

# The rate of convergence of sparse grid quadrature on the torus

Markus Hegland <sup>\*</sup>      Paul C. Leopardi <sup>†</sup>

30 January 2011

## Abstract

This paper examines sparse grid quadrature on weighted tensor products (WTP) of reproducing kernel Hilbert spaces on the unit torus  $\mathbb{T}^d$ . We describe a WTP quadrature algorithm based on an algorithm of Hegland [2], and also formulate a version of Wasilkowski and Woźniakowski's WTP algorithm [6], here called the WW algorithm. We claim that our algorithm is generally lower in cost than the WW algorithm, and therefore both algorithms have the optimal asymptotic rate of convergence given by Theorem 3 of Wasilkowski and Woźniakowski [6]. Even so, if the dimension weights decay slowly enough, both algorithms need  $2^d$  points to produce a substantial reduction in quadrature error.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Setting</b>	<b>2</b>
<b>3</b>	<b>Algorithm</b>	<b>5</b>
<b>4</b>	<b>Analysis</b>	<b>6</b>
<b>5</b>	<b>Numerical results</b>	<b>9</b>

---

<sup>\*</sup>Centre for Mathematics and its Applications, Australian National University.

<sup>†</sup>Centre for Mathematics and its Applications, Australian National University. `mailto:paul.leopardi@anu.edu.au`

# 1 Introduction

This paper examines sparse grid quadrature on weighted tensor products of reproducing kernel Hilbert spaces (RKHS) on the unit torus. The empirical rates of convergence of the quadrature rules constructed here are compared to the theory of Wasilkowski and Woźniakowski [6].

The setting is essentially the same as that used by Kuo and Sloan [4] to examine quasi-Monte Carlo (QMC) quadrature on products of the sphere  $\mathbb{S}^s$ , but is confined to the special case of the unit circle  $\mathbb{S}^1$ . As noted by Kuo and Sloan, this case corresponds to the setting of weighted Korobov spaces. Rates of convergence and criteria for strong tractability of quadrature in these spaces have been well studied by Hickernell and Woźniakowski [3] and by Sloan and Woźniakowski [5].

There are a number of closely related dimension adaptive sparse grid algorithms, including those of Wasilkowski and Woźniakowski [6], Hegland [2], and Gerstner and Griebel [1]. Of these, the algorithm of Wasilkowski and Woźniakowski has the most well-developed theory of the rate of convergence of the worst case error.

## 2 Setting

Let  $\mathcal{D} \subset \mathbb{R}^{s+1}$  be a compact manifold with probability measure  $\mu$ . It follows that the constant function  $\mathbf{1}$ , with  $\mathbf{1}(x) = 1$  for all  $x \in \mathcal{D}$ , is integrable and  $\int_{\mathcal{D}} \mathbf{1}(x) d\mu = 1$ . 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

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

where the functions  $k_x$  are given by  $k_x(y) := K(x, y)$ , and where  $\langle \cdot, \cdot \rangle_H$  denotes the scalar product in  $H$ . We recognize  $H$  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 will assume that the quadrature points  $x_i$  are given. An optimal choice of weights  $w_i$  minimizes the worst case quadrature error, which is given by the norm  $\|\mathbf{1} - q\|_H$ . The optimal  $q^*$  is thus defined as

$$q^* := \operatorname{argmin}_q \{ \|\mathbf{1} - q\|_H \mid q \in \operatorname{span}\{k_{x_1}, \dots, k_{x_n}\} \}.$$

The weights of an optimal quadrature method are thus obtained by solving a linear system of equations with a matrix with elements being the values of the reproducing kernel  $K(x_i, x_j) = \langle k_{x_i}, k_{x_j} \rangle_H$ . 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 instead of (2), it also satisfies the more specific

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

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

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

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

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

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  $1 \geq \gamma_1 \cdots \geq \gamma_d \geq 0$ . This is an RKHS of functions on  $\mathcal{D}^d$  with reproducing kernel  $\mathcal{K}_d(x, y) := \prod_{k=1}^d (1 + \gamma_k \mathcal{K}(x_k, y_k))$  where  $x_k, y_k \in \mathcal{D}$  are the components of  $x, y \in \mathcal{D}^d$ . Moreover

