Sparse grid quadrature on products of spheres

Paul Leopardi

Mathematical Sciences Institute, Australian National University.
For presentation at University of Canterbury 2012.
Joint work with Markus Hegland, ANU.

8 June 2012

AUSTRALIAN RESEARCH COUNCIL
Centre of Excellence for Mathematics and Statistics of Complex Systems

ARC

ANU
THE AUSTRALIAN NATIONAL UNIVERSITY

A simple question

How efficiently can a quadrature rule approximate an integral, especially in high dimensions?

The main result

For quadrature in the setting of weighted tensor product spaces on the unit sphere, there is a choice of points on each sphere, such that a quadrature rule constructed using a greedy dimension adaptive sparse grid algorithm breaks the curse of dimensionality.

The asymptotic rate of convergence to zero of the worst case quadrature error on a product of spheres has essentially the same order as that for a single sphere.

Outline of the talk

- Applications
- Some key concepts
- The setting
- The algorithm
- Optimality of the algorithm
- Numerical results

Applications of integration on products of spheres include quantum mechanics, and transport and multiple scattering problems in various areas, including acoustics, optical scattering, and neutron transport.

One prototypical problem is scattering by a sequence of objects, with some decay after each object.

This can be modelled using a multiple integral of a function on a product of spheres.

(Zakowicz at al. 2003; Sato 1988; Altmann 1988; Kaplan et al. 2001; Vineyard 1954)

Some key questions

- What is quadrature? How can some abstract concepts help us to understand quadrature better?
- What is the curse of dimensionality? Why does this make multidimensional quadrature difficult?
- What is a grid? What is the advantage of nesting of points? Why is there no efficient nesting of points on a sphere? What are "equally-spaced" points on a sphere?

An abstract look at quadrature

We start with some normed linear space $\mathcal{F}$ of real valued functions defined on some domain $\mathcal{D}$.

To make things simpler and more concrete, suppose $\mathcal{D}$ is a compact subset of some $\mathbb{R}^{(s+1)}$, with a probability measure μ, meaning $\mu(\mathcal{D}) = 1$.

The integral $\mathbf{Int}$ is a linear functional on $\mathcal{F}$, that is, an element of the dual space $\mathcal{F}^*$.

For $f \in \mathcal{F}$,

$$\langle \mathbf{Int}, f \rangle := \int_{\mathcal{D}} f(x)\, d\mu(x).$$

An abstract look at quadrature (cont.)

In this framework, a quadrature rule is a finite linear combination of point evaluation functionals, which approximates the integral $\mathbf{Int}$.

For $f \in \mathcal{F}$,

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

For this to make sense, each point evaluation functional must be well defined. This constrains our choice of the space $\mathcal{F}$.

The theory turns out to be simpler if we choose $\mathcal{F}$ to be a reproducing kernel Hilbert space (RKHS) of functions on $\mathcal{D}$.

An RKHS on $\mathcal{D}$

$\mathcal{F}$ is a reproducing kernel Hilbert space of functions $f : \mathcal{D} \to \mathbb{R}$ with inner product $\langle \cdot, \cdot \rangle_{\mathcal{F}}$ if $\mathcal{F}$ is a Hilbert space with this inner product *and* there is some kernel $K : \mathcal{D} \times \mathcal{D} \to \mathbb{R}$ such that for all $x \in \mathcal{D}$, the function k_x defined by $k_x(y) := K(x, y)$ satisfies

$$k_x \in \mathcal{F}, \quad \text{and, for all } f \in \mathcal{F}, \quad \langle k_x, f \rangle_{\mathcal{F}} = f(x).$$

Thus each k_x is the representer in the sense of Riesz of the point evaluation functional $\langle K_x, f \rangle := f(x)$, for every point of $\mathcal{F}$, and each of these functionals is bounded and therefore continuous.

The curse of dimensionality

There are two obstacles to efficient quadrature on high-dimensional products of domains $\mathcal{D}^d$.

The first is the curse of dimensionality. If we put just two points on each copy of the domain $\mathcal{D}$, then we need 2^d points: the number of points, and this the number of function evaluations needed for quadrature is exponential in dimension d.

In contrast, a Monte Carlo approach to quadrature distributes points "randomly" across $\mathcal{D}^d$, and the number of points is independent of the dimension d.

What is a grid?

The second obstacle involves the nesting of points on a sphere. This affects our ideas about how to create grids.

The usual concept of a grid of points in two dimensions starts with m points on (e.g.) the unit interval $\mathcal{D} = [0, 1]$, and n points on a second copy of the interval $\mathcal{D}$, and then forms a set of mn points on the unit square $\mathcal{D}^2$, called a grid.

Often the points are equally spaced on each of the two intervals, but this is not necessary to the definition.

For the purposes of this talk, we generalize this concept of a grid, by generalizing the domain $\mathcal{D}$, and by considering $\mathcal{D}^d$.

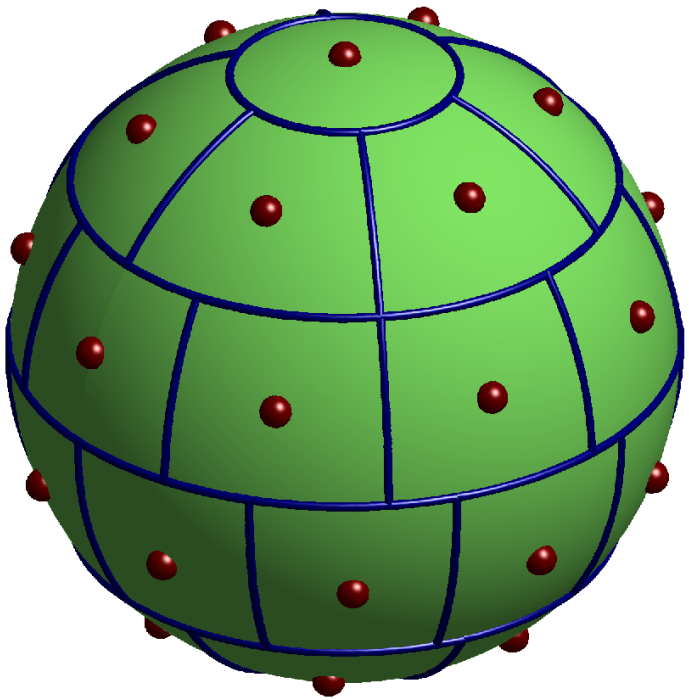
Nesting points on the unit sphere

For a unit circle $\mathbb{T} := \mathbb{S}^1 \in \mathbb{R}^2$, you can place m points at angles of $2\pi/m$ to obtain perfectly equally spaced points. Just add another m points half way in between and you have $2m$ equally spaced points.

