Cubature Formulas on Wavelet Spaces

Dissertation

zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften

(Dr. rer. nat.)

der Technischen Fakultät

der Christian-Albrechts-Universität zu Kiel

vorgelegt von

Dipl. Math. René Lindloh

Kiel

2007

1. Gutachter: Prof. Dr. Anand Srivastav

2. Gutachter: Prof. Dr. Reinhold Schneider

Datum der mündlichen Prüfung: 19. April 2007

How strong is your dream?

Preface

The word “wavelet” was introduced by Morlet and Grossman in the early 1980s. Wavelets are the result of extensive interdisciplinary research of, e.g. approximation theorists, harmonic and functional analysts, mathematical physicists. The “wavelet methods” or “wavelet techniques” are very popular in both applied areas and more theoretical ones. This thesis focuses on “wavelet methods” for numerical integration and in particular for high-dimensional integration.

The first step is to analyze given cubature rules via wavelet bases in appropriate function spaces from a combinatorial point of view. This idea is based on the observation that error bounds of Koksma Hlawka type are often related to so-called reproducing kernel Hilbert spaces which can be described via wavelet bases. For Haar wavelet spaces this yields a somehow natural measure for the quality of cubature rules which can be termed “geometric discrepancy” in the classical sense. Thus the cubature error is divided into two parts: a problem dependent one (smoothness of the function spaces) and a second part that only depends on the set of sample points used by the cubature rule (geometric discrepancy).

Using cubature rules based on so-called “low-discrepancy point sets” guarantee good error bounds. Moreover, also for tensor product rules (e.g. sparse grids) these methods allow a detailed error analysis of the given cubature rule in a relative elementary way.

The second step is the construction of almost optimal cubature rules via wavelets. We use the pleasant fact that expressing error bounds via wavelet bases often reveals the structure of good cubature rules. The analysis starts with the one-dimensional case, but since we are mainly interested in the computation of high-dimensional integrals, we expand the one-dimensional wavelet based quadrature rules in a natural way using a well known clever chosen tensor product approach. Note that it is not the aim of this thesis to “reinvent” the well known sparse grid method.

Acknowledgments

I would like to thank Prof. Dr. Anand Srivastav for drawing my interest on the interplay of geometric discrepancy and numerical integration and for supervising this thesis. I am also grateful to Prof. Dr. Reinhold Schneider for fruitful discussions and introducing me to wavelet analysis.

I am thankful to the Deutsche Forschungsgemeinschaft (DFG) for their partial financial support through the graduate school 357 “Effiziente Algorithmen und Mehrskalenmethoden”.

Especially, I would like to thank my wife Dorit for her understanding, patience and support.

Kiel, February 2007

René Lindloh

Contents

Chapter 1

Introduction

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, because the number of function calls used increases exponentially with the dimension. Since the cost of a cubature rule is essentially proportional to the number of function evaluations, the complexity of the problem grows exponentially with the dimension. This observation is known as the "curse of dimensionality", introduced by Bellman, see [8]. There is a long list of interesting methods that try to overcome the "curse of dimensionality" such as Monte Carlo and quasi-Monte Carlo methods [47], lattice rules [63], sparse grid methods [10, 34, 35], etc.

In this thesis we present wavelet based cubature methods which can be used to handle the following multivariate integration problem also in higher dimensions: we want to approximate the integral

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

of functions $f : [0,1)^d \to \mathbb{R}$ belonging to some function classes of theoretical or practical interest. From the view point of applicability of high-dimensional cubature it is most important 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 H defined by

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

where

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

As I and Q_N are linear, $\mathrm{err}(H, \mathrm{Q}_N)$ is nothing but the operator norm $\|I - Q_N\|_{\mathrm{op}}$ induced by the norm of H. Observe that for arbitrary $f \in H$ and not necessarily from the unit ball of H, we have

$$|I(f) - \mathrm{Q}_N(f)| \leq \|f\| \, \|I - \mathrm{Q}_N\|_{\mathrm{op}} \, .$$

Error bounds of this type are related to so-called Koksma-Hlawka inequalities which are typically of the form

$$\mathrm{err}(f, \mathrm{Q}_N) \leq D_\lambda(\{x_1, x_2, \ldots, x_N\}) V(f),$$

where $V(f)$ is a measure of the variation of the integrand, and the first term $D_\lambda(\{x_1, x_2, \ldots, x_N\})$ is a measure of the non-uniformity of the set of sample points, with respect to the given weights $\lambda_1, \ldots, \lambda_N$. This term is often called the (weighted) geometric discrepancy of $x_1, x_2, \ldots, x_N$. Such error bounds are for instance studied in [26, 38], where the main mathematical tool is the theory of reproducing kernel Hilbert spaces.

The function classes we consider here are certain Hilbert spaces H characterized via discrete norms defined by the asymptotic rate of decay of the wavelet coefficients. These function spaces are continuously embedded in L^2, and under proper requirements, H contains classical function spaces like tensor products of Sobolev spaces.

Our aim is to provide cubature methods that guarantee a (nearly) optimal worst case error and which are sufficiently easy to implement. The cubature rules we present here are based on one-dimensional quadratures, chosen with respect to the particular space H under consideration, and Smolyak's construction. The present Smolyak construction corresponds to tensor product wavelet expansions such that the cubature is exact on finite wavelet series up to a critical level. A comparable approach can be found in [37]. Here the considered function classes depend on Haar wavelet series and the cubature is, instead of the method we prefer, given by a quasi-Monte Carlo method based on so-called scrambled net, see, e.g. [51].

We focus on Smolyak's construction, which is also known as "Birmann interpolation", "Boolean method", "discrete blending method", "hyperbolic cross points", and "sparse grid method".

Even in the most general case it is known that this construction leads to almost optimal approximations in any dimension $d > 1$ as long as the underlying one-dimensional approximation is optimal, see, e.g. [67]. Therefore it is not surprising that the application of Smolyak's construction to numerical integration has been studied extensively, see, e.g. [34, 48, 54, 65, 67] and the literature mentioned therein. The problems studied in these papers are usually defined on spaces of functions with bounded mixed derivatives, i.e., on spaces of functions with a certain degree of smoothness. The central theme of this thesis is to provide good error bounds with respect to Hilbert spaces H of not necessarily smooth functions.

This thesis is organized as follows: In Chapter 2 we focus on the interrelation between reproducing kernel Hilbert spaces, frames, and a classical measure of non-uniformity called "geometric discrepancy". After recalling the definition and the crucial properties of reproducing kernel Hilbert spaces, we give a short discussion on the interaction of those kernels with tight error bounds for cubature formulas. This leads to so-called Koksma Hlawka type inequalities and a notion of generalized geometric discrepancy, see, e.g. [38]. The important question, how to find interesting function spaces to generalize such error bounds, can be answered by a closer look at the bases of suitable function spaces.

Meschkowski [44] showed the interaction of reproducing kernel Hilbert spaces H and orthonormal systems on H. By a generalization of this result we are able to establish the following expression for the cubature error under suitable conditions on a frame $\{\psi_\lambda\}_\lambda$ on H,

$$\mathrm{err}(H, Q) \le \left\| \sum_\lambda \tilde{\psi}_\lambda(\cdot) \left\{ \int \psi_\lambda(t)dt - Q[\psi_\lambda] \right\} \right\|_H ,$$

where $\tilde{\psi}_\lambda$ denotes the dual frame.

In Chapter 3 we describe the idea of a wavelet based error analysis for quadrature respectively cubature formulas. After giving a brief review of multiresolution analysis we point out the interplay between discrete norms defined by weighted wavelet series $\|f\|_s^2 = \sum_\lambda 2^{|\lambda|2s} |\langle f, \psi_\lambda \rangle|^2$, and some classical notion of smoothness, see, e.g. [15, 61]. Scaled wavelet spaces, e.g. $H_s := \{f \in \mathrm{L}^2 : \|f\|_s < \infty\}$, are for a suitable smoothness parameter $s > 1/2$ in line with the generalization of the result of Meschkowski we proved in Chapter 2. By the fact that the considered wavelets have vanishing moments, i.e. $\langle \psi_\lambda, x^\alpha \rangle = 0$ for a certain α, the error bounds in Chapter 2 have the concise form

$$\mathrm{err}(H_s, Q)^2 \le \sum_\lambda 2^{-|\lambda|2s} \{-Q[\psi_\lambda]\}^2 .$$

We take a closer look at the Haar wavelet case, where an optimal quadrature or respectively two-dimensional cubature is given in a intuitive way by a quasi-Monte Carlo rule using so-called binary nets $P_{net} \subset [0,1]^2$ as sample points, see, e.g. [43, 47]. For two-dimensional Haar wavelet spaces with smoothness parameter $s > 1/2$, we get for cubature rules based on binary nets, $Q_{net,N} = \frac{1}{|P_{net}|} \sum_{p \in P_{net}} f(p)$ and $|P_{net}| = N$,

$$\mathrm{err}(H_s, Q_{net}) \sim \frac{\log(N)^{1/2}}{N^s}.$$

By $a \sim b$ we always mean that $a \lesssim b$ and $b \lesssim a$ holds, where $a \lesssim b$ says that a can be bounded by a constant multiple of b. Note that here and in the sequel the constant as well as the constant in the $\mathcal{O}(\cdot)$ notation may have dependence on the dimension and the smoothness parameter s. The error analysis for the Haar wavelet case is strongly related to the classical concept of geometric discrepancy.

In Chapter 4 we recall the main ideas of Smolyak's construction that overcomes the "curse of dimensionality". This method differs from the previous ideas by using only subsets of the full tensor product formula to find an "optimal" cubature formula. For sake of completeness we recall a justification, based upon combinatorial arguments for the choice of the special index set that defines the construction, see [10]. If we consider smooth functions it is known that Smolyak's construction, e.g. based on Clenshaw-Curtis rule, is optimal up to logarithmic factors, see [48]. We discuss the difference between the optimal cubature for the two-dimensional Haar wavelets from Chapter 3 and the cubature we get for the Haar wavelet cases by Smolyak's construction. Afterwords we go into a more detailed discussion on the complexity of these algorithms and slight modifications.

Chapter 5 contains one of the main results of this thesis. We construct simple algorithms for high-dimensional numerical integration of function classes that consist of square integrable functions over the d-dimensional unit cube whose coefficients with respect to certain multiwavelet expansions decay rapidly. After a short construction of discontinuous multiwavelet bases we extend our optimal one-dimensional quadrature via Smolyak's construction to a d-dimensional cubature. More precisely we use composite quadrature rules of a fixed order n. These rules are exact for piecewise polynomials of order n. The current Smolyak construction is related to a tensor product multiwavelet expansion since the cubature $A(L, d)$ is exact on finite multiwavelet series up to a critical level L. For these multiwavelet spaces H_s^{multi}, we get for a smoothness parameter $s > 1/2$,

$$\operatorname{err}(H_s^{\text{multi}}, A(L, d)) \lesssim \frac{(L + d - 1)^{d-1}}{2^{Ls}}.$$

Furthermore we give a lower bound for the cubature error on multiwavelet spaces H_s^{multi}. For $s > 1/2$ we get for any cubature rule Q_N that uses N sample points,

$$\frac{\log(N)^{(d-1)/2}}{N^s} \lesssim \operatorname{err}(H_s^{\text{multi}}, Q_N).$$

From the abstract definition of our function space H_s^{multi} it is not immediately clear if it contains a reasonable class of interesting functions apart 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 H_s^{multi}. We implement our cubature methods and report on several numerical tests which allow comparison to established methods.

In Chapter 6 we give a generalization of the multiwavelet based cubature, we discuss so-called wavelet frames and an anisotropic version of Smolyak's construction. The aim of this chapter is to resolve the dilemma which is indicated in the chapters before. On the one hand side we are interested in wavelets that guarantee an adapted approximation order, but on the other hand this requirement leads to difficulties on the boundary of the considered domain. To evade these problems we prefer a formulation via wavelet frames, see [11]. We define a general approximation of the integral operator on function spaces characterized by discrete norms as before. To get norm equivalence to classical function spaces additional effort using so-called approximate duals is needed.

We apply the optimal approximation method to spline wavelets, and it turns out that for B-splines the general approximation is a classical quadrature rule, i.e. defined by point evaluations. The d-dimensional approximation is based on an anisotropic version of Smolyak's construction. In more detail, we allow a weighting of the directions of the cubature rule. Sometimes this is called adaptivity in the sense of a-priori knowledge, see, e.g. [34, 53]. This idea can be used to reduce the overall cost of the algorithm, if there is knowledge about the dependence on smoothness and directions. The error analysis gets more complicated in this case, and we have a larger gap between lower and upper bounds. This additional expense is justified by the norm equivalence to classical functions spaces with higher regularities. We also implement one possible spline cubature and report on several numerical tests which allow comparison to established methods.

At the end of this thesis we give a short conclusion and report on some interesting and still open problems.

Chapter 2

Frames and Reproducing Kernel Hilbert Spaces

In this chapter we focus on aspects of frame theory and the theory of reproducing kernel Hilbert spaces which are useful for our later purpose to construct wavelet based cubature rules. A reproducing Hilbert space (RKHS) is a Hilbert space of functions whose kernels have special properties. Roughly speaking a reproducing kernel requires that the space of integrands have enough regularity to insure that the evaluation functionals are bounded. Examples of those kernels have been known for a long time see, e.g. the historical introduction in [3]. Reproducing kernel Hilbert spaces play an important rule, e.g. in approximation and regularization theory, but the aspect of RKHS we are mainly interested in is to get an easy approach to generalize so-called Koksma-Hlawka type inequalities. For a generalization of error bounds of the Koksma-Hlawka type that is closely related to the generalization of the concept of geometric discrepancy see, e.g. [38].

This chapter is organized as follows: First of all we give a precise definition of reproducing kernel Hilbert spaces. Then we point out the interrelation of reproducing kernels and error bounds for numerical integration. In the end we will give a brief review of frame theory.

The concept of frames or in other words, a stable representation system was firstly introduced by Duffin and Schaeffer [28] in the context of non-harmonic Fourier series and irregular sampling. But this field of research can be described as well-investigated. With the later concept of wavelet based cubature methods in mind we are mainly interested in the connection between reproducing kernels and frames on a given Hilbert space see, e.g. [44, 50]. In the next sections we consider (separable) Hilbert spaces H of real valued function over some domain $\Omega \subset \mathbb{R}^d$.

2.1 RKHS and generalized discrepancy

The purpose of this section is to point out the interaction of RKHS and the classical theory of geometric discrepancy. We start with the definition that specifies the above phrase of having "enough" regularity. This yields the definition of so called reproducing kernel Hilbert spaces (RKHS). For those readers who are interested in some deeper details of this theory we refer to [3]. In this section we only recall the definition and some elementary properties which are useful in our context. We use the theory of reproducing kernels to verify error bounds for cubature rules in RKHS and we give an elementary example where the error bound tends to a classical and well known measure of non-uniformity, the $\mathcal{L}_2$-discrepancy for corners.

Definition 2.1.1 (RKHS)**.** A Hilbert space H with inner product $\langle \cdot, \cdot \rangle_H$ is a reproducing kernel Hilbert space (RKHS) of real valued functions defined over some domain $\Omega \subset \mathbb{R}^d$ if for each $t \in \Omega$ the evaluation functional δ_t defined as

$$\delta_t[f] = f(t) \quad \text{for all } f \in H$$

are linear bounded functionals.

Then, owing to the well known Riesz representation theorem one can state that to each RKHS there is a unique function $K : \Omega \times \Omega \to \mathbb{R}$ called the reproducing kernel of H. This function K satisfies the so called reproducing property: For all $t \in \Omega$ we get $K(\cdot, t) \in H$ and for all $f \in H$

$$f(t) = \langle f, K(\cdot, t) \rangle_H .$$

According to the well known Cauchy-Schwarz inequality we get, if such a reproducing kernel K exists, for each $t \in \Omega$

$$\begin{aligned} |\delta_t[f]| &= |f(t)| = \langle f, K(\cdot,t)\rangle_H \leq \|f\|_H \, \|K(\cdot,t)\| \\ &= \|f\|_H \, (\langle K(\cdot,t), K(\cdot,t)\rangle_H)^{1/2} \\ &= K(t,t)^{1/2} \, \|f\|_H \, , \end{aligned}$$

consequently the evaluation functionals are linear bounded functionals on H. Thus, each RKHS corresponding to a unique reproducing kernel. Note, the uniqueness of the kernel function is meant in the sense that for two given kernel K and K' on H we get for all $t \in \Omega$ that $\|K(\cdot,t) - K'(\cdot,t)\|_H = 0$. This fact can easily be verified by the reproducing property of K and K', we get

$$\begin{aligned} 0 &\leq \|K(\cdot,t) - K'(\cdot,t)\|_H^2 = \langle K(\cdot,t) - K'(\cdot,t), K(\cdot,t) - K'(\cdot,t)\rangle_H \\ &= \langle K(\cdot,t) - K'(\cdot,t), K(\cdot,t)\rangle_H - \langle K(\cdot,t) - K'(\cdot,t), K'(\cdot,t)\rangle_H \\ &= 0 \end{aligned}$$

and consequently $K = K'$.

The theory of reproducing kernels is often the main mathematical tool to generalize error bounds of the Koksma-Hlawka type. More generally, if H is a RKHS and the reproducing kernel is given, it will be straightforward to compute the representer ξ_T for any bounded linear functional T on H:

$$T[f] = \langle \xi_T, f \rangle_H \text{ for all } f \in H,$$

where

$$\xi_T(t) = \langle K(\cdot, t), \xi_T \rangle_H = T[K(\cdot, t)].$$

In relation to the cubature error one has to assure that $I - Q$ is bounded, then its representer is given by

$$(I - Q)[f] = \langle \xi, f \rangle_H \text{ for all } f \in H,$$

where

$$\begin{aligned} \xi(t) &= \langle K(\cdot, t), \xi \rangle_H \\ &= (I - Q)[K(\cdot, t)]. \end{aligned}$$

Now the Cauchy-Schwarz inequality implies the following upper bound for the cubature error. If $f \in H$ is an arbitrary function, the cubature error of f can be upper bounded by

$$|I[f] - Q[f]| = |(I - Q)[f]| = |\langle \xi, f \rangle_H| \leq \|\xi\|_H \|f\|_H , \tag{2.1.1}$$

where $\|\cdot\|_H$ constitute the norm induced by the inner product of H. Observe, the inequality is sharp in the sense that equality holds when the function f is a constant multiple of the representer ξ. In other words, the representer is the worst case integrand itself. Let us recall that by linearity of I for a given cubature Q the worst case error over H is nothing, but the operator norm $\|I - Q\|_{\text{op}}$ induced by the norm of H. More detailed, in the worst case setting, the error of a linear algorithm Q is given by

$$\text{err}(H, Q) = \sup \{|I[f] - Q[f]| : \ \|f\|_H \leq 1\} .$$

Due to linearity of I and the algorithm Q, we get

$$\text{err}(H, Q) = \|I - Q\|_{\text{op}} \, (= \|\xi\|_H) .$$

Hence, under the above requirements we can formulate the following lemma which also motivates the next section.

Lemma 2.1.2. *Let H be a RKHS with given RK. The worst case error of a given cubature Q can be upper bounded by*

$$\operatorname{err}(H, Q) \leq \|(I-Q)[K(\cdot,t)]\|_H \,,$$

where the norm $\|\cdot\|_H$ relates to the parameter t.

Remark 2.1.3. Obviously the quantity $\|\operatorname{err}(K(\cdot,t), Q)\|_H$ depends only on the given points P and weights used by the cubature rule Q and may be identified as generalized weighted discrepancy of the the set of sample points. Another important observation is that the use of Hölder's inequality in (2.1.1) instead of Cauchy-Schwarz inequality yields to the concept of the so-called $\mathcal{L}_p$-Discrepancy see, e.g. [38, 64].

As an elementary and well known example of RKHS which has the pleasant property to tend to the usual $\mathcal{L}_2$-discrepancy for corners as integration error, we consider the Hilbert space H of absolutely continuous functions $f : [0,1] \to \mathbb{R}$ such that $f(1) = 0$ and $f' \in \mathrm{L}^2([0,1])$. We define an inner product by

$$\langle f, g\rangle_H := \int_0^1 f'(x) g'(x) dx.$$

Thus, this space H with the given inner product is a RKHS and the kernel function is $K(x,y) = \min\{1-x, 1-y\}$. To verify this we have to show that for any f that satisfies the above requirements the evaluation functional is given by $K(\cdot,t)$. But this is an elementary exercise, for all $t \in [0,1]$ we get

$$\begin{aligned}
\langle f, K(\cdot,t)\rangle_H &= \int_0^1 f'(x) K'(x,t) dx \\
&= \int_0^1 f'(x) \begin{cases} -1 & \text{for all } x > t \\ 0 & \text{else} \end{cases} dx \\
&= -\int_t^1 f'(x) dx = f(t).
\end{aligned}$$

The next step is to calculate $\|\operatorname{err}(K(\cdot,t), Q)\|_H$ with the norm induced by the inner product of the considered space and a quadrature simply given by

$$Q(f) = \frac{1}{|P|} \sum_{p \in P} f(p).$$

First, we calculate

$$\int_0^1 K(x,t) dx = 1/2 - t^2/2,$$

consequently we get

$$
\begin{aligned}
\|(I-Q)[K(\cdot,t)]\|_H^2 &= \int_0^1 \left(\frac{\partial}{\partial t}(I-Q)[K(\cdot,t)]\right)^2 dt \\
&= \int_0^1 \left(\frac{\partial}{\partial t}\int_0^1 K(x,t)dx - \frac{1}{|P|}\frac{\partial}{\partial t}\sum_{p\in P}\min\{1-p,1-t\}\right)^2 dt \\
&= \int_0^1 \left(-t+\frac{1}{|P|}\left|P\cap[0,t)\right|\right)^2 dt = \mathrm{disc}_2(P).
\end{aligned}
$$

For more general examples we refer the interested reader to [64]. This tool is a great mechanism to generalize error bounds of the Koksma-Hlawka type. But the remaining problem is –one has to find interesting function spaces which have the pleasant property to be a RKHS. The next requirement one has to attend is that the kernel is relatively easy to calculate and yields to a common measure of non-uniformity. But how does in general a common measure of non-uniformity $D(P)$ for a given point set P looks like? A brief discussion on the desirable qualities were found e.g. in [38]:

- $D(P)$ should arise from an error bound for cubature, function approximation, or some other application,
- projections of P into lower dimensional space should not increase $D(P)$,
- $D(P)$ should be easy to compute,
- $D(P)$ should have an intuitive interpretation,
- it should be invariant under certain transformations of P, such as reflections about the plane $x_j = 1/2$ and permutations of the coordinates.

The first point is naturally satisfied in all the cases we are interested in. But for the other qualities we have sometimes failed. However, the generalized $\mathcal{L}_p$-discrepancy only satisfies the first two criteria. Moreover, also the classical $\mathcal{L}_p$-star discrepancy that is well-investigated and one of the classical measures of non-uniformity does not suffice to all criteria unless the special case $p = \infty$.

2.2 Reproducing kernel Hilbert spaces and their frames

One of the most important question in the previous section is how to find interesting function spaces to generalize error bounds of the Koksma-Hlawka type. A canonical method to compute those kernels on suitable spaces is to take a closer look on its basis. In this section we discuss the connection between reproducing kernels and orthonormal systems or rather frames. The following classical result is proved by Meschkowski [44] and shows how the kernel function looks like in a (separable) reproducing kernel Hilbert space where $\{\phi_\lambda\}_\lambda$ is an arbitrary orthonormal system.

Theorem 2.2.1 (Meschkowski)**.** *Let H be a separable Hilbert space with reproducing kernel and let $\{\phi_\lambda\}_\lambda$ be an orthonormal system of H. Then the kernel function is given by*

$$K(s,t) = \sum_\lambda \phi_\lambda(s)\phi_\lambda(t).$$

We are mainly interested in a generalization of this theorem and the natural question on what terms the space H is a RKHS. It turns out that the considered function system in Theorem 2.2.1 does not have to be an orthonormal system, it can even be linearly dependent see, e.g. [50]. But first of all, we give a short introduction into frame theory. The task of frame theory is to establish general conditions under which one can perfectly reconstruct a given function f in a Hilbert space from its inner product with a family of vectors belonging to the given space see, e.g. [20, 28].

Definition 2.2.2. Let $\{\psi_\lambda\}_\lambda$ be a countable system in a separable Hilbert space H. The system $\{\psi_\lambda\}_\lambda$ is a frame of H if there exist constants $A, B > 0$ such that for any $f \in H$

$$A\,\|f\|_H^2 \le \sum_\lambda |\langle f, \psi_\lambda\rangle_H|^2 \le B\,\|f\|_H^2\,. \tag{2.2.1}$$

When $A = B$ the frame is said to be a tight frame.

Remark 2.2.3. When the frame vectors are normalized $\|\psi_\lambda\| = 1$ the frame bounds A and B can be understood as a measure of redundancy of the reconstruction. As a special case a Riesz basis is a frame whose vectors a linearly independent. The frame is then an orthonormal basis if and only if $A = B = 1$.

Before we discuss how to reconstruct a function by a frame we give some elementary examples. Apparently, a given orthonormal basis $\{e_\lambda\}_\lambda$ of H is a normalized tight frame, but also the following sequences are normalized tight frames for H,

$$\{e_1, 0, e_2, 0, e_3, 0, \ldots\}\,,$$

and

$$\left\{\frac{e_1}{\sqrt{2}}, \frac{e_1}{\sqrt{2}}, \frac{e_2}{\sqrt{2}}, \frac{e_2}{\sqrt{2}}, \frac{e_3}{\sqrt{2}}, \frac{e_3}{\sqrt{2}}, \ldots\right\}.$$

If the frame condition (2.2.1) is satisfied, the frame operator $\mathcal{A}$ can be defined by

$$\mathcal{A} : H \to l^2, \quad f \mapsto \{\langle f, \psi_\lambda \rangle\}_\lambda . \tag{2.2.2}$$

The reconstruction of a function $f \in H$ from its frame coefficients $\{\langle f, \psi_\lambda \rangle\}_\lambda$ need the definition of the so called dual frame. First, we introduce the adjoint operator $\mathcal{A}^*$ of $\mathcal{A}$ which exists and is unique because $\mathcal{A}$ is a bounded linear operator on a Hilbert space. The adjoint operator is defined by

$$\mathcal{A}^* : l^2 \to H, \quad \{a_\lambda\}_\lambda \mapsto \sum_\lambda a_\lambda \psi_\lambda.$$

Thus, the dual frame is defined by

$$\tilde{\psi}_\lambda := (\mathcal{A}^*\mathcal{A})^{-1} \psi_\lambda. \tag{2.2.3}$$

The following well known theorem gives the reconstruction of a function from its frame coefficients, for proof see [20].

Theorem 2.2.4 (Daubechies)**.** *Let $\{\psi_\lambda\}_\lambda$ be a frame of H with the frame bounds A and B. Then the dual frame satisfies for all $f \in H$*

$$\frac{1}{B} \|f\|_H^2 \leq \sum_\lambda \left| \langle f, \tilde{\psi}_\lambda \rangle_H \right|^2 \leq \frac{1}{A} \|f\|_H^2 , \tag{2.2.4}$$

and we get an expansion for f

$$f = \sum_\lambda \langle f, \psi_\lambda \rangle_H \tilde{\psi}_\lambda = \sum_\lambda \langle f, \tilde{\psi}_\lambda \rangle_H \psi_\lambda. \tag{2.2.5}$$

Remark 2.2.5. If the frame is tight and this is the case we are mainly interested in the later chapter, the dual frame is a constant multiple of the frame itself,

$$\tilde{\psi}_\lambda = A^{-1} \cdot \psi_\lambda. \tag{2.2.6}$$

If the original frame is a Riesz basis, then the two frames form a biorthogonal basis system,

$$\langle \psi_\lambda, \tilde{\psi}_{\lambda'} \rangle = \delta_{\lambda,\lambda'}. \tag{2.2.7}$$

Now, after this brief introduction on frame theory, we point out under which conditions a Hilbert space H with given frame is also a RKHS.

Theorem 2.2.6. *Let H be a Hilbert space of real valued function over some domain $\Omega \subset \mathbb{R}^d$. Let $\{\psi_\lambda\}_\lambda$ be a frame on it. If the frame satisfies the property that for all $t \in \Omega$,*

$$\left\| \sum_\lambda \tilde{\psi}_\lambda(\cdot) \psi_\lambda(t) \right\|_H < \infty, \tag{2.2.8}$$

then H is a RKHS.

Proof. All we have to point out is that the evaluation functionals δ_t are linear bounded functionals. By the representation of a function f by its frame coefficients and the Cauchy-Schwarz inequality we get for all $t \in \Omega$

$$\begin{aligned} |\delta_t[f]| &= |f(t)| = \left|\sum_\lambda \langle f, \tilde{\psi}_\lambda\rangle_H \psi_\lambda(t)\right| = \left|\langle f, \sum_\lambda \tilde{\psi}_\lambda(\cdot)\psi_\lambda(t)\rangle_H\right| \\ &\leq \|f\|_H \underbrace{\left\|\sum_\lambda \tilde{\psi}_\lambda(\cdot)\psi_\lambda(t)\right\|_H}_{=:C_t} = C_t \|f\|_H . \end{aligned}$$

□

By the uniqueness of the reproducing kernel it is not difficult to calculate how the kernel looks like on a RKHS with given frame.

Theorem 2.2.7. *Let H be a RKHS and $\{\psi_\lambda\}_\lambda$ be a frame on H. The reproducing kernel is given by*

$$K(s,t) = \sum_\lambda \tilde{\psi}_\lambda(s)\psi_\lambda(t).$$

Proof. The proof based again on the expansion of a function f by its frame coefficients, we get

$$f(t) = \sum_\lambda \langle f, \tilde{\psi}_\lambda\rangle_H \psi_\lambda(t) = \langle f, \sum_\lambda \tilde{\psi}_\lambda(\cdot)\psi_\lambda(t)\rangle_H,$$

by the reproducing property of the kernel function K we get for all f,

$$K(\cdot,t) = \sum_\lambda \tilde{\psi}_\lambda(\cdot)\psi_\lambda(t) \quad \text{for all } t \in \Omega$$

and consequently $K(s,t) = \sum_{\lambda,} \tilde{\psi}_\lambda(s)\psi_\lambda(t)$. □

Remark 2.2.8. Obviously the kernel $K(s,t) = \sum_{\lambda,} \tilde{\psi}_\lambda(s)\psi_\lambda(t)$ is symmetric, this again can easily be checked by the frame representation.

The next step is to apply the knowledge of the connection between RKHS and its frame to the worst case error of a given cubature rule. But we get a little bit more information since the structure of the error bounds also give an idea how to construct more or less good cubature rules on those spaces.

Corollary 2.2.9. *Let H be a Hilbert space of real valued function over some domain Ω. Let $\{\psi_\lambda\}_\lambda$ be a frame on H that satisfies the reproducing property (2.2.8). If we assume $\int_\Omega \sum_\lambda \tilde{\psi}_\lambda(\cdot)\psi_\lambda(t)dt \in H$, we get for a given cubature rule Q an error bound*

$$\operatorname{err}(H,Q) \leq \left\|\sum_\lambda \tilde{\psi}_\lambda(\cdot)\left\{\int_\Omega \psi_\lambda(t)dt - Q[\psi_\lambda]\right\}\right\|_H . \tag{2.2.9}$$

Proof. Because of the previous results it is obvious that H is a RKHS, the reproducing kernel is given by $K(s,t) = \sum_\lambda \tilde{\psi}_\lambda(s)\psi_\lambda(t)$. We only calculate the given error bound in Lemma 2.1.2 for the cubature Q and to avoid confusion on which part Q has effect we write $Q_p(f(p)) := Q[f]$.

$$\begin{aligned}
\|(I-Q)[K(\cdot,t)]\|_H &= \left\| \int_\Omega K(\cdot,t)dt - Q_p[K(\cdot,p)]) \right\|_H \\
&= \left\| \int_\Omega \sum_\lambda \tilde{\psi}_\lambda(\cdot)\psi_\lambda(t)dt - Q_p[\sum_\lambda \tilde{\psi}_\lambda(\cdot)\psi_\lambda(p)] \right\|_H \\
&= \left\| \sum_\lambda \tilde{\psi}_\lambda(\cdot) \left\{ \int_\Omega \psi_\lambda(t)dt - Q[\psi_\lambda] \right\} \right\|_H .
\end{aligned}$$

□

Remark 2.2.10. The first observation we made is that this way of looking at general cubature errors takes us more and more of the subject of classical notion of Koksma-Hlawka bounds and the corresponding (geometric) discrepancy. It seems that good cubature rules which minimize the error term stand out due to the fact that they are exact for frames up to a special index λ'. Hence, the term $\int_\Omega \psi_\lambda(t)dt - Q[\psi_\lambda]$ vanishes for the first indices. Afterwords, considering subspaces of $\mathrm{L}^2(\Omega)$, we examine orthonormal bases or tight frames of $\mathrm{L}^2(\Omega)$ and rescale them to satisfy the reproducing property (2.2.8). We deal only with those functions of $\mathrm{L}^2(\Omega)$ whose frame coefficients decrease rapidly with respect to the rescaled frame.

The next step is to define discrete norms on this spaces with weighted frame coefficients. Thus, we can uses Cauchy-Schwarz inequality directly to get upper bounds for the cubature error instead of defining inner product and corresponding norm. Since we are interested in frames that satisfy some locality property, the error terms are indeed dominated by the first indices in (2.2.9).

Chapter 3

A Wavelet Based Error Analysis

In the previous section we have described the useful connection between Hilbert spaces with reproducing kernel and corresponding properties of a given basis, or more generally a frame on this space. A very popular method to construct bases on L^2 is a wavelet approach, the so called multiresolution analysis. In the first part of this chapter we will give a short introduction on stationary multiresolution analysis. We will discuss the interrelation between wavelet bases and one-dimensional quadrature formulas based upon the well known and often stressed example of multiresolution analysis, the Haar wavelets. Afterwords we will take a closer look on the interaction between discrete norms defined by weighted wavelet series,

$$\|f\|_s^2 = \sum_\lambda 2^{|\lambda|2s} \left|\langle f, \psi_\lambda\rangle\right|^2$$

and some classical notion of smoothness, e.g. norms on classical function spaces like Sobolev or Besov spaces. There is a general theory of analyzing these equivalences based on so called Jackson and Bernstein estimates. After we having recalled the main results that concern these correlations for biorthogonal wavelet bases we will get back from this general approach to the problem of numerical integration.

Those scaled wavelet bases are a useful tool to analyze the worst case error of a given cubature rule since they enable a useful depiction of an error function. To focus on this idea we analysis the worst case error on function spaces defined by Haar wavelet series. For the two-dimensional case a good cubature rule is naturally given by binary nets. But the d-dimensional generalizations are not strait forward see, e.g. [37]. Observe that for the Haar wavelet case it's not clear which classical notion of smoothness is considered for parameter $s > 1$. This fact will yield a generalization of the Haar wavelet.

3.1 A brief review of (stationary) multiresolution analysis

The aim of this section is to give a brief description on stationary multiresolution analysis. Generally, these schemes are based on hierarchies of nested spaces that reproduce the main features like local approximation and smoothness. One can construct wavelet (orthogonal) bases $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ of $\mathrm{L}^2(\mathbb{R})$ where $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ is the dilated and translated family

$$\left\{\psi_{j,k} = 2^{j/2}\psi(2^j \cdot -k)\right\}_{j,k\in\mathbb{Z}},$$

for a suitable choice of $\psi \in \mathrm{L}^2(\mathbb{R})$. In the literature the function ψ is often called a discrete wavelet. This definition is relatively simple, but it is not granted that such a function ψ exists. As mentioned before, in this chapter we will give a brief review of a construction called multiresolution analysis, which allows to construct bases spanned by the dilated a translated version of an fine function $\psi \in \mathrm{L}^2(\mathbb{R})$ in a rather systematical way see, e.g. [12, 15, 20, 42]. Also desirable properties like smoothness, vanishing moments and compact support are guaranteed. Afterwords, with the problem of analyze cubature rules in mind we are interested in a multiresolution analysis on bounded intervals. This leads to some difficulties on the boundaries and will give way to generalizations like multiwavelets or biorthogonal multiresolution analysis see, e.g. [2, 18, 61, 66]. However, let us start with the ordinary definition of a multiresolution analysis in a slightly different way as introduced by Mallat [42] and Meyer [45].