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

where  $\mu_d$  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 [4], with  $s := 1$ . We take our domain  $\mathcal{D}$  to be the unit circle  $\mathbb{T} := \mathbb{S}^1 := \{x \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 1\}$ , and consider the real space  $L_2(\mathbb{T})$ . We

use the real Fourier basis defined by  $Y_{0,0}(x) := \mathbf{1}$ ,  $Y_{\ell,1}((\cos \theta, \sin \theta)) = \cos \ell \theta$ ,  $Y_{\ell,2}((\cos \theta, \sin \theta)) = \sin \ell \theta$ ,  $\ell = 1, \dots, \infty$ .

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

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

For positive weight  $\gamma$ , we define the RKHS

$$\begin{aligned} \mathbb{H}_{1,\gamma}^{(r)} &:= \{f \in L_2(\mathbb{T}) \mid \|f\|_{\mathbb{H}_{1,\gamma}^{(r)}} < \infty\}, \quad \text{where} \\ \langle f, g \rangle_{\mathbb{H}_{1,\gamma}^{(r)}} &:= \hat{f}_{0,0} \hat{g}_{0,0} + \gamma^{-1} \sum_{\ell=1}^{\infty} \sum_{m=1}^2 \ell^{2r} \hat{f}_{\ell,m} \hat{g}_{\ell,m}. \end{aligned}$$

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

$$\begin{aligned} K_{1,\gamma}^{(r)}(x, y) &:= 1 + \gamma A_r(x \cdot y), \quad \text{where for } z \in [-1, 1], \\ A_r(z) &:= \sum_{\ell=1}^{\infty} \frac{2}{\ell^{2r}} T_{\ell}(z), \end{aligned}$$

with  $T_{\ell}$  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

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

This space is equivalent to the weighted Korobov space of periodic functions on the unit cube, studied by Hickernell and Woźniakowski [3] and by Sloan and Woźniakowski [5]. Quadrature is therefore strongly tractable on  $\mathbb{H}_{d,\gamma}^{(r)}$  if and only if  $\sum_{k=1}^d \gamma_{d,k} < \infty$  as  $d \rightarrow \infty$ , and, in the case of exponentially decreasing weights, as studied here, the optimal worst-case error has an upper bound of order  $O(n^{-r})$ , where  $n$  is the cost of the quadrature rule in terms of the number of points [5]. The order of the lower bound is known to be the same as that of the non-periodic setting [3].

### 3 Algorithm

Algorithm 1 studied here is an adaptation of the dimension-adaptive algorithm for the solution of variational problems suggested by Hegland [2]. 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  $\mathcal{H}^\gamma$ . We will only consider up to a maximum of  $n$  points which we denote by  $x_1, \dots, x_n \in \mathcal{D}$ . The quadrature rules for  $\mathcal{H}^\gamma$  are then defined as some element of  $V_i^\gamma = \text{span}\{k_{x_1}^\gamma, \dots, k_{x_{n_i}}^\gamma\} \subset \mathcal{H}^\gamma$ . We denote the optimal rule in  $V_i^\gamma$  by  $q_i^\gamma$ . Now define the pair-wise orthogonal spaces  $U_i^\gamma$  by  $U_0^\gamma = V_0^\gamma$ , and by the orthogonal decomposition  $V_{i+1}^\gamma = V_i^\gamma \oplus U_{i+1}^\gamma$ . Using the fact that the  $q_i^\gamma$  are optimal, one can see that

$$\delta_{i+1}^\gamma := q_{i+1}^\gamma - q_i^\gamma \in U_{i+1}^\gamma$$

and  $\delta_0 := q_0^\gamma \in U_0^\gamma = V_0^\gamma$ . Note that one has

$$U_{i+1}^\gamma \neq \text{span}\{k_{x_{n_{i+1}}}, \dots, k_{x_{n_{i+1}}}\}.$$

This is the fundamental reason why one needs the admissibility condition discussed in this section.