This nesting makes it easy to form well-behaved nested grids on the torus $\mathbb{T}^d$.

This does not happen for the unit sphere $\mathbb{S}^2 \in \mathbb{R}^3$. The closest you can get to equally spaced, well distributed points on the unit sphere is to use the vertices of the Platonic solids, and there are only five of those. Even then, these sets of vertices do not nest. We will see a solution to this problem later in the talk.

33 points on a sphere



(L, 2007)

Weighted tensor product spaces

The idea of weighted tensor product spaces is the dimensions of $\mathcal{D}^d$ may differ in importance.

Each dimension is given a weight, and depending on how the weights decay, a quadrature problem which is intractable with equal weights can become tractable.

Tractability means the minimal number of function evaluations is polynomial in $1/\epsilon$, where ϵ is the worst case quadrature error.

(Hickernell and Woźniakowski 2001; Sloan and Woźniakowski 2001; Kuo and Sloan, 2005)

An RKHS on $\mathcal{D}$ of functions with mean zero

Let $\mathcal{D} \subset \mathbb{R}^{s+1}$ be a compact s-dimensional manifold with probability measure μ, and let $\mathcal{H}$ be a reproducing kernel Hilbert space (RKHS) of functions $f : \mathcal{D} \to \mathbb{R}$, such that

$$\int_{\mathcal{D}} f(x)\, d\mu(x) = 0 \text{ for all } f \in \mathcal{H},$$

with kernel $\mathcal{K} : \mathcal{D} \times \mathcal{D} \to \mathbb{R}$ such that for all $x \in \mathcal{D}$, the function k_x defined by $k_x(y) := \mathcal{K}(x, y)$ satisfies

$$k_x \in \mathcal{H}, \quad \text{and, for all } f \in \mathcal{H}, \quad \langle k_x, f \rangle_{\mathcal{H}} = f(x).$$

(Hickernell and Woźniakowski 2001; Sloan and Woźniakowski 2001; Kuo and Sloan, 2005)

The weighted space $\mathcal{H}^\gamma$

For $0 < \gamma \leqslant 1$, extend $\mathcal{H}$ into the space $\mathcal{H}^\gamma$ of all functions of the form $a1 + f$, where $1(x) := 1$, $a \in \mathbb{R}$, and $f \in \mathcal{H}$, with inner product

$$\langle a1 + f, b1 + g \rangle_{\mathcal{H}^\gamma} := ab + \frac{1}{\gamma} \langle f, g \rangle_{\mathcal{H}}.$$

$\mathcal{H}^\gamma$ 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 $\mathcal{H}$.

(Hickernell and Woźniakowski 2001; Sloan and Woźniakowski 2001; Kuo and Sloan, 2005)

The weighted space $\mathcal{H}^\gamma$ (cont.)

Since μ is a probability measure, we have $\langle \mathbf{Int}, 1 \rangle = 1$, so

$$\langle \mathbf{Int}, b1 + g \rangle = b = \langle 1, b1 + g \rangle_{\mathcal{H}^\gamma},$$

so the function 1 represents the functional $\mathbf{Int}$ in the space $\mathcal{H}^\gamma$.

If we could evaluate the inner product exactly on all elements of the space, then we could evaluate the integral. But in our case, we assume that we can only evaluate functions from $\mathcal{H}^\gamma$ at single points of the domain $\mathcal{D}$.

The weighted tensor product space $\mathcal{H}^{d,\gamma}$

Let $\gamma := (\gamma_1, \ldots, \gamma_d)$, with $1 \geqslant \gamma_1 \geqslant \ldots \geqslant \gamma_d > 0$.

On $\mathcal{D}^d$ define the tensor product RKHS

$$\mathcal{H}^{d,\gamma} := \bigotimes_{h=1}^{d} \mathcal{H}^{\gamma_h}.$$

The reproducing kernel of $\mathcal{H}^{d,\gamma}$ is

$$\mathcal{K}^{d,\gamma}(x, y) := \prod_{h=1}^{d} \mathcal{K}^{\gamma_h}(x_h, y_h).$$

(Hickernell and Woźniakowski 2001; Sloan and Woźniakowski 2001; Kuo and Sloan, 2005)

Quadrature rules on $\mathcal{H}^{d,\gamma}$

For $\{x_1, \ldots, x_n\} \subset \mathcal{D}^d$, the quadrature rule

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

is a continuous linear functional on $\mathcal{H}^{d,\gamma}$, satisfying

$$\langle Q, f \rangle = \langle q, f \rangle_{\mathcal{H}^{d,\gamma}},$$

where

$$q := \sum_{i=1}^{n} w_i k_{x_i}^{d,\gamma}, \quad k_{x_i}^{d,\gamma}(y) := \mathcal{K}^{d,\gamma}(x_i, y).$$

(Wasilkowski and Woźniakowski 1999; Hickernell and Woźniakowski 2001)

Optimal quadrature weights on $\mathcal{H}^{d,\gamma}$

The worst case error

$$e(q) := \sup_{\|f\|_{\mathcal{H}^{d,\gamma}} \leqslant 1} \left| \int_{\mathcal{D}^d} f(x) d\mu_d(x) - \langle q, f \rangle_{\mathcal{H}^{d,\gamma}} \right|$$