Definition 3.1.1 (multiresolution analysis)**.** A sequence $\{V_j\}_{j\in\mathbb{Z}}$ of closed subspaces of $\mathrm{L}^2(\mathbb{R})$ is a multiresolution analysis (MRA) if the following six properties are satisfied:

1. For all $j \in \mathbb{Z}$, $V_j \subset V_{j+1}$.
2. $\overline{\bigcup_{j\in\mathbb{Z}} V_j} = \mathrm{L}^2(\mathbb{R})$.
3. $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$.
4. For all $j \in \mathbb{Z}$ and $f \in \mathrm{L}^2(\mathbb{R})$ it holds $f \in V_j \Leftrightarrow f(2\cdot) \in V_{j+1}$.
5. $f \in V_j$ and $k \in \mathbb{Z} \Rightarrow f(\cdot\ - 2^{-j}k) \in V_j$.
6. There exists a function $\varphi \in V_0$ such that $\{\varphi(\cdot\ - k)\}_{k\in\mathbb{Z}}$ is a Riesz basis of V_0. φ is called the scaling or generating function of $\{V_j\}_{j\in\mathbb{Z}}$.

Examples are given by the well known Haar multiresolution analysis with the scaling function given by $\varphi = \mathbf{1}_{[0,1)}$ or the Shannon MRA were V_0 is the space of bandlimited functions

$$V_0 = \left\{f \in \mathrm{L}^2(\mathbb{R}) \ : \ \mathrm{supp}\ \hat{f} \subset [-1/2, 1/2]\right\},$$

the scaling function is given by the sinc-function. For a detailed discussion on these somewhat redundant properties we refer the reader to [42] or any other reliable source. So far, there are no wavelets defined. And also in the next chapters there should be no confusion between the scaling function of a MRA and the corresponding wavelet. Denote by W_j the orthogonal complement on V_j in V_{j+1}, i.e.,

$$W_j := \{\psi \in V_{j+1} \ : \ \langle \psi, \varphi \rangle = 0 \text{ for all } \varphi \in V_j\}.$$

W_j is often called the detailed space of scale j, if there exist a function $\psi \in W_0$ such that

$$W_0 = \operatorname{span}\{\psi(\cdot - k) \ : \ k \in \mathbb{Z}\}$$

and ψ is a Riesz Basis of W_0 the function ψ will be called a wavelet. In this case we get a decomposition

$$\mathrm{L}^2(\mathbb{R}) = \overline{\bigoplus_{l=-\infty}^{\infty} W_l},$$

with $\left\{\psi_{j,k} := 2^{l/2}\psi(2^l \cdot \, -k) \ : \ k \in \mathbb{Z}, \ l \in \mathbb{Z}\right\}$ a Riesz basis of $\mathrm{L}^2(\mathbb{R})$. From a practical point of view, the locality condition of the scaling function and also the wavelet with respect to the corresponding scale j is of special interest, i.e.

$$\operatorname{supp} \varphi_{j,k} \sim 2^{-j} \quad \text{and } \operatorname{supp} \psi_{j,k} \sim 2^{-j}.$$

The final requirement we pose on a wavelet basis is very useful for the analysis of cubature error. We claim the wavelets have (e.g. d) vanishing moments, i.e.

$$\langle \psi_{j,k}, x^\alpha \rangle = 0, \quad \alpha = 0, 1, 2, \ldots, d-1.$$

This construction of a wavelet (Riesz) bases via multiresolution analysis can easily be extended to multivariate functions defined on $\mathbb{R}^d$. The simplest method is the tensor product strategy. But the adaption of MRA to a bounded domain and in particular the formulation of boundary condition is a difficult task by itself. As mentioned in the introduction of this chapter this yields to more general aspects of MRA. During the later chapter of this thesis we will consider a MRA based on discontinuous multiwavelets on intervals or a non-stationary multiresolution analysis that yields to a tight frame see, e.g. [11, 16]. However, first let us give one well known and elementary example were all these problems do not appear. Consider the Haar MRA as mentioned before and the dilated and translated version of the scaling function $\varphi = \mathbf{1}_{[0,1)}$,

$$\varphi_{j,k} = 2^{j/2}\varphi(2^j \ \cdot -k), \quad \text{for } k \in \mathbb{N}_0, \ j \in \mathbb{N}_0.$$

Observe these function have compact support

$$\operatorname{supp}\varphi_{j,k} = [2^{-j}k, 2^{-j}(k+1)] =: I_k^j, \quad \text{for } j \in \mathbb{N}_0,\ k = 0,1,2,\ldots,2^j-1$$

and orthogonality

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

Furthermore, we define the spaces of piecewise constant function

$$V_j := \operatorname{span}\left\{\varphi_{j,k} \ :\ k = 0,1,2,\ldots,2^j-1\right\}.$$

Let us use the shorthand $\Delta_j = \{0,1,2,\ldots,2^j-1\}$. It is obvious that the spaces V_j has dimension 2^j and

$$V_0 \subset V_1 \subset \ldots \subset V_j \subset V_{j+1} \subset \ldots \subset \mathrm{L}^2([0,1]).$$

For $j = 0,1,2,\ldots$ we define the 2^j-dimensional spaces W_j to be the orthogonal complement of V_j in V_{j+1}, so we get inductively the decomposition

$$V_j = V_0 \oplus W_0 \oplus W_1 \oplus \ldots \oplus W_{j-1}.$$

The functions that spanned the spaces W_j are again piecewise constant and are given by dilatation and translation of the well known Haar function,

$$\psi = \mathbf{1}_{[1,1/2)} - \mathbf{1}_{[1/2,1)}.$$

These functions are also supported on so-called canonical intervals I_k^j

$$\operatorname{supp}\varphi_{j,k} = I_k^j, \quad \text{for } j \in \mathbb{N}_0,\ k \in \nabla_j = \left\{0,1,2,\ldots,2^j-1\right\},$$

satisfies the following orthogonality conditions:

$$\langle \psi_{j,k}, \psi_{j',k'} \rangle = \delta_{j,j'}\delta_{k,k'}$$

and since the spaces W_j are orthogonal to the space V_0 we have one vanishing moment

$$\int_0^1 \psi_{j,k} dx = 0.$$

Thus, we get

$$\overline{\bigcup_{j=0}^{\infty} V_j} = \overline{V_0 \oplus \bigoplus_{j=0}^{\infty} W_j} = \mathrm{L}^2([0,1])$$

and the well known Haar wavelet bases of $\mathrm{L}^2([0,1])$ is given by

$$\varphi \cup \{\psi_{j,k} \ : \ j \geq 0, \ k \in \nabla_j\}.$$

For convenience we define $\psi_{-1} := \varphi$ and consequently for every function $f \in \mathrm{L}^2([0,1])$ we get the following unique Haar wavelet expansion

$$f = \sum_{j \geq -1} \sum_{k \in \nabla_j} \langle f, \psi_{j,k} \rangle \psi_{j,k} \quad \text{and} \quad \|f\|^2 = \sum_{j \geq -1} \sum_{k \in \nabla_j} |\langle f, \psi_{j,k} \rangle|^2.$$

Now, let us apply the expertise of the previous section to the case of Haar wavelet bases. Let $s > 0$ we rescale the Haar bases of $\mathrm{L}^2([0,1])$ by

$$\psi^s_{j,k} := 2^{-js} \psi_{j,k},$$

also, we define the inner product $\langle \cdot, \cdot \rangle_s$ by

$$\langle f, g \rangle_s := \sum_{j \geq -1} \sum_{k \in \nabla_j} 2^{j2s} \langle f, \psi_{j,k} \rangle \langle g, \psi_{j,k} \rangle.$$

Consider the spaces H_s defined by

$$H_s = \left\{ f \in \mathrm{L}^2([0,1]) \ : \ \sum_{j \geq -1} \sum_{k \in \nabla_j} 2^{j2s} |\langle f, \psi_{j,k} \rangle|^2 < \infty \right\}. \tag{3.1.1}$$

Hence, we can prove that this space is a RKHS and as pointed out in the chapter before we are also able to compute the kernel function of these space by the scaled Haar wavelet bases.

Theorem 3.1.2. *For $s > 1/2$ the space H_s endowed with the inner product $\langle \cdot, \cdot \rangle_s$ is a RKHS and the kernel function is given by*

$$K(s,t) = \sum_{j \geq -1} \sum_{k \in \nabla_j} \psi^s_{j,k}(s) \psi^s_{j,k}(t).$$

Proof. Let $f \in H_s$ obviously $\left\{\psi^s_{j,k}\right\}_{j,k \in \nabla_j}$ is a tight frame of H_s we get

$$\begin{aligned} \sum_{j \geq -1} \sum_{k \in \nabla_j} \langle f, \psi^s_{j,k} \rangle_s \psi^s_{j,k} &= \sum_{j \geq -1} \sum_{k \in \nabla_j} \sum_{l \geq -1} \sum_{m \in \nabla_l} 2^{l2s} \langle f, \psi_{l,m} \rangle \langle \psi^s_{j,k}, \psi_{l,m} \rangle \psi^s_{j,k} \\ &= \sum_{j \geq -1} \sum_{k \in \nabla_j} \sum_{l \geq -1} \sum_{m \in \nabla_l} 2^{l2s} \langle f, \psi_{l,m} \rangle 2^{-j2s} \langle \psi_{j,k}, \psi_{l,m} \rangle \psi_{j,k} \\ &= \sum_{j \geq -1} \sum_{k \in \nabla_j} \langle f, \psi_{j,k} \rangle \psi_{j,k} = f. \end{aligned}$$

We only have to show that the reproducing property in Theorem 2.2.8 in Chapter 2 is satisfied. Let $t \in [0,1]$ we consider the norm $\|\cdot\|_s$ induced by the inner product $\langle \cdot, \cdot \rangle_s$ and get

$$\begin{aligned}
\|K(\cdot,t)\|_s^2 &= \left\| \sum_{j\geq -1} \sum_{k\in\nabla_j} \psi_{j,k}^s(\cdot)\psi_{j,k}^s(t) \right\|_s^2 \\
&= \sum_{l\geq -1} \sum_{m\in\nabla_l} 2^{l2s} \langle \sum_{j\geq -1} \sum_{k\in\nabla_j} \psi_{j,k}^s(\cdot)\psi_{j,k}^s(t), \psi_{l,m}\rangle \langle \sum_{j\geq -1} \sum_{k\in\nabla_j} \psi_{j,k}^s(\cdot)\psi_{j,k}^s(t), \psi_{l,m}\rangle \\
&= \sum_{l\geq -1} \sum_{m\in\nabla_l} 2^{l2s} \sum_{j\geq -1} \sum_{k\in\nabla_j} \psi_{j,k}^s(t)\langle \psi_{j,k}^s, \psi_{l,m}\rangle \sum_{j\geq -1} \sum_{k\in\nabla_j} \psi_{j,k}^s(t)\langle \psi_{j,k}^s, \psi_{l,m}\rangle \\
&= \sum_{l\geq -1} \sum_{m\in\nabla_l} 2^{l2s} \psi_{l,m}^s(t) 2^{-ls} \psi_{l,m}^s(t) 2^{-ls} \\
&= \sum_{l\geq -1} \sum_{m\in\nabla_l} \psi_{l,m}^s(t)^2.
\end{aligned}$$

Thus it is enough to show that $\sum_{l\geq -1} \sum_{m\in\nabla_l} \psi_{l,m}^s(t)^2 < \infty$, but by some elementary calculations we get

$$\begin{aligned}
\sum_{l\geq -1} \sum_{m\in\nabla_l} \psi_{l,m}^s(t)^2 &= \sum_{l\geq -1} \sum_{m\in\nabla_l} 2^{-l2s} \psi_{l,m}(t)^2 \\
&= \sum_{l\geq -1} \sum_{m\in\nabla_l} 2^{-l2s} \left(2^{l/2}\psi(2^l t - m)\right)^2 \\
&= \sum_{l\geq -1} \sum_{m\in\nabla_l} 2^{-l(2s-1)} \psi(2^l t - m)^2 \\
&= \sum_{l\geq -1} 2^{-l(2s-1)} < \infty.
\end{aligned}$$

□

For the corresponding quadrature error on H_s we get by Corollary 2.2.9 in Chapter 2 the following error bounds.

Corollary 3.1.3. *Let $s > 1/2$ and H_s defined as above. For a given quadrature*

$$Q[f] = \frac{1}{|P|} \sum_{p\in P} f(p)$$

we get an error bound of

$$\mathrm{err}(H_s, Q) \leq \sum_{l\geq 0} 2^{-l(2s-1)} \sum_{m\in\nabla_l} \left\{ \frac{1}{|P|} \sum_{p\in P} \Delta_{l,m}^-(p) - \Delta_{l,m}^+(p) \right\}^2,$$

with

$$\Delta^-_{l,m}(p) = \mathbf{1}_{[2^{-l}m, 2^{-l}(m+1/2))}(p) \quad \textit{and } \Delta^+_{l,m}(p) = \mathbf{1}_{[2^{-l}(m+1/2), 2^{-l}(m+1))}(p).$$

Proof. Obviously

$$I[K(\cdot,t)] = \int_0^1 K(\cdot,t)dt \in H_s$$

and we can directly use Corollary 2.2.9 to calculate the discrepancy $\|(I-Q)[K(\cdot,t)]\|_s$ which the norm relates to the smoothness parameter s. Thus,

$$\begin{aligned}
\|(I-Q)[K(\cdot,t)]\|_s^2 &= \|I[K(\cdot,t)] - Q_t[K(\cdot,t)]\|_s^2 \\
&= \left\| \int_0^1 \sum_{j\geq -1} \sum_{k\in\nabla_j} \psi^s_{j,k}(\cdot)\psi^s_{j,k}(t)dt - Q_t[\sum_{j\geq -1}\sum_{k\in\nabla_j} \psi^s_{j,k}(\cdot)\psi^s_{j,k}(t)] \right\|_s^2 \\
&= \left\| \sum_{j\geq -1} \sum_{k\in\nabla_j} \psi^s_{j,k}(\cdot) \int_0^1 \psi^s_{j,k}(t)dt - \sum_{j\geq -1}\sum_{k\in\nabla_j} \psi^s_{j,k}(\cdot)Q_t[\psi^s_{j,k}(t)] \right\|_s^2 \\
&= \left\| \sum_{j\geq -1} \sum_{k\in\nabla_j} \psi^s_{j,k}(\cdot) \left\{ \int_0^1 \psi^s_{j,k}(t)dt - Q_t[\psi^s_{j,k}(t)] \right\} \right\|_s^2 .
\end{aligned}$$

Since every non-trivial quadrature is exact on V_0, we get by the fact, that we have vanishing moments

$$\|(I-Q)[K(\cdot,t)]\|_s^2 = \left\| \sum_{j\geq 0} \sum_{k\in\nabla_j} \psi^s_{j,k}(\cdot) \left\{ -Q_t[\psi^s_{j,k}(t)] \right\} \right\|_s^2,$$

consequently, since

$$\begin{aligned}
& \sum_{l\geq -1} \sum_{m\in\nabla_l} 2^{l2s} \langle \sum_{j\geq 0} \sum_{k\in\nabla_j} \psi^s_{j,k}(\cdot) \left\{ -Q_t[\psi^s_{j,k}(t)] \right\}, \psi_{l,m} \rangle \\
& \cdot \langle \sum_{j\geq 0} \sum_{k\in\nabla_j} \psi^s_{j,k}(\cdot) \left\{ -Q_t[\psi^s_{j,k}(t)] \right\}, \psi_{l,m} \rangle \\
=\ & \sum_{l\geq -1} \sum_{m\in\nabla_l} 2^{l2s} \sum_{j\geq 0} \sum_{k\in\nabla_j} \left\{ -Q_t[\psi^s_{j,k}(t)] \right\} \langle \psi^s_{j,k}, \psi_{l,m} \rangle \\
& \sum_{j\geq 0} \sum_{k\in\nabla_j} \left\{ -Q_t[\psi^s_{j,k}(t)] \right\} \langle \psi^s_{j,k}, \psi_{l,m} \rangle \\
=\ & \sum_{l\geq 0} \sum_{m\in\nabla_l} \left\{ -Q_t[\psi^s_{l,m}(t)] \right\}^2,
\end{aligned}$$

we get

$$\begin{aligned}
\|(I-Q)[K(\cdot,t)]\|_s^2 &= \sum_{l\geq 0}\sum_{m\in\nabla_l} 2^{-l2s}\left\{-Q[\psi_{l,m}]\right\}^2 \\
&= \sum_{l\geq 0}\sum_{m\in\nabla_l} 2^{-l2s}\left\{\frac{1}{|P|}\sum_{p\in P}\psi_{l,m}(p)\right\}^2 \\
&= \sum_{l\geq 0}\sum_{m\in\nabla_l} 2^{-l2s}\left\{\frac{1}{|P|}\sum_{p\in P}2^{l/2}\psi(2^l p-m)\right\}^2 \\
&= \sum_{l\geq 0}2^{-l(2s-1)}\sum_{m\in\nabla_l}\left\{\frac{1}{|P|}\sum_{p\in P}\psi(2^l p-m)\right\}^2 .
\end{aligned}$$

According to the definition of the Haar wavelet

$$\|(I-Q)[K(\cdot,t)]\|_s^2 = \sum_{l\geq 0}2^{-l(2s-1)}\sum_{m\in\nabla_l}\left\{\frac{1}{|P|}\sum_{p\in P}\psi(2^l p-m)\right\}^2$$

$$= \sum_{l\geq 0}2^{-l(2s-1)}\sum_{m\in\nabla_l}\left\{\frac{1}{|P|}\sum_{p\in P}\mathbf{1}_{[0,1/2)}(2^l p-m)-\mathbf{1}_{[1/2,1)}(2^l p-m)\right\}^2$$

$$= \sum_{l\geq 0}2^{-l(2s-1)}\sum_{m\in\nabla_l}\left\{\frac{1}{|P|}\sum_{p\in P}\Delta^-_{l,m}(p)-\Delta^+_{l,m}(p)\right\}^2 ,$$

where

$$\Delta^-_{l,m}(p)=\mathbf{1}_{[2^{-l}m,2^{-l}(m+1/2))}(p) \quad\text{and}\quad \Delta^+_{l,m}(p)=\mathbf{1}_{[2^{-l}(m+1/2),2^{-l}(m+1))}(p).$$

□

Remark 3.1.4. Obviously most of this analysis also works in a more general setting, where the spaces H_s are defined by an arbitrary stationary multiresolution analysis. But at this stage we restrict our analysis to the Haar wavelet case, because the considered spaces are spanned by a single generator and the corresponding quadrature rules have simple structure. On the general case we will focus in the forthcoming chapters.

Remark 3.1.5. This calculation seems to be quite detailed and of course we can also use the shortcut to apply Cauchy-Schwarz inequality directly without the detour to consider RKHS and the corresponding generalized geometric discrepancy. Nevertheless, the way of looking at these problems via RKHS is also of interest and poses more or less the background for our later analysis. To point out this fact we give a short alternative proof of this corollary using the Cauchy-Schwarz inequality directly without using RKHS. In the later chapters we will always make this shortcut and have to accept that the concept of geometric discrepancy will get lost in our analysis.

Proof (alternative). Let $f \in H_s$ and $s > 1/2$, we get by wavelet expansion and Cauchy-Schwarz inequality

$$\begin{aligned}
\mathrm{err}(f,Q) &= |I[f]-Q[f]| = \left| I[\sum_{j\geq -1}\sum_{k\in\nabla_j}\langle f,\psi_{j,k}\rangle\psi_{j,k}] - Q[\sum_{j\geq -1}\sum_{k\in\nabla_j}\langle f,\psi_{j,k}\rangle\psi_{j,k}] \right| \\
&= \left| \sum_{j\geq -1}\sum_{k\in\nabla_j} 2^{js}\langle f,\psi_{j,k}\rangle 2^{-js}\left\{I[\psi_{j,k}]-Q[\psi_{j,k}]\right\} \right| \\
&\leq \|f\|_s \left(\sum_{j\geq -1}\sum_{k\in\nabla_j} 2^{-j2s}\left\{I[\psi_{j,k}]-Q[\psi_{j,k}]\right\}^2 \right)^{1/2}.
\end{aligned}$$

Since the worst case error is nothing but the induced operator norm the rest of the proof is similar to the first one. $\square$

Indeed, the question of good quadrature points in this case is quite simple to respond. We have to find a point distribution that is uniformly with respect to the canonical intervals such that

$$\sum_{k\in\nabla_j}\left\{\sum_{p\in P}\Delta^-_{j,k}(p)-\Delta^+_{j,k}(p)\right\}^2 = 0, \tag{3.1.2}$$

up to a special level l i.e. $Q_P[\psi_{j,k}] = 0$ for all $j < l$ and $k \in \Delta_j$.

Another point of view, again more distant from the classical theory of non-uniformity (geometric discrepancy), is the question of quadrature rules that are exact for the spaces V_j up to a special level l. In this case we need only one point on the support of the dilated and translated version of the scaling function. Thus the simple quadrature defined for a level l by

$$Q_l[f] := \frac{1}{2^l}\sum_{k=0}^{2^l-1} f(2^{-l}\cdot k + 2^{-(l+1)}),$$

is exact for the spaces $V_0 \subset V_1 \subset \ldots \subset V_l$. We get the following error bound.

Corollary 3.1.6. *Let $s > 1/2$ and the space H_s defined as above. Then*

$$\mathrm{err}(H_s,Q_l) \leq \frac{2^{-ls}}{\sqrt{1-2^{1-2s}}}.$$

Proof. Since the quadrature Q_l is exact on V_l and $\psi_{j,k} \in W_j \subset V_{j+1}$, i.e. $Q_l[\psi_{j,k}] = 0$ for all $j < l$ and $k \in \Delta_j$ we get an upper bound for the worst case error by

$$\begin{aligned}
\operatorname{err}(H_s, Q_l)^2 &\leq \sum_{j \geq l} 2^{-j(2s-1)} \sum_{k \in \nabla_j} \left\{ \frac{1}{2^l} \sum_{\kappa=0}^{2^l-1} \psi(2^j(2^{-l} \cdot \kappa + 2^{-(l+1)}) - k) \right\}^2 \\
&= \sum_{j \geq l} 2^{-2l} 2^{-j(2s-1)} \sum_{k \in \nabla_j} \left\{ \sum_{\kappa=0}^{2^l-1} \psi(2^j(2^{-l} \cdot \kappa + 2^{-(l+1)}) - k) \right\}^2 \\
&\leq \sum_{j \geq l} 2^{-2l} 2^{-j(2s-1)} \sum_{k \in \nabla_j} \left\{ \sum_{\kappa=0}^{2^l-1} \mathbf{1}_{I_k^j}(2^{-l} \cdot \kappa + 2^{-(l+1)}) \right\}^2 .
\end{aligned}$$

For $j \geq l$ and a given $k \in \nabla_j$ let $\kappa' = \kappa'(j, k, l)$ be the unique element

$$\kappa' \in \{0, 1, 2, \ldots, 2^l - 1\}$$

such that

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

Then,

$$\begin{aligned}
\operatorname{err}(H_s, Q_l)^2 &\leq \sum_{j \geq l} 2^{-2l} 2^{-j(2s-1)} \sum_{k \in \nabla_j} \left\{ \mathbf{1}_{I_k^j}(2^{-l} \cdot \kappa' + 2^{-(l+1)}) \right\}^2 \\
&\leq \sum_{j \geq l} 2^{-2l} 2^{-j(2s-1)} \sum_{\kappa' \in \nabla_l} \left\{ \mathbf{1}_{I_{\kappa'}^j}(2^{-l} \cdot \kappa' + 2^{-(l+1)}) \right\}^2 \\
&\leq \sum_{j \geq l} 2^{-2l} 2^{-j(2s-1)} |\nabla_l| = \sum_{j \geq l} 2^{-l} 2^{-j(2s-1)} \\
&= 2^{-l2s} \sum_{j \geq 0} 2^{j(1-2s)} = \frac{2^{-l2s}}{1 - 2^{1-2s}}.
\end{aligned}$$

□

The question of optimality of such a simple quadrature rule can also respond as elementary as the upper bound. The following lower bound for the quadrature on H_s shows that the chosen quadrature is an asymptotically optimal one.

Lemma 3.1.7. *Let $s > 1/2$ and the space H_s defined as above, we get for every quadrature Q_N used N sample points, that*

$$\operatorname{err}(H_s, Q_N) \geq CN^{-s}.$$

Proof. Let $P \subset [0,1]$, $|P| = N$ be a set of sample points used by an arbitrary quadrature rule Q_N. We choose the integer l that satisfies

$$2^{l-1} \leq 2N \leq 2^l$$

and define a function

$$f_l(x) = \begin{cases} 1 & \text{for all } x \in I_k^l,\ k \in \nabla_l \quad \text{with} \quad I_k^l \cap P = \emptyset \\ 0 & \text{else} \end{cases} .$$

Hence, by some elementary calculations we get

$$\begin{aligned} \|f_l\|_s^2 &= \sum_{j \geq -1} \sum_{k \in \nabla_j} 2^{j2s} \langle f_l, \psi_{j,k} \rangle^2 \\ &= \sum_{j=-1}^{l-1} \sum_{k \in \nabla_j} 2^{j2s} \langle f_l, \psi_{j,k} \rangle^2 \\ &\leq \sum_{j=-1}^{l-1} \sum_{k \in \nabla_j} 2^{j2s} \|\psi_{j,k}\|_\infty^2 \cdot \mathrm{vol}(I_k^j)^2 \\ &= \sum_{j=-1}^{l-1} \sum_{k \in \nabla_j} 2^{j2s} 2^j \|\psi\|_\infty^2 \cdot 2^{-2j} \\ &= \sum_{j=-1}^{l-1} 2^{j2s} \leq C 2^{l2s}. \end{aligned}$$

Furthermore, we get

$$\begin{aligned} \int_0^1 f_l dx &= \int_0^1 \begin{cases} 1 & \text{for all } x \in I_k^l,\ k \in \nabla_l \quad \text{with} \quad I_k^l \cap P = \emptyset \\ 0 & \text{else} \end{cases} dx \\ &\geq \mathrm{vol}(I_k^l)(|\nabla_l| - |P|) = 2^{-l}(2^l - N) = 1 - 2^{-l}N \geq 1/2. \end{aligned}$$

Now, we consider the normalized function $f_l / \|f\|_s$ and get

$$\mathrm{err}(f_l / \|f\|_s, Q_N) \geq C 2^{-ls}.$$

□

The remaining and still open demand is the question on the nature of the considered function spaces H_s. And indeed, this type of function spaces defined by the discrete norms via wavelets are under some proper requirements, equivalent to well known classical function spaces. The close connection between classical function spaces like Sobolev spaces are given by the so-called Jackson and Bernstein inequalities.

3.1.1 Jackson and Bernstein estimates, norm equivalences

As mentioned in the introduction of this chapter and also in the previous section there are some difficulties to generate a (orthogonal) MRA in intervals. The later generalization of the Haar basis called multiwavelet basis is more or less straightforward, but in a more general context biorthogonality is useful. As mentioned before, there are correlations between weighted wavelet series and classical norms that are fundamental for later analysis. The first step to get norm equivalence of the discrete norms defined by wavelet series and classical norms like Sobolev or Besov norms is to establish estimates of the Jackson and Bernstein type (direct and inverse estimates). Afterwords the well known equivalence of the standard L^2-modulus of continuity and K-functional yield to the required equivalence see, e.g. [40]. For a detailed discussion we refer the reader to [12, 17, 58] and the further reading literature mentioned therein.

To get the equivalences in a preferably general form we start with a brief introduction in biorthogonal MRA on intervals, also let us consider the interval $[0, 1]$. Then, in general a biorthogonal MRA consists of two nested families of finite dimensional subspaces

$$V_0 \subset V_1 \subset \ldots \subset V_j \subset V_{j+1} \subset \ldots \subset \mathrm{L}^2([0,1]),$$
$$\widetilde{V}_0 \subset \widetilde{V}_1 \subset \ldots \subset \widetilde{V}_j \subset \widetilde{V}_{j+1} \subset \ldots \subset \mathrm{L}^2([0,1]),$$

such that also $\dim V_j \sim \dim \widetilde{V}_j \sim 2^j$ and

$$\overline{\bigcup_{j\geq 0} V_j} = \mathrm{L}^2([0,1]) = \overline{\bigcup_{j\geq 0} \widetilde{V}_j} = \mathrm{L}^2([0,1]).$$

The spaces V_j and $\widetilde{V}_j$ are also generated by biorthogonal single scale bases

$$V_j = \text{span } \{\varphi_{j,k}\ ,\ k \in \Delta_j\}, \quad \widetilde{V}_j = \text{span } \left\{\widetilde{\varphi}_{j,k}\ ,\ k \in \tilde{\Delta}_j\right\}$$

and

$$\langle \varphi_{j,k}, \widetilde{\varphi}_{j,k'} \rangle = \delta_{k,k'},$$

where Δ_j and $\widetilde{\Delta}_j$ denotes suitable index sets with cardinality $|\Delta_j| \sim \left|\widetilde{\Delta}_j\right| \sim 2^j$. Sometimes it is useful to understand the basis $\Phi_j := [\varphi_{j,k}]_{k\in\Delta_j}$ as a row vector. A final requirement we make is that these bases are uniformly stable (Riesz basis property) see, e.g. [15]: for any vector $\mathbf{c} \in \ell^2(\Delta_j)$ the following holds

$$\|\Phi_j \mathbf{c}\| \sim \left\|\widetilde{\Phi}_j \mathbf{c}\right\| \sim \|\mathbf{c}\|_{\ell^2(\Delta_j)}\,.$$

In the next definition we introduce the so called Jackson and Bernstein-inequalities which are important for our purpose to get norm equivalence.

Definition 3.1.8. We say that Jackson and Bernstein estimates holds for approximation spaces V_j if for $s \leq t \leq d$, $s \leq t \leq \gamma$

$$\inf_{v \in V_j} \|u - v\|_{H^s([0,1])} \lesssim 2^{j(s-t)} \|u\|_{H^t([0,1])}, \quad u \in H^t([0,1])$$

and

$$\|v\|_{H^t([0,1])} \lesssim 2^{j(t-s)} \|v\|_{H^s([0,1])}, \quad v \in V_j,$$

where d, γ are fixed constants given by

$$\begin{aligned} d &= \sup \left\{ s \in \mathbb{R} : \inf_{v \in V_j} \|u - v\| \leq 2^{-js} \|u\|_{H^s([0,1])} \right\}, \\ \gamma &= \sup \left\{ s \in \mathbb{R} : V_j \subset H^s([0,1]) \right\}. \end{aligned}$$

Usually d is the maximal degree of polynomial exactness of V_j or in other words the maximal degree polynomials that are contained in V_j. The second parameter γ denotes the regularity or smoothness of the functions contained in V_j. The biorthogonal wavelets

$$\Psi_j = [\psi_{j,k}]_{k \in \nabla_j}, \quad \widetilde{\Psi}_j = [\widetilde{\psi}_{j,k}]_{k \in \nabla_j}, \quad \langle \Psi_j, \widetilde{\Psi}_j \rangle_{\mathrm{L}^2([0,1])} = \mathbf{1},$$

where $\nabla_j = \Delta_{j+1} \backslash \Delta_j$ are the bases of uniquely determined complement spaces

$$\begin{aligned} W_j &= \text{span } \Psi_j, \\ \widetilde{W}_j &= \text{span } \widetilde{\Psi}_j, \end{aligned}$$

satisfying

$$\begin{aligned} V_{j+1} = V_j \oplus W_j, \quad V_j \cap W_j = \{0\}, \quad W_j \perp \widetilde{V}_j, \\ \widetilde{V}_{j+1} = \widetilde{V}_j \oplus \widetilde{W}_j, \quad \widetilde{V}_j \cap \widetilde{W}_j = \{0\}, \quad \widetilde{W}_j \perp V_j. \end{aligned}$$

We claim that the primal wavelets $\psi_{j,k}$ are also local with respect to the corresponding scale j, i.e.

$$\text{vol}(\text{supp } \psi_{j,k}) \sim 2^{-j},$$

we will normalize the wavelet such that

$$\|\psi_{j,k}\|_{\mathrm{L}^2(\Omega)} \sim \left\| \widetilde{\psi}_{j,k} \right\|_{\mathrm{L}^2([0,1])} \sim 1.$$

We fix the wavelet bases

$$\Psi = [\Psi_j]_{j \geq -1}, \quad \widetilde{\Psi} = [\widetilde{\Psi}_j]_{j \geq -1},$$

where $\Psi_{-1} := \Phi_0$, $\widetilde{\Psi}_{-1} := \widetilde{\Phi}_0$ are Riesz bases of $\mathrm{L}^2([0,1])$.

Under the assumption that Jackson and Bernstein type estimates holds for both, for the primal and dual multiresolution analysis with given parameters d, γ and $\widetilde{d}, \widetilde{\gamma}$ we get an appropriate tool to switch between the discrete norms and Sobolev norms.

Theorem 3.1.9. *Under the above requirements the following norm equivalence hold*

$$\|f\|_{H^s([0,1])} \sim \sum_{j \geq -1} \sum_{k \in \nabla_j} 2^{j2s} \left| \langle f, \widetilde{\psi}_{j,k} \rangle \right|^2, \quad \textit{for } s \in (-\widetilde{\gamma}, \gamma),$$

$$\|f\|_{H^s([0,1])} \sim \sum_{j \geq -1} \sum_{k \in \nabla_j} 2^{j2s} \left| \langle f, \psi_{j,k} \rangle \right|^2, \quad \textit{for } s \in (-\gamma, \widetilde{\gamma}).$$

For the proof see, e.g. [15, 61]. Later, we consider the case of non-stationary MRA that yields to a tight wavelet frame, for this case we will go into a more detailed discussion. Observe for $s = 0$ the norm equivalence implies the Riesz property of wavelet bases.

3.2 Cubature on Haar wavelet spaces

After this short excursion we resume the discussion of quadrature or in higher dimension often called cubature on Haar wavelet spaces. Let us recall, that on one-dimensional Haar wavelet spaces we get an (asymptotic) optimal quadrature by a simple midpoint rule. We recapitulate this fact by the following corollary.

Corollary 3.2.1. *Let $s > 1/2$ and H_s defined by (3.1.1) for*

$$Q_l[f] = \sum_{k=0}^{2^l - 1} f(2^{-l} \cdot k + 2^{-(l+1)})$$

we get

$$\mathrm{err}(H_s, Q_l) \sim 2^{-ls}.$$

Remark 3.2.2. In the Haar wavelet case the discrete norms only corresponds to classical Sobolev spaces for a relatively small smoothness parameter $s < 1/2$. Hence, there is no hope to get norm equivalence for the interesting case where $s > 1/2$. But we can assure by the so called approximation property that the discrete norms defined by Haar wavelet series can be upper bounded by Sobolev norms up to a smoothness parameter $s < 1$,

$$\sum_{j \geq -1} \sum_{k \in \nabla_j} 2^{j2s} \left| \langle f, \psi_{j,k} \rangle \right|^2 \lesssim \|f\|^2_{H^s([0,1])} \, .,$$

and consequently $H^s([0,1]) \subset H_s$.

This problem will be dealt with when we consider multivariate integration problems of functions with moderate smoothness and this justifies among other things the later generalization of the Haar bases the the so called multiwavelet bases.

