Institut für Numerische Simulation
Rheinische Friedrich-Wilhelms-Universität Bonn
Wegelerstraße 6 • 53115 Bonn • Germany
phone +49 228 73-3427 • fax +49 228 73-7527
www.ins.uni-bonn.de

M. Griebel, J. Oettershagen

Dimension-adaptive sparse grid quadrature for integrals with boundary singularities

INS Preprint No. 1310

2013

Dimension-adaptive sparse grid quadrature for integrals with boundary singularities

Michael Griebel and Jens Oettershagen

Abstract Classical Gaussian quadrature rules achieve exponential convergence for univariate functions that are infinitly smooth and where all derivatives are uniformly bounded. The aim of this paper is to construct generalized Gaussian quadrature rules based on non-polynomial basis functions, which yield exponential convergence even for integrands with (integrable) boundary singularities whose exact type is not a-priori known. Moreover, we use sparse tensor-products of these new formulae to compute d-dimensional integrands with boundary singularities by means of a dimension-adaptive approach. As application, we consider, besides standard model problems, the approximation of multivariate normal probabilities using the Genz-algorithm.

1 Introduction

The approximation of an integral of a function $f : \Omega^{(d)} \to \mathbb{R}, \Omega^{(d)} \subseteq \mathbb{R}^d$ using point-evaluations is an important task in numerical analysis. It is part of numerous methods and algorithms in almost every scientific area where computers are employed. Such a quadrature rule takes the general form

$$\int_{\Omega^{(d)}} f(\mathbf{x})\, \omega(\mathbf{x})\, \mathrm{d}\mathbf{x} \approx \sum_{i=1}^{n} w_i f(\mathbf{x}_i), \tag{1}$$

where $\omega : \Omega^{(d)} \to \mathbb{R}^+$ is a positive weight-function, the $\mathbf{x}_i \in \Omega^{(d)}$ are the quadrature nodes and the $w_i \in \mathbb{R}$ are the quadrature weights. The quality of such an approximation depends on the regularity of f and the specific choice of the nodes and weights of the quadrature formula. Most quadrature rules are constructed in such a way that

Michael Griebel
Institute for Numerical Simulation, Bonn, e-mail: griebel@ins.uni-bonn.de

Jens Oettershagen
Institute for Numerical Simulation, Bonn, e-mail: oettershagen@ins.uni-bonn.de

they are exact on a certain finite-dimensional subspace of $L_1(\Omega^{(d)}, \omega)$, e.g., on the polynomials up to a certain degree.

In this paper we will consider integration over the open unit cube $\Omega^{(d)} = (0,1)^d$ with respect to the uniform Lebesgue measure $\omega \equiv 1$. We treat integrands that might possess integrable boundary singularities, whose exact type is not a-priori known[1].

A common approach to deal with boundary singularities in the univariate setting are variable transformations $\tau : \mathbb{R} \supseteq \hat{\Omega} \to (0,1)$, such that

$$\int_0^1 f(x)\,\mathrm{d}x = \int_{\hat{\Omega}} f \circ \tau(y) \cdot \tau'(y)\,\mathrm{d}y.$$

If τ is properly chosen, the transformed integrand $\hat{f}(y) = f \circ \tau(y) \cdot \tau'(y)$ is no longer singular, but decays to zero as y approaches the boundary of $\hat{\Omega}$. Popular examples are the tanh [44] and double exponential [32, 45] transforms, where a truncated trapezoidal rule is applied to $\hat{f}$ on $\hat{\Omega} = \mathbb{R}$. Other examples are the so-called periodization transforms [27, 37, 39, 40], which use a mapping to $\hat{\Omega} = (0,1)$ that makes the resulting integrand periodic. Eventhough these approaches work very well in the univariate setting, their multivariate versions suffer from an exponential blowup of the norm of the transformed integrand $\hat{f}$, as it was pointed out in [27] for the case of periodizing transformations in the context of lattice rules. Thus it takes exponentially long (in d), until the good asymptotic convergence rate kicks in, which makes these approaches very costly for practial problems in higher dimensions.

For this reason we are not going to apply a trapezoidal rule to the transformed integrand $\hat{f}$, but rather use a suitable Gaussian rule tailored to the weight function $\tau'(y)$ on $\hat{\Omega}$ and then map back the associated Gaussian quadrature nodes to the unit interval $(0,1)$. This approach results in a so-called generalized Gaussian quadrature rule on $(0,1)$ that is exact not on a space of polynomials as are conventional Gaussian rules, but on a $2n$-dimensional subspace of univariate *singular* functions from $L_1(0,1)$. Its basis functions are given by powers of certain monotonous singular functions. We will prove exponential convergence for integrands with arbitrary algebraic boundary singularities. Moreover, we explicitly compute error constants for quadrature on $(0,1)$ in the Hardy space of functions that are analytic in the unit disc. In contrast to Gauss-Legendre quadrature, which only achieves an algebraic rate of convergence, our approach shows an exponential decay of the error. For the higher dimensional case we then employ these univariate quadrature rules within a sparse tensor product construction which also exhibits a certain degree of exactness on tensor products of the univariate singular functions. We give numerical evidence that, for singular problems in the unit cube, our approach significantly outperforms the conventional sparse grid quadrature methods which are based on the Gauss-Legendre or Clenshaw-Curtis rules, respectivley. Furthermore, we use our new method in combination with dimension-adaptive sparse grids for various standard model problems and for the computation of multivariate normal probabilities by the Genz-algorithm [12].

[1] Otherwise one could reformulate the problem to integration with respect to a weight function ω that resembles the singularity.