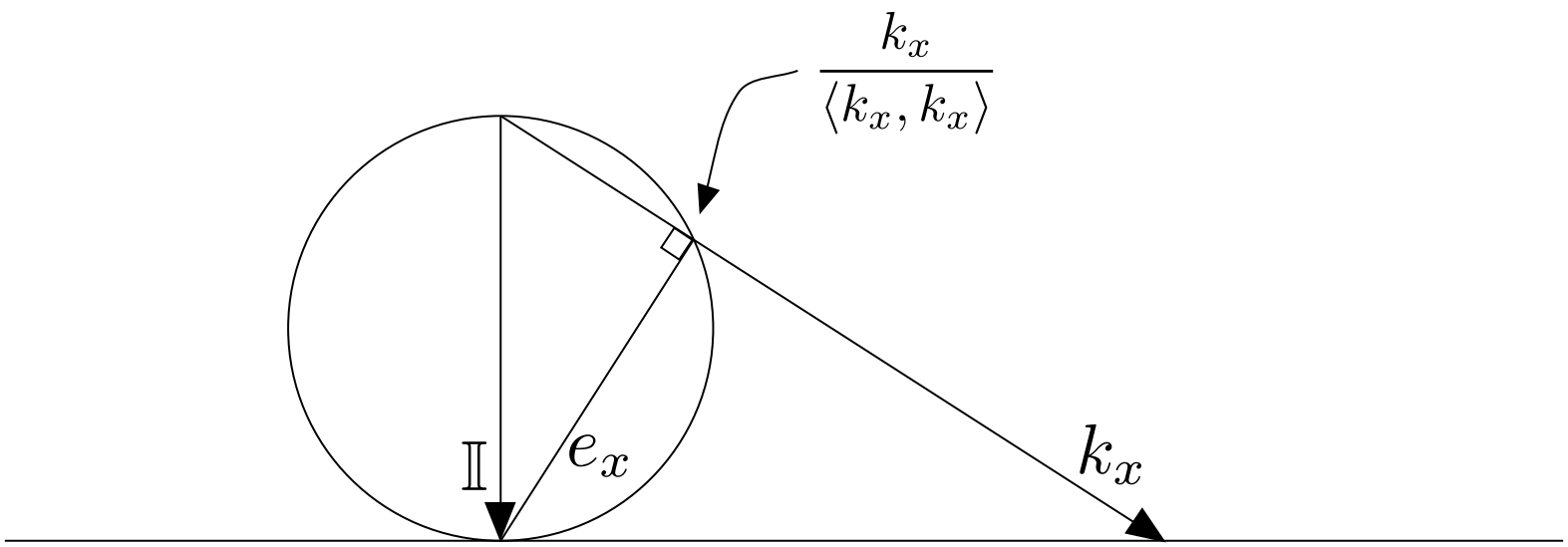
satisfies

$$\begin{aligned} e(q)^2 = \|1 - q\|^2_{\mathcal{H}^{d,\gamma}} &= \langle 1 - q, 1 - q \rangle_{\mathcal{H}^{d,\gamma}} \\ &= 1 - 2 \sum_{i=1}^{n} w_i + w^T G w, \quad \text{where} \\ G_{i,j} := \langle k^{d,\gamma}_{x_i}, k^{d,\gamma}_{x_j} \rangle_{\mathcal{H}^{d,\gamma}} &= \mathcal{K}^{d,\gamma}(x_i, x_j). \end{aligned}$$

The weights w are optimal when $Gw = [1, \ldots, 1]^T$.

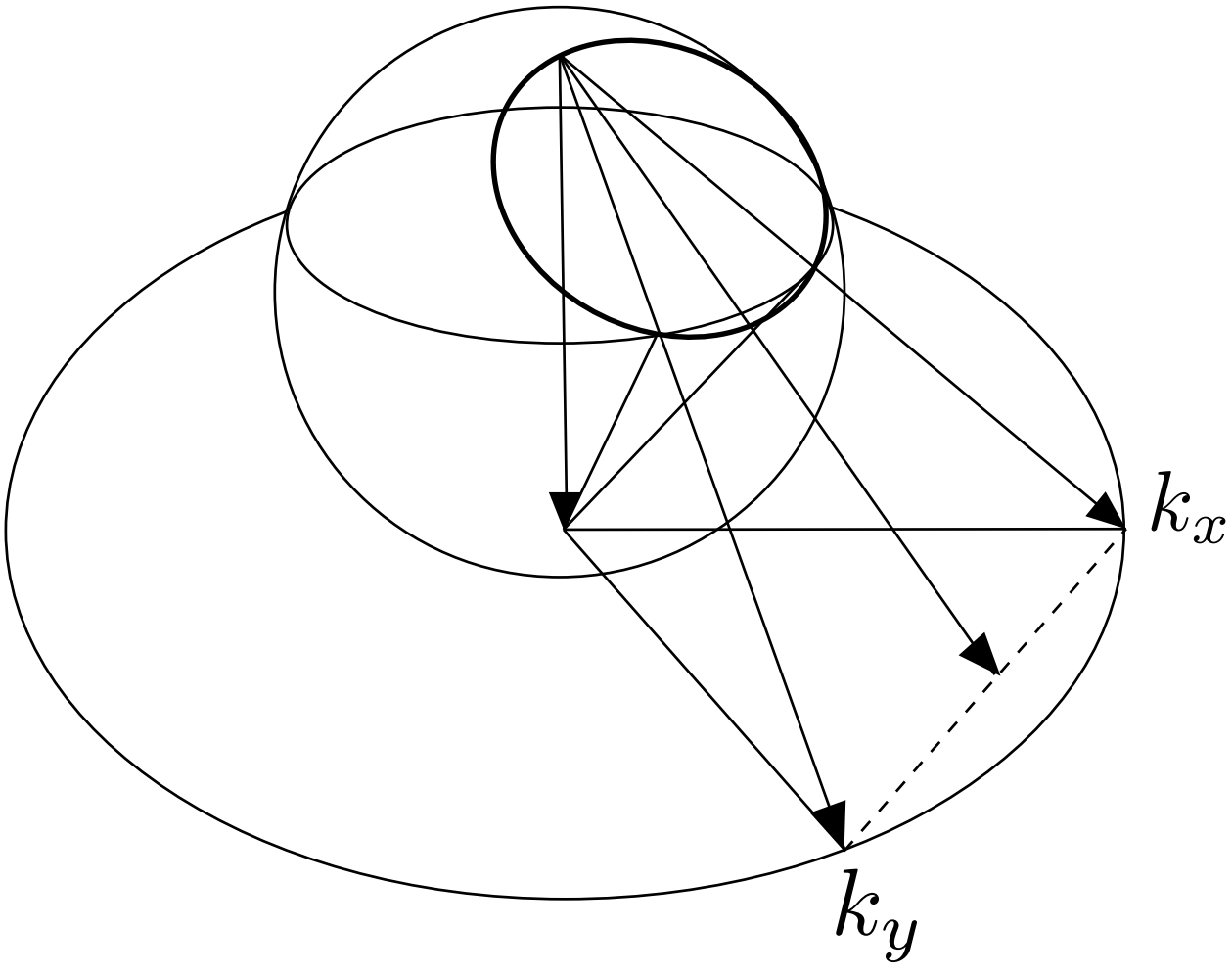
(Wasilkowski and Woźniakowski 1999)

Optimal weight for one quadrature point



(Illustration by Osborn, 2009)

Optimal weights for two quadrature points



(Illustration by Osborn, 2009)

Dimension adaptive sparse grid quadrature

Given a sequence of point sets on a single domain $\mathcal{D}$, each of which defines an optimal quadrature rule, how can these point sets be used on each copy of $\mathcal{D}$ to yield an efficient quadrature rule on $\mathcal{D}^d$?