However, let us carry on with the intuitively Haar wavelet case in two dimensions. One of the reasons we focus on the Haar case is the close connection between the integration problem on these spaces and the notion of classical geometric discrepancy. A simple way to obtain a multivariate wavelet bases is, as mentioned before, the tensor product strategy. To fix this strategy we consider a wavelet bases for the space $\mathrm{L}^2([0,1]^2)$. The complement space $W_0^{(2)}$ of $V_0^{(2)}$ in $V_1^{(2)}$ is similarly generated by the translates of three functions:

$$\varphi \otimes \psi,\ \psi \otimes \varphi,\ \psi \otimes \psi,$$

generally

$$V_{j+1}^{(2)} = V_{j+1} \otimes V_{j+1} = (V_j \oplus W_j) \otimes (V_j \oplus W_j) = \dots.$$

But notice, there are also several other extensions to higher dimensional wavelets. For a given multi-index $\mathbf{j} \in \mathbb{Z}^2$ we put $|\mathbf{j}| := j_1 + j_2$. With purpose in mind to analysis cubature rules in the two-dimensional case we consider the approximation spaces defined by

$$V_L^{(2)} := \sum_{|\mathbf{j}|=L} \bigotimes_{i=1}^{2} V_{j_i}.$$

Similar to the one-dimensional case we put

$$V^{(2)} := \bigcup_{L=0}^{\infty} V_L^{(2)}.$$

Since the union of the one-dimensional approximation spaces are dense in $\mathrm{L}^2([0,1])$, the space $V^{(2)}$ is dense in $\mathrm{L}^2([0,1]^2)$. Thus, we obtain the following expansion for $f \in \mathrm{L}^2([0,1]^2)$

$$f = \sum_{\mathbf{j} \geq -1} \sum_{\mathbf{k} \in \nabla_{\mathbf{j}}} \langle f, \Psi_{\mathbf{j},\mathbf{k}} \rangle \Psi_{\mathbf{j},\mathbf{k}},$$

where $\mathbf{j} = (j_1, j_2) \geq -1$ is meant in the way that $j_u \geq -1$ for $u = 1, 2$. (In the following all inequalities between vectors and between a vector and a scalar are meant componentwise.) Furthermore, we use the shorthands $\nabla_{\mathbf{j}} := \nabla_{j_1} \times \nabla_{j_2}$ and

$$\Psi_{\mathbf{j},\mathbf{k}} := \psi_{j_1,k_1} \otimes \psi_{j_2,k_2}.$$

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

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

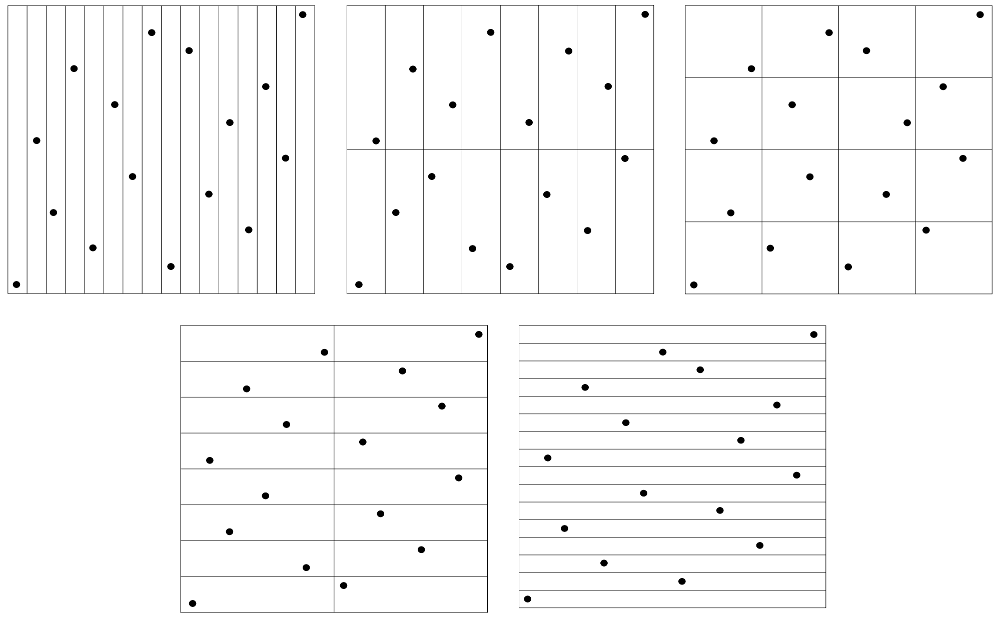

Figure 3.1: Illustration of a binary net with 1 point per canonical interval of volume 2^{-4}.

then supp $\Psi_{\mathbf{j},\mathbf{k}} = I_{\mathbf{k}}^{\mathbf{j}}$ holds. Obviously, cubature rules that contains exactly one sample point in every canonical interval of a fixed volume,

$$I_{\mathbf{k}}^{\mathbf{j}}, \text{ with } |\mathbf{j}| = L, \text{ and } \mathbf{k} \in \nabla_{\mathbf{j}},$$

are exact on $V_L^{(2)}$. This leads to the definition of so-called binary nets, or in higher dimensions the b-ary nets see, e.g. [43, 47] for more details. In figure 3.1 we show a simple example of binary nets. The corresponding cubature is a quasi Monte-Carlo method and given by

$$Q_L[f] = \frac{1}{|P_{net}|} \sum_{p \in P_{net}} f(p).$$

Lemma 3.2.3. *The cubature Q_L based on binary net with 2^L sample points (with 1 point per canonical interval on level L) is exact for the approximation space $V_L^{(2)}$.*

For error analysis we consider product spaces which are based on the one-dimensional function spaces H_s used for the one-dimensional quadrature error bounds. More detailed, we consider

$$H_s^2 = H_s \otimes H_s,$$

with $s > 1/2$ and

$$\|f_1 \otimes f_2\|_s = \|f_1\|_s \cdot \|f_2\|_s .$$

Hence, we also get

$$H_s^2 = \left\{ f \in \mathrm{L}^2([0,1]^2) : \sum_{\mathbf{j} \geq -1} \sum_{\mathbf{k} \in \nabla_{\mathbf{j}}} 2^{|\mathbf{j}|2s} \left|\langle f, \Psi_{\mathbf{j},\mathbf{k}} \rangle\right|^2 < \infty \right\}.$$

Therefore, the cubature error on H_s^2 can be upper bounded by the following elementary theorem.

Theorem 3.2.4. *Let $s > 1/2$ and Q_L a cubature based on binary nets, then*

$$\mathrm{err}(H_s^2, Q_L) \lesssim 2^{-Ls}(L+1)^{1/2}.$$

Proof. Let $f \in H_s^2$ and $s > 1/2$, we get again by the two dimensional wavelet expansion and Cauchy-Schwarz inequality

$$\mathrm{err}(f, Q_L) \leq \|f\|_s \left(\sum_{\mathbf{j} \geq -1} \sum_{\mathbf{k} \in \nabla_{\mathbf{j}}} 2^{-|j|2s} \left\{ I[\Psi_{\mathbf{j},\mathbf{k}}] - Q_L[\Psi_{\mathbf{j},\mathbf{k}}] \right\}^2 \right)^{1/2} .$$

So, by the exactness of Q_L

$$\begin{aligned}
\mathrm{err}(H_s^2, Q_L)^2 &\leq \sum_{|\mathbf{j}| \geq L} \sum_{\mathbf{k} \in \nabla_{\mathbf{j}}} 2^{-|j|2s} \left\{ Q_L[\Psi_{\mathbf{j},\mathbf{k}}] \right\}^2 \\
&= \sum_{|\mathbf{j}| \geq L} \sum_{\mathbf{k} \in \nabla_{\mathbf{j}}} 2^{-|j|2s} \left\{ 2^{-L} \sum_{p \in P_{net}} \Psi_{\mathbf{j},\mathbf{k}}(p) \right\}^2 \\
&= 2^{-2L} \sum_{|\mathbf{j}| \geq L} \sum_{\mathbf{k} \in \nabla_{\mathbf{j}}} 2^{-|j|2s} \left\{ \sum_{p \in P_{net}} 2^{|\mathbf{j}|/2} \psi(2^{j_1} \cdot - k_1) \otimes \psi(2^{j_2} \cdot - k_2)(p) \right\}^2 \\
&= 2^{-2L} \sum_{|\mathbf{j}| \geq L} 2^{|j|(1-2s)} \sum_{\mathbf{k} \in \nabla_{\mathbf{j}}} \left\{ \sum_{p \in P_{net}} \psi(2^{j_1} \cdot - k_1) \otimes \psi(2^{j_2} \cdot - k_2)(p) \right\}^2 .
\end{aligned}$$

Thus, we get with similar arguments as in the one-dimensional case,

$$\begin{aligned}
\mathrm{err}(H_s^2, Q_L)^2 &\leq 2^{-2L} \sum_{|\mathbf{j}| \geq L} 2^{|j|(1-2s)} \sum_{\mathbf{k} \in \nabla_{\mathbf{j}}} \left\{ \sum_{p \in P_{net}} \mathbf{1}_{I_{k_1}^{j_1}} \otimes \mathbf{1}_{I_{k_2}^{j_2}}(p) \right\}^2 \\
&\leq 2^{-2L} \sum_{|\mathbf{j}| \geq L} 2^{|j|(1-2s)} 2^L = 2^{-L} \sum_{j \geq L} 2^{j(1-2s)}(L+1) \\
&= 2^{-L2s} \sum_{j \geq 0} 2^{j(1-2s)}(L+1) = \frac{1}{1 - 2^{1-2s}} 2^{-L2s}(L+1).
\end{aligned}$$

□

Remark 3.2.5. In the two-dimensional case the cubature via binary nets is optimal. The lower bound proof is more or less straightforward and we skip them since in the next chapter we will consider more general cases. The generalization of these Haar wavelet cubature to higher dimensions is not straightforward and we refer the interested reader to [37]. Here, the authors make use of a randomized variation of low discrepancy point sets see, e.g. [51]. Another technical difficulty is the fact that the authors have to deal with MRA based on b-adic Haar wavelets and use analogies to the so called ANOVA (analysis of variance) decomposition see, e.g. [51, 52].

At the end of the next chapter we will pick up this example of two-dimensional cubature via nets and make a comparison to another method that makes use of optimal one-dimensional quadrature rules to construct an d-dimensional cubature by cleverly chosen tensor products. It turns out that these tensor product methods have the same order of convergence but use more sample points and consequently they do not have the optimal logarithmic order such as the cubature via nets. But we have to point out that these tensor product methods are relatively simple to implement.

Chapter 4

Numerical Integration Based on Smolyak's Construction

In [65] Smolyak studied special tensor product problems and introduced a general construction that leads to almost optimal (up to logarithmic factors) approximation for any dimension $d > 1$ starting from optimal approximations for the one-dimensional case. By now this construction is also known under different names, as e.g. "Birmann interpolation", "Boolean method", "discrete blending method", "hyperbolic cross points" and "sparse grid method". Further results and interesting application of Smolyak's construction can be found e.g. in [34, 48, 49, 67]. It is worth mentioning that the article [34] contains a bibliography including 56 references from the most significant work in this area (for a comprehensive survey article on this topic see [10]). In this chapter we are mainly interested in the construction of cubature rules via Smolyak's construction, thus we formulate Smolyak's construction not in the most general form.

This chapter is organized as follows: In the first section we give a brief review of the algorithm, that can overcome the "curse of dimension" and afterwords in Section 4.2 we discuss some negative results on this construction. The negative results are based on the not very surprising fact the the $\mathcal{L}_2$- respectively the $\mathcal{L}_\infty$-discrepancy of a given sparse grid leads to very poor error estimates via Koksma Hlawka inequalities. We give an elementary example of Smolyak's construction on two-dimensional Haar wavelet spaces and compare this algorithm with the optimal cubature based on nets. At the end of this chapter in Section 4.3 we will go into a more detailed discussion on the complexity of the algorithm and also consider the complexity of a variation of this algorithm. It is not hard to examine that the complexity of the algorithm is mainly based on the number of function evaluations which leads to an elementary problem of counting lattice points in a given simplex. In the following discussion take a closer look of a so called anisotropic version of Smolyak's construction which seems to be useful for high dimensional problems. This anisotropic algorithm uses so called a-priori knowledge of the considered function and applies less costly quadratures in smoother direction and therefore the overall cost is reduced.

4.1 Smolyak's construction

We start by giving a brief description on Smolyak's construction for multivariate numerical integration. But observe, we explain Smolyak's construction not in the most general form. However, let H be a one-dimensional function space defined for simplicity over the unit cube with given norm $\|\cdot\|_H$. And for the multivariate case we consider the tensor product

$$H^{(d)} := \underbrace{H \otimes H \otimes \ldots \otimes H}_{d \text{ times}}.$$

This means finite linear combinations of functions

$$(f_1 \otimes \ldots \otimes f_d)(t_1, \ldots t_d) = f_1(t_1) \cdot \ldots \cdot f_d(t_d)$$

with $f_i \in H$ are dense in $H^{(d)}$ and for the norm we get

$$\|f_1 \otimes \ldots \otimes f_d\|_{H^{(d)}} = \|f_1\|_H \|f_2\|_H \ldots \|f_d\|_H .$$

For example consider the one-dimensional function space of smooth function

$$F_1^r = C^r([0,1]), \quad r \in \mathbb{N}$$

with the norm

$$\|f\| = \max\left\{\|f\|_\infty, \ldots, \left\|f^{(r)}\right\|_\infty\right\}.$$

For $d > 1$ consider the tensor product space that is often regard for error bounds of variants of Smolyak's construction

$$F_d^r = \{f : [0,1] \to \mathbb{R} : \ D^\alpha f \text{ continuous if } \alpha_i \le r \text{ for all } i\} \tag{4.1.1}$$

with the norm

$$\|f\| = \max\left\{\|D^\alpha f\|_\infty : \ \alpha \in \mathbb{N}_0^d, \ \alpha_i \le r\right\}.$$

The main idea to extend a one-dimensional algorithm $A(l, 1)$ to a d-dimensional cubature is to choose clever tensor products instead of the full tensor product rule, we start by writing the quadrature formula as a telescope sum. The so called difference quadrature for level $l \ge 0$ is defined by

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

with $A(-1, 1) := 0$. For a given multi-index $\mathbf{l} \in \mathbb{N}_0^d$ we put $|\mathbf{l}| = l_1 + l_2 + \ldots + l_d$. Smolyak's construction of level L is then given by

$$A(L, d) \ := \sum_{\mathbf{l} \in \mathbb{N}_0^d, \ |\mathbf{l}| \le L} (\Delta^{l_1} \otimes \Delta^{l_2} \otimes \cdots \otimes \Delta^{l_d}). \tag{4.1.2}$$

Remark 4.1.1. Observe that we use a slightly modification of Smolyak's construction. In the literature you often find the notation $A(0,1) = 1$. But with the wavelet spaces in mind our intention is that the quadrature on level 0, $A(0,1)$ is exact on the approximation space V_0. Thus, this notation agrees with the notation we have used in the previous chapter for the definition of two-dimensional approximation spaces.

It is well known that in general Smolyak's cubature can be written in terms of the one-dimensional quadrature instead of the difference quadrature, hence we get a formula

$$A(L,d) = \sum_{L-d+1 \leq |\mathbf{l}| \leq L} (-1)^{L-|\mathbf{l}|} \binom{d-1}{L-|\mathbf{l}|} \bigotimes_{k=1}^{d} A(l_k,1). \tag{4.1.3}$$

Obviously, the one-dimensional quadrature $A(l,1)[f]$ depends on f only through the function values at a finite number of points called the sample points. Let us denote this points by

$$X^l = \left\{x_1^l, x_2^l, \ldots, x_{m_l}^l\right\} \subset [0,1].$$

The tensor product algorithm in (4.1.2) alternatively (4.1.3) is based on the sparse grid

$$X(L,d) := \bigcup_{L-d+1 \leq |\mathbf{l}| \leq L} (X^{l_1} \times X^{l_2} \times \ldots \times X^{l_d}) \subset [0,1]^d.$$

Nested sets $X^i \subset X^{i+1}$ yields to $X(L,d) \subset X(L,d+1)$ and consequently the differential grid is given by

$$X(L,d) = \bigcup_{|\mathbf{l}|=L} (X^{l_1} \times X^{l_2} \times \ldots \times X^{l_d}) \subset [0,1]^d.$$

Therefore, starting with nested sets seems to be the most economical choice for cubature rules based on those constructions.

Remark 4.1.2. The intuition we have in the later context can be described as follows. We have an expansion of a function f where the basis functions (wavelets) have small support if the index arise (locality conditions) and consequently we have a small part in the representation of f. Furthermore, it can be shown that under suitable conditions the approximation properties of a function f by its sparse grid representation is almost as good as the full grid representation of f see, e.g. [10].

Apart from our intuition there are two different ways to justify the choice of this special index set. A continuous one is based on an analytical approach where the multi-index $\mathbf{l}$ is generalized to a non-negative real one, and a discrete one which uses techniques from combinatorial optimization. And since the continuously way to generalized multi-indices seems to be bit unnatural we restrict our formulation to the discrete optimization problem. A more detailed discussion on this optimization problem can be found in [10], but for reasons of a close representation we recall the main idea. Therefore, we ask how to construct multivariate approximation spaces V^{opt} (as we had considered in the previous section) which have better properties as the full tensor product approximation spaces V^{∞}. More precise we ask how to choose V^{opt} such that the same number of invested grid points or sample points leads to a higher order of accuracy.

First of all, let us recall the setting of a general biorthogonal wavelet basis generate by a biorthogonal MRA we have considered in Chapter 3 Section 3.1.1. Thus let

$$\begin{aligned} V_0 \subset V_1 \subset \ldots \subset V_j \subset V_{j+1} \subset \ldots \subset \mathrm{L}^2([0,1]), \\ \widetilde{V}_0 \subset \widetilde{V}_1 \subset \ldots \subset \widetilde{V}_j \subset \widetilde{V}_{j+1} \subset \ldots \subset \mathrm{L}^2([0,1]), \end{aligned}$$

be the two families of finite dimensional subspaces spanned by a biorthogonal single scale bases

$$V_j = \text{span } \{\varphi_{j,k} \ , \ k \in \Delta_j\}, \quad \widetilde{V}_j = \text{span } \left\{\widetilde{\varphi}_{j,k} \ , \ k \in \tilde{\Delta}_j\right\}.$$

The biorthogonal wavelets

$$\Psi_j = [\psi_{j,k}]_{k\in\nabla_j}, \quad \widetilde{\Psi}_j = [\widetilde{\psi}_{j,k}]_{k\in\nabla_j}, \quad \langle \Psi_j, \widetilde{\Psi}_j\rangle_{\mathrm{L}^2([0,1])} = \mathbf{1},$$

where $\nabla_j = \Delta_{j+1}\backslash\Delta_j$ are the bases of uniquely determined complement spaces

$$\begin{aligned} W_j &= \text{span } \Psi_j, \\ \widetilde{W}_j &= \text{span } \widetilde{\Psi}_j, \end{aligned}$$

satisfying

$$\begin{aligned} V_{j+1} = V_j \oplus W_j, \quad V_j \cap W_j = \{0\}, \quad W_j \bot \widetilde{V}_j, \\ \widetilde{V}_{j+1} = \widetilde{V}_j \oplus \widetilde{W}_j, \quad \widetilde{V}_j \cap \widetilde{W}_j = \{0\}, \quad \widetilde{W}_j \bot V_j. \end{aligned}$$

Finally, we assume this setting satisfies all useful properties we point out in the previous chapter. For reasons of a closer presentation let us use slightly modified notations. We define $W_{-1} := V_0$ and respectively $\widetilde{W}_{-1} := \widetilde{V}_0$ with little displaces of the indices, $W_l^0 := W_{l-1}$ and also $\widetilde{W}_l^0 := \widetilde{W}_{l-1}$ we get the decomposition

$$V_j = W_0^0 \oplus W_1^0 \oplus \ldots \oplus W_{j-1}^0 \oplus W_j^0$$

and respectively for the dual

$$\widetilde{V}_j = \widetilde{W}_0^0 \oplus \widetilde{W}_1^0 \oplus \ldots \oplus \widetilde{W}_{j-1}^0 \oplus \widetilde{W}_j^0.$$

Thus, the conventional sparse grid can be written as

$$V_L^1 = \bigoplus_{|\mathbf{l}| \leq L} \bigotimes_{u=1}^{d} W_{l_u}^0 \quad \text{and the full grid } V_L^\infty = \bigoplus_{|\mathbf{l}|_\infty \leq L} \bigotimes_{u=1}^{d} W_{l_u}^0,$$

respectively

$$\widetilde{V}_L^1 = \bigoplus_{|\mathbf{l}| \leq L} \bigotimes_{u=1}^{d} \widetilde{W}_{l_u}^0 \quad \text{and the full grid } \widetilde{V}_L^\infty = \bigoplus_{|\mathbf{l}|_\infty \leq L} \bigotimes_{u=1}^{d} \widetilde{W}_{l_u}^0.$$

We consider the tensor product $H_s^{(d)} = H_s \otimes \ldots \otimes H_s$, defined as follows. For $s > 0$ we define an inner product similar to the considered Haar wavelet case

$$\langle f, g \rangle_s := \sum_{\mathbf{j} \geq -\mathbf{1}} \sum_{\mathbf{k} \in \nabla_{\mathbf{j}}} 2^{|\mathbf{j}| 2s} \langle f, \widetilde{\psi}_{\mathbf{j},\mathbf{k}} \rangle \langle g, \widetilde{\psi}_{\mathbf{j},\mathbf{k}} \rangle,$$

where $\widetilde{\psi}_{\mathbf{j},\mathbf{k}} = \otimes_{u=1}^d \widetilde{\psi}_{j_u,k_u}$ and a corresponding norm

$$\|f\|_s^2 := \langle f, f \rangle_s.$$

Thus, we define the space

$$H_s^{(d)} = \left\{ f \in \mathrm{L}^2([0,1]^d) : \ \|f\|_s < \infty \right\}.$$

Observe, we get similar definitions for the primal wavelets $\psi_{\mathbf{j},\mathbf{k}} = \otimes_{u=1}^d \psi_{j_u,k_u}$. Hence, in the following we look for an optimum V^{opt} by solving a restricted optimization problem of the type

$$\max_{u \in H_s^{(d)} : \ \|u\|_s = 1} \|u - u_{V^{\mathrm{opt}}}\| = \min_{U \subset V^\infty : \ |U| = n} \ \max_{u \in H_s^{(d)} : \ \|u\|_s = 1} \|u - u_U\| \tag{4.1.4}$$

for some cost n. Thus any potential solution V^{opt} depends on the norm $\|\cdot\|$ as well as the (semi) norm that is used to measure the error of the interpolant of the u or the smoothness of u. According to our later setting we will allow only discrete spaces of the type

$$U := \sum_{\mathbf{l} \in \mathbf{I}} V_{\mathbf{l}} = \bigoplus_{\mathbf{l} \in \mathbf{I}} W_{\mathbf{l}}^0, \tag{4.1.5}$$

where $\mathbf{I} \subset \mathbb{N}_0^d$. Quite similar to the well known approximation property for wavelets we get an upper bound for the contribution $u_{\mathbf{l}} \in W_{\mathbf{l}}^0$ to the wavelet expansion of a multivariate function $u \in H_s^{(d)}$.

Lemma 4.1.3. *Let $u \in H_s^{(d)}$ given in its biorthogonal wavelet expansion. Then*

$$\|u_{\mathbf{l}}\| \lesssim 2^{-|\mathbf{l}| s} \|u\|_s,$$

holds for the components $u_{\mathbf{l}} \in W_{\mathbf{l}}^0$.

Proof. Similar to the one-dimensional Haar wavelet case we have considered in Chapter 3 we define for $s > 0$ a scaled wavelet bases by

$$\psi^s_{\mathbf{j},\mathbf{k}} := 2^{-|\mathbf{l}|s}\psi^s_{\mathbf{j},\mathbf{k}} \quad \text{and respectively } \widetilde{\psi}^s_{\mathbf{j},\mathbf{k}} := 2^{-|\mathbf{l}|s}\widetilde{\psi}_{\mathbf{j},\mathbf{k}},$$

and get an expansion

$$\begin{aligned} u_{\mathbf{l}} &= \sum_{\mathbf{k}\in\nabla_{\mathbf{l}}} \langle u, \widetilde{\psi}_{\mathbf{l},\mathbf{k}}\rangle \psi_{\mathbf{l},\mathbf{k}} = \sum_{\mathbf{k}\in\nabla_{\mathbf{l}}} 2^{|\mathbf{l}|2s} \langle u, \widetilde{\psi}^s_{\mathbf{l},\mathbf{k}}\rangle \psi^s_{\mathbf{l},\mathbf{k}} \\ &= \sum_{\mathbf{k}\in\nabla_{\mathbf{l}}} \langle u, \widetilde{\psi}^s_{\mathbf{l},\mathbf{k}}\rangle_s \psi^s_{\mathbf{l},\mathbf{k}}. \end{aligned}$$

Thus, by Cauchy-Schwarz inequality we get for L^2-norm of the components $u_{\mathbf{l}}$

$$\begin{aligned} \|u_{\mathbf{l}}\|^2 &= \left\| \sum_{\mathbf{k}\in\nabla_{\mathbf{l}}} \langle u, \widetilde{\psi}^s_{\mathbf{l},\mathbf{k}}\rangle_s \psi^s_{\mathbf{l},\mathbf{k}} \right\|^2 \\ &\le \sum_{\mathbf{k}\in\nabla_{\mathbf{l}}} \left| \langle u, \widetilde{\psi}^s_{\mathbf{l},\mathbf{k}}\rangle_s \right|^2 \left\|\psi^s_{\mathbf{l},\mathbf{k}}\right\|^2 = \sum_{\mathbf{k}\in\nabla_{\mathbf{l}}} \left| \langle u, \widetilde{\psi}^s_{\mathbf{l},\mathbf{k}}\rangle_s \right|^2 2^{-|\mathbf{l}|2s} \\ &\le \sum_{\mathbf{k}\in\nabla_{\mathbf{l}}} \left\| u|_{\mathrm{supp}\, \widetilde{\psi}^s_{\mathbf{l},\mathbf{k}}} \right\|^2 2^{-|\mathbf{l}|2s} \lesssim \|u\|_s^2\, 2^{-|\mathbf{l}|2s}. \end{aligned}$$

□

Now, we define the so-called local cost functions $c(\mathbf{l})$ by

$$c(\mathbf{l}) := |V_{\mathbf{l}}|\,.$$

Obviously $c(\mathbf{l}) \in \mathbb{N}$ holds for all $\mathbf{l} \in \mathbb{N}^d$. The local benefit function is given by

$$b(\mathbf{l}) = \gamma\beta(\mathbf{l}),$$

where $\beta(\mathbf{l})$ is the upper bound for $\|u_{\mathbf{l}}\|^2$ we compute in the previous lemma. The factor γ is depending on the dimension and the smoothness of u, but it is constant with respect to the index set $\mathbf{l}$, such that $b(\mathbf{l}) \in \mathbb{N}$. Observe that the possibility of this choice is due to the fact that we are interested in subset of the full grid of type (4.1.5). Next, we define for a sufficiently large N, $\mathbf{I} \subset \mathbf{I}^{\max} := \{1, 2, \ldots, N\}^d$, the global cost function by

$$C(\mathbf{I}) := \sum_{\mathbf{l}\in\mathbf{I}} x(\mathbf{l})c(\mathbf{l}),$$

where

$$x(\mathbf{l}) := \begin{cases} 0 & \text{for all } \mathbf{l} \notin \mathbf{I} \\ 1 & \text{else.} \end{cases}$$

The approximation of u on the considered grid $\mathbf{I}$ provides the global benefit $B(\mathbf{I})$,

$$\begin{aligned}
\left\| u - \sum_{\mathbf{l}\in\mathbf{I}} u_{\mathbf{l}} \right\|^2 &\approx \left\| \sum_{\mathbf{l}\in\mathbf{I}^{\max}} u_{\mathbf{l}} - \sum_{\mathbf{l}\in\mathbf{I}} u_{\mathbf{l}} \right\|^2 = \left\| \sum_{\mathbf{l}\in\mathbf{I}^{\max}\setminus\mathbf{I}} u_{\mathbf{l}} \right\|^2 \\
&\leq \sum_{\mathbf{l}\in\mathbf{I}^{\max}\setminus\mathbf{I}} \|u_{\mathbf{l}}\|^2 = \sum_{\mathbf{l}\in\mathbf{I}^{\max}} (1 - x(\mathbf{l}))\, \|u_{\mathbf{l}}\|^2 \\
&\leq \sum_{\mathbf{l}\in\mathbf{I}^{\max}} (1 - x(\mathbf{l}))\beta(\mathbf{l}) \leq \sum_{\mathbf{l}\in\mathbf{I}^{\max}} (1 - x(\mathbf{l}))\gamma\beta(\mathbf{l}) \\
&= \sum_{\mathbf{l}\in\mathbf{I}^{\max}} \gamma\beta(\mathbf{l}) - \sum_{\mathbf{l}\in\mathbf{I}^{\max}} x(\mathbf{l})\gamma\beta(\mathbf{l}) \\
&=: \sum_{\mathbf{l}\in\mathbf{I}^{\max}} \gamma\beta(\mathbf{l}) - B(\mathbf{l}).
\end{aligned}$$

Of course, the considered upper bound only gives a bound for the approximation with respect to the full grid representation of u. However, since N and consequently $\mathbf{I} \subset \mathbf{I}^{\max} := \{1, 2, \ldots, N\}^d$ can be chosen as large as appropriate this is not a to serious restriction. Thus, we get the following formulation of the optimization problem (4.1.4)

$$\max_{\mathbf{I}\subset\mathbf{I}^{\max}} \sum_{\mathbf{l}\in\mathbf{I}} x(\mathbf{l})\gamma\beta(\mathbf{l}) \quad \text{with} \quad \sum_{\mathbf{l}\in\mathbf{I}} x(\mathbf{l})c(\mathbf{l}) = n.$$

Assume that the indices $\mathbf{I} \in \mathbf{I}^{\max}$ are arranged in lexicographical order with the local cost c_i and benefit b_i for $i = 1, 2, \ldots, N^d$ consequently the optimization problem is given by

$$\max_{x\in\{0,1\}^{N^d}} \mathbf{b}^T\mathbf{x} \quad \text{with} \quad \mathbf{c}^T\mathbf{x} = n,$$

where $\mathbf{b} \in \mathbb{N}^{N^d}$, $\mathbf{c} \in \mathbb{N}^{N^d}$ and without loss of generality $n \in \mathbb{N}$. In combinatorial optimization problems like those are called binary knapsack problems and are known to be NP-hard. But, a slightly change makes things easier. We consider a so called relaxation, e.g. we also allow rational solutions $\mathbf{x} \in ([0,1] \cap \mathbb{Q})^{N^d}$. In this case there exists a very simple algorithm that provides an optimal solution $\mathbf{x} \in ([0,1] \cap \mathbb{Q})^{N^d}$ see [10].

Algorithm :

(1) rearrange the order such that $\frac{b_1}{c_1} \geq \frac{b_2}{c_2} \geq \cdots \geq \frac{b_{N^d}}{c_{N^d}}$,

(2) let $r := \max \left\{ j : \sum_{i=1}^{j} c_i \leq n \right\}$,

(3) $x_1 = \ldots = x_i := 1$,

$x_{r+1} := \frac{n - \sum_{i=1}^{j} c_i}{c_{r+1}}$,

$x_{r+2} = \ldots = x_{N^d} := 0$.

Consequently, there is only one potential non-binary coefficient x_{r+1} in the rational solution. In general, this rational solution has nothing to do with the binary solution we are interested in, but in this special case our knapsack is of variable size, since the global work count n is an arbitrary chosen natural number. Therefore it is possible to force the solution of the relaxation to be a binary one which is also a solution of the original binary problem. Consequently, the optimization problem can be reduced to the question of local cost-benefit ratio b_i/c_i or $b(\mathbf{l})/c(\mathbf{l})$ of the underlying subspaces $W_{\mathbf{l}}^0$. And its seems to be a logical consequence to take into account those subspaces $W_{\mathbf{l}}^0$ with the best cost-benefit ratio first. We define

$$\mathrm{cbr}(\mathbf{l}) := \frac{b(\mathbf{l})}{c(\mathbf{l})} = \frac{\gamma \beta(\mathbf{l})}{|V_{\mathbf{l}}|} = \frac{\gamma 2^{-|\mathbf{l}| 2s} \|u\|_s^2}{2^{|\mathbf{l}|}} = \gamma 2^{-|\mathbf{l}|(2s+1)} \|u\|_s^2 ,$$

as the local cost-benefit ratio. The crucial point is the index set of an optimal grid $\mathbf{I}^{\mathrm{opt}}$ will consist of all indices $\mathbf{l}$ where the local cost-benefit ratio $\mathrm{cbr}(\mathbf{l})$ is bigger than some threshold. We chose this threshold to be of the order of $\mathrm{cbr}(\mathbf{l}^*)$ with $\mathbf{l}^* = (L, 0, \ldots, 0)$, thus we get a threshold

$$\mathrm{cbr}(\mathbf{l}^*) := \frac{b(\mathbf{l}^*)}{c(\mathbf{l}^*)} = \gamma 2^{-|\mathbf{l}^*|(2s+1)} \|u\|_s^2 = \gamma 2^{-L(2s+1)} \|u\|_s^2 .$$

Under the requirement that for a contribution $\mathbf{l}$ to the optimal index set

$$\mathrm{cbr}(\mathbf{l}) \geq \mathrm{cbr}(\mathbf{l}^*) = \gamma 2^{-L(2s+1)} \|u\|_s^2$$

we get the relation

$$|\mathbf{l}| \leq L.$$

And this finally leads to the well known definition of the standard sparse grids

$$V_L^{(1)} = \bigoplus_{|\mathbf{l}| \leq L} \bigotimes_{u=1}^{d} W_{l_u}^0 .$$

Remark 4.1.4. Let us again mention that we only recalled a slightly modification of the methods present in [10] in the special case where the approximation error is measured in the L^2-norm. Observe other norms in (4.1.4) yields to other sparse grids.

In the next section we will go into a more detailed discussion of restriction to allow only discrete spaces of the type (4.1.5) when we consider an integration problem. It turns out, this restriction is in some cases not the optimal one.

4.2 Smolyak's construction and discrepancy

The question of suitable measure of the precision or in other words the quality of a given cubature based on Smolyak's construction remains. In the previous chapters we have seen that error bounds of the Koksma-Hlawka type or generalized Koksma-Hlawka type inequalities seem to be a conventional measure of cubature errors. But it turns out that a more or less natural measure has to depend on the smoothness of the integrand. Nevertheless, with the classical Koksma-Hlawka inequality in mind, it would also be of interest to compute the $\mathcal{L}_2$- or $\mathcal{L}_\infty$-discrepancy of a given Smolyak construction. The $\mathcal{L}_2$-discrepancy is often used to compare the quality of multivariate cubature such as quasi Monte Carlo methods or (pseudo) Monte Carlo methods consequently it seems to be a suitable measure to compare Smolyak's construction and quasi Monte Carlo methods. But the results see, e.g. [31, 55] are very poor. However, since Smolyak's construction leads to good performance with respect to Monte Carlo or quasi Monte Carlo methods, especially if the integrand is smooth, it would be the more suitable choice to consider measures of the cubature error that take into account the smoothness of the integrand. If we consider e.g. the function class (4.1.1) Novak and Ritter proved in [48] that the error of Smolyak's construction based on Clenshaw-Curtis rule can be upper bounded by

$$\mathcal{O}(N^{-r}(\log N)^{(d-1)(r+1)}).$$

However, let us recall the example of two-dimension Haar wavelet spaces from Chapter 3 and the error bound on H_s^2 for a cubature based on binary nets,

$$\mathrm{err}(H_s^2, Q_L) \lesssim 2^{-Ls}(L+1)^{1/2}, \quad \text{for } s > 1/2.$$

In figure 4.1 we show both, the sample points based on binary nets and the competing Smolyak construction with the same level of exactness, with respect to the approximation spaces. As we will seen in the next chapter the cubature error for our variation of Smolyak's construction can be upper bounded by

$$\mathrm{err}(H_s^2, A(L,2)) \lesssim 2^{-Ls}(L+1).$$

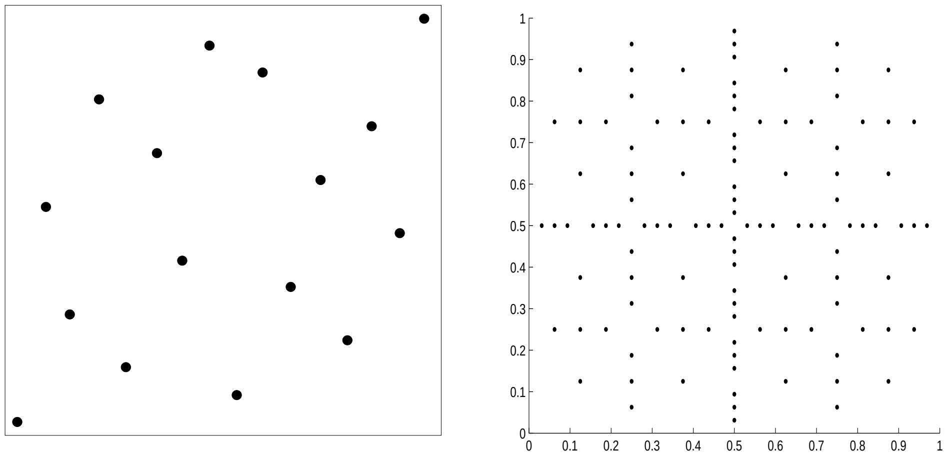

Figure 4.1: Binary net with worst case error $\mathcal{O}(2^{-4s}\sqrt{5})$ and the competing Smolyak construction.

If we consider the cost of the cubature algorithm $A(L,2)$ via formula (4.1.3) we will get

$$N := |A(L,2)| \lesssim 2^L(L+1)$$

and consequently an error bound

$$\mathrm{err}(H_s^2, A(L,2)) \lesssim \frac{\log(N)^{s+1}}{N^s},$$

instead of the optimal worst case error bound we get for the cubature based on nets

$$\mathrm{err}(H_s^2, Q_L) \lesssim \frac{\log(N)^{1/2}}{N^s}.$$

Remark 4.2.1. This comparison also shows that the term optimality used for the formulation of the optimization problem which yields to the sparse approximation spaces has to be used with the appropriate care. The cubature rule based on Smolyak's construction is an optimal one regarding tensor product rules that is based on optimal one-dimensional quadrature rules. Another important point is that the nets which lead to optimal cubature rules for Haar wavelet spaces are also known under the characteristic name "small discrepancy point sets".

