial chaos expansion (NIPCE) and stochastic collocation (SC) approaches to UQ. These methods can be used to automate the assessment of UQ convergence and improve computational efficiency through the detection of anisotropy and the preferential refinement of important stochastic dimensions.

The first class of p-refinement methods used a general-purpose metric of response covariance to guide and control uniform and adaptive refinement processes. In the adaptive refinement case, variance-based decomposition was employed and total Sobol' indices were used to define and update anisotropic dimension preference. While these methods were shown to be reasonably effective on the test problems investigated (improvement over uniform p-refinement in 2 out of 3 cases), it is recognized that variance attribution may not be the best measure of required refinement levels (large variance contributors could be low order and vice versa); rather, future capabilities for spectral coefficient decay rate estimation are expected to be more direct and effective when using anisotropic tensor and sparse grids for adaptive p-refinement of NIPCE.

The second class of p-refinement approaches replaced anisotropic sparse grids with generalized sparse grids and general-purpose refinement metrics with goal-oriented ones. In the three test problems investigated, goal-oriented p-refinement using generalized sparse grids was the top performer, rapidly achieving the desired error tolerances. However, the primary strength of this method (fine-grained precise control of tensor index sets), is also its primary weakness, in that premature convergence can occur for problems that exhibit non-monotic decay in the importance of these index sets (the Ishigami test function is an example of this, and starting the generalized sparse grid algorithm from a higher initial isotropic grid was necessary to prevent premature convergence). By comparison, p-refinement approaches from the first class tend to take larger steps (adding multiple index sets at a time) and therefore are less prone to premature convergence.

Future work will focus on assembling these components to enable adaptive reliability-based design under uncertainty using adaptive goal-oriented p-refinement of stochastic expansions and accelerated procedures for surrogate-based tail probability estimation. In addition, goal-oriented p-refinement addresses accuracy and scalability for smooth problems; future work in h- and hp-refinement is also needed for addressing stochastic

approximation of nonsmooth functions.

Acknowledgments

The authors thank John Jakeman of Australian National University for sharing his insight on generalized sparse grids. The authors also thank Paul Constantine of Sandia Labs and Clayton Webster of Florida Power and Light for discussions related to adaptive collocation, and John Burkardt of Virginia Tech for development of isotropic and anisotropic sparse grid software used for numerical integrations in this work.

References

[1]Xiu, D. and Karniadakis, G. M., "The Wiener-Askey Polynomial Chaos for Stochastic Differential Equations," *SIAM J. Sci. Comput.*, Vol. 24, No. 2, 2002, pp. 619–644.

[2]Tang, G., Iaccarino, G., and Eldred, M. S., "Global Sensitivity Analysis for Stochastic Collocation Expansion," *Proceedings of the 12th AIAA Non-Deterministic Approaches Conference*, No. AIAA-2010-2922, Orlando, FL, April 12–15, 2010.

[3]Eldred, M. S., Webster, C. G., and Constantine, P., "Design Under Uncertainty Employing Stochastic Expansion Methods," *Proceedings of the 12th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, No. AIAA-2008-6001, Victoria, British Columbia, September 10–12, 2008.

[4]Eldred, M. S. and P., S. L., "Efficient Algorithms for Mixed Aleatory-Epistemic Uncertainty Quantification with Application to Radiation-Hardened Electronics. Part I: Algorithms and Benchmark Results," Tech. Rep. SAND2009-5805, Sandia National Laboratories, Albuquerque, NM, 2009.

[5]Eldred, M. S., "Recent Advances in Non-Intrusive Polynomial Chaos and Stochastic Collocation Methods for Uncertainty Analysis and Design," *Proceedings of the 11th AIAA Nondeterministic Approaches Conference*, No. AIAA-2009-2274, Palm Springs, CA, May 4–7 2009.

[6]Tang, G., Swiler, L. P., and Eldred, M. S., "Using Stochastic Expansion Methods in Evidence Theory for Mixed Aleatory-Epistemic Uncertainty Quantification," *Proceedings of the 12th AIAA Non-Deterministic Approaches Conference*, No. AIAA-2010-2589, Orlando, FL, April 12–15, 2010.

[7]Zienkiewicz, O. C. and Zhu, J. Z., "A Simple Error Estimator and Adaptive Procedure for Practical Engineering Analysis," *Int. J. Num. Meth. Eng.*, Vol. 24, 1987, pp. 337–357.

[8]Askey, R. and Wilson, J., "Some Basic Hypergeometric Polynomials that Generalize Jacobi Polynomials," *Mem. Amer. Math. Soc. 319*, AMS, Providence, RI, 1985.

[9]Wiener, N., "The Homogoeneous Chaos," *Amer. J. Math.*, Vol. 60, 1938, pp. 897–936.

[10]Abramowitz, M. and Stegun, I. A., *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover, New York, 1965.

[11]Eldred, M. S., Webster, C. G., and Constantine, P., "Evaluation of Non-Intrusive Approaches for Wiener-Askey Generalized Polynomial Chaos," *Proceedings of the 10th AIAA Nondeterministic Approaches Conference*, No. AIAA-2008-1892, Schaumburg, IL, April 7–10 2008.

[12]Der Kiureghian, A. and Liu, P. L., "Structural Reliability Under Incomplete Probability Information," *J. Eng. Mech., ASCE*, Vol. 112, No. 1, 1986, pp. 85–104.

[13]Simpson, I., "Numerical Integration over a Semi-Infinite Interval, Using the Lognormal Distibution," *Numerische Mathematik*, Vol. 31, 1978.