The concept of efficiency here is a trade-off between worst case quadrature error, and cost in terms of number of function evaluations.

Dimension adaptive sparse sparse grid quadrature answers this question by combining the differences between optimal quadrature rules in a particular way.

Differences of optimal quadrature rules in $\mathcal{H}^\gamma$

Consider a sequence of quadrature points $x_1, x_2, \ldots \in \mathcal{D}$, and a sequence of positive integers $m_0 < m_1 < \ldots$

For $j \geqslant 0$, let q_j^γ denote the optimal quadrature rule in

$$V_j^\gamma := \mathbf{span}\{k_{x_1}^\gamma, \ldots, k_{x_{m_j}}^\gamma\} \subset \mathcal{H}^\gamma.$$

Define the pair-wise orthogonal spaces U_j^γ by $U_0^\gamma = V_0^\gamma$, by the orthogonal decomposition $V_{j+1}^\gamma = V_j^\gamma \oplus U_{j+1}^\gamma$.

Since the q_j^γ are optimal,

$$\delta_{j+1}^\gamma := q_{j+1}^\gamma - q_j^\gamma \in U_{j+1}^\gamma, \quad \text{and}$$
$$\delta_0^\gamma := q_0^\gamma \in U_0^\gamma = V_0^\gamma.$$

(Gerstner and Griebel 1998; Wasilkowski and Woźniakowski 1999)

Multi-indices and down-sets

Elements of $\mathbb{J} := \mathbb{N}^d$ are treated as multi-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 h from 1 to d.

For a multi-index $i \in \mathbb{J}$, let $\downarrow i$ denote the down-set of i, defined by $\downarrow i := \{j \in \mathbb{J} \mid j \leqslant i\}$.

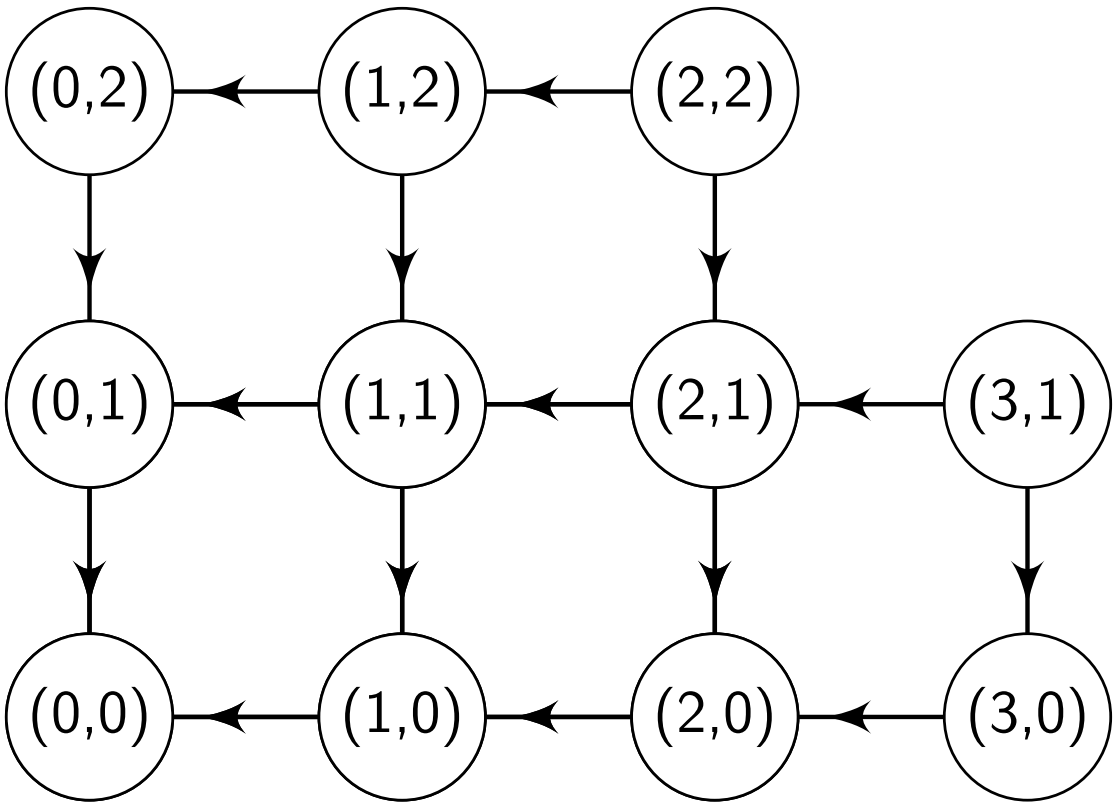
Subsets of $\mathbb{J}$ are partially ordered by set inclusion.
For a subset $I \subset \mathbb{J}$, let $\downarrow I$ denote the down-set of I, defined by $\downarrow I := \bigcup_{i \in I} \downarrow i$.

Then $\downarrow I$ is the smallest set $Y \supseteq I$ such that if $i \in Y$ and $j \leqslant i$ then $j \in Y$. Thus $\downarrow\downarrow I = \downarrow I$.

(Davey and Priestley 1990)

Example of a down-set

The down-set of the set $\{(2,2),(3,1)\}$:



(Davey and Priestley 1990, Matthes 2011 (LaTeX macros))

Sparse grid quadrature in $\mathcal{H}^{d,\gamma}$

A sparse grid quadrature rule in $\mathcal{H}^{d,\gamma}$ is of the form

$$q \in V_I := \sum_{j \in I} \bigotimes_{h=1}^{d} V_{j_h}^{\gamma_h}$$

for some index set $I \subset \mathbb{J} = \mathbb{N}^d$.

