

Efficient Solution of Ordinary Differential Equations with High-Dimensional Parametrized Uncertainty

Zhen Gao^{1,2} and Jan S. Hesthaven^{2,*}

¹ *Research Center for Applied Mathematics, Ocean University of China, Qingdao 266071, Shandong, China.*

² *Division of Applied Mathematics, Brown University, 182 George Street, Providence, RI 02912, USA.*

Received 9 January 2010; Accepted (in revised version) 8 September 2010

Available online 27 April 2011

Abstract. The important task of evaluating the impact of random parameters on the output of stochastic ordinary differential equations (SODE) can be computationally very demanding, in particular for problems with a high-dimensional parameter space. In this work we consider this problem in some detail and demonstrate that by combining several techniques one can dramatically reduce the overall cost without impacting the predictive accuracy of the output of interests. We discuss how the combination of ANOVA expansions, different sparse grid techniques, and the total sensitivity index (TSI) as a pre-selective mechanism enables the modeling of problems with hundred of parameters. We demonstrate the accuracy and efficiency of this approach on a number of challenging test cases drawn from engineering and science.

AMS subject classifications: 62F12, 65C20, 65C30

Key words: Sparse grids, stochastic collocation method, ANOVA expansion, total sensitivity index.

1 Introduction

Quantifying the impact of uncertainty in physical systems has received considerable attention during the last decade, in particular emphasizing the need to develop efficient and accurate computational techniques for high-dimensional problems. Applications of such techniques can be found across the sciences and engineering with the uncertainty being caused by insufficient or inaccessible data among other things.

*Corresponding author. *Email addresses:* zhengao.ouc@gmail.com (Z. Gao), Jan.Hesthaven@Brown.edu (J. S. Hesthaven)

In this work we consider problems of the type

$$\begin{aligned}\frac{d\mathbf{u}(t,\boldsymbol{\alpha})}{dt} &= \mathbf{f}(\mathbf{u},t,\boldsymbol{\alpha}), \quad t > 0, \\ \mathbf{u}(0,\boldsymbol{\alpha}) &= \mathbf{g}(\boldsymbol{\alpha}),\end{aligned}$$

where the state vector $\mathbf{u} : \mathbb{R}^+ \times \mathbb{R}^p \rightarrow \mathbb{R}^m$, and the flux $\mathbf{f} : \mathbb{R}^m \times \mathbb{R}^+ \times \mathbb{R}^p \rightarrow \mathbb{R}^m$ is assumed Lipschitz continuous. The solution is parameterized by $\boldsymbol{\alpha} = (\alpha^1, \dots, \alpha^p)$ which describes the system, e.g., the details of the initial conditions or parameters in the flux.

This very general problem arises in numerous applications and there is a long history of the development of accurate and efficient methods for solving them provided $\boldsymbol{\alpha}$ is known accurately. We think of this as the purely deterministic case. However, for many problems, the parameters are not known, known only with poor accuracy, or even entirely inaccessible. One approach in such cases is to endow the parameters with a confidence interval and associated probability density, hence turning the problem into a stochastic problem. We must then consider methods that enable the rapid computation of statistical measures such as the mean and variances of the state variables. This clearly has to be approached carefully since minor, but correlated, changes in some parameters may lead to major changes in the output variables. Simply freezing the parameters at expectation values is generally far from adequate.

It is reasonable to categorize the majority of methods for computationally dealing with such problems into two groups: sampling based statistical methods and probabilistic techniques. In the first category one finds the classic Monte Carlo (MC) method [5] which has the clear advantage of being simple, e.g., one needs only a deterministic solver. The simplicity, however, comes at the cost of slow convergence as $\mathcal{O}(K^{-1/2})$ where K is the number of samples. This quickly becomes prohibitive if even reasonable accuracy is needed, in particular if the interest is on higher moments such as variance/sensitivity. To accelerate convergence of the MC method, several techniques have been proposed, e.g., Latin hypercube sampling [24], 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 not general.

A particular alternative to sampling based techniques has recently received substantial attention. Known as Stochastic Galerkin or Polynomial Chaos (PC) methods, these are probabilistic in nature and based on a generalization of the Wiener-Hermite PC expansion. 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 measure of the random variables used in the Wiener expansion.

Substantial recent work has shown the accuracy and efficiency of this approach, in particular for problem of low to moderate dimensionality and problems with sufficient smoothness in probability space, enabling a very efficient representations through the Wiener expansion. However, a substantial disadvantage of the Galerkin approach is the need to develop entirely new software to solve 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 short coming of an otherwise successful approach, several authors have proposed a modification. The bottleneck in the stochastic Galerkin approach is the creation and solution of a large coupled system through the evaluation of the inner products in the Galerkin procedure. It has been proposed 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 contrast to MC based techniques where the sampling points are drawn randomly from an a priori distribution, in the collocation approach, the sampling points are deterministic and associated with integration formulas for the evaluation of high-dimensional integrals.

This approach, now known as stochastic collocation, was first proposed by [26] and more recently revisited and extended in [30] and subsequently considered in more detail by numerous authors, see [33] for a recent review. A clear advantage of this approach over the stochastic Galerkin formulation is its simple implementation, enabling one to use existing validated software much in the same way as for MC.

A central component of the efficiency and accuracy of these techniques is found in the construction of efficient and accurate integration methods for high-dimensional problems. In [30, 33] several options are discussed in detail, including Stroud's cubature points [25], resulting in a moderate accuracy but high efficiency and sparse grids constructed through Smolyak's algorithm [21] combined with a one-dimensional integration method. This latter approach improves accuracy but also adds cost.

Even with sparse integration techniques, the computational cost associated with accurately solving problems with many parameters remain a significant challenge and additional ideas have to be introduced. In this work we consider this problem and develop a strategy that often allows a substantial compression of the parameter space without impacting the accuracy of the statistics of the predicted output values of interest. A key assumption in this is that one is often not interested in accurate estimation of all output values of a system but just a few or a combination of these. This is an entirely reasonable assumption for many problems and is used widely to accelerate the numerical solution of complex systems of partial differential equations through adjoint based error estimation (see [12] and references therein).

In this work we use this basic premise to first evaluate the sensitivity of the output functional to parameter variations in a cheap fashion and subsequently compress the parameter space based on this. The main tool here will be the Total Sensitivity Index (TSI) to help identify those parameters which impact the output the strongest. This is computed efficiently through an approximate ANOVA expansion using sparse grid techniques. Compressing the parameter space further reduces the computational complexity of the system and a new and more accurate ANOVA expansion for the reduced system can be computed to efficiently evaluate the sensitivity of the output. We will demonstrate that this multi-stage approach allows us to accurately and efficiently model problems with several hundred parameters and accurately compute both expectation and sensitiv-

ity values of the output of interests.

What remains of the paper is organized as follows. We begin in Section 2 with a brief introduction to the stochastic collocation method and continue in Section 3 with some background material on sparse grid integration methods. In Section 4 we introduce the ANOVA expansion as an efficient way to represent and compress functions depending on a high-dimensional parameter space. This sets the stage for Section 5 where we introduce the total sensitivity index and demonstrates its effectivity in identifying the important reduced parameter space and enable a reduction of the parameter space without impacting the accuracy of the predicted values. The accuracy and efficiency of this approach is illustrated in Section 6 where we consider more complex problems. Section 7 contains a few concluding remarks and outlook for future work.

2 The stochastic collocation method

Let us adopt the notation of [30]. $(\Omega, \mathcal{A}, \mathcal{P})$ 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. We focus on the problem of finding a stochastic function, $u \equiv u(\boldsymbol{\alpha}(\omega), t)$, such that for \mathcal{P} -almost everywhere $\omega \in \Omega$, the following equation holds,

$$\frac{du(\boldsymbol{\alpha}(\omega), t)}{dt} = f(\mathbf{u}, t, \boldsymbol{\alpha}(\omega)), \quad t > 0, \quad (2.1)$$

subject to the initial condition

$$\mathbf{u}(\boldsymbol{\alpha}(\omega), 0) = g(\boldsymbol{\alpha}(\omega)), \quad (2.2)$$

where the state vector, \mathbf{u} , and nonlinear flux, f , are defined as previously.

We now assume that the parameters, $\boldsymbol{\alpha}$, each can be endowed with a probability density and in most cases we will simply assume that all parameters are uniformly distributed random variables with a prescribed mean and variance. This is not, however, an essential condition and can be relaxed.

To account for the impact of the uncertainty in the parameters, 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

$$I[u] = \int_{R^p} G(u(\omega)) d\mu(\omega),$$

where μ is the joint distribution and $G(u)$ is some function of u . The simplest way to accomplish this is through a Monte Carlo approach like

$$I[u] \simeq \frac{1}{M} \sum_m G(u(\omega_m))$$

with the samples being drawn from the distribution, $\mu(\omega)$. The disadvantage of this is the low convergence rate which, however, is independent of the dimension p of the random space.

Realizing that we need only evaluate integrals accurately, it is reasonable to utilize more accurate integration techniques. At least for problems of moderate dimensionality one would expect this to be superior in terms of accuracy and cost. A subtle assumption in this argument is that the joint probability density of the random variables has a certain amount of smoothness. If not, however, the Monte Carlo integration will also converge very slowly. This line of arguments were first explored in [26] for relatively simple ordinary differential equations and discussed in much more detail in [30,33].

