

The rate of convergence of sparse grid quadrature on the torus

Markus Hegland¹ Paul C. Leopardi²

(Received 30 January 2011; revised 27 June 2011)

Abstract

We examine sparse grid quadrature on Korobov spaces; that is, weighted tensor product reproducing kernel Hilbert spaces on the torus. We describe a dimension adaptive quadrature algorithm based on an algorithm of Hegland [ANZIAM J., 44(E):C335, 2003], and also formulate a version of Wasilkowski and Woźniakowski's weighted tensor product algorithm [J. Complexity, 15(3):402, 1999]. We claim that our algorithm is generally lower in cost than Wasilkowski and Woźniakowski's algorithm, and therefore both algorithms have the optimal asymptotic rate of convergence given by Theorem 3 of Wasilkowski and Woźniakowski. Even so, if the dimension weights decay slowly enough, both algorithms need a number of points exponential in the dimension to produce a substantial reduction in quadrature error.

<http://anziamj.austms.org.au/ojs/index.php/ANZIAMJ/article/view/3952> gives this article, © Austral. Mathematical Soc. 2011. Published July 29, 2011. ISSN 1446-8735. (Print two pages per sheet of paper.) Copies of this article must not be made otherwise available on the internet; instead link directly to this URL for this article.

Contents

1 Introduction	C501
2 Setting	C503
3 Algorithm	C506
4 Analysis	C508
5 Numerical results	C512
References	C514

1 Introduction

This article examines sparse grid quadrature on weighted Korobov spaces, that is, weighted tensor products (WTP) of reproducing kernel Hilbert spaces (RKHS) on the torus \mathbb{T}^d , the d -dimensional product of \mathbb{S}^1 , the circle of unit radius.

A function defined on the d -dimensional unit torus is equivalent, from the point of view of integration, to a function which is periodic on the d -dimensional unit cube. The integration of functions on a high dimensional unit cube occurs in many applications, most notably in finance. Often, such integration cannot be performed analytically, but must be approximated by quadrature, that is, by a linear combination of function values obtained at a finite number of points in the cube.

The study of rates of convergence of quadrature of periodic functions on the unit cube goes back at least as far as Korobov [5, 6], who studied classes of functions defined by the rate of decay of the coefficients in the Fourier series of the function. The rate of convergence of a quadrature rule, as the number

of points increases, generally depends on the properties of the function, or space of functions, to which the quadrature rule is applied.

More recently, attention has shifted to Korobov spaces, defined as reproducing kernel Hilbert spaces of periodic functions on the unit cube. In a Korobov space, the reproducing kernel is the product of one dimensional kernels. Weighted Korobov spaces are a generalisation of Korobov spaces where each one dimensional kernel is defined using a possibly different weight. These weighted spaces are used to study tractability of quadrature rules. The general question studied is how the rate of convergence of a quadrature rule depends on the number of dimensions of the space where it is defined. Rates of convergence and criteria for strong tractability of quadrature in weighted Korobov spaces have been well studied by Hickernell and Woźniakowski [4] and by Sloan and Woźniakowski [8]. As noted by Kuo and Sloan [7], the setting of weighted Korobov spaces is equivalent to the setting they used to examine quasi-Monte Carlo quadrature on \mathbf{d} -dimensional products of the unit sphere $\mathbb{S}^s \subset \mathbb{R}^{s+1}$, confined to the special case where $s = 1$, that is, the \mathbf{d} -dimensional torus $\mathbb{T}^{\mathbf{d}}$. This is the setting used in this article, and it is described in more detail in the next section 2.

The idea of sparse grid quadrature is based on Smolyak's quadrature rules. Smolyak's original article studied the convergence of rules for non-periodic functions on the unit cube, as well as Korobov's classes [9].

In function spaces where the dimensions may have different weights, Smolyak's original sparse grid algorithm has been modified and generalized, resulting in dimension adaptive sparse grid algorithms. Such algorithms include those of Wasilkowski and Woźniakowski [10], Hegland [3], and Gerstner and Griebel [2]. Of these, the WTP algorithm of Wasilkowski and Woźniakowski has the most well-developed theory of the rate of convergence of the worst case error, but this algorithm and its analysis are based on weighted spaces of non-periodic functions on the unit cube, and must be modified for our weighted Korobov setting.

The remainder of this article is organised as follows. Section 2 describes our

weighted Korobov space setting in detail. Section 3 introduces our dimension adaptive sparse grid quadrature algorithm. Section 4 analyses our version of the WTP algorithm of Wasilkowski and Woźniakowski, and compares its theoretical rate of convergence with that of our dimension adaptive algorithm. Section 5 contains numerical results, comparing the two algorithms, and shows how our algorithm performs as the dimension is increased.