The orthogonal decomposition $V_j^\gamma = \bigoplus_{i=1}^{j} U_i^\gamma$
yields the multidimensional orthogonal decomposition

$$V_I = \bigoplus_{j \in \downarrow I} \bigotimes_{h=1}^{d} U_{j_h}^{\gamma_h},$$

(Gerstner and Griebel 1998; Wasilkowski and Woźniakowski 1999; Hegland 2003)

Sparse grid quadrature in $\mathcal{H}^{d,\gamma}$ (cont.)

An optimal $q \in V_I$ is

$$q_I = \sum_{j \in \downarrow I} \bigotimes_{h=1}^{d} \delta_{j_h}^{\gamma_h}.$$

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

(Gerstner and Griebel 1998; Hegland 2003)

Our optimization problem

The optimization problem uses the following definitions.

Definition 1

For index $j \in \mathbb{J}$, define

$$\nu_{j_k}^{(k)} := \dim U_{j_k}^{\gamma_{d,k}}, \quad \nu_j := \prod_{k=1}^{d} \nu_{j_k}^{(k)},$$

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

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

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

Our optimization problem (cont.)

Our optimization problem is to

$$\begin{aligned} \text{Minimize} \quad & \boldsymbol{\nu(I) := \sum_{i \in I} \nu_i} \\ \text{subject to} \quad & \boldsymbol{I = {\downarrow} I, \quad p(I) := \sum_{i \in I} p_i \geqslant P,} \end{aligned} \tag{1}$$

for $\boldsymbol{P = 1 - \epsilon^2}$, where $\boldsymbol{p_i := \|\Delta_i\|^2} \in \mathbb{R}_+$ and $\boldsymbol{\nu_i} \in \mathbb{N}_+$.

This is a down-set constrained binary knapsack problem.

(Gerstner and Griebel 1998; Hegland 2003; Gerstner and Griebel 2003)

A dimension adaptive algorithm to choose I

Algorithm 1: A greedy dimension adaptive sparse grid algorithm.

Data: accuracy ϵ, incremental rules Δ_j and costs ν_j for $j \in \mathbb{J}$

Result: ϵ approximation q and index set I

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

while $\|1 - q\| > \epsilon$ **do**
  $i := \mathbf{argmax}_j \{\|\Delta_j\|^2 / \nu_j \mid I \cup \{j\}$ is a down-set$\}$;
  $I := I \cup \{i\}$; $q := q + \Delta_i$;

(Hegland 2003; Gerstner and Griebel 2003)

The related classical knapsack problem

A widely studied problem in optimisation is the knapsack problem. The knapsack problem related to our problem (1) is to

$$\begin{aligned} \text{Minimize} \quad & \boldsymbol{\nu(I) := \sum_{i \in I} \nu_i} \\ \text{subject to} \quad & \boldsymbol{p(I) := \sum_{i \in I} p_i \geqslant P,} \end{aligned} \tag{2}$$

for $\boldsymbol{P = 1 - \epsilon^2}$, where $\boldsymbol{p_i := \|\Delta_i\|^2} \in \mathbb{R}_+$ and $\boldsymbol{\nu_i} \in \mathbb{N}_+$.

This is problem (1) without the down-set constraint.

(Dantzig 1957)

Adding the lattice-constraint back in

The following proposition holds.

Proposition 1

If $\boldsymbol{I}$ *is a solution of the knapsack problem* (2)*, and satisfies the admissibility condition* $\boldsymbol{I} = \downarrow \boldsymbol{I}$*, then it is a solution of the optimization problem* (1)*.*

This justifies our calling problem (1) a lattice-constrained knapsack problem.

Monotonicity

The sequence $\boldsymbol{p} \in \mathbb{R}_+^{\mathbb{J}}$ is monotonically decreasing if $\boldsymbol{i} < \boldsymbol{j}$ implies that $\boldsymbol{p_i} \geqslant \boldsymbol{p_j}$.

If $\boldsymbol{i} < \boldsymbol{j}$ implies that $\boldsymbol{p_i} > \boldsymbol{p_j}$,
then $\boldsymbol{p} \in \mathbb{R}_+^{\mathbb{J}}$ is strictly decreasing.

The definitions of "monotonically increasing" and "strictly increasing" are similar.

Algorithm 1 is optimal, given monotonicity

The following theorem holds.

Theorem 2

If $\boldsymbol{p} \in \mathbb{R}_+^{\mathbb{J}}$ *is strictly decreasing and* $\boldsymbol{\nu} \in \mathbb{N}_+^{\mathbb{J}}$ *is monotonically increasing, each set* $\boldsymbol{I}$ *given by Algorithm 1 is a solution of the optimization problem* (1) *for* $\boldsymbol{P} = \boldsymbol{p}(\boldsymbol{I})$.

The proof involves formulating the corresponding greedy algorithm for the knapsack problem (2), then showing that if monotonicity holds, this algorithm generates the same sequence of sets $\boldsymbol{I}$ as Algorithm 1. It is known since Dantzig (1957) that the greedy algorithm for the knapsack problem is optimal when $\boldsymbol{P} = \boldsymbol{p}(\boldsymbol{I})$.

(Dantzig 1957)

