DFG-Schwerpunktprogramm 1324

„Extraktion quantifizierbarer Information aus komplexen Systemen”

Quadrature for Self-Affine Distributions on $\mathbb{R}^d$

S. Dereich, T. Müller-Gronbach

Preprint 160

Edited by

AG Numerik/Optimierung
Fachbereich 12 - Mathematik und Informatik
Philipps-Universität Marburg
Hans-Meerwein-Str.
35032 Marburg

DFG-Schwerpunktprogramm 1324

„Extraktion quantifizierbarer Information aus komplexen Systemen"

Quadrature for Self-Affine Distributions on $\mathbb{R}^d$

S. Dereich, T. Müller-Gronbach

Preprint 160

QUADRATURE FOR SELF-AFFINE DISTRIBUTIONS ON $\mathbb{R}^d$

STEFFEN DEREICH AND THOMAS MÜLLER-GRONBACH

ABSTRACT. This article presents a systematic treatment of quadrature problems for self-similar probability distributions. We introduce explicit deterministic and randomized algorithms and study their errors for integrands of varying fractional smoothness of Hölder-Zymund type. Conversely, we derive lower bounds for worst case errors of arbitrary integration schemes that prove optimality of our algorithms in many cases. In particular, we see that that the effective dimension of the quadrature problem for functions of smoothness $q > 0$ is given by the quantization dimension of order q of the fractal measure.

1. INTRODUCTION

The analysis of fractal sets and measures was initiated and popularised by Mandelbrot in the 1980s, see [Man82], and has enormously advanced since then. We refer the reader to [Fal97] and [Fal] for a rigorous introduction to this field. However, despite the abundant mathematical literature on this topic, the analysis of integration problems on fractals seems to have found almost no attention up to now. We are only aware of the article [BBCR13], which studies integration problems on string-generated Cantor sets motivated by empirical findings on the structure of brain synapses, see [Cra13]. The aim of the present paper is therefore to provide a first step towards a systematic treatment of quadrature problems with respect to fractal probability measures for function classes of varying smoothness.

Let $d \in \mathbb{N}$. The computational task is to compute an approximation to the integral

$$I(f) = \int f \, \mathrm{d}P \tag{1}$$

for a self-similar probability distribution P on $\mathbb{R}^d$ and a function $f : \mathbb{R}^d \to \mathbb{R}$ by means of a deterministic or a randomized algorithm that is based on finitely many evaluations of f. In the present article we restrict attention to the case that P is self-similar with respect to a finite number of affine contractions and that f satisfies a smoothness condition of Hölder-Zygmund type.

To be more precise, we fix $m \in \mathbb{N}\backslash\{1\}$ as well as a vector $S = (S_1, \dots, S_m)$ of affine contractions $S_j : \mathbb{R}^d \to \mathbb{R}^d$ with respect to a norm $\|\cdot\|$ on $\mathbb{R}^d$ and a vector $\rho = (\rho_1, \dots, \rho_m) \in (0,1)^m$ that satisfies $\rho_1 + \dots + \rho_m = 1$. We assume that P is self-similar with respect to (S, ρ), i.e., P is the unique probability measure on $\mathbb{R}^d$ that satisfies

$$P = \sum_{j=1}^{m} \rho_j \, S_j(P).$$

In particular, P has compact support $K \subset \mathbb{R}^d$, where K is the unique nonempty compact set in $\mathbb{R}^d$ satisfying

$$K = \bigcup_{j=1}^{m} S_j(K).$$

See, e.g., Hutchinson [Hut81] and Falconer [Fal] for these facts and further properties of self-similar distributions.

Smoothness classes of integrands f are specified by an open set $\mathcal{D} \subset \mathbb{R}^d$ with $K \subset \mathcal{D}$ and a positive number $q \in (0, \infty)$. We put $q^* = \lceil q \rceil - 1 \in \mathbb{N}_0$ and we consider the set $\mathcal{F}^q = \mathcal{F}^q(\mathcal{D})$ of all functions $f \colon \mathcal{D} \to \mathbb{R}$ that are q^*-times continuously differentiable and have finite semi-norm

$$\|f\|_{\mathcal{F}^q} = \sup\left\{ \frac{|f^{(\alpha)}(x) - f^{(\alpha)}(y)|}{\|x-y\|^{q-q^*}} : \alpha \in \mathbb{N}_0^d \text{ with } |\alpha|_1 = q^* \text{ and } x, y \in \mathcal{D} \text{ with } x \neq y \right\},$$

where $|\alpha|_1 = \alpha_1 + \cdots + \alpha_d$. Note that for a bounded set $\mathcal{D}$ the spaces $\mathcal{F}^q$ are hierarchically ordered with respect to the parameter q, i.e., $\mathcal{F}^q \subset \mathcal{F}^{q'}$ for $q' < q$.

We study the worst case error of algorithms on the unit ball $\mathcal{F}_1^q$ in $\mathcal{F}^q$ and we use the worst case average number of evaluations of integrands $f \in \mathcal{F}_1^q$ to define the cost of an algorithm. We are interested in the construction of algorithms with an optimal relation of error and cost.

We briefly describe our results with a focus on error estimates. A crucial quantity in our analysis is given by the unique solution $\beta > 0$ of the equation

$$(\rho_1 r_1^q)^{\frac{\beta}{q+\beta}} + \cdots + (\rho_m r_m^q)^{\frac{\beta}{q+\beta}} = 1, \tag{2}$$

where

$$r_j = \sup_{x \neq y} \frac{\|S_j(x) - S_j(y)\|}{\|x-y\|} \in (0,1)$$

denotes the Lipschitz constant of the contraction S_j for $j = 1, \ldots, m$. If the affine contractions $S_1, \ldots, S_m$ are similarities and satisfy the open set condition, see (S1) and (S2) in Section 6, then the parameter β equals the *quantization dimension of order* q in the context of the quantization problem for the probability measure P, see [GL01], and it turns out that β can be interpreted as the *effective dimension* for the present quadrature problem.

Based on divide and conquer strategies that are adapted to the structure of the self-similarity of P we construct easy to implement deterministic composite quadrature rules $I^{(n)}$ that use n evaluations of any integrand f and achieve errors

$$|I(f) - I^{(n)}(f)| \leq c\, \|f\|_{\mathcal{F}^q}\, n^{-\frac{q}{\beta}}$$

for every $f \in \mathcal{F}^q$, where $c > 0$ is a constant that neither depends on f nor on n, see Theorem 1. We add that the actual computational cost of the method $I^{(n)}$ is proportional to n.

Employing variance reduction based on appropriate control variates we obtain randomized composite quadrature rules $\widehat{I}^{(n)}$ that use n evaluations of any integrand f and achieve errors

$$\mathbb{E}[(I(f) - \widehat{I}^{(n)}(f))^2]^{1/2} \leq c\, \|f\|_{\mathcal{F}^q}\, n^{-(\frac{q}{\beta}+\frac{1}{2})}$$

for every $f \in \mathcal{F}^q$, where $c > 0$ is a constant that neither depends on f nor on n, see Theorem 2. The algorithms $\widehat{I}^{(n)}$ use sampling from the self-similar distribution P, which is not feasible in general. We therefore provide a modified version $\bar{I}^{(n)}$ of $\widehat{I}^{(n)}$ that uses n evaluations of any integrand f and satisfies the same error estimate as $\widehat{I}^{(n)}$ but employs only the uniform distribution on finite sets, see Theorem 3. The actual computational cost of $\bar{I}^{(n)}$ is proportional to $n \log(n)$, see Remark 8.

If the affine transformations $S_1, \ldots, S_m$ are similarities and satisfy the open set condition then the deterministic algorithms $I^{(n)}$ and the randomized algorithms $\widehat{I}^{(n)}$ and $\bar{I}^{(n)}$ are worst case optimal in the following sense. There exists a constant $c > 0$ such that for any sequence of deterministic algorithms I_n with $\mathrm{cost}(I_n) \leq n$ one has

$$\sup_{f\in\mathcal{F}_1^q} |I(f) - I_n(f)| \geq c\, n^{-\frac{q}{\beta}}$$

for all $n \in \mathbb{N}$, and the same result with $n^{-\frac{q}{\beta}}$ replaced by $n^{-(\frac{q}{\beta}+\frac{1}{2})}$ is valid for sequences of randomized algorithms, see Proposition 9.

If additionally $\rho_1 = \cdots = \rho_m$ and $r_1 = \cdots = r_m$ then β equals the Hausdorff dimension of the support K of P. In particular, one recovers the classical results for quadrature with respect to the uniform distribution on the unit cube of $\mathbb{R}^d$, see [Nov88]. We add that for many linear problems with isotropic smoothness conditions the optimal order of convergence is determined by the ratio of the smoothness and the dimension, see [Rit00, NW08, NW10, NW12] for an overview and further references.

It will be convenient to formulate and prove our results in terms of semi-norms that are slightly different from the semi-norm $\|\cdot\|_{\mathcal{F}^q}$. Let $\mathcal{G}^q = \mathcal{G}^q(\mathcal{D})$ and $\mathcal{H}^q = \mathcal{H}^q(\mathcal{D})$ consist of all functions $f\colon \mathcal{D} \to \mathbb{R}$ that are q^*-times continuously differentiable and have finite semi-norm

$$\|f\|_{\mathcal{G}^q} = \sup\left\{ \frac{|D_v^{q^*} f(x) - D_v^{q^*} f(y)|}{\|x-y\|^{q-q^*}} : v \in \mathbb{R}^d \text{ with } \|v\| = 1 \text{ and } x, y \in \mathcal{D} \text{ with } x \neq y \right\}$$

and

$$\|f\|_{\mathcal{H}^q} = \sup\left\{ \frac{|D_{x-y}^{q^*} f(x) - D_{x-y}^{q^*} f(y)|}{\|x-y\|^{q}} : x, y \in \mathcal{D} \text{ with } x \neq y \right\},$$

respectively, where $D_v^{q^*} f$ denotes the q^*-th directional derivative of f along $v \in \mathbb{R}^d$. The corresponding unit balls are denoted by $\mathcal{G}_1^q$ and $\mathcal{H}_1^q$. We have

$$\|f\|_{\mathcal{H}^q} \leq \|f\|_{\mathcal{G}^q} \leq \|f\|_{\mathcal{F}^q} \sup\{|v|_1^{q^*} : v \in \mathbb{R}^d, \|v\| = 1\} \tag{3}$$

and

$$\|f\|_{\mathcal{F}^q} \leq \|f\|_{\mathcal{G}^q} \sup\{\|v\|^{q^*} : v \in \mathbb{R}^d, |v|_2 = 1\}, \tag{4}$$

where $|v|_p$ denotes the p-norm of $v \in \mathbb{R}^d$ for $p = 1, 2$. In particular, $\mathcal{F}^q = \mathcal{G}^q \subset \mathcal{H}^q$. We add that the inequalities in (3) are straightforward to show while inequality (4) is a consequence of the fact that the polarization constant of a Hilbert space equals one. See, e.g., [Din99, Proposition 1.44] for a proof of the latter fact.

Throughout the following we restrict attention to the case $\mathcal{D} = \mathbb{R}^d$ in order to avoid technicalities and simplify notation. The reader can easily verify that the algorithms $I^{(n)}$, $\widehat{I}^{(n)}$, $\bar{I}^{(n)}$ constructed in Sections 4 and 5 are well-defined and satisfy the corresponding stated error bounds also in the case $\mathcal{D} \subsetneq \mathbb{R}^d$ for sufficiently large indices n, see also Remark 1 for the details in the case of the deterministic algorithms $I^{(n)}$. The worst case errors shrink with increasing $\mathcal{D}$, so that our lower bounds for $\mathcal{D} = \mathbb{R}^d$, see Proposition 9, immediately carry over to the general case $\mathcal{D} \subset \mathbb{R}^d$.

We briefly outline the content of the paper. In Section 2 we collect some basic notation and definitions. In Section 3 we introduce the notion of a cutset and motivate its crucial role for the quadrature problem. Sections 4 and 5 are devoted to the construction and error analysis of deterministic and randomized composite quadrature rules, respectively. Lower error bounds that hold for arbitrary methods based on finitely many function evaluations are presented in Section 6. The Appendix contains recursion formulas for the moments of a self-affine probability measure.

2. Preliminaries

We use $\#A$ to denote the cardinality of a set A. The identity mapping on A is denoted by id_A. Furthermore, $\mathrm{cl}(A)$ means the closure of a set $A \subset \mathbb{R}^d$ and

$$B(x,R) = \{y \in \mathbb{R}^d \colon \|y-x\| < R\},\ \overline{B}(x,R) = \{y \in \mathbb{R}^d \colon \|y-x\| \leq R\}$$

denote the open and the closed ball in $\mathbb{R}^d$ with center $x \in \mathbb{R}^d$ and radius $R \geq 0$, respectively.

The transpose of a $x \in \mathbb{R}^d$ and $V \in \mathbb{R}^{d\times d}$ are denoted by x^{T} and V^{T}, respectively. Furthermore, for a multi-index $\alpha \in \mathbb{N}_0^d$ we put

$$x^\alpha = x_1^{\alpha_1}\cdots x_d^{\alpha_d}, \quad \alpha! = \alpha_1!\cdots\alpha_d!.$$

For $q > 0$ we put $q^* = \lceil q\rceil - 1$ as in the introduction, and we define the fractional factorial of q by

$$q! = q(q-1)\cdot(q-q^*+1).$$

Let μ be a signed measure on the Borel sets of $\mathbb{R}^d$. Then $\mathrm{supp}(\mu)$ and $|\mu|$ denote the support and the total variation of μ, respectively, and by $\|\mu\|_{\mathrm{TV}}$ we mean the total variation norm of μ. If μ has compact support then

$$\mathrm{rad}(\mu) = \inf\{R \geq 0 \colon \exists\, x \in \mathbb{R}^d \colon\ \mathrm{supp}(\mu) \subset \overline{B}(x,R)\} \in [0,\infty)$$

is called the *effective radius* of μ. It is easy to see that the infimum is attained, i.e., there exists $x \in \mathbb{R}^d$ such that

$$\mathrm{supp}(\mu) \subset \overline{B}(x, \mathrm{rad}(\mu)).$$

The uniform distribution on a finite, non-empty set A is denoted by U_A.

Finally, $\mathcal{P}_k$ denotes the class of polynomials $p \colon \mathbb{R}^d \to \mathbb{R}$ of order at most $k \in \mathbb{N}_0$.

3. Cutsets and Self-Similarity

Consider the set

$$\mathcal{T} = \{\lambda\} \cup \bigcup_{\ell=1}^{\infty} \{1,\ldots,m\}^\ell$$

of all words over the alphabet $\{1,\ldots,m\}$ with λ denoting the empty word. We interpret $\mathcal{T}$ as a complete m-ary tree with root λ and we use this structure to encode the self-similarity of the probability measure P.

Let $\mathbf{j} = (j_1,\ldots,j_\ell) \in \mathcal{T}\backslash\{\lambda\}$. We use

$$\mathbf{j}_- = \begin{cases} (j_1,\ldots,j_{\ell-1}), & \text{if } \ell \geq 2,\\ \lambda, & \text{if } \ell = 1,\end{cases}$$

to denote the father of $\mathbf{j}$ and

$$|\mathbf{j}| = \ell$$

to denote the level of $\mathbf{j}$. We define

$$S_{\mathbf{j}} = S_{j_1} \circ \cdots \circ S_{j_\ell}$$

and we put

$$x_{\mathbf{j}} = x_{j_1}\cdots x_{j_\ell}$$

for any vector $x = (x_1,\ldots,x_m) \in \mathbb{R}^m$. Thus, $S_{\mathbf{j}}$ is an affine contraction with respect to the norm $\|\cdot\|$ and the Lipschitz constant of $S_{\mathbf{j}}$ is bounded from above by

$$r_{\mathbf{j}} = r_{j_1}\cdots r_{j_\ell}.$$

Furthermore, we put
$$|\lambda| = 0,\ S_\lambda = \mathrm{id}_{\mathbb{R}^d},\ x_\lambda = 1,$$
where x is any vector of real numbers. Finally, for any signed measure μ on the Borel sets of $\mathbb{R}^d$ and any vertex $\mathbf{j} \in \mathcal{T}$ we use
$$\mu_{\mathbf{j}} = S_{\mathbf{j}}(\mu)$$
to denote the image of $\mu_{\mathbf{j}}$ with respect to $S_{\mathbf{j}}$.

We turn to the notion of a cutset in the rooted tree $\mathcal{T}$, which will serve as a crucial tool for the construction of good quadrature rules. A finite subset $\mathcal{C} \subset \mathcal{T}$ is called a *cutset* or *tight*, if $\mathcal{C} = \{\lambda\}$ or if for any sequence $(j_k)_{k\in\mathbb{N}} \in \{1, \dots, m\}^{\mathbb{N}}$ there exists a unique $\ell \in \mathbb{N}$ such that
$$(j_1, \dots, j_\ell) \in \mathcal{C},$$
see [Hut81]. In other words, any self-avoiding path in $\mathcal{T}$ from the root to infinity meets a cutset $\mathcal{C}$ in exactly one vertex of $\mathcal{C}$.

Proposition 1. *Let $\mathcal{C} \subset \mathcal{T}$ be a cutset. Then*
$$P = \sum_{\mathbf{j}\in\mathcal{C}} \rho_{\mathbf{j}}\, P_{\mathbf{j}}.$$

Proposition 1 naturally leads to a divide and conquer strategy for the construction of quadrature rules. Roughly speaking, a cutset $\mathcal{C}$ divides the quadrature problem for P into $\#\mathcal{C}$ subproblems specified by the weighted probability measures $\rho_{\mathbf{j}}\, P_{\mathbf{j}}$ with $\mathbf{j} \in \mathcal{C}$. With increasing level $|\mathbf{j}|$ the diameter of the support $S_{\mathbf{j}}(K)$ of $P_{\mathbf{j}}$ shrinks and, consequently, the quality of a Taylor approximation of an integrand f on $S_{\mathbf{j}}(K)$ improves.