2 Setting

Let $\mathcal{D} \subset \mathbb{R}^{s+1}$ be a compact s -dimensional manifold with probability measure μ . It follows that the constant function $\mathbf{1}$, with $\mathbf{1}(x) = 1$ for all $x \in \mathcal{D}$, is integrable and $\int_{\mathcal{D}} \mathbf{1}(x) d\mu(x) = 1$. Then let H be a Hilbert space of functions $f : \mathcal{D} \rightarrow \mathbb{R}$, with a kernel K , satisfying

- for every $x \in \mathcal{D}$ there exists $k_x \in H$ such that

$$f(x) = \langle k_x, f \rangle_H \quad \text{for all } f \in H, \quad (1)$$

- every $f \in H$ is integrable and

$$\int_{\mathcal{D}} f(x) d\mu(x) = \langle \mathbf{1}, f \rangle_H, \quad (2)$$

where the functions $k_x(y) := K(x, y)$, and where $\langle \cdot, \cdot \rangle_H$ denotes the scalar product in H . We recognise H as a reproducing kernel Hilbert space. In this framework, quadrature methods Q , defined by

$$Q(f) := \sum_{i=1}^n w_i f(x_i),$$

are continuous linear functionals and $Q(f) = \langle q, f \rangle_H$ with $q = \sum_{i=1}^n w_i k_{x_i}$.

We assume that the quadrature points x_i are given. An optimal choice of weights w_i minimises the worst case quadrature error $e(\mathbf{q})$, which is given by the norm $\|\mathbf{1} - \mathbf{q}\|_H$. The optimal \mathbf{q}^* is thus defined as

$$\mathbf{q}^* := \operatorname{argmin}_{\mathbf{q}} \{ \|\mathbf{1} - \mathbf{q}\|_H \mid \mathbf{q} \in \operatorname{span}\{\mathbf{k}_{x_1}, \dots, \mathbf{k}_{x_n}\} \}. \quad (3)$$

The weights of an optimal quadrature method are thus obtained by solving a linear system of equations with a matrix whose elements are the values of the reproducing kernel $K(x_i, x_j) = \langle \mathbf{k}_{x_i}, \mathbf{k}_{x_j} \rangle_H$. The right-hand side of these equations is a vector with elements all equal to one.

We now describe our more specific reproducing kernel Hilbert space \mathcal{H} of functions on \mathcal{D} . The space \mathcal{H} satisfies (1), but as well as (2), it also satisfies the more specific

$$\int_{\mathcal{D}} f(x) \, d\mu(x) = \langle \mathbf{1}, f \rangle_{\mathcal{H}} = 0.$$

We now extend \mathcal{H} into the space \mathcal{H}^γ , which consists of all functions of the form $\mathbf{g} = \mathbf{a}\mathbf{1} + f$, where $\mathbf{a} \in \mathbb{R}$, and $f \in \mathcal{H}$ with the norm $\|\cdot\|_{\mathcal{H}^\gamma}$ defined by

$$\|\mathbf{g}\|_{\mathcal{H}^\gamma}^2 = |\mathbf{a}|^2 + \frac{1}{\gamma} \|f\|_{\mathcal{H}}^2.$$

It is easily verified that \mathcal{H}^γ is an RKHS with reproducing kernel

$$\mathcal{K}_\gamma(\mathbf{x}, \mathbf{y}) = \mathbf{1} + \gamma \mathcal{K}(\mathbf{x}, \mathbf{y}),$$

where \mathcal{K} is the reproducing kernel of the RKHS \mathcal{H} .

For functions on the domain \mathcal{D}^d we consider the tensor product space $\mathcal{H}_d := \bigotimes_{k=1}^d \mathcal{H}^{\gamma_k}$ where $\mathbf{1} \geq \gamma_1 \geq \dots \geq \gamma_d \geq 0$. This is an RKHS of functions on \mathcal{D}^d with reproducing kernel $\mathcal{K}_d(\mathbf{x}, \mathbf{y}) := \prod_{k=1}^d (1 + \gamma_k \mathcal{K}(x_k, y_k))$ where $x_k, y_k \in \mathcal{D}$ are the components of $\mathbf{x}, \mathbf{y} \in \mathcal{D}^d$. Moreover

$$\int_{\mathcal{D}^d} f(\mathbf{x}) \, d\mu_d(\mathbf{x}) = \langle \mathbf{1}, f \rangle_{\mathcal{H}_d},$$

where μ_d is the product measure, $\langle \cdot, \cdot \rangle_{\mathcal{H}_d}$ is the scalar product on the tensor product space \mathcal{H}_d , and $\mathbf{1}$ is the constant function on \mathcal{D}^d with value 1. It follows that the space \mathcal{H}_d satisfies the two conditions (1) and (2) and we can derive optimal quadrature rules for given point sets.

We now describe our specific Korobov space setting, which is the setting of Kuo and Sloan [7], with $s := 1$. We take our domain \mathcal{D} to be the unit circle $\mathbb{T} := \mathbb{S}^1 := \{\mathbf{x} \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 1\}$, and consider the real space $L_2(\mathbb{T})$. We use the real Fourier basis defined by $Y_{0,0}(\mathbf{x}) := \mathbf{1}$, $Y_{\ell,1}((\cos \theta, \sin \theta)) = \cos \ell \theta$, $Y_{\ell,2}((\cos \theta, \sin \theta)) = \sin \ell \theta$, $\ell = 1, \dots, \infty$.

