Adaptive sparse grid algorithms with applications to electromagnetic scattering under uncertainty

Meilin Liu[a,b,*], Zhen Gao[b,c,*], Jan S. Hesthaven[b,**]

[a] *College of Information Science and Technology*
Nanjing University of Aeronautics and Astronautics, China
[b] *Division of Applied Mathematics*
Brown University, USA
[c] *Research Center for Applied Mathematics*
Ocean University of China, China

Abstract

We discuss adaptive sparse grid algorithms for stochastic differential equations with a particular focus on applications to electromagnetic scattering by structures with holes of uncertain size, location, and quantity. Stochastic collocation (SC) methods are used in combination with an adaptive sparse grid approach based on nested Gauss-Patterson grids. As an error estimator we demonstrate how the nested structure allows an effective error estimation through Richardson extrapolation. This is shown to allow excellent error estimation and it also provides an efficient means by which to estimate the solution at the next level of the refinement. We introduce an adaptive approach for the computation of problems with discrete random variables and demonstrate its efficiency for scattering problems with a random number of holes. The results are compared with results based on Monte Carlo methods and with Stroud based integration, confirming the accuracy and efficiency of the proposed techniques.

Keywords: Sparse grids, stochastic collocation, adaptivity, Maxwell's equations

*The two lead authors have contributed equally and substantially to the completion of the published work
**Corresponding author
Email address: Jan.Hesthaven@Brown.edu (Jan S Hesthaven)

1. Introduction

With the increasing need to quantify the impact of random input parameters on different types of physical systems or engineering applications, often described by stochastic ordinary or partial differential equations, comes the need for new and efficient computational techniques to deal with such problems. The classic Monte Carlo based method are increasingly inadequate and a number of alternatives have begun to emerge.

It is reasonable to categorize the majority of these methods into two groups: sampling based statistical methods and probabilistic techniques. In the first category one finds the classic Monte Carlo (MC) method [5] with the clear advantage of being simple and non-intrusive, e.g., one needs only a deterministic solver. The simplicity, however, comes at the cost of very slow convergence as $\mathcal{O}(M^{-1/2})$ where M is the number of samples. This quickly becomes prohibitive even if reasonable accuracy is required, in particular if the interest is on higher moments such as variance/sensitivity. A notable exception to this is for very high dimensional problems where the advantage of the dimensionally independent convergence rate eventually becomes important. However, we shall not consider this limit here.

To accelerate convergence of the MC method, several techniques have been proposed, e.g., Latin hypercube sampling [17], quasi-MC (QMC) method [6], and the Markov chain MC (MCMC) [7] method. However, additional restrictions are often imposed by these methods and their applicability is limited.

An alternative to sampling based techniques have been recently received substantial attention. These methods, known as Stochastic Galerkin or Polynomial Chaos (PC) methods [11, 23], are probabilistic in nature and are based on a generalization of the Wiener-Hermite PC expansion [22]. In this approach, the randomness is represented by the Wiener expansion and the unknown expansion coefficients are found by a Galerkin procedure in the inner product associated with the random variables using in the Wiener expansion.

Substantial recent work has confirmed the accuracy and efficiency of this approach, in particular for problem with low to moderate dimensionality and for problems with sufficient smoothness in observation space, resulting in very efficient representations through the Wiener expansion. However, a substantial disadvantage of the Galerkin approach lies in the need to have to develop entirely new software to solve the large coupled equations resulting

from this procedure. This represents a significant problem as validated existing software can not be used directly to model the impact of randomness and uncertainty.

To address this shortcoming of an otherwise successful approach, several authors have proposed a slight modification of this traditional approach. As mentioned above, the bottleneck in the stochastic Galerkin approach is the creation of a new large coupled system through the required inner product. It is natural to attempt to satisfy the high-dimensional problem in a collocation fashion instead, resulting in a large number of decoupled small problems, much in the sense of an MC approach. However, in the collocation approach, the sampling points are deterministic and associated with integration formulas for the evaluation of high-dimensional integrals in contrast MC based techniques where the sampling points are drawn randomly from some a priori distribution. This approach, now known as stochastic collocation, was first proposed by [19] and more recently revisited and extended in [24] and subsequently considered in more detail by numerous authors, see [26] for a recent review. A clear advantage of this approach over the stochastic Galerkin formulation is its non-intrusive nature, enabling one to use validated software much in the same way as for MC based techniques.

A central component of the efficiency and accuracy of these techniques is the construction of efficient and accurate integration methods for high-dimensional problems. In [24, 26] several options are discussed in detail, including Stroud's cubature points [18], resulting in an efficient approach at moderate accuracy, and sparse grids constructed through Smolyak's algorithm [16] combined with the Clenslaw-Curtis integration method [4]. This latter approach improves accuracy but is costly due to the moderate accuracy of the quadrature.

In this work we discuss the tradeoff between accuracy and computational efficiency in a few different ways. We first consider the use of hierarchical Gauss-Patterson integration formulas as a more accurate alternative to the Clenslaw-Curtis nodes. A complementary discussion of some of this can be found in [10]. To further decrease computational cost we propose to use Richardson extrapolation between the levels to estimate errors and ultimately predict the results of an increased accuracy at very little additional cost.

Most of the past work discussed in the above focuses on problems with continuous random variables. In this work we also discuss basic approaches for the adaptive solution of problems with discrete random variables and demonstrate how to use a priori information effectively. Throughout, we il-

lustrate the efficiency and accuracy of the methods on a set of benchmark problems, often referred to the Genz problems [8, 9]. However, we conclude this work with the consideration of problems of electromagnetic scattering in which Maxwell's equations are solved in the time-domain using a discontinuous Galerkin method. It is important to note, however, that the adaptive techniques proposed here are problem independent and can be applied to general stochastic problems, solved using a variety of computational techniques.

What remains of the paper is organized as follows. In Section 2 we provide a brief overview of polynomial chaos techniques with a focus on stochastic collocation methods. This sets the stage for Section 3 where we discuss high-dimensional integration schemes and introduce Gauss-Patterson integration methods in combination with the Smolyak construction. Section 4 discusses in detail the strategies for adaptivity and the use of Richardson extrapolation as an error estimator in this context. We also briefly discuss ideas that allow the efficient adaptive solution of problems in which discrete random variables are used to model the random behavior. In Section 5 we provide a brief overview of the application cases and the computational technique used to solve Maxwell's equations and then illustrate some additional results in support of the generality of the proposed techniques. Section 6 contains a few concluding remarks.