The remainder of this article is organized as follows: In Section 2 we will shortly revise the classical (univariate) Gaussian quadrature and its generalization to Tschebyscheff-systems. Then we introduce certain classes of Tschebyscheff-systems, whose associated Gaussian quadrature formulae can be described in terms of classical Gaussian formulae on an unbounded domain $\widehat{\Omega}$ and a mapping back to $(0,1)$. This allows for an easy construction of the new generalized Gaussian quadrature. We give a few examples for this approach and prove error bounds for a special case which is related to the classical Gauss-Laguerre formula. In Section 3 we will introduce sparse tensor products of the new univariate quadrature to deal with multivariate problems. Here, we specifically employ the dimension-adaptive approach from [17]. In Section 4 we give the results of our numerical experiments. First, we demonstrate the quality of our generalized Gaussian quadrature formula in the univariate case for singular problems and quadrature in the Hardy space. Then we apply it within the dimension-adaptive sparse grid algorithm to several model problems and to the computation of multivariate normal probabilities using the algorithm of Genz. We close with some remarks in Section 5.

2 A generalized Gaussian quadrature approach

In this section we will introduce a class of generalized Gaussian quadrature rules that are exact on a certain $2n$-dimensional subspace of singular functions on $(0,1)$. First, we shortly recall the basic properties of classical Gaussian quadrature in the univariate case. Then we generalize it to Tschebyscheff-systems on $(0,1)$ and introduce a framework for which the generalized Gaussian quadrature formula can be described by a classical Gaussian formula with respect to a certain weight function on $(0,\infty)$ together with a map to $(0,1)$. We give examples and an error estimate for a class of integrands with algebraic singularities at the boundary.

2.1 Classical Gaussian quadrature

Given an interval $\Omega \subseteq \mathbb{R}$ and a positive weight function $\omega : \Omega \to \mathbb{R}^+$ the classical n-point Gaussian rule is the standard approach to approximate

$$\int_\Omega f(x)\omega(x)\,\mathrm{d}x \approx \sum_{i=1}^{n} w_i f(x_i)$$

for smooth functions $f \in C^{2n}(\Omega)$. Classical Gaussian quadrature rules are defined by the property that polynomials up to degree $2n-1$ are integrated exactly (with respect to the weight function ω) with only n quadrature points, i.e.

$$\int_\Omega p(x)\omega(x)\,\mathrm{d}x = \sum_{i=1}^{n} w_i p(x_i) \tag{2}$$

for all $p \in \operatorname{span}\{x^k, k = 0,\dots,2n-1\}$. It is well known that there can not exist a quadrature rule that yields a higher degree of polynomial exactness, thus Gaussian quadrature is optimal in this sense.

Note that (2) is a nonlinear system of $2n$ equations which defines the nodes x_i and weights w_i of the Gaussian quadrature rule. In general, the direct solution of a nonlinear system is a difficult task. But here one can resort to orthogonal polynomials and then use their recurrence relation to compute the nodes and weights of Gaussian quadrature formulae. To this end, we consider the set of polynomials which are orthogonal with respect to the weighted L_2 inner product

$$\langle p, q\rangle_{\Omega,\omega} = \int_\Omega p(x)q(x)\,\omega(x)\,\mathrm{d}x.$$

It is known that for any domain Ω and weight function ω there exists a polynomial p_n of degree n that fulfills

$$\int_\Omega x^k p_n(x)\omega(x)\,\mathrm{d}x = 0 \quad \text{for all } k = 0,1,\dots,n-1.$$

We call p_n the n-th degree orthogonal polynomial (with respect to Ω and ω). Note that the set of all $p_0, p_1, \dots$ is a complete orthogonal system of $L_2(\Omega,\omega)$. Moreover, p_n has exactly n distinct simple roots $x_1,\dots,x_n \in \Omega$, which turn out to be the quadrature nodes of the Gaussian formula. This is important because, for any given weight-function ω on a domain $\Omega \subseteq \mathbb{R}$, there exists a sequence of orthogonal polynomials and thus there exists a uniquely determined corresponding Gaussian quadrature rule. If one knows the coefficients of $p_n(x) := a_n x^n + a_{n-1}x^{n-1} + \ldots + a_0$ and its zeros $x_1,\dots,x_n$ it is easy to compute the corresponding weights by the formula

$$w_i = \frac{a_n}{a_{n-1}} \frac{\langle p_{n-1}, p_{n-1}\rangle_{\Omega,\omega}}{p_n'(x_i)p_{n-1}(x_i)} \geq 0. \tag{3}$$

In order to obtain the coefficients and roots of such orthogonal polynomials one usually employs their recurrence formula to assemble a so-called compagnion matrix, whose eigenvalues are the roots of p_n. For details see [7, 11].

Now we consider error bounds for classical Gaussian quadrature. First we define the n-th error functional as

$$R_n(f) := \int_\Omega f(x)\,\omega(x)\,\mathrm{d}x - \sum_{i=1}^{n} w_i f(x_i). \tag{4}$$

Because of the Weierstrass approximation theorem and the positivity of the quadrature weights w_i, Gaussian quadrature rules converge for any continuous function, i.e.

$$f \in C^0(\Omega) \Rightarrow \lim_{n\to\infty} R_n(f) = 0. \tag{5}$$

For integrands which possess at least $2n$ continuous derivatives there is the well-known error-bound

$$|R_n(f)| \leq \frac{f^{(2n)}(\xi)}{a_n^2(2n)!} \quad \text{for some } \xi \in \Omega. \tag{6}$$

Thus, if for all $n \in \mathbb{N}$ it holds $|f^{(2n)}(x)| \leq M_n(f) \forall x \in \Omega$ where $\frac{M_n(f)}{a_n^2}$ is bounded by a polynomial, the quantity $|R_n(f)|$ converges even exponentially to zero.