For $f \in L_2(\mathbb{T})$, we expand f in the Fourier series

$$f(\mathbf{x}) = \hat{f}_{0,0} + \sum_{\ell=1}^{\infty} \sum_{m=1}^2 \hat{f}_{\ell,m} Y_{\ell,m}(\mathbf{x}).$$

For positive weight γ , we define the RKHS

$$\mathbb{H}_{1,\gamma}^{(r)} := \{f \in L_2(\mathbb{T}) \mid \|f\|_{\mathbb{H}_{1,\gamma}^{(r)}} < \infty\}, \quad \text{where}$$

$$\langle f, g \rangle_{\mathbb{H}_{1,\gamma}^{(r)}} := \hat{f}_{0,0} \hat{g}_{0,0} + \gamma^{-1} \sum_{\ell=1}^{\infty} \sum_{m=1}^2 \ell^{2r} \hat{f}_{\ell,m} \hat{g}_{\ell,m}.$$

The reproducing kernel of $\mathbb{H}_{1,\gamma}^{(r)}$ is then

$$K_{1,\gamma}^{(r)}(\mathbf{x}, \mathbf{y}) := 1 + \gamma A_r(\mathbf{x} \cdot \mathbf{y}), \quad \text{where for } z \in [-1, 1],$$

$$A_r(z) := \sum_{\ell=1}^{\infty} \frac{2}{\ell^{2r}} T_{\ell}(z),$$

with T_{ℓ} the Chebyshev polynomial of the first kind, $T_{\ell}(\cos \theta) := \cos \ell \theta$. Convergence of A_r requires that $r > 1/2$.

For $\gamma := (\gamma_{d,1}, \dots, \gamma_{d,d})$, we now define the tensor product space

$$\mathbb{H}_{d,\gamma}^{(r)} := \bigotimes_{k=1}^d \mathbb{H}_{1,\gamma_{d,k}}^{(r)}.$$

This is a weighted Korobov RKHS on \mathbb{T}^d , with reproducing kernel

$$\mathbb{K}_{d,\gamma}^{(r)}(\mathbf{x}, \mathbf{y}) := \prod_{k=1}^d \mathbb{K}_{1,\gamma_{d,k}}^{(r)}(x_k, y_k).$$

This space is equivalent to the weighted Korobov space of periodic functions on the unit cube, studied by Hickernell and Woźniakowski [4] and by Sloan and Woźniakowski [8]. Quadrature is therefore strongly tractable on $\mathbb{H}_{d,\gamma}^{(r)}$ if and only if $\sum_{k=1}^{\infty} \gamma_{d,k} < \infty$, and, in the case of exponentially decreasing weights as studied here, the optimal worst-case error has an upper bound of order $\mathcal{O}(n^{-r})$, where n is the cost of the quadrature rule in terms of the number of points [8]. The order of the lower bound is known to be the same as that of the non-periodic setting [4].

3 Algorithm

Algorithm 1 studied here is an adaptation of the dimension adaptive algorithm for the solution of variational problems suggested by Hegland [3]. We describe our algorithm in our general RKHS setting, as given in Section 2.

We assume here that the quadrature points in \mathcal{D} are given and the same for all spaces \mathcal{H}^γ . We only consider up to a maximum of n points which we denote by $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathcal{D}$. The quadrature rules for \mathcal{H}^γ are then defined as some element of $\mathbb{V}_i^\gamma = \text{span}\{\mathbf{k}_{\mathbf{x}_1}^\gamma, \dots, \mathbf{k}_{\mathbf{x}_{n_i}}^\gamma\} \subset \mathcal{H}^\gamma$. We denote the optimal rule in \mathbb{V}_i^γ by \mathbf{q}_i^γ . Now define the pair-wise orthogonal spaces \mathbb{U}_i^γ by $\mathbb{U}_0^\gamma = \mathbb{V}_0^\gamma$, and by the orthogonal decomposition $\mathbb{V}_{i+1}^\gamma = \mathbb{V}_i^\gamma \oplus \mathbb{U}_{i+1}^\gamma$. Using that the \mathbf{q}_i^γ are optimal,

$$\delta_{i+1}^\gamma := \mathbf{q}_{i+1}^\gamma - \mathbf{q}_i^\gamma \in \mathbb{U}_{i+1}^\gamma$$

and $\delta_0 := \mathbf{q}_0^\gamma \in \mathbb{U}_0^\gamma = \mathbb{V}_0^\gamma$. Note that

$$\mathbb{U}_{i+1}^\gamma \neq \text{span}\{\mathbf{k}_{\mathbf{x}_{n_{i+1}}}, \dots, \mathbf{k}_{\mathbf{x}_{n_{i+1}}}\}.$$

Algorithm 1: The dimension adaptive (DA) algorithm.