In the next section we will go into a more detailed discussion on the complexity of the considered algorithms. Obviously, the problem to compute the cost of a cubature based on Smolyak's construction can be reduced to the problem of lattice point counting. More detailed, for the so far considered cases where the clever chosen tensor product is easily given by the standard simplex or respectively the ℓ^1-norm of the index set, this yields to the question of lattice points inside a given simplex.

4.3 Lattice point counting

The cost of algorithms based on Smolyak's or the sparse grid construction on level L is directly related to the cardinality of the underling index set given by

$$|\mathbf{l}| \leq L, \quad \text{for } \mathbf{l} \in \mathbb{N}_0^d.$$

More generally, it is of interest to consider $\mathbb{Z}^d$, the d-dimensional integer lattice in $\mathbb{R}^d$ and $\mathcal{P}$ an d-dimensional lattice polytope also in $\mathbb{R}^d$. Consider the function $\mathcal{L} : \mathbb{R}^d \times \mathbb{N} \to \mathbb{N}$ that describes the number of lattice points inside the dilated polytope $t\mathcal{P} = \{tx : \ x \in \mathcal{P}\}$,

$$\mathcal{L}(\mathcal{P}, t) = \left|\{t\mathcal{P}\} \cap \mathbb{Z}^d\right|.$$

A systematical study of properties of these functions have was introduced by Ehrhart. In [29] he proved that this function is always a polynomial in t (Ehrhart polynomials) and in fact

$$\mathcal{L}(\mathcal{P}, t) = \mathrm{vol}(\mathcal{P})t^d + \frac{1}{2}\mathrm{vol}(\partial\mathcal{P})t^{d-1} + \ldots + \chi(\mathcal{P}),$$

where $\chi(\mathcal{P})$ denotes the Euler characteristic of the closed polytope $\mathcal{P}$ and $\mathrm{vol}(\partial P)$ is the surface area of $\mathcal{P}$ with respect to the sub-lattice on each face of $\mathcal{P}$. But since the other coefficients are still open and to compute this coefficients is a nice task and a purpose of intensively study, this theory seems to be not useful from the computational point of view.

However, we are mainly interested in the special case where the polytope is given by a (standard) d-simplex Δ,

$$\Delta = \left\{(x_1, x_2, \ldots x_d) \in \mathbb{R}^d : \ x_1 + x_2 + \ldots + x_d \leq 1 \text{ and all } x_k \geq 0\right\}.$$

In this elementary case we get

$$\begin{aligned}
\mathcal{L}(\Delta, t) &= \left|\left\{(m_1, m_2, \ldots, m_d) \in \mathbb{Z}_{\geq 0}^d : \ m_1 + m_2 + \ldots + m_d \leq t\right\}\right| \\
&= \left|\left\{(m_1, m_2, \ldots, m_d, m_{d+1}) \in \mathbb{Z}_{\geq 0}^{d+1} : \ m_1 + m_2 + \ldots + m_d + m_{d+1} = t\right\}\right| \\
&= \binom{d+t}{d}.
\end{aligned}$$

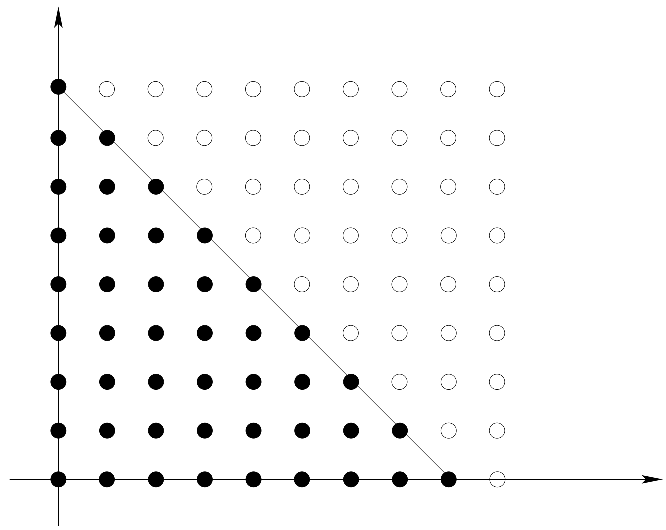

Figure 4.2: Simplex in dimension two: $\left|t\Delta \cap \mathbb{Z}^d_{\geq 0}\right| = \frac{(t+1)(t+2)}{2}$.

The proof is via induction and based on the fact

$$\begin{aligned} & \left|\left\{(m_1, m_2, \ldots, m_d, m_{d+1}) \in \mathbb{Z}^{d+1}_{\geq 0} :\ m_1 + m_2 + \ldots + m_d + m_{d+1} = t\right\}\right| \\ = & \sum_{m_{d+1}=0}^{t} \left|\left\{(m_1, m_2, \ldots, m_d) \in \mathbb{Z}^{d}_{\geq 0} :\ m_1 + m_2 + \ldots + m_d = t - m_{d+1}\right\}\right| \\ = & \sum_{m_{d+1}=0}^{t} \binom{d-1+t-m_{d+1}}{d-1} \\ = & \sum_{m_{d+1}=0}^{t} \binom{d-1+m_{d+1}}{m_{d+1}} = \binom{d-1+t+1}{t} = \binom{d+t}{d}. \end{aligned}$$

Because in the following chapter we will be interested in some anisotropic version of Smolyak's construction, let us focus on a more general index set. We consider for a parameter $\kappa \in \mathbb{R}^d_{\geq 0}$ the index set

$$\left\{\mathbf{l} \in \mathbb{N}_0^d :\ |\mathbf{l}|_\kappa = \sum_{i=1}^{d} l_i \kappa_i \leq L\right\}.$$

The version of Smolyak's construction we will focus in the last chapter is defined by

$$A(L,d) \ := \sum_{\mathbf{l} \in \mathbb{N}_0^d,\ |\mathbf{l}|_\kappa \leq L} (\Delta^{l_1} \otimes \Delta^{l_2} \otimes \cdots \otimes \Delta^{l_d}).$$

Thus, the polytope of our interest is a simplex given by

$$\Delta_\kappa = \operatorname{conv}\{0, e_1/\kappa_1, e_2/\kappa_2, \dots, e_d/\kappa_d\},$$

respectively

$$L\Delta_\kappa = \operatorname{conv}\{0, Le_1/\kappa_1, Le_2/\kappa_2, \dots, Le_d/\kappa_d\}.$$

As mentioned before, the study of Ehrhart polynomials is of interest by itself, but in our case a more elementary method yields to a accurate solution of our problem. Again, we pick up the idea of a knapsack problem or more detailed the question of the number of feasible solutions of a knapsack problem. The most familiar knapsack problem (also used in this chapter before) has posed the question of how to fill a knapsack of limited weight capacity with different items. The objective is to maximize the total utility. But a different and for us a more interesting point of view is to consider a knapsack problem in the context of so-called cutting pattern. Let $L \in \mathbb{N}$ and $\kappa_1, \kappa_2, \dots, \kappa_d > 0$ such that $\kappa_r = \min\{\kappa_1, \kappa_2, \dots, \kappa_d\} \leq L$ we are interested in the number N of feasible solutions of

$$l_1\kappa_1 + l_2\kappa_2 + \ldots + l_d\kappa_d \leq L \quad \text{where } \mathbf{l} = (l_1, \dots, l_d) \in \mathbb{N}_0^d. \tag{4.3.1}$$

It is known that we have the following upper and lower bounds on the number of integral solutions of (4.3.1) see, e.g. [7].

Lemma 4.3.1. *Let $L \in \mathbb{N}$ and $\kappa_1, \kappa_2, \dots, \kappa_d > 0$ such that $\kappa_r = \min\{\kappa_1, \kappa_2, \dots, \kappa_d\} \leq L$. The number*

$$N := \left|\left\{\mathbf{l} \in \mathbb{N}_0^d : \ l_1\kappa_1 + l_2\kappa_2 + \ldots l_d + \kappa_d \leq L\right\}\right|$$

can be bounded by

$$\frac{L^d}{d! \prod_{i=1}^d \kappa_i} \leq N \leq \frac{(L + \sum_{i=1}^d \kappa_i)^d}{d! \prod_{i=1}^d \kappa_i}.$$

Thus, in the later sections we are also able to consider more general versions of Smolyak's construction and give upper bounds for the cost of such algorithms.

Chapter 5

Cubature Formulas Based on Discontinuous Multiwavelet Bases

This chapter contains one of our main results. We construct simple algorithms for high-dimensional numerical integration of function classes with moderate smoothness, see [36]. 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, as mentioned in the previous chapters for the right choice of parameters, the d-fold tensor product of a Sobolev space $H^s([0,1])$ on the other. 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. These bounds show 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.

Let us recall that our aim is to provide a cubature method that guarantees a (nearly) optimal worst case error and which can be implemented easily. More precise we use composite quadrature rules of a fixed order n, these rules are exact for piecewise polynomials of order n. The present Smolyak construction is related to tensor product multiwavelet expansion in the way that the cubature is exact on finite multiwavelet series up to a critical level. A comparable approach can be found in [37]. Here the considered function classes depends on Haar wavelet series and the cubature is given by a so called scrambled net, see, e.g. [51].

This chapter is organized as follows: In Section 5.1 we define multiwavelets and introduce the spaces on which our cubatures of prescribed level should be exact. In Section 5.2 we present one-dimensional quadratures suited to evaluate the integrals of the univariate wavelets introduced in Section 5.1. 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 5.3 we use Smolyak's construction to obtain cubature rules for multivariate integrands from our one-dimensional quadratures. After giving a precise definition of the class of Hilbert spaces H of multivariate functions we want to consider, error bounds for our cubatures are given; 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 5.4 we report on several numerical tests which allow us to compare our method to established methods.

5.1 Discontinuous multiwavelet bases

This section can be understood as an application of the MRA we have described in Chapter 3. The generalization of the one-dimensional basis to the multivariate case is directly associated to the situation we considered in Chapter 4. Please do not confuse by the notation we use in the multiwavelet case. But since we consider a family of scaling functions we use a slightly modified notation for the index sets.

5.1.1 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, 61, 66].

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 scalingfunctions $\varphi_i(x) := h_i(x) / \|h_i\|_2$, where $\|\cdot\|_2$ is the usual norm on $\mathrm{L}^2([0,1])$. 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

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

of the scalingfunctions φ_i. Observe 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 [61, Subsec. 5.4.1].) Then it is straightforward to verify that the $2^j n$ functions

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

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

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

Next, we define the space

$$V := \bigcup_{j=0}^{\infty} V_n^j = V_n^0 \oplus \bigoplus_{j=0}^{\infty} W_n^j. \tag{5.1.1}$$

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 [61] 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

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

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

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

5.1.2 The multivariate case

In this subsection we extend the concept of multiwavelet bases to higher dimensions. Here we follow an approach we have discussed in the sparse grid chapter and which is suitable for our later analysis. Recall that 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 $|\mathbf{i}|_\infty := \max\{i_1, \ldots, i_d\}$. 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 (sparse) approximation space on level L by

$$V^{d,L} := \sum_{|\mathbf{j}|=L} \bigotimes_{i=1}^{d} V_n^{j_i}. \tag{5.1.2}$$

Similar to the one-dimensional case we put

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

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

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

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 use the shorthands $\nabla_{\mathbf{j}} = \nabla_{j_1} \times \ldots \times \nabla_{j_d}$ and

$$\Psi^{\mathbf{j}}_{\mathbf{i},\mathbf{k}} = \bigotimes_{u=1}^{d} \psi^{j_u}_{i_u,k_u}.$$

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

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

then $\mathrm{supp}\Psi^{\mathbf{j}}_{\mathbf{i},\mathbf{k}} = I^{\mathbf{j}}_{\mathbf{k}}$ holds.

5.2 One-dimensional quadrature

First of all let us recall that a general one-dimensional quadrature is given by

$$\mathrm{Q}_m(f) = \sum_{\nu=1}^{m} \lambda_\nu f(x_\nu), \tag{5.2.1}$$

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 interested in quadrature formulas with high polynomial exactness here –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. [22].

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

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

We get the requested quadrature by scaling and translating a simpler one-dimensional quadrature formula Q_m. If Q_m has the explicit form (5.2.1), then our resulting quadrature uses $2^l m$ sample points and is given by

$$A_m(l,1)(f) = \sum_{k \in \nabla_l} \sum_{\nu=1}^{m} 2^{-l} \lambda_\nu f(2^{-l} x_\nu + 2^{-l} k). \tag{5.2.2}$$

$A_m(l,1)$ 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.

5.2.1 Error analysis

The error analysis is based on the observations we pointed out in the previous chapters. However, let us start with a detailed discussion of our one-dimensional quadrature method. Thus, 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 5.1.1. For $s > 0$ we define a discrete norm

$$|f|^2 := \sum_{j\ge -1} \sum_{k\in\nabla_j} \sum_{i=0}^{n-1} 2^{j2s} \left|\langle f, \psi_{i,k}^j\rangle\right|^2 \tag{5.2.3}$$

on the space

$$H_{s,n} := \left\{f \in \mathrm{L}^2([0,1]) \mid |f| < \infty\right\}, \tag{5.2.4}$$

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 $H_{s,n}$ 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 (5.2.1) such that Q_m is exact on Π_n. Let $A_m(l,1)$ be as in (5.2.2). Then, the wavelet expansion of a function $f \in H_{s,n}$ and the Cauchy-Schwarz inequality yield the following error bound for our algorithm $A_m(l,1)$:

Theorem 5.2.1. *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*

$$\mathrm{err}(H_{s,n}, A_m(l,1)) \quad \le \quad C\, 2^{-ls}. \tag{5.2.5}$$

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

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

The Cauchy-Schwarz inequality yields

$$\mathrm{err}(f, A_m(l,1)) \quad \le \quad |f| \left(\sum_{j\ge -1} \sum_{k\in\nabla_j} \sum_{i=0}^{n-1} 2^{-j2s} \left\{I(\psi_{i,k}^j) - A_m(l,1)\psi_{i,k}^j\right\}^2\right)^{1/2}.$$

Because of the polynomial exactness and vanishing moments we get

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

By some easy calculations and by the fact that supp $\psi_{i,k}^j = I_k^j$ we get

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

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

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

Then,

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

Hence, the integration error can be upper bounded by

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

Thus we proved that (5.2.5) 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 5.2.2. The error estimate in Theorem 5.2.1 is asymptotically optimal as Theorem 5.3.9 will reveal.

5.3 Integration via Smolyak's construction

5.3.1 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: Recall that 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

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

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 difficult to show the exactness of $A_m(L,d)$ on $V_n^{d,L}$.

Theorem 5.3.1. *The cubature $A_m(L,d)$ is exact on the approximation space $V_n^{d,L}$.*

Proof. Here, we follow the lines of proof of [48, Theorem 2] and proceed via induction over the dimension. For $d = 1$ the assertion follows by the exactness of the one-dimensional quadrature formula $A_m(L,1)$. Let $\mathbf{l} \in \mathbb{N}_0^{d+1}$ satisfy $|\mathbf{l}| = L$ and $l_{d+1} = L - \tau$. Let $f_i \in V_n^{l_i}$ for $i = 1, \ldots d+1$ and

$$f = f_1 \otimes \cdots \otimes f_d \otimes f_{d+1}.$$

Then we get

$$A_m(L,d+1)f = \left(\sum_{\nu=0}^{L} A_m(\nu,d) \otimes \Delta^{L-\nu} \right)(f)$$

$$= \sum_{\nu=0}^{L} A_m(\nu,d)(f_1 \otimes \cdots \otimes f_d) \cdot (A_m(L-\nu,1) - A_m(L-\nu-1,1)) f_{d+1}.$$

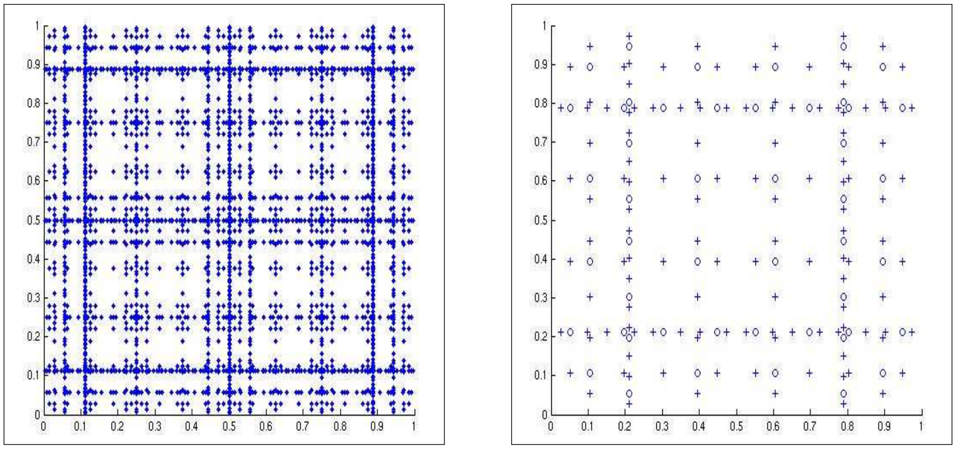

Figure 5.1: $A_3(5, 2)$ and $A_2(3, 2)$ with underlying Gauss quadrature. In the right diagram "+" denotes sample points with positive, "o" sample points with negative weights.

Since for all $\nu \leq \tau - 1$

$$A_m(L - \nu, 1)f_{d+1} = \int_{[0,1)} f_{d+1} dx = A_m(L - \nu - 1, 1)f_{d+1}$$

and, according to our assumption,

$$A_m(\nu, d)(f_1 \otimes \cdots \otimes f_d) = \int_{[0,1)^d} f_1 \otimes \cdots \otimes f_d \, dx \quad \text{for all} \ \ \tau \leq \nu \leq L,$$

we get

$$\begin{aligned} A_m(L, d+1)f &= \sum_{\nu=\tau}^{L} A_m(\nu, d)(f_1 \otimes \cdots \otimes f_d) \\ &\quad \cdot (A_m(L - \nu, 1) - A_m(L - \nu - 1, 1))f_{d+1} \\ &= \int_{[0,1)^d} f_1 \otimes \cdots \otimes f_d \, dx \cdot A_m(L - \tau, 1)f_{d+1} \\ &= \int_{[0,1)^d} f_1 \otimes \cdots \otimes f_d \, dx \cdot \int_{[0,1)} f_{d+1} \, dx = \int_{[0,1)^{d+1}} f \, dx. \end{aligned}$$

$\square$

5.3.2 Upper bounds for the cubature error

For the error analysis we consider product spaces which are based on the spaces $H_{s,n}$ used for our one-dimensional quadrature error bounds. These seem to be

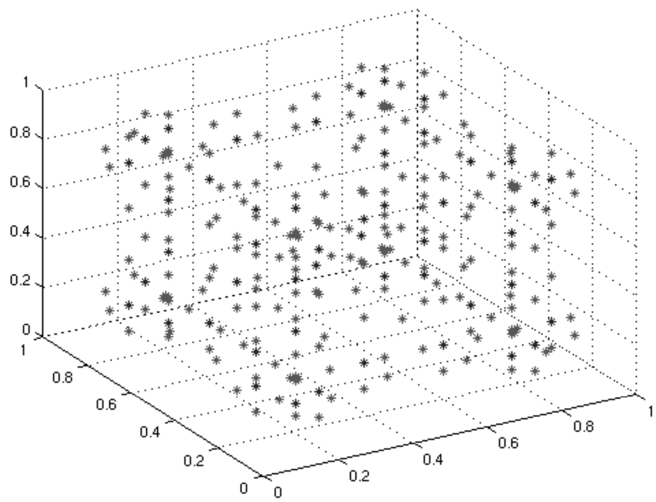

Figure 5.2: Smolyak's construction based on Gauss quadrature in dimension 3.

the natural spaces for our variation of Smolyak's construction. For a function f we define a norm

$$|f|_{d,s,n}{}^2 \quad = \quad \sum_{\mathbf{j}\geq -1} \sum_{\mathbf{k}\in\nabla_{\mathbf{j}}} \sum_{|\mathbf{i}|_\infty=0}^{n-1} 2^{|\mathbf{j}|2s} \left|\langle f, \Psi_{\mathbf{i},\mathbf{k}}^{\mathbf{j}}\rangle\right|^2 \tag{5.3.1}$$

and the space

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

Before we verify the main result of this chapter, we want to calculate the induced operator norm of the functional $I : H_{s,n} \to \mathbb{R}$: For arbitrary $f \in H_{s,n}$ we get

$$\begin{aligned} I(f) &= \sum_{j\geq -1} \sum_{k\in\nabla_j} \sum_{i=0}^{n-1} \langle f, \psi_{i,k}^j\rangle I(\psi_{i,k}^j) \leq \|f\|_{s,n} \left(\sum_{j\geq -1} \sum_{k\in\nabla_j} \sum_{i=0}^{n-1} 2^{-j2s} I(\psi_{i,k}^j)^2\right)^{1/2} \\ &= \|f\|_{s,n} 2^s. \end{aligned}$$

Consequently, we obtain for the induced operator norm

$$\|I\|_{\mathrm{op}} = \sup_{f\in H_{s,n},\ \|f\|_{s,n}=1} |I(f)| \leq 2^s.$$

On the other hand, we get for $f^* = 2^s \cdot \mathbf{1}_{[0,1)}$

$$\|f^*\|_{s,n} = |2^s| \left(\sum_{j\geq -1} \sum_{k\in\nabla_j} \sum_{i=0}^{n-1} 2^{j2s} \left|\langle \mathbf{1}_{[0,1)}, \psi_{i,k}^j\rangle\right|^2\right)^{1/2} = 2^s (2^{-2s})^{1/2} = 1.$$

Thus, the induced operator norm satisfies

$$\|I\|_{\mathrm{op}} \geq I(f^*) = I(2^s\,\mathbf{1}_{[0,1)}) = 2^s,$$

which implies $\|I\|_{\mathrm{op}} = 2^s$. But this small obliquity is due to the convention that $\psi_i^{-1} = \varphi_i$.

Theorem 5.3.2. *Let $n \in \mathbb{N}$ and let the one-dimensional quadrature Q_m be exact on Π_n. For $s > 1/2$ let C be the constant from (5.2.5). The worst case error of $A_m(L,d)$ satisfies*

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

To prove our main result we adapt the proof of [67, Lemma 2]. There, Wasilkowski and Woźniakowski verified error bounds not only for d-dimensional cubatures, but also for more general d-dimensional approximation algorithms based on Smolyak's construction.

Proof. Let n be fixed. The proof is via induction and based on the observation that

$$\begin{aligned}
A_m(L,d) &= \sum_{\tilde{\mathbf{l}}\in\mathbb{N}_0^{d-1},\,\left|\tilde{\mathbf{l}}\right|\leq L}\ \sum_{k=0}^{L-\left|\tilde{\mathbf{l}}\right|} \Delta^{\tilde{l}_1}\otimes\Delta^{\tilde{l}_2}\otimes\ldots\otimes\Delta^{\tilde{l}_{d-1}}\otimes\Delta^k \\
&= \sum_{\tilde{\mathbf{l}}\in\mathbb{N}_0^{d-1},\,\left|\tilde{\mathbf{l}}\right|\leq L} \Delta^{\tilde{l}_1}\otimes\Delta^{\tilde{l}_2}\otimes\ldots\otimes\Delta^{\tilde{l}_{d-1}}\otimes\sum_{k=0}^{L-\left|\tilde{\mathbf{l}}\right|}\Delta^k \\
&= \sum_{\tilde{\mathbf{l}}\in\mathbb{N}_0^{d-1},\,\left|\tilde{\mathbf{l}}\right|\leq L} \Delta^{\tilde{l}_1}\otimes\Delta^{\tilde{l}_2}\otimes\ldots\otimes\Delta^{\tilde{l}_{d-1}}\otimes A_m(L-\left|\tilde{\mathbf{l}}\right|,1).
\end{aligned}$$

Thus, we get

$$\begin{aligned}
I_d - A_m(L,d) &= \sum_{\tilde{\mathbf{l}}\in\mathbb{N}_0^{d-1},\,\left|\tilde{\mathbf{l}}\right|\leq L}\left(\bigotimes_{u=1}^{d-1}\Delta^{\tilde{l}_u}\right)\otimes\left(I_1 - A_m(L-\left|\tilde{\mathbf{l}}\right|,1)\right) \\
&\quad + (I_{d-1} - A_m(L,d-1))\otimes I_1.
\end{aligned}$$

Consequently,

$$\begin{aligned}
\|I_d - A_m(L,d)\|_{\mathrm{op}} &\leq \sum_{\tilde{\mathbf{l}}\in\mathbb{N}_0^{d-1},\,\left|\tilde{\mathbf{l}}\right|\leq L}\ \prod_{u=1}^{d-1}\left\|\Delta^{\tilde{l}_u}\right\|_{\mathrm{op}}\left\|I_1 - A_m(L-\left|\tilde{\mathbf{l}}\right|,1)\right\|_{\mathrm{op}} \\
&\quad + \|I_{d-1} - A_m(L,d-1)\|_{\mathrm{op}}\,\|I_1\|_{\mathrm{op}}.
\end{aligned}$$

The next step is to consider the operator norm of the difference quadratures. According to Theorem 5.2.1 we have

$$\left\|\Delta^{\tilde{l}_u}\right\|_{\mathrm{op}} \leq C\left(2^{-\tilde{l}_u s} + 2^{-(\tilde{l}_u-1)s}\right) = C\left(1+2^s\right)2^{-\tilde{l}_u s},$$

which leads to

$$\begin{aligned}
\sum_{\tilde{\mathbf{l}}\in\mathbb{N}_0^{d-1},\,\left|\tilde{\mathbf{l}}\right|\leq L}\prod_{u=1}^{d-1}\left\|\Delta^{\tilde{l}_u}\right\|_{\mathrm{op}}\left\|I_1 - A_m(L-\left|\tilde{\mathbf{l}}\right|,1)\right\|_{\mathrm{op}} \\
&\leq C^d\left(1+2^s\right)^{d-1}2^{-Ls}\sum_{\tilde{\mathbf{l}}\in\mathbb{N}_0^{d-1},\,\left|\tilde{\mathbf{l}}\right|\leq L} 1 \\
&= C^d\left(1+2^s\right)^{d-1}2^{-Ls}\binom{L+d-1}{d-1}.
\end{aligned}$$

And therefore we get for the integration error

$$\begin{aligned}
\left\|I_d - A_m(L,d)\right\|_{\mathrm{op}} \leq & \quad C\left(C\left(1+2^s\right)\right)^{d-1}\,2^{-Ls}\binom{L+d-1}{d-1} \\
& + \quad \left\|I_{d-1} - A_m(L,d-1)\right\|_{\mathrm{op}}\,2^s.
\end{aligned}$$

Inductively we get

$$\begin{aligned}
\left\|I_d - A_m(L,d)\right\|_{\mathrm{op}} &\leq C\,2^{-Ls}(2^s)^{d-1}\sum_{\nu=0}^{d-1}\left(\frac{C\left(1+2^s\right)}{2^s}\right)^{\nu}\binom{L+\nu}{\nu} \\
&\leq C\left(\max\left\{2^s, C\left(1+2^s\right)\right\}\right)^{d-1}2^{-Ls}\sum_{\nu=0}^{d-1}\binom{L+\nu}{\nu}.
\end{aligned}$$

Consequently, our proposition follows by the fact that $\sum_{\nu=0}^{d-1}\binom{L+\nu}{\nu} = \binom{L+d}{d-1}$. $\square$

From the abstract definition of our function space $H_{s,n}$ 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 $H_{s,n}$. In Chapter 3 Section 3.1.1 we have seen that the continuous embedding is established by some Jackson type inequality.

Theorem 5.3.3. *Let $(\psi_i)_{i=0}^{n-1}$ be multiwavelets of order n. For all $s < n$ the inclusion $H^s([0,1]) \subset H_{s,n}$ holds. More precisely, there exists a constant $K > 0$ such that for every $f \in H^s([0,1])$ we have*

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

For a proof of the theorem see, e.g. [12, 17, 61, 66]. Notice in general we cannot hope to prove equivalence of the norms on $H_{s,n}$ and $H^s([0,1])$. This is obvious in the case where $s > 1/2$: $H_{s,n}$ contains discontinuous functions, while $H^s([0,1])$ does not. The so-called 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 5.3.2 reads as follows:

Corollary 5.3.4. *Let $s > 1/2$ and $n > s$. Let the one-dimensional quadrature Q_m be exact on Π_n. Then for every $L \geq 0$*

$$\text{err}(H^s_{\text{mix}}, A_{m(s)}(L,d)) \leq CK \left(\max\{2^s, C(1+2^s)\}\right)^{d-1} 2^{-Ls} \binom{L+d}{d-1},$$

where the constant C is as in Theorem 5.2.1 and K as in Theorem 5.3.3.

Now we analyze the cost of the cubature algorithm $A_m(L,d)$. By mimicking the proof of [67, Lemma 1], we get

$$A_m(L,d) = \sum_{L-d+1 \leq |\mathbf{l}| \leq L} (-1)^{L-|\mathbf{l}|} \binom{d-1}{L-|\mathbf{l}|} \bigotimes_{u=1}^{d} A_m(l_u, 1), \tag{5.3.2}$$

where $A_m(l_u,1)$ is as in (5.2.2). This clearly shows 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 generally 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, we have

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

This estimate, Theorem 5.3.2, and some elementary calculations lead to the following corollary.

Corollary 5.3.5. *Let $n \in \mathbb{N}$ and let Q_m be exact on Π_n. For $s > 1/2$ the worst case error of $A_m(L,d)$ satisfies*

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

Remark 5.3.6. Recall that H^s_{mix} is continuously embedded in $H^d_{s,n}$ if $s < n$. In this situation Corollary 5.3.5 holds in particular for H^s_{mix} in place of $H^d_{s,n}$.

5.3.3 Alternative upper bound proof

Alternatively to the proof of Theorem 5.3.2 where we adapt the proof of [67, Lemma 2] it is also convenient to make use of the description in (5.3.2) and the fact that these construction is exact on the approximation spaces $V_n^{d,L}$. The main idea is to give a direct estimation of the error similar to the Haar wavelet case we considered in Chapter 3 Section 3.2. Before we are able to give the proof of our main result which has not the same logarithmic order as in the first proof we have to show the following fact.

Lemma 5.3.7. *For all $s > 0$, $s \neq 1$ and $L \geq 0$ we get*

$$\sum_{j=0}^{\infty} 2^{-js} \binom{j+L+d-1}{d-1} \lesssim \frac{(L+d-1)^{d-1}}{(d-1)!} \left(\frac{1-s^{-(d-1)}}{s-1} \right)$$

and for $s = 1$

$$\sum_{j=0}^{\infty} 2^{-js} \binom{j+L+d-1}{d-1} \lesssim \frac{d(L+d-1)^{d-1}}{(d-1)!}.$$

Proof. Obviously for all $s > 0$ and $L \geq 0$

$$\begin{aligned} \sum_{j=0}^{\infty} 2^{-js} \binom{j+L+d-1}{d-1} &\lesssim \frac{1}{(d-1)!} \int_0^{\infty} 2^{-sx} (x+L')^{d-1} dx \\ &= \frac{1}{(d-1)!} \int_0^{\infty} e^{-s'x} (x+L')^{d-1} dx, \end{aligned}$$

where $L' := L + d - 1$ and $s' := s \ln(2)$. Thus, we get

$$\sum_{j=0}^{\infty} 2^{-js} \binom{j+L+d-1}{d-1} \lesssim \frac{e^{-s(L')}}{(d-1)!} \underbrace{\int_{L'}^{\infty} e^{-s'x} x^{d-1} dx}_{=:I}.$$

The integral I can be bounded in the following way

$$\begin{aligned} I &= \frac{e^{-sx}}{s^d} [(-sx)^{d-1} - (d-1)(-sx)^{d-2} + (d-1)(d-2)(-sx)^{d-3} \\ &\quad - \cdots + (-1)^{d-1}(d-1)!]_{x=L'}^{\infty} \\ &= \frac{e^{-s(L')}}{s} (L')^{d-1} [1 + \frac{(d-1)}{s(L')} + \frac{(d-1)(d-2)}{s^2(L')^2} + \cdots + \frac{(d-1)!}{s^{d-1}(L')^{d-1}}] \\ &\leq \frac{e^{-s(L')}}{s} (L')^{d-1} [1 + \frac{1}{s} + \frac{1}{s^2} + \cdots + \frac{1}{s^{d-1}}]. \end{aligned}$$

And consequently for the case where $s \neq 1$ we get

$$\sum_{j=0}^{\infty} 2^{-js} \binom{j+L+d-1}{d-1} \lesssim 2\frac{(L+d-1)^{d-1}}{(d-1)!} \left(\frac{1-s^{-(d-1)}}{s-1} \right).$$

The case for $s = 1$ is clear. $\square$

With this elementary lemma in mind and the idea of the error analysis we used for the Haar wavelet case we state one of our main results.