On the other hand, if the derivatives of f are unbounded on Ω the bound (6) is useless. For example, for $\Omega = (-1,1)$ and $\omega(x) = 1$ one obtains the well-known Gauss-Legendre rule, for which the bound (6) takes the form [7]

$$|R_n(f)| \leq f^{(2n)}(\xi) \cdot \frac{(n!)^4}{(2n+1)((2n)!)^3}, \quad \xi \in (-1,1).$$

Now consider $f(x) = (1-x)^{-\alpha}, \alpha > 0$, which is unbounded on Ω as are all of its derivatives. In [9, 28] it was shown that the rate of convergence substantialy deteriorates with α, i.e.

$$|R_n(f)| = \mathcal{O}(n^{-1+\alpha})$$

and we only obtain an algebraic convergence rate of $(1-\alpha)$ for $\alpha < 1$, while for $\alpha \geq 1$ the integral does not exist anyway. This simple example shows that classical Gaussian rules lose their nice convergence properties when it comes to singular integrands. To deal with such integration problems efficiently, the Gaussian approach must be properly generalized.

2.2 Generalized Gaussian quadrature

The aim of this section is to find quadrature rules that achieve a maximum degree of exactness for systems of functions that are not polynomials, but inherently possess integrable singularities. To this end, we will use the notion of so-called *Tschebyscheff systems* (T-systems) [25].

For $n+1$ functions $\varphi_0, \ldots, \varphi_n$ and $n+1$ pairwise distinct points $t_0, \ldots, t_n \in [a,b]$ we define the generalized Vandermonde determinant as

$$\mathcal{D}(\varphi_0, \ldots, \varphi_n; t_0, \ldots, t_n) := \det \begin{pmatrix} \varphi_0(t_0) & \varphi_0(t_1) & \ldots & \varphi_0(t_n) \\ \varphi_1(t_0) & \varphi_1(t_1) & \ldots & \varphi_1(t_n) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_n(t_0) & \varphi_n(t_1) & \ldots & \varphi_n(t_n) \end{pmatrix} \tag{7}$$

Definition 1. A set of continuous functions $\varphi_0, \varphi_1, \ldots, \varphi_n$ is called *T-system* over a compact interval $[a,b]$ iff

$$\mathcal{D}(\varphi_0, \varphi_1, \ldots, \varphi_n; t_0, t_1, \ldots, t_n) > 0$$

for all pairwise distinct $t_0 < t_1 < \ldots < t_n \in [a,b]$.

This generalizes the concept of polynomials in the sense that it allows unique interpolation by cardinal functions and thus every linear combination $\sum_{i=0}^{n} c_i \varphi_i(x)$ has at most n zeros.

Examples for sets of functions that form a T-system on their respective domains are of course polynomials $\{1,x,x^2,x^3,\ldots,x^n\}$, fractional polynomials like $\{1,\sqrt{x},x,x\sqrt{x},x^2,x^2\sqrt{x},\ldots,x^{n/2}\}$ or certain radial basis functions like, e.g., $\{\exp(-\frac{(c_0-x)^2}{\sigma^2}),\ldots,\exp(-\frac{(c_n-x)^2}{\sigma^2})\}$ for pairwise distinct $c_i \in \mathbb{R}$.

We are now in the position to define the concept of generalized Gaussian quadrature.

Definition 2. (Generalized Gaussian quadrature rule)

Let $\Phi = \{\varphi_0,\ldots,\varphi_{2n-1}\}$ be a set of integrable functions on a compact interval $[a,b]$, i.e. $\int_a^b \varphi_i\,\omega(x)\,\mathrm{d}x < \infty$. A *generalized Gaussian quadrature rule* on $[a,b]$ is a n-point rule that is exact on Φ, i.e.

$$\sum_{i=1}^{n} w_i \varphi_j(x_i) = \int_a^b \varphi_j(x)\,\mathrm{d}x, \quad \text{for all } j = 0,\ldots,2n-1. \tag{8}$$

We have the following result from [25].

Theorem 1. *Let $\Phi = \{\varphi_0,\ldots,\varphi_{2n-1}\}$ be a T-system of integrable functions on the bounded interval $[a,b] \subsetneq \mathbb{R}$. Then there exists a generalized Gaussian quadrature rule with n nodes $x_1,\ldots,x_n \in (a,b)$ and non-negative weights $w_1,\ldots,w_n$.*

This result was generalized in [29] to the case of (semi-)open intervals.

Theorem 2. *If Φ constitutes a T-system on any closed interval $[\hat{a},\hat{b}] \subset (a,b) \subseteq \mathbb{R}$, then we call Φ a T-system on (a,b) and there exist n nodes $x_1,\ldots,x_n \in (a,b)$ and non-negative weights $w_1,\ldots,w_n$, such that* (8) *holds.*

Moreover, it is also known [19] that generalized Gaussian quadrature formulae are unique. The determination of a specific generalized Gaussian quadrature formula involves the problem of finding n nodes $(x_1,\ldots,x_n)$ and weights $(w_1,\ldots,w_n)$ such that (8) holds. This is a system of $2n$ equations in both $\mathbf{w} \in \mathbb{R}^n$ and $\mathbf{x} \in (a,b)^n$, which is highly non-linear in $x_1,\ldots,x_n$. Once more, its solution is in general a difficult task which is adressed in [29]. To avoid this issue, we will restrict ourselves in the following to a certain class of T-systems which involves powers of singular functions. This will allow us to resort to orthogonal polynomials again.

2.3 A certain class of singular Tschebyscheff-systems

In the following we will restrict ourselves to $\Omega = (a,b) = (0,1)$ for the sake of simplicity. But all results easily translate to arbitrary finite intervals (a,b) by affine linear dilation and translation.

We propose the follwing T-system on $\Omega = (0,1)$, which (as we will see) possesses a structure that reduces the solution of (8) to a certain classical Gaussian quadrature rule on $\widehat{\Omega} = (0,\infty)$.

To this end, let ψ be a $C^1(0,1)$ function that fulfills the following conditions:

1. $\psi(0) = 0$ and $\lim_{x\to 1} \psi(x) = \infty$.
2. $\int_0^1 \psi(x)^j \,\mathrm{d}x < \infty$, for all $j \in \mathbb{N}_0$.
3. ψ is strictly monotonous increasing, i.e. $\psi' > 0$.