Data: accuracy ϵ , incremental rules Δ_j and their costs ν_j for $j \in \mathbb{N}^d$

Result: ϵ approximation $\mathbf{q}^{(\text{DA})}$ and index set I

```

1  $I := \{0\}; \mathbf{q} := \Delta_0;$ 
2 while  $\|\mathbf{1} - \mathbf{q}\| > \epsilon$  do
3    $\mathbf{i} := \operatorname{argmax}_j \{\|\Delta_j\|^2 / \nu_j \mid I \cup \{j\} \text{ is a down-set}\};$ 
4    $I := I \cup \{\mathbf{i}\}; \mathbf{q} := \mathbf{q} + \Delta_{\mathbf{i}};$ 
5 end

```

A sparse grid quadrature rule is then of the form

$$\mathbf{q} \in V_I := \sum_{j \in I} \bigotimes_{k=1}^d V_{j_k}^{\gamma_{d,k}}$$

for some index set I . From the orthogonal decomposition $V_j^\gamma = \bigoplus_{i=1}^j U_i^\gamma$ one derives the multidimensional orthogonal decomposition

$$V_I = \bigoplus_{j \in \bar{I}} \bigotimes_{k=1}^d U_{j_k}^{\gamma_{d,k}},$$

where $\bar{I} = \{\mathbf{i} \mid \mathbf{i} \leq \mathbf{j} \text{ for some } \mathbf{j} \in I\}$, where the comparison $\mathbf{i} \leq \mathbf{j}$ has to hold for all components of \mathbf{i} and \mathbf{j} . When $I = \bar{I}$, we say that I is a *down-set* [1, p.13]. One can then show that an optimal $\mathbf{q} \in V_I$ is obtained as

$$\mathbf{q}_I^* = \sum_{j \in \bar{I}} \bigotimes_{k=1}^d \delta_{j_k}^{\gamma_{d,k}}.$$

Thus both V_I and \mathbf{q}_I^* are obtained in terms of the down-set \bar{I} , effectively restricting our choice of the set I to index sets which are also down-sets.

We now describe our dimension adaptive (DA) Algorithm 1 to choose the set I . We first define $\nu_{j_k}^{(k)} := \dim U_{j_k}^{\gamma_{d,k}}$ and $\delta_{j_k}^{(k)} := \delta_{j_k}^{\gamma_{d,k}}$. The algorithm then uses

the definitions

$$\nu_j := \prod_{k=1}^d \nu_{j_k}^{(k)} \quad \text{and} \quad \Delta_j := \bigotimes_{k=1}^d \delta_{j_k}^{(k)}.$$

Here j_k is the k th component of the multi-index j .

4 Analysis

We first describe the situation on a single circle where, if γ is large enough, the norm of the one point rule is less than the norm of the difference between the optimal two point rule and the one point rule, and show how this reverses the usual order of norms between successive incremental rules in many cases. In the following, we consider optimal weight rules in the sense of (3), and we abbreviate $K_{1,\gamma}^{(r)}$ to K .

The squared norm of the optimal one point rule on $\mathbb{H}_{1,\gamma}^{(r)}$ is $1/K(x, x) = 1/(1 + \gamma A_r(1))$. The optimal two point rule, with points x_1 and x_2 and weights w_1 and w_2 , has squared norm $2/(K(x_1, x_1) + K(x_1, x_2))$. Since A_r is an increasing function over $[-1, 1]$, $K(x_1, x_2)$ is minimal when $x_2 = -x_1$. (The two point rules used by the DA quadrature are of this form.) The optimal two point rule therefore has squared norm $2/\{2 + \gamma[A_r(1) + A_r(-1)]\}$. This is more than twice the squared norm of the one point rule when $A_r(-1) < -1$ and $\gamma > -1/A_r(-1)$. It can be shown that $A_r(-1) < -1$ for any $r > 1/2$. For our numerical examples, which have $r = 3$, we have $-1/A_3(-1) \simeq 0.50733$.

Consider two incremental DA rules Δ_j and $\Delta_{j'}$, on $\mathbb{H}_{d,\gamma}^{(r)}$, with $j_k = j'_k$ for all k except that $j_{k'} = 0$ and $j'_{k'} = 1$. Since the norms of the incremental rules are the products of difference rules on each circle, if $\gamma_{d,k} > -1/A_r(-1)$, then $\|\Delta_{j'}\| > \|\Delta_j\|$.

We now turn to estimates for rules on a single circle, in order to use them with an adapted version of the theory of Wasilkowski and Woźniakowski.

On a single circle, our sparse grid quadrature rule is an optimal weight rule $\mathbf{q}_j := \mathbf{q}_{r,\gamma}(\mathcal{S}_j)$, based on a set of equally spaced points \mathcal{S}_j on the unit circle, with $n_j := |\mathcal{S}_j|$. The series expansion of the function A_r then yields the following error bound for quadrature on $\mathbb{H}_{1,\gamma}^{(r)}$, namely

$$e^2(\mathbf{q}_j) \leq \frac{4r}{2r-1} \gamma n_j^{-2r}.$$

For our numerical example, we also have $n_j = 2^j$. Since $e^2(\mathbf{q}_j) \geq 0$, we can therefore show that $\|\mathbf{q}_j - \mathbf{q}_{j-1}\|_{\mathbb{H}_{1,\gamma}^{(r)}} \leq \sqrt{\gamma} C 2^{-rj}$, where

