

On ANOVA expansions and strategies for choosing the anchor point

Zhen Gao* and Jan S. Hesthaven†

August 31, 2010

Abstract

The classic Lebesgue ANOVA expansion offers an elegant way to represent functions that depend on a high-dimensional set of parameters and it often enables a substantial reduction in the evaluation cost of such functions once the ANOVA representation is constructed. Unfortunately, the construction of the expansion itself is expensive due to the need to evaluate high-dimensional integrals. A way around this is to consider an alternative formulation, known as the anchored ANOVA expansion. This formulation requires no integrals but has an accuracy that depends sensitively on the choice of a special parameter, known as the anchor point.

We present a comparative study of several strategies for the choice of this anchor point and argue that the optimal choice of this anchor point is the center point of a sparse grid quadrature. This choice comes at no cost and, as we shall show, results in a natural truncation of the ANOVA expansion. The efficiency and accuracy is illustrated through several standard benchmarks and is shown to outperform the alternatives over a range of applications.

Key words: ANOVA expansion, Anchored ANOVA expansion, Sparse Grid, High-Dimensional Integrals

*Research Center for Applied Mathematics, Ocean University of China, Qingdao, 266071, PRC & Division of Applied Mathematics, Brown University, Providence, 02912, USA. Email: zhen_gao@brown.edu

†Correspondence author: Division of Applied Mathematics, Brown University, Providence, 02912, USA. Email: Jan.Hesthaven@brown.edu

1 Introduction

The analysis-of-variation - ANOVA - expansion provides an elegant way to represent functions that depends on a high-dimensional set of parameters. As such it has been used in numerous applications during the last decade to represent and efficiently manipulate high-dimensional problems and to enable one to take advantage of the inherent low-dimensional interdependence, often found in many such problems. In [9] it was explored in the context of high-dimensional integration methods, in [2] it was demonstrated in relation with parameterized partial differential equations and in [5] the ANOVA expansion was utilized to develop a sensitivity index to enable the effective reduction of parametric dimensionality without impacting the accuracy of the predicted output function.

However, the classic ANOVA expansion is projection based and its construction requires the use of high-dimensional integration, rendering its construction very expensive. To address this challenge, an alternative formulation, named the anchored or Dirac ANOVA expansion, has been proposed [9]. It was also considered in [7] under the name CUT-HDMR. It relies on expressing a function $u(\alpha)$ as a superposition of its values along lines, planes and hyperplane passing through an anchor point $\beta = (\beta_1, \dots, \beta_p)$. As can be imagined, the choice of this anchor point is closely tied to the overall efficiency and accuracy of the expansion and making this choice correctly becomes a key element of the formulation. Unfortunately, there is no known rigorous result of how to make this choice in a optimal way for general functions.

Recently, a number of techniques for making this choice have been proposed. A straightforward choice is to use an anchor point chosen randomly in the high-dimensional space. While used widely it can not be expected to yield an optimal choice. In [17], it is suggested that the optimal anchor point is found as the trial point whose output is closest to the mean of the function, being computed from a moderate number of quasi-random samples. An alternative, based on ideas of optimal weights in quasi Monte Carlo methods, is proposed in [20] and shown to yield good results. Unfortunately, this approach is only rigorous for functions that allow a dimensional variable separation. Both of these latter methods require some computational work to identify the anchor point.

In this work, we propose to use the center point of a particular sparse grid quadrature as the anchor point and offer a comparative study with the previously introduced methods

mentioned above. An argument for this new approach is based on the structure of the Smolyak sparse grid which is closely related to the anchored ANOVA expansion [9]. This anchor point can be computed a minimal cost and we show that its use leads to a a very natural truncation of the anchored ANOVA expansion when one is computing integrals of the expansion. While most past work have assumed that the parameters are uniformly distributed random variables, we also discuss the use of this approach when applied to cases where the parameters are more general non-uniformly distributed random variables.

What remains of the paper is organized as follows. Section 2 introduces the ANOVA expansion based on the Lebesgue measure and the Dirac measure, respectively. We also discuss the structure of the Smolyak sparse grid in this part. Section 3 introduces four strategies for the choice of the anchor point and in Section 4 we demonstrate the efficiency and accuracy of the proposed anchor point through several examples. Section 5 contains a few concluding remarks.

2 The ANOVA expansion

We begin by introducing the ANOVA expansion and its two different representations based on different product measures. Without loss of generality, we take the integration domain \mathbb{D} to be $[0, 1]^p$, and $u \in \mathcal{L}^2(\mathbb{D})$. Take t to be any subset of coordinate indices $\mathcal{P} = \{1, \dots, p\}$ with $|t|$ denoting the cardinality of t . Let also α^t denote the $|t|$ -vector that contains the components of the vector $\alpha \in [0, 1]^{|t|}$ indexed by t and take $A^{|t|}$ to denote the $|t|$ -dimensional unit hypercube defined as the projection of the p -dimensional unit hypercube A^p onto the hypercube indexed by t . Assume $d\mu$ to be a probability measure on A^p . Then u can be expanded in an ANOVA expansion as [7, 17]

$$u(\alpha) = u_0 + \sum_{t \subseteq \mathcal{P}} u_t(\alpha_t), \quad (1)$$

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

$$u_t(\alpha_t) = \int_{A^{p-|t|}} u(\alpha) d\mu(\alpha_{\mathcal{P} \setminus t}) - \sum_{w \subset t} u_w(\alpha_w) - u_0 \quad (2)$$

starting with

$$u_0 = \int_{A^p} u(\alpha) d\mu(\alpha), \quad \int_{A^0} u(\alpha) d\mu(\alpha_\emptyset) = u(\alpha). \quad (3)$$

Here $d\mu(\alpha_{\mathcal{P}\setminus t})$ indicates integration over the coordinates except indices containing 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 [7, 17]. Furthermore, the individual terms in the expansion are mutually orthogonal, i.e.

$$\int_{A^p} u_t(\alpha_t) u_w(\alpha_w) d\mu(\alpha) = \delta_{tw} \quad (4)$$

and, as a natural consequence of this, each term except u_0 has a zero mean

$$\int_{A^p} u_t(\alpha_t) d\mu(\alpha) = 0, \quad |t| > 0. \quad (5)$$

The computational realization of the ANOVA expansion is achieved through the recursive expression, (2), and the use of orthogonality (4).

2.1 The Lebesgue expansion

In the classic ANOVA expansion, one assumes $d\mu$ is a Lebesgue measure in Eq. (1) and Eqs.(2)-(3) yield its realization through high dimensional integration.