[14]Gautschi, W., *Orthogonal Polynomials: Computation and Approximation*, Oxford University Press, New York, 2004.

[15]Witteveen, J. A. S. and Bijl, H., "Modeling Arbitrary Uncertainties Using Gram-Schmidt Polynomial Chaos," *Proceedings of the 44th AIAA Aerospace Sciences Meeting and Exhibit*, No. AIAA-2006-0896, Reno, NV, January 9–12 2006.

[16]Golub, G. H. and Welsch, J. H., "Caclulation of Gauss Quadrature Rules," *Mathematics of Computation*, Vol. 23, No. 106, 1969, pp. 221–230.

[17]Eldred, M. S., Agarwal, H., Perez, V. M., Wojtkiewicz, Jr., S. F., and Renaud, J. E., "Investigation of Reliability Method Formulations in DAKOTA/UQ," *Structure & Infrastructure Engineering: Maintenance, Management, Life-Cycle Design & Performance*, Vol. 3, No. 3, 2007, pp. 199–213.

[18]Eldred, M. S. and Bichon, B. J., "Second-Order Reliability Formulations in DAKOTA/UQ," *Proceedings of the 47th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, No. AIAA-2006-1828, Newport, RI, May 1–4 2006.

[19]Rosenblatt, M., "Remarks on a Multivariate Transformation," *Ann. Math. Stat.*, Vol. 23, No. 3, 1952, pp. 470–472.

[20]Box, G. E. P. and Cox, D. R., "An Analysis of Transformations," *J. Royal Stat. Soc.*, Vol. 26, 1964, pp. 211–252.

[21]Eldred, M. S. and Burkardt, J., "Comparison of Non-Intrusive Polynomial Chaos and Stochastic Collocation Methods for Uncertainty Quantification," *Proceedings of the 47th AIAA Aerospace Sciences Meeting and Exhibit*, No. AIAA-2009-0976, Orlando, FL, January 5–8, 2009.

[22]Smolyak, S., "Quadrature and interpolation formulas for tensor products of certain classes of functions," *Dokl. Akad. Nauk SSSR*, Vol. 4, 1963, pp. 240–243.

[23]Gerstner, T. and Griebel, M., "Numerical integration using sparse grids," *Numer. Algorithms*, Vol. 18, No. 3-4, 1998, pp. 209–232.

[24]Barthelmann, V., Novak, E., and Ritter, K., "High dimensional polynomial interpolation on sparse grids," *Adv. Comput. Math.*, Vol. 12, No. 4, 2000, pp. 273–288, Multivariate polynomial interpolation.

[25]Frauenfelder, P., Schwab, C., and Todor, R. A., "Finite elements for elliptic problems with stochastic coefficients," *Comput. Methods Appl. Mech. Engrg.*, Vol. 194, No. 2-5, 2005, pp. 205–228.

[26]Xiu, D. and Hesthaven, J., "High-order collocation methods for differential equations with random inputs," *SIAM J. Sci. Comput.*, Vol. 27, No. 3, 2005, pp. 1118–1139 (electronic).

[27]Nobile, F., Tempone, R., and Webster, C. G., "An Anisotropic Sparse Grid Stochastic Collocation Method for Partial Differential Equations with Random Input Data," *SIAM J. on Num. Anal.*, 2008, To appear.

[28]Wasilkowski, G. W. and Woźniakowski, H., "Explicit Cost Bounds of Algorithms for Multivariate Tensor Product Problems," *Journal of Complexity*, Vol. 11, 1995, pp. 1–56.

[29]Burkardt, J., "The "Combining Coefficient" for Anisotropic Sparse Grids," Tech. rep., Virginia Tech.

[30]Sudret, B., "Global sensitivity analysis using polynomial chaos expansions," *Reliability Engineering and System Safety*, Vol. 93, 2008.

[31]Bichon, B. J., Eldred, M. S., Swiler, L. P., ., Mahadevan, S., and McFarland, J. M.

[32]Gerstner, T. and Griebel, M., "Dimension-adaptive tensor-product quadrature," *Computing*, Vol. 71, No. 1, 2003, pp. 65–87.

[33]Eldred, M. S., Adams, B. M., Haskell, K., Bohnhoff, W. J., Eddy, J. P., Gay, D. M., Hart, W. E., Hough, P. D., Kolda, T. G., Swiler, L. P., and Watson, J.-P., "DAKOTA, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis: Version 4.2 Users Manual," Tech. Rep. SAND2006-6337, Sandia National Laboratories, Albuquerque, NM, 2008.

[34]Kuschel, N. and Rackwitz, R., "Two Basic Problems in Reliability-Based Structural Optimization," *Math. Method Oper. Res.*, Vol. 46, 1997, pp. 309–333.

[35]Sues, R., Aminpour, M., and Shin, Y., "Reliability-Based Multidisciplinary Optimization for Aerospace Systems," *Proceedings of the 42rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, No. AIAA-2001-1521, Seattle, WA, April 16–19, 2001.

[36]Wu, Y.-T., Shin, Y., Sues, R., and Cesare, M., "Safety-Factor Based Approach for Probability-Based Design Optimization," *Proceedings of the 42rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, No. AIAA-2001-1522, Seattle, WA, April 16–19, 2001.

[37]Storlie, C. B., Swiler, L. P., Helton, J. C., and Sallaberry, C. J., "Implementation and evaluation of nonparametric regression procedures for sensitivity analysis of computationally demanding models," *Reliability Engineering and System Safety*, Vol. 94, 2009, pp. 1735–1763.