$$C := 2^{1-r} \sqrt{r/(2r-1)}. \tag{4}$$

In our setting, and our notation, the criteria needed by Wasilkowski and Woźniakowski [10, Theorem 3] become

$$n_{j+1} D^{jp} \leq 1, \quad \text{for all } j \geq 1, \quad \text{and} \tag{5}$$

$$\|\mathbf{q}_j - \mathbf{q}_{j-1}\|_{\mathbb{H}_{1,\gamma}^{(r)}} \leq \sqrt{\gamma} C D^j, \quad \text{for all } j \geq 1, \tag{6}$$

for some $D \in (0, 1)$ and some positive C and p . For the points used by our DA quadrature rules, these criteria hold with C as per (4), $D = 2^{-r}$, and $p = 1/r$.

We now describe a second variant of WTP quadrature, $\mathbf{q}^{(WW)}$ on $\mathbb{H}_{d,\gamma}^{(r)}$, identical to the sequence of quadrature rules $\mathbf{q}^{(DA)}$ described in Section 3, except that the order in which the incremental rules are added to this second variant rule is essentially the order used by Wasilkowski and Woźniakowski [10, §5]. As a consequence of (6), we have

$$\|\Delta_j\|_{\mathbb{H}_{d,\gamma}^{(r)}} = \prod_{k=1}^d \left\| \delta_{jk}^{(k)} \right\|_{\mathbb{H}_{1,\gamma d,k}^{(r)}} \leq b(d, j) := \prod_{k=1}^d \left(\sqrt{\gamma_{d,k}} C D^{jk} \right)^{1-\delta_{0,jk}}.$$

Let $(\xi_{d,k})$, $k = 1, \dots, d$, be a sequence of positive numbers. In contrast to Wasilkowski and Woźniakowski [10, §5], we do not stipulate that $\xi_{d,k} = 1$. Define

$$\xi(\mathbf{d}, \mathbf{j}) := \prod_{k=1}^d \xi_{d,k}^{1-\delta_{0,j_k}}. \tag{7}$$

We therefore have $b(\mathbf{d}, \mathbf{j})/\xi(\mathbf{d}, \mathbf{j}) \rightarrow 0$ as $\|\mathbf{j}\|_1 \rightarrow \infty$. We order the incremental rules in order of non-decreasing $b(\mathbf{d}, \mathbf{j})/\xi(\mathbf{d}, \mathbf{j})$ for each multi-index \mathbf{j} , creating an order on the multi-indices $\mathbf{j}^{(WW)}(\mathbf{h})$. We adjust $\xi(\mathbf{d}, \mathbf{k})$ so that this order agrees with the lattice partial ordering of the multi-indices. For our numerical examples, we use $\xi_{d,k} := CD$, with C and D defined as above. We now define $I_N^{(WW)} := \{\mathbf{j}^{(WW)}(1), \dots, \mathbf{j}^{(WW)}(N)\}$, and define the quadrature rule

$$q_N^{(WW)} := \sum_{\mathbf{j} \in I_N^{(WW)}} \Delta_{\mathbf{j}}.$$

To obtain a quadrature error of at most $\epsilon \in (0, 1)$, we set

$$N(\epsilon, \mathbf{d}) := \left| \left\{ \mathbf{j} \mid b(\mathbf{d}, \mathbf{j})/\xi(\mathbf{d}, \mathbf{j}) > (\epsilon/C_1(\mathbf{d}, \eta))^{1/(1-\eta)} \right\} \right|,$$

where $\eta \in (0, 1)$ and

$$C_1(\mathbf{d}, \eta) := \sqrt{\frac{\xi_{d,1}^{2(1-\eta)}}{1-D^2} \prod_{k=2}^d \left(1 + (C^2 \gamma_{d,k})^\eta \xi_{d,k}^{2(1-\eta)} \frac{D^{2\eta}}{1-D^{2\eta}} \right)}.$$

Finally, we define

$$q_\epsilon^{(WW)} := \sum_{\mathbf{j} \in I_{N(\epsilon, \mathbf{d})}^{(WW)}} \Delta_{\mathbf{j}}. \tag{8}$$

We now present our version of Wasilkowski and Woźniakowski’s main theorem on the error and cost of WTP quadrature [10, Theorem 3].

Theorem 1 *Let $\eta \in (0, 1)$. Then the quadrature rule $\mathbf{q}_{\epsilon, \mathbf{d}}^{(\text{WW})}$ defined by (8) has worst-case quadrature error $e(\mathbf{q}_{\epsilon}^{(\text{WW})}) \leq \epsilon$, and its cost (in number of quadrature points) is bounded by*

$$\text{cost}(\mathbf{q}_{\epsilon, \mathbf{d}}^{(\text{WW})}) \leq C(\mathbf{d}, \epsilon) \left(\frac{1}{\epsilon}\right)^{p/(1-\eta)},$$