Let us define the truncated ANOVA expansion of order s as

$$u(\alpha; s) = u_0 + \sum_{t \subseteq \mathcal{P}, |t| \leq s} u_t(\alpha_t). \quad (6)$$

where $u_t(\alpha_t)$ and u_0 are defined above.

The concept of an effective dimension of a particular integrand was introduced in [11, 12] and also discussed in [13] as a way to reflect and utilize the observation that many high-dimensional functions are effectively low-dimensional. It was also observed that the ANOVA expansion was particularly well suited from bringing out this hidden low dimensional nature.

The effective dimension is the smallest integer p_s such that

$$\sum_{0 < |t| \leq p_s} V_t(u) \geq qV(u), \quad (7)$$

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

$$V_t(u) = \int_{A^p} (u_t(\alpha_t))^2 d\alpha, \quad V(u) = \sum_{|t|>0} V_t(u), \quad (8)$$

and can be thought of as a measure of the variability of u when considering a given set t .

The relationship between the accuracy of the truncated ANOVA expansion and the superposition dimension is made clear through the following result [15, 16, 19]

Theorem 1 *Assume that the function $u(\alpha)$ has superposition dimension p_s based on q and let $u(\alpha; p_s)$ denote the truncated ANOVA expansion of order p_s . Then*

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

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

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

This shows 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 allows one to reduce the cost of computing the expansion and reduce the cost of the subsequent evaluation of the expansion.

2.1.1 Sparse Smolyak grids

To control the computational cost of evaluating the required high-dimensional integrals, (2) (3), a high-dimensional efficient quadrature rule need to be considered. Here we use sparse grid methods based on the Smolyak construction [14]. These allow one to construct sparse multivariate quadrature formulas based on sparse tensor products of one-dimensional quadrature formulas.

Let us consider the numerical integration of a function $u(\alpha)$ over a p -dimensional unit hypercube $A^p = [0, 1]^p$,

$$I[u] := \int_{A^p} u(\alpha) d\alpha. \quad (9)$$

To introduce the algorithm, we choose a one-dimensional quadrature formula for a univariate function u as

$$Q_l^1 u = \sum_{i=1}^{n_l^1} \omega_i u(\alpha_i^1) \quad (10)$$

where ω_i represent the integration weights and α_i^1 reflect the quadrature points.

Now define a sequence

$$\Delta_i^1 u = (Q_i^1 - Q_{i-1}^1)u \quad (11)$$

with $Q_0^1 u = 0$ and for $i \in N_+$. Smolyak's algorithm for the p -dimensional quadrature formula is then given as

$$Q_l^p u = \sum_{|\mathbb{k}|_1 \leq l+p-1} (\Delta_{k_1}^1 \otimes \cdots \otimes \Delta_{k_p}^1)u \quad (12)$$

for $l \in n$ and $\mathbb{k} = (k_1, \dots, k_p) \in n^p$. An alternative form of this last expression is

$$Q_l^p u = \sum_{l \leq |\mathbb{k}|_1 \leq l+p-1} (-1)^{l+p-|\mathbb{k}|_1-1} \binom{p-1}{|\mathbb{k}|_1-l} (Q_{k_1}^1 \otimes \cdots \otimes Q_{k_p}^1)u. \quad (13)$$

For other equivalent expressions, see [18].

Equation (13) clearly only depends on function values at a finite number of points. To highlight the structure of the quadrature points, let

$$\alpha^{k_i} = \{\alpha_1^{k_i}, \dots, \alpha_{n_1^1}^{k_i}\} \subset [0, 1] \quad (14)$$

denote the one-dimensional quadrature grid corresponding to $Q_{k_i}^1 u$, $1 \leq k_i \leq p$. The tensor product in Eq. (13) depends on $\alpha^{k_1} \times \cdots \times \alpha^{k_p}$ and the union of these

$$\Omega_l^p = \bigcup_{l \leq |\mathbb{k}|_1 \leq l+p-1} (\alpha^{k_1} \times \cdots \times \alpha^{k_p}). \quad (15)$$

is called the sparse grid, used to evaluate (13). If α^k is a nested set, $\Omega_l^p \subset \Omega_{l+1}^p$ and Eq. (15) simplifies

$$\Omega_l^p = \bigcup_{|\mathbb{k}|_1 = l+p-1} (\alpha^{k_1} \times \cdots \times \alpha^{k_p}), \quad (16)$$

which is more compact than Eq. (15). In this work we use a sparse grid based on the Gauss-Patterson quadrature points when possible. This is hierarchical and the most efficient approach when one considers attainable accuracy for a given computational cost [6, 8].

To illustrate the efficiency and accuracy of the Lebesgue ANOVA expansion and the concept of the effective dimension, we consider a p -dimensional oscillatory function,

$$u_1(\alpha) = \cos(2\pi\omega_1 + \sum_{i=1}^p c_i\alpha_i). \quad (17)$$

proposed in [22, 23] as a suitable test function for high-dimensional integration schemes. Both c_i and ω_1 are generated as random numbers and we consider $p = 10$ as a test case.

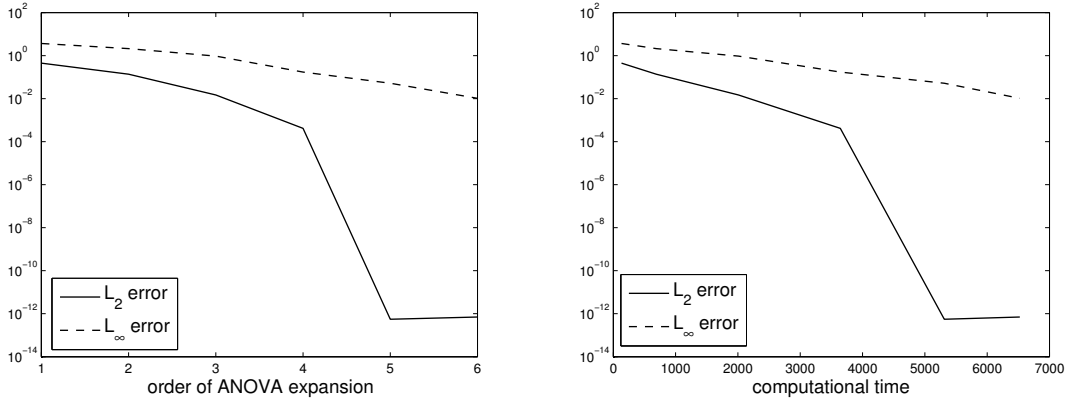


Figure 1: On the left, we show the L_2 and the L_∞ errors of the 6th-order truncated Lebesgue ANOVA expansion with increasing number of terms. The right shows the associated computational time.