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

$$\frac{du(\boldsymbol{\alpha}(\omega_k),t)}{dt} = f(\mathbf{u},t,\boldsymbol{\alpha}(\omega_k)), \quad t > 0,$$

with the initial condition

$$\mathbf{u}(\boldsymbol{\alpha}(\omega_k),0) = g(\boldsymbol{\alpha}(\omega_k)),$$

where $\omega_k \in \Gamma$ represent specific instances of the parameter values chosen with an integration formula in mind.

3 Sparse grid integration methods

An objective in identifying this integration approach is to minimize the number of samples to achieve a given accuracy when evaluating the integral. For the multi-dimensional integration, we utilize a number of different approaches, the simplest of which is the Stroud [25] cubature points. These are useful to compute integrals of the form

$$I[u] = \int_{[-1,1]^p} G(u(\boldsymbol{\alpha})) d\boldsymbol{\alpha} \tag{3.1}$$

which will be associated with p -dimensional uniformly distributed random parameters, $\boldsymbol{\alpha}$. This set of cubature points based on $(p+1)$ points is exact for polynomials of degree two, and are written as

$$I[u] \simeq \frac{2^n}{n+1} \sum_{i=1}^n G(u(\boldsymbol{\alpha}_i)), \tag{3.2}$$

where the $n = p+1$ cubature points $\boldsymbol{\alpha}_i = (\alpha_i^1, \alpha_i^2, \dots, \alpha_i^n)$ are given by

$$\alpha_i^{2r-1} = \sqrt{\frac{2}{3}} \cos\left(\frac{2r(i-1)\pi}{n+1}\right), \quad \alpha_i^{2r} = \sqrt{\frac{2}{3}} \sin\left(\frac{2r(i-1)\pi}{n+1}\right), \tag{3.3}$$

for $r=1, \dots, [n/2]$. If n is odd, $\alpha_i^n = (-1)^{(i-1)} / \sqrt{3}$. Similarly, the Stroud-3 method is based on $2p$ points and is exact for polynomials of degree three:

$$I[u] \simeq \frac{2^n}{2n} \sum_{i=1}^n G(u(\alpha_i)), \quad (3.4)$$

where the $n=2p$ cubature points α_i are defined by

$$\alpha_i^{2r-1} = \sqrt{\frac{2}{3}} \cos\left(\frac{(2r-1)i\pi}{n}\right), \quad \alpha_i^{2r} = \sqrt{\frac{2}{3}} \sin\left(\frac{(2r-1)i\pi}{n}\right), \quad (3.5)$$

for $r=1, \dots, [n/2]$. It can be shown [4, 32] that Stroud-2 and Stroud-3 methods employ the minimal number of points for their corresponding integration accuracy. These very simple schemes have recently been extended to general weights in [32].

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

A valuable and often superior alternative is the use of sparse grid methods of which the most notable one is based on the Smolyak construction [21], leading to sparse multivariate quadrature formulas based on sparse tensor products of one dimensional quadrature formulas.

Consider the numerical integration of functions $u(\alpha)$ over a p -dimensional unit hypercube $\Omega := [-1, 1]^p$,

$$I[u] := \int_{\Omega} G(u(\alpha)) d\alpha.$$

For the purpose of introduction consider first a one-dimensional quadrature formula for a univariate function u

$$Q_i^1 u = \sum_{i=1}^{n_i^1} \omega_i G(u(\alpha_i^1)), \quad (3.6)$$

with $Q_0^1 u = 0$ and for $i \in N_+$ define

$$\Delta_i^1 u = (Q_i^1 - Q_{i-1}^1)u. \quad (3.7)$$

The tensor product of p -dimensional quadrature formulas is computed by the sums of possible combinations of one-dimensional quadrature formula

$$Q_{i_1}^1 \otimes \dots \otimes Q_{i_p}^1 u = \sum_{i_1=1}^{n_{i_1}^1} \dots \sum_{i_p=1}^{n_{i_p}^1} \omega_{i_1 i_1} \dots \omega_{i_p i_p} G(u(\alpha_{i_1 i_1}^1, \dots, \alpha_{i_p i_p}^1)). \quad (3.8)$$

To decrease the computational cost of tensor product, Smolyak's algorithm for the p -dimensional quadrature formula is given by

$$Q_l^p u = \sum_{|k|_1 \leq l+p-1} (\Delta_{k_1}^1 \otimes \dots \otimes \Delta_{k_p}^1) u \tag{3.9}$$

for $l \in n$ and $k \in n^p$. The last formula can also be written as

$$Q_l^p u = \sum_{l \leq |k|_1 \leq l+p-1} (-1)^{l+p-|k|_1-1} \binom{p-1}{|k|_1-l} (Q_{k_1}^1 \otimes \dots \otimes Q_{k_d}^1) u. \tag{3.10}$$

For other forms, see [29].

3.1 Gauss-Patterson quadrature rules

An essential feature of the Smolyak construction is that the sparse quadrature formulas are nested if the corresponding one-dimensional quadrature nodes are nested. Notably, this rules out the use of classic Gauss quadratures.

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 [32, 33] which is exact for polynomials of order n when $n+1$ points are used. 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 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 considered 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). The additional $n+1$ points are the zeros of the Stieltjes polynomials F_{n+1} which satisfy

$$\int_{-1}^1 P_n(x) F_{n+1}(x) x^j dx = 0, \quad j=0,1,\dots,n, \tag{3.11}$$

where $P_n(x)$ is the n th-order Legendre polynomial. Patterson iterated Kronrod's scheme and obtained a sequence of nested quadrature formulas with maximal degree of exactness. The construction includes a sequence of polynomial $G_k(x)$ of degree $2^{k-1}(n+1)$, $k \geq 1$, satisfying

$$\int_{-1}^1 P_n(x) \left(\prod_{i=1}^{k-1} G_i(x) \right) G_k(x) x^j dx = 0, \quad j=0,1,\dots,2^{k-1}(n+1)-1. \tag{3.12}$$

Through this method $F_{n+1}(x) = G_1(x)$ and $G_k(x)$ are orthogonal to all the polynomials of degree less than $2^{k-1}(n+1)$ with respect to the weight function $P_n(x) (\prod_{i=1}^{k-1} G_i(x))$. The

combination of the zeros of $P_n(x)$ and $G_j, 1 \leq j < k$ forms the grids for the Gauss-Patterson quadrature.

When considering the efficiency of the integration measured through polynomial exactness, it is well known that using a quadrature with n points, the Clenshaw-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 Gauss-Patterson rule, one can show exactness up to order $(3n-1)/2$, confirming that this is truly a compromise [4]. However, it is worth keeping in mind that polynomial exactness is just one measure of accuracy and making other choices may well better highlight the advantages of other methods, e.g., in [19] it is shown that the Clenshaw-Curtis very often is comparable in accuracy to Gauss quadrature despite formally having lower polynomial exactness.

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. To improve the accuracy one needs only compute new additional grids from level l to level $l+1$. This nested structure is an important property, in particular for high dimensional problems to minimize the overall computational cost. 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 Clenshaw-Curtis rule at level l with the Gauss-Patterson rule at level $l-1$.

An important question to address is whether the Gauss-Patterson based approach, with its improved accuracy but with more quadrature points at a given level, are competitive with the more traditional Clenshaw-Curtis scheme when one compares cost vs accuracy, i.e., to achieve a given accuracy in the integral, which of the two schemes require the least number of function evaluations. This has been addressed recently by several authors [11, 15] with the conclusion that the Smolyak construction based on the Gauss-Patterson rule is indeed the most efficient compromise and we shall use that in what remains unless stated otherwise.

4 The ANOVA expansion

A key technique in what follows is the ANOVA expansion and the associated definition of effective dimensions of a parameter space.

Without loss of generalization, we let the integration domain \mathbb{D} be $[0,1]^p$, and $u \in \mathcal{L}^2(\mathbb{D})$. Let t be any subset of coordinate indices $\mathcal{P} = \{1, \dots, p\}$ and $|t|$ denote the cardinality of t , let α^t denote the $|t|$ -vector containing the components of the vector $\alpha \in [0,1]^{|t|}$ indexed by t . Let also $A^{|t|}$ denote the $|t|$ -dimensional unit hypercube which is the projection of the p -dimensional unit hypercube A^p onto the hypercube indexed by t . Then u can be expanded using an ANOVA expansion as [2, 13, 20]

$$u(\alpha) = u_0 + \sum_{t \subseteq \mathcal{P}} u_t(\alpha^t), \quad (4.1)$$

where $u_t(\boldsymbol{\alpha}^t), t \subseteq \mathcal{P}$ is defined recursively through

$$u_t(\boldsymbol{\alpha}^t) = \int_{A^{p-|t|}} u(\boldsymbol{\alpha}) d\boldsymbol{\alpha}_{\mathcal{P} \setminus t} - \sum_{w \subset t} u_w(\boldsymbol{\alpha}^w) - u_0, \tag{4.2}$$

starting with

$$u_0 = \int_{A^p} u(\boldsymbol{\alpha}) d\boldsymbol{\alpha}, \quad \int_{A^0} u(\boldsymbol{\alpha}) d\boldsymbol{\alpha}^\emptyset = u(\boldsymbol{\alpha}). \tag{4.3}$$

Here $d\boldsymbol{\alpha}_{\mathcal{P} \setminus t}$ indicates integration over all the coordinates without indices included in t . The total number of terms in the ANOVA expansion is 2^p .

