Sparse grid quadrature on products of spheres

Markus Hegland[a], Paul Leopardi[a,*]

[a] *Centre for Mathematics and its Applications, Australian National University.*

Abstract

This paper examines sparse grid quadrature on weighted tensor products (WTP) of reproducing kernel Hilbert spaces on products of the unit sphere $\mathbb{S}^2$. We describe a WTP quadrature algorithm based on an algorithm of Hegland [1], and also formulate a version of Wasilkowski and Woźniakowski's WTP algorithm [2], here called the WW algorithm. We prove that our algorithm is optimal and therefore 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 [2]. Even so, the initial rate of convergence can be very slow, if the dimension weights decay slowly enough.

Keywords: reproducing kernel Hilbert spaces, quadrature, tractability, sparse grids, knapsack problems, spherical designs

1. Introduction

This paper examines sparse grid quadrature on weighted tensor products of reproducing kernel Hilbert spaces (RKHS) on the unit sphere $\mathbb{S}^2 \subset \mathbb{R}^3$. As per our previous paper on sparse grid quadrature on the torus [3], the empirical rates of convergence of the quadrature rules constructed here are compared to the theory of Wasilkowski and Woźniakowski [2].

The setting is the same as that used by Kuo and Sloan [4], and Hesse, Kuo and Sloan [5] to examine quasi-Monte Carlo (QMC) quadrature on products of the sphere $\mathbb{S}^2$, but here we examine quadrature with arbitrary weights.

As noted in our previous paper [3], rates of convergence and criteria for strong tractability of quadrature with arbitrary weights are known in the case of weighted Korobov spaces on the unit torus [6, 7]. As far as we know, this paper is the first to examine the analogous questions for quadrature with arbitrary weights on the corresponding spaces on products of spheres.

The algorithms we examine are an adaptation of the algorithm used in our previous paper [1, 3], and an adaptation of the WTP algorithm of Wasilkowski and Woźniakowski [2]. We examine these algorithms theoretically, giving bounds

*Corresponding author
Email address: paul.leopardi@anu.edu.au (Paul Leopardi)

for the asymptotic convergence rate of quadrature error in the worst case. We also examine the algorithms empirically, via a small number of numerical examples.

Arbitrary weight quadrature on products of spheres is interesting, not just for purely theoretical reasons, such as comparison with the results for equal weight quadrature, but also for practical reasons. Integration over products of the unit sphere is equivalent to multiple integration over the unit sphere. Such multiple integrals can be approximated in a number of ways, including Monte Carlo methods.

Applications of tensor product spaces on spheres and approximate integration over products of spheres include quantum mechanics [8], and transport and multiple scattering problems in various topic areas, including acoustics [9], optical scattering problems [10, 11, 12], and neutron transport problems [13].

One prototypical problem to be solved is scattering by a sequence of spheres. This can be modelled using a multiple integral of a function on the product of the spheres. The decay in the weights of successive spheres could model the decreasing influence of scattering on each successive sphere, as opposed to just cutting off the calculation after an arbitrary number of scatterings.

The remainder of this paper is organized as follows. Section 2 describes our weighted tensor product space setting in detail. Section 3 describes the optimization problem involved in dimension adaptive sparse grid quadrature. Section 4 introduces the dimension adaptive (DA) algorithm and shows that it is optimal in a certain sense. Section 5 analyses a version of the WTP algorithm of Wasilkowski and Woźniakowski, and compares its theoretical rate of convergence with that the da algorithm. Section 6 contains numerical results, comparing the two algorithms, and showing how the da algorithm performs as the dimension is increased.

2. Setting

The general setting used in our previous paper [3] applies equally well here. We repeat it here for the sake of completeness.

Let $\mathcal{D} \subset \mathbb{R}^{s+1}$ be a compact 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 $f : \mathcal{D} \to \mathbb{R}$, with inner product $\langle \cdot, \cdot \rangle_H$, and kernel K, with the following properties.

1. Every $f \in H$ is integrable and

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

2. For every $x \in \mathcal{D}$, the function $k_x \in H$, given by $k_x(y) := K(x, y)$, satisfies

$$f(x) = \langle k_x, f \rangle_H, \quad \text{for all } f \in H. \tag{2}$$

We recognize H as a reproducing kernel Hilbert space (RKHS). In this framework, quadrature rules 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 $e(q)$, which is given by the norm $\|\mathbf{1} - q\|_H$. The optimal q^* is thus defined as

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

The weights of an optimal quadrature rule are thus obtained by solving a linear system of equations with a matrix whose elements are the values of the reproducing kernel $K(x_i, x_j) = \langle k_{x_i}, k_{x_j} \rangle_H$. 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 (2), but as well as (1), it is also assumed to satisfy (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 $\mathcal{H}^\gamma$, which consists 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

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

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 $\mathcal{H}$.

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 $1 \geqslant \gamma_1 \cdots \geqslant \gamma_d \geqslant 0$. 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, since $\mathcal{H}_d$ is an RKHS, one has

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

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.

The specific setting for this paper is that of Kuo and Sloan [4], with $s := 2$. We now describe this setting. We take our domain $\mathcal{D}$ to be the unit sphere $\mathbb{S}^2 := \{x \in \mathbb{R}^3 \mid x_1^2 + x_2^2 + x_3^2 = 1\}$, and consider the real space $L_2(\mathbb{S}^2)$. We use

a basis of real spherical harmonics $Y_{\ell,m}(x)$, $\ell = 0, \ldots, \infty$, $m = -\ell, \ldots, \ell$. For $f \in L_2(\mathbb{S}^2)$, we expand f in the Fourier series

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

For a positive weight γ, we define the RKHS

$$\mathbb{H}_{1,\gamma}^{(r)} := \{f : \mathbb{S}^2 \to \mathbb{R} \mid \|f\|_{\mathbb{H}_{1,\gamma}^{(r)}} < \infty\},$$

where

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

Kuo and Sloan [4] show that the reproducing kernel of $\mathbb{H}_{1,\gamma}^{(r)}$ is

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

where P_ℓ is the Legendre polynomial of degree ℓ. Convergence of A_r needs $r > 3/2$.

For $\gamma := (\gamma_{d,1}, \ldots, \gamma_{d,d})$, 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 RKHS on $(\mathbb{S}^2)^d$, with reproducing kernel

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

Kuo and Sloan [4] studied equal weight (QMC) quadrature on the space $\mathbb{H}_{d,\gamma}^{(r)}$, and found that it is strongly tractable if and only if $\sum_{k=1}^{d} \gamma_{d,k} < \infty$ as $d \to \infty$. Hesse, Kuo and Sloan [5] constructed sequences of QMC rules on this space, and proved that their worst case error converges at least as quickly as the Monte Carlo error rate of $\mathrm{O}(n^{-1/2})$, where n is the cost of the quadrature rule in terms of the number of points.