Figure 1 shows the accuracy and the computational cost of the Lebesgue ANOVA expansion measured in both the L_2 norm and the L_∞ norm. Clearly, the 4th-order truncated expansion can represent the function well down to an accuracy below 10^{-10} . However, this accuracy comes at considerable computational cost due to the need to evaluate the high dimensional integrals.

2.2 The Dirac expansion

Now assume that $d\mu$ is a Dirac measure located at the anchor point $\beta = (\beta_1, \beta_2 \cdots \beta_p) \in [0, 1]^p$. This leads to what is known as the anchored or the Dirac ANOVA expansion.

The recursive formula Eq. (2) and the initial formula Eq. (3) now takes the forms

$$\begin{aligned}
 u_t(\alpha_t) &= u(\beta_1, \dots, \beta_{i_1-1}, \alpha_1, \beta_{i_1+1}, \dots, \beta_{i_2-1}, \alpha_2, \beta_{i_2+1}, \dots, \beta_{i_{|t|}-1}, \alpha_{|t|}, \beta_{i_{|t|}+1}, \dots, \beta_p) \\
 &- \sum_{w \subset t} u_w(\alpha_w) - u_0,
 \end{aligned}
 \tag{18}$$

and

$$u_0 = u(\beta_1, \beta_2 \dots \beta_p).
 \tag{19}$$

The computational realization of the anchored ANOVA expansion is considerably more efficient than the Lebesgue ANOVA expansion as there is no need to evaluate high-dimensional integrals in Eqs. (18)-(19).

To illustrate this representation, let us again consider the example in Eq. (17). In Fig. 2.2 we illustrate that errors are again reduced to below 10^{-12} with the 4th-order anchored ANOVA expansion with the anchor point taken to be $(0, 0, \dots, 0)$. With a comparable accuracy, the anchored ANOVA expansion is achieved at a fraction of the time required for the classic ANOVA expansion. For higher dimensional problems, the gain can be expected to be even more significant.

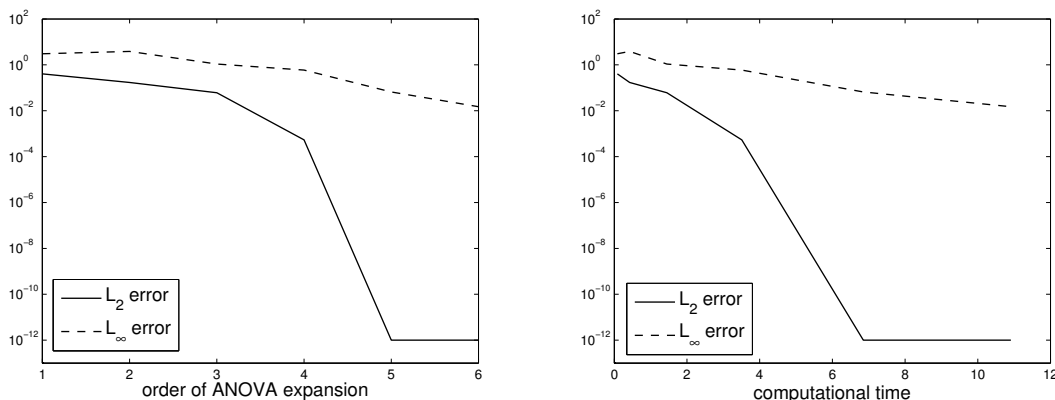


Figure 2: On the left, we show the L_2 and the L_∞ errors of the 6th-order truncated anchored ANOVA expansion with an increasing number of terms. The right shows the associated computational time.

3 Strategies for choosing the anchor point

A key element in the anchored ANOVA expansion is the choice of the anchor point as this directly impacts the accuracy of the ANOVA expansion and the truncation dimension and, hence, the total cost of evaluating the expansion.

A number of strategies have been proposed for choosing the anchor point and we will briefly summarize these below before arguing for an alternative approach.

A simple approach is to randomly choose a point as the anchor point. This is clearly straightforward and with negligible cost. However, there are no guarantees for the accuracy of this approach and, as we shall see, it generally leads to an ANOVA expansion of poor quality,

In [17] it was suggested to choose an anchor point based on a moderate number of function evaluations to estimate the mean, denoted \bar{u} , through a number of quasi-random trial points in $[0, 1]^p$. The anchor point is then chosen to be the trial point whose output is closest to the mean of the function. This guarantees that the zero order term approximates the function as accurate as possible but does not offer any guarantees for the quality of the higher order terms. While there is a cost associated with the computation of the anchor point through the sampling, an obvious advantage is that this generalizes to the case of non-uniformly distributed parameters. In the following we shall refer to the mean anchor point as one chosen using this approach.

In [20], an alternative approach for choosing the anchor point for a more restricted class of problem of the type

$$u(\alpha) = \prod_{j=1}^p u_j(\alpha_j), \quad (20)$$

was developed. This technique, based on analysis borrowed from quasi Monte Carlo methods, is expressed by defining the dimensional weights γ_j , $j = 1, \dots, p$, as

$$\gamma_j = \frac{\|u_j - u_j(\beta_j)\|_\infty}{|u_j(\beta_j)|}, \quad u(\beta) \neq 0. \quad (21)$$

where $\beta = (\beta_1, \beta_2, \dots, \beta_p)$ is the anchored point. With the goal to minimize γ_j , [20] proves the following result.

Lemma 2 *Assume that the anchored-ANOVA expansion is truncated at order \tilde{v} and that $p_{\tilde{v}}$ satisfies*

$$\sum_{m=\tilde{v}+1}^p \sum_{|S|=m} \prod_{j \in S} \gamma_j = (1 - p_{\tilde{v}}) \left(\prod_{j=1}^p (1 + \gamma_j) - 1 \right). \quad (22)$$

Then, the relative error in L_∞ -norm can be estimated as

$$\frac{\|u - \sum_{|S| \leq \tilde{v}} u_S\|_{L_\infty}}{\|u\|_{L_\infty}} \leq (1 - p_{\tilde{v}}) \left(\prod_{j=1}^p (1 + \gamma_j) - 1 \right) \left(\prod_{j=1}^p \frac{|u_j(\beta_j)|}{\|u_j\|_{L_\infty}} \right). \quad (23)$$

Furthermore, for one-signed functions with the anchor point $\beta = (\beta_1, \beta_2, \dots, \beta_p)$ selected as

$$u_j(\beta_j) = \frac{1}{2} \left(\max_{[0,1]} u_j(\alpha_j) + \min_{[0,1]} u_j(\alpha_j) \right), \quad (24)$$

the corresponding γ_j minimizes the weights in Eq. (21).

This method is limited to functions with separated variables, (20), and the computation of (24), albeit one-dimensional in nature, can be costly. While developed for uniformly distributed parameters, this approach can be extended to include more general distributions by generalizing the theory with appropriate L_1 norms. In the following, we shall refer to the extremum anchor point as one chosen using this approach.