The ANOVA expansion is a finite and exact expansion of a general high-dimensional function [2]. Furthermore, the individual terms in the expansion are mutually orthogonal, i.e.,

$$\int_{A^p} u_t(\boldsymbol{\alpha}^t) u_w(\boldsymbol{\alpha}^w) d\boldsymbol{\alpha} = \delta_{tw} \tag{4.4}$$

and, as a natural consequence of this, each terms has a zero mean

$$\int_{A^p} u_t(\boldsymbol{\alpha}^t) d\boldsymbol{\alpha} = 0, \quad |t| > 0.$$

The computational realization of the ANOVA expansion is achieved through the recursive formula (4.2) and the use of orthogonality, Eq. (4.4), as

1. Compute u_0 of $u(\boldsymbol{\alpha})$ using

$$u_0 = \int_{A^p} u(\boldsymbol{\alpha}) d\boldsymbol{\alpha}.$$

2. Compute the $p, (p-1)$ -dimensional integrals

$$u_i(\boldsymbol{\alpha}^i) = \int_{A^{p-1}} u(\boldsymbol{\alpha}) d\boldsymbol{\alpha}_i,$$

where $d\boldsymbol{\alpha}_i$ indicates all indices not including i . The first order terms in the ANOVA expansion are then computed by

$$U_i = u_i - u_0. \tag{4.5}$$

These terms each depend on one parameter.

3. Compute $p(p-1)/2, (p-2)$ -dimensional integrals

$$u_{ij}(\boldsymbol{\alpha}^i, \boldsymbol{\alpha}^j) = \int_{A^{p-2}} u(\boldsymbol{\alpha}) d\boldsymbol{\alpha}_{ij} - u_0, \quad i < j = 1, \dots, p,$$

where $d\boldsymbol{\alpha}_{ij}$ indicates all indices not including ij . The second order terms U_{ij} will be computed by the recursive formula

$$U_{ij} = u_{ij} - U_i - U_j, \quad i < j = 1, \dots, p,$$

where U_i and U_j are defined by Eq. (4.5). This accounts for all binary dependencies between the parameters.

4. The third or higher order terms in expansion are computed by using the similar recursive approach, requiring the computation of $\binom{p}{|t|}$ terms of $p - |t|$ size integrals.
5. The ANOVA expansion of $u(\boldsymbol{\alpha})$ is now expressed as

$$u(\boldsymbol{\alpha}) = u_0 + \sum_{i=1}^p U_i + \sum_{1=i < j \leq p} U_{ij} + \sum_{\substack{t \subset P \\ |t|=3}} U_t + \dots.$$

The computational bottleneck here is the need to evaluate a large number of high-dimensional integrals needed in the construction of the expansion.

Define the truncated ANOVA expansion of order s as

$$u(\boldsymbol{\alpha}; s) = u_0 + \sum_{\substack{t \subset P \\ |t| \leq s}} u_t(\boldsymbol{\alpha}^t), \quad (4.6)$$

where $u_t(\boldsymbol{\alpha}^t)$ and u_0 are as above.

The concept of effective dimension of a particular integrand was introduced in [1, 17] and also discussed in [18]. The effective dimension of u defined in the superposition sense is the smallest integer p_s such that

$$\sum_{0 < |t| \leq p_s} V_t(u) \geq qV(u), \quad (4.7)$$

where $q \leq 1$. Here $V_t(u)$ and $V(u)$ are defined by

$$V_t(u) = \int_{A^p} (u_t(\boldsymbol{\alpha}^t))^2 d\boldsymbol{\alpha}, \quad V(u) = \sum_{|t| > 0} V_t(u), \quad (4.8)$$

and can be thought of the variability of u over a given set t .

The relationship between the accuracy of the truncated ANOVA expansion and the superposition dimension is made clear in the following result [22, 23, 28].

Theorem 4.1. *Assume that the function $u(\boldsymbol{\alpha})$ has superposition dimension p_s based on q and let $u(\boldsymbol{\alpha}; p_s)$ denote the truncated ANOVA expansion of order p_s . Then*

$$Err(\boldsymbol{\alpha}, p_s) \leq (1 - q),$$

where $Err(\boldsymbol{\alpha}, p_s)$ is the normalized approximation error defined by

$$Err(\boldsymbol{\alpha}, p_s) = \frac{1}{V(u)} \int_{A^p} [u(\boldsymbol{\alpha}) - u(\boldsymbol{\alpha}; p_s)]^2 d\boldsymbol{\alpha}.$$

This highlights that if the superposition dimension is small, $p_s \ll p$, the function can be well approximated by using just a few terms in the ANOVA expansion. This promises to dramatically reduce the cost of the computation of the expansion.

To illustrate the efficiency of the ANOVA expansion in accurately and efficiently representing a high-dimensional function using a truncated expansion, let us consider a subset of the classic test functions [9, 10].

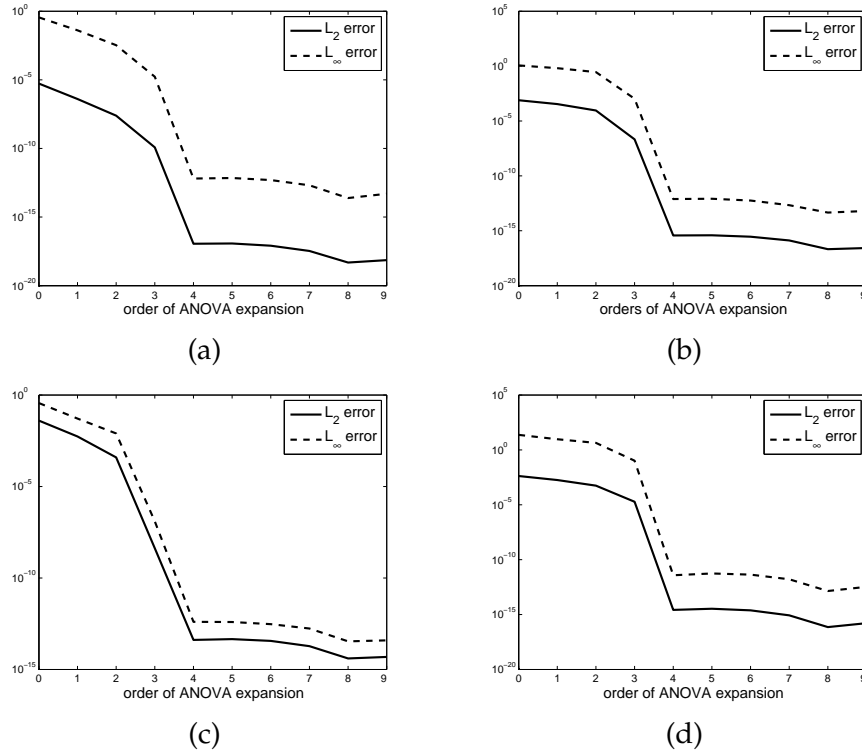


Figure 1: The accuracy of the ANOVA expansion measured in both the L_2 and the L_∞ norm as a function of the number of terms for the four test functions described in the text. (a) u_1 , (b) u_2 , (c) u_3 , (d) u_4 .

- Product Peak function: $u_1(x) = \prod_{i=1}^p (c_i^{-2} + (x_i - \omega_i)^2)^{-1}$,
- Corner Peak function: $u_2(x) = (1 + \sum_{i=1}^p c_i x_i)^{-(p+1)}$,
- Gaussian function: $u_3(x) = \exp[-\sum_{i=1}^p c_i^2 (x_i - \omega_i)^2]$,
- Continuous function: $u_4(x) = \exp[-\sum_{i=1}^p c_i |x_i - \omega_i|]$,

where the parameters $c = (c_1, \dots, c_p)$ and $\omega = (\omega_1, \dots, \omega_p)$ are generated randomly. The parameter ω acts as a shift parameter and the parameters c are constrained. See [9,10] for the details.

In Fig. 1 we show both the L_2 error and the L_∞ error of the ANOVA expansion for $p = 10$ for the four Genz test functions. The error is computed using a fine sparse grid. It is clear that once the 4th order terms in the ANOVA expansion are computed, the complete expansion approximates the full parametric variation very well to accuracies below 10^{-10} . It is encouraging that this appears to be insensitive to the choice of test-function. This has been confirmed with many other tests also, typically showing that including 2nd order terms often suffices. The observation that one often finds a small effective dimension for high-dimensional problems has also been noted [17, 28], e.g., for

large classes of problems one often finds that including 2nd or 3rd order terms in the ANOVA expansion suffices for problems of moderate dimensionality, e.g., $p \leq 25$.

5 Parameter space compression through sensitivity estimation

While the use of the ANOVA expansion allows for an efficient way to represent the high-dimensional function, the dimension of the parameter space remains unchanged. However, for many types of problems, certain parameters are likely more important than others, in particular in cases where one is mainly interested in a subset of the dependent variables or some specific output functional.