The "rounded" classical knapsack problem

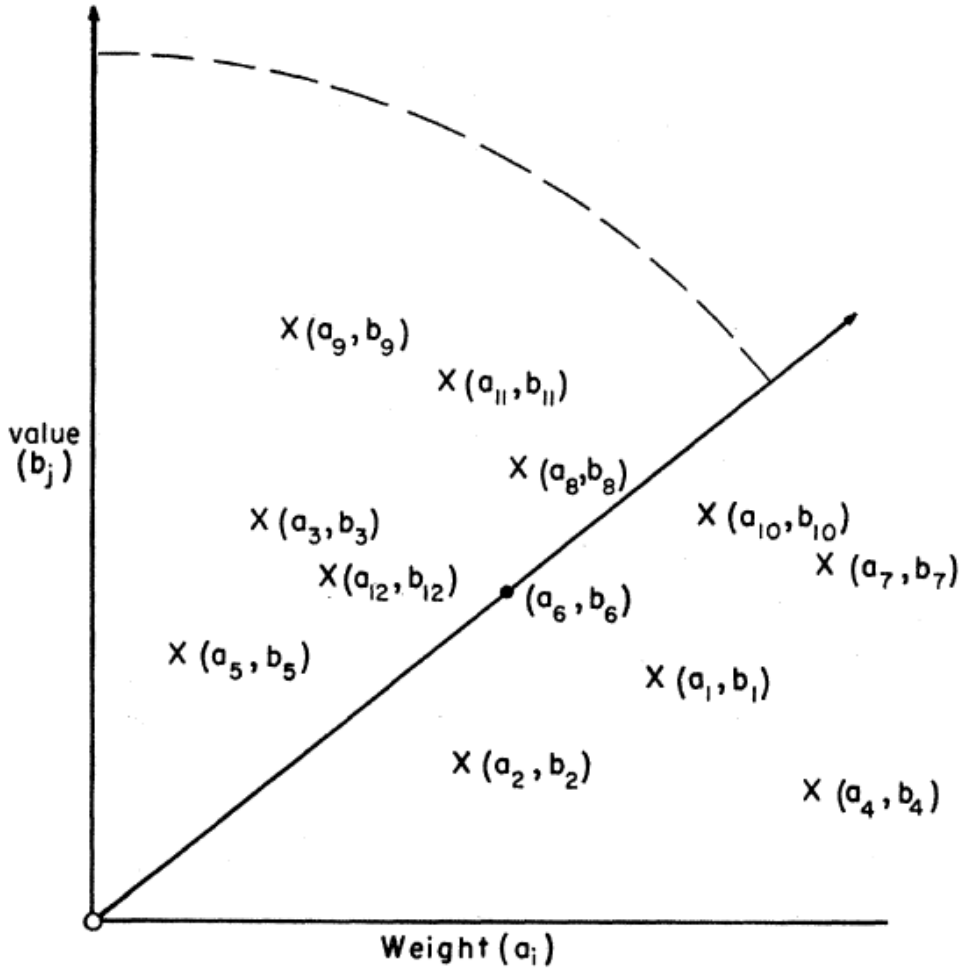


Fig. 3. Graphical solution of the knapsack problem.

(Dantzig 1957)

The work of Wasilkowski and Woźniakowski

Wasilkowski and Woźniakowski (1999) define a Weighted Tensor Product algorithm similar to Algorithm 1 and prove a relationship between cost and error bounds, given the following conditions on the difference rules δ_j^γ

$$\left\| \delta_j^\gamma \right\|_{\mathbb{H}_{1,\gamma}^{(r)}} \leqslant \sqrt{\gamma} C D^j, \quad \text{for all } j \geqslant 1. \tag{3}$$

$$(j+1)\, D^{j\rho} \leqslant 1, \quad \text{for all } j \geqslant 1, \tag{4}$$

for some $0 < D < 1$ and some positive C and ρ.

(Wasilkowski and Woźniakowski 1999)

The work of Wasilkowski and Woźniakowski

Denote the sequence of rules generated by the algorithm of Wasilkowski and Woźniakowski by $q_{\epsilon,d}^{(\mathrm{WW})}$. The following theorem holds as a corollary to a more complicated bound corresponding to Theorem 21 of Wasilkowski and Woźniakowski (1999).

Theorem 3

For every positive δ there exists a positive $c(d,\delta)$ such that the cost of the quadrature rule $q_{\epsilon,d}^{(\mathrm{WW})}$ is bounded by

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

Since the sequence of rules $q_{\epsilon,d}^{(\mathrm{DA})}$ generated by Algorithm 1 is optimal, these rules satisfy the same bound.

Specifics for $\mathbb{S}^2$ – spherical designs

We use a sequence of rules on a single sphere $\mathbb{S}^2$, which yields "good enough" worst case quadrature error with optimal weights. We choose a sequence of unions of spherical designs with increasing numbers of points, and non-decreasing strengths.

For the unit sphere $\mathbb{S}^2$, a spherical design 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

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

is exact for all spherical polynomials p of total degree at most t. Spherical designs do not nest efficiently. For the numerical examples, the unions of spherical designs have strength at most 1.

(Delsarte, Goethals and Seidel 1977)

Spherical designs used in numerical example

For the realization of the dimension adaptive algorithm, a combination of (approximate) extremal (E) and low cardinality (L) spherical designs are used. These were all provided by Womersley.

Index j	$\mathbf{0}$	$\mathbf{1}$	$\mathbf{2}$	$\mathbf{3}$	$\mathbf{4}$	$\mathbf{5}$	$\mathbf{6}$	$\mathbf{7}$	$\mathbf{8}$	$\mathbf{9}$	$\mathbf{10}$	$\mathbf{11}$
Type	$\boldsymbol{L}$	$\boldsymbol{L}$	$\boldsymbol{E}$	$\boldsymbol{L}$	$\boldsymbol{E}$	$\boldsymbol{L}$	$\boldsymbol{E}$	$\boldsymbol{L}$	$\boldsymbol{E}$	$\boldsymbol{L}$	$\boldsymbol{E}$	$\boldsymbol{L}$
Strength t	$\mathbf{0}$	$\mathbf{1}$	$\mathbf{1}$	$\mathbf{3}$	$\mathbf{3}$	$\mathbf{7}$	$\mathbf{7}$	$\mathbf{15}$	$\mathbf{15}$	$\mathbf{31}$	$\mathbf{31}$	$\mathbf{63}$
Card. m	$\mathbf{1}$	$\mathbf{2}$	$\mathbf{4}$	$\mathbf{8}$	$\mathbf{16}$	$\mathbf{32}$	$\mathbf{64}$	$\mathbf{129}$	$\mathbf{256}$	$\mathbf{513}$	$\mathbf{1024}$	$\mathbf{2049}$

For the successive unions of these designs criteria (3) and (4) hold with $\boldsymbol{D = 2^{-r/2}}$, $\boldsymbol{C \sim 1.453}$ as above, and $\boldsymbol{\rho = 2/r}$.

It is not known if the L or E designs form an infinite sequence, but it has recently been proven that a sequence exists with $\boldsymbol{m \leqslant Ct^2}$.

(Chen and Womersley 2006; Womersley 2009; Bondarenko et al. 2011)

Monotonicity holds for this choice of points

For monotonicity to hold, it need only hold for the sequence of unions of designs on a single sphere.

As each design is added to an existing union, the number of extra points does not decrease.

Also, upper and lower bounds on the worst case error show that the norm of each successive difference rule decreases monotonically.