3.1 Anchor point as center of sparse grid quadrature

An intuitive alternative is to simply choose the centroid point in the parameter space. For uniformly distributed parameters this can be expected to work well. However, for the more general situation with non-uniformly distributed variables, it is reasonable to generalize this choice of the anchor point to that of being the centroid of the lowest dimensional tensorial Gaussian quadrature in the p -dimensional space. The quadrature should be chosen to reflect the proper measure associated with the non-uniformly distributed parameter.

As simple as choosing the anchor as the centroid of the tensorial quadrature is, its utilization is highlighted when recalling that one often seeks to be able to effectively compute moments of the ANOVA expanded function using sparse grids. As we shall show in the following theorem, there is a strong connection between between the anchored ANOVA expansion, the sparse grid Smolyak construction, and the anchor point based on the centroid.

Theorem 3 Let $u(\alpha)$ be a p -dimensional integrable function which is represented by the anchored ANOVA expansion located at the anchor point $\beta = (\beta_1, \dots, \beta_p)$, chosen to be the centroid of the Smolyak sparse grid. Then, all terms of order $l + 1 \leq p$ or higher are identically zero when evaluated at the p -dimensional $l + 1$ level sparse Smolyak grid.

Proof: Let $d\mu_i$, $i = 1, 2, \dots, p$ be a Dirac measure on A^i and define the averaging operator

$$(\Gamma_i u)(\alpha) = \int_{A^i} u(\alpha_1, \dots, \alpha_p) d\mu_i(\alpha_i) = u(\alpha_1, \dots, \beta_i, \dots, \alpha_p). \quad (25)$$

Let the identity be decomposed as

$$\begin{aligned} I &= \prod_i (\Gamma_i + (I - \Gamma_i)) = \prod_i \Gamma_i + \sum_i (I - \Gamma_i) \prod_{i \neq j} \Gamma_j \\ &+ \sum_{i < j} (I - \Gamma_i)(I - \Gamma_j) \prod_{k \neq i, j} \Gamma_k + \dots + \prod_i (I - \Gamma_i) \end{aligned} \quad (26)$$

Each term of (1) is generated by each of the components of this decomposition (26) [21],

$$\begin{aligned} u_0 &= \prod_i \Gamma_i u, \\ u_1 &= (I - \Gamma_i) \prod_{i \neq j} \Gamma_j u, \\ &\vdots \\ u_l &= (I - \Gamma_{\mathcal{L}_1})(I - \Gamma_{\mathcal{L}_2}) \cdots (I - \Gamma_{\mathcal{L}_l}) \prod_{\mathcal{M} \neq \mathcal{L}} \Gamma_{\mathcal{M}} u, \\ u_{l+1} &= (I - \Gamma_{(\mathcal{L}+1)_1})(I - \Gamma_{(\mathcal{L}+1)_2}) \cdots (I - \Gamma_{(\mathcal{L}+1)_{l+1}}) \prod_{\mathcal{N} \neq (\mathcal{L}+1)} \Gamma_{\mathcal{N}} u, \\ &\vdots \\ u_p &= \prod_i (I - \Gamma_i) \end{aligned} \quad (27)$$

Without loss of generality, we consider the first term of the $l + 1$ order term of the anchored ANOVA expansion,

$$\begin{aligned} u_{l+1}(\alpha) &= (I - \Gamma_{(\mathcal{L}+1)_1})(I - \Gamma_{(\mathcal{L}+1)_2}) \cdots (I - \Gamma_{(\mathcal{L}+1)_{l+1}}) \prod_{\mathcal{N} \neq (\mathcal{L}+1)} \Gamma_{\mathcal{N}} u(\alpha), \\ &= (I - \Gamma_{(\mathcal{L}+1)_1})(I - \Gamma_{(\mathcal{L}+1)_2}) \cdots (I - \Gamma_{(\mathcal{L}+1)_{l+1}}) \int_{A^{p-n}} u(\alpha_1, \dots, \alpha_p) d\mu(\alpha_{\mathcal{N}}), \\ &= (I - \Gamma_{(\mathcal{L}+1)_1})(I - \Gamma_{(\mathcal{L}+1)_2}) \cdots (I - \Gamma_{(\mathcal{L}+1)_{l+1}}) u(\alpha_1, \dots, \alpha_{l+1}, \beta_{l+2}, \dots, \beta_p). \end{aligned} \quad (28)$$

where $n = p - l - 1$. Observe that Eq. (28) contains at most $l + 1$ variables $(\alpha_1, \dots, \alpha_{l+1})$.

The $l + 1$ level sparse grid is given by (15)

$$\Omega_{l+1}^p = \bigcup_{(l+1) \leq |\mathbb{k}|_1 \leq l+p} (\alpha^{k_1} \times \cdots \times \alpha^{k_p}). \quad (29)$$

Assume now that there are $l + 1$ variables that are not solely defined at the centroid. Then

$$\begin{aligned} k_i &\geq 2, i = 1, \dots, l + 1 \\ k_j &= 1, j = l + 2, \dots, p \\ |\mathbb{k}|_1 &= \sum_{i=1}^{l+1} k_i + \sum_{j=l+2}^p k_j, \\ &\geq \sum_{i=1}^{l+1} 2 + \sum_{j=l+2}^p 1, \\ &= 2 \times (l + 1) + p - (l + 1), \\ &= p + l + 1. \end{aligned} \quad (30)$$

which contradicts $|\mathbb{k}|_1 \leq l + p$ in (29). Therefore, at least for one we have $k_i = 1$, $i = 1, \dots, l + 1$, i.e., $\alpha_i = \beta_i$ must be the centroid.

Without loss of generality, let α_{l+1} be this one. Equation (28) becomes

$$\begin{aligned} u_{l+1}(\alpha) &= (I - \Gamma_{(\mathcal{L}+1)_1}) \cdots (I - \Gamma_{(\mathcal{L}+1)_{l+1}}) u(\alpha_1, \dots, \alpha_l, \beta_{l+1}, \beta_{l+2}, \dots, \beta_p). \\ &= (I - \Gamma_{(\mathcal{L}+1)_1}) \cdots (I - \Gamma_{(\mathcal{L}+1)_l}) [I u(\alpha_1, \dots, \alpha_l, \beta_{l+1}, \beta_{l+2}, \dots, \beta_p) \\ &\quad - \Gamma_{(\mathcal{L}+1)_{l+1}} u(\alpha_1, \dots, \alpha_l, \beta_{l+1}, \beta_{l+2}, \dots, \beta_p)]. \\ &= (I - \Gamma_{(\mathcal{L}+1)_1}) \cdots (I - \Gamma_{(\mathcal{L}+1)_l}) [u(\alpha_1, \dots, \alpha_l, \beta_{l+1}, \beta_{l+2}, \dots, \beta_p) \\ &\quad - u(\alpha_1, \dots, \alpha_l, \beta_{l+1}, \beta_{l+2}, \dots, \beta_p)]. \\ &= 0 \end{aligned} \quad (31)$$

It is not difficult to conclude that all $m > l + 1$ order terms of the expansion are zero by repeating this argument. This completes the proof.