Proof of Proposition 1. We prove the proposition by induction over the cardinality of the cutset. Clearly, $\#\mathcal{C} = 1$ implies $\mathcal{C} = \{\lambda\}$, in which case the statement is trivial. Next, let $n \in \mathbb{N}$ and suppose that the statement is true for any cutset with cardinality less than or equal to n. Let $\mathcal{C}$ be a cutset of cardinality $n + 1$. Pick a vertex $\mathbf{j}^* = (j_1^*, \dots, j_{\ell^*}^*) \in \mathcal{C}$ with
$$\ell^* = \max_{\mathbf{j}\in\mathcal{C}} |\mathbf{j}|.$$
Then $(j_1^*, \dots, j_k^*) \notin \mathcal{C}$ for $k = 1, \dots, \ell^* - 1$ and therefore
$$(\mathbf{j}_-^*, 1), \dots, (\mathbf{j}_-^*, m) \in \mathcal{C} \tag{5}$$
since otherwise one of the paths $(j_1^*, \dots, j_{\ell^*-1}^*, j, j, \dots)$ with $j \in \{1, \dots, m\}$ would not meet the cutset $\mathcal{C}$. Put
$$\mathcal{C}' = \mathcal{C}\backslash\{(\mathbf{j}_-^*, 1), \dots, (\mathbf{j}_-^*, m)\} \cup \{\mathbf{j}_-^*\}$$
Then $\mathcal{C}'$ is a cutset since for all paths to infinity it is equivalent to hit one of the vertices $(\mathbf{j}_-^*, 1), \dots, (\mathbf{j}_-^*, m)$ or the vertex $\mathbf{j}_-^*$. Furthermore, $\#\mathcal{C}' \le n$, due to (5) and $m \ge 2$.

By self-similarity of P with respect to (S, ρ) we have
$$\sum_{j=1}^{m} \rho_j\, P_{(\mathbf{j}_-^*, j)}(A) = \sum_{j=1}^{m} \rho_j\, P_j(S_{\mathbf{j}_-^*}^{-1}(A)) = P(S_{\mathbf{j}_-^*}^{-1}(A)) = P_{\mathbf{j}_-^*}(A)$$
for all Borel sets $A \subset \mathbb{R}^d$ and, consequently,
$$P = \sum_{\mathbf{j}\in\mathcal{C}'} \rho_{\mathbf{j}}\, P_{\mathbf{j}} = \sum_{\mathbf{j}\in\mathcal{C}'\backslash\{\mathbf{j}_-^*\}} \rho_{\mathbf{j}}\, P_{\mathbf{j}} + \rho_{\mathbf{j}_-^*} \sum_{j=1}^{m} \rho_j\, P_{(\mathbf{j}_-^*, j)} = \sum_{\mathbf{j}\in\mathcal{C}} \rho_{\mathbf{j}}\, P_{\mathbf{j}},$$
which completes the proof. □

4. Deterministic Composite Quadrature Rules

In this section we introduce and analyse deterministic quadrature formulas for the class $\mathcal{H}^q = \mathcal{H}^q(\mathbb{R}^d)$. For a signed measure Q on $\mathbb{R}^d$ with finite support we define a corresponding deterministic quadrature rule $I_Q \colon \mathcal{H}^q \to \mathbb{R}$ by

$$I_Q(f) = \int f\,dQ.$$

and we call Q and I_Q *exact* of order q^* if

$$I_Q(p) = I(p)$$

for every $p \in \mathcal{P}_{q^*}$.

We estimate the worst case error

$$e_Q = \sup_{f \in \mathcal{H}_1^q} |I(f) - I_Q(f)|$$

of I_Q on the unit ball $\mathcal{H}_1^q$ in terms of the effective radius and the total variation norm of $P - Q$.

Proposition 2. *If Q is exact of order q^* then*

$$e_Q \le \|P - Q\|_{\mathrm{TV}} \frac{(\mathrm{rad}(P-Q))^q}{q!}.$$

Otherwise, $e_Q = \infty$.

Proof. Since $\mathcal{P}_{q^*}$ is a linear subspace of $\mathcal{H}_1^q$ the worst case error of I_Q on $\mathcal{H}_1^q$ is infinite if Q is not exact of order q^*. Assume now that Q is exact of order q^*. Put $R = \mathrm{rad}(P-Q)$ and choose $x_0 \in \mathbb{R}^d$ such that $\mathrm{supp}(P-Q) \subset \overline{B}(x_0, R)$. Let $f \in \mathcal{H}^q$ and consider the q^*-th order Taylor-polynomial p of f at x_0. Let $x \in \overline{B}(x_0, R) \setminus \{x_0\}$. In order to obtain an estimate for $f(x) - p(x)$, we consider the function

$$h \colon [0,1] \to \mathbb{R}, \quad t \mapsto f(x_0 + (x - x_0)t) - p(x_0 + (x - x_0)t).$$

Clearly,

$$h^{(k)}(t) = D^k_{x-x_0} f(x_0 + (x-x_0)t) - D^k_{x-x_0} f(x_0)$$

and, in particular, $h^{(k)}(0) = 0$ for $k = 0, \dots, q^*$. Moreover,

$$|h^{(q^*)}(t)| = t^{-q^*} |D^{q^*}_{t(x-x_0)} f(x_0 + (x-x_0)t) - D^{q^*}_{t(x-x_0)} f(x_0)| \le t^{q-q^*} \|x - x_0\|^q \|f\|_{\mathcal{H}^q}.$$

Hence

$$|h^{(q^*-1)}(t)| = \left| \int_0^t h^{(q^*)}(s)\,ds \right| \le \frac{1}{1+q-q^*} t^{1+q-q^*} R^q \|f\|_{\mathcal{H}^q}.$$

Iterating this argument we obtain

$$|h(t)| \le \frac{1}{q!} t^q R^q \|f\|_{\mathcal{H}^q},$$

which implies

$$|f(x) - p(x)| \le \frac{R^q}{q!} \|f\|_{\mathcal{H}^q}.$$

Thus

$$\left| \int f\,\mathrm{d}P - \int f\,\mathrm{d}Q \right| = \left| \int (f - p)\,\mathrm{d}(P - Q) \right| \le \|P - Q\|_{\mathrm{TV}} \frac{R^q}{q!} \|f\|_{\mathcal{H}^q}$$

as claimed. □

For a signed measure Q on $\mathbb{R}^d$ with finite support and a cutset $\mathcal{C} \subset \mathcal{T}$ we consider the signed measure

$$Q^{\mathcal{C}} = \sum_{\mathbf{j} \in \mathcal{C}} \rho_{\mathbf{j}} Q_{\mathbf{j}},$$

which leads to the deterministic composite quadrature rule $I_{Q^{\mathcal{C}}} : \mathcal{H}^q \to \mathbb{R}$ given by

$$I_{Q^{\mathcal{C}}}(f) = \int f \,\mathrm{d}Q^{\mathcal{C}} = \sum_{\mathbf{j} \in \mathcal{C}} \rho_{\mathbf{j}} \int f \,\mathrm{d}Q_{\mathbf{j}}.$$

We provide an estimate of the error of $I_{Q^{\mathcal{C}}}$, which seperates the effects of the choice of Q and the choice of the cutset $\mathcal{C}$.

Proposition 3. *For every $f \in \mathcal{H}^q$ we have*

$$|I(f) - I_{Q^{\mathcal{C}}}(f)| \leq e_Q \, \|f\|_{\mathcal{H}^q} \sum_{\mathbf{j} \in \mathcal{C}} \rho_{\mathbf{j}} \, r_{\mathbf{j}}^q.$$

Proof. By Proposition 1,

$$I(f) - I_{Q^{\mathcal{C}}}(f) = \sum_{\mathbf{j} \in \mathcal{C}} \rho_{\mathbf{j}} \Big(\int f \,\mathrm{d}P_{\mathbf{j}} - \int f \,\mathrm{d}Q_{\mathbf{j}} \Big).$$

Let $\mathbf{j} \in \mathcal{T}$. Since $S_{\mathbf{j}}$ is affine there exist $A \in \mathbb{R}^{d \times d}$ and $b \in \mathbb{R}^d$ such that $S_{\mathbf{j}}(x) = Ax + b$ and $\|Ax\| \leq r_{\mathbf{j}} \|x\|$ for every $x \in \mathbb{R}^d$. In particular, we have $f \circ S_{\mathbf{j}} \in C^{q^*}(\mathbb{R}^d)$ and therefore

$$\left| \int f \,\mathrm{d}P_{\mathbf{j}} - \int f \,\mathrm{d}Q_{\mathbf{j}} \right| = \left| \int f \circ S_{\mathbf{j}} \,\mathrm{d}P - \int f \circ S_{\mathbf{j}} \,\mathrm{d}Q \right| \leq \|f \circ S_{\mathbf{j}}\|_{\mathcal{H}^q} \, e(Q).$$

Let $x, y \in \mathbb{R}^d$ with $x \neq y$ and put $v = x - y$ and $u = S_{\mathbf{j}}(x) - S_{\mathbf{j}}(y)$. Then

$$D_v^{q^*}(f \circ S_{\mathbf{j}}) = \big(D_u^{q^*} f\big) \circ S_{\mathbf{j}}.$$

Hence

$$|D_v^{q^*}(f \circ S_{\mathbf{j}})(x) - D_v^{q^*}(f \circ S_{\mathbf{j}})(y)| \leq \|S_{\mathbf{j}}(x) - S_{\mathbf{j}}(y)\|^q \, \|f\|_{\mathcal{H}^q} \leq r_{\mathbf{j}}^q \, \|x - y\|^q \, \|f\|_{\mathcal{H}^q}.$$

We conclude that

$$\|f \circ S_{\mathbf{j}}\|_{\mathcal{H}^q} \leq r_{\mathbf{j}}^q \, \|f\|_{\mathcal{H}^q}, \tag{6}$$

which finishes the proof. □

The error estimate from Proposition 3 suggests to consider cutsets $\mathcal{C}$ such that all of the terms $\rho_{\mathbf{j}} r_{\mathbf{j}}^q$ with $\mathbf{j} \in \mathcal{C}$ are approximately of the same size. To this end we put

$$s_j = \rho_j r_j^q$$

for $j \in \{1, \ldots, m\}$ and we define cutsets $\mathcal{C}(T)$ for $T \geq 1$ by

$$\mathcal{C}(T) = \{\mathbf{j} \in \mathcal{T} \backslash \{\lambda\} \colon 1/T \leq s_{\mathbf{j}_-} \text{ and } s_{\mathbf{j}} < 1/T\}. \tag{7}$$

Recall the definition (2) of the effective dimension $\beta > 0$ and put

$$\theta = \frac{\beta}{\beta + q}. \tag{8}$$

Thus, $\theta \in (0, 1)$ is the unique positive number that satisfies

$$s_1^\theta + \cdots + s_m^\theta = 1. \tag{9}$$

The following result states that the size of the cutsets $\mathcal{C}(T)$ essentially behaves like T^θ. Put

$$s_{\min} = \min_{j=1,\ldots,m} s_j, \ \ s_{\max} = \max_{j=1,\ldots,m} s_j.$$

Proposition 4. *For every $T \geq 1$ we have*

$$T^\theta < \#\mathcal{C}(T) \leq s_{\min}^{-\theta}\, T^\theta.$$

Proof. Put $\mathcal{C}(T) = \{\lambda\}$ for $T \in (0,1)$. The proof of the proposition is based on the recursion

$$\forall\, T \geq 1\colon\ \#\mathcal{C}(T) = \sum_{k=1}^m \#\mathcal{C}(s_k T). \tag{10}$$

For a proof of (10) let $T \geq 1$ and put

$$\mathcal{C}_k(T) = \{\mathbf{j} \in \mathcal{C}(T)\colon j_1 = k\}$$

for $k = 1, \ldots, m$. Clearly,

$$\mathcal{C}(T) = \bigcup_{k=1}^m \mathcal{C}_k(T)$$

and $\mathcal{C}_{\tilde{k}}(T) \cap \mathcal{C}_k(T) = \emptyset$ if $k \neq \tilde{k}$. Moreover, it is easy to check that $\mathbf{j} \mapsto (k, \mathbf{j})$ defines a bijection from $\mathcal{C}(s_k T)$ to $\mathcal{C}_k(T)$, and therefore

$$\#\mathcal{C}(T) = \sum_{k=1}^m \#\mathcal{C}_k(T) = \sum_{k=1}^m \#\mathcal{C}(s_k T).$$

We prove by induction that

$$\forall\, T \in [s_{\min}, s_{\max}^{-n})\colon\ T^\theta < \#\mathcal{C}(T) \leq s_{\min}^{-\theta}\, T^\theta \tag{11}$$

for every $n \in \mathbb{N}_0$, which clearly implies the statement of the proposition.

For $n = 0$ and $T \in [s_{\min}, s_{\max}^{-n}) = [s_{\min}, 1)$ we have

$$T^\theta < 1 = \#\mathcal{C}(T) \leq (s_{\min}^{-1} T)^\theta$$

as claimed. Next, assume that (11) holds for some $n \in \mathbb{N}_0$ and let $T \in [s_{\max}^{-n}, s_{\max}^{-n-1})$. Then $s_k T \in [s_{\min}, s_{\max}^{-n})$ for all $k \in \{1, \ldots, m\}$ and by (9) and (11) we obtain

$$T^\theta = \sum_{k=1}^m (s_k T)^\theta < \sum_{k=1}^m \#\mathcal{C}(s_k T) \leq \sum_{k=1}^m s_{\min}^{-\theta}\, (s_k T)^\theta = s_{\min}^{-\theta}\, T^\theta.$$

Apply (10) to complete the proof of the induction step. □

We turn to the analysis of the quadrature rules $I_{Q^{\mathcal{C}(T)}}$.

Proposition 5. *Let Q be exact of order q^* and $T \geq 1$. The quadrature rule $I_{Q^{\mathcal{C}(T)}}$ uses*

$$\#\operatorname{supp}(Q^{\mathcal{C}(T)}) \leq \#\operatorname{supp}(Q)\, s_{\min}^{-\theta}\, T^\theta$$

function evaluations and satisfies for every $f \in \mathcal{H}^q$,

$$|I(f) - I_{Q^{\mathcal{C}(T)}}(f)| \leq \|P - Q\|_{\mathrm{TV}}\, \frac{(\operatorname{rad}(P-Q))^q}{q!}\, s_{\min}^{-\theta}\, \|f\|_{\mathcal{H}^q}\, T^{-(1-\theta)}.$$

Proof. By Proposition 4,

$$\#\operatorname{supp}(Q^{\mathcal{C}(T)}) \leq \#\mathcal{C}(T)\,\#\operatorname{supp}(Q) \leq \#\operatorname{supp}(Q)\, s_{\min}^{-\theta}\, T^{\theta}.$$

By Proposition 2 and Proposition 3,

$$|I(f) - I_{Q^{\mathcal{C}(T)}}(f)| \leq \|P - Q\|_{\mathrm{TV}} \frac{(\operatorname{rad}(P-Q))^q}{q!} \|f\|_{\mathcal{H}^q} \sum_{\mathbf{j}\in\mathcal{C}(T)} s_{\mathbf{j}}.$$

By the definition of $\mathcal{C}(T)$ and Proposition 4,

$$\sum_{\mathbf{j}\in\mathcal{C}(T)} s_{\mathbf{j}} < T^{-1}\,\#\mathcal{C}(T) \leq s_{\min}^{-\theta}\, T^{-(1-\theta)},$$

which completes the proof. □

Put

$$T_n = \max\bigl(1, s_{\min}\,(\#\operatorname{supp}(Q))^{-1/\theta}\, n^{1/\theta}\bigr)$$

and define

$$I_Q^{(n)} = I_{Q^{\mathcal{C}(T_n)}}$$

for every $n \in \mathbb{N}$. The following error estimate is an immediate consequence of Proposition 5.

Theorem 1. *Let Q be exact of order q^*. For every $n \in \mathbb{N}$ with $n \geq \#\operatorname{supp}(Q)\, s_{\min}^{-\theta}$ the quadrature rule $I_Q^{(n)}$ uses at most n function evaluations and satisfies for all $f \in \mathcal{H}^q$,*

$$|I(f) - I_Q^{(n)}(f)| \leq \|P - Q\|_{\mathrm{TV}} \frac{(\operatorname{rad}(P-Q))^q}{q!}\, s_{\min}^{-1}\,(\#\operatorname{supp}(Q))^{q/\beta}\, \|f\|_{\mathcal{H}^q}\, n^{-q/\beta}.$$

Let us now explain how the arguments have to be changed in order to cover the quadrature problem on general domains $\mathcal{D}$.

Remark 1. In order to obtain error estimates on a general domain $\mathcal{D}$ one chooses a ball B that contains the support of $P - Q$ and defines

$$e_Q = \sup_{f\in\mathcal{H}_1^q(B)} |I(f) - I_Q(f)|.$$

Proposition 2 remains true as is and Proposition 3 remains true for $f \in \mathcal{H}^q(\mathcal{D})$ provided that

$$\bigcup_{\mathbf{j}\in\mathcal{C}} S_{\mathbf{j}}(B) \subset \mathcal{D}. \tag{12}$$

Consequently, Theorem 1 is also true for $f \in \mathcal{H}^q(\mathcal{D})$, if (12) is fulfilled for the corresponding cutsets. This is always the case for sufficiently large n since $S_{\mathbf{j}}(B) \not\subset \mathcal{D}$ holds only for a finite number of $\mathbf{j} \in \mathcal{T}$. Indeed, this follows from the estimate

$$d(S_{\mathbf{j}}(B), K) := \sup_{x\in S_{\mathbf{j}}(B)} \min_{y\in K} \|x - y\| \leq r_{\mathbf{j}}\, d(B, K) \qquad (\mathbf{j} \in \mathcal{T})$$

and the fact that there is a $\varepsilon > 0$ such that $K + B(0,\varepsilon) \subset \mathcal{D}$ due to compactness of K.