This suggests that if we could reliably and at modest cost quantify the importance of the individual parameters on an output of interest, this could be used to compress the parameter space without adversely impacting the accuracy of the prediction of the output and its sensitivity.

To accomplish this we consider the Total Sensitivity Indices (TSI) used in the Fourier Amplitude Sensitivity Test (FAST) and Sobol' method [14, 22, 23]. The TSI of parameter i is defined to measure the combined sensitivity of all terms depending on this particular parameter. To realize how to most efficiently compute this, we define the sensitivity measure

$$S(t) = \frac{V_t}{V}, \quad (5.1)$$

where V_t and V are defined in (4.8). It should be noted that this can be defined based on any output function of interest or on a particular entry in a vector valued function.

Summing up all the terms $S(t)$, $|t| > 0$ we recover

$$\sum_{|t|>0} S(t) = \sum_{|t|>0} \frac{V_t}{V} = 1. \quad (5.2)$$

We now express this as

$$\sum_{i \in t} S(t) + \sum_{i \notin t} S(t) = 1, \quad (5.3)$$

where $i = 1, \dots, p$ is the index of variable α^i . The first term in this expression is TSI(i) of variable α^i while the second term reflect all interactions not involving α^i .

The individual elements in the TSI are computable directly from the truncated ANOVA expansion. However, this may be quite expensive to compute for a high-dimensional case. To overcome this bottleneck, we observe that the use of the TSI is just as an indicator and low accuracy of this will likely be adequate.

We therefore propose an approach in which the ANOVA expansion is first computed for the output of interest using a low order Stroud based integration scheme. This enables the computation of the TSI for the full parameter space at low cost and the identification of the parameters of importance. With this information, we compress the dimensionality of the problem, retaining only the important parameters and freezing less important ones

at their mean value. This results in a compressed system which contain the parameters of dynamic importance. We now proceed and recompute the ANOVA expansion of this problem at a higher accuracy to enable the accurate modeling of the sensitivity of the output of interests. In the following we shall illustrate in more detail the validity of this approach on a non-trivial problem with intuitive behavior.

5.1 An intuitive example

We consider a situation with p particles, each held fixed at a random position in a two-dimensional space $[-a, a]^2$. Let us furthermore assume that a single particle of unit mass is initially positioned at $(0,0)$ and feels an acceleration through Newtonian gravitational forces from all the other particles. This leads to a simple dynamical equation

$$\ddot{\mathbf{x}}(t) = \sum_{i=1}^p \frac{m_i \hat{\mathbf{r}}_i}{r_i^2}, \quad \mathbf{x}(\mathbf{t}_0) = \mathbf{x}_0. \quad (5.4)$$

Here $\hat{\mathbf{r}}_i$ is the distance vector between the fixed particle i and the moving particle and r_i is the Eulerian distance between the fixed particle i and the moving particle.

To endow this problem with a high-dimensional characteristic, we assume that all the masses, m_i , are uniformly distributed random variables with a mean of $1/(p+1)$ and a 10% variance. The goal is to predict the mean trajectory of the moving particle as well as its sensitivity due to the variation in mass.

Intuitive understanding of the problem suggests that just a small number of fixed masses will contribute significantly to the dynamics of the moving particle. Hence, we expect that the parametric compression computed using the TSI approach will work well in this case and identify particles situated close to the moving particle. As the function of interest we consider the kinetic energy but this is not a unique choice.

5.1.1 25 dimensional problem

We first consider a small problem with $p = 25$. In Fig. 2 we show the first part of the approach in which the ANOVA expansion of the kinetic energy for the full problem is computed using the Stroud-3 integration and TSI computed based on that.

We notice in Fig. 2 that including only 2nd order terms suffice in the ANOVA expansion to accurately represent the output function. Furthermore, and as expected, the TSI clearly indicates that only a fraction of the fixed particles are of significant importance for computing the sensitivity.

The next step is to reduce the number of parameters by freezing those of minimal influence at their expectation and the specification of the threshold is a question of judgement. Experimentation has shown that parameters with a TSI of 2% or less can typically be frozen without any substantial effect and we shall use this in what remains. In Fig. 3 we illustrate which particles have been identified by the TSI approach based on the kinetic energy, confirming that it identifies those particles which are closest to the particle track as one would intuitively expect.

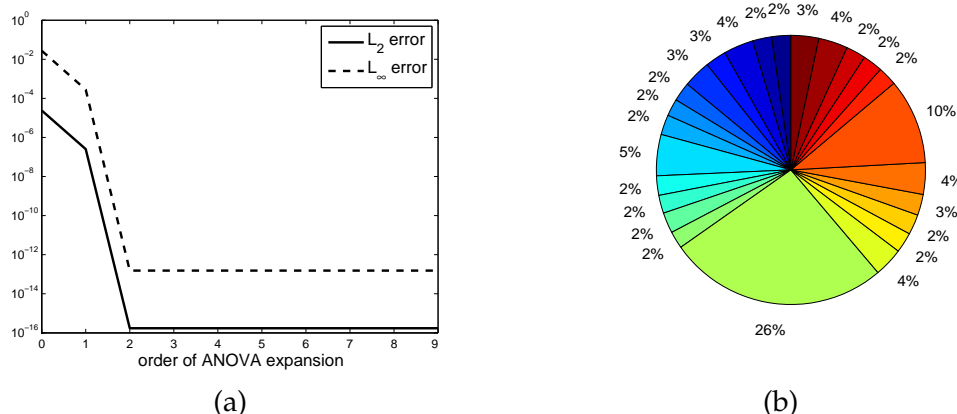


Figure 2: Gravitational problem with $p = 25$, (a): L_2 error and L_∞ error for the ANOVA expansion. (b): TSI for the kinetic energy of the moving particle based on the parametrized masses of the fixed particles.

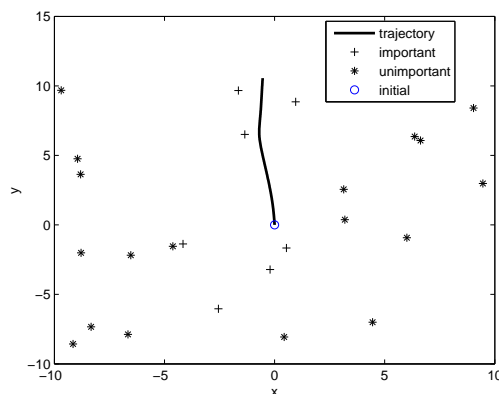


Figure 3: Gravitational problem with $p = 25$. Illustration of the 7 most important particles as identified by the TSI approach.

In Fig. 4 we show the convergence of the ANOVA expansion based on the reduced set of parameters as well as the computed solution and the sensitivity obtained by using the compressed set of parameters. For comparison we also show the result based on the full set of 25 parameters. A third order ANOVA expansion remains sufficient and the agreement between the full problem and the compressed problem is excellent, both for the mean and the sensitivity of the problem. The sensitivities are computed using Monte Carlo in both cases.

To further validate the accuracy of the approach, we show in Fig. 5 a direct comparison between the computed results and those obtained using a direct Monte Carlo approach. The agreement is excellent.

The value at which we choose to truncate the number of parameters based on TSI has several implications. The most immediate one is naturally the accuracy of the re-

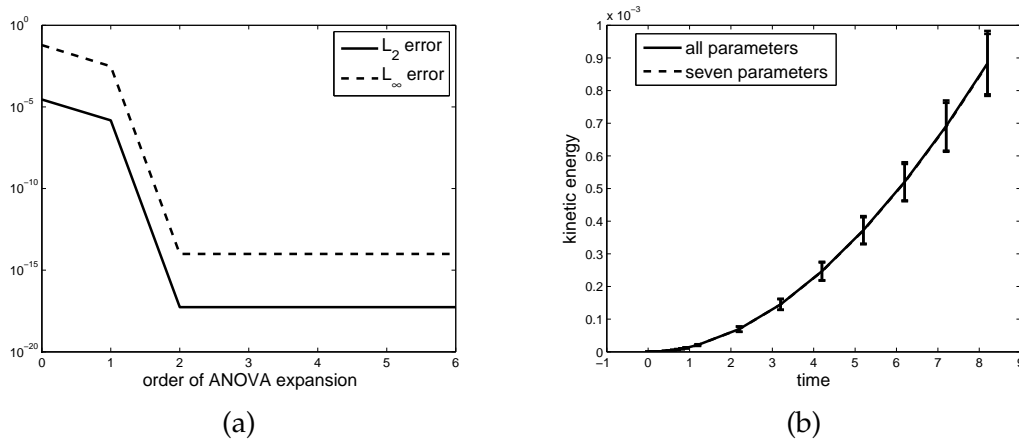


Figure 4: Gravitational problem with $p=25$. (a): We show the accuracy of the ANOVA expansion for the compressed problem based on 7 parameters. (b): Computed solution and sensitivity obtained using the compressed set of parameters as well as the full set of 25 parameters.