2. Stochastic Collocation Methods

Let us adopt the notation of [24]. $(\Omega, \mathcal{A}, \mathcal{P})$ is a complete probability space, where Ω is the event space, $\mathcal{A} \in 2^{\Omega}$ the σ-algebra, and $\mathcal{P}$ the probability measure. Assume a d-dimensional bounded domain $D \subset R^d$ $(d = 1, 2, 3)$, with boundary ∂D, and focus on the following problem: find a stochastic function, $u \equiv u(\omega, x) : \Omega \times \bar{D} \rightarrow R$, such that for $\mathcal{P}$-almost everywhere $\omega \in \Omega$, the following equation holds,

$$\mathcal{L}(\omega, x; u) = f(\omega, x), \quad x \in D, \tag{1}$$

subject to the boundary condition

$$\mathcal{B}(\omega, x; u) = g(\omega, x), \quad x \in \partial D, \tag{2}$$

where $x = (x_1, ..., x_d, t)$, $\mathcal{L}$ is a differential operator, and $\mathcal{B}$ is a boundary operator. Note that we do not designate between spatial and temporal dimensions at this stage to keep the notation simple.

We assume that the randomness can be represented by p independent variables with zero mean and unit variance, each depending on the random event ω. There are several ways to achieve this depending on details of the problem [26]. We assume validity of the standard technique of a Karhunen-Loéve expansion to express the randomness as

$$\Xi(\omega) = \mu_0 + \sum_{k=1}^{p} \sqrt{\lambda_k}\psi_k\xi_k(\omega),$$

where μ_0 represents the mean of the random field and ψ_k are the orthogonal eigenfunctions associated with the eigenvalues, λ_k, of the correlation function of the random process being represented, and ξ_k are the random variables (see e.g. [11, 26] for details).

Representing the randomness in this generic form in Eqns. (1)-(2) yields an $p + d$ dimensional differential equation in strong form as

$$\mathcal{L}(\xi, x; u) = f(\xi, x), \quad (\xi, x) \in \Gamma \times D,$$

subject to the boundary condition

$$\mathcal{B}(\xi, x; u) = g(\xi, x), \quad (\xi, x) \in \Gamma \times \partial D,$$

where Γ is the p-dimensional random space.

To account for the impact of the uncertainty, it is natural to consider moments of the solutions over the probabilistic space. In other words, we need to evaluate multi-dimensional integrals of the form

$$I[f] = \int_{\Gamma} f(\xi) d\mu(\xi).$$

The simplest way to achieve this is through a Monte Carlo approach like

$$I[f] \simeq \frac{1}{M} \sum_{m=1}^{M} f(\xi_m)$$

with the M instances, ξ_m, being drawn from the distribution, $\mu(\xi)$. As mentioned previously, the disadvantage of this approach is its low convergence rate which, however, is independent on the dimension p of the random space.

Realizing that all we need is to be able to evaluate integrals accurately, it seems reasonable to utilize more accurate integration techniques. At least for

problems of moderate dimensionality one would expect these to be superior in terms of accuracy vs cost. This line of arguments were first explored in [19] for relatively simply ordinary differential equations and discussed in much more detail in [24]. We refer to these and [26] for further aspects of this.

The essence of the stochastic collocation approach is to abandon the random sampling approach and consider the use of more advanced integration approached and, in this work, adaptive hierarchical integration techniques. In other words we shall solve the deterministic problems

$$\mathcal{L}(\xi_k, x; u) = f(\xi_k, x), \quad x \in D,$$

with boundary condition

$$\mathcal{B}(\xi_k, x; u) = g(\xi_k, x), \quad x \in \partial D.$$

where $\xi_k \in \Gamma$ are specific instances chosen with an integration formula in mind. Clearly, an objective in identifying this integration approach is to minimize the number of samples to achieve a given accuracy in evaluating the integral.

3. Integration for high-dimensional problems

For the multi-dimensional integration, we consider a number of different approaches, the simplest of which is the Stroud [18] cubature points. These are useful when computing integrals of the form

$$I[f] = \int_{[-1,1]^p} f(\mathbf{x}) d\mathbf{x}. \tag{3}$$

This set of cubature points based on $(p+1)$ points is exact for polynomials of degree two, and are given as

$$I[f] \simeq \sum_{i=1}^{n} \omega_i f(\mathbf{x}_i), \tag{4}$$

where the $n = p + 1$ cubature points $\mathbf{x}_i = (x_i^1, x_i^2, ..., x_i^n)$ are given by

$$\begin{aligned} x_i^{2r-1} &= \sqrt{\frac{2}{3}} \cos\left(\frac{2r(i-1)\pi}{n+1}\right), \\ x_i^{2r} &= \sqrt{\frac{2}{3}} \sin\left(\frac{2r(i-1)\pi}{n+1}\right), \end{aligned} \tag{5}$$

for $r = 1, ..., \lfloor n/2 \rfloor$. If n is odd, $x_i^n = (-1)^{(i-1)}/\sqrt{3}$. The weights in (4) are all equal to $2^n/(n+1)$.

Similarly, we have the Stroud-3 method based on $2p$ points which is exact for polynomials of degree three :

$$I[f] \simeq \sum_{i=1}^{n} \omega_i f(\mathbf{x}_i), \tag{6}$$

where the $n = 2p$ cubature points $\mathbf{x}_i = (x_i^1, x_i^2, ..., x_i^n)$ are now defined by

$$\begin{aligned} x_i^{2r-1} &= \sqrt{\frac{2}{3}} \cos\left(\frac{(2r-1)i\pi}{n}\right), \\ x_i^{2r} &= \sqrt{\frac{2}{3}} \sin\left(\frac{(2r-1)i\pi}{n}\right), \end{aligned} \tag{7}$$

for $r = 1, ..., \lfloor n/2 \rfloor$. Again, if n is odd, we have, $x_i^n = (-1)^i/\sqrt{3}$. The weights in (6) are all equal to $2^n/2n$. It can be shown [4, 25] that Stroud-2 and Stroud-3 methods use the minimal number of points for their corresponding integration accuracy. The very simple schemes have recently been extended to general weights in [25].

While the Stroud schemes are efficient and may suffice to compute the expectation, their limited accuracy is often a problem. The most straightforward way to extend the many known one-dimensional integration methods to higher (p) dimensions is through the use of simple tensor products. However, this quickly becomes prohibitive with the number of samples growing like N^p for a quadrature of order N used in p dimensions.