Remark 2. If Q is exact of order q^* then, by Theorem 1 and (3), the worst case errors of $I_Q^{(n)}$ on $\mathcal{G}_1^q$ and $\mathcal{F}_1^q$ converge to zero as $n \to \infty$ at least with order q/β in terms of the number of function evaluations. In Section 6 we show that this order of convergence is optimal within the class of all deterministic algorithms based on finitely many function evaluations if the contractions $S_1, \ldots, S_m$ are similarities and satisfy the open set condition, see Proposition 9 and Theorem 4.

Remark 3. Assume that

$$r_1 = \cdots = r_m = r, \quad \rho_1 = \cdots = \rho_m = 1/m.$$

Then $s_1 = \cdots = s_m = r^q/m$ and the cutsets $\mathcal{C}(T)$ are given by

$$\mathcal{C}(T) = \{\mathbf{j} \in \mathcal{T} \colon |\mathbf{j}| = \lfloor \ln(T)/(\ln(m) + q\ln(1/r)) \rfloor + 1\}.$$

Moreover,

$$\beta = \frac{\ln(m)}{\ln(1/r)}, \quad \theta = \frac{\ln(m)}{\ln(m/r^q)}.$$

Assume that $\|\cdot\| = |\cdot|_2$ and $S_1, \dots, S_m$ are similarities, see (S1) in Section 6. Then β equals the similarity dimension of S since

$$\sum_{j=1}^m r_j^\beta = m\, r^\beta = 1.$$

If, additionally, $S_1, \dots, S_m$ satisfy the open set condition, see (S2) in Section 6, then β coincides with the Hausdorff dimension of $\mathrm{supp}(P)$. See [Hut81] for these facts.

In the particular case of the classical quadrature problem for the uniform distribution on the unit cube $[0,1]^d$ we have

$$m = 2^d, \quad r = 1/2, \quad \beta = d,$$

and $I_Q^{(n)}$ is a classical composite quadrature rule with the well-known order of convergence q/d.

Remark 4. By a result of Tchakaloff there exists a probability measure Q on $\mathbb{R}^d$, which is exact of order q^* and satisfies

$$\mathrm{supp}(Q) \subset \mathrm{supp}(P), \ \#\,\mathrm{supp}(Q) \le \binom{d+q^*}{d}, \tag{13}$$

see [Tch57, BT06]. In particular, $\mathrm{rad}(P-Q) = \mathrm{rad}(P)$ and $\|P-Q\|_{\mathrm{TV}} = 2$ can be used for the error estimates in Proposition 5 and Theorem 1 for this choice of Q.

Explicit constructions of signed measures Q with finite support that are exact of order q^* can be obtained, e.g., by polynomial interpolation methods, see Section 5, since the moments

$$\int x^\alpha \,\mathrm{d}P(x), \quad \alpha \in \mathbb{N}_0^d,$$

of the self-affine measure P can easily be computed in a recursive way, see Proposition 10 in the appendix.

For example, if $d = 1$ and $\#\,\mathrm{supp}(P) = \infty$ then Gauss-Christoffel quadrature rules can be determined by using orthonormal polynomials with respect to P, see, e.g., [MM08]. If $q^* \in 2\mathbb{N}-1$ we obtain a probability measure Q that is exact of order q^* and satisfies

$$\mathrm{supp}(Q) \subset \mathrm{Co}(\mathrm{supp}(P)), \ \#\,\mathrm{supp}(Q) = (q^*+1)/2,$$

where $\mathrm{Co}(A)$ denotes the convex hull of a set $A \subset \mathbb{R}$. In particular, $\mathrm{rad}(P-Q) = \mathrm{rad}(P)$. Thus, by Theorem 1 the corresponding composite Gauss-Christoffel quadrature rules $I_Q^{(n)}$ satisfy for all $f \in \mathcal{H}^q$ and $n \ge m(q^*+1)/2$,

$$|I(f) - I_Q^{(n)}(f)| \le 2\,\frac{(\mathrm{rad}(P))^q}{q!} \left(\frac{q^*+1}{2}\right)^{q/\beta} \frac{m}{r^q}\, \|f\|_{\mathcal{H}^q}\, n^{-q/\beta}. \tag{14}$$

Example 1. Consider a generalized Cantor distribution P on $\mathbb{R}$, i.e.,

$$d = 1,\ m = 2,\ \rho_1 = \rho_2 = 1/2$$

and

$$S_1(x) = rx,\ S_2(x) = (1 - r) + rx$$

for $x \in \mathbb{R}$, where $r \in (0, 1/2)$. Then

$$\mathrm{rad}(P) = 1/2,\ \beta = \frac{\ln(2)}{\ln(1/r)}.$$

Let $q^* \in 2\mathbb{N} - 1$. Then the corresponding composite Gauss-Christoffel quadrature rules satisfy for all $f \in \mathcal{H}^q$ and $n \geq q^* + 1$,

$$|I(f) - I_Q^{(n)}(f)| \leq \frac{4(q^*+1)^{q \ln(1/r)/\ln(2)}}{q!\,2^q} \|f\|_{\mathcal{H}^q}\, n^{-q \ln(1/r)/\ln(2)},$$

see (14).

Consider the classical case $r = 1/3$. Then the moments

$$\nu_k = \int x^k \,\mathrm{d}P, \quad k \in \mathbb{N}_0,$$

satisfy the recursive relation

$$\nu_k = \frac{2^{k-1}}{3^k - 1} \sum_{i=0}^{k-1} \binom{k}{i} 2^{-i}\, \nu_i$$

for $k \in \mathbb{N}$, see Remark 9 in the appendix. We determine the Gauss-Christoffel quadrature rule I_Q for $q^* = 3$. We have

$$\nu_0 = 1,\ \nu_1 = 1/2,\ \nu_2 = 3/8,\ \nu_3 = 5/16$$

and the first three orthonormal polynomials with respect to the Cantor distribution P are given by $p_0 = 1$ and

$$p_1(x) = \sqrt{8}(x - 1/2),\ p_2(x) = \sqrt{160}(x^2 - x + 1/8)$$

for $x \in \mathbb{R}$. The support of Q is given by the zeros of p_2, i.e.,

$$\mathrm{supp}(Q) = \{1/2 - \sqrt{1/8}, 1/2 + \sqrt{1/8}\},$$

and the corresponding weights are $1/2$ each. Thus

$$Q = \frac{1}{2}\delta_{1/2-\sqrt{1/8}} + \frac{1}{2}\delta_{1/2+\sqrt{1/8}}.$$

Clearly, $\mathrm{supp}(Q) \not\subset \mathrm{supp}(P)$. A probability measure Q that is exact of order 3 and satisfies $\mathrm{supp}(Q) \subset \mathrm{supp}(P)$ is, e.g., given by

$$Q = \frac{63}{288}(\delta_0 + \delta_1) + \frac{81}{288}(\delta_{1/3} + \delta_{2/3}).$$

Example 2. We consider a self-affine distribution on the Koch curve. Thus

$$d = 2,\ \|\cdot\| = |\cdot|_2,\ m = 4,$$

and the affine contractions $S_1, \dots, S_4$ are given by

$$S_j(x) = A_j x + b_j$$

with

$$A_1 = A_2 = \frac{1}{3}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},\ A_3 = A_4^{\mathrm{T}} = \frac{1}{6}\begin{pmatrix} 1 & -\sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix}$$

and

$$b_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix},\ b_2 = \begin{pmatrix} 2/3 \\ 0 \end{pmatrix},\ b_3 = \begin{pmatrix} 1/3 \\ 0 \end{pmatrix},\ b_4 = \begin{pmatrix} 1/2 \\ \sqrt{3}/6 \end{pmatrix}.$$

In particular,

$$r_1 = r_2 = r_3 = r_4 = 1/3.$$

Take

$$\rho_1 = \rho_2 = \rho_3 = \rho_4 = 1/4$$

and consider the probability measure P that is self-similar with respect (ρ, S). Using Proposition 10 in the appendix we obtain

$$\int x^\alpha \, dP(x) = \begin{cases} 1/2, & \text{if } \alpha = (1,0), \\ \sqrt{3}/18, & \text{if } \alpha = (0,1), \\ 19/60, & \text{if } \alpha = (2,0), \\ 1/60, & \text{if } \alpha = (0,2), \\ \sqrt{3}/36, & \text{if } \alpha = (1,1). \end{cases}$$

Define a probability measure Q on $\mathbb{R}^2$ by

$$Q = \sum_{i=1}^{6} w_i \, \delta_{x_i}$$

with

$$x_i = \begin{cases} (1/6, \sqrt{3}/18), & \text{if } i = 1, \\ (1/3, 0), & \text{if } i = 2, \\ (1/3, \sqrt{3}/9), & \text{if } i = 3, \\ (2/3, \sqrt{3}/9), & \text{if } i = 4, \\ (5/6, \sqrt{3}/18), & \text{if } i = 5, \\ (1, 0), & \text{if } i = 6, \end{cases} \qquad w_i = \begin{cases} 1/10, & \text{if } i = 1, \\ 3/10, & \text{if } i = 2, \\ 2/10, & \text{if } i = 3, \\ 2/10, & \text{if } i = 4, \\ 1/10, & \text{if } i = 5, \\ 1/10, & \text{if } i = 6. \end{cases}$$

Then $\operatorname{supp}(Q) \subset \operatorname{supp}(P)$ and it is easy to check that

$$\int x^\alpha \, dQ(x) = \int x^\alpha \, dP(x)$$

for all $\alpha \in \mathbb{N}_0^2$ with $|\alpha|_1 \leq 2$. Thus Q is exact of order 2 and satisfies (13).

We have

$$\operatorname{rad}(P) = 1/2,\ \beta = \frac{\ln(4)}{\ln(3)},\ s_{\min} = 3^{-3}/4,$$

and therefore

$$|I(f) - I_Q^{(n)}(f)| \leq 8 \frac{(3/2)^q}{q!} \, 6^{q \ln(3)/\ln(4)} \, \|f\|_{\mathcal{H}^q} \, n^{-q \ln(3)/\ln(4)}$$

for all $n \geq 24$ and $f \in \mathcal{H}^q$ with $2 < q \leq 3$, see Theorem 1 and Remarks 3 and 4.

5. Randomized Composite Quadrature Formulas

We combine the cutset technique from the previous section with a variance reduction approach to construct randomized quadrature rules.

Let $q > 0$ and consider a set $\mathfrak{X} \subset \mathbb{R}^d$ of cardinality

$$\#\mathfrak{X} = \dim(\mathcal{P}_{q^*}) = \binom{q^* + d}{d}$$

such that the Lagrange interpolation problem with respect to $\mathfrak{X}$ is poised in $\mathcal{P}_{q^*}$, i.e., for every function $f\colon \mathbb{R}^d \to \mathbb{R}$ there exists a unique polynomial

$$f_{\mathfrak{X}} \in \mathcal{P}_{q^*}$$

such that for all $x \in \mathfrak{X}$,

$$f(x) = f_{\mathfrak{X}}(x).$$

Let

$$\{p_x \colon x \in \mathfrak{X}\} \subset \mathcal{P}_{q^*}$$

denote the corresponding basis of $\mathcal{P}_{q^*}$ of Lagrange polynomials, i.e.,

$$f_{\mathfrak{X}} = \sum_{x \in \mathfrak{X}} f(x)\, p_x$$

for every function $f\colon \mathbb{R}^d \to \mathbb{R}$. Then

$$Q_{\mathfrak{X}} = \sum_{x \in \mathfrak{X}} \Big(\int p_x \,\mathrm{d}P \Big)\, \delta_x$$

is a signed measure on the Borel sets of $\mathbb{R}^d$, which satisfies

$$\int f_{\mathfrak{X}} \,\mathrm{d}P = \int f \,\mathrm{d}Q_{\mathfrak{X}} \tag{15}$$

for every function $f\colon \mathbb{R}^d \to \mathbb{R}$. In particular, $Q_{\mathfrak{X}}$ is exact of order q^*.

Consider a cutset $\mathcal{C} \subset \mathcal{T}$ and let $N \in \mathbb{N}$. Take independent random variables

$$\mathbf{J}_1, \dots, \mathbf{J}_N, X_1, \dots, X_N$$

such that for $k = 1, \dots, N$,

$$\mathbf{J}_k \sim U_{\mathcal{C}}, \; X_k \sim P,$$

and define a corresponding randomized composite quadrature rule $\widehat{I}_{\mathfrak{X},\mathcal{C},N}$ by

$$\widehat{I}_{\mathfrak{X},\mathcal{C},N}(f) = I_{Q^{\mathcal{C}}_{\mathfrak{X}}}(f) + \frac{\#\mathcal{C}}{N} \sum_{k=1}^{N} \rho_{\mathbf{J}_k} \big(f \circ S_{\mathbf{J}_k} - (f \circ S_{\mathbf{J}_k})_{\mathfrak{X}} \big)(X_k)$$

for $f \in \mathcal{H}^q$.

Remark 5. We briefly comment on the construction of $\widehat{I}_{\mathfrak{X},\mathcal{C},N}$. For $f \in \mathcal{H}^q$ we have

$$I(f) = \#\mathcal{C} \int_{\mathcal{C}} \int_{\mathbb{R}^d} \rho_{\mathbf{j}}\, (f \circ S_{\mathbf{j}})(x) \,\mathrm{d}P(x) \,\mathrm{d}U_{\mathcal{C}}(\mathbf{j}) \tag{16}$$

by Proposition 1. Furthermore,

$$I_{Q^{\mathcal{C}}_{\mathfrak{X}}}(f) = \#\mathcal{C} \int_{\mathcal{C}} \int_{\mathbb{R}^d} \rho_{\mathbf{j}}\, (f \circ S_{\mathbf{j}})_{\mathfrak{X}}(x) \,\mathrm{d}P(x) \,\mathrm{d}U_{\mathcal{C}}(\mathbf{j}), \tag{17}$$

due to (15), such that $\widehat{I}_{\mathfrak{X},\mathcal{C},N}$ is obtained by variance reduction based on the control variate

$$(\mathbf{j}, x) \mapsto \rho_{\mathbf{j}}\, (f \circ S_{\mathbf{j}})_{\mathfrak{X}}(x).$$

For the analysis of the mean squared error of $\widehat{I}_{\mathfrak{X},\mathcal{C},N}$ we use

$$\mathbf{e}(\mathfrak{X}) = \sup_{f\in\mathcal{H}_1^q} \left(\int (f(x) - f_{\mathfrak{X}}(x))^2 \,\mathrm{d}P \right)^{1/2}$$

to denote the worst case mean squared error of Lagrange interpolation in $\mathcal{P}_q$ based on the set of nodes $\mathfrak{X}$ for the unit ball $\mathcal{H}_1^q$ in $\mathcal{H}^q$ with respect to P.

Proposition 6. *For every cutset $\mathcal{C}$ and every $N \in \mathbb{N}$ the randomized quadrature rule $\widehat{I}_{\mathfrak{X},\mathcal{C},N}$ is unbiased and satisfies for $f \in \mathcal{H}^q$,*

$$\big(\mathbb{E}(I(f) - \widehat{I}_{\mathfrak{X},\mathcal{C},N}(f))^2\big)^{1/2} \leq \mathbf{e}(\mathfrak{X})\, \|f\|_{\mathcal{H}^q} \left(\frac{\#\mathcal{C}}{N} \sum_{\mathbf{j}\in\mathcal{C}} s_{\mathbf{j}}^2 \right)^{1/2}.$$

Proof. Clearly,

$$\mathbb{E}(\widehat{I}_{\mathfrak{X},\mathcal{C},N}(f)) = I_{Q_{\mathfrak{X}}^{\mathcal{C}}}(f) + \#\mathcal{C}\, \mathbb{E}\big(\rho_{\mathbf{J}_1}(f\circ S_{\mathbf{J}_1} - (f\circ S_{\mathbf{J}_1})_{\mathfrak{X}})(X_1)\big),$$

and by (16) and (17) we have

$$\#\mathcal{C}\, \mathbb{E}\big(\rho_{\mathbf{J}_1}(f\circ S_{\mathbf{J}_1})(X_1)\big) - \#\mathcal{C}\, \mathbb{E}\big(\rho_{\mathbf{J}_1}(f\circ S_{\mathbf{J}_1})_{\mathfrak{X}}\,(X_1)\big) = I(f) - I_{Q_{\mathfrak{X}}^{\mathcal{C}}}(f),$$

which shows the unbiasedness of $\widehat{I}_{\mathfrak{X},\mathcal{C},N}$.

It remains to estimate the variance of $\widehat{I}_{\mathfrak{X},\mathcal{C},N}$. Clearly,

$$\mathrm{Var}(\widehat{I}_{\mathfrak{X},\mathcal{C},N}(f)) \leq \frac{(\#\mathcal{C})^2}{N}\, \mathbb{E}\big(\rho_{\mathbf{J}_1}(f\circ S_{\mathbf{J}_1} - (f\circ S_{\mathbf{J}_1})_{\mathfrak{X}})(X_1)\big)^2.$$

Using the independence of $\mathbf{J}_1$ and X_1 as well as (6) we obtain

$$\begin{aligned}\mathbb{E}\big(\rho_{\mathbf{J}_1}(f\circ S_{\mathbf{J}_1} - (f\circ S_{\mathbf{J}_1})_{\mathfrak{X}})(X_1)\big)^2 &= \frac{1}{\#\mathcal{C}} \sum_{\mathbf{j}\in\mathcal{C}} \rho_{\mathbf{j}}^2\, \mathbb{E}\big((f\circ S_{\mathbf{j}} - (f\circ S_{\mathbf{j}})_{\mathfrak{X}})(X_1)\big)^2 \\ &\leq \frac{1}{\#\mathcal{C}} \sum_{\mathbf{j}\in\mathcal{C}} \rho_{\mathbf{j}}^2\, \|f\circ S_{\mathbf{j}}\|_{\mathcal{H}^q}^2\, \mathbf{e}^2(\mathfrak{X}) \leq \frac{\mathbf{e}^2(\mathfrak{X})}{\#\mathcal{C}}\, \|f\|_{\mathcal{H}^q}^2 \sum_{\mathbf{j}\in\mathcal{C}} \rho_{\mathbf{j}}^2\, r_{\mathbf{j}}^{2q},\end{aligned}$$

which completes the proof. □

Now we consider the particular choice of cutsets $\mathcal{C} = \mathcal{C}(T)$, see (7). Recall the definition (8) of the parameter $\theta \in (0,1)$.