Theorem 5.3.8. *Let $n \in \mathbb{N}$ and let the one-dimensional quadrature Q_m be exact on Π_n. For $s > 1/2$, $L > 0$ the worst case error of $A_m(L,d)$ satisfies*

$$\operatorname{err}(H^d_{s,n}, A_m(L,d)) \leq C_{d,n,s} \cdot \frac{(L+d-1)^{3/2(d-1)}}{2^{Ls}},$$

where

$$C_{d,n,s} = C \cdot g(s,d) \cdot \frac{d^{3/2}}{(d-1)!} \left(\sum_{i=0}^{n-1} \|\Psi_i\|_\infty^2 \right)^{1/2}$$

and

$$g(s,d)^2 = \begin{cases} d & : \quad s = 1 \\ 1/2 \cdot \frac{(2s-1)^{d-1}-1}{(2s-1)^{d-1}} \frac{1}{s-1} & : \quad \textit{else.} \end{cases}$$

Proof. Similar to the one-dimensional case the cubature error can be bounded by

$$\operatorname{err}(f, A_m(L,d)) \leq |f|_{d,s,n} \left(\sum_{|\mathbf{j}|\geq 0} \sum_{\mathbf{k}\in\nabla_{\mathbf{j}}} \sum_{i=0}^{n-1} 2^{-|\mathbf{j}|2s} \left\{ A_m(L,d)\Psi^{\mathbf{j}}_{i,\mathbf{k}} \right\}^2 \right)^{1/2} .$$

Form Theorem 5.3.1 we know that our variation of Smolyak's construction $A_m(L,d)$ is exact on the piecewise polynomial spaces $V_n^{d,L}$, and consequently exact for the first multiwavelets up to the critical level L,

$$A_m(L,d)\Psi^{\mathbf{j}}_{i,\mathbf{k}} = \int_{[0,1]^d} \Psi^{\mathbf{j}}_{i,\mathbf{k}} dx = 0 \quad \text{for all} \quad |\mathbf{j}| < L.$$

Hence we get an error bound

$$\operatorname*{err}_{|f|_{d,s,n}=1}(f, A_m(L,d))^2 \quad \leq \quad \sum_{|\mathbf{j}|\geq L} \sum_{\mathbf{k}\in\nabla_{\mathbf{j}}} \sum_{i=0}^{n-1} 2^{-|\mathbf{j}|2s} \left\{ A_m(L,d)\Psi^{\mathbf{j}}_{i,\mathbf{k}} \right\}^2 .$$

If we use the notation in (5.3.2) and the shorthand A^{l_u} for $A_m(l_u,1)$ it is easy to verify that the error is bounded by

$$\operatorname*{err}_{|f|_{d,s,n}=1}(f, A_m(L,d))^2 \leq \sum_{|\mathbf{j}|\geq L} \sum_{i=0}^{n-1} \sum_{\mathbf{k}\in\nabla_{\mathbf{j}}} 2^{-|\mathbf{j}|2s} \left\{ \sum_{L-d+1\leq|\mathbf{l}|\leq L} (-1)^{L-|\mathbf{l}|} \binom{d-1}{L-|\mathbf{l}|} \bigotimes_{u=1}^{d} A^{l_u}\Psi^{\mathbf{j}}_{i,\mathbf{k}} \right\}^2$$

Observe that in this equation only the case where $l_u \le j_u$ is of interest because in the other case $\bigotimes_{u=1}^{d} A^{l_u} \Psi^{\mathbf{j}}_{i,\mathbf{k}} = 0$. Consequently we get

$$
\begin{aligned}
& \sum_{|\mathbf{j}|\ge L} \sum_{i=0}^{n-1} \sum_{\mathbf{k}\in\nabla_{\mathbf{j}}} 2^{-|\mathbf{j}|2s} \left\{ \sum_{\substack{L-d+1\le|\mathbf{l}|\le L \\ l_u\le j_u \;u=1,\ldots,d}} (-1)^{L-|\mathbf{l}|} \binom{d-1}{L-|\mathbf{l}|} \sum_{\mathbf{k}'\in\nabla_{\mathbf{l}}} \bigotimes_{u=1}^{d} A^{l_u}_{k'_u} \Psi^{\mathbf{j}}_{i,\mathbf{k}} \right\}^2 \\
\le \; & \sum_{|\mathbf{j}|\ge L} \sum_{i=0}^{n-1} \sum_{\mathbf{k}\in\nabla_{\mathbf{j}}} 2^{-|\mathbf{j}|2s} \left\{ \sum_{\substack{L-d+1\le|\mathbf{l}|\le L \\ l_u\le j_u \;u=1,\ldots,d}} \binom{d-1}{L-|\mathbf{l}|} \sum_{\mathbf{k}'\in\nabla_{\mathbf{l}}} \bigotimes_{u=1}^{d} |A|^{l_u}_{k'_u} \left\|\Psi^{\mathbf{j}}_{i,\mathbf{k}}\right\|_\infty \mathbf{1}_{\operatorname{supp}\Psi^{\mathbf{j}}_{i,\mathbf{k}}} \right\}^2 \\
= \; & \sum_{i=0}^{n-1} \|\Psi_i\|_\infty^2 \sum_{|\mathbf{j}|\ge L} \sum_{\mathbf{k}\in\nabla_{\mathbf{j}}} 2^{|\mathbf{j}|(1-2s)} \left\{ \sum_{\substack{L-d+1\le|\mathbf{l}|\le L \\ l_u\le j_u \;u=1,\ldots,d}} \binom{d-1}{L-|\mathbf{l}|} \sum_{\mathbf{k}'\in\nabla_{\mathbf{l}}} \bigotimes_{u=1}^{d} |A|^{l_u}_{k'_u} \mathbf{1}_{\operatorname{supp}\Psi^{\mathbf{j}}_{i,\mathbf{k}}} \right\}^2 ,
\end{aligned}
$$

where $|A|^{l_u}_{k'_u}$ denote the quadrature $A^{l_u}_{k'_u}$ with the absolute value of the given weights. Let us take a closer look to square term. We define a function

$$
\begin{aligned}
\phi_j(\mathbf{l},\mathbf{k}) &:= \sum_{\mathbf{k}'\in\nabla_{\mathbf{l}}} \bigotimes_{u=1}^{d} |A|^{l_u}_{k'_u} \mathbf{1}_{\operatorname{supp}\Psi^{\mathbf{j}}_{i,\mathbf{k}}} \\
&= \sum_{\mathbf{k}'\in\nabla_{\mathbf{l}}} \prod_{u=1}^{d} |A|^{l_u}_{k'_u} \mathbf{1}_{\operatorname{supp}\Psi^{j_u}_{i,k_u}} \\
&= \sum_{\mathbf{k}'\in\nabla_{\mathbf{l}}} \prod_{u=1}^{d} \sum_{m=1}^{n} 2^{-l_u} |\omega_m| \, \mathbf{1}_{I^{j_u}_{k_u}} (2^{-l_u} x_m + 2^{-l_u} k'_u) \\
&= 2^{-|\mathbf{l}|} \sum_{\mathbf{k}'\in\nabla_{\mathbf{l}}} \prod_{u=1}^{d} \sum_{m=1}^{n} |\omega_m| \, \mathbf{1}_{I^{j_u}_{k_u}} (2^{-l_u} x_m + 2^{-l_u} k'_u) = 2^{-|\mathbf{l}|} \tilde{\phi}_j(\mathbf{l},\mathbf{k}).
\end{aligned}
$$

If the underlying one-dimensional quadrature have positive weights it is not hard to see that for the choices of $\mathbf{l},\mathbf{k}$ we are interested in the function $\tilde{\phi}_j(\mathbf{l},\mathbf{k}) \le 1$ and consequently $\tilde{\phi}_j(\mathbf{l},\mathbf{k})^2 \le \tilde{\phi}_j(\mathbf{l},\mathbf{k})$ otherwise we get

$$
\tilde{\phi}_j(\mathbf{l},\mathbf{k})^2 \le \left(\sum_{m=1}^{n} |\omega_m| \right)^d \tilde{\phi}_j(\mathbf{l},\mathbf{k}).
$$

However, let us assume we have positive weights only. The Cauchy-Schwarz inequality yields

$$
\sum_{\substack{L-d+1\le|\mathbf{l}|\le L \\ l_u\le j_u \;u=1,\ldots,d}} \binom{d-1}{L-|\mathbf{l}|} \sum_{\mathbf{k}'\in\nabla_{\mathbf{l}}} \bigotimes_{u=1}^{d} |A|^{l_u}_{k'_u} \mathbf{1}_{\operatorname{supp}\Psi^{\mathbf{j}}_{i,\mathbf{k}}} = \sum_{\substack{L-d+1\le|\mathbf{l}|\le L \\ l_u\le j_u \;u=1,\ldots,d}} \binom{d-1}{L-|\mathbf{l}|} \phi(\mathbf{l},\mathbf{k})
$$

$$
\begin{aligned}
&\leq \left(\sum_{L-d+1\leq|\mathbf{l}|\leq L} 1\right)^{1/2} \left(\sum_{\substack{L-d+1\leq|\mathbf{l}|\leq L\\ l_u\leq j_u \ u=1,\ldots,d}} \binom{d-1}{L-|\mathbf{l}|}^2 \phi(\mathbf{l},\mathbf{k})^2\right)^{1/2} \\
&\leq \left(\frac{d}{(d-1)!}\cdot(L+d-1)^{d-1}\right)^{1/2} \left(\sum_{\substack{L-d+1\leq|\mathbf{l}|\leq L\\ l_u\leq j_u \ u=1,\ldots,d}} \binom{d-1}{L-|\mathbf{l}|}^2 \left(2^{-|\mathbf{l}|}\tilde{\phi}(\mathbf{l},\mathbf{k})\right)^2\right)^{1/2}.
\end{aligned}
$$

By our assumption that $\tilde{\phi}(\mathbf{l},\mathbf{k}) \leq 1$ we get

$$
\begin{aligned}
\sum_{\substack{L-d+1\leq|\mathbf{l}|\leq L\\ l_u\leq j_u \ u=1,\ldots,d}} \binom{d-1}{L-|\mathbf{l}|}^2 \left(2^{-|\mathbf{l}|}\tilde{\phi}(\mathbf{l},\mathbf{k})\right)^2 &= \sum_{\substack{L-d+1\leq|\mathbf{l}|\leq L\\ l_u\leq j_u \ u=1,\ldots,d}} 2^{-2|\mathbf{l}|}\binom{d-1}{L-|\mathbf{l}|}^2 \tilde{\phi}(\mathbf{l},\mathbf{k})^2 \\
&\leq \sum_{\substack{L-d+1\leq|\mathbf{l}|\leq L\\ l_u\leq j_u \ u=1,\ldots,d}} 2^{-2|\mathbf{l}|}\binom{d-1}{L-|\mathbf{l}|}^2 \tilde{\phi}(\mathbf{l},\mathbf{k}).
\end{aligned}
$$

Hence, we get

$$
\begin{aligned}
&\sum_{\mathbf{k}\in\nabla_{\mathbf{j}}} \left\{ \sum_{\substack{L-d+1\leq|\mathbf{l}|\leq L\\ l_u\leq j_u \ u=1,\ldots,d}} \binom{d-1}{L-|\mathbf{l}|} \sum_{\mathbf{k}'\in\nabla_{\mathbf{l}}} \bigotimes_{u=1}^{d} |A|_{k'_u}^{l_u}\, \mathbf{1}_{\operatorname{supp}\Psi_{i,\mathbf{k}}^{\mathbf{j}}} \right\}^2 \qquad (5.3.3)\\
&\leq \frac{d}{(d-1)!}\cdot(L+d-1)^{d-1} \sum_{\mathbf{k}\in\nabla_{\mathbf{j}}} \sum_{\substack{L-d+1\leq|\mathbf{l}|\leq L\\ l_u\leq j_u \ u=1,\ldots,d}} 2^{-2|\mathbf{l}|}\binom{d-1}{L-|\mathbf{l}|}^2 \tilde{\phi}(\mathbf{l},\mathbf{k}) \\
&= \frac{d}{(d-1)!}\cdot(L+d-1)^{d-1} \sum_{\substack{L-d+1\leq|\mathbf{l}|\leq L\\ l_u\leq j_u \ u=1,\ldots,d}} 2^{-2|\mathbf{l}|}\binom{d-1}{L-|\mathbf{l}|}^2 \sum_{\mathbf{k}\in\nabla_{\mathbf{j}}} \tilde{\phi}(\mathbf{l},\mathbf{k}).
\end{aligned}
$$

And since

$$
\begin{aligned}
\sum_{\mathbf{k}\in\nabla_{\mathbf{j}}} \tilde{\phi}(\mathbf{l},\mathbf{k}) &= \sum_{\mathbf{k}\in\nabla_{\mathbf{j}}} \sum_{\mathbf{k}'\in\nabla_{\mathbf{l}}} \prod_{u=1}^{d} \sum_{m=1}^{n} |\omega_m|\, \mathbf{1}_{\operatorname{supp}\Psi_{i,k_u}^{j_u}} (p_m + 2^{-l_u}k'_u) \\
&= \sum_{\mathbf{k}'\in\nabla_{\mathbf{l}}} \sum_{\mathbf{k}\in\nabla_{\mathbf{j}}} \prod_{u=1}^{d} \sum_{m=1}^{n} |\omega_m|\, \mathbf{1}_{\operatorname{supp}\Psi_{i,k_u}^{j_u}} (p_m + 2^{-l_u}k'_u) \\
&= \sum_{\mathbf{k}'\in\nabla_{\mathbf{l}}} \left(\sum_{m=1}^{n} |\omega_m|\right)^d \\
&= |\nabla_{\mathbf{l}}|,
\end{aligned}
$$

we get that inequality (5.3.3) can be bounded by

$$\begin{aligned}
&\frac{d}{(d-1)!}\cdot(L+d-1)^{d-1}\sum_{L-d+1\le|\mathbf{l}|\le L}2^{-2|\mathbf{l}|}\binom{d-1}{L-|\mathbf{l}|}^2|\nabla_{\mathbf{l}}|\\
&\quad=\frac{d}{(d-1)!}\cdot(L+d-1)^{d-1}\sum_{L-d+1\le|\mathbf{l}|\le L}2^{-|\mathbf{l}|}\binom{d-1}{L-|\mathbf{l}|}^2\\
&\le\frac{d}{(d-1)!}\cdot(L+d-1)^{d-1}\sum_{l=L-d+1}^{L}2^{-l}\binom{l+d-1}{d-1}\binom{d-1}{L-l}^2.
\end{aligned}$$

If the dimension is not to small an easy version of the well known Stirling formula yields to the following estimation for the last term,

$$\begin{aligned}
\sum_{l=L-d+1}^{L}2^{-l}\binom{l+d-1}{d-1}\binom{d-1}{L-l}^2 &\le \sum_{l=L-d+1}^{L}2^{-l}\frac{(l+d-1)!}{l!}\frac{(d-1)!(2e)^{2(d-1)}}{(d-1)^{2(d-1)}}\\
&\le d\cdot 2^{-L}(L+d-1)^{d-1}.
\end{aligned}$$

Now, we can finish the proof and get an error bound

$$\begin{aligned}
\operatorname*{err}_{|f|_{d,s,n}=1}(f,A_m(L,d))^2 &\le \frac{d^2}{(d-1)!}\cdot\sum_{i=0}^{n-1}\|\Psi_i\|_\infty^2\cdot(L+d-1)^{2(d-1)}\cdot 2^{-L}\sum_{|\mathbf{j}|\ge L}2^{|\mathbf{j}|(1-2s)}\\
&= \frac{d^2}{(d-1)!}\cdot\sum_{i=0}^{n-1}\|\Psi_i\|_\infty^2\cdot(L+d-1)^{2(d-1)}\cdot 2^{-L}\sum_{j=L}^{\infty}2^{j(1-2s)}\binom{j+d-1}{d-1}\\
&= \frac{d^2}{(d-1)!}\cdot\sum_{i=0}^{n-1}\|\Psi_i\|_\infty^2\cdot(L+d-1)^{2(d-1)}\cdot 2^{-L2s}\sum_{j=0}^{\infty}2^{j(1-2s)}\binom{j+L+d-1}{d-1}\\
&\lesssim \frac{d^3}{(d-1)!^2}\sum_{i=0}^{n-1}\|\Psi_i\|_\infty^2\cdot(L+d-1)^{3(d-1)}\cdot 2^{-L2s}\left(\frac{1-(2s-1)^{-(d-1)}}{(2s-1)-1}\right)\\
&= C_d\cdot\sum_{i=0}^{n-1}\|\Psi_i\|_\infty^2\cdot(L+d-1)^{3(d-1)}\cdot 2^{-L2s}\left(\frac{(2s-1)^{d-1}-1}{(2s-1)^{d-1}}\right)1/(2s-2).
\end{aligned}$$

□

5.3.4 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 $H^d_{s,n}$ and H^s_{mix}. For the considered spaces $H^d_{s,n}$ there is a general method to prove lower bounds for the worst case error of any cubature Q_N. In [37] Heinrich, Hickernell and

Yue presented a lower bound for Haar wavelet spaces that works similar for the spaces $H^d_{s,n}$. 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 $\mathcal{L}^2$-discrepancy [56].)

Theorem 5.3.9. *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*

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

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, for the norm of our candidate we get

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

Due to (5.1.1) 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 $M := \max_{i=1}^n \{\|\varphi_{\mathbf{i}}\|_\infty, \|\psi_{\mathbf{i}}\|_\infty\}$, 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

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

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

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

Furthermore, we have

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

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

$$
\mathrm{err}(f^*, \mathrm{Q}_N) = \frac{\left| \int_{[0,1)^d} f dx \right|}{|f|^d_{d,s,n}} \ge C \frac{L^{(d-1)/2}}{2^{Ls}}
$$

with a constant C not depending on L, but depending on d and s. □

5.4 Numerical examples

We have implemented our cubature method and computed the integrals of certain test functions in dimension 5 and 10. The families of test functions we have considered were selected from the testing package of Genz [32, 33], and they are named as follows:

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

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

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

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

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

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

This choice of test functions is obviously unfavorable with regard to our cubature rule and the corresponding function classes, but it enables us to compare our results directly to the results of the algorithms studied in [48] and [63]. The algorithm in [48] is based on Smolyak's construction and the Clenshaw-Curtis rule in dimension $d = 1$. The algorithms in [63, 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 [27], 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 followed the conventions from [48, 63]: 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, $j = 1, \dots, 6$. 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 [63] and in [48], 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 [63] 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 5.3 to 5.10 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 and additionally the two families 2 and 5 in dimension 5. In Figure 5.3 to 5.8 we also plotted the median error of the lattice rule COPY taken from the diagrams in [63] and the median error of the algorithm considered by Novak and Ritter taken from the diagrams in [48].

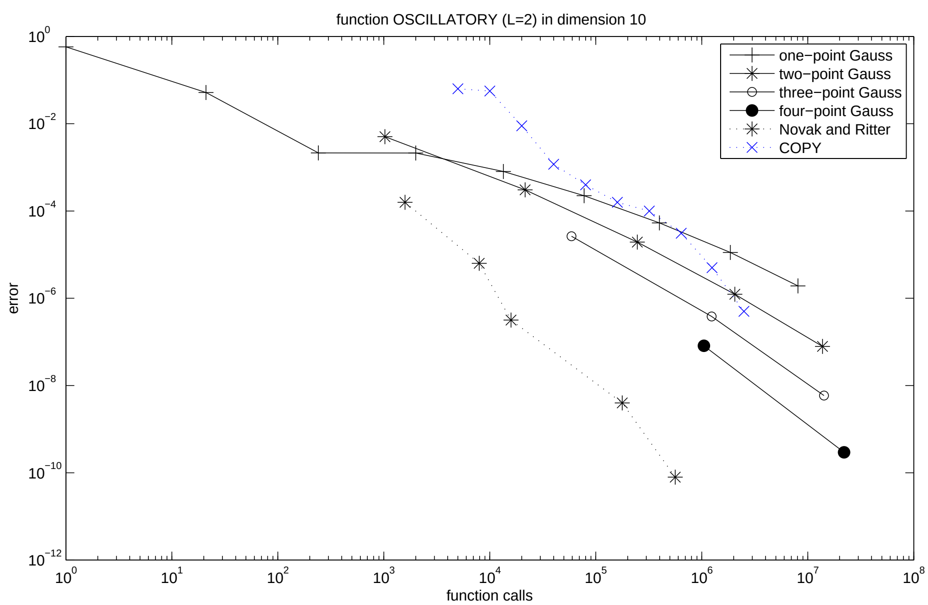

Figure 5.3: Median of absolute error of family (1), 20 integrands.

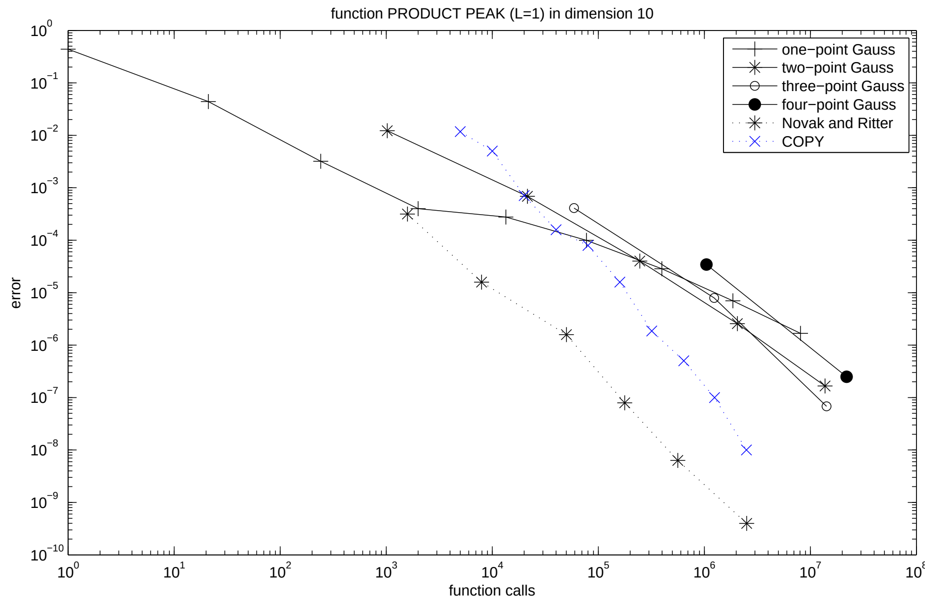

Figure 5.4: Median of absolute error of family (2), 20 integrands.

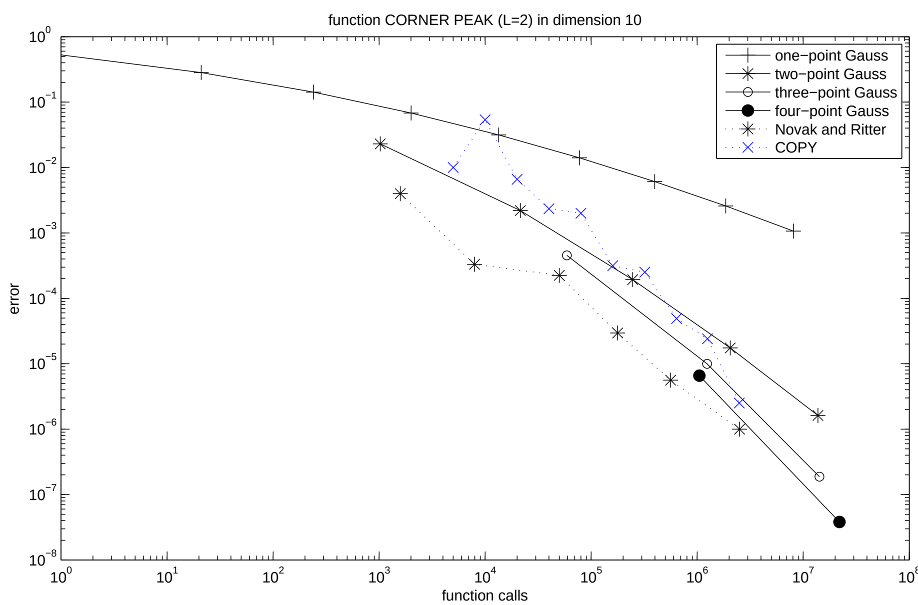

Figure 5.5: Median of absolute error of family (3), 20 integrands.

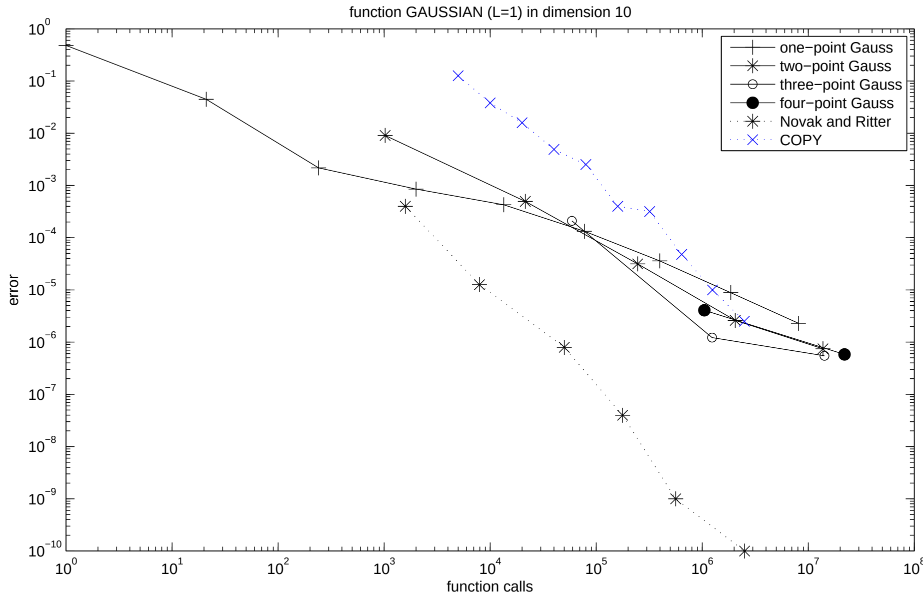

Figure 5.6: Median of absolute error of family (4), 20 integrands.

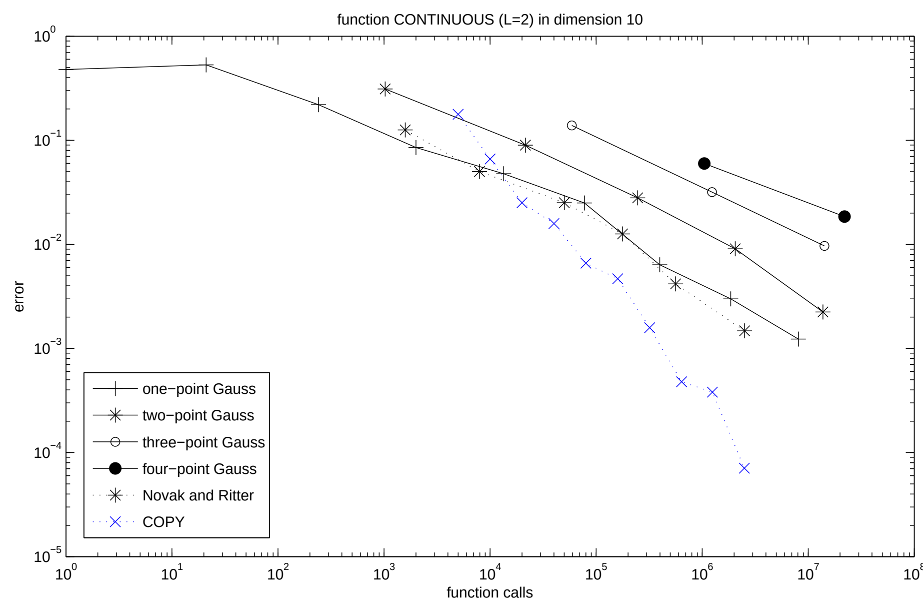

Figure 5.7: Median of absolute error of family (5), 20 integrands.

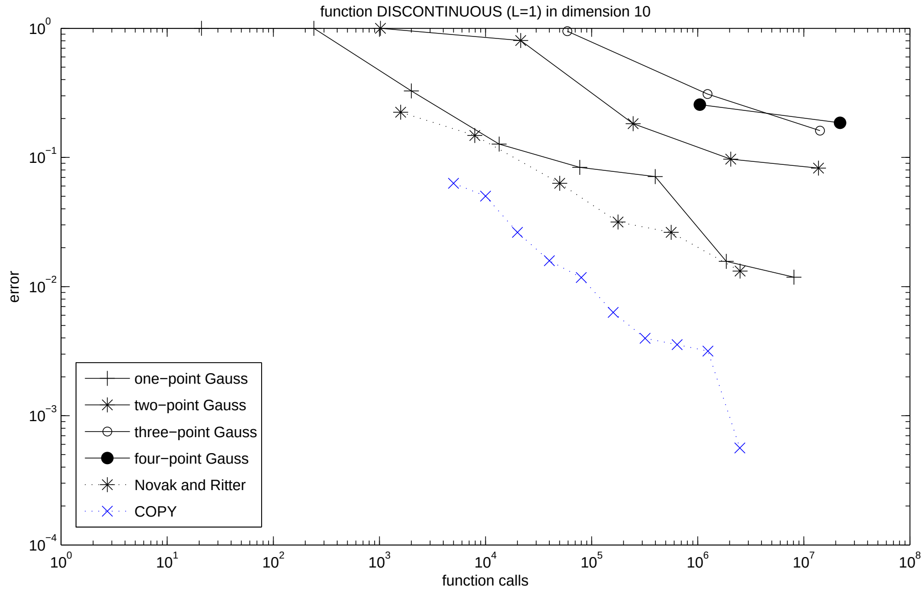

Figure 5.8: Median of absolute error of family (6), 20 integrands.

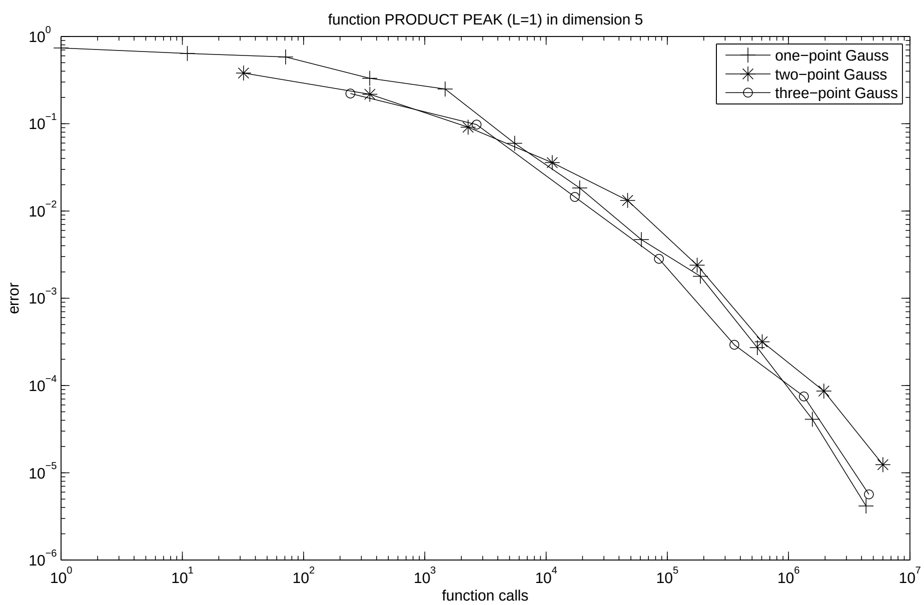

Figure 5.9: Median of absolute error of family (2), 20 integrands.

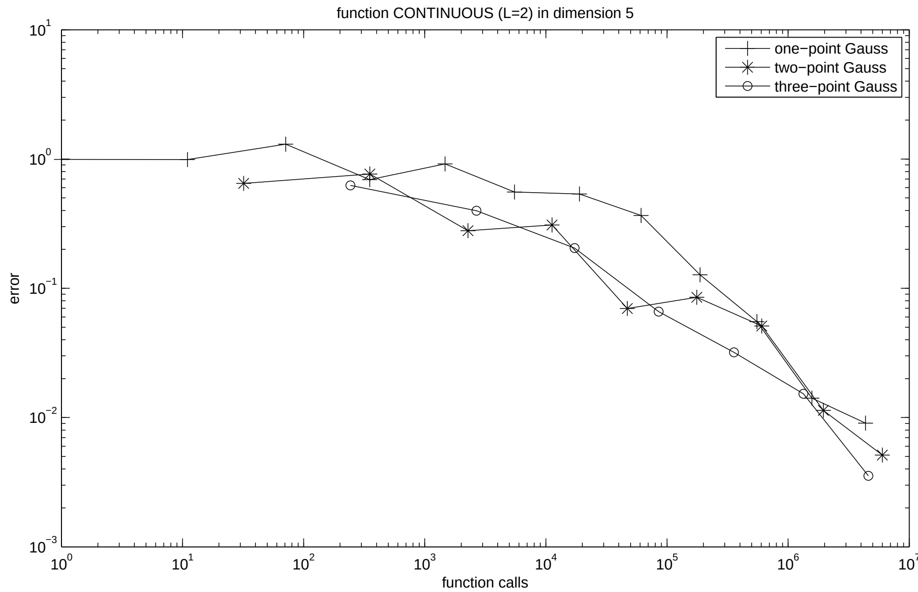

Figure 5.10: Median of absolute error of family (5), 20 integrands.

We have tested our method by using Gauss rules as underlying one-dimensional quadrature Q_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 [63], 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 [63].

It would be of interest to test our algorithm by considering functions which are more favorable with regard to our cubature rule, e.g. some discontinuous functions lying in our multiwavelet spaces, and compare the results to other methods.

Chapter 6

Cubature Formulas Based on Wavelet Frames

The aim of this chapter is to generalize the results and cubature methods we have discussed in Chapter 5. As mentioned in Chapter 3 the adaption of a general multiresolution analysis that provides higher regularity to bounded domains is a difficult task by itself. So the following dilemma appears: on the one hand we are interested in wavelets that guaranty an adapted approximation order, but on the other hand this request yields to difficulties on the boundary of the domains. To circumvent this problems and to get more general propositions we prefer a formulation via tight frames. A so called frame multiresolution analysis, respectively a non-stationary multiresolution analysis (NMRA), seems to be a useful tool for the generalization of the wavelet based cubature methods we have presented and analyzed before.

We give a general approximation of the integral operator which is optimal for one-dimensional function spaces defined via discrete norms. Because of the weak formulation of the considered MRA we have to take a little extra care, but the error analysis is quite similar to the analysis we established in the multiwavelet case. By the higher regularity of the associated wavelets we get norm equivalences to classical Sobolev spaces which also have more regularity. Consequently the corresponding approximation yields to a higher order of convergence on classical function spaces. We consider an example of NMRA based on univariate B-splines and take a closer look to the corresponding spline quadrature. The multivariate approximation we discuss is based on a variation of Smolyak's construction. More detailed, we extend the one-dimensional approximation to a d-dimensional one via an anisotropic version of Smolyak's construction, sometimes this is called adaptivity in the sense of a-priori knowledge. The idea behind this is quite simple, if there is a a-priori knowledge of the dependence of smoothness and directions we can reduce the overall cost of the algorithm and nevertheless we get accurate error bounds. Our aim is to provide an anisotropic cubature that guaranties a (nearly) optimal worst case error on spaces that consists of function with different smoothness in different directions.

This chapter is organized as follows: In the first section we give a brief description of frame multiresolution analysis. We recall the main requirements discussed in [11] that guaranty the existence of an so-called non-stationary multiresolution tight frame. In Section 6.2 we point out the equivalences of discrete norms defined via wavelet frames and classical function spaces like Sobolev spaces. Afterwords, in Section 6.3 we give the definition of an asymptotic optimal approximation (quadrature) for one-dimensional wavelet spaces. In Section 6.4 we give the explicit definition of a spline quadrature. In Section 6.5 we use an anisotropic version of Smolyak's construction to obtain from the asymptotic optimal one-dimensional approximation a nearly optimal d-dimensional approximation, with respect to the considered function spaces. We also provide lower bounds for the worst case error of any cubature rule on those spaces. In Section 6.6 we report on several numerical tests which allow us to compare our spline cubature to established methods and the cubature based on multiwavelets.

6.1 A frame multiresolution analysis

In this section we focus on those aspects of frame multiresolution analysis and a corresponding construction of wavelet frames which are useful for our purpose to construct optimal quadrature rules. For a more detailed discussion on this topic we refer the reader to [9, 11, 21] and the literature mentioned therein. First of all we give the definition of a NMRA, where we follow the lines and frequency the notation used in [11]. Subsequently we define an approximate dual which guaranties in the later subsections the existence of wavelet frames with vanishing moments. Observe that the order of the given approximate dual is directly related to the number of vanishing moments. The definition of a tight frame we will use in this section and respectively in the rest of this chapter is slightly modified with respect to the approximate duals.

6.1.1 Non-stationary multiresolution analysis

The definition of a non-stationary multiresolution analysis on bounded intervals is in some aspects quite similar to the definition of the MRA we have used in Chapter 3 see, e.g. [11]. But due to the fact that we do not consider orthogonal complement spaces nor a biorthogonal one we have to do a little extra work to realize wavelets with vanishing moments. Hence, the concept of so called approximate duals play the important part in the construction of those wavelet frames. Observe that the requirement we make in this subsection will hold for the rest of this chapter.

Definition 6.1.1. Let $\Omega \subset \mathbb{R}$ be a bounded interval. In general, a non-stationary multiresolution analysis (NMRA) consists of a nested family of finite dimensional subspaces

$$V_0 \subset V_1 \subset \cdots \subset V_j \subset V_{j+1} \subset \cdots \subset \mathrm{L}^2(\Omega) \tag{6.1.1}$$

such that the following holds

$$\overline{\bigcup_j V_j} = \mathrm{L}^2(\Omega) \quad \text{and the space } \; V_j \text{ is spanned by } \Phi_j := [\phi_{j,k}]_{k\in\Delta_j}\,, \tag{6.1.2}$$

where Δ_j is a suitable index set with cardinality $|\Delta_j| \geq \dim V_j$. The generator $[\phi_{0,k}]_{k\in\Delta_0}$ that spanned the space V_0 is called scaling function or a refinable function.

For our intention it is useful to assume that Φ_j is a (Riesz-) basis of the space V_j and $\phi_{j,k}$ satisfy the locality condition that $\max_{k\in\Delta_j} \mathrm{vol}(\mathrm{supp}\,\phi_{j,k})$ converges to zero if $j \to \infty$, for simplicity assume $\mathrm{vol}(\mathrm{supp}\,\phi_{j,k}) \sim 2^{-j}$. Note that here and in the sequel $\Phi_j := [\phi_{j,k}]_{k\in\Delta_j}$ has to be understood as a row vector. Furthermore, we assume

$$\|\phi_{j,k}\| \lesssim 1, \quad \text{for all } k \in \Delta_j \tag{6.1.3}$$

and that there is a moderate overlap of the supports $\phi_{j,k}$, more precisely we define

$$I_k^j := \mathrm{supp}\,\phi_{j,k}, \text{ for all } k \in \Delta_j, \tag{6.1.4}$$

and assume

$$\left| k' \in \Delta_j \;:\; I_{k'}^j \cap I_k^j \neq \emptyset \right| \lesssim 1, \text{ for all } k \in \Delta_j. \tag{6.1.5}$$

Similar to the previous cases where we considered Haar wavelets or multiwavelets we are interested in tight wavelet frames with vanishing moments. Hence we assume that the space V_0 contains the set Π_n of polynomials of order n, i.e., of degree strictly smaller than n, on Ω. Now, the concept of wavelet frames is to choose functions (wavelets) $\Psi_j := [\psi_{j,k}]_{k\in\nabla_j}$, where $\nabla_j \sim \Delta_{j+1} \setminus \Delta_j$, from V_{j+1} that compose the j^{th} level W_j of the tight frame of $\mathrm{L}^2(\Omega)$, such that

$$V_{j+1} = V_j + W_j \tag{6.1.6}$$

and the $\psi_{j,k}$ are also local with respect to the corresponding scale j. As mentioned in the introduction we follow the lines and frequently the expedient notation of Chui, He and Stöckler [11] and formulate the wavelets in terms of matrices. Let $M_j \in \mathbb{R}^{|\Delta_{j+1}|\times|\Delta_j|}$ be a matrix that describes the so called refinement relation

$$\Phi_j = \Phi_{j+1} M_j. \tag{6.1.7}$$

The complete scale relation is given by the refinement relation described by the matrices M_j and matrices $Q_j \in \mathbb{R}^{|\Delta_{j+1}| \times |\nabla_j|}$ such that the following holds,

$$\Psi_j = \Phi_{j+1} Q_j. \tag{6.1.8}$$

We are mainly interested in matrices $Q_j = [q_{i,k}^{(j)}]_{i \in \Delta_{j+1}, k \in \nabla_j}$ with the comfortable property that the resulting Ψ_j also satisfies the localization property

$$\text{vol}(\text{supp } \psi_{j,k}) \to 0.$$

Thus, we consider so called banded matrices Q_j with

$$q_{i,k}^{(j)} = 0 \text{ for all } i < i_j(k) \text{ and } i > i_j(k) + m, \tag{6.1.9}$$

where $\{i_j(k)\}_{k \in \nabla_j}$, are nondecreasing sequences. The crucial point is this condition assure that every wavelet $\psi_{j,k}$ is a linear combination of at most $m+1$ elements of Φ_{j+1}. Furthermore, we assume

$$\left| q_{i,k}^{(j)} \right| \lesssim 1$$

for all $i \in \Delta_{j+1}, k \in \nabla_j$.