A valuable and often superior alternative to this is the use of sparse grid methods of which the most notable ones are those based on the Smolyak construction. In [16], Smolyak proposed a construction of sparse multivariate quadrature formulas based on sparse tensor products of one dimensional quadrature formulas.

Let us consider the numerical integration of functions $f(\mathbf{x})$ over a p-dimensional unit hypercube $\Gamma := [-1, 1]^p$,

$$I[f] := \int_{\Gamma} f(\mathbf{x}) d\mathbf{x},$$

by a sequence of n_l^p-point quadrature formulas with level $l \in N$ and $n_l^p <$

n_{l+1}^p,

$$Q_l^p f := \sum_{i=1}^{n_l^p} \omega_{li} f(\mathbf{x}_{li}) \tag{8}$$

using the weights ω_{li} and abscissas $\mathbf{x}_{li}$. Moreover, we define the underlying grids of a quadrature formula by

$$\Gamma_l^p := \{\mathbf{x}_{li} \in [-1,1]^p : 1 \leq i \leq n_l^p\}. \tag{9}$$

Now, define the difference quadrature as

$$\Delta_k^1 f := (Q_k^1 - Q_{k-1}^1) f \quad \text{with} \quad Q_0^1 f := 0. \tag{10}$$

Smolyak's construction for the integration of p-dimensional functions f is

$$Q_l^p f := \sum_{|\mathbf{k}|_1 \leq l+p+1} (\Delta_{k_1}^1 \otimes ... \otimes \Delta_{k_d}^1) f \tag{11}$$

where $l \in N$ and $\mathbf{k} \in N^p$. An essential feature of this construction is that the sparse quadrature formulas are nested if the corresponding one-dimensional quadrature nodes are nested. Notably, this rules out classic Gauss quadratures which are not nested and the impact of this is illustrated in Fig. 1 and will be discussed shortly.

3.1. Gauss-Patterson quadrature rules

Seeking nested one-dimensional integration formulas, simple trapezoidal rules immediately comes to mind. However, the limited accuracy of these makes this a less interesting choice. A more appropriate, and widely used, approach is based on the Clenshaw-Curtis rule [4] which is exact for polynomials of order n when using $n+1$ points. This is considerably better than the second order accuracy of the trapezoidal rules but falls short of the $2n+1$ polynomial exactness of the Gaussian quadrature.

The natural question to raise is whether there are nested quadratures which are better than the Clenshaw-Curtis rules, but perhaps not quite as good as the classic quadratures. This question was first addressed by Kronrod [4] who extended an n-point Gauss-Legendre quadrature formula by $n+1$ points such that the quadrature formulas completed the polynomial degree of the exactness with degree $3n+1$(n even) or $3n+2$(n odd). Patterson iterated Kronrod's scheme recursively and obtained a sequence of nested quadrature

formulas with maximal degree of exactness. We refer to [4] for a discussion of the details of this construction.

For $f \in \mathcal{C}^r$, the error bounds of integration f using Gauss-Patterson sparse grid formula will be [10]

$$\left| E_l^1 f \right| = O(2^{-lr}),$$

where l refers to the number of levels in the hierarchical construction.

When considering the efficiency of the integration measured through polynomial exactness, it is well known that using a quadrature with n points, the Clenslaw-Curtis is exact for polynomials up to order $n-1$ and the Gauss-Legendre quadrature is exact for orders up to $2n-1$. For the general Gauss-Patterson rule, one can show exactness up to order $(3n-1)/2$, confirming that this is truly a compromise between the two alternatives [4].

The nested structure of the Gauss-Patterson quadrature grids in combination with the Smolyak's construction results in a natural hierarchical structure for computing the integrals. Hence, to improve the accuracy one needs only compute those new additional grids required to increase from level l to level $l+1$. This is an important property, in particular for high dimensional problems. For the one-dimensional Clenshaw-Curtis rule, the number of points grows like $2^{l-1}+1$, whereas the growth for the Gauss-Patterson rule is $2^l - 1$ since the rule is based on the Gauss quadrature. Hence, when comparing cost of the two methods, it is most appropriate to compare the Clenslaw-Curtis rule at level l with the Gauss-Patterson rule at level $l-1$.

When the dimensionality of the problem increases this becomes more pronounced because the number of quadrature points grows as $O(2^l l^{d-1})$ in the sparse grid [2]. This is illustrated in Table 1 listing the number grid points required for the two-dimensional example with an increasing number of levels. The importance of the nested structure is evident for all levels but the first one.

As a further illustration of this aspect, we plot in Fig.1, the Smolyak based sparse grids Clenshaw-Curtis, Gauss-Patterson and Gauss-Legendre formules for two variables and with 6 levels in the Clenslaw-Curtis based scheme and 5 levels in the two other cases.

The final question to address is whether the Gauss-Patterson based approach, with its improved accuracy but with more quadrature points at a given level, is competitive with the more traditional Clenslaw-Curtis scheme when one compares cost vs accuracy, i.e., to achieve a given accuracy in

Table 1: The number of grid points required in a two-dimensional Smolyak sparse grid based Clenslaw-Curtis (C-C), Gauss-Patterson (G-P) or the Gauss-Legendre (G-L) integration nodes.

Level	C-C	G-P	G-L
1	5	5	5
2	13	17	21
3	29	49	73
4	65	129	221
5	145	321	609
6	321	769	1573

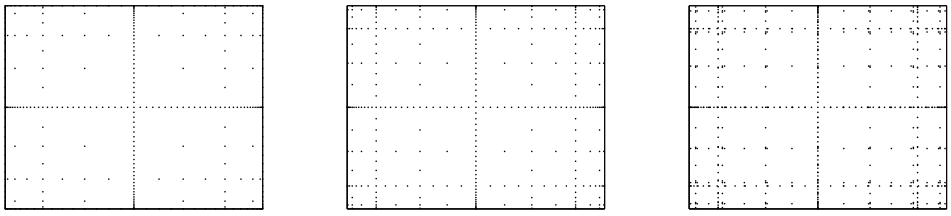

Figure 1: Two-dimensional Smolyak based sparse grids based on Clenshaw-Curtis (left), Gauss-Patterson (middle) and Gauss-Legendre (right) for 6 levels in the hierarchical integration.

the integral, which of the two schemes require the least number of function evaluations.