Remark 1. From the conditions 1.–3. the following results can be derived:

- $\psi : [0,1) \leftrightarrow [0,\infty)$ is a C^1-bijection.
- $\lim_{y\to\infty} \psi^{-1}(y) = 1$.
- $\int_0^\infty \psi'(\psi^{-1}(y))^{-1} \,\mathrm{d}y = 1$.
- ψ is the inverse of a cumulative distribution function whose underlying distribution has finite moments on $[0,\infty)$.

Since polynomials form a T-system over any subset of $\mathbb{R}$, the following lemma proves, that

$$\varphi_j(x) := \psi(x)^j, \quad j = 0,1,2,\ldots \tag{9}$$

is a complete T-system over $(0,1)$, if ψ fulfills the conditions 1.–3.

Lemma 1. *If $\widehat{\varphi}_0,\ldots,\widehat{\varphi}_n$ is a T-system on some domain $\widehat{\Omega} \subseteq \mathbb{R}$ and $\psi : \Omega \leftrightarrow \widehat{\Omega}$ is a bijection between $\Omega \subseteq \mathbb{R}$ and $\widehat{\Omega}$, then the set*

$$\varphi_j := \widehat{\varphi}_j \circ \psi : \Omega \to \mathbb{R}, \quad j = 0,1,\ldots,n$$

is a T-system on Ω.

Proof. Suppose that there exist pairwise distinct points $t_0,\ldots,t_n \in \Omega$, such that $\mathcal{D}(\varphi_0,\ldots,\varphi_n;t_0,\ldots,t_n) \le 0$ holds. Because of the bijectivity of ψ there are points $\widehat{t}_j := \psi(t_j) \in \widehat{\Omega}$, such that $\mathcal{D}(\widehat{\varphi}_0,\ldots,\widehat{\varphi}_n;\widehat{t}_0,\ldots,\widehat{t}_n) \le 0$, which is a contradiction to the assumption that $\widehat{\varphi}_0,\ldots,\widehat{\varphi}_n$ is a T-system on $\widehat{\Omega}$. □

Next we will describe the relationship between the generalized Gaussian quadrature with respect to the system $\Phi = \{\psi(\cdot)^j\}_{j=0}^\infty$ on $\Omega = (0,1)$ and classical Gaussian quadrature on the unbounded domain $\widehat{\Omega} = (0,\infty)$. To this end, we set

$$\omega(y) := \frac{d}{dy}\psi^{-1}(y) = \frac{1}{\psi'(\psi^{-1}(y))},$$

which is non-negative on $\widehat{\Omega}$. We know that there exists a sequence of polynomials $p_0, p_1, p_2, \ldots$ on $\widehat{\Omega}$ which are orthogonal with respect to ω, i.e. $\langle p_i, p_j\rangle_{\widehat{\Omega},\omega} = \delta_{i,j}$.

Remark 2. The orthogonality of the p_i translates to the set of functions $q_j : (0,1) \to \mathbb{R}$

$$q_j(x) := p_j \circ \psi(x), \quad j = 0,1,2,\ldots,$$

i.e.

$$\int_0^1 q_i(x)q_j(x)\,\mathrm{d}x = \int_0^\infty p_i(y)p_j(y)\,\omega(y)dy = \delta_{i,j}.$$

Analogously, the n distinct zeros $y_j \in \widehat{\Omega}$ of p_n carry over to the zeros $x_j \in (0,1)$ of q_n as

$$x_j := \psi^{-1}(y_j), \tag{10}$$

since $q_n(x_j) = p_n(y_j) = 0$.

We finally arrive at the following result:

Theorem 3. *With $\Omega = (0,1)$ and $\widehat{\Omega} = (0,\infty)$ let $\psi : \Omega \to \widehat{\Omega}$ fulfill the conditions 1.–3., and let $y_j \in \hat{\Omega}$, $w_j \in \mathbb{R}^+$ ($j = 1,\dots,n$) be the nodes and weights of the classical Gaussian quadrature rule on $\widehat{\Omega}$ with respect to the weight function $\omega(y) = \frac{d}{dy}\psi^{-1}(y)$. Then the quadrature rule*

$$Q_n(f) := \sum_{j=1}^n w_j f(x_j), \quad \textit{where } x_j = \psi^{-1}(y_j) \in (0,1), \tag{11}$$

is exact on the span of $\Phi = \{\varphi_0,\dots,\varphi_{2n-1}\}$ defined in (9)*, i.e.*

$$Q_n(\varphi_j) = \int_0^1 \varphi_j(x)\,\mathrm{d}x, \quad \textit{for } j = 0,\dots,2n-1.$$

Proof. Because of $\varphi_k(x_j) = y_j^k$ and

$$\sum_{j=1}^n w_j y_j^k = \int_0^\infty y^k \omega(y)\,dy = \int_0^1 \varphi_k(x)\,\mathrm{d}x$$

it holds for $k = 0,\dots,2n-1$ that

$$Q_n(\varphi_k) = \sum_{j=1}^n w_j \varphi_k(x_j) = \int_0^1 \varphi_k(x)\,\mathrm{d}x.$$

□

By now we have shown that the generalized n-point Gaussian quadrature with respect to the set of singular functions $\varphi_0, \varphi_1, \dots, \varphi_{2n-1}$ from (9) can simply be computed by mapping the nodes of a certain classical Gaussian quadrature in $(0,\infty)$ back to $(0,1)$. Moreover, because of (5) the quadrature rule (11) converges for any continuous function on $(0,1)$. This kind of quadrature rule is especially suited for integrands on $(0,1)$ with a boundary singularity located at $x = 1$. It is possible to extend this approach to integrands with singularities at both endpoints of the domain $(-1,1)$.