where

$$C(\mathbf{d}, \epsilon) := \frac{\xi_{\mathbf{d}, 1}^p \prod_{k=2}^{\mathbf{d}} (1 + C^p \gamma_{\mathbf{d}, k}^{p/2} / \xi_{\mathbf{d}, k}^p) f(\mathbf{k}, \epsilon)^p}{(1 - D^p)(1 - D^2)^{p/(2(1-\eta))}},$$

$$f(\mathbf{i}, \epsilon) := \left(1 + C^{2\eta} \gamma_{\mathbf{d}, \mathbf{i}}^{\eta} \xi_{\mathbf{d}, \mathbf{i}}^{2(1-\eta)} \frac{D^{2\eta}}{1 - D^{2\eta}}\right)^{1/(2(1-\eta))},$$

$$g(\mathbf{k}, \epsilon) := \left[\frac{\log \left(C \gamma_{\mathbf{d}, \mathbf{k}}^{1/2} / (\xi_{\mathbf{d}, \mathbf{k}} (1 - D^2))^{1/(2(1-\eta))} \prod_{i=2}^{\mathbf{k}} (f(\mathbf{i}, \epsilon)) \epsilon^{-1/(1-\eta)} \right)}{\log D^{-1}} \right]_+.$$

By $[x]_+$, we mean $\max(0, x)$.

Wasilkowski and Woźniakowski’s proof, with $s := 2$ and $\alpha := 1$, applies directly to our Theorem 1, once the change in $\xi_{\mathbf{d}, 1}$ is taken into account. For exponentially decreasing dimension weights $\gamma_{\mathbf{d}, k}$, Theorem 4 of Wasilkowski and Woźniakowski [10] shows that the $\mathbf{q}^{(\text{WW})}$ rules are strongly polynomial.

We conjecture that our sequence of rules $\mathbf{q}^{(\text{DA})}$ is in general more efficient than $\mathbf{q}^{(\text{WW})}$, since $\mathbf{q}^{(\text{DA})}$ relies on an ordering of multi-indices in terms of decreasing average squared norm W_j/N_j , and this ordering is greedy with respect to minimising the error of the overall quadrature rule with respect to its cost in terms of function evaluations. Our conjecture is true when $\gamma_{\mathbf{d}, 1} < -1/A_r(-1)$, since then the ordering with respect to nondecreasing norm W_j , nondecreasing average squared norm W_j/N_j , and nonincreasing number of

points agree for each single dimension, and therefore the ordering of their products respects the lattice partial ordering. This makes the sequence $\mathbf{q}^{(\text{DA})}$ optimal. When $\gamma_{\mathbf{d},1} \geq -1/A_r(-1)$, our numerical results indicate that the conjecture is still true.

5 Numerical results

With the estimates given by our analysis in hand, we now compare these to our numerical results. Our numerical results use $r = 3$ and $\gamma_k = \mathbf{g}^k$, for $\mathbf{g} = 0.1, 0.5$, and 0.9 , to see how our rules $\mathbf{q}^{(\text{DA})}$ and $\mathbf{q}^{(\text{WW})}$ behave as the decay of the dimension weights is varied.

For the DA and WW weighted tensor product algorithms, each program run used $r = 3$; $\mathbf{g} = 0.1, 0.5$, or 0.9 ; a particular dimension \mathbf{d} , from $\mathbf{d} = 1$ to 16 ; a particular maximum 1-norm for multi-indices, typically 20 ; and a particular maximum number of points, up to $1\,000\,000$. The numerical results are potentially affected by two problems. First, if γ is close to zero, and the number of points is large, then the matrix used to compute the weights becomes ill-conditioned, and the weights may become inaccurate. Second, if the current squared error is close to zero, and the squared norm for the current multi-index is close to machine epsilon, then severe cancellation may occur. If either problem is detected, the calculation of the quadrature rule is terminated.

Figure 1 displays the typical convergence behaviour of the DA and WW rules for the cases examined. The particular case shown is that of \mathbb{T}^4 , $r = 3$, $\gamma_{\mathbf{d},k} = 0.5^k$. The number of points used varies from $\mathbf{n} = 1$ to $1\,000\,000$. In general, the DA algorithm has a cost no greater than that of the WW algorithm. Both are bounded by the WW bound of Theorem 1, and judging from the plot, the rates of convergence of both algorithms appear consistent with that of the bound. The WW cost bound itself has an asymptotic rate of convergence of $\mathcal{O}(\epsilon^{-1/3})$ for all of our cases.