To address this question, we consider the test function [9] for $\mathbf{x} \in [0,1]^p$

$$f(x_1, \ldots, x_p) = \cos\left(2\pi\omega_1 + \sum_{i=1}^{p} c_i x_i\right),$$

where ω_1 and $\mathbf{c} = (c_1, \ldots, c_p)$ are randomly generated and the c_i's sum to 9.

In Fig. 2 we show the convergence of the different high-dimensional integration schemes discussed previously. We note that the accuracy of the Gauss-Patterson scheme is comparable to that of the scheme based on the full Gauss integration. We also note that the sparse grid based on the Clenslaw-Curtis formula experiences convergence problems for the high-dimensional case. This is a known problem [10] and is likely associated with the kink-phenomena discussed in detail in [21]. It is noteworthy that there are numerous other examples of high-dimensional cases where the Clenslaw-Curtis

formula converges as expected.

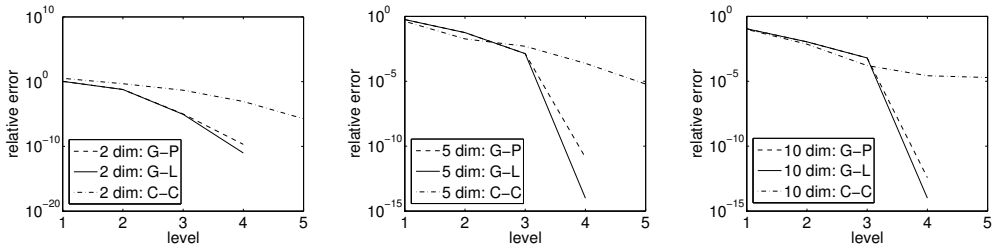

Figure 2: Accuracy vs level for integration of the test-function for dimensions $p = 2, 5, 10$.

A more important question is, however, which of these methods are preferred when taking into account the work needed to achieve a specific accuracy. In Fig. 3 we address this question by showing the time taken to achieve a specific accuracy. From this, it is clear that the sparse grid scheme based on the Gauss-Patterson scheme is superior except when a very moderate accurate is needed in which case the Stroud based schemes may well be preferred. Similar conclusions have been reached by considering other test functions and we will therefore focus on Gauss-Patterson based sparse grid schemes going forward. This conclusion was also reached in [10] although the problems being considered were different and less focused on the cost vs accuracy considerations introduced here.

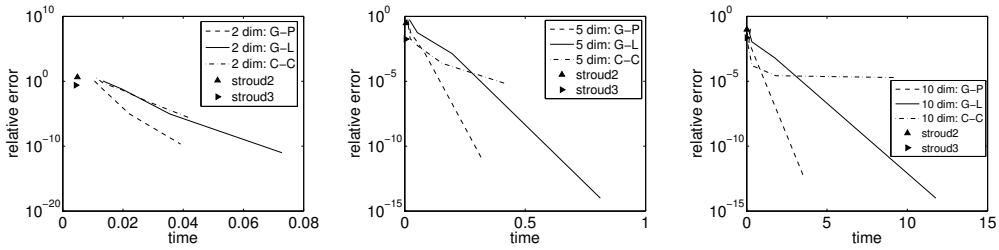

Figure 3: Cost vs accuracy for integration of the test function for dimensions $p = 2, 5, 10$.

4. Error estimation and adaptivity in the sparse grid construction

With the computational work growing exponentially with the dimensionality, it is essential that we seek to minimize the cost by carefully using the required degrees of freedom only where they are needed. This suggests that

an adaptive approach is warranted and therefore the need to formulate an effective error estimator.

The hierarchical nature of the Smolyak based integration immediately suggests that some kind of extrapolation be use to predict the value of the integral at the next level using results obtained at the previous levels [1]. We here propose the use of Richardson extrapolation to achieve this.

It is known that for functions $f \in \mathcal{C}^r$, the error in the Smolyak algorithm is $|E_p^l f| = O(n^{-r/p}(\log n)^{(p-1)(r/p+1)})$ when the integration utilize $n = 2^l l^{p-1}$ as the number of integration point sparse grid at level l to evaluate a p-dimensional integral[14]. With this knowledge of the local error behavior as a function of the level in the Smolyak grid, we can apply classic Richardson extrapolation in the following way.

Assume we have computed an estimate to the integral, $I[f]$, at two different levels. We then know that

$$I[f]^l = I[f] + C_p^r(2^l l^{p-1})^{-r/p}(\log(2^l l^{p-1}))^{(p-1)(r/p+1)},$$
$$I[f]^{l+1} = I[f] + C_p^r(2^{l+1}(l+1)^{p-1})^{-r/p}(\log(2^l(l+1)^{p-1}))^{(p-1)(r/p+1)}.$$

Through simple algebra one recovers that

$$I[f] = \frac{\alpha_1 I(f)^{l+1} - \alpha_2 I(f)^l}{\alpha_1 - \alpha_2},$$

where

$$\alpha_1 = (2^l \cdot l^{(p-1)})^{(-r/p)} \cdot (\log(2^l \cdot l^{(p-1)}))^{(p-1)\cdot(r/p+1)}$$
$$\alpha_2 = (2^{(l+1)} \cdot (l+1)^{(p-1)})^{(-r/p)} \cdot (\log(2^{(l+1)} \cdot (l+1)^{(p-1)}))^{(p-1)\cdot(r/p+1)}.$$

This improved estimate of the integral depends on r which we recall is a measure of the smoothness of $f \in C^r$. This is generally unknown. However, extensive tests have shown that the efficiency of the error estimator has limited sensitivity to this number and we have found that taking it to values of $2-4$ generally yields excellent results as we will illustrate shortly. In some cases, there may be benefits to estimating a different value and this can simply be done by evaluating the accuracy/efficiency of the extrapolated values at the coarse grid levels and identify the optimal value of r.

In our approach we will not compute the next level in the Smolyak grid if the extrapolation result is close to the result at the current level to within a given tolerance. This is vastly advantageous over previously used methods

which simply compare solutions at the two levels, often resulting in having to compute an additional level at substantial cost - typically the cost of a new level is comparable to the combined cost of all previous levels. Instead of this, we use the Richardson extrapolation to estimate the result at the next level, hence dramatically reducing the overall cost without impacting the accuracy.