Proposition 7. *Let $T \geq 1$ and $N \in \mathbb{N}$. The randomized quadrature rule $\widehat{I}_{\mathfrak{X},\mathcal{C}(T),N}$ uses at most*

$$\#\,\mathrm{supp}\big(Q_{\mathfrak{X}}^{\mathcal{C}(T)}\big) + N \leq \binom{q^*+d}{d} s_{min}^{-\theta}\, T^{\theta} + N$$

function evaluations and satisfies for every $f \in \mathcal{H}^q$,

$$\big(\mathbb{E}(I(f) - \widehat{I}_{\mathfrak{X},\mathcal{C}(T),N}(f))^2\big)^{1/2} \leq \mathbf{e}(\mathfrak{X})\, s_{min}^{-\theta}\, N^{-1/2}\, T^{-(1-\theta)}.$$

Proof. By Proposition 4,

$$\#\,\mathrm{supp}(Q_{\mathfrak{X}}^{\mathcal{C}(T)}) \leq \#\mathcal{C}(T)\,\mathrm{supp}(Q_{\mathfrak{X}}) \leq s_{\min}^{-\theta}\, T^{\theta}\, \#\mathfrak{X}.$$

Furthermore, by the definition of $\mathcal{C}(T)$ and by Proposition 4,

$$\#\mathcal{C}(T) \sum_{\mathbf{j}\in\mathcal{C}(T)} s_{\mathbf{j}}^2 < (\#\mathcal{C}(T))^2\, T^{-2} \leq s_{\min}^{-2\theta}\, T^{-(2-2\theta)}. \tag{18}$$

Now apply Proposition 6 to obtain the error estimate, which completes the proof. □

We adjust the parameter T to the number N of Monte Carlo repetitions by taking

$$T^{(N)} = \max\left(1, s_{\min} \binom{q^* + d}{d}^{-1/\theta} N^{1/\theta}\right)$$

and we put

$$\widehat{I}_{\mathfrak{X}}^{(n)} = \widehat{I}_{\mathfrak{X},\mathcal{C}(T^{(n/2)}),\lfloor n/2 \rfloor}$$

for every $n \in \mathbb{N}$ with $n \geq 2$.

Using Proposition 7 we immediately obtain the following result.

Theorem 2. *For every $n \in \mathbb{N}$ with $n \geq 2s_{min}^{-\theta}\binom{q^*+d}{d}$ the randomized quadrature rule $\widehat{I}_{\mathfrak{X}}^{(n)}$ uses at most n function evaluations and satisfies for every $f \in \mathcal{H}^q$,*

$$\left(\mathbb{E}(I(f) - \widehat{I}_{\mathfrak{X}}^{(n)}(f))^2\right)^{1/2} \leq 2^{q/\beta+1/2}\binom{q^* + d}{d}^{q/\beta} \mathbf{e}(\mathfrak{X})\, s_{min}^{-1}\, \|f\|_{\mathcal{H}^q}\, n^{-(q/\beta+1/2)}.$$

Remark 6. Due to Theorem 2 the worst case errors of $\widehat{I}_{\mathfrak{X}}^{(n)}$ on the unit balls $\mathcal{G}_1^q$ and $\mathcal{F}_1^q$ converge to zero as $n \to \infty$ with order at least $q/\beta + 1/2$ in terms of the number of function evaluations. In Section 6 we show that this order of convergence is optimal within the class of all randomized algorithms based on finitely many function evaluations if the contractions $S_1, \dots, S_m$ are similarities and satisfy the open set condition, see Proposition 9 and Theorem 4. As for the classical quadrature problem on the d-dimensional unit cube, we gain a power of $1/2$ compared with the best possible order of convergence q/β for deterministic methods.

Remark 7. We provide an estimate of the quantity $\mathbf{e}(\mathfrak{X})$ in terms of the Lagrange polynomials p_x, $x \in \mathfrak{X}$, and the effective radius of $P + |Q_{\mathfrak{X}}|$.

Put $R = \mathrm{rad}(P + |Q_{\mathfrak{X}}|)$ and choose $x_0 \in \mathbb{R}^d$ such that $\mathrm{supp}(P + |Q_{\mathfrak{X}}|) \subset \overline{B}(x_0, R)$. Let $f \in \mathcal{H}_1^q$ and consider the q^*-th order Taylor-polynomial p of f at x_0. Since $p \in \mathcal{P}_{q^*}$ we have $p_{\mathfrak{X}} = p$ and therefore

$$|f - f_{\mathfrak{X}}| \leq |f - p| + |(f - p)_{\mathfrak{X}}| \leq |f - p| + \max_{x \in \mathfrak{X}} |f(x) - p(x)| \sum_{x \in \mathfrak{X}} |p_x|.$$

From the proof of Proposition 2 we know that

$$\sup_{z \in \overline{B}(x_0,R)} |f(z) - p(z)| \leq \frac{R^q}{q!},$$

and therefore we have

$$\mathbf{e}(\mathfrak{X}) \leq \left(1 + \sup_{z \in \mathrm{supp}(P)} \sum_{x \in \mathfrak{X}} |p_x(z)|\right) \frac{(\mathrm{rad}(P + |Q_{\mathfrak{X}}|))^q}{q!}. \tag{19}$$

To implement the randomized quadrature rule $\widehat{I}_{\mathfrak{X}}^{(n)}$ requires an algorithm for sampling from the self-similar distribution P. We present an alternative method, which overcomes this disadvantage and has the same level of accuracy as $\widehat{I}_{\mathfrak{X}}^{(n)}$.

For convenience we assume that $Q_{\mathfrak{X}}$ is a probability measure. Let $T_2 \geq T_1 \geq 1$ and consider independent random variables

$$Z, \mathbf{J}, \bar{J}_1, \bar{J}_2, \dots$$

on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$ such that

$$Z \sim Q_{\mathfrak{X}}, \quad \mathbf{J} \sim U_{\mathcal{C}(T_1)},$$

and for $k \in \mathbb{N}$ and $j \in \{1, \dots, m\}$,

$$\mathbb{P}(\bar{J}_k = j) = \rho_j. \tag{20}$$

Put
$$\bar{\mathbf{J}}_k = \begin{cases} (\bar{J}_1, \dots, \bar{J}_k), & \text{if } k \in \mathbb{N}, \\ \lambda, & \text{if } k = 0, \end{cases}$$
and define
$$\tau = \min\{k \in \mathbb{N}_0 \colon s_{\mathbf{J}}\, s_{\bar{\mathbf{J}}_k} < 1/T_2\}.$$

We briefly discuss the law of the random vertex $(\mathbf{J}, \bar{\mathbf{J}}_\tau)$. Note that $s_{\mathbf{j}} T_2 \geq 1$ for $\mathbf{j} \in \mathcal{C}(T_1) \backslash \mathcal{C}(T_2)$ and put
$$\mathcal{D}_{\mathbf{j}} = \begin{cases} \mathcal{C}(s_{\mathbf{j}} T_2), & \text{if } \mathbf{j} \in \mathcal{C}(T_1) \backslash \mathcal{C}(T_2), \\ \{\lambda\}, & \text{if } \mathbf{j} \in \mathcal{C}(T_1) \cap \mathcal{C}(T_2). \end{cases}$$

Lemma 1. *The sets of vertices $\{(\mathbf{j}, \bar{\mathbf{j}}) \colon \bar{\mathbf{j}} \in \mathcal{D}_{\mathbf{j}}\}$, $\mathbf{j} \in \mathcal{C}(T_1)$, are pairwise disjoint with*
$$\mathcal{C}(T_2) = \bigcup_{\mathbf{j} \in \mathcal{C}(T_1)} \{(\mathbf{j}, \bar{\mathbf{j}}) \colon \bar{\mathbf{j}} \in \mathcal{D}_{\mathbf{j}}\}.$$
For all $\mathbf{j} \in \mathcal{C}(T_1)$ and $\bar{\mathbf{j}} \in \mathcal{D}_{\mathbf{j}}$,
$$\mathbb{P}((\mathbf{J}, \bar{\mathbf{J}}_\tau) = (\mathbf{j}, \bar{\mathbf{j}})) = \frac{\rho_{\bar{\mathbf{j}}}}{\#\mathcal{C}(T_1)}.$$

Proof. For $\mathbf{j} \in \mathcal{C}(T_1)$ we put
$$G_{\mathbf{j}} = \{(\mathbf{j}, \bar{\mathbf{j}}) \colon \bar{\mathbf{j}} \in \mathcal{D}_{\mathbf{j}}\}.$$

Pairwise disjointness of the sets $G_{\mathbf{j}}$ follows immediately from the fact that $\mathcal{C}(T_1)$ is a cutset.

Let $\mathbf{j} \in \mathcal{C}(T_1)$ and $\bar{\mathbf{j}} \in \mathcal{D}_{\mathbf{j}}$. If $\mathbf{j} \in \mathcal{C}(T_2)$ then $(\mathbf{j}, \bar{\mathbf{j}}) = \mathbf{j} \in \mathcal{C}(T_2)$. Otherwise $\bar{\mathbf{j}} \in \mathcal{C}(s_{\mathbf{j}} T_2)$, and we have $s_{(\mathbf{j}, \bar{\mathbf{j}})} = s_{\mathbf{j}}\, s_{\bar{\mathbf{j}}} < s_{\mathbf{j}}/(s_{\mathbf{j}} T_2) = 1/T_2$ as well as $s_{(\mathbf{j}, \bar{\mathbf{j}})-} = s_{\mathbf{j}}\, s_{\bar{\mathbf{j}}-} \geq s_{\mathbf{j}}/(s_{\mathbf{j}} T_2) = 1/T_2$, which yields $(\mathbf{j}, \bar{\mathbf{j}}) \in \mathcal{C}(T_2)$. Thus, $G_{\mathbf{j}} \subset \mathcal{C}(T_2)$.

Let $\ell \in \mathbb{N}$ and $(j_1, \dots, j_\ell) \in \mathcal{C}(T_2)$. Put $i^* = \min\{i \in \{1, \dots, \ell\} \colon s_{(j_1, \dots, j_i)} < 1/T_1\}$ and let
$$\mathbf{j} = (j_1, \dots, j_{i^*}), \; \bar{\mathbf{j}} = (j_{i^*+1}, \dots, j_\ell).$$
Clearly, $(j_1, \dots, j_\ell) = (\mathbf{j}, \bar{\mathbf{j}})$ and $s_{\mathbf{j}} \in \mathcal{C}(T_1)$. Assume $s_{\mathbf{j}} \in \mathcal{C}(T_2)$. Then $i^* = \ell$, and therefore $\bar{\mathbf{j}} = \lambda$, since otherwise $1/T_2 > s_{\mathbf{j}} \geq s_{(j_1, \dots, j_\ell)-} \geq 1/T_2$. On the other hand, if $s_{\mathbf{j}} \notin \mathcal{C}(T_2)$ then $i^* < \ell$ and we have $s_{\bar{\mathbf{j}}} = s_{(j_1, \dots, j_\ell)}/s_{\mathbf{j}} < 1/(s_{\mathbf{j}} T_2)$ as well as $s_{\bar{\mathbf{j}}-} = s_{(j_1, \dots, j_\ell)-}/s_{\mathbf{j}} \geq 1/(s_{\mathbf{j}} T_2)$, which yields $s_{\bar{\mathbf{j}}} \in \mathcal{C}(s_{\mathbf{j}} T_2)$. Hence $(j_1, \dots, j_\ell) \in G_{\mathbf{j}}$, which completes the proof of the first statement in the lemma.

Let $\mathbf{j} \in \mathcal{C}(T_1)$ and $\bar{\mathbf{j}} \in \mathcal{D}_{\mathbf{j}}$, and put $\ell = |\bar{\mathbf{j}}| \in \mathbb{N}_0$. It is easy to see that
$$\{\mathbf{J} = \mathbf{j}, \bar{\mathbf{J}}_\tau = \bar{\mathbf{j}}\} = \{\mathbf{J} = \mathbf{j}, \bar{\mathbf{J}}_\ell = \bar{\mathbf{j}}\}.$$
Furthermore, $\mathbf{J}$ and $\bar{\mathbf{J}}_\ell$ are independent and $\mathbb{P}(\bar{\mathbf{J}}_\ell = \bar{\mathbf{j}}) = \rho_{\bar{\mathbf{j}}}$, due to (20), which yields the second statement in the lemma and completes the proof. □

Take N independent copies
$$(Z_1, \mathbf{J}_1, \bar{\mathbf{J}}_{\tau,1}), \dots, (Z_N, \mathbf{J}_N, \bar{\mathbf{J}}_{\tau,N})$$
of $(Z, \mathbf{J}, \bar{\mathbf{J}}_\tau)$ and define a corresponding randomized composite quadrature formula by
$$\bar{I}_{\mathfrak{X}, T_1, T_2, N}(f) = I_{Q_{\mathfrak{X}}^{\mathcal{C}(T_1)}}(f) + \frac{\#\mathcal{C}(T_1)}{N} \sum_{k=1}^{N} \rho_{\mathbf{J}_k} \left(f \circ S_{\mathbf{J}_k} - (f \circ S_{\mathbf{J}_k})_{\mathfrak{X}}\right)(S_{\bar{\mathbf{J}}_{\tau,k}} \circ Z_k).$$

For the error analysis of $\bar{I}_{\mathfrak{X}, T_1, T_2, N}$ we put
$$\Lambda(\mathfrak{X}) = \sup\Big\{\sum_{x \in \mathfrak{X}} |p_x(z)| \colon |z|_2 \leq \operatorname{rad}(P + |Q_{\mathfrak{X}}|)\Big\}.$$

Lemma 2. *Assume that $Q_{\mathfrak{X}}$ is a probability measure. Let $T_2 \geq T_1 \geq 1$ and $N \in \mathbb{N}$. The randomized quadrature rule $\bar{I}_{\mathfrak{X},T_1,T_2,N}$ satisfies for every $f \in \mathcal{H}^q$,*

$$\mathbb{E}\big(\bar{I}_{\mathfrak{X},T_1,T_2,N}(f)\big) = I_{Q_{\mathfrak{X}}^{\mathcal{C}(T_2)}}(f)$$

and

$$Var\big(\bar{I}_{\mathfrak{X},T_1,T_2,N}(f)\big) \leq (1+\Lambda(\mathfrak{X}))^2 \, \frac{(3\,\mathrm{rad}(P+|Q_{\mathfrak{X}}|))^{2q}}{(q!)^2} \, s_{\min}^{-2\theta} \, \|f\|_{\mathcal{H}^q}^2 \, N^{-1} \, T_1^{-2(1-\theta)}.$$

Proof. Let $f \in \mathcal{H}^q$. Clearly,

$$\mathbb{E}\big(\bar{I}_{\mathfrak{X},T_1,T_2,N}(f)\big) = I_{Q_{\mathfrak{X}}^{\mathcal{C}(T_1)}}(f) + \#\mathcal{C}(T_1)\,\mathbb{E}\big(\rho_{\mathbf{J}}\,(f\circ S_{\mathbf{J}} - (f\circ S_{\mathbf{J}})_{\mathfrak{X}})(S_{\bar{\mathbf{J}}_\tau}\circ Z)\big).$$

By the independence of Z and $(\mathbf{J},\bar{\mathbf{J}}_\tau)$ and by Lemma 1,

$$\mathbb{E}\big(\rho_{\mathbf{J}}\,(f\circ S_{\mathbf{J}} - (f\circ S_{\mathbf{J}})_{\mathfrak{X}})(S_{\bar{\mathbf{J}}_\tau}\circ Z)\big) = \sum_{\mathbf{j}\in\mathcal{C}(T_1)}\sum_{\bar{\mathbf{j}}\in\mathcal{D}_{\mathbf{j}}} \frac{\rho_{\mathbf{j}}\,\rho_{\bar{\mathbf{j}}}}{\#\mathcal{C}(T_1)}\,\mathbb{E}\big((f\circ S_{\mathbf{j}} - (f\circ S_{\mathbf{j}})_{\mathfrak{X}})(S_{\bar{\mathbf{j}}}\circ Z)\big).$$

Moreover,

$$\sum_{\mathbf{j}\in\mathcal{C}(T_1)}\sum_{\bar{\mathbf{j}}\in\mathcal{D}_{\mathbf{j}}} \rho_{\mathbf{j}}\,\rho_{\bar{\mathbf{j}}}\,\mathbb{E}(f\circ S_{\mathbf{j}}\circ S_{\bar{\mathbf{j}}}\circ Z) = \sum_{\mathbf{j}\in\mathcal{C}(T_2)} \rho_{\mathbf{j}} \int f\circ S_{\mathbf{j}}\,\mathrm{d}Q_{\mathfrak{X}} = \int f\,\mathrm{d}Q_{\mathfrak{X}}^{\mathcal{C}(T_2)} = I_{Q_{\mathfrak{X}}^{\mathcal{C}(T_2)}}(f)$$

and

$$\sum_{\mathbf{j}\in\mathcal{C}(T_1)}\sum_{\bar{\mathbf{j}}\in\mathcal{D}_{\mathbf{j}}} \rho_{\mathbf{j}}\,\rho_{\bar{\mathbf{j}}}\,\mathbb{E}((f\circ S_{\mathbf{j}})_{\mathfrak{X}}(S_{\bar{\mathbf{j}}}\circ Z)) = \sum_{x\in\mathfrak{X}}\sum_{\mathbf{j}\in\mathcal{C}(T_1)} \rho_{\mathbf{j}}\,f(S_{\mathbf{j}}(x)) \sum_{\bar{\mathbf{j}}\in\mathcal{D}_{\mathbf{j}}} \rho_{\bar{\mathbf{j}}}\,\mathbb{E}(p_x\circ S_{\bar{\mathbf{j}}}(Z)).$$

Since $S_{\bar{\mathbf{j}}}$ is affine we have $p_x\circ S_{\bar{\mathbf{j}}} \in \mathcal{P}_{q^*}$ for all $\bar{\mathbf{j}} \in \mathcal{D}_{\mathbf{j}}$. Thus, by Proposition 1,

$$\sum_{\bar{\mathbf{j}}\in\mathcal{D}_{\mathbf{j}}} \rho_{\bar{\mathbf{j}}}\cdot\mathbb{E}(p_x\circ S_{\bar{\mathbf{j}}}(Z)) = \sum_{\bar{\mathbf{j}}\in\mathcal{D}_{\mathbf{j}}} \rho_{\bar{\mathbf{j}}} \int p_x\circ S_{\bar{\mathbf{j}}}\,\mathrm{d}Q_{\mathfrak{X}} = \sum_{\bar{\mathbf{j}}\in\mathcal{D}_{\mathbf{j}}} \rho_{\bar{\mathbf{j}}} \int p_x\circ S_{\bar{\mathbf{j}}}\,\mathrm{d}P = \int p_x\,\mathrm{d}P$$

for every $\mathbf{j} \in \mathcal{C}(T_1)$, and we conclude that

$$\begin{aligned}\sum_{\mathbf{j}\in\mathcal{C}(T_1)}\sum_{\bar{\mathbf{j}}\in\mathcal{D}_{\mathbf{j}}} \rho_{\mathbf{j}}\,\rho_{\bar{\mathbf{j}}}\,\mathbb{E}((f\circ S_{\mathbf{j}})_{\mathfrak{X}}(S_{\bar{\mathbf{j}}}\circ Z)) &= \sum_{\mathbf{j}\in\mathcal{C}(T_1)} \rho_{\mathbf{j}} \sum_{x\in\mathfrak{X}} f(S_{\mathbf{j}}(x)) \int p_x\,\mathrm{d}P \\ &= \sum_{\mathbf{j}\in\mathcal{C}(T_1)} \rho_{\mathbf{j}} \int f\circ S_{\mathbf{j}}\,\mathrm{d}Q_{\mathfrak{X}} = I_{Q_{\mathfrak{X}}^{\mathcal{C}(T_1)}}(f),\end{aligned}$$

which finishes the proof of the first statement in the lemma.