The work of Hickernell and Woźniakowski [6], and Sloan and Woźniakowski [7], on the weighted Korobov space of periodic functions on the unit cube, and the work of Wasilkowski and Woźniakowski [2] on WTP quadrature on non-periodic functions on the unit cube, suggests bounds on the worst case error for our

case of quadrature with arbitrary weights on the space $\mathbb{H}^{(r)}_{d,\gamma}$. We might expect quadrature with arbitrary weights on this space to be strongly tractable if and only if $\sum_{k=1}^{d} \gamma_{d,k} < \infty$ as $d \to \infty$, and, in the case of exponentially decreasing weights, as studied here, we might therefore expect the optimal worst-case error to have an upper bound of order $\mathrm{O}(n^{-r})$, and a lower bound of order $\Omega(n^{-r})$, for $r > 3/2$.

Our analysis below shows that these expectations for quadrature with arbitrary weights on $\mathbb{H}^{(r)}_{d,\gamma}$ can be met, and specifically, that our algorithm satisfies the upper bound suggested here.

3. Optimization Problem

We first describe the optimization problem in the general RKHS setting, as given in Section 2.

We assume that a sequence of quadrature points $x_1, x_2, \ldots \in \mathcal{D}$, and a sequence of positive integers $n_0 < n_1 < \ldots$ are given and are the same for all spaces $\mathcal{H}^\gamma$. The quadrature rules for $\mathcal{H}^\gamma$ are then defined as some element of $V_j^\gamma := \mathrm{span}\{k^\gamma_{x_1}, \ldots, k^\gamma_{x_{n_j}}\} \subset \mathcal{H}^\gamma$. We denote the optimal rule in V_j^γ by q_j^γ. Now define the pair-wise orthogonal spaces U_j^γ by $U_0^\gamma := V_0^\gamma$, and by the orthogonal decomposition $V_{j+1}^\gamma = V_j^\gamma \oplus U_{j+1}^\gamma$. Using the fact that the q_j^γ are optimal, one can see that

$$\delta_{j+1}^\gamma := q_{j+1}^\gamma - q_j^\gamma \in U_{j+1}^\gamma$$

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

$$U_{j+1}^\gamma \neq \mathrm{span}\{k_{x_{1+n_j}}, \ldots, k_{x_{n_{j+1}}}\}.$$

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

We use the notation $\mathbb{J} := \mathbb{N}^d$, treating elements of $\mathbb{J}$ as indices, with a partial order such that for $i, j \in \mathbb{J}$, $i \leqslant j$ if and only if $i_h \leqslant j_h$ for all components.

For a index $i \in \mathbb{J}$, let $\downarrow i$ denote the *down-set* of i, defined by $\downarrow i := \{j \in \mathbb{J} \mid j \leqslant i\}$ [14, p. 13]. Subsets of $\mathbb{J}$ are partially ordered by set inclusion. For a subset $I \subset \mathbb{J}$, let $\downarrow I$ denote the down-set of I, defined by $\downarrow I := \bigcup_{i \in I} \downarrow i$. Then $\downarrow I$ is the smallest set $J \supseteq I$ such that if $i \in J$ and $j \leqslant i$ then $j \in J$. Thus $\downarrow\downarrow I = \downarrow I$.

A 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

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

One can then show that an optimal $q \in V_I$ is obtained as

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

Thus both V_I and q_I^* are obtained in terms of the down-set $\downarrow I$, effectively restricting the choice of the set I to index sets which are also down-sets.

This leads us to defining the concept of an *optimal* index set. An optimal index set is one which minimizes the error for a given cost, or minimizes the cost for a given error. Here, the cost is the number of quadrature points, which is the dimension of V_I.

We now make this definition precise. We first define $\nu_{j_k}^{(k)} := \dim U_{j_k}^{\gamma_{d,k}}$ and $\delta_{j_k}^{(k)} := \delta_{j_k}^{\gamma_{d,k}}$. For the remainder of this section, we use $\epsilon \in (0,1)$ to denote the required upper bound on quadrature error. The optimization problem then uses the following definitions.

Definition 1. *For index $j \in \mathbb{J}$, define*

$$\nu_j := \prod_{k=1}^{d} \nu_{j_k}^{(k)}, \quad \Delta_j := \bigotimes_{k=1}^{d} \delta_{j_k}^{(k)}, \quad p_j := \|\Delta_j\|^2, \quad r_j := p_j/\nu_j.$$

For subset $I \subset \mathbb{J}$, define

$$\nu(I) := \sum_{j \in I} \nu_j, \quad p(I) := \sum_{j \in I} p_j.$$

Also, define $P := 1 - \epsilon^2$.

Here, j_k is the kth component of the index j.

Due to the properties of $\nu_{j_k}^{(k)}$ and Δ_j, ν and p satisfy

$$\nu_j, \nu(I) \in \mathbb{N}_+, \quad 0 < p_j < 1, \quad 0 < p(I) < 1, \quad p(\mathbb{J}) = 1. \tag{3}$$

We now consider the following optimization problem, posed as a minimization problem on the variable $I \subset \mathbb{J}$.

Optimization Problem 1.

$$\textit{Minimize} \quad \nu(\downarrow I), \quad \textit{subject to} \quad p(I) \geqslant P,$$

for some $0 < P < 1$, where ν and p satisfy (3).

In other words, given a required upper bound ϵ on the quadrature error, the problem is to find the subset $I \subset \mathbb{J}$ with the smallest cost $\nu(\downarrow I) = \sum_{j \in \downarrow I} \nu_j$, satisfying the constraint $1 - \sum_{j \in I} p_j \leqslant \epsilon^2$.

Optimization Problem 1 can have multiple solutions, since for $H, I, J \subset \mathbb{J}$ if $J = \downarrow H = \downarrow I$ and both $p(H) \geqslant P$ and $p(I) \geqslant P$ then both I and J are solutions to Optimization Problem 1. The following problem breaks this tie.

Optimization Problem 2.

$$\textit{Maximize} \quad p(I) \quad \textit{subject to } I \textit{ solving Optimization Problem 1.}$$

The solution of Optimization Problem 2 satisfies an admissibility condition.

