Cubature formulas for function spaces with moderate smoothness

Michael Gnewuch, René Lindloh,* Reinhold Schneider, and Anand Srivastav
Institut für Informatik
Christian-Albrechts-Universität zu Kiel
Christian-Albrechts-Platz 4, 24098 Kiel, Germany
Tel.: +49-431/880-{7451, 7452, 7470, 7252}, Fax: +49-431/880-1725
E-mail: {mig, rha, rs, asr}@informatik.uni-kiel.de

To Henryk Woźniakowski on the occasion of his 60th birthday

Abstract

We construct simple algorithms for high-dimensional numerical integration of function classes with moderate smoothness. These classes consist of square-integrable functions over the d-dimensional unit cube whose coefficients with respect to certain multiwavelet expansions decay rapidly. Such a class contains discontinuous functions on the one hand and, for the right choice of parameters, the quite natural d-fold tensor product of a Sobolev space $H^s[0,1]$ on the other hand.

The algorithms are based on one-dimensional quadrature rules appropriate for the integration of the particular wavelets under consideration and on Smolyak's construction. We provide upper bounds for the worst-case error of our cubature rule in terms of the number of function calls. We additionally prove lower bounds showing that our method is optimal in dimension $d = 1$ and almost optimal (up to logarithmic factors) in higher dimensions. We perform numerical tests which allow the comparison with other cubature methods.

1 Introduction

The computation of high-dimensional integrals is a difficult task arising, e.g. from applications in physics, quantum chemistry, and finance. The traditional methods used in lower dimensions, such as product rules of one-dimensional quadratures, are usually too costly in high dimensions, since the number of function calls increases exponentially with the dimension.

*Partially supported by the DFG-Graduiertenkolleg 357 "Effiziente Algorithmen und Mehrskalenmethoden"

In this paper we present a cubature method which can be used to handle the following multivariate integration problem also in higher dimensions:

Problem definition. We want to approximate the integral

$$I(f) = \int_{[0,1]^d} f(x)dx$$

for functions $f : [0,1)^d \to \mathbb{R}$ belonging to function classes $\mathcal{H}$ of theoretical or practical interest. It is important from the view point of applicability of high-dimensional cubature that the function class is general and rich and contains important classes arising in numerical mathematics. A general cubature formula with N sample points $\{x_1, x_2, \ldots, x_N\} \subset [0,1]^d$ is given by

$$\mathrm{Q}_N(f) = \sum_{\nu=1}^{N} \lambda_\nu f(x_\nu),$$

where $\{\lambda_1, \ldots, \lambda_N\}$ is some suitable set of weights. To measure the quality of a given cubature Q_N we use the worst case error over $\mathcal{H}$ defined by

$$\mathrm{err}(\mathcal{H}, \mathrm{Q}_N) := \sup_{f \in \mathcal{H},\, \|f\|=1} \mathrm{err}(f, Q_N),$$

where

$$\mathrm{err}(f, Q_N) := |I(f) - Q_N(f)|.$$

As I and Q_N are linear, $\mathrm{err}(\mathcal{H}, \mathrm{Q}_N)$ is nothing but the operator norm $\|I - Q_N\|_{\mathrm{op}}$ induced by the norm of $\mathcal{H}$.

Results. The function classes we consider in this paper are certain Hilbert spaces $\mathcal{H}_s = \{f \in \mathrm{L}^2 \mid \|f\|_s < \infty\}$ which are spanned by multiwavelets $\{\psi_\lambda\}_\lambda$ and are characterized by discrete norms $\|f\|_s^2 = \sum_\lambda 2^{|\lambda|2s} \langle f, \psi_\lambda \rangle^2$. The functions in $\mathcal{H}_s$ are continuously embedded in $L^2[0,1]^d$ and under proper requirements $\mathcal{H}_s$ contains classical function spaces like Sobolev spaces. Our aim is to provide a cubature method that guarantees a (nearly) optimal worst case error and which is easy to implement.

For arbitrary parameters $s > 1/2$ we show that its worst case error over $\mathcal{H}_s$ is of the form $\mathcal{O}\left(\frac{\log(N)^{(d-1)(s+1/2)}}{N^s}\right)$, where N denotes the number of sample points used. We also prove a lower bound $\Omega\left(\frac{\log(N)^{(d-1)/2}}{N^s}\right)$ for all cubatures on $\mathcal{H}_s$ using N sample points. This shows that the presented integration method converges on $\mathcal{H}_s$ asymptotically almost optimal. Our cubatures are based on one-dimensional quadratures chosen with respect to the particular space $\mathcal{H}_s$ under consideration, and Smolyak's construction. More precisely, we use composite quadrature rules of a fixed order n. These rules are exact for piecewise polynomials of order n. The presented Smolyak construction is related to tensor product multiwavelet expansions in the way that the cubature is exact on finite multiwavelet series up to a critical level.

Related work. To some extent our work is motivated by [13], where the considered function classes depend on Haar wavelet series and a randomized cubature given by a

quasi-Monte Carlo rule using so-called scrambled nets (see, e.g. [15]) is studied. These classes of Haar wavelets are included in the classes of multiwavelets that we consider. Notice that cubature rules using scrambled nets are not exact for (piecewise) polynomials of higher degree, in contrast to our method.

It is known that Smolyak's construction leads in general to almost optimal approximations in any dimension $d > 1$ as long as the underlying one-dimensional quadrature rule is optimal. The application of Smolyak's construction to numerical integration has been studied in a number of papers so far, see, e.g. [3, 4, 11, 12, 14, 16, 20, 24] and the literature mentioned therein. The error bounds provided in these papers were usually proved on Korobov spaces or spaces of functions with bounded mixed derivatives, i.e., on spaces of functions with a certain degree of smoothness. For our method we provide good error bounds with respect to the Hilbert spaces $\mathcal{H}_s$ of not necessarily smooth functions. Note that the power of the logarithm of N in our upper bound is $(d-1)/2$ less than the power in the corresponding upper bounds appearing in the papers mentioned above.

This paper is organized as follows: In Section 2 we define multiwavelets and introduce the spaces on which our cubatures of prescribed level should be exact.

In Section 3 we present one-dimensional quadratures suited to evaluate the integrals of the univariate wavelets introduced in Section 2. We define a scale of Hilbert spaces of square integrable functions over $[0,1)$ via wavelet coefficients and prove an optimal error bound for our quadrature with respect to these spaces.

In Section 4 we use Smolyak's construction to obtain from our one-dimensional quadratures cubature rules for multivariate integrands. After giving a precise definition of the class of Hilbert spaces $\mathcal{H}_s$ of multivariate functions we want to consider error bounds for our cubatures; first in terms of the level of our cubatures, then in terms of the number of function calls. We provide also lower bounds for the worst case error of any cubature Q_N using N sample points. These lower bounds show that our cubature method is asymptotically almost optimal (up to logarithmic factors).

In Section 5 we report on several numerical tests which allow us to compare our method with known methods.

In Section 6 we provide a conclusion and make some remarks concerning future work.

2 Discontinuous multiwavelet bases

2.1 The one-dimensional case

We start by giving a short construction of a class of bases in $\mathrm{L}^2[0,1]$ that are called discontinuous multiwavelet bases. This topic has already been studied in the mathematical literature, see , e.g. [2, 18, 23].

By Π_n we denote the set of polynomials of order n, i.e., of degree strictly smaller than n, on $[0,1)$. Let $h_0, h_1, \ldots, h_{n-1}$ denote the set of the first n Legendre polynomials on the

interval $[0,1)$; an explicit expression of these polynomials is given by

$$h_j(x) = (-1)^j \sum_{k=0}^{j} \binom{j}{k} \binom{j+k}{k} (-x)^k$$

for all $x \in [0,1)$, see, e.g. [1]. These polynomials build an orthogonal basis of Π_n and are orthogonal on lower order polynomials,

$$\int_0^1 h_j(x) x^i dx = 0, \quad i = 0, 1, \ldots, j-1.$$

For convenience we extend the polynomials h_j by zero to the whole real line. With the help of these (piecewise) polynomials we define for $i = 0, 1, \ldots, n-1$ a set of scaling functions $\varphi_i(x) := h_i(x) / \|h_i\|_2$, where $\|\cdot\|_2$ is the usual norm on $L^2[0,1]$. For arbitrary $j \in \mathbb{N}_0$ we use the shorthand

$$\nabla_j := \left\{0, 1, 2, \ldots, 2^j - 1\right\}.$$

We consider dilated and translated versions

$$\varphi_{i,k}^j := 2^{j/2} \varphi_i(2^j \cdot - k), \quad i = 0, 1, \ldots, n-1,\ j \in \mathbb{N}_0,\ k \in \nabla_j,$$

of the scaling functions φ_i. Observe that these functions have compact support

$$\operatorname{supp} \varphi_{i,k}^j = [2^{-j}k, 2^{-j}(k+1)] =: I_k^j$$

and

$$\langle \varphi_{i,k}^j, \varphi_{i',k'}^j \rangle = \delta_{i,i'} \delta_{k,k'}.$$

Furthermore, we define spaces of piecewise polynomial functions of order n,

$$V_n^j := \operatorname{span}\left\{\varphi_{i,k}^j \mid \quad i = 0, 1, \ldots, n-1,\ k \in \nabla_j\right\}.$$

It is obvious that the spaces V_n^j have dimension $2^j n$ and that they are nested in the following way:

$$\Pi_n = V_n^0 \subset V_n^1 \subset \cdots \subset \mathrm{L}^2[0,1].$$

For $j = 0, 1, 2, \ldots$ we define the $2^j n$-dimensional space W_n^j to be the orthogonal complement of V_n^j in V_n^{j+1}, i.e.,

$$W_n^j := \left\{\psi \in V_n^{j+1} \mid \langle \psi, \varphi \rangle = 0 \text{ for all } \varphi \in V_n^j\right\}.$$

This leads to the orthogonal decomposition

$$V_n^j = V_n^0 \oplus W_n^0 \oplus W_n^1 \oplus \cdots \oplus W_n^{j-1}$$

of V_n^j.