We proceed with the proof of the variance estimate. By Lemma 1,

$$\begin{aligned}\mathrm{Var}\big(\bar{I}_{\mathfrak{X},T_1,T_2,N}(f)\big) &\leq \frac{(\#\mathcal{C}(T_1))^2}{N}\,\mathbb{E}\big(\rho_{\mathbf{J}}\,(f\circ S_{\mathbf{J}} - (f\circ S_{\mathbf{J}})_{\mathfrak{X}})(S_{\bar{\mathbf{J}}_\tau}\circ Z)\big)^2 \\ &= \frac{\#\mathcal{C}(T_1)}{N} \sum_{\mathbf{j}\in\mathcal{C}(T_1)} \rho_{\mathbf{j}}^2 \sum_{\bar{\mathbf{j}}\in\mathcal{D}_{\mathbf{j}}} \rho_{\bar{\mathbf{j}}}\,\mathbb{E}\big((f\circ S_{\mathbf{j}} - (f\circ S_{\mathbf{j}})_{\mathfrak{X}})^2(S_{\bar{\mathbf{j}}}(Z))\big) \\ &\leq \frac{\#\mathcal{C}(T_1)}{N} \sum_{\mathbf{j}\in\mathcal{C}(T_1)} \rho_{\mathbf{j}}^2\,\|f\circ S_{\mathbf{j}}\|_{\mathcal{H}^q}^2 \sum_{\bar{\mathbf{j}}\in\mathcal{D}_{\mathbf{j}}} \rho_{\bar{\mathbf{j}}} \sup_{g\in\mathcal{H}_1^q} \int (g-g_{\mathfrak{X}})^2(S_{\bar{\mathbf{j}}})\,\mathrm{d}Q_{\mathfrak{X}}. \qquad (21)\end{aligned}$$

Put $R = \mathrm{rad}(P + |Q_{\mathfrak{X}}|)$ and choose $x_0 \in \mathbb{R}^d$ such that $\mathrm{supp}(P + |Q_{\mathfrak{X}}|) \subset \overline{B}(x_0, R)$. Fix $\mathbf{j} \in \mathcal{C}(T_1)$ as well as $\bar{\mathbf{j}} \in \mathcal{D}_{\mathbf{j}}$ and let $g \in \mathcal{H}^q_1$. Consider the q^*-th order Taylor-polynomial p of g at x_0 and let $x \in \mathfrak{X}$. Similar to the proofs of Proposition 2 and (19),

$$\begin{aligned}|g(S_{\bar{\mathbf{j}}}(x)) - g_{\mathfrak{X}}(S_{\bar{\mathbf{j}}}(x))| &\leq |g(S_{\bar{\mathbf{j}}}(x)) - p(S_{\bar{\mathbf{j}}}(x))| + \sup_{z\in\mathfrak{X}} |g(z) - p(z)| \sum_{z\in\mathfrak{X}} |p_z(S_{\bar{\mathbf{j}}}(x))| \\ &\leq \frac{1}{q!}\Big(\|S_{\bar{\mathbf{j}}}(x) - x_0\|^q + R^q \sum_{z\in\mathfrak{X}} |p_z(S_{\bar{\mathbf{j}}}(x))|\Big).\end{aligned}$$

If $\bar{\mathbf{j}} = \lambda$ then

$$\|S_{\bar{\mathbf{j}}}(x) - x_0\| = \|x - x_0\| \leq R.$$

Next, assume $\bar{\mathbf{j}} \neq \lambda$. Then $S_{\bar{\mathbf{j}}}$ has a fixed point $x^* \in \mathbb{R}^d$ and we have $x^* \in \mathrm{supp}(P)$, see [Hut81]. Hence

$$\|S_{\bar{\mathbf{j}}}(x) - x_0\| \leq \|S_{\bar{\mathbf{j}}}(x) - S_{\bar{\mathbf{j}}}(x^*)\| + \|x^* - x_0\| \leq r_{\bar{\mathbf{j}}}\, \|x - x^*\| + R \leq 3R.$$

It follows

$$\|S_{\bar{\mathbf{j}}}(x) - x_0\|^q + R^q \sum_{z\in\mathfrak{X}} |p_z(S_{\bar{\mathbf{j}}}(x))| \ \leq (3R)^q (1 + \Lambda(\mathfrak{X})),$$

which yields

$$\int (g - g_{\mathfrak{X}})^2 (S_{\bar{\mathbf{j}}})\, \mathrm{d}Q_{\mathfrak{X}} \leq \frac{(3R)^{2q}}{(q!)^2} (1 + \Lambda(\mathfrak{X}))^2. \tag{22}$$

Combine (21) with (22) and use (6) to conclude that

$$\mathrm{Var}\big(\bar{I}_{\mathfrak{X},T_1,T_2,N}(f)\big) \leq \frac{\#\mathcal{C}(T_1)}{N} \frac{(3R)^{2q}}{(q!)^2} (1 + \Lambda(\mathfrak{X}))^2\, \|f\|^2_{\mathcal{H}^q} \sum_{\mathbf{j}\in\mathcal{C}(T_1)} s^2_{\mathbf{j}} \sum_{\bar{\mathbf{j}}\in\mathcal{D}_{\mathbf{j}}} \rho_{\bar{\mathbf{j}}}.$$

By Proposition 1 and (18)

$$\#\mathcal{C}(T_1) \sum_{\mathbf{j}\in\mathcal{C}(T_1)} s^2_{\mathbf{j}} \sum_{\bar{\mathbf{j}}\in\mathcal{D}_{\mathbf{j}}} \rho_{\bar{\mathbf{j}}} = \#\mathcal{C}(T_1) \sum_{\mathbf{j}\in\mathcal{C}(T_1)} s^2_{\mathbf{j}} \leq s_{\min}^{-2\theta}\, T_1^{-(2-2\theta)},$$

which finishes the proof of the lemma. □

As a direct consequence of Lemma 2 and Proposition 5 we obtain the following estimate of the error of $\bar{I}_{\mathfrak{X},T_1,T_2,N}$. For the number of function evaluations of this method we obviously have the same upper bound as for the method $\widehat{I}_{\mathfrak{X},\mathcal{C}(T_1),N}$, see Proposition 7.

Proposition 8. *Assume that $Q_{\mathfrak{X}}$ is a probability measure. Let $T_2 \geq T_1 \geq 1$ and $N \in \mathbb{N}$. The randomized quadrature rule $\bar{I}_{\mathfrak{X},T_1,T_2,N}$ uses at most*

$$\#\,\mathrm{supp}\big(Q_{\mathfrak{X}}^{\mathcal{C}(T_1)}\big) + N \leq \binom{q^* + d}{d} s_{min}^{-\theta}\, T_1^{\theta} + N$$

function evaluations and satisfies for every $f \in \mathcal{H}^q$,

$$\begin{aligned}&\big(\mathbb{E}(I(f) - \bar{I}_{\mathfrak{X},T_1,T_2,N}(f))^2\big)^{1/2} \\ &\qquad \leq (1 + \Lambda(\mathfrak{X}))\, \frac{(3\,\mathrm{rad}(P + |Q_{\mathfrak{X}}|))^q}{q!}\, s_{\min}^{-\theta}\, \|f\|_{\mathcal{H}^q} \big(T_2^{-(1-\theta)} + N^{-1/2}\, T_1^{-(1-\theta)}\big).\end{aligned}$$

We adjust the parameters T_1 and T_2 to the number N of Monte Carlo repetitions by taking

$$
\begin{aligned}
T_1^{(N)} &= \max\left(1, s_{\min}\binom{q^*+d}{d}^{-1/\theta} N^{1/\theta}\right),\\
T_2^{(N)} &= N^{1/(2(1-\theta))}\, T_1^{(N)},
\end{aligned}
\tag{23}
$$

and we put

$$
\bar{I}_{\mathfrak{X}}^{(n)} = \bar{I}_{\mathfrak{X}, T_1^{(n/2)}, T_2^{(n/2)}, \lfloor n/2 \rfloor}
$$

for all $n \in \mathbb{N}$ with $n \geq 2$.

The following result immediately follows from Proposition 8.

Theorem 3. *Assume that $Q_{\mathfrak{X}}$ is a probability measure. For every $n \in \mathbb{N}$ with $n \geq 2s_{min}^{-\theta}\binom{q^*+d}{d}$ the randomized quadrature rule $\bar{I}_{\mathfrak{X}}^{(n)}$ uses at most n function evaluations and satisfies for every $f \in \mathcal{H}^q$,*

$$
\left(\mathbb{E}(I(f) - \bar{I}_{\mathfrak{X}}^{(n)}(f))^2\right)^{1/2} \leq (1 + \Lambda(\mathfrak{X}))\, \frac{(3\,\mathrm{rad}(P + |Q_{\mathfrak{X}}|))^q}{q!}\, s_{min}^{-1}\, 2^{q/\beta+1}\, \|f\|_{\mathcal{H}^q}\, n^{-(q/\beta+1/2)}.
$$

Remark 8. We briefly discuss the computational cost that is needed to compute a realization of $\bar{I}_{\mathfrak{X}}^{(n)}(f)$ for $f \in \mathcal{H}^q$. Consider, more generally, the randomized quadrature rule $\bar{I}_{\mathfrak{X},T_1,T_2,N}$ and define

$$
\mathrm{cost}^{\mathrm{comp}}(\bar{I}_{\mathfrak{X},T_1,T_2,N}) = \sup_{f \in \mathcal{H}_1^q} \mathbb{E}\, \mathrm{cost}^{\mathrm{comp}}(\bar{I}_{\mathfrak{X},T_1,T_2,N}, f),
$$

where $\mathrm{cost}^{\mathrm{comp}}(\bar{I}_{\mathfrak{X},T_1,T_2,N}, f)$ is given by the (random) sum of

1) the number n_1 of calls to a random number generator for the uniform distribution on $\mathfrak{X}$ or $\{1, \ldots, m\}$,
2) the number n_2 of evaluations of f at points in $\mathbb{R}^d$, and
3) the number n_3 of basic arithmetic operations (summation, subtraction, multiplication, division)

that are needed to compute a realization of $\bar{I}_{\mathfrak{X},T_1,T_2,N}(f)$.

Since

$$
\mathcal{C}(T) \subset \{1, \ldots, m\}^{\lceil \ln(T)/\ln(s_{\max}^{-1}) \rceil}
$$

for any $T \geq 1$, it is clear that the random vector $(Z, \mathbf{J}, \bar{\mathbf{J}}_\tau)$ can be simulated with at most $c_1 \ln(T_2)$ calls to a random number generator for the uniform distribution on $\mathfrak{X}$ or $\{1, \ldots, m\}$, where the constant $c_1 > 0$ only depends on $\#\mathfrak{X}$ and m. Hence $n_1 \leq c_1 N \ln(T_2)$. By Proposition 8 we have $n_2 \leq c_2(\#C(T_1) + N)$ where the constant $c_2 > 0$ only depends on $\#\mathfrak{X}$. Finally, it is easy to see that $n_3 \leq c_3 N \ln(T_2)$, where the constant c_3 only depends on d and $\#\mathfrak{X}$. Consequently,

$$
\mathrm{cost}^{\mathrm{comp}}(\bar{I}_{\mathfrak{X},T_1,T_2,N}, f) \leq c\,(\#\mathcal{C}(T_1) + N\, \ln(T_2))
$$

for every input $f \in \mathcal{H}^q$, where the constant $c > 0$ neither depends on T_1 nor on T_2 nor on N, and therefore

$$
\mathrm{cost}^{\mathrm{comp}}(\bar{I}_{\mathfrak{X}}^{(n)}) \leq c\, \ln(n)\, n,
$$

where the constant $c > 0$ does not depend on n.

6. Lower bounds and optimality

We study generalized deterministic and randomised algorithms for the quadrature problem given by (1) that are based on finitely many evaluations of an integrand $f \in \mathcal{G}^q$ at points in $\mathbb{R}^d$. Our goal is to provide sharp lower bounds for the worst case mean squared error on the unit ball $\mathcal{G}_1^q$ of any such algorithm in terms of its worst case average number of function evaluations.

A generalized randomised algorithm is specified by a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and a triple

$$(\psi, \nu, \varphi),$$

where

- $\psi = (\psi_k)_{k \geq 1}$ is a sequence of mappings
$$\psi_k \colon \mathbb{R}^{k-1} \times \Omega \to \mathbb{R}^d,$$
which are used to sequentially determine random evaluation nodes in $\mathbb{R}^d$ for a given integrand $f \in \mathcal{G}^q$,
- the mapping
$$\nu \colon \mathcal{G}^q \times \Omega \to \mathbb{N}$$
determines the random total number of evaluations of f, and
- $\varphi = (\varphi_k)_{k \geq 1}$ is a sequence of mappings
$$\varphi_k \colon \mathbb{R}^k \times \Omega \to \mathbb{R},$$
which are used to obtain the random approximation to $I(f)$ based on the observed function values of f.

To be more precise, we define

$$N_k^\psi \colon \mathcal{G}^q \times \Omega \to \mathbb{R}^k$$

for $k \in \mathbb{N}$ by

$$N_k^\psi(f, \omega) = (y_1(f, \omega), \dots, y_k(f, \omega)),$$

where

$$y_1(f, \omega) = f(\psi_1(\omega))$$

and

$$y_\ell(f, \omega) = f\big(\psi_\ell(y_1(f, \omega), \dots, y_{\ell-1}(f, \omega), \omega)\big), \quad \ell = 2, \dots, k.$$

For given $\omega \in \Omega$ and input $f \in \mathcal{G}^q$ the algorithm sequentially performs $\nu(f, \omega)$ evaluations of f at the points

$$\psi_1(\omega), \psi_2(y_1(f, \omega)), \dots, \psi_{\nu(f,\omega)}(y_1(f, \omega), \dots, y_{\nu(f,\omega)-1}(f, \omega)) \in \mathbb{R}^d$$

and finally applies the mapping $\varphi_{\nu(f,\omega)}(\cdot, \omega) \colon \mathbb{R}^{\nu(f,\omega)} \to \mathbb{R}$ to the data $N^\psi_{\nu(f,\omega)}(f, \omega)$ to obtain the real number

$$\widehat{I}_{\psi,\nu,\varphi}(f, \omega) = \varphi_{\nu(f,\omega)}\big(N^\psi_{\nu(f,\omega)}(f, \omega), \omega\big),$$

as an approximation to $I(f)$. The induced mapping

$$\widehat{I}_{\psi,\nu,\varphi} \colon \mathcal{G}^q \times \Omega \to \mathbb{R}$$

is called a generalized randomized algorithm if the mappings

$$\widehat{I}_{\psi,\nu,\varphi}(f, \cdot) \colon \Omega \to \mathbb{R} \ \text{ and } \ \nu(f, \cdot) \colon \Omega \to \mathbb{N},$$

are random variables for all $f \in \mathcal{G}^q$.

We use $\mathcal{A}_q^{\mathrm{ran}}$ to denote the class of all generalized randomized algorithms. The error and the cost of $\widehat{I} \in \mathcal{A}_q^{\mathrm{ran}}$ are defined by

$$e(\widehat{I}) = \sup_{f \in \mathcal{G}_1^q} \left(\mathbb{E}|I(f) - \widehat{I}(f,\cdot)|^2 \right)^{1/2}$$

and

$$\mathrm{cost}(\widehat{I}) = \inf \left\{ \sup_{f \in \mathcal{G}_1^q} \mathbb{E}(\nu(f,\cdot)) \colon \widehat{I} = \widehat{I}_{\psi,\nu,\varphi} \right\},$$

respectively. Note that the definition of the cost of $\widehat{I}$ takes into account that the representation $\widehat{I} = \widehat{I}_{\psi,\nu,\varphi}$ is not unique in general.