Apart from making the connection between the ANOVA expansion and the Smolyak sparse grid clear, an important implication of this result follows for evaluation of the moments of the anchored ANOVA expansion since one can decide exactly how many levels of the sparse grid is meaningful for an expansion of a certain length. Note, however, that the above result does not offer any measure of the accuracy of the expansion and, hence, the resulting moment.

4 Numerical examples

In the following we consider a comparative study of the different approaches for choosing the anchor point. We do this using standard high-dimensional test functions and also offer a direct comparison of the accuracy of the anchored ANOVA expansion to that of the Lebesgue ANOVA expansion for a high-dimensional system of ordinary differential equations.

4.1 Integration of high-dimensional functions

To measure the accuracy of the ANOVA expansion we define a measure of relative error of an integral as

$$\epsilon_{tr} = \frac{|\int_{A^p} u(\alpha) d\alpha - \int_{A^p} u_{tr}(\alpha) d\alpha|}{|\int_{A^p} u(\alpha) d\alpha|} \quad (32)$$

where $u_{tr}(\alpha)$ is the truncated ANOVA expansion.

We consider the classic test functions [22, 23] and one additional test example:

- Product Peak function: $u_2(\alpha) = \prod_{i=1}^p (c_i^{-2} + (\alpha_i - \xi_i)^2)^{-1}$,
- Corner Peak function: $u_3(\alpha) = (1 + \sum_{i=1}^p c_i \alpha_i)^{-(p+1)}$,
- Gaussian function: $u_4(\alpha) = \exp(-\sum_{i=1}^p c_i^2 (\alpha_i - \xi_i)^2)$,
- Continuous function: $u_5(\alpha) = \exp(-\sum_{i=1}^p c_i |\alpha_i - \xi_i|)$,
- Quadrature test example: $u_6(\alpha) = (1 + \frac{1}{p})^p \prod_{i=1}^p (\alpha_i)^{\frac{1}{p}}$.

where the parameters $c = (c_1, \dots, c_p)$ and $\xi = (\xi_1, \dots, \xi_p)$ are generated randomly. The parameter ξ acts as a shift parameter and the parameters c are constrained. See [22, 23] for the details. Note that the test function u_1 is defined in Eq.(17).

4.1.1 Uniformly distributed variables

In the first set of tests, we assume that all variables, α_i , $i = 1, 2, \dots, 10$, are uniformly distributed random variables defined on $[0, 1]^p$. We use a 10-dimensional 7-level sparse grids based on the one-dimensional Gauss-Patterson quadrature points to compute the integrals

and consider this to be the exact solution. 6-level sparse grids are used to integrate the anchored ANOVA expansion based on different choices of the anchor point.

In Fig. 3 we illustrate the relative error of the integrals recovered with different choices of the anchor point. Note that for most cases, the accuracy reaches 10^{-6} with the exception of the fifth test function where all choices lead to less accurate result. This is associated with this particular test function, is caused by a limited smoothness in the random variables and has also been reported by other authors [24].

While there are differences among the results, the choice of the centroid as the anchor appears to be superior to the alternative techniques in all cases. We also note that the results confirm the result in Theorem 3, i.e., with a 6-level sparse grid we should not expect any additional improvements in the accuracy of the integrals when using more than 5 terms in the ANOVA expansion,

4.1.2 Non-uniformly distributed elements

To further evaluate the accuracy and flexibility of using the center point of the associated sparse grid, let us again consider the problems listed above but now assume that the variables are beta-distributed variables with $\gamma = 1/2$ and $\tau = 1/3$, [25, 26]. Here the standard probability density of a beta-distributed random variable x is given as

$$f(x, \gamma, \tau) = x^\gamma(1-x)^\tau/B(\gamma, \tau), \quad 0 \leq x \leq 1, \quad \gamma > 0, \quad \tau > 0. \quad (33)$$

where $B(\gamma, \tau)$ is the normalizing beta function.

A 10-dimensional 6-level sparse grid is used to compute the integrals as the reference solution and a 5-level sparse grid is applied to compute the integral of the anchored ANOVA expansion. In Fig. 4 we show the results of the direct comparison with the alternatives that are most immediately applicable. The overall conclusion remains the same as in the uniform case above and confirms the accuracy and flexibility of the approach suggested here.

4.1.3 A higher dimensional problem

Let us again consider the oscillatory function, (17), but this time with more dimensions. The sums of the coefficients of the function are given in Table 1. We assume 10^{-4} to be an acceptable error in order to compare three different integration methods.