Let $(\psi_i)_{i=0}^{n-1}$ be an orthonormal basis of W_n^0. (An explicit construction of such a basis in more general situations is, e.g. given in [18, Subsec. 5.4.1].) Then it is straightforward to verify that the $2^j n$ functions

$$\psi_{i,k}^j := 2^{j/2}\psi_i(2^j \cdot -k), \quad i = 0, \dots, n-1,\ k \in \nabla_j,$$

form an orthonormal basis of W_n^j. The functions $(\psi_i)_{i=0}^{n-1}$ are called multiwavelets and are obviously also piecewise polynomials of degree strictly less than n. Multiwavelets are supported on canonical intervals

$$\operatorname{supp} \psi_{i,k}^j = I_k^j$$

and satisfy the orthogonality condition

$$\langle \psi_{i,k}^j, \psi_{l,n}^m \rangle = \delta_{i,l}\delta_{j,m}\delta_{k,n}.$$

Since the spaces W_n^j are orthogonal to $V_n^0 = \Pi_n$, we have vanishing moments

$$\int_0^1 \psi_{i,k}^j(x) x^\nu dx = 0, \quad \nu = 0, 1, \dots, n-1.$$

Next we define the space

$$V := \bigcup_{j=0}^{\infty} V_n^j = V_n^0 \oplus \bigoplus_{j=0}^{\infty} W_n^j. \tag{2.1}$$

Notice that V contains all elements of the well-known Haar basis; therefore V is dense in $\mathrm{L}^2[0,1]$.

We follow the convention from [18] and define $\psi_i^{-1} := \varphi_i$ (please do not confuse this notation with the notation of inverse functions), $\nabla_{-1} := \{0\}$ and $I_0^{-1} := [0,1]$. A so-called multiwavelet basis of order n for $\mathrm{L}^2[0,1]$ is given by

$$\left\{ \psi_{i,k}^j \mid \quad i = 0, 1, \dots, n-1,\ j \geq -1,\ k \in \nabla_j \right\},$$

and for every $f \in \mathrm{L}^2[0,1]$ we get the following unique multiwavelet expansion

$$f = \sum_{j \geq -1} \sum_{k \in \nabla_j} \sum_{i=0}^{n-1} \langle f, \psi_{i,k}^j \rangle \psi_{i,k}^j.$$

2.2 The multivariate case

In this subsection we extend the concept of multiwavelet bases to higher dimensions. Here we follow an approach that is suitable for our later analysis. For a given multi-index $\mathbf{j} \in \mathbb{Z}^d$ we put $|\mathbf{j}| := j_1 + j_2 + \cdots + j_d$, and for $\mathbf{i} \in \mathbb{N}_0^d$ let $|\mathbf{i}|_\infty := \max\{i_1, \dots, i_d\}$. A

multivariate multiwavelet basis of $\mathrm{L}^2[0,1]^d$ is given by so-called tensor product wavelets. For $n \in \mathbb{N}$, we define the approximation space on level L by

$$V_n^{d,L} := \sum_{|\mathbf{j}|=L} \bigotimes_{i=1}^{d} V_n^{j_i}. \tag{2.2}$$

Similarly to the one-dimensional case we put

$$V^d := \bigcup_{L=0}^{\infty} V^{d,L}.$$

Since $V = V^1$ is dense in $L^2[0,1]$, the space V^d is dense in $L^2[0,1]^d$. Thus we obtain the following expansion for $f \in \mathrm{L}^2[0,1]^d$

$$f = \sum_{\mathbf{j} \geq -1} \sum_{\mathbf{k} \in \nabla_{\mathbf{j}}} \sum_{|\mathbf{i}|_\infty = 0}^{n-1} \langle f, \Psi_{\mathbf{i},\mathbf{k}}^{\mathbf{j}} \rangle \Psi_{\mathbf{i},\mathbf{k}}^{\mathbf{j}},$$

where $\mathbf{j} = (j_1, \ldots, j_d) \geq -1$ is meant in the way that $j_u \geq -1$ for all $u = 1, \ldots, d$. (In the following all inequalities between vectors and between a vector and a scalar are meant componentwise.) Furthermore, we used the shorthands $\nabla_{\mathbf{j}} = \nabla_{j_1} \times \ldots \times \nabla_{j_d}$ and

$$\Psi_{\mathbf{i},\mathbf{k}}^{\mathbf{j}} := \bigotimes_{u=1}^{d} \psi_{i_u,k_u}^{j_u}.$$

If the d-dimensional canonical interval $I_{\mathbf{k}}^{\mathbf{j}}$ is defined by

$$I_{\mathbf{k}}^{\mathbf{j}} := I_{k_1}^{j_1} \times I_{k_2}^{j_2} \times \ldots \times I_{k_d}^{j_d},$$

then $\mathrm{supp}\Psi_{\mathbf{i},\mathbf{k}}^{\mathbf{j}} = I_{\mathbf{k}}^{\mathbf{j}}$ holds.

3 One-dimensional integration

3.1 One-dimensional quadrature formulas

Recall that a general one-dimensional quadrature is given by

$$\mathrm{Q}_m(f) = \sum_{\nu=1}^{m} \lambda_\nu f(x_\nu), \tag{3.1}$$

where $x_1, \ldots, x_m \subset [0,1]$ are the sample points, and $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ are the weights. Since we are here interested in quadrature formulas with high polynomial exactness—like the Newton-Cotes, Clenshaw-Curtis or Gauss formulas—we confine ourselves to the case

$\sum_{\nu=1}^{m} \lambda_\nu = 1$. For a detailed discussion of one-dimensional quadrature formulas see, e.g. [7].

Our aim is to give a simple construction of quadrature formulas Q_N which satisfy for a given polynomial order n and a so-called critical level l

$$\mathrm{err}(h, \mathrm{Q}_N) = 0 \quad \text{for all } h \in V_n^l.$$

We get the requested quadrature by scaling and translating a simpler one-dimensional quadrature formula Q_m that is exact for all polynomials of order n on $[0,1]$. If Q_m has the explicit form (3.1), then our resulting quadrature uses $2^l m$ sample points and is given by

$$A_m(l,1)(f) := \sum_{k \in \nabla_l} \sum_{\nu=1}^{m} 2^{-l} \lambda_\nu f(2^{-l} x_\nu + 2^{-l} k). \tag{3.2}$$

$A_m(l,1)$ is exact for polynomials on canonical intervals I_k^j, $j \le l$, $k \in \nabla_j$, of degree strictly less than n and therefore also on the whole space V_n^l.

Let us call a sequence of quadratures or cubatures $(Q_N)_N$ nested if the corresponding sets of sample points $(X_N)_N$ are nested, i.e., if $X_N \subseteq X_{N+1}$ for all N. Whether our quadratures $(A_m(l,1))_l$ are nested or not depends of course on the set of sample points X of the underlying quadrature Q_m. If we, e.g., consider the case $n = 1$, then we may choose Q_m to be the mid point rule $Q_m(f) = f(1/2)$, which results in the non-nestedness of our quadratures $(A_m(l,1))_l$. If we choose on the other hand the rule $Q_m(f) = f(0)$, then our quadratures are indeed nested. (Notice that in the latter case $A_m(l,1)$ is nothing but the iterated trapezoidal rule for periodic functions.)

3.2 Error analysis

For the error analysis of our one-dimensional quadrature method let $n \in \mathbb{N}$, and let

$$\left\{ \psi_{i,k}^j \mid \quad i = 0, 1, 2, \ldots, n-1,\ j \ge -1,\ k \in \nabla_j \right\}$$

be the multiwavelet basis of order n defined in Section 2.1. For $s > 0$ we define a discrete norm

$$|f|_{s,n}^2 := \sum_{j \ge -1} \sum_{k \in \nabla_j} \sum_{i=0}^{n-1} 2^{j 2s} \langle f, \psi_{i,k}^j \rangle^2 \tag{3.3}$$

on the space

$$\mathcal{H}_{s,n} := \left\{ f \in \mathrm{L}^2[0,1] \mid \ |f|_{s,n} < \infty \right\}, \tag{3.4}$$

consisting of functions whose wavelet coefficients decrease rapidly. Point evaluations are obviously well defined on the linear span of the functions $\psi_{i,k}^j$, $i = 0, 1, \ldots, n-1$, $j \ge -1$,

$k \in \nabla_j$. Moreover, it is easy to see that they can be extended to bounded linear functionals on $\mathcal{H}_{s,n}$ as long as $s > 1/2$. On these spaces quadrature formulas are therefore well defined.

Now we choose an $m = m(n)$ and an underlying quadrature rule Q_m as in (3.1) such that Q_m is exact on Π_n. Let $A_m(l,1)$ be as in (3.2). Then the wavelet expansion of a function $f \in \mathcal{H}_{s,n}$ and the Cauchy-Schwarz inequality yield the following error bound for our algorithm $A_m(l,1)$:

Theorem 3.1. *Let $s > 1/2$ and $n \in \mathbb{N}$. Let Q_m and $A_m(l,1)$ as above. Then there exists a constant $C > 0$ such that*

$$\operatorname{err}(\mathcal{H}_{s,n}, A_m(l,1)) \;\leq\; C\, 2^{-ls}. \tag{3.5}$$

Proof. Let $f \in \mathcal{H}_{s,n}$. The quadrature error is given by

$$\begin{aligned} \operatorname{err}(f, A_m(l,1)) &= |I(f) - A_m(l,1)f| \\ &= \left| \sum_{j\geq -1} \sum_{k\in\nabla_j} \sum_{i=0}^{n-1} \langle f, \psi_{i,k}^j \rangle \left\{ I(\psi_{i,k}^j) - A_m(l,1)\psi_{i,k}^j \right\} \right|. \end{aligned}$$

The Cauchy-Schwarz inequality yields

$$\operatorname{err}(f, A_m(l,1)) \;\leq\; |f|_{s,n} \left(\sum_{j\geq -1} \sum_{k\in\nabla_j} \sum_{i=0}^{n-1} 2^{-j2s} \left\{ I(\psi_{i,k}^j) - A_m(l,1)\psi_{i,k}^j \right\}^2 \right)^{1/2}.$$

Recall that the Cauchy-Schwarz inequality leads to a tight worst case error bound. Because of the polynomial exactness and vanishing moments we get therefore

$$\operatorname{err}(\mathcal{H}_{s,n}, A_m(l,1))^2 = \sum_{j\geq l} \sum_{k\in\nabla_j} \sum_{i=0}^{n-1} 2^{-j2s} \left\{ A_m(l,1)\psi_{i,k}^j \right\}^2.$$

By some easy calculations and with the identities $\operatorname{supp} \psi_{i,k}^j = I_k^j$ and $\|\psi_{i,k}^j\|_\infty = 2^{j/2}\|\psi_i\|_\infty$ we get

$$\begin{aligned} & \operatorname{err}(\mathcal{H}_{s,n}, A_m(l,1))^2 \\ \leq\; & \sum_{j\geq l} \sum_{k\in\nabla_j} \sum_{i=0}^{n-1} 2^{-j2s} \left\{ \sum_{k'\in\nabla_l} \sum_{\nu=1}^{m} 2^{-l} |\lambda_\nu| \left\|\psi_{i,k}^j\right\|_\infty \mathbf{1}_{I_k^j}(2^{-l}x_\nu + 2^{-l}k') \right\}^2 \\ =\; & \sum_{j\geq l} 2^{-2l} 2^{j(1-2s)} \sum_{i=0}^{n-1} \|\psi_i\|_\infty^2 \sum_{k\in\nabla_j} \left\{ \sum_{k'\in\nabla_l} \sum_{\nu=1}^{m} |\lambda_\nu| \, \mathbf{1}_{I_k^j}(2^{-l}x_\nu + 2^{-l}k') \right\}^2. \end{aligned}$$