Remark 3. If we change the first and second condition to

$1'$. $\lim_{x\to\pm 1} \psi(x) = \pm\infty$.
$2'$. $\int_{-1}^1 \psi(x)^j\,\mathrm{d}x < \infty$, for all $j \in \mathbb{N}_0$.

we obtain a T-system of functions which have singularities at both $x = -1$ and $x = +1$. Moreover, it follows that

- $\psi : (-1,1) \leftrightarrow (-\infty,\infty)$ is a C^1 bijection.
- $\lim_{x\to\pm\infty} \psi^{-1}(x) = \pm 1$.
- $\int_{-\infty}^{\infty} \psi'(\psi^{-1}(y))^{-1}\,\mathrm{d}y = 2$.
- ψ is the inverse of a cumulative distribution function whose underlying distribution has finite moments on $(-\infty,\infty)$.

All previous statements also hold true for this case, with obvious modifications in their proofs.

2.4 Examples

We now give three examples for the choice of the function ψ, where the first one relates to the classical Gauss-Laguerre quadrature on $(0,\infty)$ with respect to the weight-function $\omega(y) = e^{-y}$ and the second one to a non-classical rule with respect to $\omega(y) = \cosh(y)\cdot\cosh(\sinh(y))^{-2}$ on $(0,\infty)$ as well. The third example refers to Remark 3 and is related to the classical Gauss-Hermite quadrature on the double-infinite interval $(-\infty,\infty)$ with respect to the weight function $\omega(y) = \exp(-y^2)$.

2.4.1 Powers of logarithms

Our first example involves the classical Gauss-Laguerre quadrature rule, which is well-known and its nodes and weights are available from numerous libraries.

We choose $\psi : (0,1) \to (0,\infty)$ as

$$\psi_{\log}(x) := -\log(1-x).$$

With (9), this results in the T-system defined by

$$\varphi_k(x) := (-\log(1-x))^k, \quad k = 0,1,2,\ldots. \tag{12}$$

The inverse is given as $\psi^{-1}(y) = 1 - \exp(-y)$ and its derivative is

$$\omega(y) := \frac{d}{dy}\psi^{-1}(y) = \exp(-y). \tag{13}$$

Because of $\int_0^1 (-\log(1-x))^k = k!$ all φ_k are integrable. Thus the conditions 1.–3. are fulfilled and Theorem 3 relates to the well-known Laguerre polynomials which define the Gauss-Laguerre quadrature rules on $(0,\infty)$ with respect to the weight-function $\omega(y) = e^{-y}$. Let w_i and y_i denote the n quadrature weigths and nodes of the Gauss-Laguerre formula on $(0,\infty)$. Then, by mapping the Gaussian nodes y_i back to $(0,1)$, we obtain the quadrature rule

$$\int_0^1 f(x)\,\mathrm{d}x \approx \sum_{i=1}^{n} w_i f(x_i), \qquad \text{with } x_i = 1-\exp(-y_i) \tag{14}$$

which is exact for $\varphi_0, \varphi_1, \ldots, \varphi_{2n-1}$ defined in (12) on the interval $(0,1)$. Moreover, one can prove that the generalized Gaussian quadrature (11) on $(0,1)$ with φ_i given by (12) achieves basically exponential convergence for certain integrands with algebraic singularities.

Theorem 4. *Let $f(x) = (1-x)^{-\alpha}$ with $\alpha < 1$. Then, the generalized Gaussian formula with respect to the T-system* (12) *converges faster than any polynomial, i.e.*

$$\left| \int_0^1 f(x)\,\mathrm{d}x - \sum_{i=1}^{n} w_i f(x_i) \right| \leq c_{\alpha,k} n^{-k},$$

where $c_{\alpha,k}$ is a constant depending on both α and k, but not on n.

Proof. By Theorem 1 from [30] the quadrature error of the Gauss-Laguerre formula for a k-times continuous differentiable function g is bounded by

$$\left| \int_0^\infty g(y)\, e^{-y}\, dy - \sum_{i=1}^{n} w_i g(y_i) \right| \leq c n^{-k} \cdot \int_0^\infty |y^{\frac{k}{2}} g^{(k)}(y)| e^{-y}\, dy. \tag{15}$$

Mapping this result to the unit interval yields

$$\begin{aligned}
\left| \int_0^1 f(x)\,\mathrm{d}x - \sum_{i=1}^{n} w_i f(x_i) \right| &= \left| \int_0^\infty f(1-e^{-y})\, e^{-y}\, dy - \sum_{i=1}^{n} w_i f(1-e^{-y_i}) \right| \\
&= \left| \int_0^\infty e^{\alpha y}\, e^{-y}\, dy - \sum_{i=1}^{n} w_i e^{\alpha y_i} \right| \\
&\overset{(15)}{\leq} c n^{-k} \cdot \alpha^k \int_0^\infty |y^{\frac{k}{2}} e^{\alpha y}| e^{-y}\, dy \\
&= c n^{-k} \cdot \frac{\alpha^k}{(1-\alpha)^{\frac{k}{2}+1}} \Gamma\left(\frac{k}{2}+1\right)
\end{aligned}$$

which holds for any $k = 1, 2, \ldots$. □

Recall that for this class of integrands the Gauss-Legendre quadrature only achieves an algebraic rate of convergence of $n^{-1+\alpha}$.