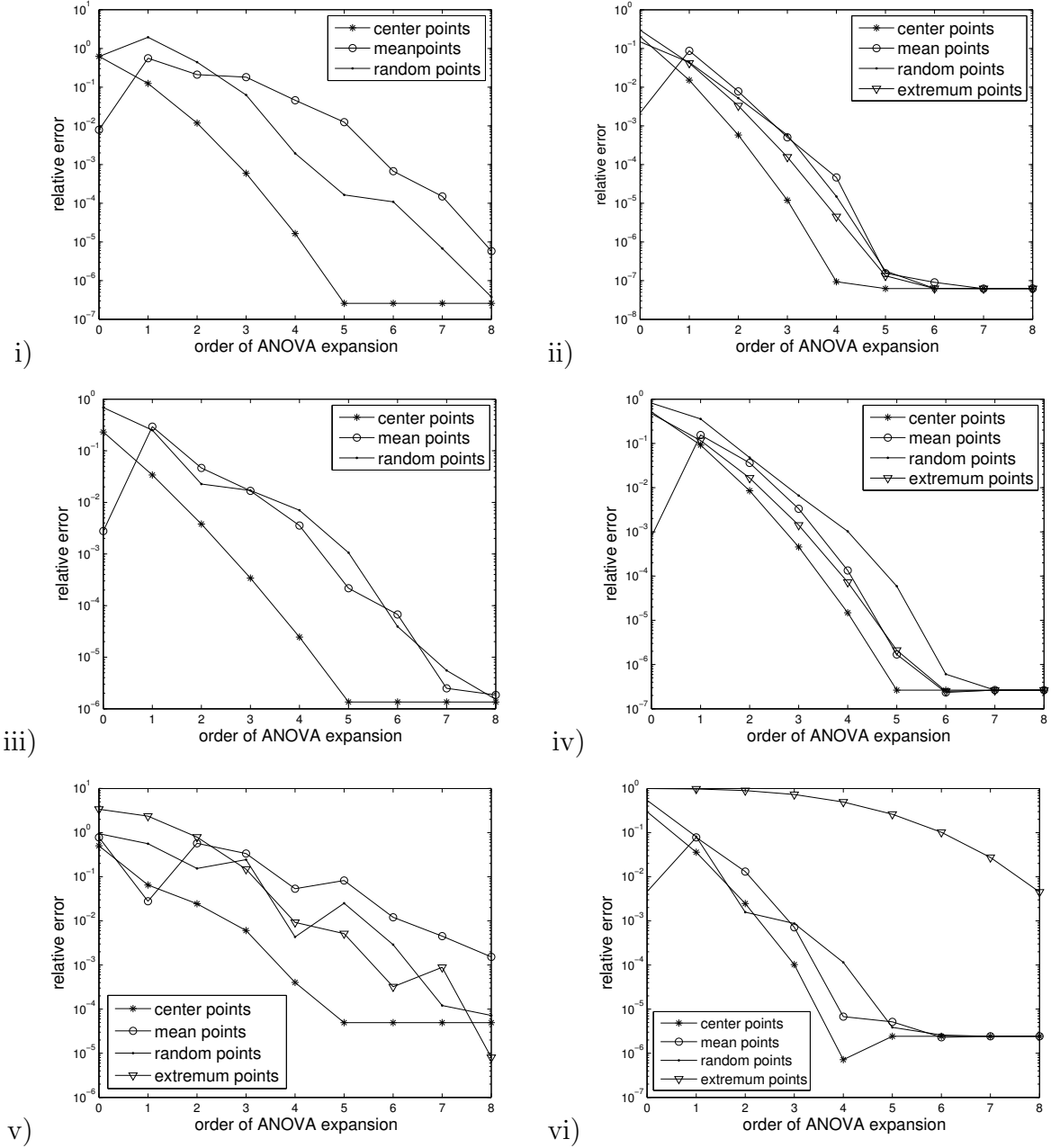


Figure 3: Relative errors of the test functions computed using different strategies for choosing the anchor point. All variables are assumed to be uniformly distributed. i) u_1 , ii) u_2 , iii) u_3 , iv) u_4 , v) u_5 , vi) u_6 .

In Fig. 5 we show results that confirm that the anchored ANOVA expansion is the most efficient method to integrate the test function until the dimension of the problem becomes sufficiently high. When this happens naturally depends on the problem at hand, in particular

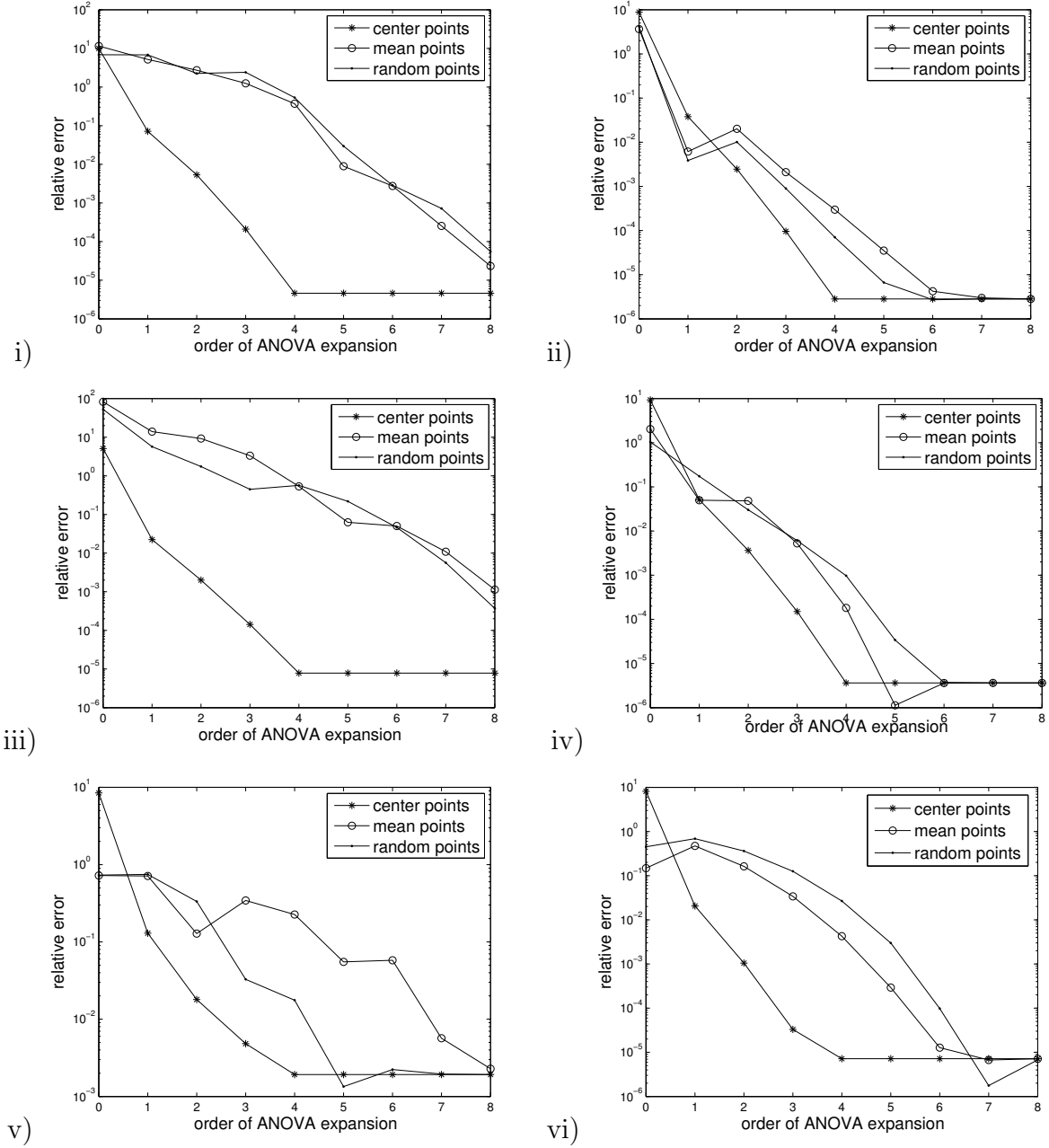


Figure 4: Relative errors of the test functions computed using different strategies for choosing the anchor point. All variables are assumed to be beta-distributed with $\gamma = 1/2$, $\tau = 1/3$.

i) u_1 , ii) u_2 , iii) u_3 , iv) u_4 , v) u_5 , vi) u_6 .

the cost of evaluating the function, i.e., for more complex and expensive function evaluations one can expect there to be further advantages in using the ANOVA expansion over a Monte Carlo based technique.