For $j \geq l$ and $k \in \nabla_j$ let $\kappa = \kappa(j,k,l)$ be the unique element $\kappa \in \nabla_l$ such that

$$2^{-l}\kappa \leq 2^{-j}k < 2^{-j}(k+1) \leq 2^{-l}(\kappa+1).$$

Then

$$\begin{aligned}
&\operatorname{err}(\mathcal{H}_{s,n}, A_m(l,1))^2 \\
\leq\ & \sum_{j\geq l} 2^{-2l} 2^{j(1-2s)} \sum_{i=0}^{n-1} \|\psi_i\|_\infty^2 \sum_{k\in\nabla_j} \left\{ \sum_{\nu=1}^{m} |\lambda_\nu| \, \mathbf{1}_{I_k^j}(2^{-l}x_\nu + 2^{-l}\kappa) \right\}^2 \\
\leq\ & \sum_{j\geq l} 2^{-2l} 2^{j(1-2s)} \sum_{i=0}^{n-1} \|\psi_i\|_\infty^2 \sum_{\kappa\in\nabla_l} \left\{ \sum_{\nu=1}^{m} |\lambda_\nu| \, \mathbf{1}_{I_\kappa^l}(2^{-l}x_\nu + 2^{-l}\kappa) \right\}^2 \\
=\ & \sum_{j\geq l} 2^{-2l} 2^{j(1-2s)} \sum_{i=0}^{n-1} \|\psi_i\|_\infty^2 \, |\nabla_l| \left(\sum_{\nu=1}^{m} |\lambda_\nu| \right)^2 .
\end{aligned}$$

Note that $|\nabla_l| = 2^l$. We can upper bound the integration error by

$$\begin{aligned}
\operatorname{err}(\mathcal{H}_{s,n}, A_m(l,1))^2 &\leq \sum_{i=0}^{n-1} \|\psi_i\|_\infty^2 \left(\sum_{\nu=1}^{m} |\lambda_\nu| \right)^2 2^{-l} \sum_{j\geq l} 2^{j(1-2s)} \\
&= \sum_{i=0}^{n-1} \|\psi_i\|_\infty^2 \left(\sum_{\nu=1}^{m} |\lambda_\nu| \right)^2 2^{-l2s} \sum_{j\geq 0} 2^{j(1-2s)} \\
&= \sum_{i=0}^{n-1} \|\psi_i\|_\infty^2 \left(\sum_{\nu=1}^{m} |\lambda_\nu| \right)^2 \frac{2^{-l2s}}{1-2^{(1-2s)}}.
\end{aligned}$$

Thus we proved that (3.5) holds with the constant

$$C = \frac{1}{\sqrt{1-2^{1-2s}}} \left(\sum_{i=0}^{n-1} \|\psi_i\|_\infty^2 \right)^{1/2} \sum_{\nu=1}^{m} |\lambda_\nu|.$$

□

Remark 3.2. The error estimate in Theorem 3.1 is asymptotically optimal as Theorem 4.9 will reveal.

4 Multivariate numerical integration

4.1 The d-dimensional cubature method

Now we extend our one-dimensional algorithm $A_m(l,1)$ to a d-dimensional cubature. This should be done via Smolyak's construction:

The so-called difference quadrature of level $l \geq 0$ is defined by

$$\Delta^l := A_m(l,1) - A_m(l-1,1),$$

with $A_m(-1,1) := 0$. Smolyak's construction of level L is then given by

$$A_m(L,d) \;:=\; \sum_{\mathbf{l}\in\mathbb{N}_0^d,\,|\mathbf{l}|\le L} (\Delta^{l_1}\otimes\Delta^{l_2}\otimes\cdots\otimes\Delta^{l_d}).$$

Notice that we have $\Delta^0 = \mathrm{Q}_m$. Let us recall that in the one-dimensional case $A_m(l,1)$ is exact on V_n^l. In the d-dimensional case, it is not too difficult to show the exactness of $A_m(L,d)$ on $V_n^{d,L}$.

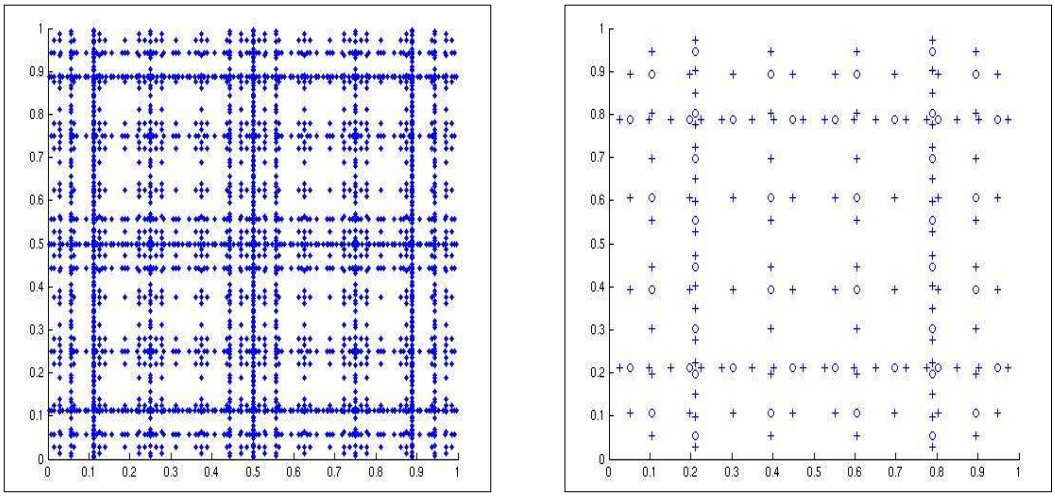

Figure 1: $A_3(5,2)$ and $A_2(3,2)$ with underlying Gauss quadrature. In the right diagram "+" denotes sample points with positive, "o" sample points with negative weights.

Theorem 4.1. *The cubature $A_m(L,d)$ is exact on the approximation space $V_n^{d,L}$.*

The proof follows the lines of the proof of [14, Theorem 2] and proceeds via induction over the dimension.

4.2 Upper bounds for the cubature error

For the error analysis we consider product spaces which are based on the spaces $\mathcal{H}_{s,n}$ used for our one-dimensional quadrature error bounds. These seem to be the natural spaces for our variation of Smolyak's construction. For a function f we define a norm

$$|f|_{d,s,n}{}^2 \;:=\; \sum_{\mathbf{j}\ge -1}\sum_{\mathbf{k}\in\nabla_{\mathbf{j}}}\sum_{|\mathbf{i}|_\infty=0}^{n-1} 2^{|\mathbf{j}|2s}\langle f, \Psi_{\mathbf{i},\mathbf{k}}^{\mathbf{j}}\rangle^2 \tag{4.1}$$

and the space

$$\mathcal{H}_{s,n}^d := \{f\in \mathrm{L}^2[0,1]^d \mid |f|_{d,s,n} < \infty\}.$$

In [24, Lemma 2] Wasilkowski and Woźniakowski provided an error bound that is valid not only for d-dimensional cubatures, but also for more general d-dimensional approximation algorithms based on Smolyak's construction. Adapting the corresponding proof, we see that our one dimensional error bound from Theorem 3.1 implies the following result.

Theorem 4.2. *Let $d, n \in \mathbb{N}$, and let the one-dimensional quadrature Q_m be exact on Π_n. For $s > 1/2$ let C be the constant from (3.5). The worst case error of $A_m(L,d)$ satisfies*

$$\operatorname{err}(\mathcal{H}^d_{s,n}, A_m(L,d)) \le C\left(\max\{2^s, C\,(1+2^s)\}\right)^{d-1} 2^{-Ls} \binom{L+d}{d-1}.$$

Instead of explaining the proof in detail we want to provide a better upper bound in which essentially the term $\binom{L+d}{d-1} \sim L^{d-1}$ is replaced by $L^{(d-1)/2}$. Before establishing the corresponding theorem, we state a simple helpful lemma and a well-known identity.

Lemma 4.3. *Let $\mathbf{i}, \mathbf{j} \in \mathbb{Z}^d$ with $n-1 \ge \mathbf{i} \ge 0$, $\mathbf{j} \ge -1$, and let $\mathbf{k} \in \nabla_{\mathbf{j}}$. Assume that $j_d = -1$ and $i_d = 0$. If $\mathbf{i}', \mathbf{j}'$ and $\mathbf{k}'$ denote the $(d-1)$-dimensional vectors consisting of the first $d-1$ components of $\mathbf{i}, \mathbf{j}$ and $\mathbf{k}$ respectively, then we have for all $L \in \mathbb{N}_0$*

$$A_m(L,d)\Psi^{\mathbf{j}}_{\mathbf{i},\mathbf{k}} = A_m(L,d-1)\Psi^{\mathbf{j}'}_{\mathbf{i}',\mathbf{k}'}\,.$$

Proof. We have $\psi^{j_d}_{i_d,k_d} = \psi_0^{-1} = \varphi_0 = \mathbf{1}_{[0,1)}$, implying $\Delta^0\psi^{j_d}_{i_d,k_d} = 1$ and $\Delta^\nu\psi^{j_d}_{i_d,k_d} = 0$ for all $\nu \ge 1$. Now the lemma follows immediately from the definition of $A_m(L,d)$. □

A well-known formula expressing $A_m(L,d)$ solely in terms of tensor quadratures is

$$A_m(L,d) = \sum_{L-d+1 \le |\mathbf{l}| \le L} (-1)^{L-|\mathbf{l}|} \binom{d-1}{L-|\mathbf{l}|} \bigotimes_{u=1}^{d} A_m(l_u, 1). \tag{4.2}$$

A proof of this identity can, e.g., be found in [24, Lemma 1].

Theorem 4.4. *Let $d, n \in \mathbb{N}$, and let the one-dimensional quadrature Q_m be exact on Π_n. For $s > 1/2$ there exists a constant $C > 0$ such that for all $L > 0$ the worst case error of $A_m(L,d)$ satisfies*

$$\operatorname{err}(\mathcal{H}^d_{s,n}, A_m(L,d)) \le C 2^{-Ls} L^{\frac{d-1}{2}}.$$

Proof. For the sake of brevity we do not try to give a reasonably good bound for the constant C in the theorem; instead we use rather rough estimates and a generic constant C, which may depend on n, m, s and d, but not on the given level L.