To test the validity of this approach as a way to accelerate the adaptive sparse grid algorithm and reduce the computational cost we consider a set of high-dimensional test functions proposed in [8, 9]. These function are all defined on $[0,1]^p$ and we seek to integrate them as accurately as possible. The functions being considered have different characteristics:

1. Oscillatory:

$$f_1(x) = \cos(2\pi\omega_1 + \sum_{i=1}^{p} c_i x_i); \tag{12}$$

2. Product peak:

$$f_2(x) = \prod_{i=1}^{p} (c_i^{-2} + (x_i - \omega_i)^2)^{-1}; \tag{13}$$

3. Corner peak:

$$f_3(x) = (1 + \sum_{i=1}^{p} c_i x_i)^{-(d+1)}; \tag{14}$$

4. Gaussian:

$$f_4(x) = \exp(-\sum_{i=1}^{p} c_i^2 (x_i - \omega_i)^2); \tag{15}$$

5. Continuous:

$$f_5(x) = \exp(-\sum_{i=1}^{p} c_i \left| x_i - \omega_i \right|); \tag{16}$$

Different test functions can be obtained by varying the parameters $c = (c_1, ..., c_p)$ and $\omega = (\omega_1, ..., \omega_p)$. The parameters ω_i act as shift parameters, and the difficulty of the functions is increasing with $c_i > 0$. We test the integration with the dimension $p = 10$ and use parameters c_i such that

$$\sum_{i=1}^{p} c_i = b_j, \tag{17}$$

Table 2: Parameter b_j for the five different test functions, f_j.

j	1	2	3	4	5
b_j	9.0	7.25	1.85	7.03	20.4

Table 3: Integrals computed for functions $f_1 - f_4$ for $p = 10$ using sparse grid (SG) and Richardson extrapolation (RE) at levels one to six. Note that the integral was also computed at level zero to enable the extrapolation but is not shown due to its poor accuracy.

		level 1	level 2	level 3	level 4	level 5	level 6
$f1$	SG	-0.1191	-0.1378	-0.1359	-0.1360	-0.1360	-0.1360
	RE	–	-0.1191	-0.1378	-0.1359	-0.1360	-0.1360
$f2$	SG	4.1472e-06	4.3598e-06	4.3452e-06	4.3458e-06	4.3458e-06	4.3458e-06
	RE	–	4.1472e-06	4.3601e-06	4.3451e-06	4.3458e-06	4.3458e-06
$f3$	SG	0.0014	0.0017	0.0018	0.0018	0.0018	0.0018
	RE	–	0.0014	0.0017	0.0018	0.0018	0.0018
$f4$	SG	0.3970	0.4085	0.4091	0.4091	0.4091	0.4091
	RE	–	0.3970	0.4085	0.4091	0.4091	0.4091

where b_j depends on the family f_j and is given by Table 2.

In Table 3 we illustrate the accuracy of the extrapolation for $p = 10$ for the first four test functions above. It is evident that the extrapolation works very well and offers an accurate estimate of the integral at the next level once it is reasonably well approximated at previous levels. We have used $r = 3$ as an estimator of the smoothness in this example but similar results very obtained with $r = 4$, confirming the relative insensitivity of the estimate on the performance of the scheme.

The final test function, f_5, is more challenging and in fact violates the smoothness assumption required to derive the error estimates for the Richardson extrapolation. We show in Table 4 the results for the extrapolation for different dimensions and observe that in spite of the more complex function, the results confirm the accuracy of the proposed approach for local error estimation and, thus, enables accurate adaptive integration. Naturally, the exact same idea can be explored as a dimensional error estimator to enable anisotropic adaptivity, although we have not explored this further in this work. Note that in the above results, we used $r = 3$ for all tests, confirming the lack of sensitivity to this parameter.

Table 4: Integrals computed for function f_5 for dimensions $p = 2, 5, 10$ using sparse grid (SG) and Richardson extrapolation (RE) at levels zero to seven.

		level 0	level 1	level 2	level 3	level 4	level 5	level 6	level 7
p=2	SG	0.0029	0.0208	0.0201	0.0452	0.0326	0.0330	0.0338	0.0338
	RE	–	–	0.0214	0.1337	0.0099	0.0336	0.0349	0.0338
p=5	SG	0.0047	0.0114	0.0102	0.0066	0.0068	0.0073	0.0072	0.0072
	RE	–	–	0.0114	0.0112	0.0063	0.0056	0.0078	0.0072
p=10	SG	0.0092	0.0069	0.0021	0.0034	0.0031	0.0034	0.0035	0.0035
	RE	–	–	0.0069	0.0021	0.0035	0.0030	0.0033	0.0035

4.1. Predictive Sampling for Discrete Variables

For discrete random variables, sparse grids and extrapolation are not directly applicable and we need to seek an alternative approach. We will, in this work, simply use the knowledge of the a priori density of the discrete random variables. Let us, for instance, assume a discrete random variable k with a density

$$f(k;\lambda) = e^{-\lambda}\frac{\lambda^k}{k!}, k = 0, 1, 2, ... \tag{18}$$

where the mean and variance of k are: $\mu_k = \overline{k} = \lambda > 0$ and $\sigma_k^2 = \lambda$, respectively. We use the a priori density to select which samples to compute first. Hence, if $k = 2$ is the most likely case, this will be computed first, followed by the next most likely value of the discrete variable. The convergence of the integral is used to decide whether additional instances are needed. We will demonstrate the computational advantage of this simple idea in the following section.

5. Applications to Electromagnetic Scattering Problems

To evaluate the benefits of the techniques discussed in the above for problems of a more practical character, we consider electromagnetic scattering by a two-dimensional cylinder with holes where the size, the location, and ultimately, the number of holes, are considered as uncertain and described by continuous and discrete random variables.

The physical model is Maxwell's equations

$$\epsilon \frac{\partial \mathbf{E}}{\partial t} = \nabla \times \mathbf{H}, \tag{19}$$

$$\mu \frac{\partial \mathbf{H}}{\partial t} = -\nabla \times \mathbf{E}, \tag{20}$$

in the general three-dimensional domain Λ without sources. $\mathbf{E}$ is electric field intensity, $\mathbf{H}$ is the magnetic field, ϵ and μ are the permittivity and the permeability of the domain.

As an output measure of interest we use the Radar Cross Section (RCS) defined as

$$RCS_{db}(\phi) = 10 \log \left(2\pi \frac{|\mathbf{F}(\phi)|^2}{|\mathbf{E}_i|^2} \right), \tag{21}$$

where $\mathbf{E}_i$ is the incident field and $\mathbf{F}(\phi)$ is a function of $\mathbf{E}$ and $\mathbf{H}$, computing the scattered far field as a function of the polar angle, ϕ. In this particular case, $\mathbf{F}(\phi)$, is near-to-far-field transformation along some closed contour [20].