Table 1: Sums of coefficients of the Oscillatory function

p : number of dimension. $b_p = \sum_{i=1}^p c_i$.										
p	5	10	15	20	25	30	35	40	45	50
b_p	9.0	9.0	9.0	9.0	9.0	27	31.5	36	40.5	45

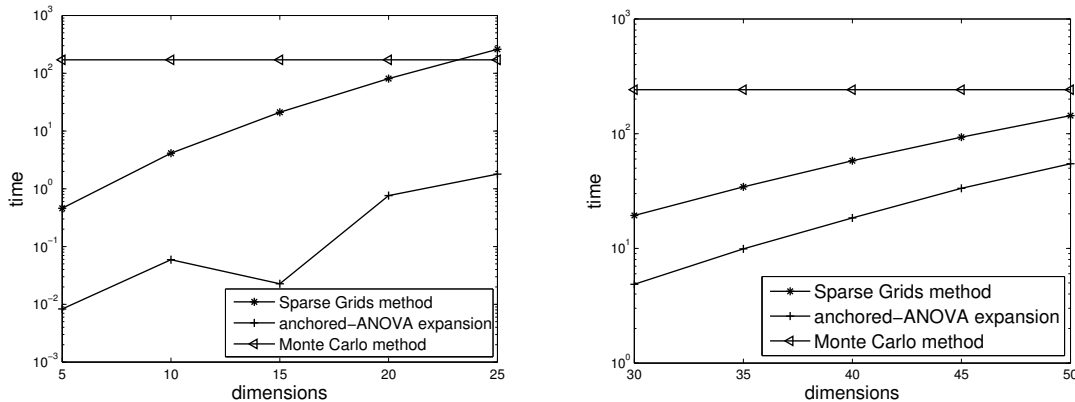


Figure 5: Computational costs of computing the integral of the oscillatory function using a sparse grids, the anchored ANOVA expansion with a sparse grid, and Monte Carlo method. On the left are shown results based on a 5-level sparse grid and on the right a 4-level sparse grid is used. The computational cost of the Monte Carlo is estimated from 15-dimensional and 40-dimensional test functions, respectively.

4.2 Higher dimensional ODE

Let us finally consider a more complex problem and also use this to compare the accuracy of the Lebesgue and the anchored ANOVA expansions.

We consider a situation with $p = 25$ 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 that it feels an acceleration through Newtonian gravitational forces from all the other particles. This leads to a simple dynamical equation

$$\ddot{u}(t) = \sum_{i=1}^p m_i \hat{r}_i / r_i^2, \quad u(t_0) = u_0. \quad (34)$$

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

is the Euclidian 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_j , are uniformly distributed random variables with a mean of $1/(p + 1)$ and a 10% variance.

As a high-dimensional function of interest, we consider the kinetic energy at a fixed time ($t = 8$) and built an ANOVA expansion of this. This is achieved by following the approach of [2] in which a second order polynomial between the kinetic energy of the moving particle and the random masses of fixed particles is constructed through a least squares approximation. This has been studied previously in [5] in a related context using a Lebesgue ANOVA expansion and we refer to that for further details.

To validate the accuracy and efficiency of the anchored ANOVA expansion, we also compute the Lebesgue ANOVA expansion using a Stroud-3 method [5] and a 25-dimensional level three sparse grid. We have used a 25-dimensional 3-level sparse grid to implement the anchored ANOVA expansion.

In Fig. 6 we show that there is only a slight difference in the L_2 and the L_∞ errors in the Lebesgue ANOVA expansion based on the Stroud-3 method and the sparse grid, confirming that the integration has converged and that the effective dimension of the kinetic energy is indeed very low

Figure 7 confirms that the second order truncated Lebesgue ANOVA expansion and the second order truncated anchored ANOVA expansion have the same accuracy. However, the latter is obtained at a cost which is more than two order of magnitude less. The three level Lebesgue ANOVA expansion is taken as the exact solution.

5 Concluding remarks

We have discussed two representations of high dimensional functions using ANOVA expansions, resulting in the classic Lebesgue ANOVA expansion and the anchored ANOVA expansion. Both of these can represent high dimensional functions well and often expose a low effective dimension, allowing for the effective evaluation of moments of the high-dimensional functions without impacting the accuracy. However, the classic ANOVA expansion is expensive to construct due to the need to accurately evaluate high dimensional integrals. We therefore consider the anchored ANOVA expansion in more detail.

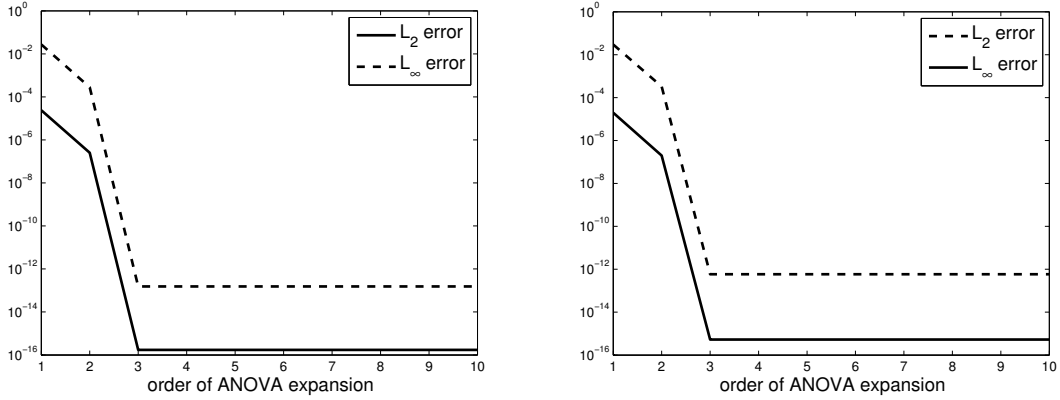


Figure 6: Errors of the Lebesgue ANOVA expansion computed using a Stroud 3 method (left) and a 3-level sparse grid (right).