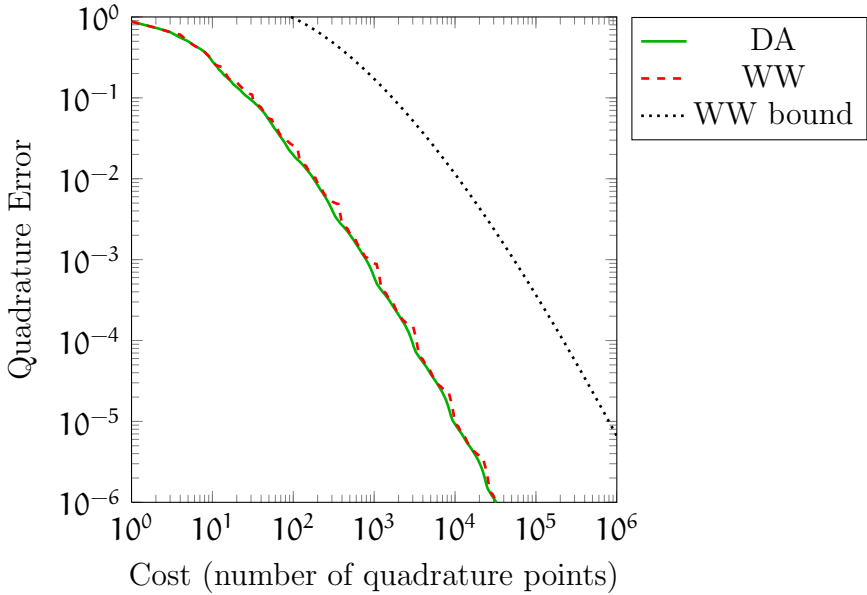


FIGURE 1: Error of DA and WW rules vs WW bound for \mathbb{T}^4 , $r = 3$, $\gamma_{d,k} = 0.5^k$.

For $\gamma_{d,k} = 0.1^k$, Figure 2 shows how the convergence rate of the error of the DA quadrature rules varies with dimension d , for $d = 1, 2, 4, 8$, and 16 . The cases $d = 8$ and $d = 16$ are almost indistinguishable on this graph. This is an example of the convergence in dimension.

Figure 3 shows the equivalent results for the DA quadrature rules for $\gamma_{d,k} = 0.9^k$. As d increases to 16 , the initial rate of convergence to zero of the error becomes much slower than that for $\gamma_{d,k} = 0.1^k$. This behaviour is expected, given the WW bound.

The most remarkable novel feature of Figure 3 is the series of distinct bumps and kinks, evident for $d = 4$ and $d = 8$. The main reason for these bumps and kinks for these values of d is the interaction between the lattice partial ordering constraint and the reversal of the order of norms described in Section 4. For $d = 8$, this results in the first 256 incremental rules each adding one point to

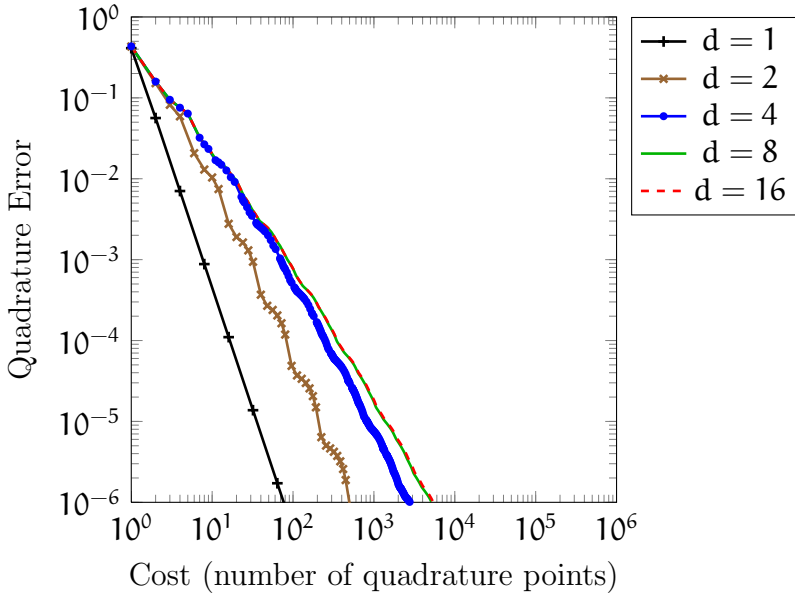


FIGURE 2: Error of DA rules for \mathbb{T}^d , $d = 1, 2, 4, 8, 16$; $r = 3$, $\gamma_{d,k} = 0.1^k$.

the main rule, with each successive multi-index corresponding to the binary expansion of the total number of points minus one. Figure 4 shows the average norm squared of each incremental rule, up to 512 points.

Acknowledgements The support of the Australian Research Council under its Centre of Excellence program is gratefully acknowledged.

References

- [1] B. Davey and H. A. Priestley. *Introduction to Lattices and Order*. Cambridge University Press, Cambridge, 1990. **C507**