5.1. Nodal Discontinuous Galerkin Finite Element Method

Discontinuous Galerkin methods is a general and flexible way of solving Maxwell's equations. We follow the formulation in [12] and refer to [13] for a more general and in depth account of these techniques.

We assume the computational domain Ω is approximated by K elements D^k as

$$\Lambda \cong \bigcup_{k=1}^{K} D^k, \tag{22}$$

where D^k is a two-dimensional simplex. In each element, we approximate $\mathbf{E}$ and $\mathbf{H}$ by Lagrange polynomials as

$$[\mathbf{E}_h, \mathbf{H}_h] = \sum_{j=1}^{N} [\mathbf{E}(\mathbf{x}_j, t), \mathbf{H}(\mathbf{x}_j, t)] l_j(\mathbf{x}), \tag{23}$$

where $\mathbf{x}_j$ are the interpolation points. The number of nodes N is given as

$$N = \frac{(n+1)(n+2)}{2}$$

for the nth order polynomial in two dimensions. We insert the approximate solution into Maxwell's equation and require that the local residual is orthogonal to all nth order polynomials. Integrating by part twice, this yields the scheme

$$\int_{D^k} (\epsilon \frac{\partial \mathbf{E}_h}{\partial t} - \nabla \times \mathbf{H}_h) l_i(\mathbf{x}) d\mathbf{x} = -\int_{\partial D^k} \hat{\mathbf{n}}^k \times (\mathbf{H}_k - \mathbf{H}^*) l_i(\mathbf{x}) d\mathbf{x}, \quad (24)$$

$$\int_{D^k} (\mu \frac{\partial \mathbf{H}_h}{\partial t} + \nabla \times \mathbf{E}_h) l_i(\mathbf{x}) d\mathbf{x} = \int_{\partial D^k} \hat{\mathbf{n}}^k \times (\mathbf{E}_k - \mathbf{E}^*) l_i(\mathbf{x}) d\mathbf{x}, \quad (25)$$

where $[\mathbf{E}^*, \mathbf{H}^*]$ denote the numerical flux of the corresponding vector quantities and $\hat{\mathbf{n}}$ is the unit outward normal vector along ∂D^k. The numerical flux is responsible for the coupling of the elements, for the stability of the scheme, and for the imposition of boundary conditions. We use an standard upwind flux with the explicit form given in [12]. We use a low storage Runge-Kutta scheme [12] for the temporal integration.

It should be noted that the method used to solve Maxwell's equations in this work is less important since the focus is on the efficiency of the techniques dealing with the stochastic elements of the problem.

The computational setup of the problem is a plane wave that impinges on a 2-D cylinder from some specific direction. The cylinder has an uncertain number of holes in it and their sizes and locations may also be uncertain. Examples of meshes of the computational model with different number of holes, hole sizes and locations are shown in Fig. 4. The bistatic RCS has an exact solution for the case of a plane wave impinging on a 2-D cylinder without holes. We consider the case with $ka = \pi$ where a is the radius and $k = 2\pi/\lambda$ is the wave number associated with wavelength of the incoming wave. We employ a perfectly matched layer (PML) to absorb the reflected wave in our computation [20]. The comparison of exact solution and numerical solution is shown in Fig. 5, confirming the accuracy of the solver. A complete analysis of the solution approach can be found in [12].

5.2. Low dimensional examples

To verify our algorithm, we first compare the results using the three methods for low-dimensional problems, i.e., we restrict the problem to having one or two random parameters, all assumed to be uniformly distributed random variables. The three methods are a standard Monte Carlo (MC) method, Stroud's method at second and third order, and the sparse grid method. For all case we recover the mean of the RCS as

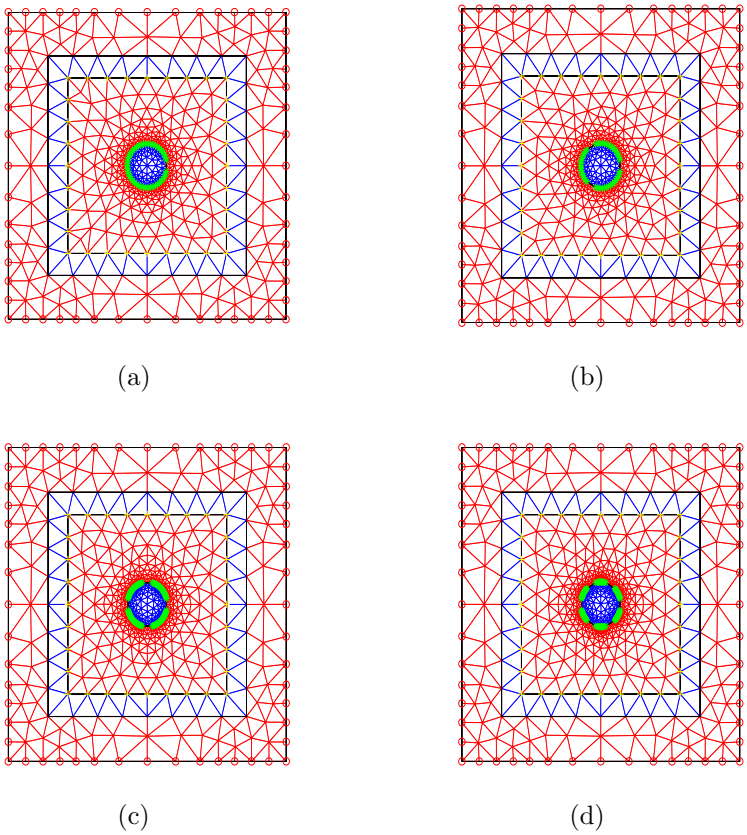

Figure 4: Examples of meshes of the cylinder with different number of holes, hole size, and hole location. a) has one hole, b) three holes, c) four holes, and d) six holes.

$$< RCS >= \sum_{i=1}^{Q} \omega_i RCS_i, \tag{26}$$

and the variance of RCS as

$$var(RCS) = \sum_{i=1}^{Q} \omega_i (RCS_i - < RCS >)^2, \tag{27}$$

where the number of terms, Q, and the integration weights, ω_i, depend on the specific integration technique used.