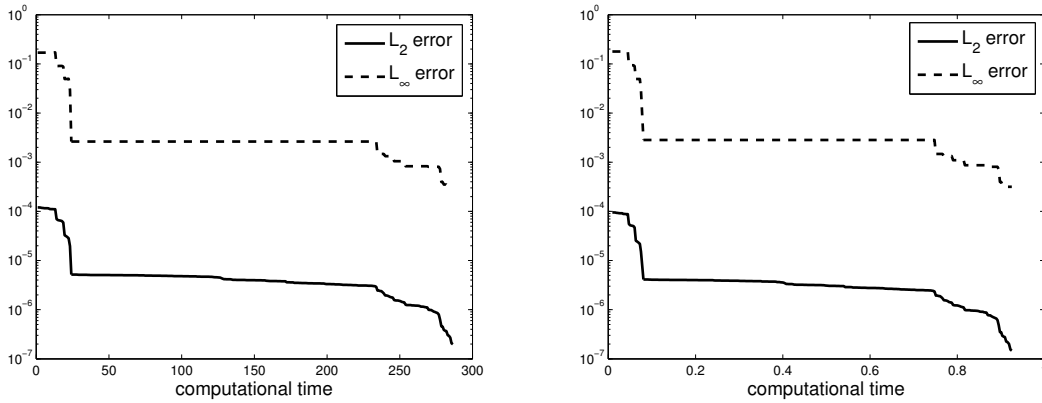


Figure 7: Errors of the Lebesgue ANOVA expansion (left) and the anchored ANOVA expansion (right).

A key element in the anchored ANOVA expansion is the need to choose an anchor point as this choice impacts the accuracy of the expansion, or rather the number of terms needed in the expansion to achieve a desired accuracy. This choice is therefore important from a practical point of view as longer expansions results in increased computational cost when evaluating moments.

We proposed a simple but efficient method for choosing the anchor point based on the structure of the Smolyak sparse grid. The computation of the anchor point is straightforward and avoids any additional cost. The accuracy and flexibility of this approach has been

demonstrated for a number of standard test functions and found to compare favorably with several alternative techniques. An additional advantage of this approach is that it generalizes to problems with non-uniformly distributed random variables.

The method was applied to study a more complex high-dimensional system of ordinary differential equations, yielding excellent agreement with results obtained through a Lebesgue ANOVA expansion, yet achieved at considerable less cost.

The derivation of more rigorous error estimates for the anchored ANOVA expansion largely remains an open question but with the close connection to the sparse grid integration discussed here, we hope to be able to report on this in the future.

6 Acknowledgement

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

References

- [1] Alis O. F. Rabitz H., General Foundation of High Dimensional Model Representations. *J. Math. Chem.* **25**(1999), 197-233.
- [2] Y. Cao, Z. Chen and M. Gunzburger, ANOVA expansions and efficient sampling methods for parameter dependent nonlinear PDEs, *Inter. J. Numer. Anal. Model.* **6**(2009), 256-273.
- [3] C. Gu, *Smoothing Spline ANOVA Models*, Springer, Berlin, 2002.
- [4] A. Saltelli, K. Chan and E. Scott, *Sensitivity Analysis*, Wiley & Sons, Chichester, 2000.
- [5] Z. Gao, J. S. Hesthaven, 2010, Efficient solution of ordinary differential equations with high-dimensional parametrized uncertainty, *Commun. Comput. Phys.* submitted.

- [6] M. Liu, Z. Gao and J. S. Hesthaven, 2010, Adaptive sparse grid algorithms with applications to electromagnetic scattering under uncertainty, *Appl. Numer. Math.* - to appear.
- [7] O.F. Alis and H. Rabitz, General Foundation of High Dimensional Model Representations. *J. Math. Chem.* **25**(1999), 197-233.
- [8] T. Gerstner and M. Griebel, Numerical integration using sparse grids, *Numer. Algorithms*, **18**(1998), 209-232
- [9] M. Griebel. Sparse grids and related approximation schemes for higher dimensional problems. In L. Pardo, A. Pinkus, E. Suli and M. Todd, editors, *Foundations of Computational Mathematics (FoCM05)*, Santander, 106-161. Cambridge University Press, 2006.
- [10] F. Kuo, I. Sloan, G. Wasilkowski and H. Wozniakowski. On decompositions of multivariate functions. Technical Report, University of New South Wales, 2008.
- [11] R. E. Caflishch, W. Morokoff, and A. Owen, Valuation of mortgage backed securities using Brownian bridges to reduce effective dimension, *J. Comput. Finance* **1**(1997), 27-46
- [12] A. B. Owen, The dimension distribution and quadrature test functions, Technical Report, Stanford University, 2001.
- [13] S. Paskov and J. Traub, Faster valuation of financial derivatives, *J. Portfolio Managm.* **22**(1995), 113-120.
- [14] S.A.Smolyak, Quadrature and interpolation formulas for tensor products of certain classes of functions, *Dokl. Akad. Nauk SSSR* **4**(1963), 240-243.
- [15] I.M. Sobol' , Sensitivity estimates for nonlinear mathematical models. *Mathematicheskoe Modelirovanie* **2**(1990), 112-118 (in Russian).
- [16] I.M. Sobol', Sensitivity estimates for nonlinear mathematical models, *Math. Model. Comput. Exp.* **1**(1993), 407-414.

- [17] I. M. Sobol, Theorems and examples on high dimensional model representation, *Reliab. Engng. Syst. Safety* **79**(2003), 187-193.
- [18] G.W. Wasilakowski and H. Wozniakowski, Explicit cost bounds of algorithms for multivariate tensor product problems, *J. Complexity*, **11**(1995), 1-56.
- [19] X. Wang and K. Fang, The effective dimension and quasi-Monte Carlo integration, *J. Complex* **19**(2003), 101-124.
- [20] Z. Zhang, M. Choi and G. E. Karniadakis, Anchor Points Matter in ANOVA Decomposition, *Proceedings of ICOSAHOM'09*, 2010.
- [21] G. Wahba, An Introduction to (Smoothing Spline) ANOVA Models in RKHS, with Examples in Geographical Data, Medicine, Atmospheric Science and Machine Learning, the Proceeding of the 13th IFAC Symposium on System Identificatin 2003, Rotterdam, 549-559.
- [22] 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.
- [23] A.C. Genz, A package for testing multiple integration subroutines, in: *Numerical Integration*, eds. P. Keast and G. Fairweather (Kluwer, Dordrecht, 1987, 337-340.
- [24] V. Barthelmann, E. Novak and K. Ritter, High dimensional polynomial interpolation on sparse grids, *Adv. Comput. Math.* **12**(2000), 273-288.
- [25] D. L. Keefer and S. E. Bodily, Three-point Approximations for Continuous Random variables. *Management Science* **29**(1983), 595-609, 1983.
- [26] D. L. Keefer and S. E. Bodily, Better Estimation of PERT Activity Time Parameters. *Management Science* **39**(1993), 1086-1091.