2.4.2 Powers of inverse hyperbolic functions

Next, we consider a choice for ψ that consists of certain hyperbolic functions. It is given by

$$\psi_{\mathrm{hyp}}(x) := \operatorname{arc\,sinh}\left(\frac{2}{\pi}\operatorname{arc\,tanh}(x)\right), \quad x \in (0,1), \tag{16}$$

which is inspired by the so-called double exponential (DE) quadrature [32, 45] that has gained substantial interest within the last years. It leads to the problem of constructing a Gaussian quadrature rule for the weight function

$$\omega(y) = \frac{\cosh(y)}{\cosh(\frac{\pi}{2}\sinh(y))^2}$$

on the infinite interval $(0,\infty)$. For this purpose we use the algorithm proposed in [11] and map the resulting Gaussian nodes $y_1,\dots,y_n$ back to $(0,1)$. This results in the quadrature formula

$$\int_0^1 f(x)\,\mathrm{d}x \approx \sum_{i=1}^{n} w_i f(x_i), \qquad \text{with } x_i = \tanh\left(\frac{\pi}{2}\sinh(y_i)\right) \tag{17}$$

which is exact on the T-system

$$\varphi_k(x) = \operatorname{arc\,sinh}\left(\frac{2}{\pi}\operatorname{arc\,tanh}(x)\right)^k, \quad k = 0,1,\dots 2n-1.$$

Note that the predecessor of the DE-rule was the Tanh-quadrature which was introduced in [44]. In our setting it relates to

$$\psi(x) = \operatorname{arc\,tanh}(x)$$

which also fulfills the conditions 1.–3. and leads to orthogonal polynomials with respect to

$$\omega(y) = \cosh(y)^{-2}.$$

Remark 4. Both the double exponential and the tanh approach rely on the quick decay of $f\circ\psi^{-1}(y)\cdot D\psi^{-1}(y)$ as $y\to\pm\infty$, which allows for an exponential rate of convergence for the trapezoidal rule. This approach does not work well in a multivariate setting, because the factor $\prod_{j=1}^{d} D\psi^{-1}(y_j)$ exponentially blows up the norm of the transformed integrand [27]. Our approach is different in the sense that we do not apply a trapezoidal rule directly to $f\circ\psi^{-1}(y)\cdot D\psi^{-1}(y)$ but use a Gaussian rule tailored to $D\psi^{-1}$, which is applied to $f\circ\psi^{-1}$ only.

2.4.3 Powers of the inverse error function

Our third example illustrates Remark 3 and relates to Gauss-Hermite quadrature on the whole $\mathbb{R}$. We choose $\psi:(-1,1)\to(-\infty,\infty)$ as

$$\psi_{\mathrm{erf}}(x) := \mathrm{erf}^{-1}(x),$$

where $\mathrm{erf}(x)$ denotes the error-function[2]. This leads to the T-system of functions

[2] The error-function $\mathrm{erf}:\mathbb{R}\to(-1,1)$ is defined as $\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}\,\mathrm{d}t$.

$$\varphi_k(x) := \left(\mathrm{erf}^{-1}(x)\right)^k, \quad k = 0,1,2,\ldots \tag{18}$$

which have singularities at both $x = -1$ and $x = 1$. Since $\psi(x)^k$ is integrable on $(-1,1)$ for all $k = 0,1,\ldots$, the conditions $1'$.–$3'$. are fulfilled and Theorem 3 relates to the well-known Hermite polynomials which define the Gauss-Hermite quadrature rules on $(-\infty,\infty)$ with respect to $\omega(y) = \frac{2}{\sqrt{\pi}} e^{-y^2}$. Then, if w_i and y_i denote the weights and nodes of the n-point Gauss-Hermite quadrature on $\mathbb{R}$, the resulting quadrature rule on $(-1,1)$ with respect to the T-system (18) is given by

$$\int_{-1}^{1} f(x)\,\mathrm{d}x \approx \sum_{i=1}^{n} w_i f(x_i), \qquad \text{with } x_i = \mathrm{erf}(y_i).$$

Note that it is possible to derive an analogous error bound as in Theorem 4 by an estimate for Gauss-Hermite quadrature which is given in [30].

3 Dimension-adaptive sparse grids

In this section we recall the sparse grid method for the integration of multivariate functions. Sparse grid quadrature formulae are constructed using certain combinations of tensor products of one-dimensional quadrature rules, see, e.g., [6, 15, 16, 20, 21, 36, 43, 46]. This way, sparse grid methods can exploit the mixed smoothness of f, if present, and may overcome the curse of dimension to a certain extent. Moreover, they can be employed in a dimension-adaptive fashion.

3.1 Classical construction

For a continuous univariate function $g : (0,1) \to \mathbb{R}$ let

$$Q_{n_k} g := \sum_{i=1}^{n_k} w_{i,k}\, g(x_{i,k})$$

denote a sequence of univariate quadrature rules with $n_{k+1} > n_k$, $Q_0 f = 0$ and $Q_{n_k} g \to \int_0^1 g(x)\,\mathrm{d}x$ for $k \to \infty$. Using the difference quadrature formulae

$$\Delta_k = Q_{n_k} - Q_{n_{k-1}}$$

one has

$$\sum_{k=1}^{\infty} \Delta_k g = \int_0^1 g(x)\,\mathrm{d}x.$$

Then, for a d-variate function $f : (0,1)^d \to \mathbb{R}$, its integral can be represented by the infinite telescoping series

$$\int_{(0,1)^d} g(\mathbf{x})\, d\mathbf{x} = \sum_{\mathbf{k}\in\mathbb{N}^d} \Delta_{\mathbf{k}} f \tag{19}$$

which collects the products of each possible combination of the univariate difference formula. Here, $\mathbf{k} \in \mathbb{N}^d$ denotes a multi-index with $k_j > 0$ and

$$\Delta_{\mathbf{k}} f := \left(\Delta_{k_1} \otimes \ldots \otimes \Delta_{k_d}\right) f.$$

For a given level $l \in \mathbb{N}$ and the choice $n_k = 2^k - 1$ the classical sparse grid method[3], see, e.g., [16, 36, 43], is then defined by

$$\mathrm{SG}_l f := \sum_{|\mathbf{k}|_1 \le l+d-1} \Delta_{\mathbf{k}} f \tag{20}$$

where $|\mathbf{k}|_1 := \sum_{j=1}^d k_j$. Here, from the set of all possible indices $\mathbf{k} \in \mathbb{N}^d$, only those are considered whose $|\cdot|_1$-norm is smaller than a certain value. Note that the product integration rule is recovered if the norm $|\cdot|_\infty := \max\{k_j : j = 1,\ldots,d\}$ is used for the selection of indices instead of the $|\cdot|_1$-norm in (20).