Lemma 1. *If I is a solution of Optimization Problem 2, then*

$$I = \downarrow I. \tag{4}$$

Proof. Let $J = \downarrow I$, where I is a solution of Optimization Problem 2, and therefore of Optimization Problem 1. Then $I \subset J$ and thus J satisfies the constraints of Optimization Problem 1, since $p_i > 0$. Therefore J is also a solution of Optimization Problem 1, since $\nu(\downarrow J) = \nu(\downarrow I)$. If $I \subsetneq J$, it follows from $p_i > 0$ that $p(J) > p(I)$, and so I cannot be optimal. Therefore $I = J$. □

In view of the admissibility condition, we reformulate Optimization Problem 2 as:

Optimization Problem 3.

$$\textit{Minimize} \quad \nu(I), \quad \textit{subject to} \quad I = \downarrow I, \quad \textit{and} \quad p(I) \geqslant P,$$

for some $0 < P < 1$, where ν and p satisfy (3).

As pointed out by (e.g.) Griebel and Knapek [15], some sparse grid problems can be formulated and solved as knapsack problems. The resulting solution is optimal in terms of estimated profit for a given "weight".

We call Optimization Problem 3 a *down-set-constrained* binary knapsack problem. Each item in the knapsack is a product difference rule. The profit for each item is just the squared norm of a product difference rule, and this can be calculated precisely. The "weight" of each item is the number of extra points the product difference rule contributes to the overall quadrature rule, assuming that the admissibility condition applies. The relationships between Optimization Problem 3 and other more well-known knapsack problems are described in more detail Section 4.

4. Algorithm

The dimension adaptive (DA) algorithm to choose the set I for $\mathcal{D} = \mathbb{S}^1$, described in [3] is quite general, and applies equally well to the case here, where

Algorithm 1: The dimension adaptive (DA) algorithm.

Data: error ϵ, incremental rules Δ_j and their costs ν_j for $j \in \mathbb{J}$
Result: ϵ approximation q and index set I
$I := \{0\}$; $q := \Delta_0$;
while $\|\mathbf{1} - q\| > \epsilon$ **do**
  $j := \operatorname{argmax}_i\{r_i \mid i \notin I \text{ and } I \cup \{i\} \text{ is a down-set}\}$;
  $I := I \cup \{j\}$; $q := q + \Delta_j$;

$\mathcal{D}$ is $\mathbb{S}^2$. For the sake of completeness, we repeat the algorithm here, with a slight change in notation, as Algorithm 1.

This is a greedy algorithm for Optimization Problem 3.

Under certain conditions on ν and p, Algorithm 1 solves Optimization Problem 3. To see this, consider *monotonicity* with respect to the lattice partial ordering of $\mathbb{J}$.

Definition 2. *The function $p \in \mathbb{R}_+^{\mathbb{J}}$ is* monotonically decreasing *if $i < j$ implies that $p_i \geqslant p_j$. If $i < j$ implies that $p_i > p_j$, then $p \in \mathbb{R}_+^{\mathbb{J}}$ is* strictly decreasing*. The definitions of "monotonically increasing" and "strictly increasing" are similar.*

Using Definition 2, the following theorem holds.

Theorem 2. *If $p \in \mathbb{R}_+^{\mathbb{J}}$ is strictly decreasing and $\nu \in \mathbb{N}_+^{\mathbb{J}}$ is monotonically increasing, then Algorithm 1 yields a quadrature rule q and index set I such that I solves the down-set-constrained knapsack Optimization Problem 3, for $P = p(I) = \|\mathbf{1} - q\|^2$.*

The proof of Theorem 2 presented below proceeds in these stages.

1. We introduce a related binary knapsack problem, and show that if I is a solution of the binary knapsack problem, and I is also a down-set, then I is a solution of Optimization Problem 3.
2. We define the efficiency r_i, describe a greedy algorithm for the binary knapsack problem (Algorithm 2 below), and show that if the efficiency is strictly decreasing, then each set I produced by the greedy algorithm is a solution of the binary knapsack problem, and I is also a down-set, and is therefore also a solution of Optimization Problem 3.
3. We show that if the efficiency is strictly decreasing, then Algorithm 2 produces the same sequence of sets as Algorithm 1.

A binary (0/1) knapsack problem [16] related to the Optimization Problem 3 is:

Optimization Problem 4.

$$\textit{Minimize} \quad \nu(I), \quad \textit{subject to} \quad p(I) \geqslant P,$$

for some $0 < P < 1$, where ν and p satisfy (3)*.*

Usually a binary knapsack problem is posed as a maximization problem, where the selection is from a finite set of items. Here we have a minimization problem and a countably infinite set. A finite minimization problem can always be posed as an equivalent maximization problem [17, p. 15]. In the case of Problem 4 this cannot be done, because the quantity to be maximized (the sum of the costs of the elements not in the knapsack) would be infinite. Instead, we must deal directly with the minimization form.

We now formulate a converse of Lemma 1.

Lemma 3. *If I is a solution of the Optimization Problem 4, and I also satisfies the admissibility condition $I = {\downarrow} I$, then I is a solution of Optimization Problem 3.*

Proof. I satisfies the admissibility condition ${\downarrow} I = I$ and consequently $p(I) \geqslant P$. It follows that I satisfies the constraints of Optimization Problem 3 and thus minimizes p under these constraints, i.e., is a solution of Optimization Problem 3. □

This justifies our calling Optimization Problem 3 a down-set-constrained knapsack problem.

If, in Optimization Problem 4 we identify each set $I \subset \mathbb{J}$ with its *indicator function* $\mathcal{I} \in \{0,1\}^{\mathbb{J}}$, where $\mathcal{I}_i = 1$ if and only if $i \in I$, we obtain a more usual formulation of the binary knapsack problem:

Optimization Problem 5.

$$\textit{Minimize} \quad \sum_{i \in \mathbb{J}} \nu_i \mathcal{I}_i, \quad \textit{subject to} \quad \sum_{i \in \mathbb{J}} p_i \mathcal{I}_i \geqslant P, \quad \mathcal{I} \in \{0,1\}^{\mathbb{J}},$$

for some $0 < P < 1$, where ν and p satisfy (3).

Solving the binary knapsack problem is hard in general, but for certain values of the constraint P, a greedy algorithm yields the solution. These values are exactly the values for which the solution of the binary knapsack problem equals the solution of the *continuous* knapsack problem, which uses the same objective function ν as Optimization Problem 5, and relaxes the constraints $\mathcal{I}_i \in \{0,1\}$ to $\mathcal{I}_i \in [0,1]$. Dantzig [16] gives a graphical proof of this for the classical binary knapsack problem – the finite maximization problem. Martello and Toth [17, Theorem 2.1, p. 16] give an explicit solution for the continous problem, and a more formal proof.