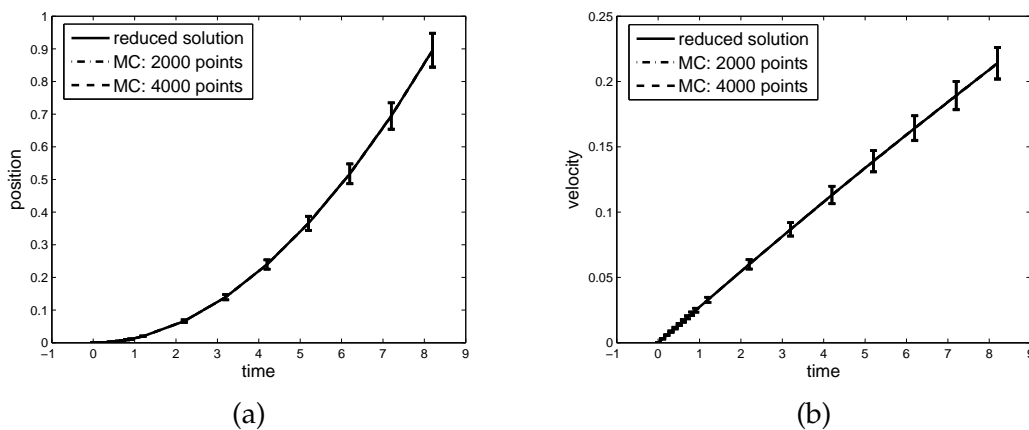


Figure 5: Gravitational problem with $p=25$. We show the computed mean and sensitivity of the position (a) and velocity (b) using different methods as marked.

duced model and for this we find that 2% is a suitable value for all test cases we have considered. However, there may be reasons for having to include additional parameters beyond what is suggested by the cut-off value. Once the parameters are chosen, we build a full response surface following [2] by fitting a second order polynomial to available data through a least squares approximation. Higher order approximations may at times be advantageous but we have not found strong arguments for doing so in the cases considered here.

Once the least squares approximation is built, it can be sampled a low cost and we use this to compute sensitivities and other statistical measures. However, if the parameter space is reduced too aggressively and/or the ANOVA expansion is short, the least

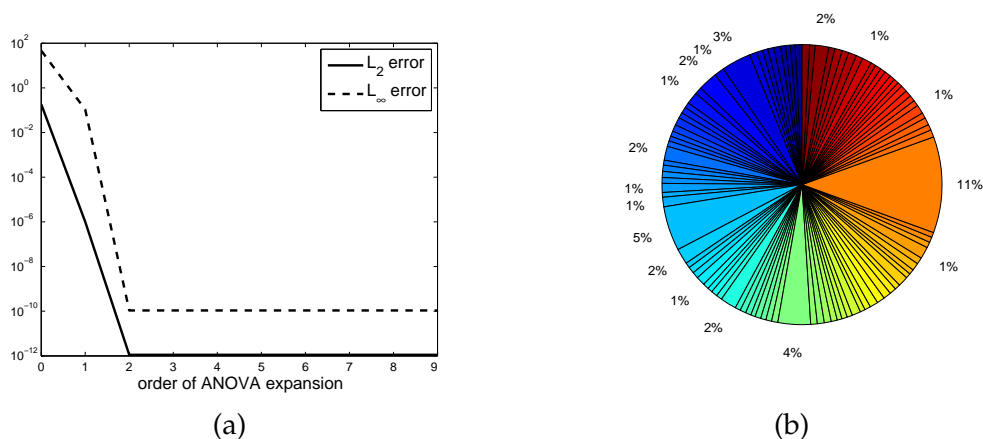


Figure 6: Gravitational problem with $p = 100$, (a): L_2 error and L_∞ error for the ANOVA expansion. (b): Sensitivity index for the kinetic energy of the moving particle based on the parametrized masses of the fixed particles. Unmarked slices contribute less than 1%.

squares approximation may not exist. In such case, we add parameters drawn randomly from those just below the 2% cut-off until the least squares approximation is computable. This typically involves just adding a few additional parameters.

5.1.2 100 dimensional problem

Let us make the problem more challenging and consider a problem with $p = 100$ particles. In Fig. 6 we illustrate the values of the ANOVA expansion of the full problem computed using the Stroud-3 integration and the TSI for the kinetic energy computed based on that.

We notice in Fig. 6 that including all 2nd order terms suffice in the ANOVA expansion to accurately represent the output function. Furthermore, we see clear indications in the TSI that only a fraction of the fixed particles are of significant importance. Using the previously discussed threshold, we find that as little as 10 parameters suffice to accurately compute the dynamics of the moving particle.

In Fig. 7 we show the convergence of the ANOVA expansion based on the reduced set of parameters as well as the computed solution and sensitivity obtained using the compressed set of parameters and the full set of parameters. A third order ANOVA expansion is sufficient and the agreement between the full problem and the compressed problem is excellent, both for the mean and the sensitivity of the problem. The sensitivities are computed using Monte Carlo in both cases.

To further validate the accuracy of the approach, we show in Fig. 8 a direct comparison between the computed results and those obtained using a direct Monte Carlo solution. The agreement remains excellent and supports the validity of this approach for high-dimensional problems.

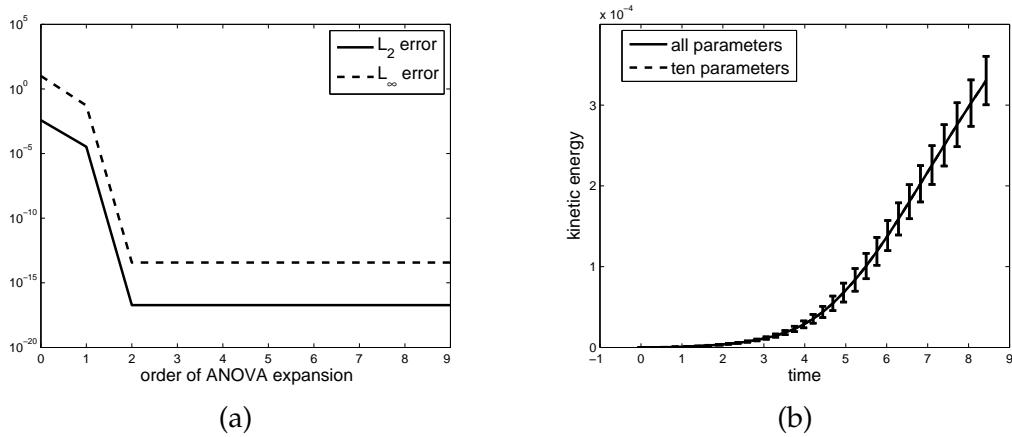


Figure 7: Gravitational problem with $p = 100$. (a): We show the accuracy of the ANOVA expansion for the compressed problem based on 10 parameters. (b): Computed solution and sensitivity obtained using the compressed set of parameters as well as the full set of 100 parameters based on Monte Carlo.

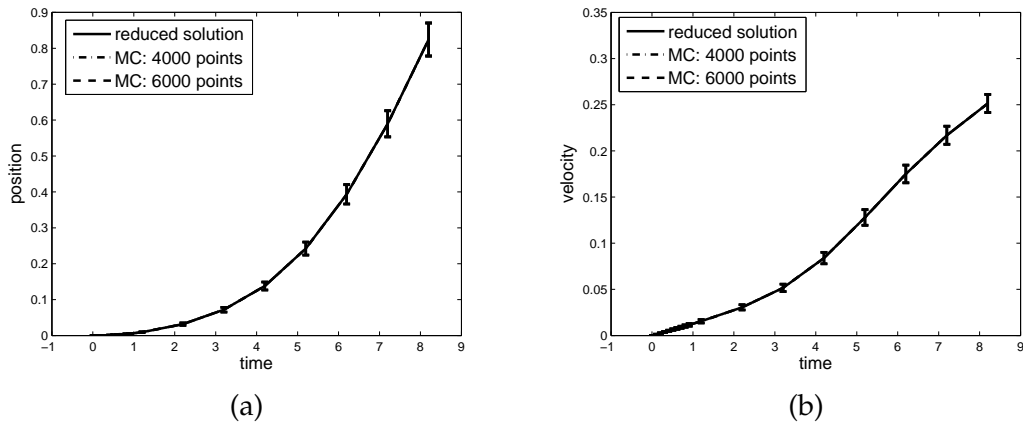


Figure 8: Gravitational problem with $p=100$. We show the computed mean and sensitivity of the position (a) and velocity (b) using different methods as marked.

6 Numerical examples

In the following we evaluate the ANOVA expansion and the approach to parametric compression on two more challenging test cases, both of which has been studied previously, albeit using different techniques.

6.1 Genetic toggle switch

We first consider the genetic toggle switch

$$\frac{du}{dt} = \frac{\alpha_1}{1+v^\beta} - u, \tag{6.1a}$$

$$\frac{dv}{dt} = \frac{\alpha_2}{1 + \omega^\gamma} - v, \quad (6.1b)$$

$$\omega = \frac{u}{(1 + [IPTG]/\mathcal{K})^\eta}, \quad (6.1c)$$

where $\alpha_1, \alpha_2, \beta, \gamma, \eta, \mathcal{K}$ are parameters and $[IPTG]$ is a system input that controls the behavior of the steady state solution. This system of equations describes a genetic switch in *Escherichia coli* [8,30,33].