A generalized randomized algorithm $\widehat{I} \in \mathcal{A}_q^{\mathrm{ran}}$ is called deterministic if the mapping $\widehat{I}(f,\cdot)$ is constant for all $f \in \mathcal{G}^q$. In this case we have $\widehat{I} = \widehat{I}_{\psi,\nu,\varphi}$ with mappings

$$\psi_k \colon \mathbb{R}^{k-1} \to \mathbb{R}^d,\ \nu \colon \mathcal{H}^q \to \mathbb{N},\ \varphi_k \colon \mathbb{R}^k \to \mathbb{R}, \tag{24}$$

and it is easy to see that

$$\mathrm{cost}(\widehat{I}) = \inf \{ \sup_{f \in \mathcal{G}_1^q} \nu(f) \colon \widehat{I} = \widehat{I}_{\psi,\nu,\varphi} \},$$

where the infimum extends over all triples (ψ,ν,φ) that satisfy (24). The class of all generalized deterministic algorithms is denoted by $\mathcal{A}_q^{\mathrm{det}}$.

Note that the deterministic quadrature rules introduced in Section 4 and the randomized quadrature rules introduced in Section 5 belong to $\mathcal{A}_q^{\mathrm{det}}$ and $\mathcal{A}_q^{\mathrm{ran}}$, respectively.

Let $n \in \mathbb{N}$. The crucial quantities for our analysis are the n-th minimal errors

$$e_n^{\mathrm{det}}(\mathcal{G}_1^q) = \inf\{ e(\widehat{I}) \colon \widehat{I} \in \mathcal{A}_q^{\mathrm{det}},\ \mathrm{cost}(\widehat{I}) \le n \}$$

and

$$e_n^{\mathrm{ran}}(\mathcal{G}_1^q) = \inf\{ e(\widehat{I}) \colon \widehat{I} \in \mathcal{A}_q^{\mathrm{ran}},\ \mathrm{cost}(\widehat{I}) \le n \},$$

i.e., the smallest possible error that can be achieved by generalized deterministic algorithms based on at most n function evaluations and the smallest possible error that can be achieved by generalized randomized algorithms that use at most n function evaluations on average, respectively.

By Theorem 1, Theorem 2 and Remark 4 we have

$$e_n^{\mathrm{det}}(\mathcal{G}_1^q) \le c\, n^{-q/\beta}, \quad e_n^{\mathrm{ran}}(\mathcal{G}_1^q) \le c\, n^{-(q/\beta + 1/2)}$$

for n sufficiently large, where the constant $c > 0$ does not depend on n. We show that both bounds are sharp if the following two conditions are satisfied.

(S1) The contractions $S_1, \dots, S_m$ are similarities, i.e.,

$$S_j(x) = r_j V_j\, x + b_j, \quad x \in \mathbb{R}^d,$$

where $r_j \in (0,1)$, $b_j \in \mathbb{R}^d$ and $V_j \in \mathbb{R}^{d \times d}$ satisfies

$$\|V_j x\| = r_j\, \|x\|$$

for all $x \in \mathbb{R}^d$.

(S2) The contractions $S_1, \dots, S_m$ satisfy the open set condition, i.e.,

$$\exists\, \emptyset \neq O \subset \mathbb{R}^d, \text{open} \colon S_1(O), \dots, S_m(O) \text{ are pairwise disjoint and } \bigcup_{i=1}^m S_i(O) \subset O.$$

Proposition 9. *Assume that the contractions $S_1, \ldots, S_m$ are similarities and satisfy the open set condition. Then there exists $c > 0$ such that for every $n \in \mathbb{N}$,*

$$e_n^{det}(\mathcal{G}_1^q) \geq c\, n^{-q/\beta}$$

and

$$e_n^{ran}(\mathcal{G}_1^q) \geq c\, n^{-(q/\beta+1/2)}.$$

The proof of Proposition 9 will be based on the following lemma, which itself is a consequence of more general results on lower bounds for minimal errors in the context of linear problems with standard information, see [Nov88].

Lemma 3. *Let $n, k \in \mathbb{N}$ with $k > n$ and let $\varepsilon > 0$. Assume that there exist Borel-measurable functions*

$$h_1, \ldots, h_k : \mathbb{R}^d \to [0, \infty)$$

such that

(a) *the sets $\{h_i \neq 0\}$ are pairwise disjoint,*
(b) *for all $\sigma_1, \ldots, \sigma_k \in \{-1, 1\}$,*

$$\sum_{i=1}^{k} \sigma_i\, h_i \in \mathcal{G}_1^q,$$

(c) *for every $i = 1, \ldots, k$*

$$\int h_i \,\mathrm{d}P \geq \varepsilon.$$

Then

$$e_n^{det}(\mathcal{G}_1^q) \geq (k - n)\, \varepsilon.$$

Furthermore, if $k > 4n$ then

$$e_n^{ran}(\mathcal{G}_1^q) \geq (k/4 - n)^{1/2}\, \varepsilon.$$

Proof of Proposition 9. Fix a non-empty open set $O \subset \mathbb{R}^d$ according to condition (S2). According to [Sch94] we may assume that $O \cap K \neq \emptyset$. Choose $x \in O \cap K$ and note that $P(B(x, \varepsilon)) > 0$ for every $\varepsilon > 0$. Take any non-negative function $h \in \mathcal{G}_1^q$ with

$$\emptyset \neq \{h \neq 0\} \subset O$$

and choose $\varepsilon > 0$ such that $B(x, \varepsilon) \subset O$ and $\inf_{y \in B(x,\varepsilon)} h(y) > 0$. Then

$$\int h \,\mathrm{d}P > 0.$$

Consider a cutset $\mathcal{C} \subset \mathcal{T} \backslash \{\lambda\}$ and put

$$O_{\mathbf{j}} = S_{\mathbf{j}}(O)$$

as well as

$$h_{\mathbf{j}} = \frac{r_{\mathbf{j}}^q}{2}\, h \circ S_{\mathbf{j}}^{-1}$$

for all $\mathbf{j} \in \mathcal{C}$. Below we will prove that the set of functions $\{h_{\mathbf{j}} \colon \mathbf{j} \in \mathcal{C}\}$ satisfies the conditions (a) to (c) in Lemma 3 with

$$\varepsilon = \frac{1}{2} \min_{\mathbf{j} \in \mathcal{C}} s_{\mathbf{j}} \int h \,\mathrm{d}P.$$

For the particular choice of $\mathcal{C} = \mathcal{C}(T)$ with $T \geq 1$ we have $\min_{\mathbf{j}\in\mathcal{C}(T)} s_{\mathbf{j}} \geq s_{\min} T^{-1}$. Moreover, $\#\mathcal{C}(T) \geq T^\theta$, see Proposition 4. Let $n \in \mathbb{N}$. Choose $T = (2n)^{1/\theta}$ and apply Lemma 3 with $k = \#\mathcal{C}(T) \geq 2n$ to obtain

$$e_n^{\mathrm{det}}(\mathcal{G}_1^q) \geq 2^{-1/\theta-1} s_{\min} \int h \,\mathrm{d}P \, n^{-(1/\theta-1)}.$$

Choose $T = (6n)^{1/\theta}$ and apply Lemma 3 with $k = \#\mathcal{C}(T) \geq 6n$ to conclude

$$e_n^{\mathrm{ran}}(\mathcal{G}_1^q) \geq 2^{-3/2}\, 6^{-1/\theta} s_{\min} \int h \,\mathrm{d}P \, n^{-(1/\theta-1/2)}.$$

It remains to verify properties (a) to (c) in Lemma 3 for the functions $h_{\mathbf{j}}$ with $\mathbf{j} \in \mathcal{C}$. We first note that the sets $O_{\mathbf{j}}$ are pairwise disjoint since $\mathcal{C}$ is a cutset, see [Hut81, Sec. 5.2]. By definition of h we have

$$\{h_{\mathbf{j}} \neq 0\} \subset O_{\mathbf{j}}, \tag{25}$$

whence condition (a) is satisfied.

Furthermore, $S_{\mathbf{j}}(\mathrm{supp}(P)) \subset \mathrm{cl}(S_{\mathbf{j}}(O))$ for every $\mathbf{j} \in \mathcal{T}$, see [Hut81, Sec. 5.2]. Hence

$$P_{\mathbf{j}'}(O_{\mathbf{j}}) = P_{\mathbf{j}'}(O_{\mathbf{j}} \cap \mathrm{cl}(O_{\mathbf{j}'})) = P_{\mathbf{j}'}(\emptyset) = 0$$

for all $\mathbf{j}, \mathbf{j}' \in \mathcal{C}$ with $\mathbf{j} \neq \mathbf{j}'$. Therefore, by Proposition 1,

$$\int h_{\mathbf{j}} \,\mathrm{d}P = \sum_{\mathbf{j}'\in\mathcal{C}} \rho_{\mathbf{j}'} \int_{O_{\mathbf{j}}} h_{\mathbf{j}} \,\mathrm{d}P_{\mathbf{j}'} = \frac{1}{2}\, \rho_{\mathbf{j}}\, r_{\mathbf{j}}^q \int_{O_{\mathbf{j}}} h \circ S_{\mathbf{j}}^{-1} \,\mathrm{d}P_{\mathbf{j}} = \frac{1}{2}\, \rho_{\mathbf{j}}\, r_{\mathbf{j}}^q \int_O h \,\mathrm{d}P = \frac{1}{2}\, s_{\mathbf{j}} \int h \,\mathrm{d}P,$$

which shows property (c).

Finally, let $\sigma = (\sigma_{\mathbf{j}})_{\mathbf{j}\in\mathcal{C}} \in \{-1, 1\}^{\#\mathcal{C}}$ and consider the function

$$f = \sum_{\mathbf{j}\in\mathcal{C}} \sigma_{\mathbf{j}}\, h_{\mathbf{j}}.$$

Clearly, $f \in C^{q^*}(\mathbb{R}^d)$ and for $x, y, v \in \mathbb{R}^d$ with $\|v\| = 1$ we have

$$D_v^{q^*} f(x) - D_v^{q^*} f(y) = \sum_{\mathbf{j}\in\mathcal{C}} \sigma_{\mathbf{j}}\, (D_v^{q^*} h_{\mathbf{j}}(x) - D_v^{q^*} h_{\mathbf{j}}(y)). \tag{26}$$

For $\mathbf{j} = (j_1, \ldots, j_\ell) \in \mathcal{C}$ we put

$$V_{\mathbf{j}} = V_{j_1} \cdots V_{j_\ell},\ u_{\mathbf{j}} = V_{\mathbf{j}}^{-1} v.$$

By (S1) we have

$$S_{\mathbf{j}}^{-1}(x) = r_{\mathbf{j}}^{-1} V_{\mathbf{j}}^{-1} x + a_{\mathbf{j}}$$

for some $a_{\mathbf{j}} \in \mathbb{R}^d$, and therefore,

$$D_v^{q^*} h_{\mathbf{j}}(x) = \frac{1}{2}\, r_{\mathbf{j}}^q D_v^{q^*} (h \circ S_{\mathbf{j}}^{-1})(x) = \frac{1}{2}\, r_{\mathbf{j}}^{q-q^*} D_{u_{\mathbf{j}}}^{q^*} h(S_{\mathbf{j}}^{-1}(x)).$$

Consequently, for every $\mathbf{j} \in \mathcal{C}$,

$$|D_v^{q^*} h_{\mathbf{j}}(x) - D_v^{q^*} h_{\mathbf{j}}(y)| \leq \frac{1}{2}\, r_{\mathbf{j}}^{q-q^*} \|S_{\mathbf{j}}^{-1}(x) - S_{\mathbf{j}}^{-1}(y)\|^{q-q^*} = \frac{1}{2}\, \|x - y\|^{q-q^*}. \tag{27}$$

We show that for every $z \in \mathbb{R}^d$,

$$\#\{\mathbf{j} \in \mathcal{C} \colon D_v^{q^*} h_{\mathbf{j}}(z) \neq 0\} \leq 1, \tag{28}$$

which jointly with (26) and (27) implies $|D_v^q f(x) - D_v^q f(y)| \leq \|x - y\|^{q-q^*}$ and hereby completes the proof of property (b).

For the proof of (28) we consider the functions

$$g_{\mathbf{j}}\colon \mathbb{R} \to \mathbb{R},\ t \mapsto h_{\mathbf{j}}(z+tv),$$

for $\mathbf{j} \in \mathcal{C}$. Clearly, $g_{\mathbf{j}} \in C^{q^*}(\mathbb{R}^d)$ and

$$D_v^{q^*} h_{\mathbf{j}}(z) = g_{\mathbf{j}}^{(q^*)}(0).$$

Let $\mathbf{j} \in \mathcal{C}$ with $g_{\mathbf{j}}^{(q^*)}(0) \neq 0$. Assume that there exists a sequence $(t_\ell)_{\ell\in\mathbb{N}}$ in $\mathbb{R}\backslash\{0\}$ such that $\lim_{\ell\to\infty} t_\ell = 0$ and $z + t_\ell v \in \mathbb{R}^d\backslash O_{\mathbf{j}}$ for all $\ell \in \mathbb{N}$. Then $g_{\mathbf{j}}(t_\ell) = 0$ for every $\ell \in \mathbb{N}$, and, using the mean value theorem, we obtain by induction that there exists a sequence $(\tilde{t}_\ell)_{\ell\in\mathbb{N}}$ in $\mathbb{R}\backslash\{0\}$ such that $\lim_{\ell\to\infty} \tilde{t}_\ell = 0$ and $g_{\mathbf{j}}^{(q^*)}(\tilde{t}_\ell) = 0$ for every $\ell \in \mathbb{N}$. The latter contradicts $g_{\mathbf{j}}^{(q^*)}(0) \neq 0$. Hence there exists $t_0 > 0$ such that

$$\{z + tv\colon t \in (-t_0, t_0)\backslash\{0\}\} \subset O_{\mathbf{j}}.$$

Consequently, $g_{\mathbf{j}'}(t) = 0$ for all $\mathbf{j}' \neq \mathbf{j}$ and $t \in (-t_0, t_0)\backslash\{0\}$, which implies $g_{\mathbf{j}'}^{(q^*)}(0) = 0$ for all $\mathbf{j}' \neq \mathbf{j}$ and finishes the proof. □

Combining Theorem 1 with Proposition 9 we conclude that the deterministic quadrature rules $I_Q^{(n)}$ perform asymptotically optimal in the class $\mathcal{A}_q^{\mathrm{det}}$ of all deterministic methods for quadrature with respect to P, if the conditions (S1) and (S2) are satisfied. Similarly, from Theorem 2 and Proposition 9 we obtain that the randomized quadrature rules $\widehat{I}_{\mathfrak{X}}^{(n)}$ perform asymptotically optimal in the class $\mathcal{A}_q^{\mathrm{ran}}$ of all randomized methods for quadrature with respect to P, if the conditions (S1) and (S2) are satisfied.

Theorem 4. *Assume that the contractions $S_1, \ldots, S_m$ are similarities and satisfy the open set condition. Then there exist $c_2 \geq c_1 > 0$ such that for sufficiently large $n \in \mathbb{N}$,*

$$c_1\, n^{-q/\beta} \leq e_n^{det}(\mathcal{H}_1^q) \leq e(I_Q^{(n)}) \leq c_2\, n^{-q/\beta}$$

and

$$c_1\, n^{-(q/\beta+1/2)} \leq e_n^{ran}(\mathcal{H}_1^q) \leq e(\widehat{I}_{\mathfrak{X}}^{(n)}) \leq c_2\, n^{-(q/\beta+1/2)}.$$

Appendix: Moments of self-affine measures

We provide a recursion formula for the computation of moments of P in the case of affine linear contractions

$$S_j(x) = A_j x + b_j$$

with $A_j \in \mathbb{R}^{d\times d}$ and $b_j \in \mathbb{R}^d$ for $j = 1, \ldots, m$.

Put $V = \mathbb{R}^d$ and consider a d-dimensional random vector X with

$$X \sim P.$$

For every $\ell \in \mathbb{N}$ the mapping

$$V^\ell \ni (v_1, \ldots, v_\ell) \mapsto \mathbb{E}(v_1^{\mathrm{T}} X \cdots v_\ell^{\mathrm{T}} X) \in \mathbb{R}$$

is multilinear and hence it defines a real-valued linear mapping M_ℓ on the ℓ-th tensor power $V^{\otimes \ell}$ via

$$M_\ell(v_1 \otimes \cdots \otimes v_\ell) = \mathbb{E}(v_1^{\mathrm{T}} X \cdots v_\ell^{\mathrm{T}} X).$$

Proposition 10. *For every $\ell \in \mathbb{N}$ the mapping*

$$\mathrm{id}_{V^{\otimes l}} - \sum_{j=1}^{m} \rho_j \, (A_j^{\mathrm{T}})^{\otimes \ell} \colon V^{\otimes l} \to V^{\otimes l}$$

is a bijection and for every $\mathbf{v} = v_1 \otimes \cdots \otimes v_\ell \in V^{\otimes l}$ we have

$$M_\ell\Big(\mathrm{id}_{V^{\otimes l}} - \sum_{j=1}^{m} \rho_j \, (A_j^{\mathrm{T}})^{\otimes \ell}\Big)(\mathbf{v}) = \sum_{j=1}^{m} \rho_j \Big(\prod_{i=1}^{\ell} v_i^{\mathrm{T}} b_j + \sum_{\emptyset \neq I \subsetneq \{1,\dots,\ell\}} \Big(\prod_{i \in I^c} v_i^{\mathrm{T}} b_j \Big) \, M_{\# I}\Big(\bigotimes_{i \in I} A_j^{\mathrm{T}} v_i \Big)\Big).$$

Proof. Let $\ell \in \mathbb{N}$ and $v_1, \dots, v_\ell \in V$. By the self-similarity of P and the particular form of the contractions S_j we have

$$\begin{aligned}
\mathbb{E}(v_1^{\mathrm{T}} X \cdots v_\ell^{\mathrm{T}} X) &= \sum_{j=1}^{m} \rho_j \, \mathbb{E}(v_1^{\mathrm{T}} S_j X \cdots v_\ell^{\mathrm{T}} S_j X) \\
&= \sum_{j=1}^{m} \rho_j \sum_{I \subset \{1,\dots,l\}} \Big(\prod_{i \in I^c} v_i^{\mathrm{T}} b_j \Big) \, \mathbb{E}\Big(\prod_{i \in I} v_i^{\mathrm{T}} A_j X \Big),
\end{aligned}$$

which implies the recursion formula.