The greedy algorithm for Optimization Problem 4 is based on the *efficiency* $r_j := p_j/\nu_j$. The algorithms generates the initial values of an enumeration $j^{(t)}$ of $\mathbb{J}$, $t \in \mathbb{N}_+$, satisfying

$$r_{j^{(t)}} \geqslant r_{j^{(t+1)}}.$$

The algorithm recursively generates $I_{(t)}$ from $I_{(t-1)}$, until for some T the condition

$$p(I_{(T-1)}) < P \leqslant p(I_{(T)})$$

holds, where

$$I_{(t)} := \bigcup_{s=1}^{t} j^{(s)}.$$

The greedy algorithm is therefore as follows.

Algorithm 2: The greedy algorithm for Optimization Problem 4.

Data: error ϵ, incremental rules Δ_j and their costs ν_j for $j \in \mathbb{J}$
Result: ϵ approximation q and index set I
$I := \{0\}$; $q := \Delta_0$;
while $\|\mathbf{1} - q\| > \epsilon$ **do**
  $j := \text{argmax}_i\{r_i \mid i \notin I\}$;
  $I := I \cup \{j\}$; $q := q + \Delta_j$;

This algorithm has the following properties.

Lemma 4. *For any $0 < P < 1$, Algorithm 2 terminates for some $t = T$.*

For each $t \geqslant 1$, the set generated by Algorithm 2, $I_{(t)}$ is the solution to Optimization Problem 4 for $P = p(I_{(t)})$.

Proof. The algorithm terminates because $j^{(t)}$ is an enumeration of $\mathbb{J}$ and therefore $\sum_{t=1}^{\infty} p_{j^{(t)}} = 1$, since $\sum_{t=1}^{\infty} \Delta_{j^{(t)}} = \mathbf{1}$, but $P < 1$.

When $p(I_{(t)}) = P$ the constraints of Optimization Problem 4 are satisfied. Furthermore, as the method used the largest r_i, the objective function ν is minimised for Optimization Problem 4. A more detailed proof can be constructed along the lines of the proof of Theorem 2.1 of Martello and Toth [17]. □

The construction of the enumeration used in Algorithm 2 requires sorting an infinite sequence and is thus not feasible in general, but, in the case where p is strictly decreasing and ν is monotonically increasing, the enumeration can be done recursively in finite time.

Here and in the following we say that i is a *minimal element* of a subset of $\mathbb{J}$ if there are no elements $j < i$ in that subset. The minimum is thus with respect to the lattice defined by the partial order in $\mathbb{J}$.

Lemma 5. *If p is strictly decreasing and ν is monotonically increasing, at each step $t > 1$ of Algorithm 2, the index $j^{(t)}$ produced by the algorithm is a minimal element of the set $I^C_{(t-1)} := \mathbb{J} \backslash I_{(t)}$. Also $j^{(1)} = 0$, and therefore $I_{(t)}$ is a down-set.*

Proof. If p is strictly decreasing and ν is monotonically increasing, then r is monotonically decreasing. By construction, $r_{j^{(t-1)}} \geqslant r_{j^{(t)}}$, so the enumeration must have $j^{(t-1)} < j^{(t)}$. It follows that $j^{(1)} = 0$.

For $t > 1$, since $j^{(t)}$ is an enumeration of $\mathbb{J}$, no element occurs twice, and so $j^{(t)} \in I^C_{(t-1)}$. Any later element $j^{(t+s)}$ in the enumeration cannot be smaller than

$j^{(t)}$, so $j^{(t)}$ is a minimal element of $I_{(t-1)}^C$. Since all elements smaller than $j^{(t)}$ occur earlier in the enumeration, we must have $\downarrow j \subset I_{(t-1)} \cup \{j^{(t)}\}$. Therefore, if $I_{(t-1)}$ is a down-set, then so is $I_{(t)}$. Since $I_{(t-1)} = \{0\}$, by induction, $I_{(t)}$ is always a down-set. □

Corollary 6. *For each $t \geqslant 1$, the set generated by Algorithm 2, $I_{(t)}$ is the solution to Optimization Problem 4 for $P = p(I_{(t)})$.*

Proof. This is an immediate consequence of Lemmas 3, 4 and 5. □

The set $M_{(t)}$ of minimal elements of $I_{(t)}^C$ is finite. One can thus find $j^{(t+1)} = i$ with largest r_i in this set. This is how Algorithm 1 finds the index $j := \operatorname{argmax}_i\{r_i \mid i \notin I \text{ and } I \cup \{i\} \text{ is a down-set}\}$, even in the case where the efficiency r is not strictly decreasing.

In the case where r is strictly decreasing, we have the following result.

Lemma 7. *If the efficiency r is strictly decreasing, then Algorithm 2 produces the same sequence of sets I as Algorithm 1.*

Proof. From the proof of Lemma 5, we have that if r is strictly decreasing, then for $t > 1$, $j^{(t)}$ as per Algorithm 2 is always the element of $M_{(t)}$ which maximizes r. This is exactly $j := \operatorname{argmax}_i\{r_i \mid i \notin I \text{ and } I \cup \{i\} \text{ is a down-set}\}$, as per Algorithm 1. For both algorithms, $I_{(1)} = \{0\}$. □

All the pieces are now in place for the main proof of this Section.

Proof of Theorem 2. From Lemma 5 we see that if the efficiency r is strictly decreasing, then each set I produced by the greedy algorithm (Algorithm 2) is a solution of Optimization Problem 3. From Lemma 7 we see that if the efficiency r is strictly decreasing, then Algorithm 2 produces the same sequence of sets I as Algorithm 1.

If p is strictly decreasing and ν is monotonically increasing, then since $r_j = p_j/\nu_j$ then r is strictly decreasing. Therefore each set I in the sequence produced by Algorithm 1 is a solution of Optimization Problem 3. □

It remains to show how to construct the set of minimal elements of $I_{(t)}^C$. To do so, we define $S(i)$, the *forward neighbourhood* of $i \in \mathbb{J}$ [18, p. 71] as

$$S(i) := \{j \in \mathbb{J} \mid i < j \text{ and } (i \leqslant \ell < j \Rightarrow \ell = i)\},$$

that is, $S(i)$ is the set of minimal elements of $\{j \in \mathbb{J} \mid i < j\}$.

Let e be the standard basis of $\mathbb{R}^{\mathbb{J}}$. To construct $M_{(t)}$, start with $M_{(1)} = S(i^{(1)}) = S(0) = \{e_1, \ldots, e_d\}$. Then given $M_{(t-1)}$ and $i^{(t)}$, one obtains $M_{(t)} = \left(M_{(t-1)} \setminus \{i^{(t)}\}\right) \cup S(i^{(t)})$.