We model the parameters $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_6) = (\alpha_1, \alpha_2, \beta, \gamma, \eta, \mathcal{K})$ as random variables on the form $\boldsymbol{\alpha} = \langle \boldsymbol{\alpha} \rangle (1 + \sigma y)$, where $\langle \boldsymbol{\alpha} \rangle = (156.25, 15.6, 2.5, 1, 2.0015, 2.9618 \times 10^{-5})$ are the expectation values. The $y = (y^{(1)}, \dots, y^{(6)})$ are uniformly distributed random variables in $[-1, 1]^6$ and $\sigma = 0.1$, i.e., a 10% variation. See [31, 33] for further details.

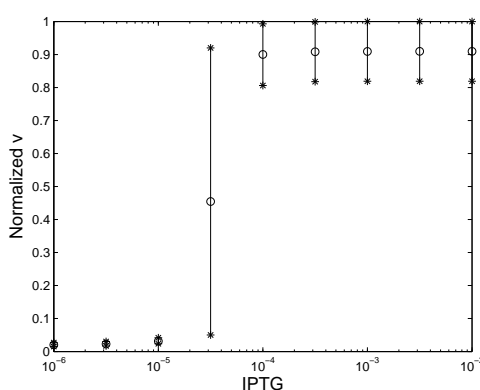


Figure 9: Reference steady state solution of the genetic toggle problem using a 3 level Smolyak sparse grid.

We employ a 4th-order Runge-Kutta method to integrate the system and a Gauss-Patterson based Smolyak sparse integration to obtain a reference solution. Fig. 9 shows the steady state solution and its sensitivity as a function of IPTG. This solution is obtained using 3 levels in the Smolyak grid with 545 function evaluations and have been verified against a 4-level computation with 2561 function evaluations.

The sensitivity of the solution to the value of IPTG is noteworthy and suggests different dynamic behavior away from and close to the critical value. This is clearly confirmed when the TSI is computed for different values of IPTG. In Fig. 10 we show the TSI computed for two values of IPTG with one being very close to the critical value.

The results in Fig. 10 nicely illustrate that only three of the parameters are important for $IPTG = 10^{-6}$ whereas they all enter into the dynamics in the highly sensitive range around $IPTG = 10^{-4.5}$. Depending on the area of interest this provides guidance to a possible parameter reduction. An important observation is, however, that the TSI approach is able to pick up these subtleties of the different regimes at low computational cost.

The computational expense of computing the TSI and, ultimately, evaluating the output of interest depends on the efficiency of the ANOVA expansion. Indeed, one could be concerned that the truncation dimension depends on the value of $IPTG$.

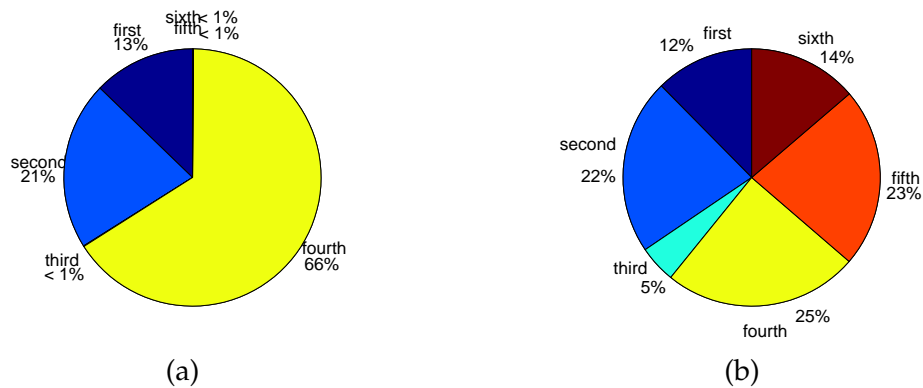


Figure 10: Total sensitivity index of each random variable. (a): $IPTG=10^{-6}$. (b): $IPTG=10^{-4.5}$.

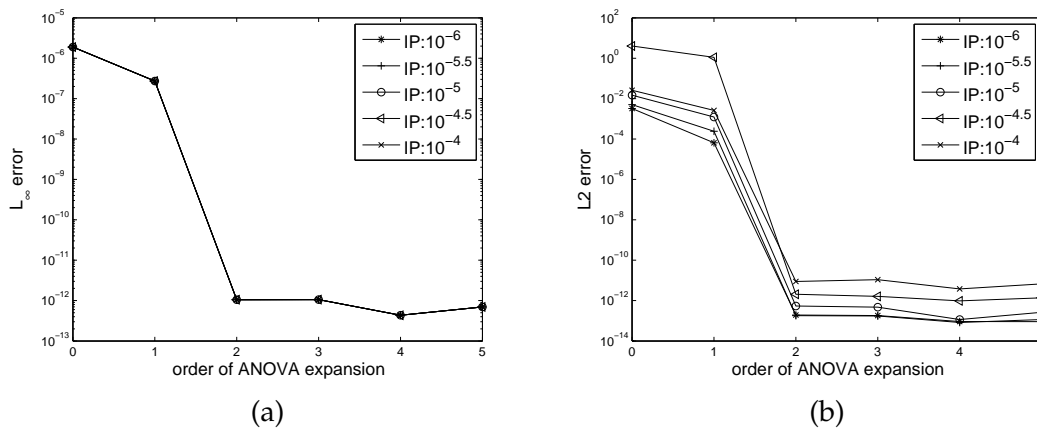


Figure 11: The numerical results computed from ANOVA expansion for the first five IPTG. (a): L_∞ error. (b): L_2 error.

In Fig. 11 we show the L_2 and L_∞ errors for the ANOVA expansion computed using all six variables for different values of $IPTG$. The results confirm the efficiency of the ANOVA expansion, requiring only 2nd order terms and a total of 21 terms, and also the insensitivity of the truncation dimension to the value of $IPTG$. The combination of this and the TSI hence provides an efficient and accurate way of dealing with this otherwise challenging problem.

6.2 Pollution problem

We next consider a pollution problem, developed by The Dutch National Institute of Public Health and Environmental Protection (RIVM). It is a chemical model consisting of 25 reactions and 20 reacting compounds. We refer to [16, 27] for detailed information regarding the identification of variables with species and the reaction scheme.

The problem is given in a 20 dimensional space $\mathbf{u} \in \mathbb{R}^{20}$ and with a right hand side on

the form

$$\mathbf{f}(\mathbf{u}) = \begin{cases} -\sum_{j \in \{1,10,14,23,24\}} r_j + \sum_{j \in \{2,3,9,11,12,22,25\}} r_j, \\ -r_2 - r_3 - r_9 - r_{12} + r_1 + r_{21}, \\ -r_{15} + r_1 + r_{17} + r_{19} + r_{22}, \\ -r_2 - r_{16} - r_{17} - r_{23} + r_{15}, \\ -r_3 + 2 \cdot r_4 + r_6 + r_7 + r_{13} + r_{20}, \\ -r_6 - r_8 - r_{14} - r_{20} + r_3 + 2 \cdot r_{18}, \\ -r_4 - r_5 - r_6 + r_{13}, \\ r_4 + r_5 + r_6 + r_7, \\ -r_7 - r_8, \\ -r_{12} + r_7 + r_9, \\ -r_9 - r_{10} + r_8 + r_{11}, \\ r_9, \\ -r_{11} + r_{10}, \\ -r_{13} + r_{12}, \\ r_{14}, \\ -r_{18} - r_{19} + r_{16}, \\ -r_{20}, \\ r_{20}, \\ -r_{21} - r_{22} - r_{24} + r_{23} + r_{25}, \\ -r_{25} + r_{24}. \end{cases} \quad (6.2)$$

The auxiliary variables, r_i , are connected to the state variables as defined in Table 1 and the initial condition is prescribed as

$$\mathbf{u}(0) = (0, 0.2, 0, 0.04, 0, 0, 0.1, 0.3, 0.01, 0, 0, 0, 0, 0, 0, 0, 0, 0.007, 0, 0, 0)^T.$$

The auxiliary variables in Table 1 depends on 25 coefficients, k_i , with mean values given in Table 2.

In the following we assume that all 25 parameters are uncertain as $k_i = \langle k_i \rangle (1 + \sigma y_i)$, where $\sigma = 0.1$, and y_i are uniformly distributed independent random variables in $[-1, 1]$.

Table 1: Auxiliary variables for the pollution problem.

$r_1 = k_1 \cdot u_1$	$r_{10} = k_{10} \cdot u_1 \cdot u_{11}$	$r_{19} = k_{19} \cdot u_{16}$
$r_2 = k_2 \cdot u_2 \cdot u_3$	$r_{11} = k_{11} \cdot u_{13}$	$r_{20} = k_{20} \cdot u_6 \cdot u_{17}$
$r_3 = k_3 \cdot u_2 \cdot u_5$	$r_{12} = k_{12} \cdot u_2 \cdot u_{10}$	$r_{21} = k_{21} \cdot u_{19}$
$r_4 = k_4 \cdot u_7$	$r_{13} = k_{13} \cdot u_{14}$	$r_{22} = k_{22} \cdot u_{19}$
$r_5 = k_5 \cdot u_7$	$r_{14} = k_{14} \cdot u_1 \cdot u_6$	$r_{23} = k_{23} \cdot u_1 \cdot u_4$
$r_6 = k_6 \cdot u_6 \cdot u_7$	$r_{15} = k_{15} \cdot u_3$	$r_{24} = k_{24} \cdot u_1 \cdot u_{19}$
$r_7 = k_7 \cdot u_9$	$r_{16} = k_{16} \cdot u_4$	$r_{25} = k_{25} \cdot u_{20}$
$r_8 = k_8 \cdot u_6 \cdot u_9$	$r_{17} = k_{17} \cdot u_4$	
$r_9 = k_9 \cdot u_2 \cdot u_{11}$	$r_{18} = k_{18} \cdot u_{16}$	

Table 2: Mean values of parameters in the pollution problem.