We first consider the problem with one random parameter, taken to be a cylinder with one hole of random size. The size of the hole is assumed to be

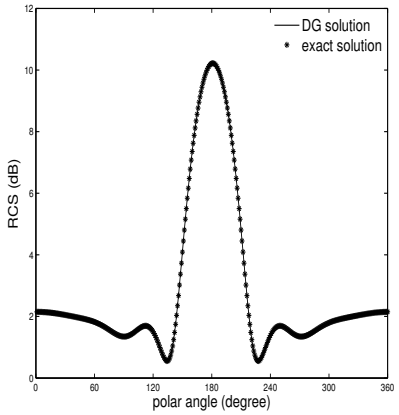

Figure 5: Comparison between numerical and exact bistatic RCS for a $ka = \pi$ metallic cylinder.

in the range of $(\frac{\pi}{12}, \frac{\pi}{6})$ in the polar angle. In Fig. 6 we show the mean RCS computed using the four methods. The level of the sparse grid method is 6 (31 quadrature points) and the Monte Carlo method employs 2000 uniformly distributed points. The difference between level 5 and level 6 results is of order 10^{-3}. The Stroud method of degree 2 and 3 yield the same results since they are identical in this simple case. We observe excellent agreement between the four methods and, in particular, observe excellent agreement between the sparse grid result and the MC computation.

Next, we consider 2 random parameters; the hole size and the angle of the incidence plane wave. The incident wave impinges from the left on the cylinder and changes in the range of $\left(\frac{35}{36}\pi, \frac{37}{36}\pi\right)$. The RCS mean and it plus/minus one standard deviation computed using the sparse grid method are presented in Fig. 7, illustrating the value of these techniques by enabling the computation of sensitivities of output measures of interest. We use 5 levels in the sparse grid computation and estimate the error to be of the order of 10^{-3} based on the extrapolation of the RCS.

5.3. Example of a higher dimensional problem

To simulate a more realistic random wave problem, we need to consider higher dimensional problems. We achieve this in two different ways. In the first example we assume the number of holes is deterministic but that their size and illumination are uniformly distributed random variables.

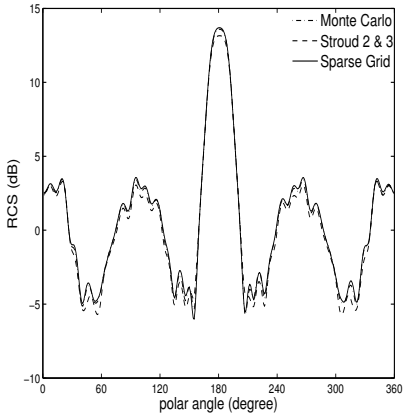

Figure 6: Mean RCS computed using four different sampling methods

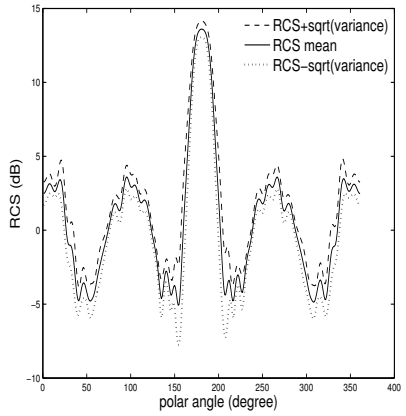

Figure 7: Mean RCS and sensitivity for a two-parameter random problem computed using the adaptive sparse grid method.

In Fig. 8 we present the results for the cylinder with 2 to 4 holes obtained with results obtained using the sparse grid method with 5 levels in the integration. We note in particular the impact of the number of holes on the sensitivity of the RCS.

To further add to the complexity, let us also assume that the number holes is a discrete random variable with a Poisson probability distribution, illustrated in Fig. 9. We assume that the number of holes is ranging between one and nine but that their location is fixed.

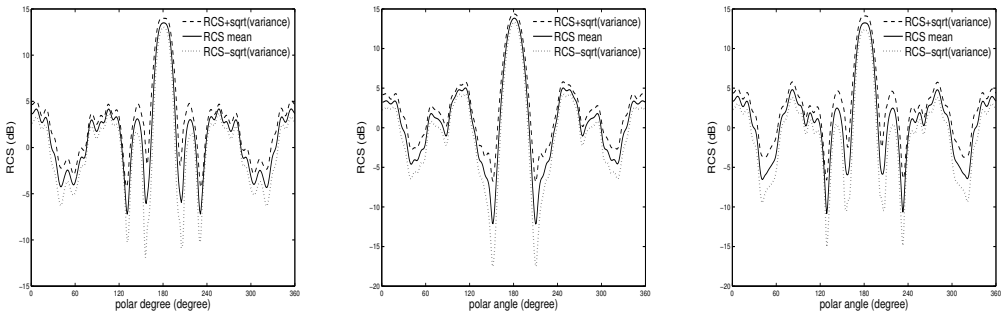

Figure 8: RCS computed for cylinders illustrated in Fig. 4, with 2 (left), 3(middle), and 4(right) holes with uniformly distributed hole size and angle of incidence of the illuminating wave.

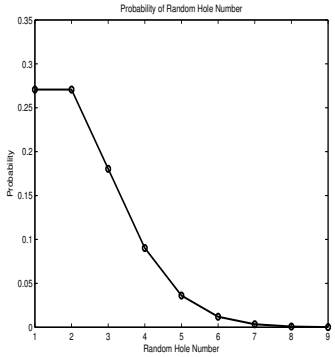

Figure 9: Poisson density associated with the number of holes in the cylinder.

The RCS mean and it sensitivity computed with the adaptive sparse grid method is presented in Fig. 10. We use this case to demonstrate the value of the approach discussed in Section 4.1. To illustrate this, we compute the L^2 error of the RCS with different numbers of holes and list the error as the number of holes increases. Here the reference solution is assumed to be the one where all variations of one to nine holes are accounted for, i.e., for each discrete number of holes, the randomness is in the hole size is accounted for as discussed previously and the moments in the discrete variables are computed subsequently.

The results, simply reflecting the error in the L^2-error or the mean on the RCS, are shown in Table 5. As expected, we quickly see convergence with as little as three to four holes and can terminate the computation and, thus

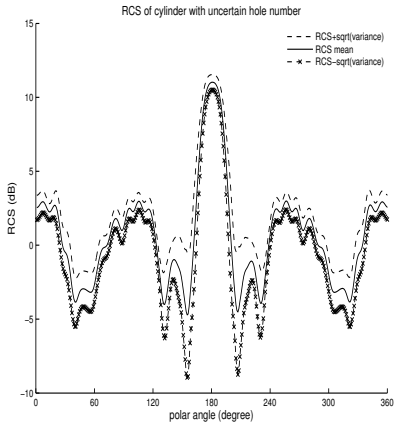

Figure 10: RCS mean and sensitivity for problem with random number of holes and uniformly distributed hole size.

saving substantially in the overall computational cost.