A generalized sparse grid quadrature rule is then of the form

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

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  $\bar{I} = \{i \mid i \leq j \text{ for some } j \in I\}$ , where the comparison  $i \leq j$  has to hold for all components of  $i$  and  $j$ . One can then show that an optimal  $q \in V_I$  is obtained as

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

We will now describe our algorithm 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

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

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

---

**Algorithm 1:** The dimension adaptive algorithm.

---

**Data:** accuracy  $\epsilon$ , incremental rules  $\Delta_j$  and their costs  $\nu_j$  for  $j \in \mathbb{N}^d$

**Result:**  $\epsilon$  approximation  $q^{(HL)}$  and index set  $I$

$I := \{0\}; q := \Delta_0;$

**while**  $\|\mathbf{1} - q\| > \epsilon$  **do**

$i := \operatorname{argmax}_j \{\|\Delta_j\|^2 / \nu_j \mid I \cup \{j\} \text{ is a down-set}\};$   
 $I := I \cup \{i\}; q := q + \Delta_i;$

---

## 4 Analysis

We first describe the situation on a single circle where, if  $\gamma$  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 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 HL 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 HL rules  $\Delta_j$  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  $q_j := q_{r,\gamma}(S_j)$ , based on a set of equally spaced points  $S_j$  on the unit circle, with  $n_j := |S_j|$ . The series expansion of the function  $A_r$  then yields the following error bound for quadrature on  $\mathbb{H}_{1,\gamma}^{(r)}$ , namely

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

For our numerical example, we also have  $n_j = 2^j$ . Since  $e^2(q_j) \geq 0$ , we can

therefore show that  $\|q_j - q_{j-1}\|_{\mathbb{H}_{1,\gamma}^{(r)}} \leq \sqrt{\gamma}C 2^{-rj}$ , where

$$C := 2^{1-r} \sqrt{r/(2r-1)}. \quad (3)$$

In our setting, and our notation, the criteria needed by Wasilkowski and Woźniakowski [6, Theorem 3] criteria become

$$n_{j+1} D^{jp} \leq 1, \quad \text{for all } j \geq 1, \quad (4)$$

and

$$\|q_j - q_{j-1}\|_{\mathbb{H}_{1,\gamma}^{(r)}} \leq \sqrt{\gamma}CD^j, \quad \text{for all } j \geq 1. \quad (5)$$

for some  $D \in (0, 1)$  and some positive  $C$  and  $p$ . For the points used by our HL quadrature rules, these criteria hold with  $C$  as per (3),  $D = 2^{-r}$ , and  $p = 1/r$ .

We will now describe a second variant of WTP quadrature,  $q^{(WW)}$  on  $\mathbb{H}_{d,\gamma}^{(r)}$ , identical to the sequence of quadrature rules  $q^{(HL)}$  described in Section 3 above, 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 [6, Section 5]. As a consequence of (5), we have

$$\begin{aligned} \|\Delta_j\|_{\mathbb{H}_{d,\gamma}^{(r)}} &= \prod_{k=1}^d \left\| \delta_{j_k}^{(k)} \right\|_{\mathbb{H}_{1,\gamma_{d,k}}^{(r)}} \leq b(d, j), \quad \text{where} \\ b(d, j) &:= \prod_{k=1}^d (\sqrt{\gamma_{d,k}} CD^{j_k})^{1-\delta_{0,j_k}}. \end{aligned}$$

Let  $(\xi_{d,k})$ ,  $k = 1, \dots, d$ , be a sequence of positive numbers. In contrast to Wasilkowski and Woźniakowski [6, Section 5], we do not stipulate that  $\xi_{d,k} = 1$ . Define

$$\xi(d, j) := \prod_{k=1}^d \xi_{d,k}^{1-\delta_{0,j_k}}. \quad (6)$$

We therefore have  $b(d, j)/\xi(d, j) \rightarrow 0$  as  $\|j\|_1 \rightarrow \infty$ . We order the incremental rules in order of non-decreasing  $b(d, j)/\xi(d, j)$  for each multi-index  $j$ , creating an order on the multi-indices  $j^{(WW)}(h)$ . We adjust  $\xi(d, k)$  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  $I_N^{(WW)} := \{j^{(WW)}(1), \dots, j^{(WW)}(N)\}$ , and define the quadrature rule

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

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

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

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

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

Finally, we define

$$q_\epsilon^{(WW)} := \sum_{j \in I_{N(\epsilon, d)}^{(WW)}} \Delta_j. \quad (7)$$

We can now present our version of Wasilkowski and Woźniakowski's main theorem on the error and cost of WTP quadrature [6, Theorem 3].