A direct consequence of the assumption that the family of functions $\{\phi_{j,k}\}_{k \in \Delta_j}$ build a (Riesz-) basis of the space V_j is that the Gramian matrix

$$\Gamma_j = \left[\langle \phi_{j,k}, \phi_{j,\tilde{k}} \rangle \right]_{k, \tilde{k} \in \Delta_j} \tag{6.1.10}$$

is symmetric positive definite and the dual basis $\tilde{\Phi}_j$ is given by

$$\tilde{\Phi}_j = \Phi_j \Gamma_j^{-1}. \tag{6.1.11}$$

Remark 6.1.2. By the assumption that the supports of the functions have moderate overlap (this will specified later) it is clear that the Gramian matrix is banded, bounded with bounded inverse. Generally, banded matrices do not have a banded inverse. But the entries $\gamma_j^{l,m}$ of Γ_j decay exponentially, see i.e. [25], there exist constants $c > 0$ and $0 < \lambda < 1$ such that

$$\left| \gamma_j^{m,l} \right| \leq c \lambda^{|l-m|}. \tag{6.1.12}$$

Since the dual basis is given by (6.1.11) it follows that the functions $\tilde{\phi}_{j,k}$ decay exponentially.

The fact that the dual basis decays exponentially under proper requirements motivates the following definition of an approximate dual.

Definition 6.1.3. For any symmetric positive semi-definite matrix $S_j = [s_{l,m}^{(j)}]_{l,m \in \Delta_j}$ we define the quadratic form T_j by

$$T_j f := [\langle f, \phi_{j,k} \rangle]_{k \in \Delta_j} S_j \, [\langle f, \phi_{j,k} \rangle]_{k \in \Delta_j}^T \ , \quad f \in \mathrm{L}^2(\Omega), \tag{6.1.13}$$

and the function vector

$$\Phi_j^{S_j} = \left\{ \phi_{j,k}^{S_j} \right\}_{k \in \Delta_j} = \Phi_j \cdot S^j. \tag{6.1.14}$$

$\Phi_j^{S_j}$ is called an approximate dual of order n if

$$\langle f, \Phi_j^{S_j} \rangle = \langle f, \tilde{\Phi}_j \rangle, \tag{6.1.15}$$

for all $f \in \Pi_n$.

Of special interest are banded symmetric positive semi-definite matrices S_j that approximate Γ_j^{-1}, such that $\Phi_j^{S_j}$ are approximate duals of Φ_j. In this case the approximate duals also satisfies a locally condition. In particular those approximate duals are fundamental to assure the existence of NMRA tight frames such that the corresponding wavelets have a pre-set number of vanishing moments.

6.1.2 Existence of an NMRA tight frame

The usual definition of a tight frame we recalled for reasons of close a representation in Chapter 2 will get a slightly modification such that the modified definition is closer to the problem of compute the dual scaling functions. Roughly speaking we only demand the exactness for the polynomial space Π_n on level -1.

We follow the line of [11] and say the family $\{\Psi_j\}_{j \geq 0}$ defined above, constitutes an MRA tight frame of $\mathrm{L}^2(\Omega)$ with respect to the quadratic form T_0, if

$$T_0 f + \sum_{j \geq 0} \sum_{k \in \nabla_j} |\langle f, \psi_{j,k} \rangle|^2 = A \, \|f\|^2 \,, \quad \text{for all } f \in \mathrm{L}^2(\Omega). \tag{6.1.16}$$

Assume that S_0 is an symmetric positive semi-definite matrix that define an approximate dual of Φ_0 of order n such that $T_0 f \leq \|f\|^2$ for all $f \in \mathrm{L}^2(\Omega)$. Also, let $\{\Psi_j\}_{j \geq 0} = \{\Phi_{j+1} Q_j\}_{j \geq 0}$ and $\Psi_j = [\psi_{j,k}]_{k \in \nabla_j}$. It is known [11, Corollary 3.3] that in this case the wavelet have n vanishing moments and define a tight frame in the sense of (6.1.16), if and only if there exist symmetric positive semi-definite matrices $S_j \in \mathbb{R}^{|\Delta_j| \times |\Delta_j|}$ such that:

- The quadratic forms T_j satisfies

$$\lim_{j \to \infty} T_j f = \|f\|^2 \ , \quad \text{for all } f \in \mathrm{L}^2(\Omega), \tag{6.1.17}$$

- for each $j \geq 0$, the Q_j, S_j and S_{j+1} satisfy the identity

$$S_{j+1} - M_j S_j M_j^T = Q_j Q_j^T \tag{6.1.18}$$

and S_j defines an approximate dual of Φ_j of order n.

Consequently, we need a slightly modification for the $\phi_{0,k}$'s to give an expansion for $f \in \mathrm{L}^2(\Omega)$. We define

$$\Phi_0^{S'} = [\phi_{0,k}^{S'}]_{k\in\Delta_0} := \Phi_0 S_0^{1/2},$$

where $S_0^{1/2}$ is the square root of S_0 in the sense of symmetric positive operators. In this setting the frame condition in (6.1.16) yields to the bounded frame operator

$$\mathcal{A} : \mathrm{L}^2(\Omega) \to l^2, \quad f \mapsto \{\langle f, \Psi_j\rangle\}_{j\geq -1},$$

where $\Psi_{-1} := \Phi_0 S_0^{1/2}$. Thus we get the representation of $f \in \mathrm{L}^2(\Omega)$,

$$\begin{aligned} f &= \sum_{j\geq -1} \langle f, \tilde{\Psi}_j\rangle \Psi_j^T = \langle f, \tilde{\Phi}_0^{S'}\rangle (\Phi_0^{S'})^T + \sum_{j\geq 0} \langle f, \tilde{\Psi}_j\rangle \Psi_j^T \\ &= \frac{1}{A}\left(\langle f, \Phi_0 S^{1/2}\rangle (\Phi_0 S^{1/2})^T + \sum_{j\geq 0} \langle f, \Psi_j\rangle \Psi_j^T \right). \end{aligned}$$

6.2 Norm equivalences

As mentioned in Chapter 3 there are some well known and extensively studied possibilities in the characterization of classical function spaces like Sobolev or Besov spaces under the appropriation of orthogonal or bi-orthogonal wavelet expressions, see i.e. [17, 18]. Anyway, this characterization also holds in a more general setting, see i.e. [13, 30, 58] and the literature mentioned therein.

First of all we show that the approximation property also holds in our setting if there exists a banded symmetric positive semi-definite matrix S_j such that $\Phi_j S_j$ are approximate duals of Φ_j. The proof is a slightly modification of [18, Lemma 2.1], more detailed we get the following lemma:

Lemma 6.2.1. *Let $\{\phi_{0,k}\}_{k\in\Delta_0}$ be the generator of a (stable) frame multiresolution analysis $V_0 \subset V_1 \subset \ldots$ such that $\Pi_n \subset V_0$ and $\{\phi_{0,k}\}_{k\in\Delta_0}$. Under the above requirements that approximate duals exist, we get for all $f \in H^s$, $s \leq n$ that*

$$\inf_{v\in V_j} \|f - v\|_{\mathrm{L}^2(\Omega)} \lesssim 2^{-js} \|f\|_{H^s(\Omega)}.$$

Proof. Let $f \in H^s$ and $p \in \Pi_n(\Omega)$. Let S_j be a banded symmetric positive semi-definite matrix such that $\Phi_j^{S_j} = \Phi_j S_j$ are approximate duals of Φ_j. Apparently we get

$$\begin{aligned}
\left\| f - \langle f, \Phi_j^{S_j} \rangle \Phi_j^T \right\|_{\mathrm{L}^2(I_k^j)} &\leq \|f - p\|_{\mathrm{L}^2(I_k^j)} + \left\| p - \langle f, \Phi_j^{S_j} \rangle \Phi_j^T \right\|_{\mathrm{L}^2(I_k^j)} \\
&= \|f - p\|_{\mathrm{L}^2(I_k^j)} + \left\| \langle (p - f), \Phi_j^{S_j} \rangle \Phi_j^T \right\|_{\mathrm{L}^2(I_k^j)} \\
&\lesssim \|f - p\|_{\mathrm{L}^2(I_k^j)} + \|f - p\|_{\mathrm{L}^2(\mathbf{I}_k^j)} \lesssim \|f - p\|_{\mathrm{L}^2(\mathbf{I}_k^j)},
\end{aligned}$$

where

$$\mathbf{I}_k^j := \bigcup_{k' \colon I_{k'}^j \cap I_k^j \neq \emptyset} \mathrm{supp} \phi_{j,k\prime}^{S_j}.$$

Since p was arbitrary, the well known Bramble-Hilbert Theorem yields

$$\|f - p\|_{\mathrm{L}^2(\mathbf{I}_k^j)} \lesssim \mathrm{vol}(\mathbf{I}_k^j)^s \|f\|_{H^s(\mathbf{I}_k^j)} \lesssim \max_{k'} \mathrm{vol}(\tilde{I}_{k'}^j)^s \|f\|_{H^s(\mathbf{I}_k^j)} \lesssim 2^{-js} \|f\|_{H^s(\mathbf{I}_k^j)}.$$

Now, squaring and summing over $k \in \Delta_j$, and taking into account that only constant of the I_k^j overlap and the fact that the approximate duals satisfies also a locally condition finish the proof. $\square$

Under the consideration of the approximation property (Lemma 6.2.1) it is straight forward to prove a Jackson type estimate. We have to define the so-called K-functional, observe that we don't do this in the most general form, we set

$$K_s(f, 2^{-j}, \Omega)_2 = \inf_{v \in H^s(\Omega)} \left\{ \|f - v\|_{\mathrm{L}^2(\Omega)} + 2^{-js} \|v\|_{H^s(\Omega)} \right\}.$$

It is well known that in this case the K-functional is equivalent to the modulus of smoothness, see, e.g. [39, 40] and can be used to defined classical Sobolev spaces, more detailed, we have

$$K_s(f, 2^{-j}, \Omega)_2 \sim \omega_s(f, 2^{-j}, \Omega)_2, \quad \text{and} \quad \|f\|_{B_{2,2}^s(\Omega)} \sim \|f\|_{H^s(\Omega)}.$$

The estimate we require is

$$\inf_{f_j \in V^j} \|f - f_j\|_2 \lesssim \omega_s(f, 2^{-j}, \Omega)_2. \tag{6.2.1}$$

Corollary 6.2.2. *Let $\{\phi_{0,k}\}_{k \in \Delta_0}$ be the generator of a (stable) frame multiresolution analysis $V_0 \subset V_1 \subset \ldots$ such that $\Pi_n \subset V_0$ and $\{\phi_{0,k}\}_{k \in \Delta_0}$ comply with the requirements in Lemma 6.2.1. Then,*

$$\inf_{f_j \in V_j} \|f - f_j\|_2 \lesssim K_s(f, 2^{-j}, \Omega)_2. \tag{6.2.2}$$

Proof. Let $s \leq n$. For $f \in H^s$ and arbitrary $v \in H^s$ we get

$$\begin{aligned}
\inf_{f_j \in V_j} \|f - f_j\|_2 &\leq \left\| f - \langle f, \tilde{\Phi}_j \rangle \Phi_j^T \right\|_2 \leq \|f - v\|_2 + \left\| \langle f, \tilde{\Phi}_j \rangle \Phi_j^T - v \right\|_2 \\
&\leq \|f - v\|_2 + \left\| \langle v, \tilde{\Phi}_j \rangle \Phi_j^T - v \right\|_2 + \left\| \langle (f - v), \tilde{\Phi}_j \rangle \Phi_j^T \right\|_2 \\
&\leq 2 \|f - v\|_2 + \left\| \langle v, \tilde{\Phi}_j \rangle \Phi_j^T - v \right\|_2 .
\end{aligned}$$

The approximation property told us

$$\left\| \langle v, \tilde{\Phi}_j \rangle \Phi_j^T - v \right\|_2 \lesssim 2^{-js} \|v\|_{H^s}$$

and consequently we get

$$\left\| f - \langle f, \tilde{\Phi}_j \rangle \Phi_j^T \right\|_2 \lesssim \|f - v\|_2 + 2^{-js} \|v\|_{H^s} .$$

If we take the infimum over all $v \in H^s$ the estimate follows. □

Hence, a so-called Jackson type estimate we have discussed in Chapter 3 is satisfied, and it remains to show that also the inverse (Bernstein) estimate holds. But due to the assumption that the generator satisfies the localization property, the generator is normalized $\|\phi_{j,k}\| \lesssim 1$ and the Riesz basis property (stability) the following lemma is a direct consequence of [14].

Lemma 6.2.3. *For the above setting the inverse estimate*

$$\|f_j\|_{H^s(\Omega)} \lesssim 2^{js} \|f_j\| , \quad f_j \in V_j,$$

holds where

$$s < \sup \{ s : \ \phi_{0,k} \in H^s(\Omega), \ k \in \Delta_0 \} .$$

Combining these facts and the fact that the Sobolev regularity of the generator is proportional to n we will get the following corollary. Since we considered only slightly modifications of the well known classical settings we skip the proof again and refer the reader to the relevant literature we cited in Chapter 3, e.g. [12, 17, 58].

Corollary 6.2.4. *Under the requirements of Corollary 6.2.2 and Lemma 6.2.3 we get the following norm equivalence,*

$$\|f\|^2_{H^s(\Omega)} \sim T_0 f + \sum_{j \geq 0} \sum_{k \in \nabla_j} 2^{j2s} \left| \langle f, \tilde{\psi}_{j,k} \rangle \right|^2 ,$$

where $s < \sup \{ s : \ \phi_{0,k} \in H^s(\Omega), \ k \in \Delta_0 \}$.

Remark 6.2.5. Observe that the case where the smoothness parameter s is chosen to be zero the norm equivalence in Corollary 6.2.4 corresponds to the definition of tight frame we use in (6.1.16).

6.3 Optimal quadrature for wavelet spaces

For this section let us assume that the family $\{\Psi_j\}_{j\geq 0}$ together with the generator $\{\phi\}_{k\in\Delta_0}$ form a MRA tight frame with n vanishing moments and respect to T_0, with all requirements made in Section 6.1. With the essential outcome of the previous sections in mind, where we have characterized function spaces with wavelet expressions we define for $s > 0$ the discrete norm

$$\begin{aligned} \|f\|_s^2 &:= T_0 f + \sum_{j\geq 0}\sum_{k\in\nabla_j} 2^{j2s}\left|\langle f, \tilde{\psi}_{j,k}\rangle\right|^2 && (6.3.1)\\ &= T_0 f + |f|_s^2, && (6.3.2)\end{aligned}$$

on the space

$$H_s := \left\{f \in \mathrm{L}^2(\Omega)\ :\ \|f\|_s < \infty\right\}, \tag{6.3.3}$$

consisting of functions whose wavelet coefficients decrease rapidly. Point evaluations are obviously well defined on the linear span of the functions $\phi_{0,k'}$ and $\psi_{j,k}$, $j \geq 0$, $k' \in \Delta_0$ and $k \in \nabla_j$. Moreover, it is not hard to see that they can be extended to bounded linear functionals on H_s as long as $s > 1/2$. On these spaces quadrature formulas are therefore well defined. Our aim is to give a general construction of one-dimensional quadrature formulas which are optimal for the spaces H_s. Those formulas seems to be a natural approximation of the integral operator, with respect to the structure of the finite- dimensional subspaces of H_s. We give an approximation that is exact on the approximation spaces $V_0 \subset V_1 \subset \cdots \subset V_j \subset V_{j+1} \subset \cdots \subset \mathrm{L}^2(\Omega)$ up to a so called critical level.

Definition 6.3.1. For a given critical level l we define an operator

$$\mathcal{A}_k^l : \mathrm{L}^2([0,1]) \times \Delta_l \to \mathbb{R}$$

by

$$\mathcal{A}_k^l v = \langle v, \tilde{\phi}_{j,k}\rangle_q. \tag{6.3.4}$$

And we define the approximation of the integral by

$$A(l,1)v := \sum_{k\in\Delta_l}\omega_k\langle v, \tilde{\phi}_{l,k}\rangle_q = \sum_{k\in\Delta_l}\omega_k\mathcal{A}_k^l v, \tag{6.3.5}$$

where

$$\omega_k = \int_\Omega \phi_{l,k}dx \sim 2^{-l/2} \tag{6.3.6}$$

and

$$\langle v, \tilde{\phi}_{l,k}\rangle_q = \langle v, \tilde{\phi}_{l,k}\rangle, \quad \text{for all } v \in V_l.$$

Observe, that these definition make sense also for the case where the approximate dual $\phi_{l,k}^S$ is used instead of the dual $\tilde{\phi}_{l,k}$. Later, we are primarily interested in the case where $\mathcal{A}_k^l$ is a point evaluation on V_l and the definition in (6.3.5) gives us a one-dimensional quadrature rule on level l. In this case the number of point evaluations used by the algorithm $A(l,1)$ is given by $|A(l,1)| = |\Delta_l| \sim 2^l$. Apparently this approximation is exact up to the space V_l.

Lemma 6.3.2. *The approximation (quadrature) $A(l,1)$ is exact on the space V_l.*

Proof. Let $v \in V_l$. We get the representation

$$v = \sum_{k \in \Delta_l} \langle v, \tilde{\phi}_{l,k} \rangle \phi_{l,k}, \tag{6.3.7}$$

consequently we get

$$\int_\Omega v\, dx = \int_\Omega \sum_{k \in \Delta_l} \langle v, \tilde{\phi}_{l,k} \rangle \phi_{l,k} dx = \sum_{k \in \Delta_l} \langle v, \tilde{\phi}_{l,k} \rangle \int_\Omega \phi_{l,k} dx = \sum_{k \in \Delta_l} \omega_k \mathcal{A}_k^l v.$$

□

For the second case, where the approximate dual is considered, we get

Lemma 6.3.3. *The approximation $A(l,1)$ defined by*

$$A(l,1)f := \sum_{k \in \Delta_l} \langle v, \phi_{l,k}^S \rangle_q \int_\Omega \phi_{l,k} dx$$

is exact on the difference spaces $\sum_{j \geq 0} W_{l-1}$.

Proof. Let $W \in \sum_{j \geq 0} W_{l-1}$. We get the representation

$$\begin{aligned} w &= \sum_{l=1}^{L} \langle w, \Psi_{l-1} \rangle \Psi_{l-1}^T + \langle w, \Phi_0 S_0 \rangle \Phi_0^T \\ &= \sum_{l=1}^{L} \langle w, \Phi_l Q_{l-1} Q_{l-1}^T \rangle \Phi_l^T + \langle w, \Phi_0 S_0 \rangle \Phi_0^T \\ &= \sum_{l=1}^{L} \langle w, \Phi_l S_L \rangle \Phi_l^T - \langle w, \Phi_{l-1} S_{l-1} \rangle \Phi_{l-1}^T + \langle w, \Phi_0 S_0 \rangle \Phi_0^T \\ &= \langle w, \Phi_L S_L \rangle \Phi_L^T. \end{aligned}$$

Hence

$$\int_0^1 w(x) dx = \int_0^1 \sum_{k \in \Delta_L} \langle w, \phi_{L,k}^{S_L} \rangle \phi_{L,k}(x) dx = \sum_{k \in \Delta_L} \langle w, \phi_{L,k}^{S_L} \rangle_q \int_0^1 \phi_{L,k}(x) dx.$$

□

6.3.1 Error bounds

Quite similar to the cases we have considered before we analyze the worst case error of the approximation $A(l,1)$ with the representation of a function $f \in H_s$. To accept that the given approximation is an asymptotic optimal one, we also give a lower bound for the quadrature error on H_s.

Theorem 6.3.4. *For $s > 1/2$ we get*

$$\mathrm{err}(H_s, A(l,1)) \lesssim 2^{-ls} \sim |(A(l,1)|^{-s}\,. \tag{6.3.8}$$

Proof. Let $f \in H_s$, the quadrature error is given by

$$\mathrm{err}(f, A(l,1)) = \left| \int_\Omega f\,dx - A(l,1)f \right|$$

$$= \left| \sum_{k\in\Delta_0} \langle f, \phi^S_{0,k}\rangle (I - A(l,1))[\phi_{0,k}] + \sum_{j\geq 0}\sum_{k\in\nabla_j} \langle f, \tilde{\psi}_{j,k}\rangle (I - A(l,1))[\psi_{j,k}] \right|$$

Since the quadrature $A(l,1)$ is exact on V_l and we have vanishing moments of $\psi_{j,k}$ we get the quadrature error

$$\begin{aligned}\mathrm{err}(f, A(l,1)) &= \left| \sum_{j\geq 0}\sum_{k\in\nabla_j} \langle f, \tilde{\psi}_{j,k}\rangle (I - A(l,1))[\psi_{j,k}] \right| \\ &= \left| \sum_{j\geq 0}\sum_{k\in\nabla_j} \langle f, \tilde{\psi}_{j,k}\rangle A(l,1)[\psi_{j,k}] \right|.\end{aligned}$$

The Cauchy-Schwarz inequality yields

$$\begin{aligned}\mathrm{err}(f, A(l,1)) &\leq \left(\sum_{j\geq 0}\sum_{k\in\nabla_j} 2^{j2s} \left|\langle f, \tilde{\psi}_{j,k}\rangle\right|^2 \right)^{1/2} \left(\sum_{j\geq 0}\sum_{k\in\nabla_j} 2^{-j2s} \left|A(l,1)\psi_{j,k}\right|^2 \right)^{1/2} \\ &= |f|_s \left(\sum_{j\geq 0}\sum_{k\in\nabla_j} 2^{-j2s} \left|A(l,1)\psi_{j,k}\right|^2 \right)^{1/2}.\end{aligned}$$

Consequently,

$$\begin{aligned}\mathrm{err}(H_s, A(l,1))^2 &\leq \sum_{j\geq 0}\sum_{k\in\nabla_j} 2^{-j2s} \left|A(l,1)\psi_{j,k}\right|^2 \\ &= \sum_{j\geq l}\sum_{k\in\nabla_j} 2^{-j2s} \left(\sum_{k_1\in\Delta_l} \omega_{k_1} \langle \psi_{j,k}, \tilde{\phi}_{l,k_1}\rangle_q \right)^2 \\ &= \sum_{j\geq l}\sum_{k\in\nabla_j} 2^{-j2s} \left(\sum_{k_1\in\Delta_l} \omega_{k_1} \sum_{i\in\Delta_{j+1}} q^{(j)}_{i,k} \langle \phi_{j+1,i}, \tilde{\phi}_{l,k_1}\rangle_q \right)^2.\end{aligned}$$

By some easy calculations and by the fact we have exponentially decay for the $\gamma_l^{k_2,k_1}$ we get for the critical term

$$\begin{aligned}
& \sum_{k\in\nabla_j}\left(\sum_{k_1\in\Delta_l}\omega_{k_1}\sum_{i\in\Delta_{j+1}} q_{i,k}^{(j)}\langle\phi_{j+1,i},\tilde{\phi}_{l,k_1}\rangle_q\right)^2 \\
\leq\ & \sum_{k\in\nabla_j}\left(\sum_{k_1\in\Delta_l}\omega_{k_1}\sum_{i=i_j(k)}^{i_j(k)+m}\left|q_{i,k}^{(j)}\right|\|\phi_{j+1,i}\|_\infty\langle\mathbf{1}_{\mathrm{supp}\phi_{j+1,i}},\tilde{\phi}_{l,k_1}\rangle_q\right)^2 \\
=\ & \sum_{k\in\nabla_j}\left(\sum_{k_1\in\Delta_l}\omega_{k_1}\sum_{i=i_j(k)}^{i_j(k)+m}\left|q_{i,k}^{(j)}\right|\|\phi_{j+1,i}\|_\infty\sum_{k_2\in\Delta_l}\gamma_l^{k_2,k_1}\langle\mathbf{1}_{I_i^{j+1}},\phi_{l,k_2}\rangle_q\right)^2 \\
\leq\ & \sum_{k\in\nabla_j}\left(\sum_{k_1\in\Delta_l}\omega_{k_1}\sum_{i=i_j(k)}^{i_j(k)+m}\left|q_{i,k}^{(j)}\right|\|\phi_{j+1,i}\|_\infty\sum_{k_2\in\Delta_l}c\lambda^{|k_2-k_1|}\langle\mathbf{1}_{I_i^{j+1}},\phi_{l,k_2}\rangle_q\right)^2 \\
\lesssim\ & \max_{i\in\Delta_{j+1}}\|\phi_{j+1,i}\|_\infty^2\sum_{k\in\nabla_j}\left(\sum_{k_1\in\Delta_l}\omega_{k_1}\sum_{k_2\in\Delta_l}\lambda^{|k_2-k_1|}\langle\sum_{i=i_j(k)}^{i_j(k)+m}\mathbf{1}_{I_i^{j+1}},\phi_{l,k_2}\rangle_q\right)^2 .
\end{aligned}$$

Since we have moderate overlap in the sense of (6.1.5) we get for

$$\Delta^k := \left\{k\in\Delta_j\ :\ \left(\bigcup_{i=i_j(k)}^{i_j(k)+m} I_i^{j+1}\right)\bigcap I_k^l\neq\emptyset\right\},$$

that $\left|\Delta^k\right|\lesssim 1$. Let $\kappa=\kappa(k)\in\Delta_l$ be the unique element such that

$$\lambda^{|\kappa-k_1|}\langle\sum_{i=i_j(k)}^{i_j(k)+m}\mathbf{1}_{I_i^{j+1}},\phi_{l,\kappa}\rangle_q=\max_{k_2\in\Delta_l}\lambda^{|k_2-k_1|}\langle\sum_{i=i_j(k)}^{i_j(k)+m}\mathbf{1}_{I_i^{j+1}},\phi_{l,k_2}\rangle_q.$$

Then,

$$\begin{aligned}
& \sum_{k\in\nabla_j}\left(\sum_{k_1\in\Delta_l}\omega_{k_1}\sum_{k_2\in\Delta_l}\lambda^{|k_2-k_1|}\langle\sum_{i=i_j(k)}^{i_j(k)+m}\mathbf{1}_{I_i^{j+1}},\phi_{l,k_2}\rangle_q\right)^2 \\
=\ & \sum_{k\in\nabla_j}\left(\sum_{k_1\in\Delta_l}\omega_{k_1}\sum_{k_2\in\Delta^k}\lambda^{|k_2-k_1|}\langle\sum_{i=i_j(k)}^{i_j(k)+m}\mathbf{1}_{I_i^{j+1}},\phi_{l,k_2}\rangle_q\right)^2 \\
\leq\ & \sum_{k\in\nabla_j}\left(\sum_{k_1\in\Delta_l}\omega_{k_1}\left|\Delta^k\right|\lambda^{|\kappa(k)-k_1|}\langle\sum_{i=i_j(k)}^{i_j(k)+m}\mathbf{1}_{I_i^{j+1}},\phi_{l,\kappa(k)}\rangle_q\right)^2 ,
\end{aligned}$$

taking again the locality of $\phi_{l,\kappa}$ and (6.1.5) into account we get

$$\lesssim \sum_{\kappa\in\Delta_l}\left(\sum_{k_1\in\Delta_l}\omega_{k_1}\lambda^{|\kappa-k_1|}\langle \mathbf{1}_{I^l_\kappa},\phi_{l,\kappa}\rangle_q\right)^2 \lesssim \sum_{\kappa\in\Delta_l}\omega_\kappa^4.$$

Since $\|\phi_{j+1,k}\|\sim 1$ it is obviously $1\sim\|\phi_{j+1,k}\|_\infty^2\,\mathrm{vol}(I_k^{j+1})$ and we get

$$\begin{aligned}\mathrm{err}(H_s,A(l,1))^2 &\lesssim \sum_{j\geq l}2^{-j2s}\max_{i\in\Delta_{j+1}}\|\phi_{j+1,i}\|_\infty^2\,2^{-l}\\ &\lesssim 2^{-l}\sum_{j\geq l}2^{j(1-2s)}\lesssim 2^{-l2s}\sum_{j\geq 0}2^{j(1-2s)}\lesssim 2^{-l2s}.\end{aligned}$$

□

Remark 6.3.5. This error bound also hold if the quasi projector $f\mapsto\langle f,\Phi^s\rangle\Phi^T$ is considered and for the exactness, Lemma 6.3.3 is used. Some parts of the proof will by quite easier in this case, and so we skip the proof.

By the fact that we have n vanishing moments the proof of Theorem 6.3.4 also shows that $\mathrm{err}(f,A(l,1))\leq|f|_s\,2^{-ls}$ with the given semi-norm $|\cdot|_s$ that vanish if $f\in\Pi_n$. This fact naturally corresponds to the assumption that $\Pi_n\subset V_0$. Anyway, the error estimate in Theorem 6.3.4 is asymptotically optimal as the next theorem will reveal. We get the following lower bound.

Theorem 6.3.6. *Let $s>1/2$. There exists a constant $C>0$ such that for any quadrature rule Q_N using N sample points we have*

$$\mathrm{err}(H_s,\mathrm{Q}_N)\geq CN^{-s}.$$

Proof. Let $P\subset\Omega$, $|P|=N$ be a set of sample points used by an arbitrary quadrature rule A_N. We choose the integer L that satisfies

$$|\nabla_{L-1}|\sim 2^{L-1}\leq(K+1)N\leq 2^L\sim|\nabla_L|,$$

where K is the number of maximal overlap of the supports $J_k^j:=\mathrm{supp}\,\psi_{j,k}$, $k\in\nabla_j$. We define a function

$$f_L(x)=\begin{cases}1 & \text{for all } x\in J_k^L,\ k\in\nabla_L \quad\text{with}\quad J_k^L\cap P=\emptyset\\ 0 & \text{else}\end{cases}.$$

Hence, by some elementary calculations we get for the semi-norm of our candidate

$$\begin{aligned}|f_L|_s^2 &= \sum_{j\geq 0}\sum_{k\in\nabla_j}2^{j2s}\left|\langle f_L,\tilde\psi_{j,k}\rangle\right|^2 = A^{-2}\sum_{j\geq 0}\sum_{k\in\nabla_j}2^{j2s}\,|\langle f_L,\psi_{j,k}\rangle|^2\\ &= A^{-2}\sum_{0\leq j\leq L-1}\sum_{k\in\nabla_j}2^{j2s}\,|\langle f_L,\psi_{j,k}\rangle|^2\\ &\leq A^{-2}\sum_{0\leq j\leq L-1}\sum_{k\in\nabla_j}2^{j2s}\,\|\psi_{j,k}\|_\infty^2\,(\mathrm{vol}(\mathrm{supp}\,\psi_{j,k}))^2,\end{aligned}$$

by the assumption we made for the matrix Q^j that describes the scale relation we get

$$|f_l|_s^2 \leq A^{-2} \sum_{0\leq j\leq L-1} \sum_{k\in\nabla_j} 2^{j2s}(m+1) \max_{k'\in\Delta_{j+1}} \left| q_{k',k}^{(j)} \right| \|\phi_{j+1,k'}\|_\infty^2 \, (\mathrm{vol}(J_k^j))^2,$$

where J_k^j denote the support of $\psi_{j,k}$. Consequently,

$$\begin{aligned} |f_l|_s^2 &\lesssim \sum_{0\leq j\leq L-1} \sum_{k\in\nabla_j} 2^{j2s} \max_{k'\in\Delta_{j+1}} \mathrm{vol}(I_{k'}^{j+1})^{-1} (\mathrm{vol}(J_k^j))^2 \\ &\lesssim \sum_{0\leq j\leq L-1} \sum_{k\in\nabla_j} 2^{j2s} \mathrm{vol}(J_k^j) \lesssim 2^{L2s}. \end{aligned}$$

Thus, we get also for the norm

$$\|f_L\|_s^2 \lesssim 2^{L2s}.$$

Furthermore, we have

$$\int_\Omega f_L dx \geq \min_{k\in\nabla_L} \mathrm{vol}(J_k^L) \left(|\nabla_L| - KN\right) \geq C/(K+1).$$

Now we consider the normalized function $f_L/\left\|f_L\right\|_s$. The estimates above results in

$$\mathrm{err}(f_L/\left\|f_L\right\|_s, Q_N) = \frac{\left|\int_\Omega f_L dx - Q_N f_L\right|}{\left\|f_L\right\|_s} \gtrsim 2^{-Ls}.$$

□

6.4 Spline quadrature

To focus on practical aspects we will apply the general approximation (quadrature) methods in Section 6.3 to the case of tight frames of spline wavelets. Before we are able to define an optimal spline quadrature we have to recall some well known facts about B-splines. The general considerations in [11], which we recalled in Section 6.3 can be applied to the case of univariate B-splines for the construction of tight frames and therefore we are able to define an optimal and more or less natural quadrature rule on spline spaces.

6.4.1 Some facts about splines

For a better overview, we introduce some current shortcuts and necessary notations. For a function $f \in C^m(\mathbb{R})$ and a sequence of knots $t_i, \ldots, t_{i+m}$ we denote by $[t_i, \ldots, t_{i+m}]f$ the divided difference of order m at $t_i, \ldots, t_{i+m}$, defined by the leading coefficient of the Lagrange polynomial which interpolates the function f at $t_i, \ldots, t_{i+m}$. By definition,

$$[t_i, \ldots, t_{i+m}]f = 0 \quad \text{if } f \in \Pi_m \text{ and } [t_i, \ldots, t_{i+m}]x^m = 1.$$

Let $m, N \in \mathbb{N}$ and

$$\mathbf{t} = \{t_k \ : \ -m+1 \leq k \leq N+m\}$$

be a knot vector such that

$$t_k \leq t_{k+1} \text{ and } t_k < t_{k+m} \text{ for all } k, \tag{6.4.1}$$

$$t_{-m+1} = \cdots = t_0 = a \text{ and } t_{N+1} = \cdots = t_{N+m} = b. \tag{6.4.2}$$

The multiplicity μ_k of a knot $t_k \in \mathbf{t}$ is the number of times this knot is repeated in $\mathbf{t}$. The number m will denote the order of the spline function, N is the number of so called interior knots. Observe that $\mu_k \leq m$ for all k.

Definition 6.4.1. The normalized B-spline $N_{\mathbf{t},,m,k}$ of order m (or of degree $m-1$) is a function on $\mathbb{R}$ defined by

$$N_{\mathbf{t},m,k}(x) = (t_{k+m} - t_k)[t_k, \ldots, t_{k+m}](\cdot - x)_+^{m-1}, \ k \in \Delta,$$

where $\Delta = \{-m+1, \cdots, N\}$ denotes the corresponding index set.

It is well known that the $N_{\mathbf{t},m,k}$ has compact support $[t_k, t_{k+m}]$, $N_{\mathbf{t},m,k}$ is strictly positive inside this interval and is a polynomial of degree $m-1$ in each interval (t_i, t_{i+1}), $k \leq i \leq k+m-1$. It has $m - \mu_i - 1$ continues derivatives at t_i and the integral of $N_{\mathbf{t},m,k}$ is given by

$$\int_{\mathbb{R}} N_{\mathbf{t},m,k}(x)dx = \frac{t_{k+m} - t_k}{m} =: d_{\mathbf{t},m,k}.$$

By $S_{\mathbf{t},m}$ we denote the spline space, consists of all piecewise polynomials of degree $m-1$ on $\Omega = [a,b]$ with so called breakpoints $t_k \in \mathbf{t}$ and smoothness $m - \mu_k - 1$ at every knot t_k. It is also known that the row vector of the normalized B-splines

$$\Phi_{\mathbf{t},m} := [N_{\mathbf{t},,m,k}]_{k\in\Delta},$$

is a bases of the spline space $S_{\mathbf{t},m}$. Moreover, under normalization

$$\Phi^B_{\mathbf{t},m} := [N^B_{\mathbf{t},m,k}]_{k\in\Delta} = [d^{-1/2}_{\mathbf{t},m,k} N_{\mathbf{t},,m,k}]_{k\in\Delta},$$

this family defines a Riesz basis of the spline space $S_{\mathbf{t},m}$ such that there exist a constant $D_m > 0$ with

$$D_m \left\| \{c_k\}_{k\in\Delta} \right\|_{\ell^2}^2 \leq \left\| \sum_{k\in\Delta} c_k N_{\mathbf{t},m,k}^B \right\|^2 \leq \left\| \{c_k\}_{k\in\Delta} \right\|_{\ell^2}^2 .$$

To make use of the previous characterization of tight frames we rely on the existence of a approximate dual. Therefore we have to take a closer look on the Gramian matrix of the given Riesz basis. It is well known that the Gramian matrix Γ^B of $\Phi_{\mathbf{t},m}^B$ is given by

$$\Gamma^B = \int_\Omega \Phi_{\mathbf{t},m}^B(x)^T \Phi_{\mathbf{t},m}^B(x) dx = \left[(d_{\mathbf{t},m,k} d_{\mathbf{t},m,l})^{-1/2} \langle N_{\mathbf{t},m,k}, N_{\mathbf{t},m,l} \rangle \right]_{k,l\in\Delta} ,$$

and Γ^B is symmetric positive definite banded matrix. Thus we define the dual basis

$$\tilde{\Phi}^B := \Phi^B {\Gamma^B}^{-1}.$$

It is also known that a banded positive semi-definite matrix S_L exists that defines an approximate dual of order $1 \leq L \leq m$ such that S_L has bandwidth at most L. In other words, there is a minimal supported approximate dual, that satisfies also a locality condition like the primal basis.