Note that

$$\left(M_{(t-1)} \setminus \{i^{(t)}\}\right) \cup S(i^{(t)}) = \left(M_{(t-1)} \cup S(i^{(t)})\right) \setminus \{i^{(t)}\}.$$

As the minimal elements of $I_{(t)}^C$ are either elements of $M_{(t-1)}$ (but not $i^{(t)}$) or elements of $S(i^{(t)})$ we see that this set is equal to $M_{(t)}$.

5. Error bounds

We will now describe a second variant of WTP quadrature, $q^{(\mathrm{WW})}$ on $\mathbb{H}^{(r)}_{d,\gamma}$, identical to the sequence of quadrature rules $q^{(\mathrm{DA})}$ described in Section 3 above, except that the order in which the incremental rules are added to this second variant is essentially the order used by Wasilkowski and Woźniakowski [2, Section 5]. This variant uses criteria similar to those used by Wasilkowski and Woźniakowski [2, Theorem 3], but adapted to our setting. These criteria are

$$\|q_j - q_{j-1}\|_{\mathbb{H}^{(r)}_{1,\gamma}} \leqslant \sqrt{\gamma} C D^j, \quad \text{for all } j \geqslant 1. \tag{5}$$

(corresponding to Wasilkowski and Woźniakowski [2, (39)]), and

$$(j+1)\ D^{j\rho} \leqslant 1, \quad \text{for all } j \geqslant 1, \tag{6}$$

(corresponding to Wasilkowski and Woźniakowski [2, (36)]),
for some $D \in (0,1)$ and some positive C and ρ.

As a consequence of (5), we have

$$\|\Delta_j\|_{\mathbb{H}^{(r)}_{d,\gamma}} = \prod_{k=1}^{d} \left\| \delta^{(k)}_{j_k} \right\|_{\mathbb{H}^{(r)}_{1,\gamma_{d,k}}} \leqslant b(d,j), \quad \text{where}$$

$$b(d,j) := \prod_{k=1}^{d} \left(\sqrt{\gamma_{d,k}}\ C D^{j_k} \right)^{1-\delta_{0,j_k}}.$$

Let $(\xi_{d,k})$, $k = 1, \ldots, d$, be a sequence of positive numbers. In contrast to Wasilkowski and Woźniakowski [2, Section 5], we do not stipulate that $\xi_{d,k} = 1$. Define

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

We therefore have $b(d,j)/\xi(d,j) \to 0$ as $\|j\|_1 \to \infty$. We order the incremental rules in order of non-decreasing $b(d,j)/\xi(d,j)$ for each index j, creating an order on the indices $j^{(\mathrm{WW})}(h)$. We adjust $\xi(d,k)$ so that this order agrees with the lattice partial ordering of the indices. We now define $I^{(\mathrm{WW})}_N := \{j^{(\mathrm{WW})}(1), \ldots, j^{(\mathrm{WW})}(N)\}$, and define the quadrature rule

$$q^{(\mathrm{WW})}_N := \sum_{j \in I^{(\mathrm{WW})}_N} \Delta_j.$$

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

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

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

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

We can now present our version of Wasilkowski and Woźniakowski's main theorem on the error and cost of WTP quadrature [2, Theorem 3].

Theorem 8. *Let $\eta \in (0,1)$. Assume that a sequence of quadrature points $x_1, x_2, \ldots \in \mathbb{S}^2$, and a sequence of positive integers $n_0 < n_1 < \ldots$ are given such that the corresponding optimal weight quadrature rules $q_j := q_j^1 \in \mathbb{H}_{1,1}^{(r)}$ satisfy (5) and (6) for some $D \in (0,1)$ and some positive C and ρ. Then the quadrature rule $q_{\epsilon,d}^{(\mathrm{WW})}$ defined by (8) has worst-case quadrature error $e(q_{\epsilon,d}^{(\mathrm{WW})}) \leqslant \epsilon$, and its cost (in number of quadrature points) is bounded by*

$$\mathrm{cost}(q_{\epsilon,d}^{(\mathrm{WW})}) \leqslant C(d,\epsilon)\left(\frac{1}{\epsilon}\right)^{\rho/(1-\eta)},$$

where

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

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

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

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

Wasilkowski and Woźniakowski's proof, with $s := 2, \alpha := 1$, applies directly to our Theorem 8, once the change in $\xi_{d,1}$ is taken into account.

Corollary 1 of Wasilkowski and Woźniakowski [2, p. 434] presents a simpler bound for the cost of their WTP algorithm, and their simplification also applies here.

Corollary 9. *For every positive δ there exists a positive $c(d,\delta)$ such that the cost of the quadrature rule $q^{(\mathrm{WW})}_{\epsilon,d}$ defined by (8) is bounded by*

$$\mathrm{cost}(q^{(\mathrm{WW})}_{\epsilon,d}) \leqslant c(d,\delta)\left(\frac{1}{\epsilon}\right)^{\rho+\delta}.$$

For exponentially decreasing dimension weights $\gamma_{d,k}$, Theorem 4 of Wasilkowski and Woźniakowski [2] shows that the $q^{(\mathrm{WW})}$ rules are strongly polynomial.

Our sequence of rules $q^{(\mathrm{DA})}$ is more efficient than $q^{(\mathrm{WW})}$, in the sense that $q^{(\mathrm{DA})}$ is based on the optimal solution of the corresponding down-set-constrained continuous knapsack problem, as explained in Section 4. As a direct consequence of Theorem 8 and Corollary 9, we therefore have the following result.

Theorem 10. *Let $\eta \in (0,1)$. Assume that a sequence of quadrature points $x_1, x_2, \ldots \in \mathbb{S}^2$, and a sequence of positive integers $n_0 < n_1 < \ldots$ are given such that the corresponding optimal weight quadrature rules $q_j := q^1_j \in \mathbb{H}^{(r)}_{1,1}$ satisfy (5) and (6) for some $D \in (0,1)$ and some positive C and ρ. Let $I_{(t)}, q^{(\mathrm{DA})}_{(t)}$ be the index set and corresponding quadrature rule generated by iteration t of Algorithm 1, based on the rules q_j, for sufficiently small error $\epsilon = \epsilon_0$.*

Then the quadrature rule $q^{(\mathrm{DA})}_{(t)}$ has worst-case quadrature error $e(q^{(\mathrm{DA})}_{(t)}) = \epsilon_{(t)} := \sqrt{1 - p(I_{(t)})}$, and its cost $\nu(I_{(t)})$ is bounded by

$$\nu(I_{(t)}) \leqslant C(d,\epsilon_{(t)})\left(\frac{1}{\epsilon_{(t)}}\right)^{\rho/(1-\eta)},$$

where $C(d,\epsilon_{(t)})$ is defined as per Theorem 8. As a consequence, for every positive δ there exists a positive $c(d,\delta)$ such that the cost of the quadrature rule $q^{(\mathrm{DA})}_{(t)}$ is bounded by

$$\nu(I_{(t)}) \leqslant c(d,\delta)\left(\frac{1}{\epsilon_{(t)}}\right)^{\rho+\delta}.$$

6. Numerical results

With the estimates given by our analysis in hand, we now compare these to our numerical results.