**Theorem 1.** *Let  $\eta \in (0, 1)$ . Then the quadrature rule  $q_{\epsilon, d}^{(WW)}$  defined by (7) has worst-case quadrature error  $e(q_\epsilon^{(WW)}) \leq \epsilon$ , and its cost (in number of quadrature points) is bounded by*

$$\begin{aligned} \text{cost}(q_{\epsilon, d}^{(WW)}) &\leq C(d, \epsilon) \left(\frac{1}{\epsilon}\right)^{p/(1-\eta)}, \quad \text{where} \\ C(d, \epsilon) &:= \frac{\xi_{d,1}^p \prod_{k=2}^d (1 + C^p \gamma_{d,k}^{p/2} / \xi_{d,k}^p g(k, \epsilon)) f(k, \epsilon)^p}{(1-D^p)(1-D^2)^{p/(2(1-\eta))}}, \\ f(i, \epsilon) &:= \left(1 + C^{2\eta} \gamma_{d,i}^\eta \xi_{d,i}^{2(1-\eta)} \frac{D^{2\eta}}{1-D^{2\eta}}\right)^{1/(2(1-\eta))}, \\ g(k, \epsilon) &:= \left\lfloor \frac{\log \left( C \gamma_{d,k}^{1/2} / (\xi_{d,k} (1-D^2))^{1/(2(1-\eta))} \prod_{i=2}^k (f(i, \epsilon)) \epsilon^{-1/(1-\eta)} \right)}{\log D^{-1}} \right\rfloor_+. \end{aligned}$$

By  $\lfloor x \rfloor_+$ , we mean  $\max(0, x)$ .



Wasilkowski and Woźniakowski's proof, with  $s := 2, \alpha := 1$ , applies directly to our Theorem 1, once the change in  $\xi_{d,1}$  is taken into account.

For exponentially decreasing dimension weights  $\gamma_{d,k}$ , Theorem 4 of Wasilkowski and Woźniakowski [6] shows that the  $q^{(WW)}$  rules are strongly polynomial.

We conjecture that our sequence of rules  $q^{(HL)}$  is in general more efficient than  $q^{(WW)}$ , since  $q^{(HL)}$  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 minimizing the error of the overall quadrature rule with respect to its cost in terms of function evaluations. Our conjecture is true when  $\gamma_{d,1} < -1/A_r(-1)$ , 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  $q^{(HL)}$  optimal. When  $\gamma_{d,1} \geq -1/A_r(-1)$ , our numerical results indicate that the conjecture is still true.

## 5 Numerical results

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

For the HL and WW WTP algorithms, each program run used  $r = 3$ ;  $g = 0.1, 0.5$ , or  $0.9$ ; a particular dimension  $d$ , from  $d = 1$  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  $\gamma$  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 HL and WW WTP rules for the cases examined. The particular case shown is that of  $\mathbb{T}^4$ ,  $r = 3$ ,  $\gamma_{d,k} = 0.5^k$ . The number of points used varies from  $n = 1$  to  $1\,000\,000$ . In general, the HL WTP algorithm has a cost no greater than that of the WW WTP algorithm. Both are bounded by the WW bound of Theorem 1, and judging from the plot, the rates of convergence of both algorithms appears consistent with that of the bound. The WW cost bound itself has