6.4.2 Construction of tight frames of spline wavelets

The aim of this section is to give only a brief description of the construction of spline wavelet frames. Since we do not need the explicit construction of the wavelets we only rely on the existence of those frames, we will skip all the proofs. Anyway, the main problem is to establish a refinement equation and verify the existence of an NMRA tight spline wavelet frame, but as mentioned before, for a detailed discussion we refer the reader again to [11]. We only give a short conclusion of the most important facts in our application.

Let $\mathbf{t}_1 \subset \mathbf{t}_2$ be knot vectors satisfying (6.4.1)-(6.4.2), where the subset notation is used for ordered sets:

- new knots of multiplicity $\leq m$ can be insert into $\mathbf{t}_1$, or the multiplicity of an existing knot $t_k \in \mathbf{t}_1$ can be increased.

The index set of the corresponding basis $\Phi_{\mathbf{t}_1,m}^B$ and $\Phi_{\mathbf{t}_2,m}^B$ are denoted by Δ_1 and respectively by Δ_2. The normalized basis satisfies the refinement equation

$$\Phi_{\mathbf{t}_1,m} = \Phi_{\mathbf{t}_2,m} M_{\mathbf{t}_1,\mathbf{t}_2,m},$$

where the matrix $M_{\mathbf{t}_1,\mathbf{t}_2,m}$ has non-negative entries, with each row summing to 1 and the matrix is parse in the sense that only those B-splines in $\Phi_{\mathbf{t}_2,m}^B$ appear in the refinement relation (entries in the k-column), whose support is contained in the support of $N^B_{\mathbf{t}_1,m,k}$. Now let $\mathbf{t}_j$, $j \geq 0$ be a nested sequence of knots vectors, such that (6.4.1)-(6.4.2) are satisfied and $\max_{k\in\Delta_j}\left\{t^{(j)}_{k+1} - t^{(j)}_k\right\}$ converge to zero, approximately like $\sim 2^{-j}$, in other words a new knot on level j is placed between to existing knots in the interior on level j. With the idea of quadrature rule in Section 6.3 in mind it is reasonable to assume that $\mathbf{t}_0$ has no interior points, $\mathbf{t}_1$ has one interior point and so one $\mathbf{t}_j$ has $2^j - 1$ interior points. Also as discussed before, let the $\Phi^B_{\mathbf{t}_j,m} = [N^B_{\mathbf{t}_j,m,k}]_{k\in\Delta_j}$ provide the bases of the MRA spline space $V_j \subset \mathrm{L}^2(\Omega)$, S^j_L defines an approximate dual of order $1 \leq L \leq m$ that complies with the requirements in Subsection 6.1.2, see [11, Theorem 6.1], more detailed the quadratic forms

$$T_j f := \left[\langle f, N^B_{\mathbf{t}_j,m,k}\rangle\right]_{k\in\Delta_j} S^j_L \left[\langle f, N^B_{\mathbf{t}_j,m,k}\rangle\right]^T_{k\in\Delta_j}, \quad f \in \mathrm{L}^2(\Omega) \tag{6.4.3}$$

are uniformly bounded on $\mathrm{L}^2(\Omega)$ and

$$\lim_{j\to\infty} T_j f = \|f\|^2 \quad f \in \mathrm{L}^2(\Omega). \tag{6.4.4}$$

The next result is essentially for the characterization of NMRA spline tight frames, see [11, Theorem 6.2].

Theorem 6.4.2. *Under the above assumption there is a factorization*

$$S^{j-1}_L - M_{\mathbf{t}_j,\mathbf{t}_{j+1},m} S^j_L M^T_{\mathbf{t}_j,\mathbf{t}_{j+1},m} = Q_j Q^T_j,$$

where $Q_j \in \mathbb{R}^{|\Delta_{j+1}|\times|\Delta_{j+1}|-L}$. The families $\Psi_j := \Phi^B_{\mathbf{t}_{j+1},m} Q_j$, $j \leq 0$, of cardinality $|\Delta_{j+1}| - L$, constitute a tight frame of $\mathrm{L}^2(\Omega)$ *relative to T_0, such that all wavelets $\phi_{j,k}$ have L vanishing moments.*

With this theoretical result in mind, we are able to give an explicit formulation of one possibility spline approximation and it will became apparent that this one is a quadrature rule.

6.4.3 One possible spline quadrature

Let $\Omega = [a, b]$ be an interval. For the explicit construction of a quadrature rule based on spline spaces let $m \in \mathbb{N}_{\geq 2}$ be the order of the underling normalized B-spline. Consider for level $l = 0$ the quadrature $A^{spline}(0, 1)$ with no interior points, introduced by the knot vector

$$\mathbf{t}_0 = \left\{\underbrace{a, \ldots, a}_{m\text{-times}}, \underbrace{b, \ldots, b}_{m\text{-times}}\right\},$$

with no interior points and multiplicity m on both boundary knots a and b. Here, we make the restriction that we consider only knot vectors with multiplicity $\mu_k = 0$ for the interior knots. Let t_j, $j \geq 1$ be the knot vector with $2^j - 1$ simple interior knots

$$\mathbf{t}_j = \{a, \ldots, a, t_1, \ldots, t_{2^j-1}, b, \ldots, b\}.$$

The sequence $\mathbf{t}_j$ is nested in the sense that a new knot in $\mathbf{t}_{j+1}$ is placed between two adjacent knots in t_j,

$$a = \underbrace{t^{(j)}_{-m+1}}_{=t^{(j-1)}_{-m+1}} = \ldots = \underbrace{t^{(j)}_0}_{=t^{(j-1)}_0} < t^{(j)}_1 < \underbrace{t^{(j)}_2}_{=t^{(j-1)}_1} \ldots$$

$$\underbrace{t^{(j)}_{2^j-2}}_{=t^{(j-1)}_{2^{j-1}-1}} < t^{(j)}_{2^j-1} < \underbrace{t^{(j)}_{2^j}}_{=t^{(j-1)}_{2^{j-1}}} = \ldots = \underbrace{t^{(j)}_{2^j-1+m}}_{=t^{(j-1)}_{2^{j-1}-1+m}} = b.$$

Consequently we get a family of nested subspaces

$$V_j = \operatorname{span}\Phi^B_{\mathbf{t}_j,m},$$

that defines a MRA tight frame of spline wavelets. By the Definition 6.3.1 we get the approximation (quadrature)

$$\begin{aligned}
A^{spline}(l,1)v &= \langle v, \tilde{\Phi}_{\mathbf{t}_l,m}\rangle \left[\int_\Omega N^B_{\mathbf{t}_l,m,k}dx\right]^T_{k\in\Delta_l} \\
&= \langle v, \tilde{\Phi}_{\mathbf{t}_l,m}\rangle \left[d^{1/2}_{\mathbf{t}_l,m,k}\right]^T_{k\in\Delta_l} \\
&= \langle v, \tilde{\Phi}_{\mathbf{t}_l,m}\rangle \left[\left(\frac{t_{k+m}-t_k}{m}\right)^{1/2}\right]^T_{k\in\Delta_l},
\end{aligned}$$

where $\langle v, \tilde{\Phi}_{\mathbf{t}_l,m}\rangle$ has to be understood as a row vector. The corresponding approximation is known as a so called interpolatory quadrature rule and in the following we have to choose for a given knot vector an adapted set of sample points. Let us choose for a given knot vector $\mathbf{t}_l$ a strictly increasing sequence of sample points $\tau_l = [\tau^l_k]_{k\in\Delta_l}$ that assure the interpolation problem: to find for a given function f a spline $\sum_{k\in\Delta_l} c_k N^B_{\mathbf{t}_l,m,k}$ that agrees with f at τ_l has unique solution. This interpolation problem has exactly one solution if and only if the linear system of equations

$$\sum_{k\in\Delta_l} c_k N^B_{\mathbf{t}_l,m,k}(\tau^l_i) = f(\tau^l_i), \quad i \in \Delta_l,$$

has a unique solution. It is well known, see, e.g. [23, 24, 62] that the coefficient matrix $[N^B_{\mathbf{t}_l,m,k}(\tau^l_i)]_{k,i\in\Delta_l}$ is invertible if and only if

$$N^B_{\mathbf{t}_l,m,k}(\tau^l_k) \neq 0, \quad k \in \Delta_l,$$

i.e., if and only if $t_k^l < \tau_k^l < t_{k+m}^l$ for all $k \in \Delta_l$. The last condition that characterize the interplay between the sample points (sometimes called: interpolation points or data point) and knots is also known under the Schoenberg-Whitney condition or Schoenberg-Whitney theorem. By the fact that we have a unique decomposition for $v \in V_l$ it is obviously that for a adapted sequence of sample points τ_l

$$\langle v, \tilde{\Phi}_{\mathbf{t}_l,m} \rangle = [c_k]_{k\in\Delta_l},$$

where the $[c_k]_{k\in\Delta_l}$ is the solution of the linear system of equations

$$\sum_{k\in\Delta_l} c_k N^B_{\mathbf{t}_l,m,k}(\tau_i^l) = v(\tau_i^l), \quad i \in \Delta_l. \tag{6.4.5}$$

Remark 6.4.3. The so-called collocation matrix $[N^B_{\mathbf{t}_l,m,k}(\tau_i^l)]_{k,i\in\Delta_l}$ is known to has maximal bandwidth m. A second important property is that the matrix is total positive, see, e.g. [23, 41]. This is also important for the cost analysis of the algorithms. A practical way to solve the interpolation problem is to solve the linear system of equations (6.4.5). This could be done by Gauss elimination without pivoting, and observe the decomposition is again banded with maximal bandwidth $2m - 1$.

Now we are able to rephrase the error bound in Theorem 6.3.4 in terms of interpolatory B-splines quadrature.

Corollary 6.4.4. *For $s > 1/2$ we get*

$$\mathrm{err}(H_s, A^{spline}(l,1)) \lesssim 2^{-ls},$$

for the spline quadrature defined above and $\left|A^{spline}(l,1)\right| = \mathcal{O}(2^l)$.

Remark 6.4.5. The first observation is that there is some freedom to choose the set of sample points that defines the interpolatary B-spline quadrature, e.g. randomized points or so-called Chebyshev-Demko points. But note, that the Schoenberg-Whitney theorem only ensure the existence of a unique solution of the interpolation problem. For the question of numerical stability we refer the reader to [19, 57, 59, 60]. The second possibility to define a spline quadrature in our context is to compute $\langle v, \tilde{\Phi}_{\mathbf{t}_l,m} \rangle$ via the inverse Gramian matrix and a Gauss quadrature rule with piecewise B-splines as weight functions see, e.g. [4, 5].

6.5 Multivariate numerical integration

In this section we will extend the concept of NMRA tight frames to higher dimensions. We follow the approach which was also used in the previous chapter. A multivariate NMRA tight frame of $\mathrm{L}^2([0,1]^d)$ is given by so-called tensor product wavelets. Under similar requirements we have posed in Section 6.3 we define the approximation space on level L by

$$V^{d,L} := \sum_{|\mathbf{j}|=L} \bigotimes_{i=1}^{d} V_{j_i} = \sum_{|\mathbf{j}|\leq L} \bigotimes_{i=1}^{d} W_{j_i-1}. \tag{6.5.1}$$

Similar to the one-dimensional case we put

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

Since $V = V^1$ is dense in $\mathrm{L}^2([0,1])$, the space V^d is dense in $\mathrm{L}^2([0,1]^d)$. Thus, we obtain with use of the shorthand $\Psi_{-1} = \Phi_0 S_0^{1/2}$ the following expansion for $f \in \mathrm{L}^2([0,1]^d)$

$$f = \frac{1}{A^d} \left(\sum_{\mathbf{j}\geq -1} \sum_{\mathbf{k}\in\nabla_{\mathbf{j}}} \langle f, \Psi_{\mathbf{j},\mathbf{k}} \rangle \Psi_{\mathbf{j},\mathbf{k}} \right),$$

where $\mathbf{j} = (j_1, \ldots, j_d) \geq -1$ is meant in the way $j_u \geq -1$ for all $u = 1, \ldots, d$. Furthermore, we use the shorthands $\nabla_{\mathbf{j}} = \nabla_{j_1} \times \ldots \times \nabla_{j_d}$ and

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

The function spaces we are interested in are more or less simple tensor products of the spaces considered in the previous sections. For a better overview let us make the assumption $\Omega = [0,1)$.

6.5.1 The d-dimensional cubature method

Now, we extend our one-dimensional algorithm $A(l,1)$ to a d-dimensional cubature. This should be done via an anisotropic version of Smolyak's construction, sometimes this is called adaptivity in the sense of a-priori knowledge, see, e.g. [34, 53]. The meantime multiple mentioned difference quadrature of level $l \geq 0$ is defined similar to the previous case by

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

with $A(-1,1) := 0$. We define an anisotropic version of Smolyak's construction of level L by

$$A_\kappa(L,d) \;:=\; \sum_{\mathbf{l}\in\mathbb{N}_0^d,\ |\mathbf{l}|_\kappa \leq L} (\Delta^{l_1} \otimes \Delta^{l_2} \otimes \cdots \otimes \Delta^{l_d}), \tag{6.5.2}$$

where $|\mathbf{l}|_\kappa = \sum_{i=1}^d l_i \kappa_i$, $\kappa_i > 0$. Notice for $\kappa = 1$ we have the well known Smolyak construction. The parameter κ in (6.5.2) can be understood as a weighting of the directions of the cubature. If there is a-priori knowledge of the dependence of smoothness and directions then the parameter κ can be used to reduce the overall cost of the algorithm. The main idea is to apply less costly quadrature formulas in smoother directions of the integrand to get a accurate asymptotic behavior, see, e.g. [34, 35]. Let us recall that in the one-dimensional case the approximation $A(l,1)$ is exact on V^l. In the d-dimensional case it is not that difficult to show the exactness of $A_1(L,d)$ on $V^{d,L}$.

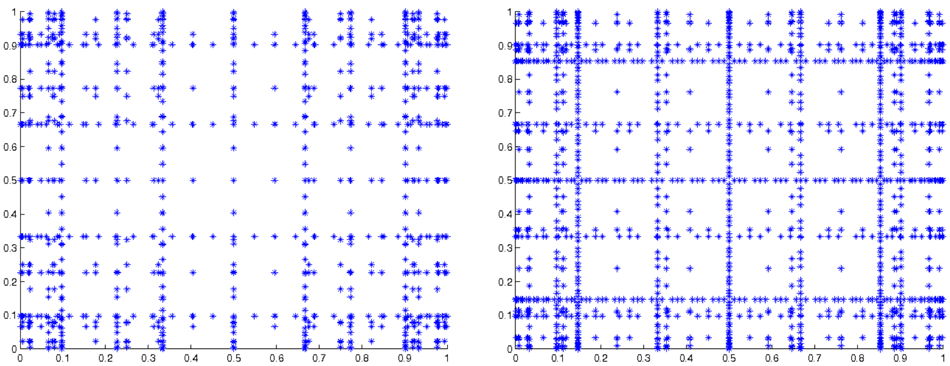

Figure 6.1: Smolyak's construction in dimension two. Both cubatures with non-equidistant knot vectors and corresponding sample points. In the left diagram the order of the underlying spline is four and the level is five. In the right diagram the cubature is on level 6 and the order of the underlying spline is three.

Theorem 6.5.1. *The approximation (cubature) $A_1(L,d)$ is exact on the approximation space $V^{d,L}$.*

The proof also follows the lines of the proof of [48, Theorem 2] and proceed via induction over the dimension.

6.5.2 Error bounds for the cubature error

For error analysis we consider product spaces which are based on the one dimensional function spaces H_s used for our one-dimensional approximation error bounds. These seem to be the natural spaces for the variation of Smolyak's construction. According to the Sobolev spaces with dominating mixed derivatives, defined by the tensor product

$$H^{\mathbf{s}}_{\text{mix}} = H^{s_1} \otimes H^{s_2} \otimes \ldots \otimes H^{s_d}, \tag{6.5.3}$$

we consider the spaces

$$H_{\mathbf{s}} = H_{s_1} \otimes H_{s_2} \otimes \ldots \otimes H_{s_d},$$

where $s_i > 1/2$ and $\|f_1 \otimes \ldots \otimes f_1\|_{\mathbf{s}} = \|f_1\|_{s_1} \cdots \|f_d\|_{s_d}$. Obviously, we also get

$$H_{\mathbf{s}} = \left\{ f \in \mathrm{L}^2 : \sum_{\mathbf{j} \geq -1} \sum_{\mathbf{k} \in \nabla_{\mathbf{j}}} 2^{|\mathbf{j}|_\kappa 2s} \left| \langle f, \widetilde{\Psi}_{\mathbf{j},\mathbf{k}} \rangle \right|^2 < \infty \right\},$$

where $\kappa_i = s_i/s$, and s can be chosen for example as $\min_{i=1}^d s_i$. Under this assumption we get for arbitrary linear operator $T : H_s \to \mathbb{R}$ that the induced operator norm is given by

$$\|T\|_{op} = \|T_1 \otimes T_2 \otimes \ldots \otimes T_d\|_{op} = \prod_{i=1}^{d} \|T_i\|_{op},$$

where

$$\|T_i\|_{op} = \sup_{f_i \in H_{s_i},\, \|f_i\|_{s_i}=1} |T_i(f_i)|.$$

Before we verify our main result, we want to calculate the induced operator norm of the functional $I : H_{s_i} \to \mathbb{R}$. For an arbitrary $f \in H_{s_i}$ we get

$$\begin{aligned}
I(f) &= I\left(\langle f, \Phi_0 S^{1/2} \rangle \Phi_0 S^{1/2} + \sum_{j \geq 0} \sum_{k \in \nabla_j} \langle f, \tilde{\psi}_{j,k} \rangle \psi_{j,k} \right) \\
&= \langle f, \Phi_0 S^{1/2} \rangle I(\Phi_0 S^{1/2}) + \sum_{j \geq 0} \sum_{k \in \nabla_j} \langle f, \tilde{\psi}_{j,k} \rangle I(\psi_{j,k}) \\
&\leq \left(T_0 f + \sum_{j \geq 0} \sum_{k \in \nabla_j} 2^{j 2 s_i} \langle f, \tilde{\psi}_{j,k} \rangle^2 \right)^{1/2} \left(\sum_{k \in \Delta_0} I\left(\phi_{0,k}^{S^{1/2}} \right)^2 + 0 \right)^{1/2} \\
&= \|f\|_{s_i} \left(I[\Phi_0 S^{1/2}] I[\Phi_0 S^{1/2}]^T \right)^{1/2} \\
&= \|f\|_{s_i} \left(\langle \mathbf{1}_{[0,1)}, \Phi_0 S^{1/2} \rangle \langle \mathbf{1}_{[0,1)}, \Phi_0 S^{1/2} \rangle^T \right)^{1/2} \\
&= \|f\|_{s_i} \left(\langle \mathbf{1}_{[0,1)}, \Phi_0 \rangle S \langle \mathbf{1}_{[0,1)}, \Phi_0 \rangle^T \right)^{1/2} \leq \|f\|_{s_i}.
\end{aligned}$$

Consequently, we obtain for the induced operator norm

$$\|I\|_{\mathrm{op}} = \sup_{f \in H_{s_i},\, \|f\|_{s_i}=1} |I(f)| \leq 1.$$

So, we get for $f^* = \mathbf{1}_{[0,1)}$

$$\|f^*\|_{s_i} = \left(T_0 f^* + \sum_{j \geq 0} \sum_{k \in \nabla_j} 2^{j2s} \langle \mathbf{1}_{[0,1)}, \tilde{\psi}_{j,k} \rangle^2 \right)^{1/2} = T_0 \mathbf{1}_{[0,1)} \leq 1.$$

Thus, the induced operator norm satisfies

$$\|I\|_{\text{op}} \geq I(f^*)/\|f^*\| \geq I(f^*) = 1,$$

which implies $\|I\|_{\text{op}} = 1$.

Theorem 6.5.2. *Let $A(l,1)$ be the approximation appropriate to a given MRA tight frame with given vanishing moments. For $\mathbf{s} > 1/2$ the worst case error of $A_\kappa(L,d)$, where $\kappa_i = s_i / min_{i=1}^d s_i$ satisfies*

$$\operatorname{err}(H_{\mathbf{s}}, A_\kappa(L,d)) = \mathcal{O}\left(\prod_{i=1}^{d-1} C(1+2^{s_u}) \cdot 2^{-Ls} \frac{(L + \sum_{u=1}^{d-1} \kappa_u)^{d-1}}{(d-1)! \prod_{u=1}^{d-1} \kappa_u}\right).$$

To prove our main result we adapt the proof of [67, Lemma 2].

Proof. The proof is via induction and based on the observation

$$\begin{aligned}
A_\kappa(L,d) &= \sum_{\tilde{l} \in \mathbb{N}_0^{d-1},\, \sum_{i=1}^{d-1} \tilde{l}_i \kappa_i \leq L} \; \sum_{l_d \kappa_d = 0}^{L - \sum_{i=1}^{d-1} \tilde{l}_i \kappa_i} \Delta^{\tilde{l}_1} \otimes \Delta^{\tilde{l}_2} \otimes \ldots \otimes \Delta^{\tilde{l}_{d-1}} \otimes \Delta^{l_d} \\
&= \sum_{\tilde{\mathbf{l}} \in \mathbb{N}_0^{d-1},\, |\tilde{\mathbf{l}}|_\kappa \leq L} \Delta^{\tilde{l}_1} \otimes \Delta^{\tilde{l}_2} \otimes \ldots \otimes \Delta^{\tilde{l}_{d-1}} \otimes \sum_{l_d \kappa_d = 0}^{L - |\tilde{\mathbf{l}}|_\kappa} \Delta^{l_d} \\
&= \sum_{\tilde{\mathbf{l}} \in \mathbb{N}_0^{d-1},\, |\tilde{\mathbf{l}}|_\kappa \leq L} \Delta^{\tilde{l}_1} \otimes \Delta^{\tilde{l}_2} \otimes \ldots \otimes \Delta^{\tilde{l}_{d-1}} \otimes A\left((L - \left|\tilde{\mathbf{l}}\right|_\kappa)/\kappa_d, 1\right).
\end{aligned}$$

Thus we get

$$\begin{aligned}
I_d - A_\kappa(L,d) &= \sum_{\tilde{\mathbf{l}} \in \mathbb{N}_0^{d-1},\, |\tilde{\mathbf{l}}|\kappa \leq L} \left(\bigotimes_{u=1}^{d-1} \Delta^{\tilde{l}_u}\right) \otimes \left(I_1 - A\left((L - \left|\tilde{\mathbf{l}}\right|_\kappa)/\kappa_d, 1\right)\right) \\
&\quad + (I_{d-1} - A_\kappa(L, d-1)) \otimes I_1.
\end{aligned}$$

Consequently

$$\begin{aligned}
\|I_d - A_\kappa(L,d)\|_{\text{op}} &\leq \sum_{\tilde{\mathbf{l}} \in \mathbb{N}_0^{d-1},\, |\tilde{\mathbf{l}}|_\kappa \leq L} \prod_{u=1}^{d-1} \left\|\Delta^{\tilde{l}_u}\right\|_{\text{op}} \left\|I_1 - A\left((L - \left|\tilde{\mathbf{l}}\right|_\kappa)/\kappa_d, 1\right)\right\|_{\text{op}} \\
&\quad + \|I_{d-1} - A_\kappa(L, d-1)\|_{\text{op}} \|I_1\|_{\text{op}}.
\end{aligned}$$

The next step is to consider the induced operator norm of the difference quadratures in different directions. According to Theorem 6.3.4 we have

$$\left\|\Delta^{\tilde{l}_u}\right\|_{\text{op}} \leq C\left(2^{-\tilde{l}_u s_u} + 2^{-(\tilde{l}_u - 1)s_u}\right) = C\left(1 + 2^{s_u}\right) 2^{-\tilde{l}_u s_u},$$

which leads to

$$\sum_{\tilde{\mathbf{l}}\in\mathbb{N}_0^{d-1},\,\left|\tilde{\mathbf{l}}\right|_\kappa\le L}\prod_{u=1}^{d-1}\left\|\Delta^{\tilde{l}_u}\right\|_{\mathrm{op}}\left\|I_1-A\left((L-\left|\tilde{\mathbf{l}}\right|_\kappa)/\kappa_d,1\right)\right\|_{\mathrm{op}}$$

$$\le\sum_{\tilde{\mathbf{l}}\in\mathbb{N}_0^{d-1},\,\left|\tilde{\mathbf{l}}\right|_\kappa\le L}\prod_{u=1}^{d-1}C(1+2^{s_u})2^{-\tilde{l}_u s_u}\cdot C2^{-\frac{L-|\tilde{l}|_\kappa}{\kappa_d}\cdot s_d}$$

$$=\sum_{\tilde{\mathbf{l}}\in\mathbb{N}_0^{d-1},\,\left|\tilde{\mathbf{l}}\right|_\kappa\le L}\prod_{u=1}^{d-1}C(1+2^{s_u})2^{-\tilde{l}_u\kappa_u s}\cdot C2^{-Ls+\left|\tilde{l}\right|_\kappa\cdot s}$$

$$=\prod_{u=1}^{d-1}C^2(1+2^{s_u})2^{-Ls}\sum_{\tilde{\mathbf{l}}\in\mathbb{N}_0^{d-1},\,\left|\tilde{\mathbf{l}}\right|_\kappa\le L}1.$$

Therefore, we get for the integration error

$$\begin{aligned}\|I_d-A_\kappa(L,d)\|_{\mathrm{op}}\le\quad&\prod_{u=1}^{d-1}C^2(1+2^{s_u})2^{-Ls}\sum_{\tilde{\mathbf{l}}\in\mathbb{N}_0^{d-1},\,\left|\tilde{\mathbf{l}}\right|_\kappa\le L}1\\&+\quad\|I_{d-1}-A_\kappa(L,d-1)\|_{\mathrm{op}}.\end{aligned}$$

Since by Lemma 4.3.1 the cardinality of the index set used for the cubature rule can be upper bounded by

$$\begin{aligned}\sum_{\tilde{\mathbf{l}}\in\mathbb{N}_0^{d-1},\,\left|\tilde{\mathbf{l}}\right|_\kappa\le L}1\quad&=\quad\left\{z\in\mathbb{N}^{d-1}\ :\ \langle z,\kappa\rangle\le L\right\}\\&=\quad\mathrm{vol}(\mathrm{conv}\left\{L/\kappa_1\cdot e_1,\ldots L/\kappa_{d-1}\cdot e_{d-1}\right\})\\&\le\quad\frac{(L+\sum_{u=1}^{d-1}\kappa_u)^{d-1}}{(d-1)!\prod_{u=1}^{d-1}\kappa_u},\end{aligned}$$

we get inductively

$$\|I_d-A_m(L,d)\|_{\mathrm{op}}\quad\le\quad C2^{-Ls}\sum_{\nu=0}^{d-1}\prod_{u=1}^{\nu}C(1+2^s_\nu)\frac{(L+\sum_{u=1}^{\nu}\kappa_u)^\nu}{\nu!\prod_{u=1}^{\nu}\kappa_u}.$$

□

Quite similar to the lower bound we have discussed in the previous chapter we can prove a lower bound for the spaces H_s considered in this one. As indicated by the one-dimensional lower bound the proof will be more technical as in the multiwavelet case. But the main idea is still the same. Observe the constant that depends one the dimension is more unfavorable since the supports of the wavelets have moderate overlap. For the lower bound we also consider the spaces

$$H_{\mathbf{s}} = H_{s_1} \otimes H_{s_2} \otimes \ldots \otimes H_{s_d},$$

where $s_i > 1/2$ and $\|f_1 \otimes \ldots \otimes f_1\|_{\mathbf{s}} = \|f_1\|_{s_1} \cdots \|f_d\|_{s_d}$. Recall that we also get

$$H_{\mathbf{s}} = \left\{ f \in \mathrm{L}^2 : \sum_{\mathbf{j} \geq -1} \sum_{\mathbf{k} \in \nabla_{\mathbf{j}}} 2^{|\mathbf{j}|_\kappa 2s} \left| \langle f, \Psi_{\mathbf{j},\mathbf{k}} \rangle \right|^2 < \infty \right\},$$

for suitable κ. It is unsurprising that the lower bound gets smaller for spaces which have different smoothness parameters in different directions. This fact we formulate in the next theorem.

Theorem 6.5.3. *Let $\mathbf{s} > 1/2$. There exists a constant $C > 0$ such that for any d-dimensional cubature rule Q_N using N sample points we have*

$$\mathrm{err}(H_{\mathbf{s}}, Q_N) \geq C 2^{-Ls} \left(\frac{1}{(d-1)! \prod_{u=1}^{d-1} \kappa_u} \right)^{1/2} \frac{(L+1)^{d-1}}{(L + \sum_{u=1}^{d-1} \kappa_u)^{\frac{d-1}{2}}}.$$

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

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

where $I_{\mathbf{k}}^{\mathbf{l}} := \operatorname{supp} \Psi_{\mathbf{l},\mathbf{k}}$. Now, similar to the one-dimensional case we choose the integer L that satisfies

$$2^{L-1} \leq (K+1)N \leq 2^L,$$

where K denotes the number of maximal overlap of the supports of $\Psi_{\mathbf{l},\mathbf{k}}$, $\mathbf{k} \in \nabla_{\mathbf{l}}$. We define a function

$$f_L := \sum_{|\mathbf{l}|_\kappa = L} f_{\mathbf{l}}.$$

Without loss of generality we restrict the estimate of $\|f_L\|_{\mathbf{s}}$ to the semi-norm $|f_L|_{\mathbf{s}}$ to verify this, see the proof of the lower bound in Chapter 5. We get for the semi-norm of our candidate

$$\begin{aligned} |f_L|_{\mathbf{s}}^2 &= \sum_{\mathbf{j} \geq 0} \sum_{\mathbf{k} \in \nabla_{\mathbf{j}}} 2^{|\mathbf{j}|_\kappa 2s} \langle f_L, \widetilde{\Psi}_{\mathbf{j},\mathbf{k}} \rangle^2 \\ &= A^{-2} \sum_{\mathbf{j} \geq 0} \sum_{\mathbf{k} \in \nabla_{\mathbf{j}}} 2^{|\mathbf{j}|_\kappa 2s} \langle f_L, \Psi_{\mathbf{j},\mathbf{k}} \rangle^2 \\ &= A^{-2} \sum_{|\mathbf{l}|_\kappa = |\mathbf{l}'|_\kappa = L} \sum_{\mathbf{j} \geq 0} \sum_{\mathbf{k} \in \nabla_{\mathbf{j}}} 2^{|\mathbf{j}|_\kappa 2s} \langle f_{\mathbf{l}}, \Psi_{\mathbf{j},\mathbf{k}} \rangle \langle f_{\mathbf{l}'}, \Psi_{\mathbf{j},\mathbf{k}} \rangle. \end{aligned}$$

Now similar to the multiwavelet case the inner product $\langle f_{\mathbf{l}}, \Psi_{\mathbf{j},\mathbf{k}} \rangle$ respectively $\langle f_{\mathbf{l}'}, \Psi_{\mathbf{j},\mathbf{k}} \rangle$ vanish if one of the indices satisfies $j_\nu \geq l_\nu \geq 0$ or $j_\nu \geq l'_\nu \geq 0$. Furthermore we have

$$|\langle f_{\mathbf{l}}, \Psi_{\mathbf{j},\mathbf{k}} \rangle| \leq \|\Psi_{\mathbf{j},\mathbf{k}}\|_\infty \|f_{\mathbf{l}}\| \,\mathrm{vol}(\mathrm{supp}\ \Psi_{\mathbf{j},\mathbf{k}}) \lesssim 2^{-|\mathbf{j}/2|}.$$

Therefore, we get

$$\begin{aligned}
|f_L|_{\mathbf{s}}^2 &\lesssim \sum_{|\mathbf{l}|_\kappa = |\mathbf{l}'|_\kappa = L} \sum_{\mathbf{j} \geq 0} \sum_{\mathbf{k} \in \nabla_{\mathbf{j}}} 2^{|\mathbf{j}|_\kappa 2s} 2^{-|\mathbf{j}|} \\
&\lesssim \sum_{|\mathbf{l}|_\kappa = |\mathbf{l}'|_\kappa = L} \sum_{\mathbf{j} \geq 0} 2^{|\mathbf{j}|_\kappa 2s} \\
&\lesssim \sum_{\nu=0}^{L} \sum_{|\mathbf{j}|_\kappa = \nu} 2^{\nu 2s} \left(\sum_{|\mathbf{l}|_\kappa = L,\, \mathbf{l} > \mathbf{j}} 1 \right)^2 \\
&\lesssim \sum_{\nu=0}^{L} \sum_{|\mathbf{j}|_\kappa = \nu} 2^{\nu 2s} \left(\sum_{|\mathbf{l}|_\kappa = L - \nu} 1 \right)^2 .
\end{aligned}$$

Hence, we get by Lemma 4.3.1

$$|f_L|_{\mathbf{s}}^2 \lesssim \sum_{\nu=0}^{L} \frac{(\nu + \sum_{u=1}^{d-1} \kappa_u)^{d-1}}{(d-1)! \prod_{u=1}^{d-1} \kappa_u} 2^{\nu 2s} \left(\frac{(L - \nu + \sum_{u=1}^{d-1} \kappa_u)^{d-1}}{(d-1)! \prod_{u=1}^{d-1} \kappa_u} \right)^2 .$$

Summing over the new index $m := L - \nu$ yields

$$\begin{aligned}
|f_L|_{\mathbf{s}}^2 &\lesssim \sum_{m=0}^{L} \frac{(L - m + \sum_{u=1}^{d-1} \kappa_u)^{d-1}}{(d-1)! \prod_{u=1}^{d-1} \kappa_u} 2^{(L-m)2s} \left(\frac{(m + \sum_{u=1}^{d-1} \kappa_u)^{d-1}}{(d-1)! \prod_{u=1}^{d-1} \kappa_u} \right)^2 \\
&\leq \sum_{m=0}^{\infty} 2^{-m2s} \left(\frac{(m + \sum_{u=1}^{d-1} \kappa_u)^{d-1}}{(d-1)! \prod_{u=1}^{d-1} \kappa_u} \right) 2^{L2s} \frac{(L + \sum_{u=1}^{d-1} \kappa_u)^{d-1}}{(d-1)! \prod_{u=1}^{d-1} \kappa_u} \\
&\lesssim 2^{L2s} \frac{(L + \sum_{u=1}^{d-1} \kappa_u)^{d-1}}{(d-1)! \prod_{u=1}^{d-1} \kappa_u}.
\end{aligned}$$

Furthermore, we have quite similar to the one-dimensional case

$$\begin{aligned}
\int_\Omega f_L dx &= \sum_{|\mathbf{l}|_\kappa = L} \int_\Omega f_{\mathbf{l}} dx \geq \sum_{|\mathbf{l}|_\kappa = L} \min_{\mathbf{k} \in \nabla_{\mathbf{l}}} \mathrm{vol}(\mathbf{I}_{\mathbf{k}}^{\mathbf{l}})(|\nabla_{\mathbf{l}}| - KN) \\
&\geq \frac{C}{K+1} \sum_{|\mathbf{l}|_\kappa = L} \geq C' \frac{L^{d-1}}{(d-1)! \prod_{u=1}^{d-1} \kappa_u}.
\end{aligned}$$

Consequently, for the new function $f := f_L / \|f_L\|_{\mathbf{s}}$ the estimates above result in

$$\mathrm{err}(f, Q_N) \geq C'' \cdot 2^{-Ls} \left(\frac{1}{(d-1)! \prod_{u=1}^{d-1} \kappa_u} \right)^{1/2} \frac{(L+1)^{d-1}}{(L + \sum_{u=1}^{d-1} \kappa_u)^{\frac{d-1}{2}}}.$$

$\square$

Remark 6.5.4. Let us recall the essence of Corollary 6.2.4. We require that there exist approximate duals of order n and that the smoothness parameter $\mathbf{s}$ satisfies both, $\mathbf{s} < n$ and $s_i < \sup \{s : \phi_{0,k} \in H^s, \ k \in \Delta_0\}$ for any direction $i = 1, \ldots, d$. Thus, we get norm equivalence and consequently $H^{\mathbf{s}}_{\mathrm{mix}} \sim H_{\mathbf{s}}$. In this situation Theorem 6.5.2 as well as the lower bound in Theorem 6.5.3 holds in particular for mixed Sobolev spaces $H^{\mathbf{s}}_{\mathrm{mix}}$ in place of $H_{\mathbf{s}}$. Note that this is the crucial refinement compared to the multiwavelet case we have considered in Chapter 5.