We proceed via induction on d. The case $d = 1$ has already been treated in Theorem 3.1. Let now $d \ge 2$, and let the induction hypothesis hold for $d-1$. Similarly as in the one-dimensional case we can use the Cauchy-Schwarz inequality and the exactness of $A_m(L,d)$ on $V_n^{d,L}$ to get

$$\operatorname{err}(\mathcal{H}^d_{s,n}, A_m(L,d))^2 = \sum_{|\mathbf{j}| \ge L-d+1} \sum_{\mathbf{k} \in \nabla_{\mathbf{j}}} \sum_{|\mathbf{i}|_\infty = 0}^{n-1} 2^{-|\mathbf{j}|2s} \{A_m(L,d)\Psi^{\mathbf{j}}_{\mathbf{i},\mathbf{k}}\}^2\,;$$

hereby note that $\Psi_{\mathbf{i},\mathbf{k}}^{\mathbf{j}} \in \otimes_{\nu=1}^{d} V_n^{l_\nu}$ if and only if $l_\nu > j_\nu$ for all $\nu \in \{1,\dots,d\}$, i.e., $\Psi_{\mathbf{i},\mathbf{k}}^{\mathbf{j}} \in V_n^{d,L}$ if and only $|\mathbf{j}| < L-d+1$.

To avoid technical difficulties, we now show that the summation over the index sets

$$U(\nu) := \{(\mathbf{i},\mathbf{j}) \,|\, n-1 \geq |\mathbf{i}|_\infty \geq 0,\, i_\nu = 0,\, |\mathbf{j}| \geq L-d+1,\, j_\nu = -1\}$$

for all $\nu \in \{1,\dots,d\}$ contributes not essentially to the square of the worst case error. Indeed, if $\mathbf{i}'$, $\mathbf{j}'$, and $\mathbf{k}'$ denote $(d-1)$-dimensional vectors, then Lemma 4.3 yields

$$\begin{aligned}
&\sum_{(\mathbf{i},\mathbf{j})\in U(\nu)} \sum_{\mathbf{k}\in\nabla_{\mathbf{j}}} 2^{-|\mathbf{j}|2s} \{A_m(L,d)\Psi_{\mathbf{i},\mathbf{k}}^{\mathbf{j}}\}^2 \\
&= \sum_{|\mathbf{i}'|_\infty=0}^{n-1} \sum_{|\mathbf{j}'|\geq L-(d-2)} \sum_{\mathbf{k}'\in\nabla_{\mathbf{j}'}} 2^{-(|\mathbf{j}'|-1)2s} \{A_m(L,d-1)\Psi_{\mathbf{i}',\mathbf{k}'}^{\mathbf{j}'}\}^2 \\
&= 2^{2s}\,\mathrm{err}(\mathcal{H}_{s,n}^{d-1}, A_m(L,d-1))^2 \leq C 2^{-2Ls} L^{d-2},
\end{aligned}$$

where in the last step we used the induction hypothesis.

So let us now consider solely pairs $(\mathbf{i},\mathbf{j})$ where for all $\nu \in \{1,\dots,d\}$ we have $i_\nu \geq 1$ or $j_\nu \geq 0$. For such pairs $(\mathbf{i},\mathbf{j})$, for $\mathbf{k} \in \nabla_{\mathbf{j}}$ and $\mu \in \{L-d+1,\dots,L\}$ let us define

$$S_{\mathbf{i},\mathbf{k}}^{\mathbf{j},\mu} := \left\{ \mathbf{l} \in \mathbb{N}_0^d \,|\, |\mathbf{l}| = \mu \,\wedge\, \bigotimes_{u=1}^{d} A_m(l_u,1)\psi_{i_u,k_u}^{j_u} \neq 0 \right\}.$$

If $j_u < l_u$, then, due to the exactness of $A_m(l_u,1)$ on $V_n^{l_u}$, we have $A_m(l_u,1)\psi_{i_u,k_u}^{j_u} = 0$ (since $j_u \neq -1$ or $i_u \neq 0$). Thus

$$S_{\mathbf{i},\mathbf{k}}^{\mathbf{j},\mu} \subseteq \tilde{S}_{\mathbf{j},\mu} := \left\{ \mathbf{l} \in \mathbb{N}_0^d \,|\, |\mathbf{l}| = \mu \,\wedge\, \forall u \in \{1,\dots,d\} : l_u \leq j_u \right\}.$$

A coarse estimate of the cardinality of $\tilde{S}_{\mathbf{j},\mu}$ is

$$|\tilde{S}_{\mathbf{j},\mu}| \leq \binom{|\mathbf{j}| - \mu + d - 1}{d-1}.$$

(One can verify this bound by starting with $\mathbf{j}$ and counting the ways to distribute the difference $\mu - |\mathbf{j}|$ to the components of $\mathbf{j}$ to get an $\mathbf{l} \in \mathbb{Z}^d$ with $|\mathbf{l}| = \mu$ and $l_u \leq j_u$ for $1 \leq u \leq d$.)

With these observations and with identity (4.2) we get

$$\begin{aligned}
\mathrm{err}(\mathcal{H}_{s,n}^{d}, A_m(L,d))^2 \leq\, & C\, 2^{-2Ls} L^{d-2} + \\
& C \sum_{|\mathbf{i}|_\infty=0}^{n-1} \sum_{|\mathbf{j}|\geq L-d+1} \sum_{\mathbf{k}\in\nabla_{\mathbf{j}}} 2^{-|\mathbf{j}|2s} \left\{ \sum_{\mu=L-d+1}^{L} \sum_{\mathbf{l}\in\tilde{S}_{\mathbf{j},\mu}} \left| \bigotimes_{u=1}^{d} A_m(l_u,1)\psi_{i_u,k_u}^{j_u} \right| \right\}^2.
\end{aligned}$$

Since for $\mathbf{l} \in \tilde{S}_{\mathbf{j},\mu}$ the tensor quadrature $\bigotimes_{u=1}^{d} A_m(l_u,1)$ uses at most m^d sample points from $\mathrm{supp}(\Psi_{\mathbf{i},\mathbf{k}}^{\mathbf{j}})$, we have

$$\left| \bigotimes_{u=1}^{d} A_m(l_u,1)\psi_{i_u,k_u}^{j_u} \right| \leq 2^{-|\mathbf{l}|} m^d \prod_{u=1}^{d} \|\psi_{i_u,k_u}^{j_u}\|_\infty \leq 2^{-\mu} m^d 2^{|\mathbf{j}|/2} M,$$

where

$$M := \left(\max_{i=0}^{n-1}\{\|\varphi_i\|_\infty, \|\psi_i\|_\infty\}\right)^d.$$

Since each of the tensor quadratures $\bigotimes_{u=1}^{d} A_m(l_u, 1)$ uses not more than $m^d 2^L$ points, we have to make at most $C\binom{|\mathbf{j}|-(L-d+1)+d-1}{d-1}2^L$ function evaluations to calculate the term inside the parentheses. For fixed $\mathbf{i}$ and $\mathbf{j}$ all the $\Psi_{\mathbf{i},\mathbf{k}}^{\mathbf{j}}$, $\mathbf{k} \in \nabla_{\mathbf{j}}$, have pairwise disjoint support and thus only the summation over some subset $\tilde{\nabla}_{\mathbf{j}}$ of $\nabla_{\mathbf{j}}$ with

$$|\tilde{\nabla}_{\mathbf{j}}| \leq C\binom{|\mathbf{j}| - (L-d+1) + d - 1}{d-1}2^L$$

yields a non-trivial contribution to our estimate. Altogether we get (suppressing the lower order term $C2^{-2Ls}L^{d-2}$)

$$\mathrm{err}(\mathcal{H}_{s,n}^d, A_m(L,d))^2 \leq C \sum_{\nu=L-d+1}^{\infty} \sum_{|\mathbf{j}|=\nu} |\tilde{\nabla}_{\mathbf{j}}| 2^{-\nu 2s} \left(\binom{\nu - (L-d+1) + d - 1}{d-1} 2^{-L} 2^{\nu/2}\right)^2.$$

Our estimate for $|\tilde{\nabla}_{\mathbf{j}}|$ and $|\{\mathbf{j} \,|\, \mathbf{j} \geq -1, |\mathbf{j}| = \nu\}| = \binom{\nu+2d-1}{d-1}$ lead to

$$\begin{aligned}
\mathrm{err}(\mathcal{H}_{s,n}^d, A_m(L,d))^2 &\leq C\, 2^{-L} \sum_{\nu=L-d+1}^{\infty} 2^{\nu(1-2s)} \binom{\nu+2d-1}{d-1} \binom{\nu-(L-d+1)+d-1}{d-1}^3 \\
&\leq C\, 2^{-L} 2^{(L-d+1)(1-2s)} \sum_{\nu=0}^{\infty} 2^{\nu(1-2s)} \binom{L+\nu+d}{d-1} \binom{\nu+d-1}{d-1}^3 \\
&\leq C \left(\sum_{\nu=0}^{\infty} 2^{\nu(1-2s)} \binom{\nu+d-1}{d-1}^4\right) 2^{-2Ls} L^{d-1}.
\end{aligned}$$

The sum inside the parentheses converges as $s > 1/2$. □

From the abstract definition of our function space $\mathcal{H}_{s,n}$ it is not immediately clear if it contains a reasonable class of interesting functions away from the piecewise polynomials. At least in the case where the parameter n is strictly larger than s, the Sobolev space $H^s[0,1]$ is continuously embedded in $\mathcal{H}_{s,n}$. There are several ways to define Sobolev spaces with non-integer index $s \in \mathbb{R}$, one can use for example the Fourier transform

$$\hat{f}(\xi) := \int_{\mathbb{R}} f(x) e^{-ix\xi} dx$$

to define the norm

$$\|f\|_s = \int_{\mathbb{R}} (1 + |y|^2)^s \left|\hat{f}(y)\right|^2 dy$$

and the space

$$H^s(\mathbb{R}) = \left\{ f \in \mathrm{L}^2 |\ \|f\|_s < \infty \right\}.$$

For the interval $[0,1]$ we define

$$H^s[0,1] = H^s(\mathbb{R})|_{[0,1]}$$

by restriction , i.e., $f \in H^s[0,1]$ if there exists a function $g \in H^s(\mathbb{R})$ such that in the sense of distributions $g|_{[0,1]} = f$ and

$$\|f\|_{H^s[0,1]} = \inf_{g:\, f=g|_{[0,1]}} \|g\|_s .$$

The continuous embedding of $H^s[0,1]$ into $\mathcal{H}_{s,n}$ is established by some Jackson type inequality.