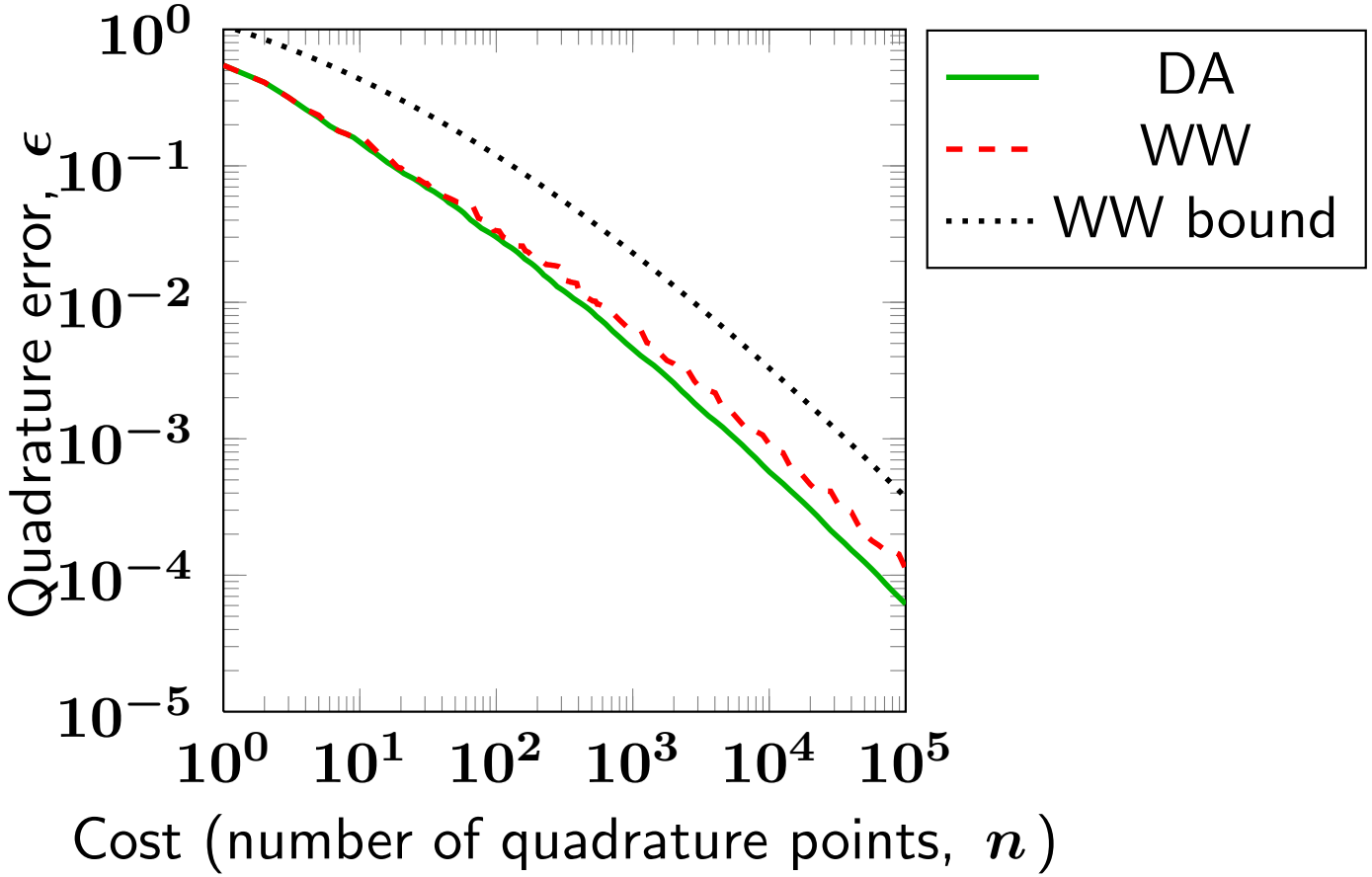
Parameters used in numerical examples

Our numerical examples use $r = 3$ and $\gamma_k = g^k$, for $g = 0.1$, 0.5, and 0.9.

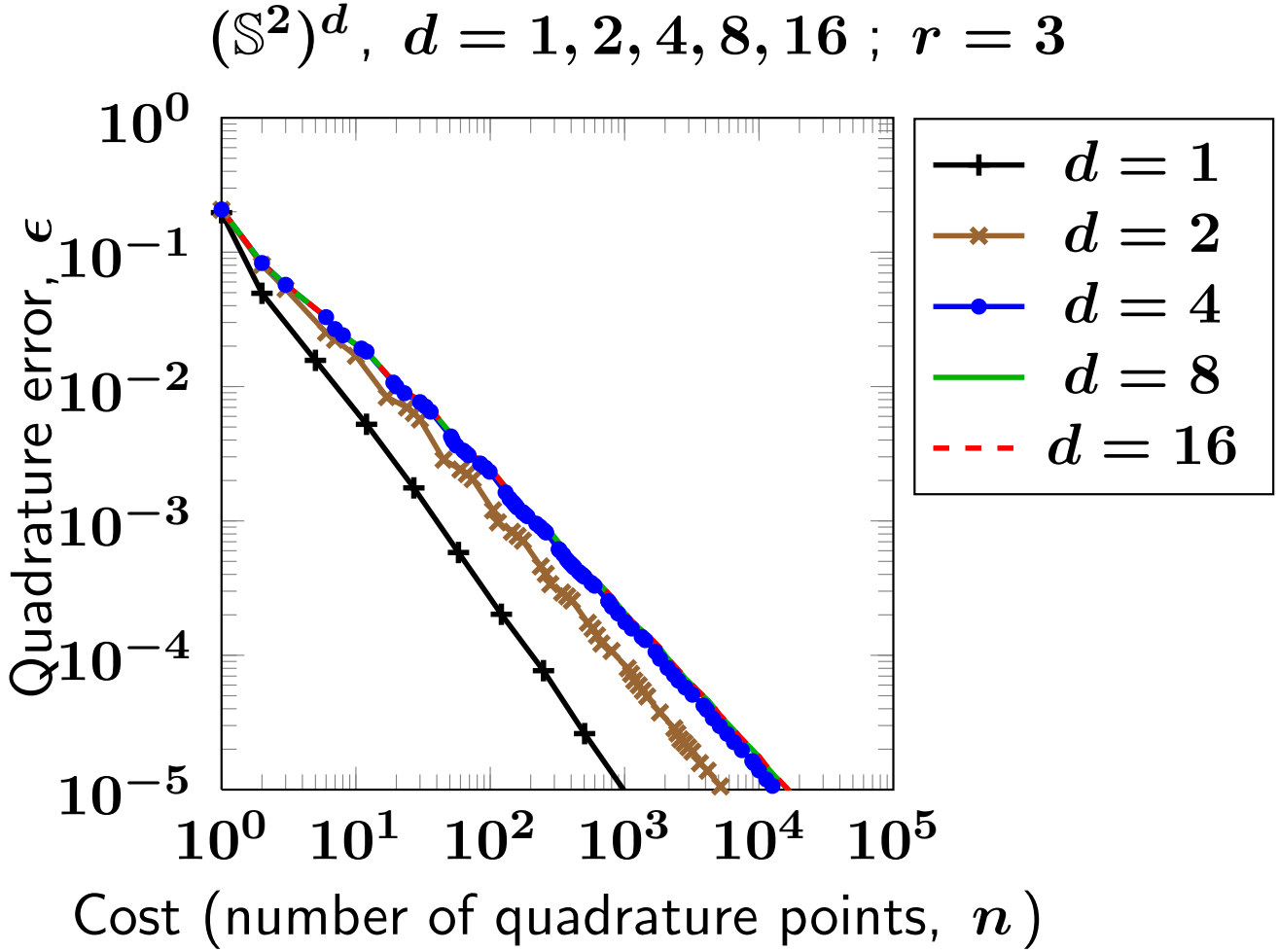
For the DA and WW weighted tensor product algorithms, each program run uses $r = 3$; $g = 0.1$, 0.5, or 0.9; dimension d, from $d = 1$ to 16; maximum 1-norm for indices, typically 20; and maximum number of points, up to $100\,000$.