Consider any norm $\|\cdot\|_{V^{\otimes l}}$ on $V^{\otimes l}$ such that $\|v_1 \otimes \cdots \otimes v_\ell\|_{V^{\otimes l}} = \|v_1\| \cdots \|v_\ell\|$ for $v_1, \dots, v_\ell \in V$. Then $v_1 \otimes \cdots \otimes v_\ell = \sum_{j=1}^{m} \rho_j (A_j^{\mathrm{T}} v_1 \otimes \cdots \otimes A_j^{\mathrm{T}} v_\ell)$ implies

$$\begin{aligned}
\|v_1\| \cdots \|v_\ell\| &= \Big\| \sum_{j=1}^{m} \rho_j (A_j^{\mathrm{T}} v_1 \otimes \cdots \otimes A_j^{\mathrm{T}} v_\ell) \Big\|_{V^{\otimes l}} \\
&\leq \sum_{j=1}^{m} \rho_j \|A_j^{\mathrm{T}} v_1\| \cdots \|A_j^{\mathrm{T}} v_\ell\| \leq \|v_1\| \cdots \|v_\ell\| \sum_{j=1}^{m} \rho_j r_j^{\ell},
\end{aligned}$$

and, consequently, $v_1 \otimes \cdots \otimes v_\ell = 0$. Hence the mapping $\mathrm{id}_{V^{\otimes l}} - \sum_{j=1}^{m} \rho_j \, (A_j^{\mathrm{T}})^{\otimes \ell}$ is injective, which completes the proof. ☐

Remark 9. The recursion formula in Proposition 10 simplifies significantly in the case of $d = 1$. Taking $v_1 = \cdots = v_\ell = 1$ we immediately obtain

$$\mathbb{E}(X^\ell) = \Big(1 - \sum_{j=1}^{m} \rho_j A_j^l\Big)^{-1} \sum_{j=1}^{m} \rho_j \sum_{k=0}^{l-1} \binom{l}{k} b_j^{l-k} \, A_j^k \, \mathbb{E}(X^k).$$

References

[BBCR13] D. H. Bailey, J. M. Borwein, R. E. Crandall, and M. G. Rose. Expectations on fractal sets. *Appl. Math. Comput.*, 220:695–721, 2013.

[BT06] C. Bayer and J. Teichmann. The proof of Tchakaloff's theorem. *Proc. Amer. Math. Soc.*, 134(10):3035–3040, 2006.

[Cra13] R. Crandall. On the fractal distribution of brain synapses. In *Computational and Analytical Mathematics*, volume 50 of *Springer Proceedings in Mathematics & Statistics*, pages 325–348. 2013.

[Din99] S. Dineen. *Complex Analysis on Infinite Dimensional Spaces.* Springer, London, 1999.

[Fal] K. Falconer. *Fractal geometry.* Second edition.

[Fal97] K. Falconer. *Techniques in fractal geometry.* John Wiley & Sons Ltd., Chichester, 1997.

[GL01] S. Graf and H. Luschgy. Asymptotics of the quantization errors for self-similar probabilities. *Real Anal. Exchange*, 26(2):795–810, 2000/01.

[Hut81] J. E. Hutchinson. Fractals and self-similarity. *Indiana Univ. Math. J.*, 30(5):713–747, 1981.

[Man82] B. B. Mandelbrot. *The fractal geometry of nature.* W. H. Freeman and Co., San Francisco, Calif., 1982.

[MM08] G. Mastroianni and G. V. Milovanović. *Interpolation processes.* Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2008. Basic theory and applications.

[Nov88] E. Novak. *Deterministic and stochastic error bounds in numerical analysis.* Lecture Notes in Mathematics, 1349. Berlin etc.: Springer-Verlag., 1988.

[NW08] E. Novak and H. Woźniakowski. *Tractability of multivariate problems. Vol. 1: Linear information*, volume 6 of *EMS Tracts in Mathematics.* European Mathematical Society (EMS), Zürich, 2008.

[NW10] E. Novak and H. Woźniakowski. *Tractability of multivariate problems. Volume II: Standard information for functionals*, volume 12 of *EMS Tracts in Mathematics.* European Mathematical Society (EMS), Zürich, 2010.

[NW12] E. Novak and H. Woźniakowski. *Tractability of multivariate problems. Volume III: Standard information for operators*, volume 18 of *EMS Tracts in Mathematics.* European Mathematical Society (EMS), Zürich, 2012.

[Rit00] K. Ritter. *Average-case analysis of numerical problems*, volume 1733 of *Lecture Notes in Mathematics.* Springer-Verlag, Berlin, 2000.

[Sch94] A. Schief. Separation properties for self-similar sets. *Proc. Amer. Math. Soc.*, 122:111–115, 1994.

[Tch57] V. Tchakaloff. Formules de cubatures mécaniques à coefficients non négatifs. *Bull. Sci. Math. (2)*, 81:123–134, 1957.

Steffen Dereich, Institut für Mathematische Statistik, Fachbereich 10: Mathematik und Informatik, Westfälische Wilhelms-Universität Münster, Einsteinstr. 62, 48149 Münster, Germany
E-mail address: steffen.dereich@wwu.de

Thomas Müller-Gronbach, Fakultät für Informatik und Mathematik, Universität Passau, Innstrasse 33, 94032 Passau, Germany
E-mail address: thomas.mueller-gronbach@uni-passau.de

Preprint Series DFG-SPP 1324

http://www.dfg-spp1324.de

Reports

[1] R. Ramlau, G. Teschke, and M. Zhariy. A Compressive Landweber Iteration for Solving Ill-Posed Inverse Problems. Preprint 1, DFG-SPP 1324, September 2008.

[2] G. Plonka. The Easy Path Wavelet Transform: A New Adaptive Wavelet Transform for Sparse Representation of Two-dimensional Data. Preprint 2, DFG-SPP 1324, September 2008.

[3] E. Novak and H. Woźniakowski. Optimal Order of Convergence and (In-) Tractability of Multivariate Approximation of Smooth Functions. Preprint 3, DFG-SPP 1324, October 2008.

[4] M. Espig, L. Grasedyck, and W. Hackbusch. Black Box Low Tensor Rank Approximation Using Fibre-Crosses. Preprint 4, DFG-SPP 1324, October 2008.

[5] T. Bonesky, S. Dahlke, P. Maass, and T. Raasch. Adaptive Wavelet Methods and Sparsity Reconstruction for Inverse Heat Conduction Problems. Preprint 5, DFG-SPP 1324, January 2009.

[6] E. Novak and H. Woźniakowski. Approximation of Infinitely Differentiable Multivariate Functions Is Intractable. Preprint 6, DFG-SPP 1324, January 2009.

[7] J. Ma and G. Plonka. A Review of Curvelets and Recent Applications. Preprint 7, DFG-SPP 1324, February 2009.

[8] L. Denis, D. A. Lorenz, and D. Trede. Greedy Solution of Ill-Posed Problems: Error Bounds and Exact Inversion. Preprint 8, DFG-SPP 1324, April 2009.

[9] U. Friedrich. A Two Parameter Generalization of Lions' Nonoverlapping Domain Decomposition Method for Linear Elliptic PDEs. Preprint 9, DFG-SPP 1324, April 2009.

[10] K. Bredies and D. A. Lorenz. Minimization of Non-smooth, Non-convex Functionals by Iterative Thresholding. Preprint 10, DFG-SPP 1324, April 2009.

[11] K. Bredies and D. A. Lorenz. Regularization with Non-convex Separable Constraints. Preprint 11, DFG-SPP 1324, April 2009.

[12] M. Döhler, S. Kunis, and D. Potts. Nonequispaced Hyperbolic Cross Fast Fourier Transform. Preprint 12, DFG-SPP 1324, April 2009.

[13] C. Bender. Dual Pricing of Multi-Exercise Options under Volume Constraints. Preprint 13, DFG-SPP 1324, April 2009.

[14] T. Müller-Gronbach and K. Ritter. Variable Subspace Sampling and Multi-level Algorithms. Preprint 14, DFG-SPP 1324, May 2009.

[15] G. Plonka, S. Tenorth, and A. Iske. Optimally Sparse Image Representation by the Easy Path Wavelet Transform. Preprint 15, DFG-SPP 1324, May 2009.

[16] S. Dahlke, E. Novak, and W. Sickel. Optimal Approximation of Elliptic Problems by Linear and Nonlinear Mappings IV: Errors in L_2 and Other Norms. Preprint 16, DFG-SPP 1324, June 2009.

[17] B. Jin, T. Khan, P. Maass, and M. Pidcock. Function Spaces and Optimal Currents in Impedance Tomography. Preprint 17, DFG-SPP 1324, June 2009.

[18] G. Plonka and J. Ma. Curvelet-Wavelet Regularized Split Bregman Iteration for Compressed Sensing. Preprint 18, DFG-SPP 1324, June 2009.

[19] G. Teschke and C. Borries. Accelerated Projected Steepest Descent Method for Nonlinear Inverse Problems with Sparsity Constraints. Preprint 19, DFG-SPP 1324, July 2009.

[20] L. Grasedyck. Hierarchical Singular Value Decomposition of Tensors. Preprint 20, DFG-SPP 1324, July 2009.

[21] D. Rudolf. Error Bounds for Computing the Expectation by Markov Chain Monte Carlo. Preprint 21, DFG-SPP 1324, July 2009.

[22] M. Hansen and W. Sickel. Best m-term Approximation and Lizorkin-Triebel Spaces. Preprint 22, DFG-SPP 1324, August 2009.

[23] F.J. Hickernell, T. Müller-Gronbach, B. Niu, and K. Ritter. Multi-level Monte Carlo Algorithms for Infinite-dimensional Integration on $\mathbb{R}^{\mathbb{N}}$. Preprint 23, DFG-SPP 1324, August 2009.

[24] S. Dereich and F. Heidenreich. A Multilevel Monte Carlo Algorithm for Lévy Driven Stochastic Differential Equations. Preprint 24, DFG-SPP 1324, August 2009.

[25] S. Dahlke, M. Fornasier, and T. Raasch. Multilevel Preconditioning for Adaptive Sparse Optimization. Preprint 25, DFG-SPP 1324, August 2009.

[26] S. Dereich. Multilevel Monte Carlo Algorithms for Lévy-driven SDEs with Gaussian Correction. Preprint 26, DFG-SPP 1324, August 2009.

[27] G. Plonka, S. Tenorth, and D. Roşca. A New Hybrid Method for Image Approximation using the Easy Path Wavelet Transform. Preprint 27, DFG-SPP 1324, October 2009.

[28] O. Koch and C. Lubich. Dynamical Low-rank Approximation of Tensors. Preprint 28, DFG-SPP 1324, November 2009.

[29] E. Faou, V. Gradinaru, and C. Lubich. Computing Semi-classical Quantum Dynamics with Hagedorn Wavepackets. Preprint 29, DFG-SPP 1324, November 2009.

[30] D. Conte and C. Lubich. An Error Analysis of the Multi-configuration Time-dependent Hartree Method of Quantum Dynamics. Preprint 30, DFG-SPP 1324, November 2009.

[31] C. E. Powell and E. Ullmann. Preconditioning Stochastic Galerkin Saddle Point Problems. Preprint 31, DFG-SPP 1324, November 2009.

[32] O. G. Ernst and E. Ullmann. Stochastic Galerkin Matrices. Preprint 32, DFG-SPP 1324, November 2009.

[33] F. Lindner and R. L. Schilling. Weak Order for the Discretization of the Stochastic Heat Equation Driven by Impulsive Noise. Preprint 33, DFG-SPP 1324, November 2009.

[34] L. Kämmerer and S. Kunis. On the Stability of the Hyperbolic Cross Discrete Fourier Transform. Preprint 34, DFG-SPP 1324, December 2009.

[35] P. Cerejeiras, M. Ferreira, U. Kähler, and G. Teschke. Inversion of the noisy Radon transform on $SO(3)$ by Gabor frames and sparse recovery principles. Preprint 35, DFG-SPP 1324, January 2010.

[36] T. Jahnke and T. Udrescu. Solving Chemical Master Equations by Adaptive Wavelet Compression. Preprint 36, DFG-SPP 1324, January 2010.

[37] P. Kittipoom, G. Kutyniok, and W.-Q Lim. Irregular Shearlet Frames: Geometry and Approximation Properties. Preprint 37, DFG-SPP 1324, February 2010.

[38] G. Kutyniok and W.-Q Lim. Compactly Supported Shearlets are Optimally Sparse. Preprint 38, DFG-SPP 1324, February 2010.

[39] M. Hansen and W. Sickel. Best m-Term Approximation and Tensor Products of Sobolev and Besov Spaces – the Case of Non-compact Embeddings. Preprint 39, DFG-SPP 1324, March 2010.

[40] B. Niu, F.J. Hickernell, T. Müller-Gronbach, and K. Ritter. Deterministic Multi-level Algorithms for Infinite-dimensional Integration on $\mathbb{R}^{\mathbb{N}}$. Preprint 40, DFG-SPP 1324, March 2010.

[41] P. Kittipoom, G. Kutyniok, and W.-Q Lim. Construction of Compactly Supported Shearlet Frames. Preprint 41, DFG-SPP 1324, March 2010.

[42] C. Bender and J. Steiner. Error Criteria for Numerical Solutions of Backward SDEs. Preprint 42, DFG-SPP 1324, April 2010.

[43] L. Grasedyck. Polynomial Approximation in Hierarchical Tucker Format by Vector-Tensorization. Preprint 43, DFG-SPP 1324, April 2010.

[44] M. Hansen und W. Sickel. Best m-Term Approximation and Sobolev-Besov Spaces of Dominating Mixed Smoothness - the Case of Compact Embeddings. Preprint 44, DFG-SPP 1324, April 2010.

[45] P. Binev, W. Dahmen, and P. Lamby. Fast High-Dimensional Approximation with Sparse Occupancy Trees. Preprint 45, DFG-SPP 1324, May 2010.

[46] J. Ballani and L. Grasedyck. A Projection Method to Solve Linear Systems in Tensor Format. Preprint 46, DFG-SPP 1324, May 2010.

[47] P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova, and P. Wojtaszczyk. Convergence Rates for Greedy Algorithms in Reduced Basis Methods. Preprint 47, DFG-SPP 1324, May 2010.

[48] S. Kestler and K. Urban. Adaptive Wavelet Methods on Unbounded Domains. Preprint 48, DFG-SPP 1324, June 2010.

[49] H. Yserentant. The Mixed Regularity of Electronic Wave Functions Multiplied by Explicit Correlation Factors. Preprint 49, DFG-SPP 1324, June 2010.

[50] H. Yserentant. On the Complexity of the Electronic Schrödinger Equation. Preprint 50, DFG-SPP 1324, June 2010.

[51] M. Guillemard and A. Iske. Curvature Analysis of Frequency Modulated Manifolds in Dimensionality Reduction. Preprint 51, DFG-SPP 1324, June 2010.

[52] E. Herrholz and G. Teschke. Compressive Sensing Principles and Iterative Sparse Recovery for Inverse and Ill-Posed Problems. Preprint 52, DFG-SPP 1324, July 2010.

[53] L. Kämmerer, S. Kunis, and D. Potts. Interpolation Lattices for Hyperbolic Cross Trigonometric Polynomials. Preprint 53, DFG-SPP 1324, July 2010.

[54] G. Kutyniok and W.-Q Lim. Shearlets on Bounded Domains. Preprint 54, DFG-SPP 1324, July 2010.

[55] A. Zeiser. Wavelet Approximation in Weighted Sobolev Spaces of Mixed Order with Applications to the Electronic Schrödinger Equation. Preprint 55, DFG-SPP 1324, July 2010.

[56] G. Kutyniok, J. Lemvig, and W.-Q Lim. Compactly Supported Shearlets. Preprint 56, DFG-SPP 1324, July 2010.

[57] A. Zeiser. On the Optimality of the Inexact Inverse Iteration Coupled with Adaptive Finite Element Methods. Preprint 57, DFG-SPP 1324, July 2010.

[58] S. Jokar. Sparse Recovery and Kronecker Products. Preprint 58, DFG-SPP 1324, August 2010.

[59] T. Aboiyar, E. H. Georgoulis, and A. Iske. Adaptive ADER Methods Using Kernel-Based Polyharmonic Spline WENO Reconstruction. Preprint 59, DFG-SPP 1324, August 2010.

[60] O. G. Ernst, A. Mugler, H.-J. Starkloff, and E. Ullmann. On the Convergence of Generalized Polynomial Chaos Expansions. Preprint 60, DFG-SPP 1324, August 2010.

[61] S. Holtz, T. Rohwedder, and R. Schneider. On Manifolds of Tensors of Fixed TT-Rank. Preprint 61, DFG-SPP 1324, September 2010.

[62] J. Ballani, L. Grasedyck, and M. Kluge. Black Box Approximation of Tensors in Hierarchical Tucker Format. Preprint 62, DFG-SPP 1324, October 2010.

[63] M. Hansen. On Tensor Products of Quasi-Banach Spaces. Preprint 63, DFG-SPP 1324, October 2010.

[64] S. Dahlke, G. Steidl, and G. Teschke. Shearlet Coorbit Spaces: Compactly Supported Analyzing Shearlets, Traces and Embeddings. Preprint 64, DFG-SPP 1324, October 2010.

[65] W. Hackbusch. Tensorisation of Vectors and their Efficient Convolution. Preprint 65, DFG-SPP 1324, November 2010.

[66] P. A. Cioica, S. Dahlke, S. Kinzel, F. Lindner, T. Raasch, K. Ritter, and R. L. Schilling. Spatial Besov Regularity for Stochastic Partial Differential Equations on Lipschitz Domains. Preprint 66, DFG-SPP 1324, November 2010.

[67] E. Novak and H. Woźniakowski. On the Power of Function Values for the Approximation Problem in Various Settings. Preprint 67, DFG-SPP 1324, November 2010.