Theorem 4.5. *Let $(\psi_i)_{i=0}^{n-1}$ be multiwavelets of order n. For all $s < n$ the inclusion $H^s[0,1] \subset \mathcal{H}_{s,n}$ holds. More precisely, there exists a constant $K > 0$ such that for every $f \in H^s[0,1]$ we have*

$$\sum_{j\geq -1} \sum_{k\in\nabla_j} \sum_{i=0}^{n-1} 2^{j2s} \langle f, \psi_{i,k}^j \rangle^2 \leq K^2 \|f\|^2_{H^s[0,1]} .$$

For a proof of the theorem see, e.g. [5, 18, 23]. Notice that in general we cannot hope to prove equivalence of the norms on $\mathcal{H}_{s,n}$ and $H^s[0,1]$. This is obvious in the case where $s > 1/2$: $\mathcal{H}_{s,n}$ contains discontinuous functions, while $H^s[0,1]$ does not.

The mixed Sobolev space H^s_{mix} is defined by

$$H^s_{\text{mix}} = \underbrace{H^s[0,1] \otimes H^s[0,1] \otimes \cdots \otimes H^s[0,1]}_{d \text{ times}},$$

i.e., it is the complete d-fold tensor product of the Hilbert space $H^s[0,1]$. In terms of H^s_{mix} Theorem 4.4 reads as follows:

Corollary 4.6. *Let $s > 1/2$ and $n > s$. Let the one-dimensional quadrature Q_m be exact on Π_n. Then there exists a constant $C > 0$ such that for every $L > 0$*

$$\text{err}(H^s_{\text{mix}}, A_m(L,d)) \leq C 2^{-Ls} L^{\frac{d-1}{2}}.$$

Now we analyze the cost of the cubature algorithm $A_m(L,d)$. Identity (4.2) shows clearly that the number of multiplications and additions performed by the algorithm $A_m(L,d)$ is more or less proportional to the number of function evaluations. Since the cost of one function evaluation is in general much greater than the cost of an arithmetic operation, we concentrate here on the number of sample points $N = N_m(L,d)$ used by $A_m(L,d)$. Since for $\mathbf{l} \in \mathbb{N}_0^d$ and a general d-variate function f the operator $\bigotimes_{u=1}^d A_m(l_u, 1)$ uses $2^{|\mathbf{l}|} m^d$ function values, identity (4.2) gives us

$$\begin{aligned} N &\leq \sum_{L-d+1\leq|\mathbf{l}|\leq L} 2^{|\mathbf{l}|} m^d \\ &\leq m^d 2^L \sum_{j=0}^{d-1} 2^{j-d+1} \binom{L+j}{d-1} \leq m^d 2^{L+1} \binom{L+d-1}{d-1}. \end{aligned}$$

The bound on N can be improved if our cubatures $(A_m(L,d))_L$ are nested, i.e., if the set of sample points used by $A_m(L,d)$ is a subset of the set of sample points of $A_m(L+1,d)$ for all L. As pointed out in Subsection 3.1, the right choice of the underlying quadrature Q_m implies that the quadratures $(A_m(l,1))_l$ are nested, which again implies—see (4.2)—that the cubatures $(A_m(L,d))_L$ are nested. Although we get for our cubatures, regardless if they are nested or not, the asymptotic estimate $N \leq \mathcal{O}(2^L L^{d-1})$, the hidden constants in the big-$\mathcal{O}$-notation are reasonably smaller if we have nestedness.

The upper bound on N, Theorem 4.4, and some elementary calculations lead to the following corollary.

Corollary 4.7. *Let $d, n \in \mathbb{N}$ and let Q_m be exact on Π_n. For $s > 1/2$ the worst case error of $A_m(L,d)$ satisfies*

$$\operatorname{err}(\mathcal{H}^d_{s,n}, A_m(L,d)) = \mathcal{O}\left(\frac{\log(N_m(L,d))^{(d-1)(s+1/2)}}{(N_m(L,d))^s}\right)$$

Remark 4.8. Recall that H^s_{mix} is continuously embedded in $\mathcal{H}^d_{s,n}$ if $s < n$. In this situation Corollary 4.7 holds in particular for H^s_{mix} in place of $\mathcal{H}^d_{s,n}$.

4.3 Lower bounds for the cubature error

In the previous section we discussed error bounds for our d-dimensional cubature rule based on Smolyak's construction with respect to the spaces $\mathcal{H}^d_{s,n}$ and H^s_{mix}. For the considered spaces $\mathcal{H}^d_{s,n}$ there is a general method to prove lower bounds for the worst case error of any cubature Q_N. In [13] Heinrich, Hickernell and Yue presented a lower bound for Haar wavelet spaces that can be extended to the spaces $\mathcal{H}^d_{s,n}$. (It is not hard to verify that their spaces $\mathcal{H}_{\mathrm{wav},s}$ coincide (for base $b = 2$) with our spaces $\mathcal{H}^d_{s,1}$.) The idea is to construct a finite linear combination f of weighted (multi)wavelet series that is zero on all canonical intervals of a fixed chosen level which contain a sample point of Q_N. This should be done in such a way that the d-dimensional integral $I(f)$ is large while the norm $|f|_{d,s,n}$ should remain small. (Similar proof ideas had been appeared in the mathematical literature before; cf, e.g. the well-known proof of Roth of the lower bound for the L^2-discrepancy [17].)

Theorem 4.9. *Let $s > 1/2$ and $n \in \mathbb{N}$. There exists a constant $C > 0$ such that for any d-dimensional cubature rule Q_N using N sample points we have*

$$\operatorname{err}(\mathcal{H}^d_{s,n}, \mathrm{Q}_N) \geq C \frac{(\log N)^{(d-1)/2}}{N^s}.$$

Proof. Let $P \subset [0,1]^d$, $|P| = N$ be the set of sample points used by the cubature rule Q_N. For all $\mathbf{l} \in \mathbb{N}_0^d$ we define a function

$$f_{\mathbf{l}}(x) = \begin{cases} 1 & \text{for all } x \in I^{\mathbf{l}}_{\mathbf{k}},\ \mathbf{k} \in \nabla_{\mathbf{l}} \quad \text{with} \quad I^{\mathbf{l}}_{\mathbf{k}} \cap P = \emptyset \\ 0 & \text{else} \end{cases}.$$

Now we choose the uniquely determined integer L that satisfies

$$2^{L-1} < 2N \leq 2^L$$

and define a function

$$f = \sum_{|\mathbf{l}|=L} f_{\mathbf{l}}.$$

Hence we get for the norm of our candidate

$$\begin{aligned} |f|^2_{d,s,n} &= \sum_{\mathbf{j}\geq -1} \sum_{\mathbf{k}\in\nabla_{\mathbf{j}}} \sum_{|\mathbf{i}|_\infty=0}^{n-1} 2^{|\mathbf{j}|2s} \langle f, \Psi^{\mathbf{j}}_{\mathbf{i},\mathbf{k}}\rangle^2 \\ &= \sum_{|\mathbf{l}|=|\mathbf{l}'|=L} \sum_{\mathbf{j}\geq -1} \sum_{\mathbf{k}\in\nabla_{\mathbf{j}}} \sum_{|\mathbf{i}|_\infty=0}^{n-1} 2^{|\mathbf{j}|2s} \langle f_{\mathbf{l}}, \Psi^{\mathbf{j}}_{\mathbf{i},\mathbf{k}}\rangle \langle f_{\mathbf{l}'}, \Psi^{\mathbf{j}}_{\mathbf{i},\mathbf{k}}\rangle. \end{aligned}$$

Due to (2.1) the inner product $\langle f_{\mathbf{l}}, \Psi^{\mathbf{j}}_{\mathbf{i},\mathbf{k}}\rangle$ vanishes if one of the indices j_ν satisfies $j_\nu \geq l_\nu \geq 0$. Furthermore, if we put $M := \left(\max_{i=0}^{n-1}\{\|\varphi_i\|_\infty, \|\psi_i\|_\infty\}\right)^d$, we have

$$\left|\langle f_{\mathbf{l}}, \Psi^{\mathbf{j}}_{\mathbf{i},\mathbf{k}}\rangle\right| \leq \left\|\Psi^{\mathbf{j}}_{\mathbf{i},\mathbf{k}}\right\|_\infty \|f_{\mathbf{l}}\|_\infty \operatorname{vol}(I^{\mathbf{j}}_{\mathbf{k}}) \leq M|\nabla_{\mathbf{j}}|^{-1/2}.$$

Therefore we get

$$\begin{aligned} |f|^2_{d,s,n} &\leq n^d M^2 \sum_{|\mathbf{l}|=|\mathbf{l}'|=L} \sum_{-1\leq\mathbf{j}<\mathbf{l},\mathbf{l}'} \sum_{\mathbf{k}\in\nabla_{\mathbf{j}}} 2^{|\mathbf{j}|2s} |\nabla_{\mathbf{j}}|^{-1} \\ &\leq n^d M^2 \sum_{|\mathbf{l}|=|\mathbf{l}'|=L} \sum_{-1\leq\mathbf{j}<\mathbf{l},\mathbf{l}'} 2^{|\mathbf{j}|2s} \\ &\leq n^d M^2 \sum_{|\mathbf{l}|=|\mathbf{l}'|=L} \sum_{\nu=0}^{d} 2^{-2s\nu} \binom{d}{\nu} \sum_{0\leq\mathbf{j}<\mathbf{l},\mathbf{l}'} 2^{|\mathbf{j}|2s} \\ &\leq n^d M^2 (1+2^{-2s})^d \sum_{|\mathbf{l}|=|\mathbf{l}'|=L} \sum_{0\leq\mathbf{j}<\mathbf{l},\mathbf{l}'} 2^{|\mathbf{j}|2s} \\ &\leq n^d M^2 (1+2^{-2s})^d \sum_{\nu=0}^{L-d} \sum_{|\mathbf{j}|=\nu,\mathbf{j}\geq 0} 2^{2\nu s} \left(\sum_{|\mathbf{l}|=L,\mathbf{l}>\mathbf{j}} 1\right)^2 \\ &= n^d M^2 (1+2^{-2s})^d \sum_{\nu=0}^{L-d} \binom{\nu+d-1}{d-1} 2^{2\nu s} \binom{L-\nu-1}{d-1}^2. \end{aligned}$$