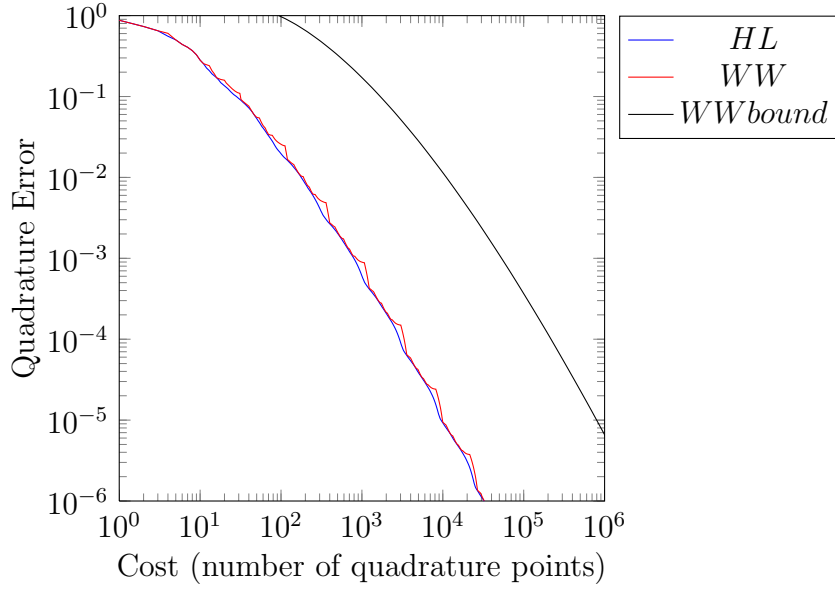


Figure 1: Error of HL and WW WTP rules vs WW bound for  $\mathbb{T}^4$ ,  $r = 3$ ,  $\gamma_{d,k} = 0.5^k$ .

an asymptotic rate of convergence of  $O(\epsilon^{-1/3})$  for all of our cases.

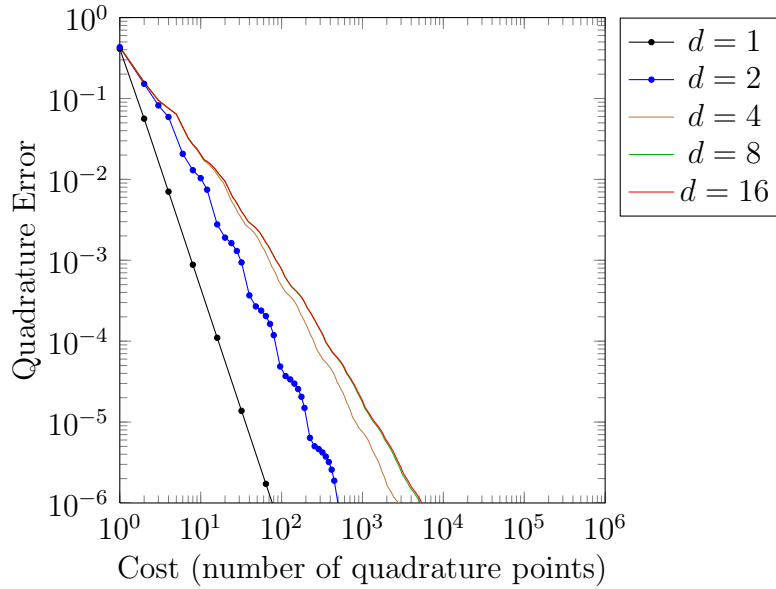


Figure 2: Error of HL WTP rules for  $\mathbb{T}^d$ ,  $d = 1, 2, 4, 8, 16$ ;  $r = 3$ ,  $\gamma_{d,k} = 0.1^k$ .

For  $\gamma_{d,k} = 0.1^k$ , Figure 2 shows how the convergence rate of the error of

the HL WTP 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.