$k_1=0.350$	$k_{10}=0.900 \cdot 10^4$	$k_{19}=0.444 \cdot 10^{12}$
$k_2=0.266 \cdot 10^2$	$k_{11}=0.220 \cdot 10^{-1}$	$k_{20}=0.124 \cdot 10^4$
$k_3=0.123 \cdot 10^5$	$k_{12}=0.120 \cdot 10^5$	$k_{21}=0.210 \cdot 10$
$k_4=0.860 \cdot 10^{-3}$	$k_{13}=0.188 \cdot 10$	$k_{22}=0.578 \cdot 10$
$k_5=0.820 \cdot 10^{-3}$	$k_{14}=0.163 \cdot 10^5$	$k_{23}=0.474 \cdot 10^{-1}$
$k_6=0.150 \cdot 10^5$	$k_{15}=0.480 \cdot 10^7$	$k_{24}=0.178 \cdot 10^4$
$k_7=0.130 \cdot 10^{-5}$	$k_{16}=0.350 \cdot 10^{-3}$	$k_{25}=0.312 \cdot 10$
$k_8=0.240 \cdot 10^5$	$k_{17}=0.175 \cdot 10^{-1}$	
$k_9=0.165 \cdot 10^5$	$k_{18}=0.100 \cdot 10^9$	

To demonstrate the efficiency of the proposed approach, we randomly select $u_8(t)$ as the output of interest. The first step is to compute an approximation of the ANOVA expansion in the full 25-dimensional space and use this to recover the TSI. This is illustrated in Fig. 12 where we observe that we can safely use a truncation dimension of only two in the ANOVA expansion. Using this to compute the associated TSI shows that only a small number of the random variables are important; these can be identified as being $(k_2, k_3, k_4, k_6, k_{16}, k_{19})$. We subsequently keep these as random variables and freeze the remaining 19 variables at their expectation value.

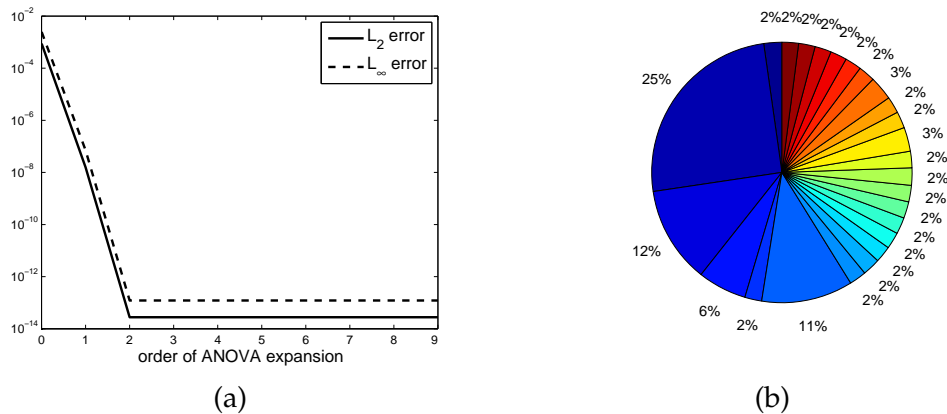


Figure 12: Results for the analysis of u_8 , (a): L_2 error and L_∞ error for the full ANOVA expansion. (b): The TSI for the parameter space for u_8 .

We proceed by computing the ANOVA expansion for the reduced parametric function and use this to predict the output of interest. The results are shown in Fig. 13, illustrating the efficiency of the ANOVA expansion on the reduced parameter space, requiring only a 2nd order expansion, and the accuracy of prediction of the reduced model, obtained by using only the six parameters identified as being important.

We finally show in Fig. 14 the computational solution of u_8 compared to Monte Carlo results, confirming the accuracy of the solution for both expectation and sensitivity results.

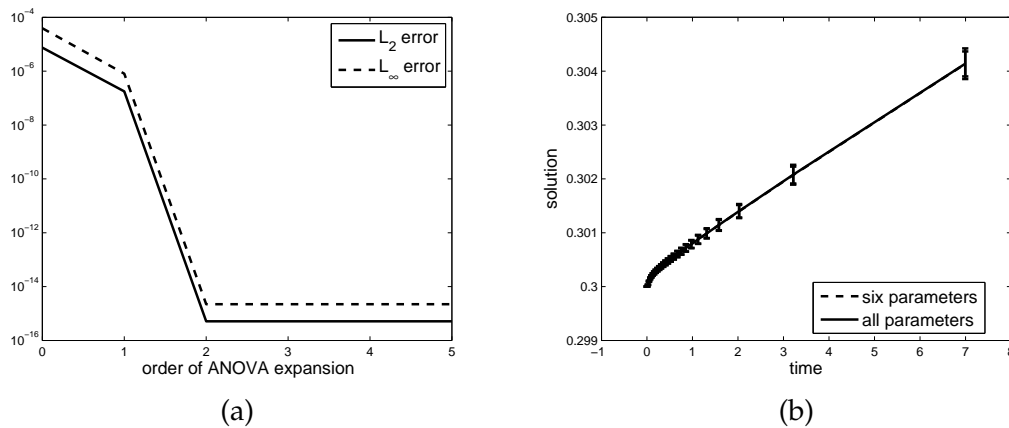


Figure 13: On the left we show the convergence of the ANOVA expansion of the compressed parametrized problem, while the right shows the solution obtained using both the full system and the compressed set of parameters.

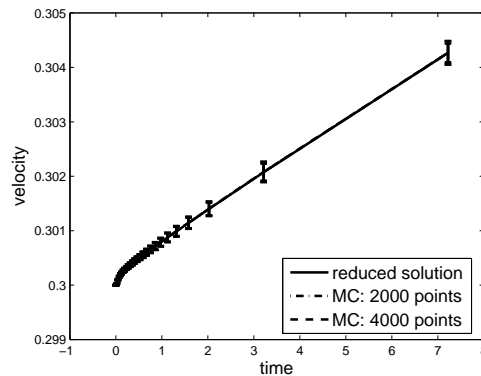


Figure 14: Computational solution and sensitivities of u_8 computed using the reduced approach and a Monte Carlo method.

To illustrate the importance of performing this analysis if a new output function is chosen, we show in Fig. 15 the TSI computed for the two variables, u_{14} and u_{17} . The analysis not only highlights that the important parameters change, but also that the actual number of parameters required may change, i.e., for u_{14} two parameters suffice while for u_{17} at least 8 parameters are needed.

The ANOVA expansions for the reduced parameter space have truncation dimension of two for both variables, resulting in a highly compressed approach to compute the solutions, shown in Fig. 16 where we show the computed solutions and their sensitivities, confirming the validity of the approach with a reduced parameter space.

To validate the accuracy of the overall approach, we show in Fig. 17 a direct comparison with the computed results and those obtained with a Monte Carlo approach. The agreement is excellent and the results in Fig. 16 are obtained at a fraction of the overall computational time.

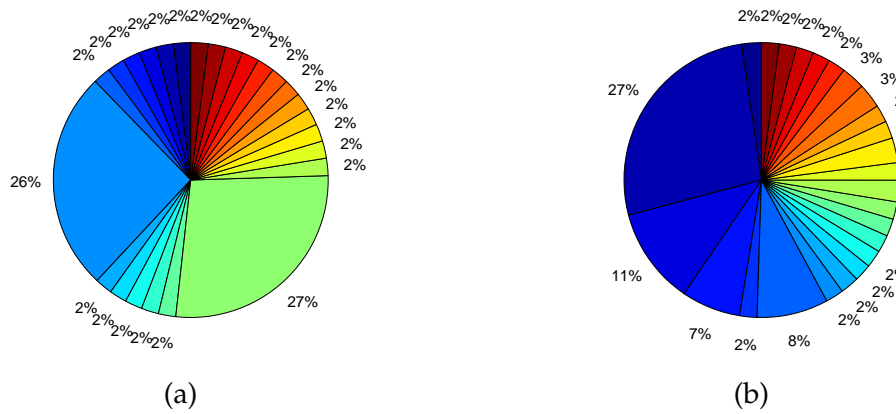


Figure 15: (a): TSI for u_{14} . (b): TSI for u_{17} .