We upper-bound $\binom{\nu+d-1}{d-1}2^{2\nu s}$ by $\binom{L-1}{d-1}2^{2(L-d)s}$. Furthermore, we use the new index $m := L-d-\nu$ and majorize the resulting sum by taking the infinite sum instead. Using the short hand $C' := n^d M^2 (1+2^{-2s})^d$ leads to

$$|f|^2_{d,s,n} \leq C' \left(\sum_{m=0}^{\infty} 2^{-m2s} \binom{m+d-1}{d-1}^2\right) \binom{L-1}{d-1} 2^{2(L-d)s} \leq C'' \binom{L-1}{d-1} 2^{2(L-d)s},$$

with a constant C'' not depending on L, but on d and s. Furthermore, we have

$$\int_{[0,1)^d} f dx \;=\; \sum_{|\mathbf{l}|=L} \int_{[0,1)^d} f_{\mathbf{l}} dx \geq \sum_{|\mathbf{l}|=L} 2^{-L}(2^L - N) \geq \sum_{|\mathbf{l}|=L} \frac{1}{2} = \frac{1}{2}\binom{L+d-1}{d-1}.$$

Let us now consider the function $f^* = f/\left|f\right|_{d,s,n}^d$. Since $Q_N(f) = 0$ the estimates above result in

$$\mathrm{err}(f^*, \mathrm{Q}_N) = \frac{\left|\int_{[0,1)^d} f dx\right|}{|f|_{d,s,n}^d} \geq \frac{2^{ds-1}}{\sqrt{C''}} \frac{\binom{L+d-1}{d-1}}{\sqrt{\binom{L-1}{d-1}}} 2^{-Ls}\,.$$

Using the asymptotic estimates

$$\binom{L+d-1}{d-1} \sim L^{d-1} \quad \text{and} \quad \sqrt{\binom{L-1}{d-1}} \sim L^{(d-1)/2}\,,$$

we finally get $\mathrm{err}(f^*, \mathrm{Q}_N) \geq C 2^{-Ls} L^{(d-1)/2}$, with a constant C not depending on L, but depending on d and s. □

5 Numerical examples

We implemented our cubature method and computed the integrals of certain test functions in dimension 5 and 10. The families of test functions we considered were selected from the testing package of Genz [9, 10], and they are named as follows:

$$(1)\quad \text{OSCILLATORY} \qquad f_1(x) = \cos\left(2\pi w_1 + \sum_{i=1}^{d} c_i x_i\right),$$

$$(2)\quad \text{PRODUCT PEAK} \qquad f_2(x) = \prod_{i=1}^{d} \left(c_i^{-2} + (x_i - w_i)^2\right)^{-1},$$

$$(3)\quad \text{CORNER PEAK} \qquad f_3(x) = \left(1 + \sum_{i=1}^{d} c_i x_i\right)^{-(d+1)},$$

$$(4)\quad \text{GAUSSIAN} \qquad f_4(x) = \exp\left(-\sum_{i=1}^{d} c_i^2 (x_i - w_i)^2\right),$$

$$(5)\quad \text{CONTINUOUS} \qquad f_5(x) = \exp\left(-\sum_{i=1}^{d} c_i \left|x_i - w_i\right|\right),$$

$$(6)\quad \text{DISCONTINUOUS} \qquad f_6(x) = \begin{cases} 0, & \text{if } x_1 > w_1 \text{ or } x_2 > w_2 \\ \exp\left(\sum_{i=1}^{d} c_i x_i\right), & \text{otherwise} \end{cases}.$$

This choice of test functions is obviously unfavorable with regard to our cubature rule and the corresponding function classes, but enables us to compare our results directly to the results of the algorithms studied in [14] and [19]. The algorithm in [14] is based on Smolyak's construction and the Clenshaw-Curtis rule in dimension $d = 1$. The algorithms in [19, Chapter 11] consist of an embedded sequence of lattice rules named COPY, an algorithm using rank-1 lattice rules, an adaptive Monte Carlo method, and an adaptive method by van Dooren and De Ridder [22], for which the short hand ADAPT is used. With respect to the six test families, COPY and ADAPT are the best performing algorithms of these four.

We should mention here that there exist more recent algorithms which improve in some applications on the algorithms we chose as benchmark methods, see, e.g., [4, 12, 16] and the literature mentioned therein. So, e.g., the use of Kronrod-Patterson formulas as one-dimensional quadratures in Smolyak's construction seems to be a very powerful tool for the treatment of smooth functions as reported in [11, 16]. These cubatures have the advantage to lead to a higher degree of polynomial exactness than the Smolyak construction of the same level based on the Clenshaw-Curtis rules, while on the other hand the number of sample points used increases faster. Although the use of Kronrod-Patterson formulas leads for some examples to reasonably better performance than the use of Clenshaw-Curtis rules, one can see in [16] that for the testing package of Genz the first were not clearly better than the latter. As observed in [11, 16] the numerical advantage of the algorithms based on the Kronrod-Patterson rules over the ones based on

Clenshaw-Curtis rules decreases with growing dimension d. The improvements of Petras by using delayed basis sequences of Kronrod-Patterson formulas can be seen as a further "fine tuning" of Smolyak's algorithm for smooth functions. The use of delayed basis sequences may also help to improve our approach for classes of smooth functions. We have not studied this so far.

We followed the conventions from [14, 19]: All the functions were normalized so that the true integrals over the unit cube equaled 1. By varying the parameters $\mathbf{c} = (c_1, \ldots, c_d)$ and $\mathbf{w} = (w_1, \ldots, w_d)$ we got different test integrals. For each family of functions we performed 20 tests in which we chose the vectors independently and uniformly distributed in $[0, 1]^d$. The vectors $\mathbf{c}$ were renormalized such that

$$\sum_{i=1}^{d} c_i = b_j$$

holds for predetermined parameters b_j, $j = 1, \ldots, 6$. Since, in general, the difficulty of the integrals increases as the (Euclidean) norm $\|\mathbf{c}\|$ increases, the choice of the b_j determines the level of difficulty. As in [19] and in [14], we chose in dimension $d = 10$ the following values of b_j:

j	1	2	3	4	5	6
b_j	9.0	7.25	1.85	7.03	20.4	4.3

.

In the notion of [19] this corresponds to the level of difficulty $L = 1$ for the families 2, 4, and 6, and to the level $L = 2$ for the families 1, 3, and 5. In dimension $d = 5$ we chose $b_2 = 29$ and $b_5 = 43.4$, which corresponds to the level $L = 1$ for family 2 and $L = 2$ for family 5.

The diagrams in Figure 2 to 7 show the median of the absolute error of our cubatures in 20 tests for each of the considered families. We treated all six families in dimension 10. In Figure 2 to 7 we plotted the median error of the lattice rule COPY taken from the diagrams in [19] and the median error of the algorithm considered by Novak and Ritter taken from the diagrams in [14].

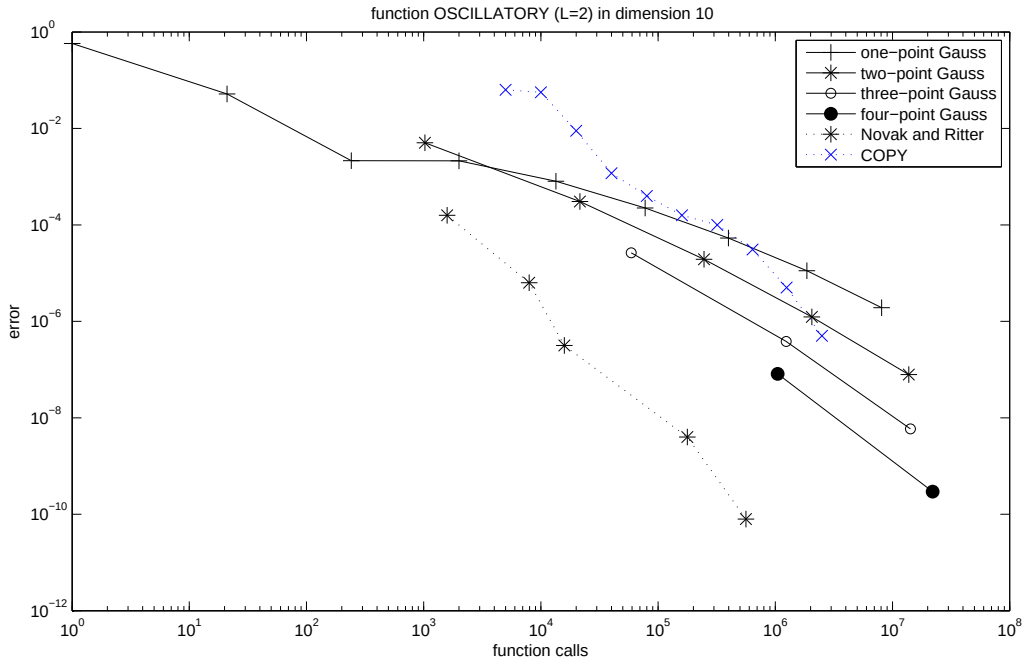

Figure 2: Median of absolute error of family (1), 20 integrands.

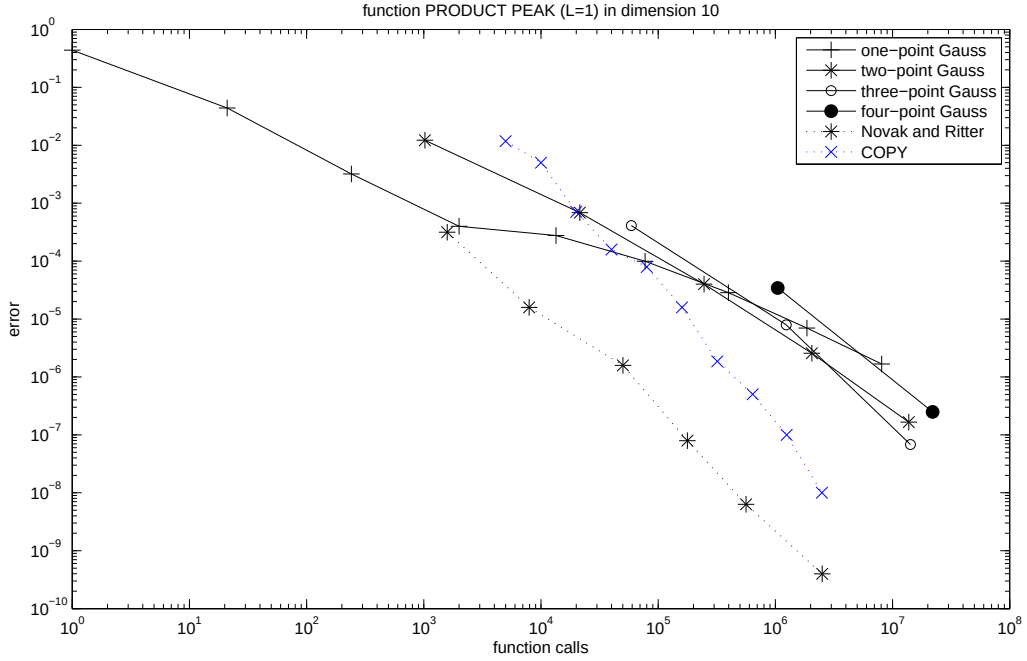

Figure 3: Median of absolute error of family (2), 20 integrands.