3.2 Generalized sparse grids

The sparse grid construction can be tailored to certain classes of integrands if some information on the importance of the dimensions or the importance of the interactions between the dimensions is a priori known. This is achieved by choosing appropriate finite index sets $\mathcal{I} \subset \mathbb{N}^d$ in the representation (19) such that a given accuracy is attained with as few as possible function evaluations.

To ensure the validity of the hierarchical expansion the index set $\mathcal{I}$ has to satisfy the admissibility condition

$$\mathbf{k} \in \mathcal{I} \text{ and } \mathbf{l} \le \mathbf{k} \Rightarrow \mathbf{l} \in \mathcal{I}.$$

In this way, the generalized sparse grid method

$$\mathrm{SG}_{\mathcal{I}} f := \sum_{\mathbf{k}\in\mathcal{I}} \Delta_{\mathbf{k}} f \tag{21}$$

is defined, see, e.g., [16]. Note that the product rule, the classical sparse grid construction (20), sparse grids with delayed basis sequences [38] or nonisotropic sparse grids based on the weighted norms $|\mathbf{k}|_{1,\mathbf{a}} := \sum_{j=1}^d a_j k_j$ with weight factor $\mathbf{a} \in \mathbb{R}_+^d$ for the different coordinate directions [16, 21] are just special cases of this general approach.

[3] Often denoted as Smolyak's method, see [43].

Algorithm 1: Dimension-adaptive construction of the index set $\mathcal{I}$.

Initialize:

1. set of active indices: $\mathcal{I} = (1,\dots,1)$.
2. $s = \Delta_{(1,\dots,1)} f$.

repeat

1. Determine the set of admissible indices $\mathcal{A} = \{\mathcal{I} + \mathbf{e}_i : i = 1,\dots,d\}$.
2. For all $\mathbf{k} \in \mathcal{A}$ compute $\Delta_{\mathbf{k}} f$.
3. Determine (some) $\hat{\mathbf{k}} = \arg\max_{\mathbf{k}\in\mathcal{A}} \Delta_{\mathbf{k}} f$.
4. Add the index $\hat{\mathbf{k}}$ to $\mathcal{I}$.
5. Update the sum $s = s + \Delta_{\hat{\mathbf{k}}} f$.

until $|\Delta_{\hat{\mathbf{k}}} f| < \varepsilon$;

Output: $\mathrm{SG}_{\mathcal{I}} f = s$.

3.3 Dimension-adaptive sparse grids

In practice, usually no a priori information on the dimension structure of the integrand is available. In this case algorithms are required which can construct appropriate index sets $\mathcal{I}$ automatically during the actual computation. Such algorithms were presented in [17, 23] where the index sets are found in a dimension-adaptive way by the use of suitable error indicators. The adaptive methods start with the smallest index set $\mathcal{I} = \{(1,\dots,1)\}$. Then, step-by-step the index $\mathbf{k}$ from the set of all admissible indices is added which has the largest value $|\Delta_{\mathbf{k}} f|$ and is therefore expected to provide the largest error reduction, see [15, 17, 21, 33] for details. Altogether, the algorithm allows for an adaptive detection of the important dimensions and heuristically constructs optimal index sets $\mathcal{I}$ in the sense of [5, 22] which is closely related to best N-term approximation [8].

For the sake of completeness, we give a simplified version[4] of the dimension-adaptive algorithm from [17, 21] in Algorithm 1. In our numerical experiments of Section 4 we will use this approach with the three generalized Gaussian quadrature rules from Subsection 2.4.

Note finally that, besides the dimension-wise adaption, also a purely local adaptivity based on the trapezoidal rule or higher order composite Newton-Cotes formulae is possible [2, 4], which leads to algebraic convergence. However, since our aim in this paper is to explicitly deal with boundary singularities by means of a special generalized Gaussian approach that allows for exponential convergence, we will stick to the dimension-adaptive approach here.

[4] The original algorithm from [17] which we employed in our computations in Section 4 uses a more sophisticated error criterion than the one described in Algorithm 1.

3.4 Degree of exactness

Now we have a look at the degree of exactness of the sparse grid method. As before, let $\Phi = \{\varphi_j\}_{j=0}^{\infty}$ be a complete Tschebyscheff-system on $(0,1)$ and let the univariate quadrature rule, on which a particular sparse grid algorithm is based, have a certain degree of exactness $\deg(n_k)$ with respect to Φ, i.e.

$$Q_{n_k}\varphi_j = \int_0^1 \varphi_j(x)\,\mathrm{d}x, \quad \text{for all } j = 0,\ldots,\deg(n_k).$$

If one defines $P_k = \operatorname{span}\{\varphi_0,\ldots,\varphi_{\deg(n_k)}\}$, the sparse grid algorithm $\mathrm{SG}_{\mathcal{I}}$ is exact on the space

$$\{P_{k_1}\otimes\ldots\otimes P_{k_d} : \mathbf{k}\in\mathcal{I}\}.$$

This is similar to a result from [36]. There it was shown that a regular Clenshaw-Curtis sparse grid (i.e. $\mathcal{I}_\ell = \{\mathbf{k}\in\mathbb{N}_+^d : |\mathbf{k}|_1 \le d+\ell-1\}$) is exact on