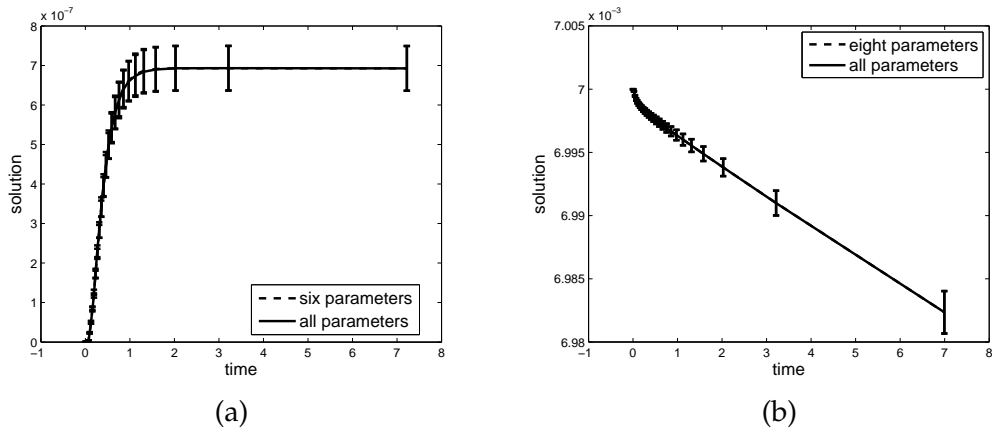


Figure 16: On the left is shown the computed solution and the sensitivity of it for u_{14} . A similar result is shown on the right for u_{17} .

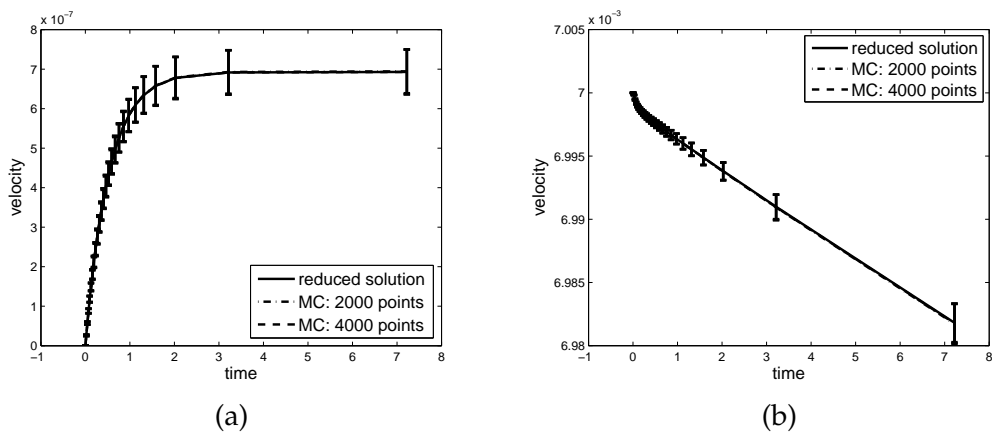


Figure 17: Computational solution and sensitivities of u_{14} (a) and u_{17} (b) computed using the reduced approach and a Monte Carlo method.

7 Concluding remarks

The goal of this paper has been to present a systematic approach to accurately and efficiently model the sensitivity of parametrized differential equations with a high degree of uncertainty associated with the parameters. While a straightforward approach is possible in principle, the curse of dimensionality requires one to explore alternative ways to deal with problems of this nature.

We have shown, through a number of problems of increasing complexity, that a combination of a low cost approximate ANOVA expansions of output functions of interest, to evaluate importance of the individual parameters through a sensitivity index analysis, followed by a more accurate ANOVA expansion of the reduced problems offers a viable and accurate approach. Since the first ANOVA expansion is to understand the parametric sensitivity, this does not have to be computed accurately and a Stroud based integration approach suffices. We also showed that the total sensitivity index is sensitive enough to identify important parameters even in highly sensitive and very dynamic parts of parameter space and that it correctly identifies the important parameters for different outputs of interest. Once the important parameters have been identified, all other are frozen at the mean value and the reduced system can be analyzed more accurately at reduced cost. Numerous examples have shown the efficiency and accuracy of this general approach.

The focus in this work has been on ordinary differential equations with high-dimensional parametric uncertainty. A natural next step is to consider the use of similar techniques for partial differential equations where the reduction of computational complexity is even more important and the dimensionality of the problem often much higher. However, in a method-of-lines approach, one recovers a large coupled system of ordinary differential equations after spatial discretization and we are optimistic that the ideas presented here will transfer to this case. We hope to report on this in the near future.

Acknowledgments

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

References

- [1] R. E. Caflisch, W. Morokoff, and A. Owen, Valuation of mortgage backed securities using Brownian bridges to reduce effective dimension, *J. Comput. Finance.*, 1 (1997), 27–46.
- [2] Y. Cao, Z. Chen, and M. Gunzburger, ANOVA expansions and efficient sampling methods for parameter dependent nonlinear PDEs, *Int. J. Numer. Anal. Model.*, 6 (2009), 256–273.
- [3] C. W. Clenshaw, and A. R. Curtis, A method for numerical integration on an automatic computer, *Numer. Math.*, 2 (1960), 197.

- [4] P. J. Davis, and P. Rabinowitz, *Methods of Numerical Integration*, Academic Press, NY, 1975.
- [5] G. Fishman, *Monte Carlo: Concepts, Algorithms, and Applications*, Springer-Verlag, New York, 1996.
- [6] B. Fox, *Strategies for Quasi-Monte Carlo*, Kluwer, Dordrecht, The Netherlands, 1999.
- [7] D. Gamerman, *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*, Chapman and Hall, London, 1997.
- [8] T. Gardner, C. Cantor, and J. Collins, Construction of a genetic toggle switch in *Escherichia coli*, *Nature.*, 403 (2000), 339–342.
- [9] A. C. Genz, Testing multidimensional integration routines, in: *Tools, Methods, and Languages for Scientific and Engineering Computation*, eds. B. Ford, J. C. Rault, and F. Thomasset (North-Holland, Amsterdam, 1984), 81–94.
- [10] A. C. Genz, A package for testing multiple integration subroutines, in: *Numerical Integration*, eds. P. Keast, and G. Fairweather (Kluwer, Dordrecht, 1987), 337–340.
- [11] T. Gerstner, and M. Griebel, Numerical integration using sparse grid, *Numer. Algor.*, 18 (1998), 209–232.
- [12] M. B. Giles, and E. Suli, Adjoint methods for PDEs: a posteriori error analysis and postprocessing by duality, *Acta. Numer.*, 11 (2002), 145–206.
- [13] C. Gu, *Smoothing Spline ANOVA Models*, Springer, Berlin, 2002.
- [14] T. Homma, and A. Saltelli, Importance measures in global sensitivity analysis of nonlinear models, *Reliab. Eng. Syst. Safe.*, 52 (1996), 1–17.
- [15] M. Liu, Z. Gao, and J. S. Hesthaven, Adaptive sparse grid algorithms with applications to electromagnetic scattering under uncertainty, *Appl. Numer. Math.*, 2010, submitted.
- [16] F. Mazzia, and C. Magherini, Test Set for Initial Values Problem Solvers, Release 2.4, Department of Mathematics, University of Bari and INdAM, Research Unit of Bari, February 2008.
- [17] A. B. Owen, The dimension distribution and quadrature test functions, Technical Report, Stanford University, 2001.
- [18] S. Paskov, and J. Traub, Faster valuation of financial derivatives, *J. Portfolio. Manag.*, 22 (1995), 113–120.
- [19] L. N. Trefethen, Is Gauss quadrature better than Clenshaw-Curtis?, *SIAM Rev.*, 50 (2008), 67–87.
- [20] A. Saltelli, K. Chan, and E. Scott, *Sensitivity Analysis*, Wiley & Sons, Chichester, 2000.
- [21] S. A. Smolyak, Quadrature and interpolation formulas for tensor products of certain classes of functions, *Dokl. Akad. Nauk. SSSR.*, 4 (1963), 240–243.
- [22] I. M. Sobol', Sensitivity estimates for nonlinear mathematical models, *Math. Model.*, 2 (1990), 112–118 (in Russian).
- [23] I. M. Sobol', Sensitivity estimates for nonlinear mathematical models, *Math. Model. Comput. Exp.*, 1 (1993), 407–414.
- [24] M. Stein, Large sample properties of simulations using Latin hypercube sampling, *Technometrics*, 29 (1987), 143–151.
- [25] A. Stroud, Remarks on the disposition of points in numerical integration formulas, *Math. Comput.*, 11 (1957), 257–261.
- [26] 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. Geophys. Res.*, 102 (1997), 21925–21932.
- [27] J. G. Verwer, Gauss-Seidel iteration for stiff ODEs from chemical kinetics, *SIAM J. Sci. Comput.*, 15 (1994), 1243–1259.

- [28] X. Wang, and K. Fang, The effective dimension and quasi-Monte Carlo integration, *J. Complex.*, 19 (2003), 101–124.
- [29] G. W. Wasilakowski, and H. Wozniakowski, Explicit cost bounds of algorithms for multivariate tensor product problems, *J. Complex.*, 11 (1995), 1–56.
- [30] D. Xiu, and J. S. Hesthaven, High-order collocation methods for differential equations with random inputs, *SIAM J. Sci. Comput.*, 27 (2005), 1118–1139.
- [31] D. Xiu, Efficient collocational approach for parametric uncertainty analysis, *Commun. Comput. Phys.*, 2 (2007), 293–309.
- [32] D. Xiu, Numerical integration formula of degree two, *Appl. Numer. Math.*, 58 (2008), 1515–1520.
- [33] D. Xiu, Fast numerical methods for stochastic computations: a review, *Commun. Comput. Phys.*, 5 (2009), 242–272.