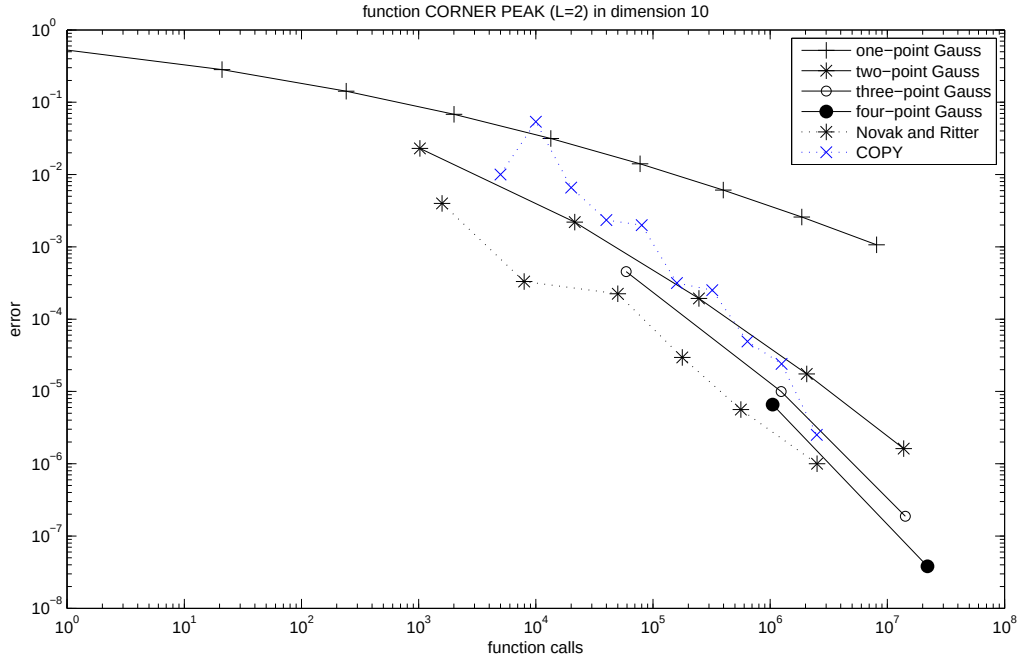

Figure 4: Median of absolute error of family (3), 20 integrands.

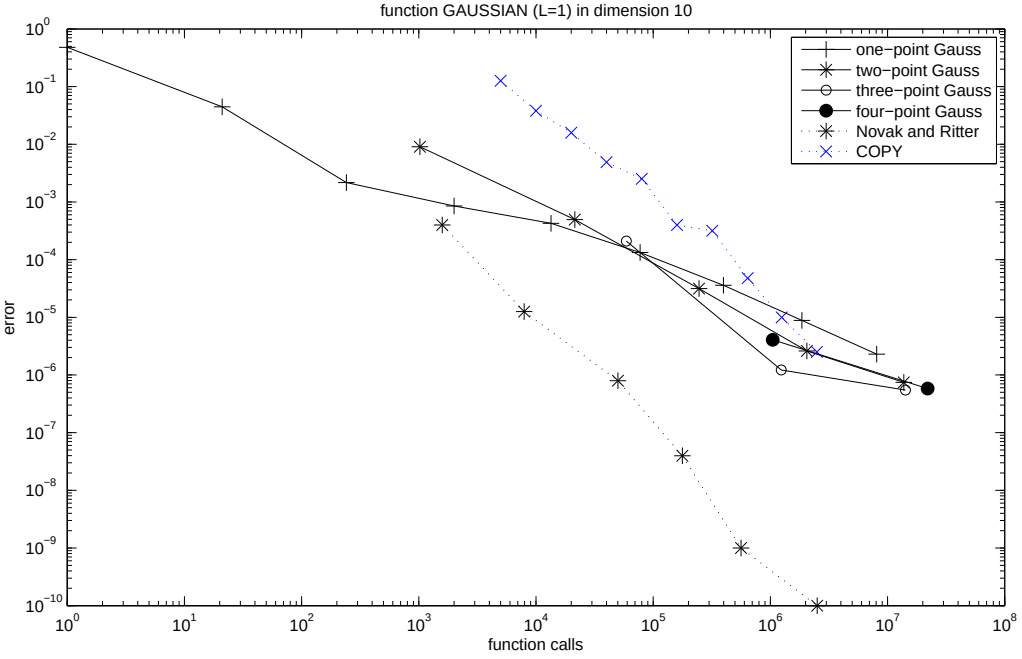

Figure 5: Median of absolute error of family (4), 20 integrands.

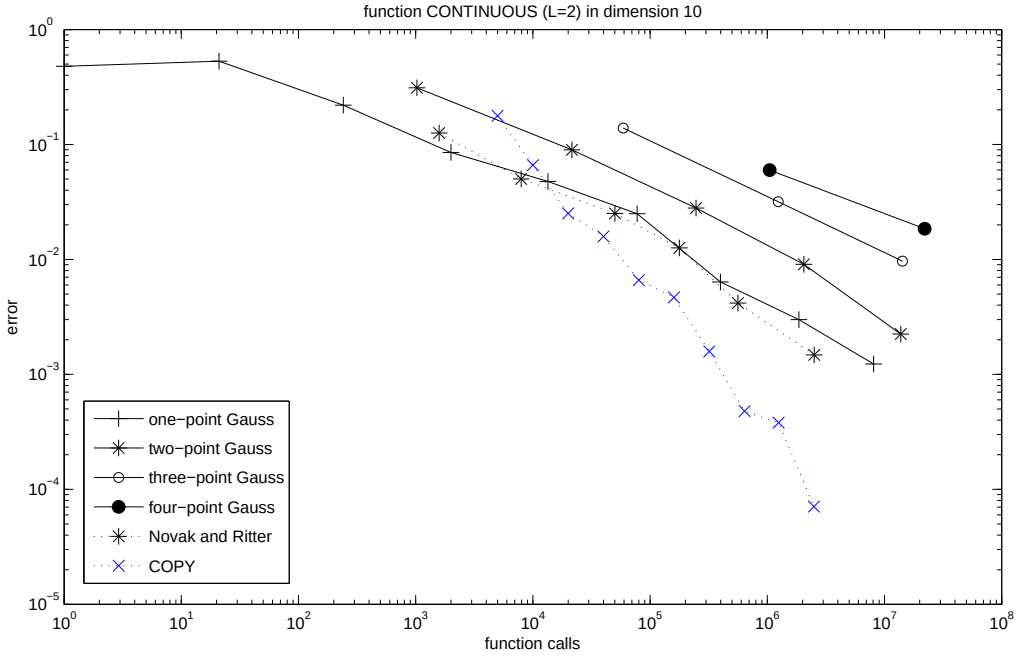

Figure 6: Median of absolute error of family (5), 20 integrands.

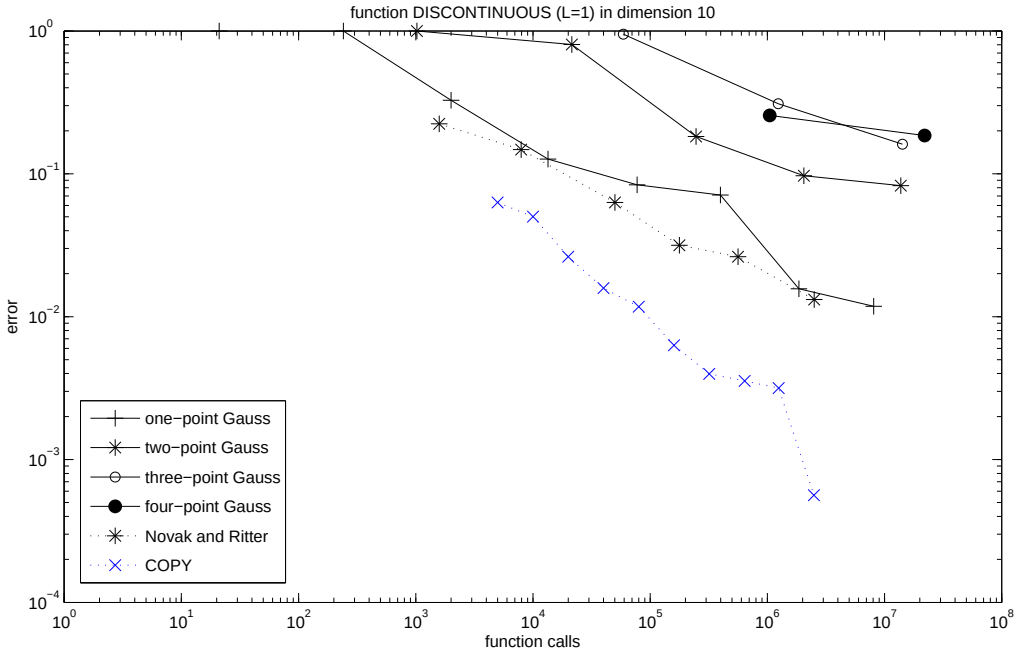

Figure 7: Median of absolute error of family (6), 20 integrands.

We tested our method by using Gauss rules as underlying one-dimensional quadrature Q_m. Notice that in the case of the one-point Gauss rule the resulting algorithm is identical to the so-called Boolean midpoint rule which has, e.g., been studied in [3]. (There in a numerical example this algorithm behaved well for a smooth test function lying in some Korobov space.) We decided to use Gauss rules since they achieve the maximal degree of polynomial exactness. This makes it easy to study the dependence of the numerical results on the degree of (piecewise) polynomial exactness by considering only few sample points in the underlying quadrature. Unfortunately, the choice of Gauss rules leads to non-nested cubatures $(A_m(L,d))_L$. As mentioned in Section 3.1 and 4.2 one may use nested one-dimensional quadratures to reduce the number of sample points in dimension d at a given level L (of course at the cost of a reduced parameter n for fixed m).

For smooth integrands one would in general expect Gauss rules Q_m with larger m superior to Gauss rules with smaller m, while for non-smooth integrands one would expect the contrary behavior. These prediction is supported by the numerical results for the families 1, 3, 5, and 6. The results for family 2 and 4 however do not display such a clear tendency.