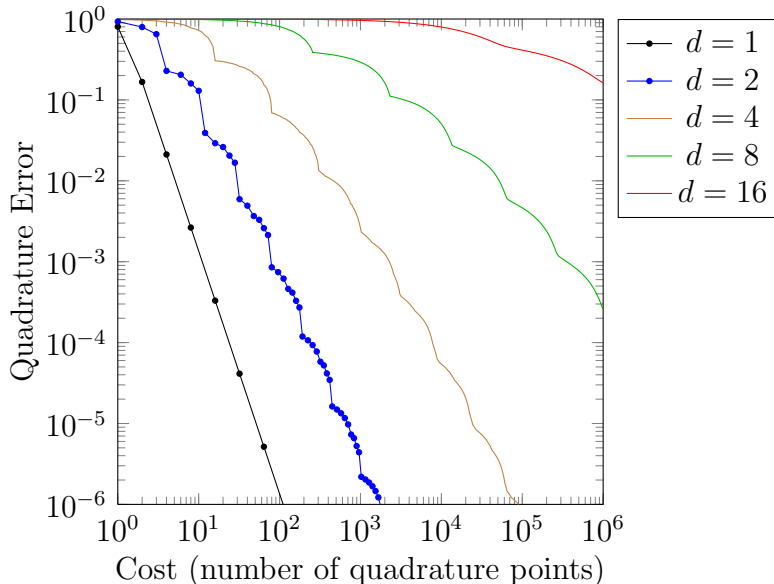


Figure 3: Error of HL WTP rules for  $\mathbb{T}^d$ ,  $d = 1, 2, 4, 8, 16$ ;  $r = 3$ ,  $\gamma_{d,k} = 0.9^k$ .

Figure 3 shows the equivalent results for the HL WTP quadrature rules for  $\gamma_{d,k} = 0.9^k$ . As  $d$  increases to 16, the initial rate of convergence to zero of the error much becomes 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 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.

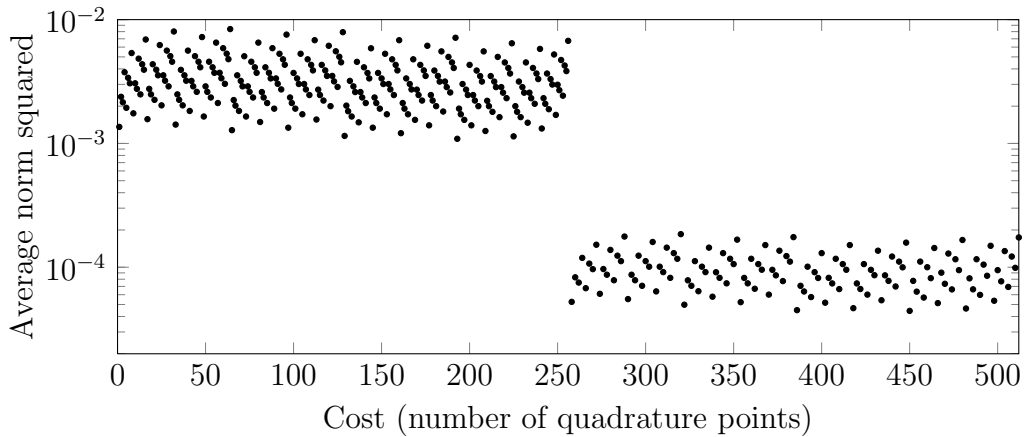


Figure 4: Average norm squared of HL WTP incremental rules for  $\mathbb{T}^8$ ,  $r = 3$ ,  $\gamma_{8,k} = 0.9^k$ .

## References

- [1] T. Gerstner and M. Griebel. Dimension-adaptive tensor product quadrature. *Computing*, 71:65–87, 2003.
- [2] M. Hegland. Adaptive sparse grids. *ANZIAM Journal*, 44 (E):C335–C353, 2003.
- [3] F. J. Hickernell and H. Woźniakowski. Tractability of multivariate integration for periodic functions. *Journal of Complexity*, 17:660–682, 2001.
- [4] 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, April 2005.
- [5] I. H. Sloan and H. Woźniakowski. Tractability of multivariate integration for weighted korobov classes. *Journal of Complexity*, 17:697–721, 2001.
- [6] G. W. Wasilkowski and H. Woźniakowski. Weighted tensor product algorithms for linear multivariate problems. *Journal of Complexity*, 15(3):402–447, September 1999.