Hole Number	L^2 error of the mean RCS
1	4.2345
2	2.4427
3	1.0191
4	0.4481
5	0.1181
6	0

Table 5: The L^2 error of the mean RCS as a function of the number of holes, assuming that the number of holes is a Poisson distributed random discrete variable.

6. Concluding remarks

In the paper we have discussed the development of adaptive sparse grid methods in the context of stochastic collocation methods for solving partial differential equations with uncertainty. The emphasis has been on identifying methods which delivers maximum accuracy at minimal cost.

We find that the combination of Gauss-Patterson quadratures and Smolyak sparse grid constructions is an effective way to reach this and results in a computationally robust approach. In particular we confirm, in agreement

with related work, that the widely used Clenslaw-Curtis Smolyak based approach may have problems with convergence for certain high-dimensional test functions. No such problem was observed with the Gauss-Patterson based scheme. Similar behavior has also been observed by other authors [10] but the detailed study of the cost vs accuracy question is a new direction of research.

Furthermore, we demonstrate how the strict hierarchical structure of the Smolyak sparse grid lends itself very well to the use of Richardson extrapolation as an efficient estimator for the computed value of the integral at the next level. This is important as it allows for significant savings in the overall computational expense. We demonstrated that this is a robust and general approach as illustrated for several standard test functions.

This suggests an accurate and fast adaptive sparse grid scheme and we illustrated its use and flexibility by considering electromagnetic scattering problems with the scatterer having randomized holes. We also demonstrare a simple but effective way to deal with discrete random variables in an error controlled manner. In combination with the adaptive sparse grid approach, this yields substantial computational savings.

An immediate next step is to look more carefully at the theory of the overall procedure and, in particular, the accuracy of the Richardson extrapolation procedure. We hope to report on this in the near future.

7. Acknowlegement

The first author acknowledges the support of the China Scholarship Committee and National Science Foundation of China (No. 60771017) for this research. The second author also acknowledges the support of China Scholarship Committee and National Science Foundation of China (No. 2008633049) for this research. The last author acknowledges partial support by by OSD/AFOSR FA9550-09-1-0613, and by NSF, and DoE.

[1] H. Bungartz, M. Griebel, and U. Rüde, Extrapolation, Combination, And Sparse Grid Techniques For Elliptic Boundary Value Problems, Comput. Methods Appl. Mech. Engrg, 116(1994), 243-252.

[2] H. Bungartz and M. Griebel, Sparse Grids, Acta Numerica 13(2004), 147–269.

[3] C. Chauviere, J.S. Hesthaven, and L. Wilcox, Efficient Computation of RCS from Scatterers of Uncertain Shapes. IEEE Trans. Antennas Propagat. 55(2007), 1437-1448.

[4] P.J. Davis and P. Rabinowitz, Methods of Numerical Integration. Academic Press, New York, 1975.

[5] G. Fishman, Monte Carlo: Concepts, Algorithms, and Applications, Springer-Verlag, New York, 1996

[6] B. Fox, Strategies for Quasi-Monte Carlo, Kluwer, Dordrect, The Netherlands, 1999.

[7] D. Gamerman, Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference, Chapman and Hall, London, 1997.

[8] A.C. Genz, Testing multidimensional integration routines. In: Tools, Methods, and Languages for Scientific and Engineering Computation. B. Ford, J.C. Rault and F. Thomasset (Ed.), North-Holland, Amsterdam, 1984, 81-94

[9] A.C. Genz, A package for testing multiple integration subroutines. In: Numerical Integration. P. Keast and G. Fairweather (Ed.), Kluwer, Dordrecht, 1987, 337-340.

[10] T. Gerstner and M. Griebel, Numerical integration using sparse grid, Numer. Algor. 18(1998), 209-232.

[11] R.G. Ghanem and P.D. Spanos, Stochastic Finite Elements: A Spectral Approach. Springer-Verlag, New York, 1991.

[12] J. S. Hesthaven and T. Warburton, Nodal High-Order Methods on Unstructured Grids I. Time-Domain Solution of Maxwell's Equations, J. Comput. Phys,181(2002), 186-221.

[13] J. S. Hesthaven and T. Warburton, Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications. Springer Verlag, New York, 2007.

[14] E. Novak and K. Ritter, The curse of dimension and a universal method for numerical integration. In: Multivariate Approximation and Splines, G. Nurnberger, J.W. Schmidt, and G. Walz (Ed.),1997, 177-188.

[15] W.H Reed and T.R.Hill, Triangular mesh methods for the neutron transport equation, Los Alamos Scientific Laboratory Report, LA-UR-73-479, 1973.

[16] S.A.Smolyak, Quadrature and interpolation formulas for tensor products of certain classes of funcitions, Dokl. Akad. Nauk SSSR 4(1963), 240-243.

[17] M. Stein, Large sample properties of simulations using Latin hypercube sampling, Technometrics, 29(1987), 143-151.

[18] A. Stroud, Remarks on the disposition of points in numerical integration formulas, Math. Comput. 11(1957), 257-261.

[19] M.A. Tatang, W.W. Pan, R.G. Prinn, and G.J. McRae. An efficient method for parametric uncertainty analysis of numerical geophysical model. J. Geophy. Res. 102(1997), 2192521932.

[20] A. Taflove, Computational Electrodynamics - The Finite-Difference Time-Domain Method, Aztech House, Boston, 1995.

[21] J. A. C. Weideman and L. N. Trefethen. The Kink Phenomenon in Fejer and Clenshaw-Curtis Quadrature., Report no.6/16, Oxford University Computing Laboratory, 2006.

[22] N. Wiener, The homogeneous chaos, American Journal of Mathematics, 60(1938), 897-936.

[23] D. Xiu and G.E. Karniadakis, The Wiener-Askey Polynomial Chaos for stochastic differential equations, SIAM J. Sci. Comput. 24(2002), 619-644.

[24] D.Xiu and J.S.Hesthaven, High-order collocation methods for differential equations with random inputs, SIAM J. Sci.Comput. 27(2005), 1118-1139.

[25] D. Xiu, Numerical Integration Formulas of Degree Two, Appl. Numer. Math. 58(2008), 1515-1520.

[26] D. Xiu, Fast numerical methods for stochastic computations: A review, Comm. Comput. Phys. 5(2009), 242-272.