If we compare our results to the ones of the algorithm of Novak and Ritter, we see that for the families 1, 2, and 4 their results are clearly better than ours, while for the families 3, 5, and 6 the results are comparable. The results for the families 1, 2, and 4 reflect that the algorithm of Novak and Ritter was constructed to make the best use of smoothness properties, while our method was not.

If we compare our cubature method with the algorithms considered in [19], it turns out that for the families 1, 3, and 4 our method is comparable to ADAPT and the two lattice rules. The adaptive Monte Carlo method is in non of these cases competitive. In case of family 2 our cubature is not as good as COPY, but comparable with the rank-1 lattice rule and ADAPT and better than the Monte Carlo method. For family 5 our method is comparable to ADAPT, but worse than Monte Carlo and both lattice rules. Our results for family 6 however are not as good as the results of any of the four algorithms in [19].

We performed additional numerical tests. Here we considered the six dimensional function

$$f_\alpha(x,y) = \|(x+\alpha)-y\| = \left(((x_1+\alpha_1)-y_1)^2 + ((x_2+\alpha_2)-y_2)^2 + ((x_3+\alpha_3)-y_3)^2\right)^{1/2},$$

where α is a random vector. This function is used as a typical prototype electron-electron cusp of the solution of the electronic Schrödinger equation, see, e.g. [8].

Since we do not know the exact value of the integral of f_α over $[0,1]^6$, we consider the following (normalized) error for a given level L:

$$\mathrm{err}(L,\alpha) := \frac{A_m(L,6)f_\alpha - A_m(L-1,6)f_\alpha}{A_m(L,6)f_\alpha}\,. \tag{5.1}$$

We performed 20 tests in which we chose the vectors α independently and uniformly distributed in $[0,1]^6$. Figure 8 shows the median of the error defined by (5.1) for the cubature $A_m(L,6)$ with underlying one-, two-, and three-point Gauss formulas. The

cubature induced by the one-point Gauss formula shows a good convergence up to an error of 10^{-4}. For a precision beyond 10^{-4} we seem to have numerical instabilities, as is also confirmed by the behavior of the cubatures induced by the two- and three-point Gauss formulas. Here adaptiveness may help to overcome these problems.

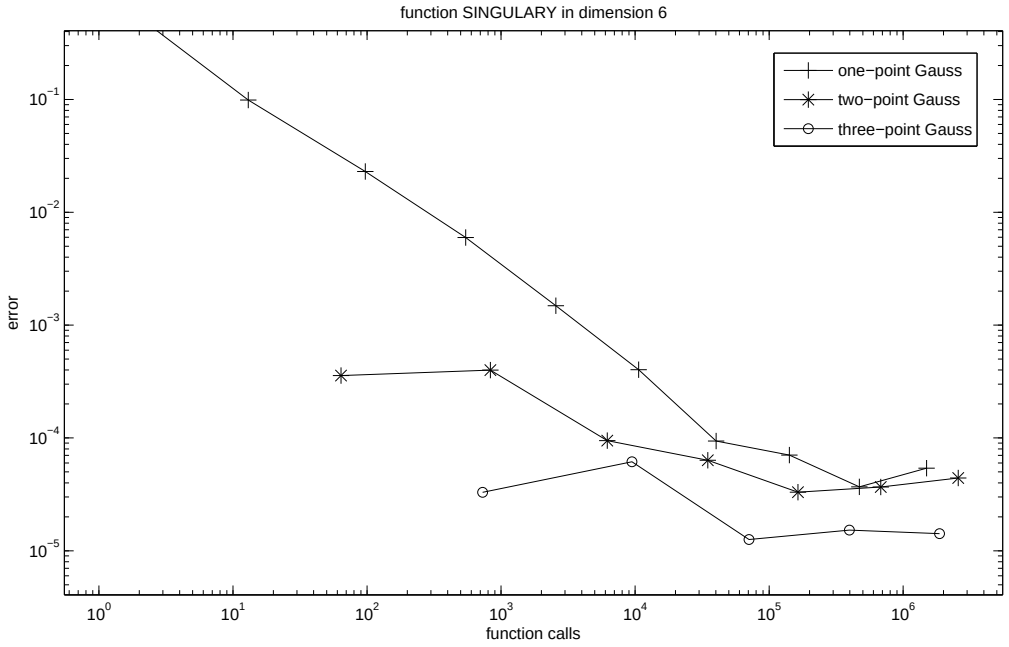

Figure 8: Median of the error defined in (5.1) of 20 integrands.

Altogether, the numerical experiments show that our algorithms induced by Gauss formulas do behave reasonably well, although in every comparison (except for family (3)) the algorithm of Novak and Ritter or COPY perform clearly better. We may create algorithms following our approach by using, e.g., underlying quadratures resulting in nested cubatures or delayed basis sequences, cf. [16]. This may lead to an improved numerical performance. We have not studied this in detail so far, since the focus of our work was on the theoretical aspects of our multiwavelet approach.

6 Conclusion and Outlook

We provide explicit algorithms for multivariate integration based on Smolyak's construction and iterated one-dimensional quadrature rules. These quadrature rules can be arbitrary as long as they satisfy a given degree of polynomial exactness. The resulting algorithms are simple and easy to implement. We consider certain multiwavelet function spaces and derive upper and lower bounds for the worst case error over the unit ball of these spaces. These bounds reveal that our algorithms are optimal up to logarithmic factors.

We have chosen the presented multiwavelet approach, because translates of scaling functions do not overlap here. The treatment of overlapping scaling functions like Daubechies functions, B-splines or smoother multiwavelets requires special care at the end points of intervals [6]. These wavelet functions can be treated with an approach based on frame concepts; this will be discussed in detail in a forthcoming paper.

The function spaces we consider are Hilbert spaces of functions whose coefficients with respect to multiwavelet expansions exhibit a prescribed decay. These spaces are related to Sobolev spaces. A natural question is whether one can combine our multiwavelet approach with other function spaces. One may consider spaces with discrete norms related to Besov spaces (e.g., by using different weights and an l_p-metric rather than an l_2-metric in (3.3)). Now Besov spaces are usually useful for adaptive non-linear approximation. It is known that for our error criterion adaptiveness and non-linearity does essentially not help (see, e.g., [21]). In our opinion it is more promising to consider weighted (tensor product) Sobolev spaces. This would require some modifications with respect to the choice of cubature points. We have not studied this approach in detail so far, but think that our multiwavelet approach seems to be the appropriate choice here and would lead to good results.

Another interesting question is if it is possible to achieve a better numerical performance than we have seen in Section 5 with algorithms following our approach. This should be the case if we consider functions which are more favorable with regard to our theoretical analysis. A way of improving our algorithms for sufficiently smooth integrands may be to use delayed basis sequences which result in nested cubatures.

Acknowledgment

We would like to thank two anonymous referees for their comments which helped to improve the representation reasonably.
Furthermore, the first author would like to thank Henryk Woźniakowski for the warm hospitality and for interesting discussions during his stay at the University of Warsaw in March 2005.

References

[1] M. Abramowitz, I. A. Stegun, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, Dover, New York, 1964.

[2] B. Alpert, A class of bases in L^2 for the sparse representation of integral operators, SIAM J. Math. Anal. 24 (1993) 246–262.

[3] G. Baszenski, F. J. Delvos, Multivariate Boolean midpoint rules, in: Numerical Integration IV, Birkhäuser, Basel, 1–11, 1993.

[4] H. J. Bungartz, M. Griebel, Sparse grids, Acta Numerica 13 (2004) 147–269.

[5] A. Cohen, Numerical Analysis of Wavelet Methods, North-Holland, Amsterdam, 2003.

[6] W. Dahmen, B. Han, R.-Q. Jia, A. Kunoth, Biorthogonal multiwavelets on the interval: cubic Hermite splines, Constr. Approx. 16 (2000) 221–259.

[7] P. J. Davis, P. Rabinowitz, Methods of Numerical Integration, 2nd edition, Academic Press, New York, 1984.

[8] H. J. Flad, W. Hackbusch, R. Schneider, Best N-term approximation in electronic structure calculations. II. Jastrow factors, Max Planck Institute for Mathematics in the Sciences, Leipzig, Preprint 80/2005, to appear in M2AN Math. Model. Numer. Anal..

[9] A. Genz, Testing multidimensional integration routines, in: B. Ford, J. C. Rault, F. Thomasset, eds., Tools, methods and languages for scientific and engineering computation, 81–94, North-Holland, Amsterdam, 1984.

[10] A. Genz, A package for testing multiple integration subroutines, in: P. Keast, G. Fairweather, eds., Numerical Integration, Kluwer, Dordrecht, 337-340, 1987.

[11] T. Gerstner, M. Griebel, Numerical integration using sparse grids, Numerical Algorithms 18 (1998) 209–232.

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

[13] S. Heinrich, F. J. Hickernell, R.-X. Yue, Optimal quadrature for Haar wavelet spaces, Math. Comput. 73 (2004) 259–277.

[14] E. Novak, K. Ritter, High-dimensional integration of smooth functions over cubes, Numer. Math. 75 (1996) 79–97.

[15] A. B. Owen, Monte Carlo variance of scrambled net quadrature, SIAM J. Numer. Anal. 34 (1997) 1884–1910.

[16] K. Petras, Smolyak cubature of given polynomial degree with few nodes for increasing dimension, Numer. Math. 93 (2003) 729–753.

[17] K. F. Roth, On irregularities of distribution, Mathematika 1 (1954) 73–79.

[18] R. Schneider, Multiskalen- und Wavelet-Matrixkompression. (German) [Multiscale and wavelet matrix compression] Analysisbasierte Methoden zur effizienten Loesung grosser vollbesetzter Gleichungssysteme. [Analysis-based methods for the efficient solution of large nonsparse systems of equations], Advances in Numerical Mathematics, B. G. Teubner, Stuttgart, 1998.

[19] I. H. Sloan, S. Joe, Lattice Methods for Multiple Integration, Oxford University Press, New York, 1994.

[20] S. A. Smolyak, Quadrature and interpolation formulas for tensor products of certain classes of functions, Soviet Math. Dokl. 4 (1963) 240–243.

[21] J. F. Traub, G. W. Wasilkowski, H. Woźniakowski, Information-Based Complexity, Academic Press, New York, 1988.

[22] P. van Dooren, L. de Ridder, An adaptive algorithm for numerical integration over an n-dimensional cube, J. Comp. Appl. Math. 2 (1976) 207–217.

[23] T. von Petersdorff, C. Schwab, R. Schneider, Multiwavelets for second kind integral equations, SIAM J. Numer. Anal. 34 (1997) 2212–2227.

[24] G. W. Wasilkowski, H. Woźniakowski, Explicit cost bounds of algorithms for multivariate tensor product problems, J. Complexity 11 (1995) 1–56.