Since our underlying domain $\mathcal{D}$ is $\mathbb{S}^2$ rather than $\mathbb{S}^1$, we need to change some of the details of the algorithm in comparison to the algorithm used for the torus [3]. Specifically, we need a sequence of rules on a single sphere, which yields "good enough" worst case quadrature error with optimal weights. Our choice of points for our numerical examples is a sequence of point sets, consisting of unions of spherical designs with increasing numbers of points, and non-decreasing strengths.

For the unit sphere $\mathbb{S}^2$, a spherical design [19] of strength t and cardinality m is a set of m points $X = \{x_1, \ldots, x_m\} \subset \mathbb{S}^2$ such that the equal weight quadrature rule

$$Q_X(p) := \frac{1}{m} \sum_{h=1}^{m} p(x_h)$$

is exact for all spherical polynomials p of total degree at most t.

One difference between the constructions for $\mathbb{S}^1$ and $\mathbb{S}^2$ is that the nesting of spherical designs is not efficient. The union of two spherical designs of strengths t_1 and t_2 is in general, a spherical design whose strength is the *minimum* of t_1 and t_2. In the case of our numerical examples, the first design of strength 0 is a single point. The next design of strength 1 consists of two antipodal points, so nesting is possible in this case. After this, the resulting unions of spherical designs, in general, have strength no greater than 1.

For the numerical examples, a combination of (approximate) extremal (E) and low cardinality (L) spherical designs are used, according to Table 1. These approximate spherical designs were all provided by Womersley [20, 21].

Index j	0	1	2	3	4	5	6	7	8	9	10	11
Type	L	L	E	L	E	L	E	L	E	L	E	L
Strength t	0	1	1	3	3	7	7	15	15	31	31	63
Cardinality m	1	2	4	8	16	32	64	129	256	513	1024	2049

Table 1: Strength and cardinality of approximate spherical designs used with Algorithm 1 in the numerical examples.

If we let $m_j := |X_j|$, the cardinality of X_j, and let t_j be the strength of X_j, then, for the sequence of spherical designs chosen for our numerical examples, the extremal spherical designs have $m_j = (t_j + 1)^2$, and the low cardinality spherical designs have $m_j = (t_j + 1)^2/2$ or $m_j = (t_j + 1)^2/2 + 1$, and in all cases $t_j \geqslant \sqrt{m_j} - 1$. It is not yet known a sequence of spherical designs satisfying this lower bound on strength can be extended indefinitely, but there is rigorous computational proof for t_j up to 100 [22]. Also, it is now known that an infinite sequence of spherical designs exists with the required asymptotic order of strength, that is, there is a sequence of spherical designs of cardinality $m_j = \mathcal{O}(t_j^2)$, [23] but the proof is not constructive and the corresponding implied constant is still unknown.

We now turn to estimates for rules on a single sphere, in order to use them with Theorem 8. On a single unit sphere, our sparse grid quadrature rule is an optimal weight rule $q_j = q_r^\gamma(S_j)$, based on an increasing union of spherical designs, $S_j := \bigcup_{i=0}^{j} X_i$. Its worst case error is therefore smaller than that of the optimal weight rule $q_r^\gamma(X_j)$ based on X_j, the largest spherical design contained in the union, which is, in turn no greater than the worst case error of the equal

weight rule $q_r^{\gamma,(\mathrm{QMC})}(X_j)$ based on X_j, with weights $1/|X_j|$,

$$e^2(q_j) < e^2\big(q_r^{\gamma}(X_j)\big) \leqslant e^2\big(q_r^{\gamma,(\mathrm{QMC})}(X_j)\big).$$

According to Hesse, Kuo and Sloan [5, Theorem 4], for $\gamma = 1$, we have the bound

$$e^2\big(q_r^{1,(\mathrm{QMC})}(X_j)\big) \leqslant c t_j^{-2r},$$

and therefore

$$e^2\big(q_r^{1}(X_j)\big) \leqslant c(\sqrt{m_j} - 1)^{-2r} \leqslant C_1 m_j^{-r},$$

for some $C_1 > 0$.

For general γ, as per Kuo and Sloan [4], we have

$$\begin{aligned} e^2\big(q_r^{\gamma,(\mathrm{QMC})}(X_j)\big) &= -1 + \frac{1}{m_j^2} \sum_{h=1}^{m_j} \sum_{i=1}^{m_j} K_{1,\gamma}^{(r)}(x_{j,h}, x_{j,i}) \\ &= \gamma \frac{1}{m_j^2} \sum_{h=1}^{m_j} \sum_{i=1}^{m_j} A_r(x_{j,h} \cdot x_{j,i}) \leqslant \gamma C_1 m_j^{-r}. \end{aligned} \tag{9}$$

For the sequence of spherical designs chosen for our numerical examples, we also have $2^j \leqslant m_j \leqslant 2^j + 1$. We therefore have

$$e^2(q_j) \leqslant \gamma C_2\ 2^{-rj}, \tag{10}$$

for some $C_2 > 0$.

Recall that

$$\begin{aligned} e^2(q_j) &= 1 - \|q_j\|^2_{\mathbb{H}^{(r)}_{1,\gamma}} \\ &= 1 - \|q_{j-1}\|^2_{\mathbb{H}^{(r)}_{1,\gamma}} - \|q_j - q_{j-1}\|^2_{\mathbb{H}^{(r)}_{1,\gamma}} \\ &= e^2(q_{j-1}) - \|q_j - q_{j-1}\|^2_{\mathbb{H}^{(r)}_{1,\gamma}}, \end{aligned}$$

for $j \geqslant 1$, since $q_j - q_{j-1}$ is orthogonal to q_{j-1}. Therefore

$$e^2(q_{j-1}) = e^2(q_j) + \|q_j - q_{j-1}\|^2_{\mathbb{H}^{(r)}_{1,\gamma}}.$$

Since $e^2(q_j) \geqslant 0$, using (10) we obtain $\|q_j - q_{j-1}\|^2_{\mathbb{H}^{(r)}_{1,\gamma}} \leqslant \gamma C_2\ 2^{-rj-1}$ This, in turn implies that

$$\|q_j - q_{j-1}\|^2_{\mathbb{H}^{(r)}_{1,\gamma}} \leqslant \gamma C_3\ 2^{-rj},$$

where $C_3 = 2^{-r} C_2$.