[68] A. Hinrichs, E. Novak, and H. Woźniakowski. The Curse of Dimensionality for Monotone and Convex Functions of Many Variables. Preprint 68, DFG-SPP 1324, November 2010.

[69] G. Kutyniok and W.-Q Lim. Image Separation Using Shearlets. Preprint 69, DFG-SPP 1324, November 2010.

[70] B. Jin and P. Maass. An Analysis of Electrical Impedance Tomography with Applications to Tikhonov Regularization. Preprint 70, DFG-SPP 1324, December 2010.

[71] S. Holtz, T. Rohwedder, and R. Schneider. The Alternating Linear Scheme for Tensor Optimisation in the TT Format. Preprint 71, DFG-SPP 1324, December 2010.

[72] T. Müller-Gronbach and K. Ritter. A Local Refinement Strategy for Constructive Quantization of Scalar SDEs. Preprint 72, DFG-SPP 1324, December 2010.

[73] T. Rohwedder and R. Schneider. An Analysis for the DIIS Acceleration Method used in Quantum Chemistry Calculations. Preprint 73, DFG-SPP 1324, December 2010.

[74] C. Bender and J. Steiner. Least-Squares Monte Carlo for Backward SDEs. Preprint 74, DFG-SPP 1324, December 2010.

[75] C. Bender. Primal and Dual Pricing of Multiple Exercise Options in Continuous Time. Preprint 75, DFG-SPP 1324, December 2010.

[76] H. Harbrecht, M. Peters, and R. Schneider. On the Low-rank Approximation by the Pivoted Cholesky Decomposition. Preprint 76, DFG-SPP 1324, December 2010.

[77] P. A. Cioica, S. Dahlke, N. Döhring, S. Kinzel, F. Lindner, T. Raasch, K. Ritter, and R. L. Schilling. Adaptive Wavelet Methods for Elliptic Stochastic Partial Differential Equations. Preprint 77, DFG-SPP 1324, January 2011.

[78] G. Plonka, S. Tenorth, and A. Iske. Optimal Representation of Piecewise Hölder Smooth Bivariate Functions by the Easy Path Wavelet Transform. Preprint 78, DFG-SPP 1324, January 2011.

[79] A. Mugler and H.-J. Starkloff. On Elliptic Partial Differential Equations with Random Coefficients. Preprint 79, DFG-SPP 1324, January 2011.

[80] T. Müller-Gronbach, K. Ritter, and L. Yaroslavtseva. A Derandomization of the Euler Scheme for Scalar Stochastic Differential Equations. Preprint 80, DFG-SPP 1324, January 2011.

[81] W. Dahmen, C. Huang, C. Schwab, and G. Welper. Adaptive Petrov-Galerkin methods for first order transport equations. Preprint 81, DFG-SPP 1324, January 2011.

[82] K. Grella and C. Schwab. Sparse Tensor Spherical Harmonics Approximation in Radiative Transfer. Preprint 82, DFG-SPP 1324, January 2011.

[83] D.A. Lorenz, S. Schiffler, and D. Trede. Beyond Convergence Rates: Exact Inversion With Tikhonov Regularization With Sparsity Constraints. Preprint 83, DFG-SPP 1324, January 2011.

[84] S. Dereich, M. Scheutzow, and R. Schottstedt. Constructive quantization: Approximation by empirical measures. Preprint 84, DFG-SPP 1324, January 2011.

[85] S. Dahlke and W. Sickel. On Besov Regularity of Solutions to Nonlinear Elliptic Partial Differential Equations. Preprint 85, DFG-SPP 1324, January 2011.

[86] S. Dahlke, U. Friedrich, P. Maass, T. Raasch, and R.A. Ressel. An adaptive wavelet method for parameter identification problems in parabolic partial differential equations. Preprint 86, DFG-SPP 1324, January 2011.

[87] A. Cohen, W. Dahmen, and G. Welper. Adaptivity and Variational Stabilization for Convection-Diffusion Equations. Preprint 87, DFG-SPP 1324, January 2011.

[88] T. Jahnke. On Reduced Models for the Chemical Master Equation. Preprint 88, DFG-SPP 1324, January 2011.

[89] P. Binev, W. Dahmen, R. DeVore, P. Lamby, D. Savu, and R. Sharpley. Compressed Sensing and Electron Microscopy. Preprint 89, DFG-SPP 1324, March 2011.

[90] P. Binev, F. Blanco-Silva, D. Blom, W. Dahmen, P. Lamby, R. Sharpley, and T. Vogt. High Quality Image Formation by Nonlocal Means Applied to High-Angle Annular Dark Field Scanning Transmission Electron Microscopy (HAADF-STEM). Preprint 90, DFG-SPP 1324, March 2011.

[91] R. A. Ressel. A Parameter Identification Problem for a Nonlinear Parabolic Differential Equation. Preprint 91, DFG-SPP 1324, May 2011.

[92] G. Kutyniok. Data Separation by Sparse Representations. Preprint 92, DFG-SPP 1324, May 2011.

[93] M. A. Davenport, M. F. Duarte, Y. C. Eldar, and G. Kutyniok. Introduction to Compressed Sensing. Preprint 93, DFG-SPP 1324, May 2011.

[94] H.-C. Kreusler and H. Yserentant. The Mixed Regularity of Electronic Wave Functions in Fractional Order and Weighted Sobolev Spaces. Preprint 94, DFG-SPP 1324, June 2011.

[95] E. Ullmann, H. C. Elman, and O. G. Ernst. Efficient Iterative Solvers for Stochastic Galerkin Discretizations of Log-Transformed Random Diffusion Problems. Preprint 95, DFG-SPP 1324, June 2011.

[96] S. Kunis and I. Melzer. On the Butterfly Sparse Fourier Transform. Preprint 96, DFG-SPP 1324, June 2011.

[97] T. Rohwedder. The Continuous Coupled Cluster Formulation for the Electronic Schrödinger Equation. Preprint 97, DFG-SPP 1324, June 2011.

[98] T. Rohwedder and R. Schneider. Error Estimates for the Coupled Cluster Method. Preprint 98, DFG-SPP 1324, June 2011.

[99] P. A. Cioica and S. Dahlke. Spatial Besov Regularity for Semilinear Stochastic Partial Differential Equations on Bounded Lipschitz Domains. Preprint 99, DFG-SPP 1324, July 2011.

[100] L. Grasedyck and W. Hackbusch. An Introduction to Hierarchical (H-) Rank and TT-Rank of Tensors with Examples. Preprint 100, DFG-SPP 1324, August 2011.

[101] N. Chegini, S. Dahlke, U. Friedrich, and R. Stevenson. Piecewise Tensor Product Wavelet Bases by Extensions and Approximation Rates. Preprint 101, DFG-SPP 1324, September 2011.

[102] S. Dahlke, P. Oswald, and T. Raasch. A Note on Quarkonial Systems and Multilevel Partition of Unity Methods. Preprint 102, DFG-SPP 1324, September 2011.

[103] A. Uschmajew. Local Convergence of the Alternating Least Squares Algorithm For Canonical Tensor Approximation. Preprint 103, DFG-SPP 1324, September 2011.

[104] S. Kvaal. Multiconfigurational time-dependent Hartree method for describing particle loss due to absorbing boundary conditions. Preprint 104, DFG-SPP 1324, September 2011.

[105] M. Guillemard and A. Iske. On Groupoid C*-Algebras, Persistent Homology and Time-Frequency Analysis. Preprint 105, DFG-SPP 1324, September 2011.

[106] A. Hinrichs, E. Novak, and H. Woźniakowski. Discontinuous information in the worst case and randomized settings. Preprint 106, DFG-SPP 1324, September 2011.

[107] M. Espig, W. Hackbusch, A. Litvinenko, H. Matthies, and E. Zander. Efficient Analysis of High Dimensional Data in Tensor Formats. Preprint 107, DFG-SPP 1324, September 2011.

[108] M. Espig, W. Hackbusch, S. Handschuh, and R. Schneider. Optimization Problems in Contracted Tensor Networks. Preprint 108, DFG-SPP 1324, October 2011.

[109] S. Dereich, T. Müller-Gronbach, and K. Ritter. On the Complexity of Computing Quadrature Formulas for SDEs. Preprint 109, DFG-SPP 1324, October 2011.

[110] D. Belomestny. Solving optimal stopping problems by empirical dual optimization and penalization. Preprint 110, DFG-SPP 1324, November 2011.

[111] D. Belomestny and J. Schoenmakers. Multilevel dual approach for pricing American style derivatives. Preprint 111, DFG-SPP 1324, November 2011.

[112] T. Rohwedder and A. Uschmajew. Local convergence of alternating schemes for optimization of convex problems in the TT format. Preprint 112, DFG-SPP 1324, December 2011.

[113] T. Görner, R. Hielscher, and S. Kunis. Efficient and accurate computation of spherical mean values at scattered center points. Preprint 113, DFG-SPP 1324, December 2011.

[114] Y. Dong, T. Görner, and S. Kunis. An iterative reconstruction scheme for photoacoustic imaging. Preprint 114, DFG-SPP 1324, December 2011.

[115] L. Kämmerer. Reconstructing hyperbolic cross trigonometric polynomials by sampling along generated sets. Preprint 115, DFG-SPP 1324, February 2012.

[116] H. Chen and R. Schneider. Numerical analysis of augmented plane waves methods for full-potential electronic structure calculations. Preprint 116, DFG-SPP 1324, February 2012.

[117] J. Ma, G. Plonka, and M.Y. Hussaini. Compressive Video Sampling with Approximate Message Passing Decoding. Preprint 117, DFG-SPP 1324, February 2012.

[118] D. Heinen and G. Plonka. Wavelet shrinkage on paths for scattered data denoising. Preprint 118, DFG-SPP 1324, February 2012.

[119] T. Jahnke and M. Kreim. Error bound for piecewise deterministic processes modeling stochastic reaction systems. Preprint 119, DFG-SPP 1324, March 2012.

[120] C. Bender and J. Steiner. A-posteriori estimates for backward SDEs. Preprint 120, DFG-SPP 1324, April 2012.

[121] M. Espig, W. Hackbusch, A. Litvinenkoy, H.G. Matthiesy, and P. Wähnert. Efficient low-rank approximation of the stochastic Galerkin matrix in tensor formats. Preprint 121, DFG-SPP 1324, May 2012.

[122] O. Bokanowski, J. Garcke, M. Griebel, and I. Klompmaker. An adaptive sparse grid semi-Lagrangian scheme for first order Hamilton-Jacobi Bellman equations. Preprint 122, DFG-SPP 1324, June 2012.

[123] A. Mugler and H.-J. Starkloff. On the convergence of the stochastic Galerkin method for random elliptic partial differential equations. Preprint 123, DFG-SPP 1324, June 2012.

[124] P.A. Cioica, S. Dahlke, N. Döhring, U. Friedrich, S. Kinzel, F. Lindner, T. Raasch, K. Ritter, and R.L. Schilling. On the convergence analysis of Rothe's method. Preprint 124, DFG-SPP 1324, July 2012.

[125] P. Binev, A. Cohen, W. Dahmen, and R. DeVore. Classification Algorithms using Adaptive Partitioning. Preprint 125, DFG-SPP 1324, July 2012.

[126] C. Lubich, T. Rohwedder, R. Schneider, and B. Vandereycken. Dynamical approximation of hierarchical Tucker and Tensor-Train tensors. Preprint 126, DFG-SPP 1324, July 2012.

[127] M. Kovács, S. Larsson, and K. Urban. On Wavelet-Galerkin methods for semilinear parabolic equations with additive noise. Preprint 127, DFG-SPP 1324, August 2012.

[128] M. Bachmayr, H. Chen, and R. Schneider. Numerical analysis of Gaussian approximations in quantum chemistry. Preprint 128, DFG-SPP 1324, August 2012.

[129] D. Rudolf. Explicit error bounds for Markov chain Monte Carlo. Preprint 129, DFG-SPP 1324, August 2012.

[130] P.A. Cioica, K.-H. Kim, K. Lee, and F. Lindner. On the $L_q(L_p)$-regularity and Besov smoothness of stochastic parabolic equations on bounded Lipschitz domains. Preprint 130, DFG-SPP 1324, December 2012.

[131] M. Hansen. $n-$term Approximation Rates and Besov Regularity for Elliptic PDEs on Polyhedral Domains. Preprint 131, DFG-SPP 1324, December 2012.

[132] R. E. Bank and H. Yserentant. On the H^1 -stability of the L_2 -projection onto finite element spaces. Preprint 132, DFG-SPP 1324, December 2012.

[133] M. Gnewuch, S. Mayer, and K. Ritter. On Weighted Hilbert Spaces and Integration of Functions of Infinitely Many Variables. Preprint 133, DFG-SPP 1324, December 2012.

[134] D. Crisan, J. Diehl, P.K. Friz, and H. Oberhauser. Robust Filtering: Correlated Noise and Multidimensional Observation. Preprint 134, DFG-SPP 1324, January 2013.

[135] Wolfgang Dahmen, Christian Plesken, and Gerrit Welper. Double Greedy Algorithms: Reduced Basis Methods for Transport Dominated Problems. Preprint 135, DFG-SPP 1324, February 2013.

[136] Aicke Hinrichs, Erich Novak, Mario Ullrich, and Henryk Wozniakowski. The Curse of Dimensionality for Numerical Integration of Smooth Functions. Preprint 136, DFG-SPP 1324, February 2013.

[137] Markus Bachmayr, Wolfgang Dahmen, Ronald DeVore, and Lars Grasedyck. Approximation of High-Dimensional Rank One Tensors. Preprint 137, DFG-SPP 1324, March 2013.

[138] Markus Bachmayr and Wolfgang Dahmen. Adaptive Near-Optimal Rank Tensor Approximation for High-Dimensional Operator Equations. Preprint 138, DFG-SPP 1324, April 2013.

[139] Felix Lindner. Singular Behavior of the Solution to the Stochastic Heat Equation on a Polygonal Domain. Preprint 139, DFG-SPP 1324, May 2013.

[140] Stephan Dahlke, Dominik Lellek, Shiu Hong Lui, and Rob Stevenson. Adaptive Wavelet Schwarz Methods for the Navier-Stokes Equation. Preprint 140, DFG-SPP 1324, May 2013.

[141] Jonas Ballani and Lars Grasedyck. Tree Adaptive Approximation in the Hierarchical Tensor Format. Preprint 141, DFG-SPP 1324, June 2013.

[142] Harry Yserentant. A short theory of the Rayleigh-Ritz method. Preprint 142, DFG-SPP 1324, July 2013.

[143] M. Hefter and K. Ritter. On Embeddings of Weighted Tensor Product Hilbert Spaces. Preprint 143, DFG-SPP 1324, August 2013.

[144] M. Altmayer and A. Neuenkirch. Multilevel Monte Carlo Quadrature of Discontinuous Payoffs in the Generalized Heston Model using Malliavin Integration by Parts. Preprint 144, DFG-SPP 1324, August 2013.

[145] L. Kämmerer, D. Potts, and T. Volkmer. Approximation of multivariate functions by trigonometric polynomials based on rank-1 lattice sampling. Preprint 145, DFG-SPP 1324, September 2013.

[146] C. Bender, N. Schweizer, and J. Zhuo. A primal-dual algorithm for BSDEs. Preprint 146, DFG-SPP 1324, October 2013.

[147] D. Rudolf. Hit-and-run for numerical integration. Preprint 147, DFG-SPP 1324, October 2013.

[148] D. Rudolf and M. Ullrich. Positivity of hit-and-run and related algorithms. Preprint 148, DFG-SPP 1324, October 2013.

[149] L. Grasedyck, M. Kluge, and S. Krämer. Alternating Directions Fitting (ADF) of Hierarchical Low Rank Tensors. Preprint 149, DFG-SPP 1324, October 2013.

[150] F. Filbir, S. Kunis, and R. Seyfried. Effective discretization of direct reconstruction schemes for photoacoustic imaging in spherical geometries. Preprint 150, DFG-SPP 1324, November 2013.

[151] E. Novak, M. Ullrich, and H. Woźniakowski. Complexity of Oscillatory Integration for Univariate Sobolev Spaces. Preprint 151, DFG-SPP 1324, November 2013.

[152] A. Hinrichs, E. Novak, and M. Ullrich. A Tractability Result for the Clenshaw Curtis Smolyak Algorithm. Preprint 152, DFG-SPP 1324, November 2013.

[153] M. Hein, S. Setzer, L. Jost, and S. Rangapuram. The Total Variation on Hypergraphs - Learning on Hypergraphs Revisited. Preprint 153, DFG-SPP 1324, November 2013.

[154] M. Kovács, S. Larsson, and F. Lindgren. On the Backward Euler Approximation of the Stochastic Allen-Chan Equation. Preprint 154, DFG-SPP 1324, November 2013.

[155] S. Dahlke, M. Fornasier, U. Friedrich, and T. Raasch. Multilevel preconditioning for sparse optimization of functionals with nonconvex fidelity terms. Preprint 155, DFG-SPP 1324, December 2013.

[156] T. Müller-Gronbach, K. Ritter, and L. Yaroslavtseva. On the complexity of computing quadrature formulas for marginal distributions of SDEs. Preprint 156, DFG-SPP 1324, January 2014.

[157] M. Giles, T. Nagapetyan, and K. Ritter. Multi-Level Monte Carlo Approximation of Distribution Functions and Densities. Preprint 157, DFG-SPP 1324, February 2014.

[158] F. Dickmann and N. Schweizer. Faster Comparison of Stopping Times by Nested Conditional Monte Carlo . Preprint 158, DFG-SPP 1324, February 2014.

[159] L. Kämmerer, D. Potts, and T. Volkmer. Approximation of multivariate periodic functions by trigonometric polynomials based on sampling along rank-1 lattice with generating vector of Korobov form . Preprint 159, DFG-SPP 1324, February 2014.

[160] S. Dereich and T. Müller-Gronbach. Quadrature for Self-Affine Distributions on $\mathbb{R}^d$. Preprint 160, DFG-SPP 1324, March 2014.