6.6 Numerical examples

For the sake of completeness we implemented a special case of our general cubature method based on spline quadratures and computed the integrals of certain test functions in dimension 10. The families of test functions and the considered level of difficulty were the same as in Section 5.4. This enables us to compare our spline cubature directly to the results of Chapter 5. Our one-dimensional quadrature is based on B-splines defined by non-equidistant knot vectors. The set of sample points were essentially chosen to be the extrema of the corresponding Chebyshev spline, and computed via the well known Remes algorithm. We extended our one-dimensional quadrature via a common version of Smolyak's construction to a d-dimensional cubature. The diagrams in Figure 6.2 to 6.7 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 and additionally the two families 2 and 5 in dimension 5, Figure 6.10, 6.11. The last function we considered is a function whose variation grows exponentially, see Figure 6.8, 6.9. This type of function is often mentioned for comparing Quasi-Monte-Carlo or Sparse Grids methods, see, e.g. [34, 46]. It is defined by

$$f(x) = (1 + 1/d)^d \prod_{i=1}^{d} (x_i)^{1/d}.$$

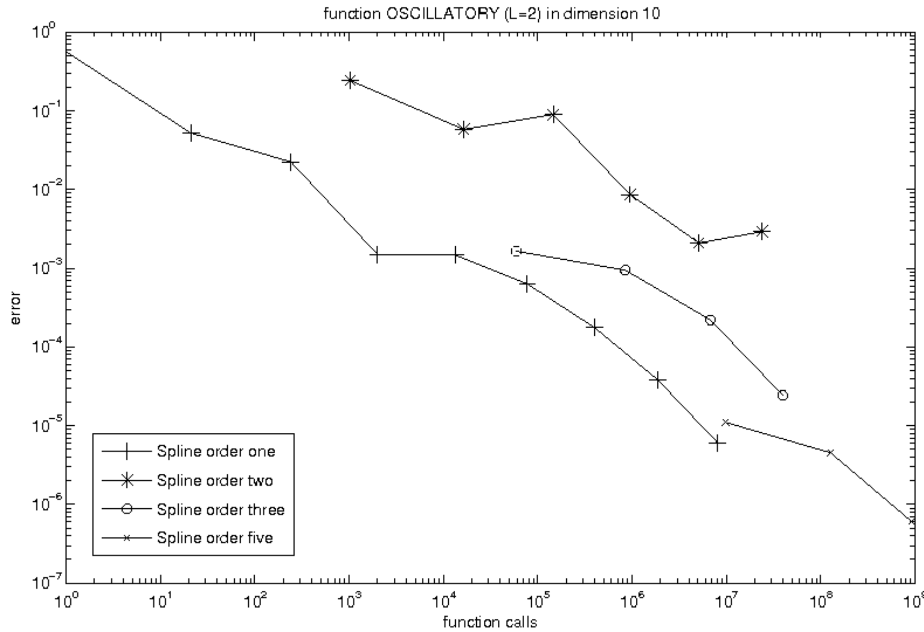

Figure 6.2: Median of absolute error of family (1), 20 integrands.

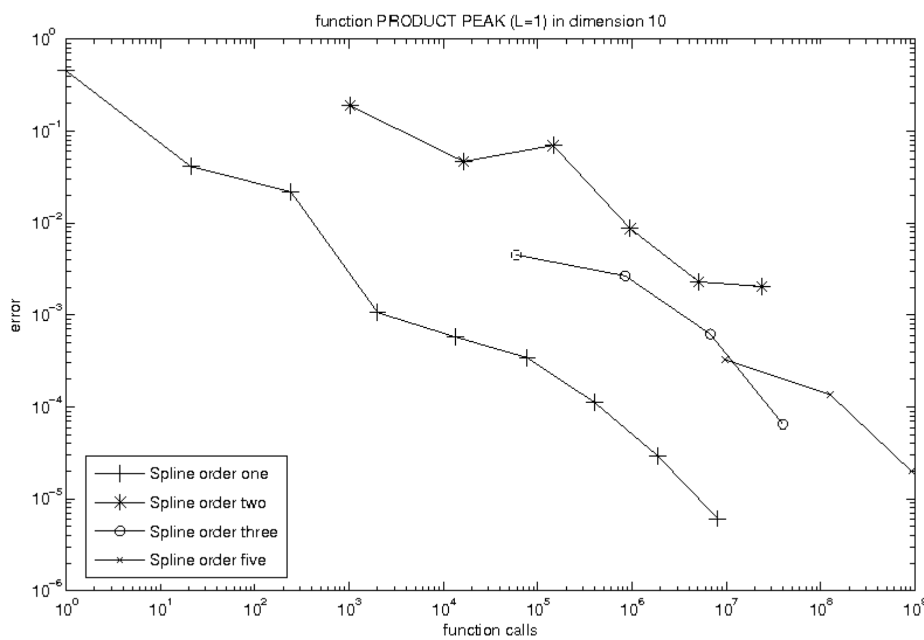

Figure 6.3: Median of absolute error of family (2), 20 integrands.

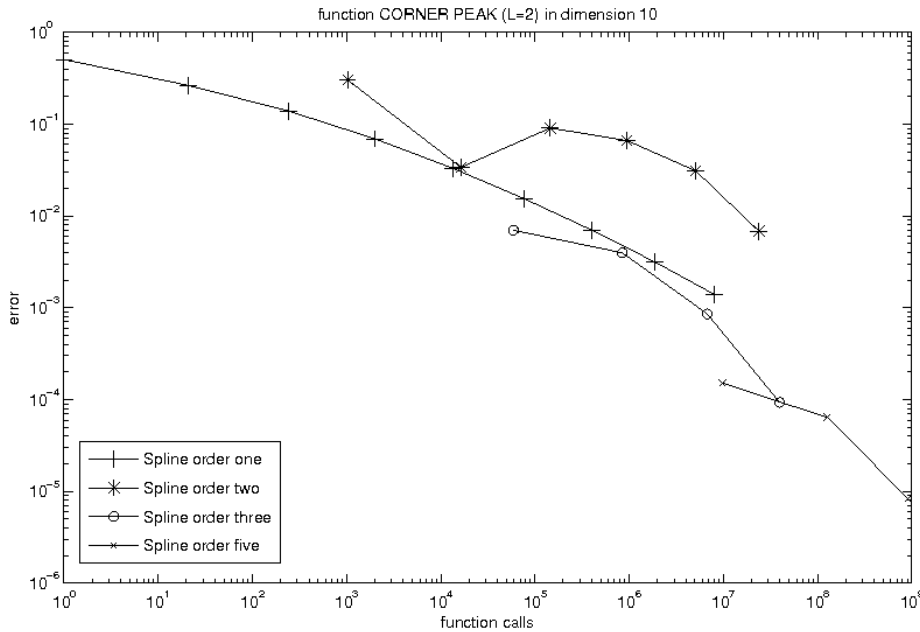

Figure 6.4: Median of absolute error of family (3), 20 integrands.

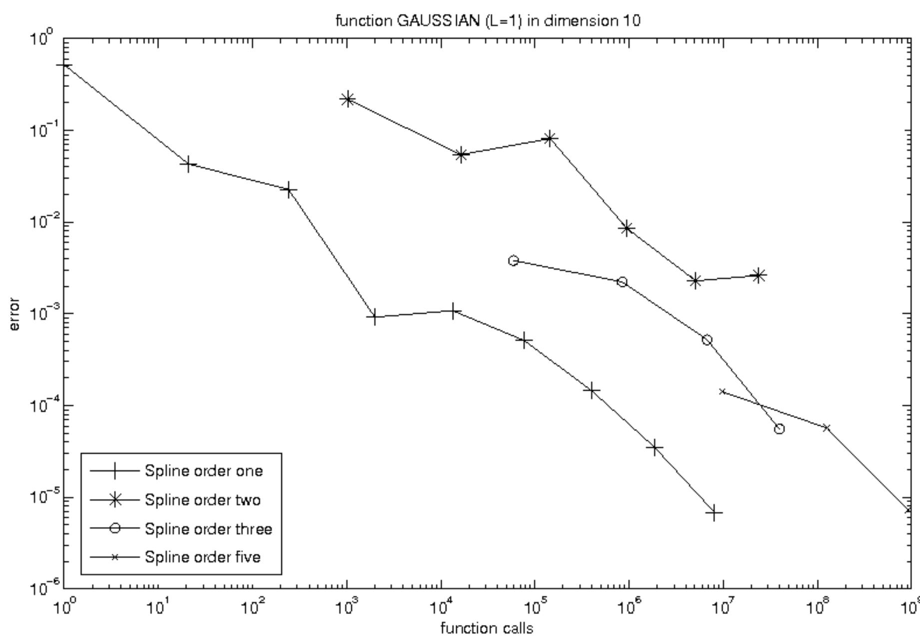

Figure 6.5: Median of absolute error of family (4), 20 integrands.

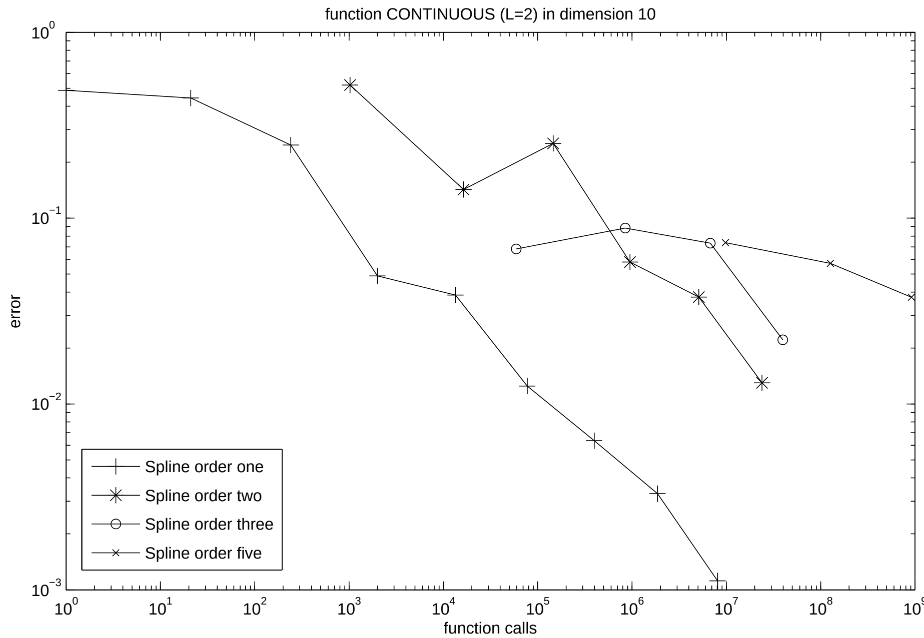

Figure 6.6: Median of absolute error of family (5), 20 integrands.

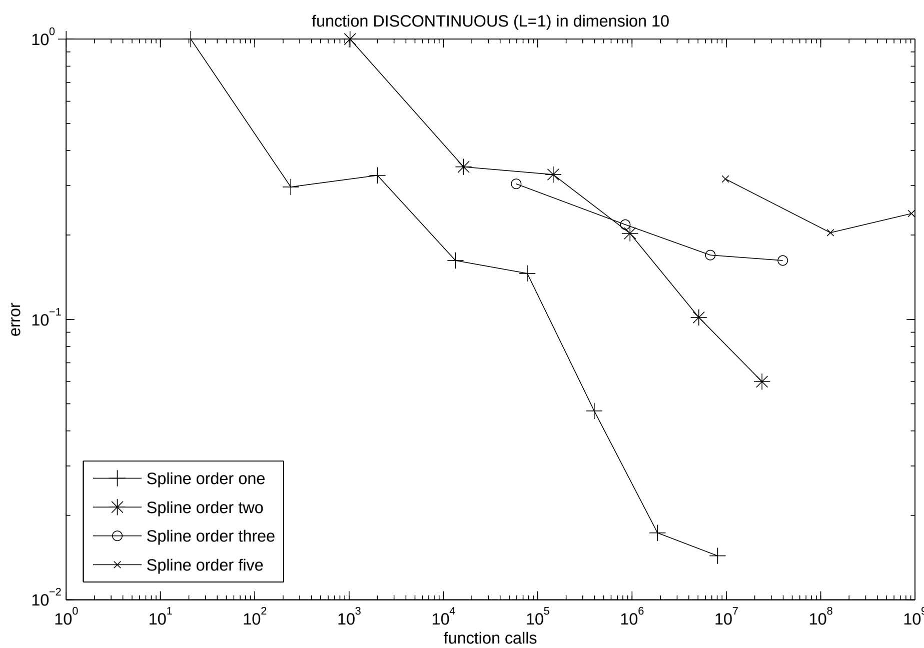

Figure 6.7: Median of absolute error of family (6), 20 integrands.

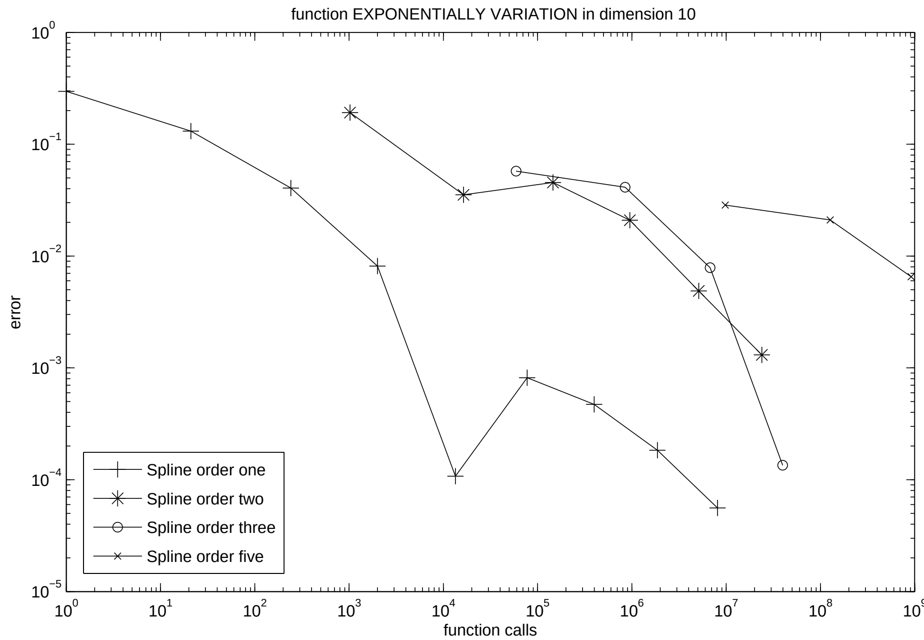

Figure 6.8: Funktion Exponentially in dimension 10.

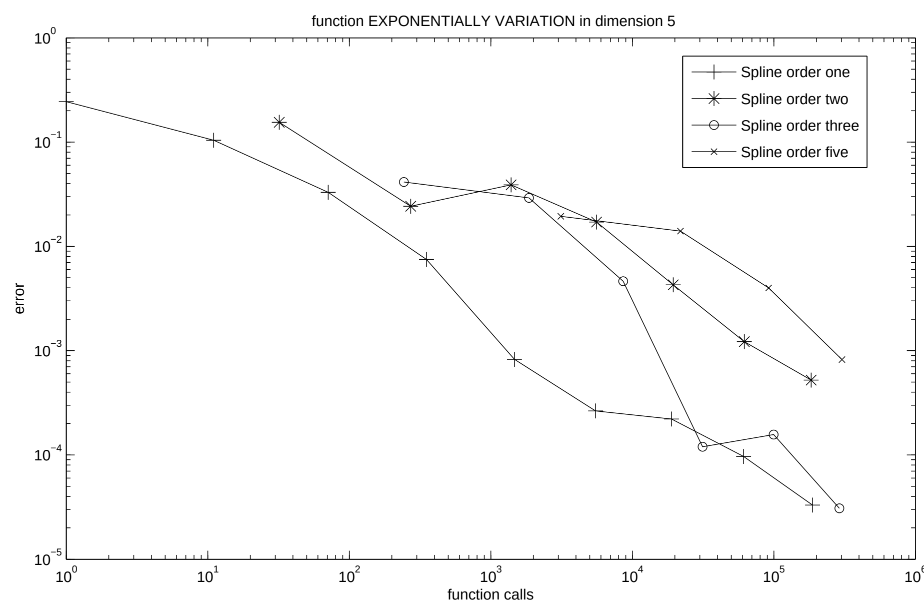

Figure 6.9: Funktion Exponentially in dimension 5.

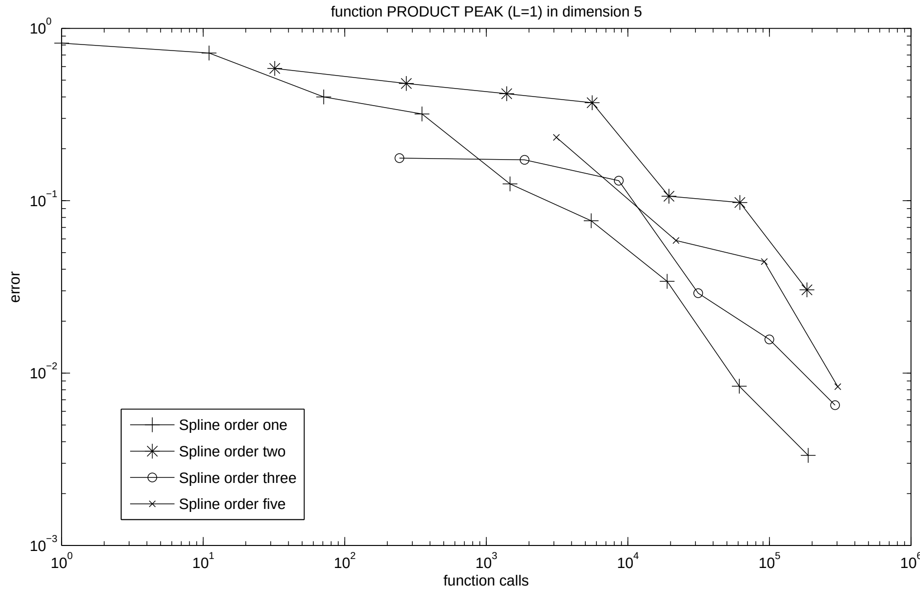

Figure 6.10: Median of absolute error of family (2), 20 integrands.

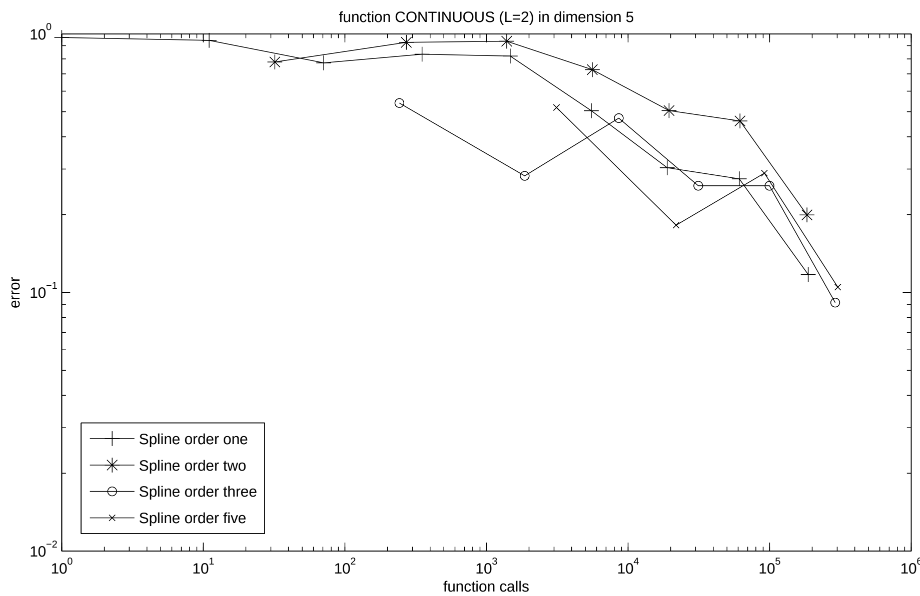

Figure 6.11: Median of absolute error of family (5), 20 integrands.

For smooth integrands one would in general expect that the order of the underlying spline has a significant effect on the convergence rate of the corresponding cubature rule. This prediction is not supported clearly by the numerical results.

We believe that this tendency is due to the unfavorable constant in the error term of cubatures based on splines with higher order. Also an important fact is, that these approximation is not optimized with respect to the considered integration problem. Namely the considered Smolyak construction was optimized with respect to the approximation problem in Chapter 4 that mainly depends on the L^2-norm. Another important fact is that the sample points were chosen to be the extrema of the Chebyshev polynomials. This choice yields a good performance for high dimensional interpolation respectively approximation problems, see, e.g. [6] but it seems to be unfavorable for our integration problem. If we compare this B-spline method with the algorithms in Chapter 5 or with the algorithms of Gerstner and Griebel, Novak and Ritter, and Petras, see [34, 48, 54], it turns out that for the families (1) to (4) the results of Novak and Ritter, and of Petras are clearly better than ours. However, the main purpose of these examples is to point out that the general method via wavelet frames yields to expedient cubature rules.

Chapter 7

Concluding remarks

In this thesis we have given an overview of a general method to construct cubature rules via wavelet techniques. The first part focused on crucial tools we used in our later error-analysis. We pointed out the basics of wavelets on bounded intervals as well as the principle of Smolyak's construction. We provided explicit algorithms for multivariate integration based on Smolyak's construction and particular one-dimensional quadrature rules. The essential aspect is that these algorithms are simple and easy to implement. We considered certain multiwavelet spaces that under proper requirements contain tensor products of classical Sobolev spaces and derived lower and upper bounds for the worst case error on these spaces. These bounds reveal that the algorithms based on multiwavelets are optimal up to logarithmic factors.

In the last part of this thesis we have shown that this approach can be extended to derive adequate cubature rules also for more general wavelet spaces. We have given a general approximation of the integral operator based on wavelet frames and Smolyak's construction. This generalization is of interest because it solves the dilemma that wavelets techniques, which guarantee an adapted approximation order and norm-equivalence to classical notions of smoothness, lead to difficulties on the boundaries of the considered domain. For wavelet frames based on univariate B-splines we have shown that the related one-dimensional approximation is a classical quadrature, i.e. defined by point evaluations. Note, we used an anisotropic version of Smolyak's construction to extend the one-dimensional approximation to a d-dimensional one. Irrespective of the error bounds we proved for this anisotropic approximation, we believe that in practice this leads to good numerical results if there is a-priori knowledge about the dependence on smoothness and direction of the integrand.

Another important question is whether there exist wavelet classes such that we find cubature rules with exactness up to a critical level L without using Smolyak's construction. The idea is based on a construction, close to the well known nets, such that the number of sample points used by the cubature is less than the number of functions that spanned the approximation spaces V_L. This takes us more and more of the subject of approximation theory, but we believe that the error analysis will be quite similar. This approach is closer to the Haar wavelet case we have considered in the first part of the thesis and the hope is that it may lead to better logarithmic factors or even optimal error bounds as in the Haar wavelet case.

Bibliography

[1] Milton Abramowitz and Irene A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, Dover, New York, 1964.

[2] Bradley Alpert, *A class of bases in l^2 for the sparse representation of integral operators*, SIAM J. Math. Anal. **24** (1993), no. 1, 246–262.

[3] Nachman Aronszajn, *Theory of reproducing kernels.*, Trans. Am. Math. Soc. **68** (1950), 337–404.

[4] Arne Barinka, Titus Barsch, Stephan Dahlke, and Michael Konik, *Some remarks on quadrature formulas for refinable functions and wavelets.*, ZAMM, Z. Angew. Math. Mech. **82** (2001), no. 12, 839–855.

[5] Arne Barinka, Titus Barsch, Stephan Dahlke, Mario Mommer, and Michael Konik, *Quadrature formulas for refinable functions and wavelets. II: Error analysis.*, J. Comput. Anal. Appl. **4** (2002), no. 4, 339–361.

[6] Volker Barthelmann, Erich Novak, and Klaus Ritter, *High dimensional polynomial interpolation on sparse grids.*, Adv. Comput. Math. **12** (2000), no. 4, 273–288.

[7] Aharon Gavriel Beged-Dov, *Lower and upper bounds for the number of lattice points in a simplex.*, SIAM J. Appl. Math. **22** (1972), 106–108.

[8] Richard E. Bellman, *Dynamic programming*, Princeton University Press, Princeton, 1957.

[9] John J. Benedetto and Oliver M. Treiber, *Wavelet frames: Multiresolution analysis and extension principles.*, Debnath, Lokenath, Wavelet transforms and time-frequency signal analysis. Boston, MA: Birkhuser. Applied and Numerical Harmonic Analysis, 3-36 , 2001.

[10] Hans-Joachim Bungartz and Michael Griebel, *Sparse grids*, Acta Numerica **13** (2004), 1–123.

[11] Charles K. Chui, Wenjie He, and Joachim Stöckler, *Nonstationary tight wavelet frames. I: Bounded intervals.*, Appl. Comput. Harmon. Anal. **17** (2004), no. 2, 141–197.

[12] Albert Cohen, *Numerical analysis of wavelet methods*, vol. 32, Elsevier, North-Holland, 2003.

[13] Stephan Dahlke, Gabriele Steidl, and Gerd Teschke, *Coorbit spaces and Banach frames on homogeneous spaces with applications to the sphere.*, Adv. Comput. Math. **21** (2004), no. 1-2, 147–180.

[14] Wolfgang Dahmen, *Multiscale analysis, approximation, and interpolation spaces.*, Chui, C. K. (ed.) et al., Approximation theory VIII. Vol. 2. Wavelets and multilevel approximation. Papers from the 8th Texas international conference, College Station, TX, USA, January 8–12, 1995. Singapore: World Scientific. Ser. Approx. Decompos. 6, 47-88 (1995)., 1995.

[15] ______, *Wavelet and multiscale methods for operator equations*, Acta Numerica **6** (1997), 55–228.

[16] Wolfgang Dahmen, Bin Han, Rong-Qing Jia, and Angela Kunoth, *Biorthogonal multiwavelets on the interval: Cubic Hermite splines.*, Constructive Approximation **16** (2000), no. 2, 221–259.

[17] Wolfgang Dahmen and Angela Kunoth, *Multilevel preconditioning*, Numerische Mathematik **63** (1992), no. 3, 315–344.

[18] Wolfgang Dahmen, Angela Kunoth, and Karsten Urban, *Biorthogonal spline wavelets on the interval – stability and moment conditions.*, Appl. Comput. Harmon. Anal. **6** (1999), no. 2, 132–196.

[19] Wolfgang Dahmen, Siegfried Prössdorf, and Reinhold Schneider, *Wavelet approximation methods for pseudodifferential equations. I: Stability and convergence.*, Math. Z. **215** (1994), no. 4, 583–620.

[20] Ingrid Daubechies, *Ten lectures on wavelets.*, CBMS-NSF Regional Conference Series in Applied Mathematics. 61. Philadelphia, PA: SIAM, Society for Industrial and Applied Mathematics. xix, 357 p. , 1992.

[21] Ingrid Daubechies, Bin Han, Amos Ron, and Zuowei Shen, Appl. Comput. Harmon. Anal. **14** (2003), no. 1, 1–46.

[22] Philip J. Davis and Philip Rabinowitz, *Methods of numerical integration. 2nd ed.*, Computer Science and Applied Mathematics. Orlando etc.: Academic Press, Inc. (Harcourt Brace Jovanovich, Publishers). XIV, 612 p., 1984.

[23] Carl de Boor, *Total positivity of the spline collocation matrix.*, Indiana Univ. Math. J. **25** (1976), 541–551.

[24] Carl De Boor, *A practical guide to splines.*, Applied Mathematical Sciences. Vol. 27. New York-Heidelberg-Berlin: Springer-Verlag. XXIV, 392 p., 1978.

[25] Stephen Demko, *Inverses of band matrices and local convergence of spline projections.*, SIAM J. Numer. Anal. **14** (1977), 616–619.

[26] Benjamin Doerr, Michael Gnewuch, and Anand Srivastav, *Bounds and constructions for the star-discrepancy via δ-covers.*, J. Complexity **21** (2005), no. 5, 691–709.

[27] Paul Van Dooren and Luc de Ridder, *An adaptive algorithm for numerical integration over an n-dimensinal cube*, Journal of Computational and Applied Mathematics (1976), 207–217.

[28] R.J. Duffin and A.C. Schaeffer, *A class of nonharmonic Fourier series.*, Trans. Am. Math. Soc. **72** (1952), 341–366.

[29] Eugene Ehrhart, *Sur un problème de géométrie diophantienne linéaires. I, II. (On à problem in linear diophantine geometry).*, J. Reine Angew. Math. **226** (1967), 1–29 (French).

[30] H. G. Feichtinger and K. Gröchenig, *Nonorthogonal wavelet and Gabor expansions, and group representations*, Wavelets and their applications, Jones and Bartlett, Boston, MA, 1992, pp. 353–375.

[31] Karin Frank and Stefan Heinrich, *Computing discrepancies of Smolyak quadrature rules.*, J. Complexity **12** (1996), no. 4, 287–314.

[32] Alan Genz, *Testing multidimensional integration routines*, Proc. of international conference on Tools, methods and languages for scientific and engineering computation (New York, NY, USA), Elsevier North-Holland, Inc., 1984, pp. 81–94.

[33] ______, *A package for testing multiple integration subroutines*, Numerical Integration, P. Keast and G. Fairweather (Eds.), D. Riedel, 1987, pp. 337–340.

[34] Thomas Gerstner and Michael Griebel, *Numerical integration using sparse grids*, Numerical Algorithms **18** (1998), no. 3–4, 209–232.

[35] Thomas Gerstner and Michael Griebel, *Dimension-adaptive tensor-product quadrature.*, Computing **71** (2003), no. 1, 65–87.

[36] Michael Gnewuch, Rene Lindloh, Reinhold Schneider, and Anand Srivastav, *Cubature formulas for function spaces with moderate smoothness.*, Preprint (2006).

[37] Stefan Heinrich, Fred J. Hickernell, and Rong-Xian Yue, *Optimal quadrature for haar wavelet spaces*, Math. Comput. **73** (2004), no. 245, 259–277.

[38] Fred J. Hickernell, *A generalized discrepancy and quadrature error bounds*, Math. Comp. **67** (1998), 299–322.

[39] Hans Johnen, *Inequalities connected with the moduli of smoothness*, Mat. Vesn., N. Ser. 9(24), 289-303, 1972.

[40] Hans Johnen and Karl Scherer, *On the equivalence of the K-functional and moduli of continuity and some applications.*, Constr. Theory Funct. several Variables, Proc. Conf. Oberwolfach 1976, Lect. Notes Math. 571, 119-140 , 1977.

[41] Samuel Karlin, *Total positivity*, Stanford University Press, Stanford, CA, 1968.

[42] Stéphane Mallat, *A wavelet tour of signal processing. 2nd ed.*, San Diego, CA: Academic Press. xxiv, 637 p. , 1999.

[43] Jiří Matoušek, *Geometric discrepancy. An illustrated guide.*, Algorithms and Combinatorics. 18. Berlin: Springer. xi, 288 p., 1999.

[44] Herbert Meschkowski, *Hilbertsche Räume mit Kernfunktion.*, Die Grundlehren der mathematischen Wissenschaften. 113. Berlin- Göttingen-Heidelberg: Springer-Verlag, VIII, 256 S. , 1962 (German).

[45] Yves Meyer, *Wavelets and operators. Translated by D. H. Salinger.*, Cambridge Studies in Advanced Mathematics. 37. Cambridge: Cambridge University Press. xv, 223 p. , 1992.

[46] W. J. Morokoff and Russel E. Caflisch, *Quasi-Monte Carlo integration*, J. Comput. Phys. **122** (1995), no. 2, 218–230.

[47] Harald Niederreiter, *Random number generation and quasi-Monte Carlo methods.*, CBMS-NSF Regional Conference Series in Applied Mathematics. 63. Philadelphia, PA: SIAM, Society for Industrial and Applied Mathematics. vi, 241 p. , 1992.

[48] Erich Novak and Klaus Ritter, *High-dimensional integration of smooth functions over cubes*, Numer. Math. **75** (1996), no. 1, 79–97.

[49] ———, *Simple cubature formulas with high polynomial exactness.*, Constructive Approximation **15** (1999), no. 4, 499–522.

[50] Roland Opfer, *Tight frame expansions of multiscale reproducing kernels in Sobolev spaces.*, Appl. Comput. Harmon. Anal. **20** (2006), no. 3, 357–374.

[51] Art B. Owen, *Monte carlo variance of scrambled net quadrature*, SIAM J. Numer. Anal. **34** (1997), no. 5, 1884–1910.

[52] Art B. Owen, *Scrambled net variance for intergrals of smooth functions.*, Ann. Stat. **25** (1997), no. 4, 1541–1562.

[53] S.H. Paskov, *Average case complexity of multivariate integration for smooth functions.*, J. Complexity **9** (1993), no. 2, 291–312.

[54] K. Petras, *Asymptotically minimal Smolyak cubature*, Tech. report, Technische Universität Braunschweig, 1999, Report.

[55] Leszek Plaskota and Grzegorz W. Wasilkowski, *The exact exponent of sparse grid quadratures in the weighted case.*, J. Complexity **17** (2001), no. 4, 840–849.

[56] Klaus F. Roth, *On irregularities of distribution*, Mathematika **1** (1954), 73–79.

[57] K. Scherer and A.Yu. Shadrin, *New upper bound for the B-spline basis condition number.*, East J. Approx. **2** (1996), no. 3, 331–342.

[58] Karl Scherer, *Weighted norm-equivalences for preconditioning.*, Kornhuber, Ralf (ed.) et al., Domain decomposition methods in science and engineering. Selected papers of the 15th international conference on domain decomposition, Berlin, Germany, July 21-25, 2003. Berlin: Springer. Lecture Notes in Computational Science and Engineering 40, 405-413 , 2005.

[59] Karl Scherer and A.Yu. Shadrin, *New upper bound for the B-spline basis condition number II. A proof of de Boor's* 2^k*-conjecture.*, J. Approximation Theory **99** (1999), no. 2, 217–229.

[60] Reinhold Schneider, *Stability of a spline collocation method for strongly elliptic multidimensional singular integral equations.*, Numer. Math. **58** (1991), no. 8, 855–873.

[61] ———, *Multiskalen- und Wavelet-Matrixkompression. (German) [Multiscale and wavelet matrix compression] Analysisbasierte Methoden zur effizienten Loesung grosser vollbesetzter Gleichungssysteme. [analysisbased methods for the efficent solution of large nonsparse systems of equations]*, Advances in Numerical Mathematics. B. G. Teubner, Stuttgart, 1998.

[62] I.J. Schoenberg and Anne Whitney, *On Plya frequency functions. III. The positivity of translation determinants with an application to the interpolation problem by spline curves.*, Trans. Am. Math. Soc. **74** (1953), 246–259.

[63] Ian H. Sloan and Stephen Joe, *Lattice methods for multiple integration*, Oxford Science Publications, The Clarendon Press Oxford University Press, New York, 1994.

[64] Ian H. Sloan and Henryk Woźniakowski, *When are quasi-Monte Carlo algorithms efficient for high dimensional integrals?*, Journal of Complexity **14** (1998), no. 1, 1–33.

[65] S. A. Smolyak, *Quadrature and interpolation formulas for tensor products of certain classes of functions*, Dokl. Akad. Nauk SSSR **148** (1963), 1042–1043, Russian, Engl. Transl.: Soviet Math. Dokl. 4:240–243, 1963.

[66] Tobias von Petersdorff, Christoph Schwab, and Reinhold Schneider, *Multiwavelets for second kind integral equations*, SIAM Journal on Numerical Analysis **34** (1997), no. 6, 2212–2227.

[67] Grzegorz W. Wasilkowski and Henryk Woźniakowski, *Explicit cost bounds of algorithms for multivariate tensor product problems*, Journal of Complexity **11** (1995), no. 1, 1–56.