All the approximate spherical designs listed in Table 1 have one point in common, the "north pole" $(0, 0, 1)$. Therefore, in our numerical examples, the

number of points $n_j = |S_j|$ satisfies $n_j = 1+\sum_{i=0}^{j}(m_i-1) = -j+\sum_{i=0}^{j} m_i$. Since $2^i \leqslant m_i \leq 2^i + 1$, we have $2^{j+1} - j - 1 \leqslant n_j \leqslant 2^{j+1}$, and so $m_j \leqslant n_j \leqslant m_{j+1}$, for $j \geqslant 0$. Also, our numerical examples obtain a value for C_3 of approximately 1.453. In view of (9), and the preceding argument, criteria (5) and (6) hold with $D = 2^{-r/2}$, $C = C_3 \sim 1.453$ as above, and $\rho = 2/r$.

Our numerical examples use $r = 3$ and $\gamma_k = g^k$, for $g = 0.1$, 0.5, and 0.9, to see how our rules $q^{(\mathrm{DA})}$ and $q^{(\mathrm{WW})}$ behave as the decay of the dimension weights is varied. For the $q^{(\mathrm{WW})}$ rules, we use $\xi_{d,k} := CD$, with C and D defined as above.

For the DA and WW weighted tensor product algorithms, each program run uses $r = 3$; $g = 0.1$, 0.5, or 0.9; a particular dimension d, from $d = 1$ to 16; a particular maximum 1-norm for indices, typically 20; and a particular maximum number of points, up to 100 000. The numerical results are potentially affected by three 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. In this case, a least squares solution is used to obtain a best approximation to the weights. Second, if the current squared error is close to zero, and the squared norm for the current index is close to machine epsilon, then severe cancellation may occur. Third, the sequence of spherical designs used in our numerical examples is finite, so it is quite possible that our algorithm generates an index corresponding to a spherical design which is not included in our finite set. In these last two cases, 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 $(\mathbb{S}^2)^4$, $r = 3$, $\gamma_{4,k} = 0.5^k$. The number of points used varies from $n = 1$ to 100 000. The cost axis is horizontal and the error axis is vertical, to match the figures shown in the torus paper [3]. The curve in Figure 1 labelled "WW bound" is actually the minimum of the bounds given by Theorem 8, as the parameter η is varied over a finite number of values between 0 and 1.

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 8. The WW cost bound itself has an asymptotic rate of convergence of $\mathcal{O}(\epsilon^{-\rho}) = \mathcal{O}(\epsilon^{-2/r}) = \mathcal{O}(\epsilon^{-2/3})$ for all of our cases. In other words, the asymptotic bound has quadrature error of order $\mathcal{O}(n^{-3/2})$. Judging from Figure 1, the rates of convergence of both algorithms appear consistent with that of the bound, but the asymptotic rate is not achieved by either algorithm or by the bound itself, for the number of quadrature points displayed in the plot.

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 curve for $d = 1$ appears consistent with the asymptotic error rate $\epsilon = \mathcal{O}(N^{-3/2})$. The cases $d = 8$ and $d = 16$ are almost indistinguishable on this Figure. This is an example of the convergence in dimension.

Figure 3 shows the equivalent results for the DA quadrature rules for $\gamma_{d,k} = 0.9^k$. The curve for $d = 1$ again appears consistent with the asymptotic error rate $\epsilon = \mathcal{O}(N^{-3/2})$, but as d increases to 16, the initial rate of convergence to

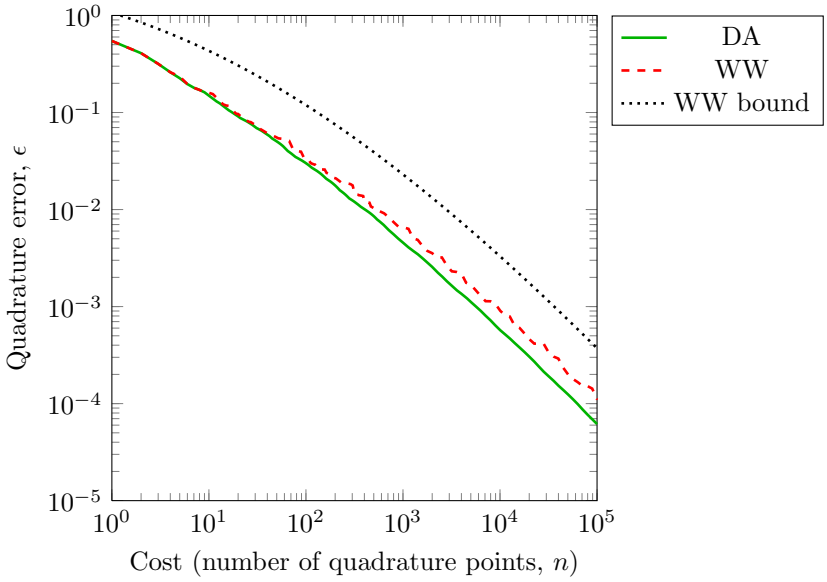

Figure 1: Error of DA and WW rules vs WW bound for $(\mathbb{S}^2)^4$, $r = 3$, $\gamma_{4,k} = 0.5^k$.

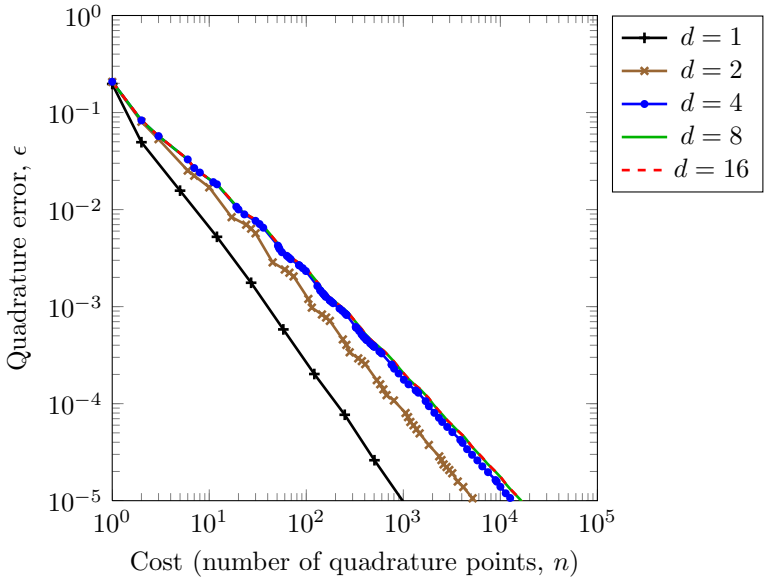

Figure 2: Error of DA rules for $(\mathbb{S}^2)^d$, $d = 1, 2, 4, 8, 16$; $r = 3$, $\gamma_{d,k} = 0.1^k$.

zero of the error becomes much slower than that for $\gamma_{d,k} = 0.1^k$. This behaviour is expected, given the WW bound.