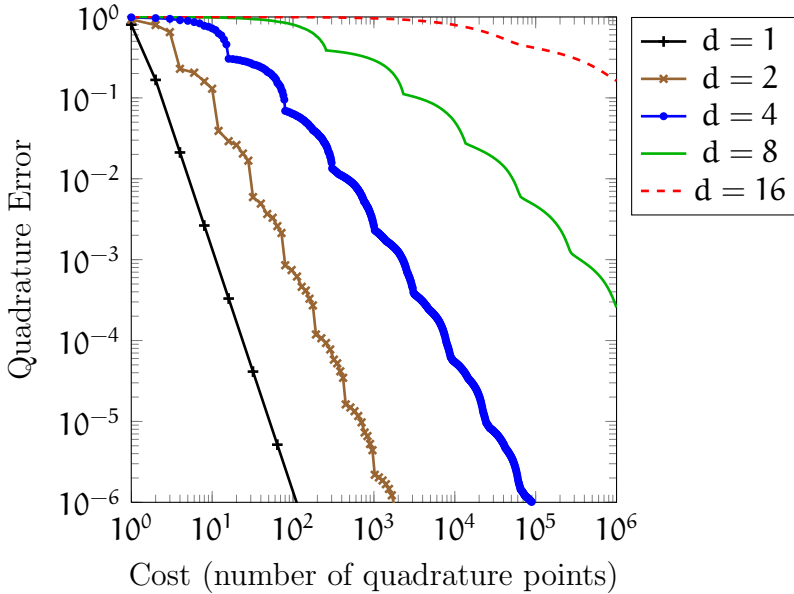


FIGURE 3: Error of DA rules for \mathbb{T}^d , $d = 1, 2, 4, 8, 16$; $r = 3$, $\gamma_{d,k} = 0.9^k$.

- [2] T. Gerstner and M. Griebel. Dimension-adaptive tensor product quadrature. *Computing*, 71:65–87, 2003. doi:10.1007/s00607-003-0015-5 C502
- [3] M. Hegland. Adaptive sparse grids. *ANZIAM Journal*, 44 (E):C335–C353, 2003. <http://anziamj.austms.org.au/ojs/index.php/ANZIAMJ/article/view/685> C502, C506
- [4] F. J. Hickernell and H. Woźniakowski. Tractability of multivariate integration for periodic functions. *Journal of Complexity*, 17(4):660–682, 2001. doi:10.1006/jcom.2001.0592 C502, C506
- [5] N. M. Korobov. Approximate evaluation of repeated integrals. (Russian). *Doklady Akademii Nauk SSSR*, 124(6):1207–1210, 1959. C501
- [6] N. M. Korobov. Computation of multiple integrals by the method of

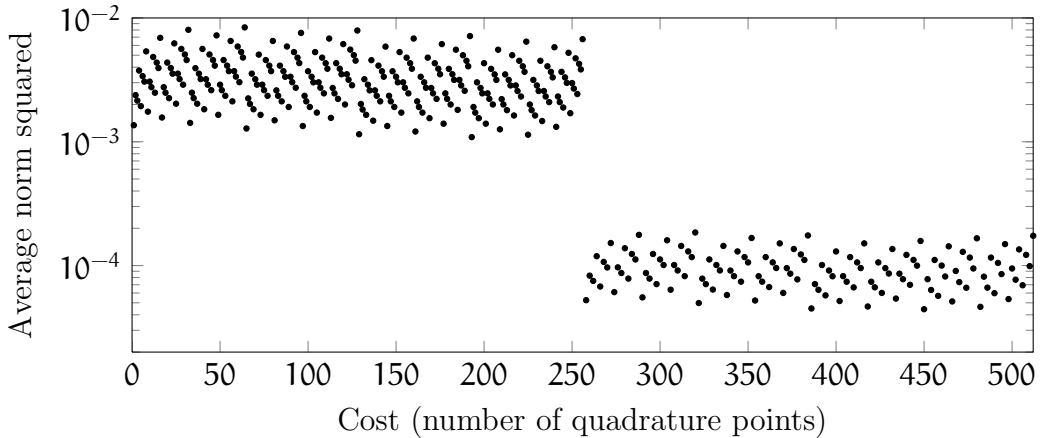


FIGURE 4: Average norm squared of DA incremental rules for \mathbb{T}^8 , $r = 3$, $\gamma_{8,k} = 0.9^k$.

- optimal coefficients. (Russian). *Vestnik Moskov. Univ. Ser. Mat. Meh. Astr. Fiz. Him.*, (4):19–25, 1959. [C501](#)
- [7] F. Y. Kuo and I. H. Sloan. Quasi-Monte Carlo methods can be efficient for integration over products of spheres. *Journal of Complexity*, 21(2):196–210, 2005. [doi:10.1016/j.jco.2004.07.0015](#) [C502](#), [C505](#)
- [8] I. H. Sloan and H. Woźniakowski. Tractability of multivariate integration for weighted Korobov classes. *Journal of Complexity*, 17(4):697–721, 2001. [doi:10.1006/jcom.2001.0599](#) [C502](#), [C506](#)
- [9] S. A. Smolyak. Quadrature and interpolation formulas for tensor products of certain classes of functions. *Dokl. Akad. Nauk SSSR*, 4:240–243, 1963. [C502](#)
- [10] G. W. Wasilkowski and H. Woźniakowski. Weighted tensor product algorithms for linear multivariate problems. *Journal of Complexity*, 15(3):402–447, 1999. [doi:10.1006/jcom.1999.0512](#) [C502](#), [C509](#), [C510](#), [C511](#)

Author addresses

1. **Markus Hegland**, Centre for Mathematics and its Applications, Australian National University.
2. **Paul C. Leopardi**, Centre for Mathematics and its Applications, Australian National University.
<mailto:paul.leopardi@anu.edu.au>