Typical convergence behaviour

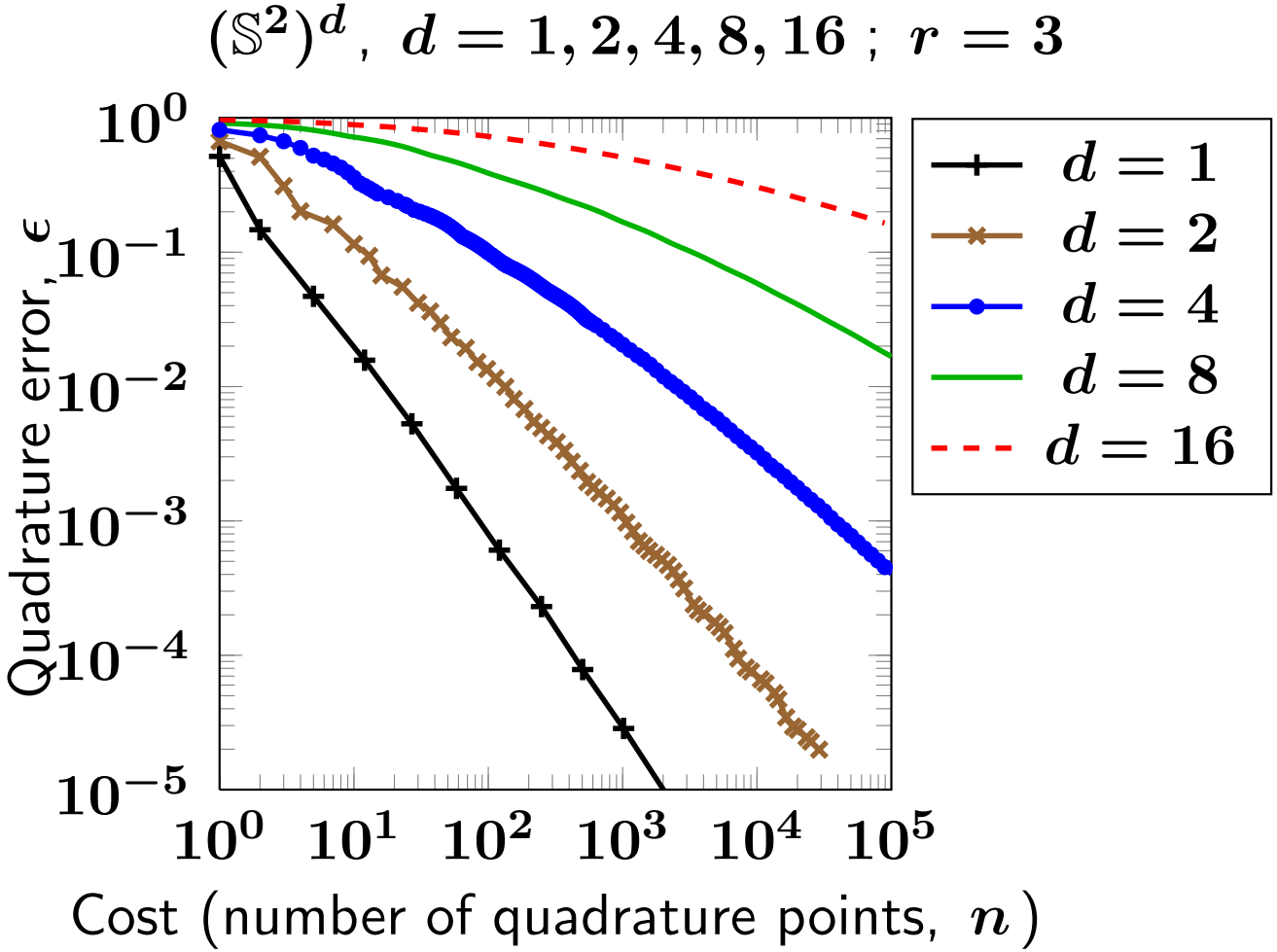
$(\mathbb{S}^2)^4$, $r = 3$, $\gamma_{4,k} = 0.5^k$



Error of DA rules for $\gamma_{d,k} = 0.1^k$.



Error of DA rules for $\gamma_{d,k} = 0.9^k$.



The main result, again

For quadrature in the setting of weighted tensor product spaces on the unit sphere, there is a choice of points on each sphere, such that a quadrature rule constructed using a greedy dimension adaptive sparse grid algorithm breaks the curse of dimensionality.

The asymptotic rate of convergence to zero of the worst case quadrature error on a product of spheres has essentially the same order as that for a single sphere.