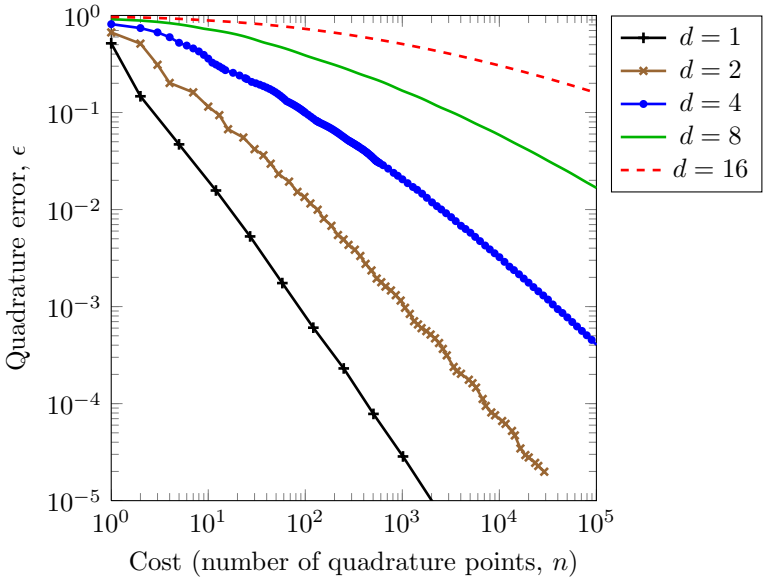

Figure 3: Error of DA rules for $(\mathbb{S}^2)^d$, $d = 1, 2, 4, 8, 16$; $r = 3$, $\gamma_{d,k} = 0.9^k$.

Acknowledgements

Thanks to Rob Womersley for the spherical designs, and to Gary Froyland for discussions on precedence-constrained knapsack problems. The hospitality of the Hausdorff Research Institute for Mathematics (HIM) in Bonn is much appreciated. The support of the Australian Research Council under its Centre of Excellence program is gratefully acknowledged.

References

[1] M. Hegland, Adaptive sparse grids, ANZIAM Journal 44 (E) (2003) C335–C353.

[2] G. W. Wasilkowski, H. Woźniakowski, Weighted tensor product algorithms for linear multivariate problems, Journal of Complexity 15 (1999) 402–447.

[3] M. Hegland, P. C. Leopardi, The rate of convergence of sparse grid quadrature on the torus, in: W. McLean, A. J. Roberts (Eds.), Proceedings of the 15th Biennial Computational Techniques and Applications Conference, CTAC-2010, volume 52 of *ANZIAM J.*, pp. C500–C517.

[4] F. Y. Kuo, I. H. Sloan, Quasi-Monte Carlo methods can be efficient for integration over products of spheres, Journal of Complexity 21 (2005) 196–210.

[5] K. Hesse, F. Y. Kuo, I. H. Sloan, A component-by-component approach to efficient numerical integration over products of spheres, Journal of Complexity 23 (2007) 25–51.

[6] F. J. Hickernell, H. Woźniakowski, Tractability of multivariate integration for periodic functions, Journal of Complexity 17 (2001) 660–682.

[7] I. H. Sloan, H. Woźniakowski, Tractability of multivariate integration for weighted Korobov classes, Journal of Complexity 17 (2001) 697–721.

[8] S. Zakowicz, Z. Harman, N. Grün, W. Scheid, Angular distribution of hypersatellite and satellite radiation emitted after resonant transfer and excitation into u^{91+} ions, Physical Review A 68 (2003) 042711.

[9] H. Sato, Is the single scattering model invalid for the coda excitation at long lapse time?, Pure And Applied Geophysics 128 (1988) 43–47.

[10] K. Altmann, Space-time distortion of laser pulses due to multiple scattering in particulate media, Applied Optics 27 (1988) 2451–2460.

[11] A. V. Starkov, M. Noormohammadian, U. G. Oppel, A stochastic model and a variance-reduction Monte-Carlo method for the calculation of light transport, Applied Physics B: Lasers and Optics 60 (1995) 335–340.

[12] B. Kaplan, G. Ledanois, B. Drévillon, Mueller matrix of dense polystyrene latex sphere suspensions: Measurements and monte carlo simulation, Applied Optics 40 (2001) 2769–2777.

[13] G. H. Vineyard, Multiple scattering of neutrons, Physical Review 96 (1954) 93–98.

[14] B. Davey, H. A. Priestley, Introduction to Lattices and Order, Cambridge University Press, Cambridge, 1990.

[15] M. Griebel, S. Knapek, Optimized general sparse grid approximation spaces for operator equations, Mathematics of Computations 78 (2009) 2223–2257. Also available as SFB611 preprint No 402.

[16] G. B. Dantzig, Discrete-variable extremum problems, Operations Research 5 (1957) 266–277.

[17] S. Martello, P. Toth, Knapsack problems, Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons Ltd., Chichester, 1990.

[18] T. Gerstner, M. Griebel, Dimension-adaptive tensor product quadrature, Computing 71 (2003) 65–87.

[19] P. Delsarte, J. M. Goethals, J. J. Seidel, Spherical codes and designs, Geometriae Dedicata 6 (1977) 363–388.

[20] X. Chen, R. Womersley, Existence of solutions to underdetermined equations and spherical designs, SIAM Journal of Numerical Analysis 44 (2006) 2326–2341.

[21] R. S. Womersley, Spherical designs with close to the minimal number of points, Applied Reports AMR09/26, The University of New South Wales, 2009. Preprint submitted to Elsevier.

[22] X. Chen, A. Frommer, B. Lang, Computational existence proofs for spherical t-designs, Numerische Mathematik 117 (2011) 289–305. 10.1007/s00211-010-0332-5.

[23] A. Bondarenko, D. Radchenko, M. Viazovska, Optimal asymptotic bounds for spherical designs, 2011. ArXiv:1009.4407v3 [math.MG].