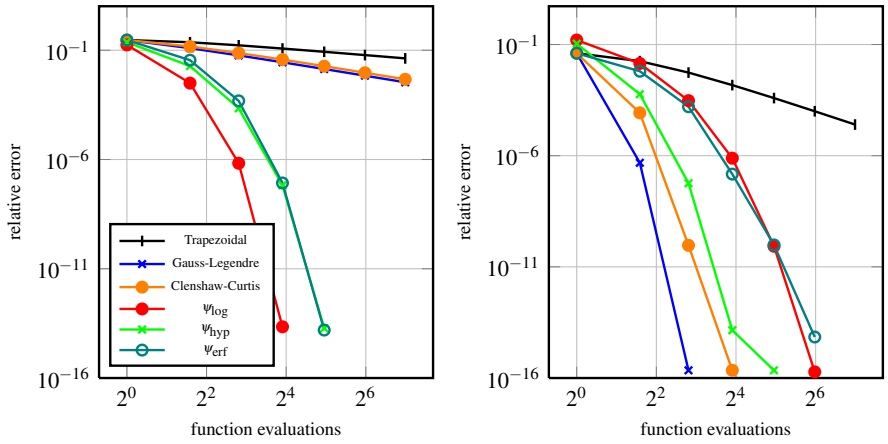
$$\{P_{k_1}\otimes\ldots\otimes P_{k_d} : |\mathbf{k}|_1 = d+\ell-1\},$$

where the φ_i are polynomials of degree i and $\deg(n_k) = n_k$. In our case, we have $\deg(n_k) = 2n-1$ since we are in a (generalized) Gaussian setting.

4 Numerical results

In this section we give results for several numerical experiments. First, we study the behaviour of our generalized Gaussian quadrature formulae in the univariate case for both, smooth and singular integrands, as well as the worst-case error in the Hardy space. Then, we deal with the higher-dimensional case where we employ the dimension-adaptive sparse grid approach that is based on our new univariate approach and compare it with dimension-adaptive sparse grids based on classical univariate rules like Gauss-Legendre or Clenshaw-Curtis quadrature. For the sake of completeness, we also compare with plain Monte Carlo and the Quasi-Monte Carlo method that is based on the Sobol sequence. Note that in all experiments the term $\psi_{\log}$ refers to the generalized Gaussian quadrature-formula with (12), while ψ_{hyp} refers to the construction with (16). ψ_{erf} refers to the construction from (18) with the additional linear transformation to the unit-interval $(0,1)$.

Note that, because of $\int_0^1 f(x)\,\mathrm{d}x = \int_0^1 f(1-x)\,\mathrm{d}x$, it is advantageous to transform integrands $f(x)$ with a singularity located at $x=+1$ to $f(1-x)$, which has the same singularity at $x=0$. Since double floating point arithmetic is more precise in a neighbourhood of 0 than in a neighbourhood of ± 1 it is possible to resolve the singularity up to a higher precision, after this transformation is employed. Of course the quadrature nodes have to undergo the same transformation, i.e. $x_i \mapsto 1-x_i$.



(a) $f(x) = (1-x)^{-1/2}$ (b) $f(x) = \exp(x)$

Fig. 1 Convergence behaviour for the singular integrand $f(x) = (1-x)^{-1/2}$ (left) and smooth integrand $f(x) = \exp(x)$ (right).

4.1 Univariate results

As our first univariate test case we consider functions with algebraic singularity of the order $0 < \alpha < 1$ at $x = 1$, i.e. we consider the problem

$$\int_0^1 \frac{1}{(1-x)^\alpha} \, \mathrm{d}x, \tag{22}$$

for various values of $0 < \alpha < 1$. As can be seen in Figure 1, the generalized Gaussian approaches achieve exponential convergence not only for the smooth example with $f(x) = \exp(x)$ but also for the singular integrand (22) with $\alpha = 1/2$. Furthermore it can be observed that the rate of convergence for the trapezoidal, the Clenshaw-Curtis and the classical Gauss-Legendre quadrature are only algebraic. We clearly see that the Gauss-Legendre rule loses its exponential convergence when it comes to the treatment of singular functions. This is however not the case for our new generalized Gaussian approach which exhibits exponential convergence for both, the smooth and the singular test function.

Next, we consider how the type of singularity of the integrand (22) affects the convergence. In Figure 2 one can see the performance of the univariate generalized Gaussian approach based on $\psi_{\log} = -\log(1-x)$ for several values of $\alpha \in (0,1)$. The convergence is always exponential, even though it takes longer to reach the asymptotic exponential regime for big values of α . Note that the integrands are not in $L^2(0,1)$ anymore for $\alpha \geq 1/2$. For the actual computation of the results displayed in Figure 2 we used long double[5] floating point arithmetic for the evaluation of

[5] 64 bit significant precision and 14 bit exponent precision.

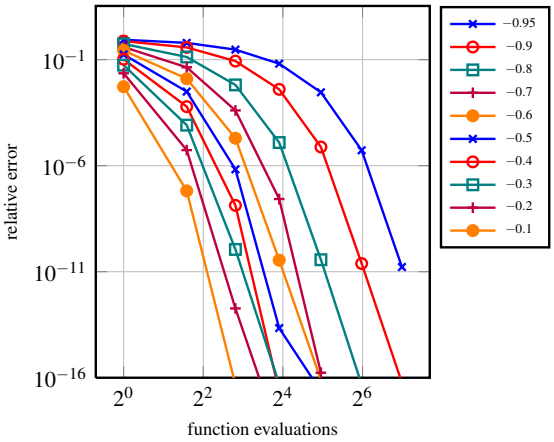


Fig. 2 Convergence behaviour of the generalized Gaussian quadrature w.r.t. $\psi_{\log}(x)$ for the integrand $f(x) = (1-x)^{\alpha}$ with algebraic singularities for different value of α.

the integrand. We stress that the quadrature nodes and weights were only stored in standard double[6] precision.

Another example concerns the quadrature error in a Hardy space. It consists of functions that are analytic in the unit-circle $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$ and is given by

$$H^2 = \{f : \mathbb{D} \to \mathbb{C} : \|f\|_{H^2}^2 < \infty\}, \quad \text{where } \|f\|_{H^2}^2 = \int_{\mathbb{D}} |f(z)|^2 \, |\mathrm{d}z|.$$

Since H^2 is a reproducing kernel Hilbert space with kernel $K(x,y) = \frac{1}{1-xy}$, see e.g. [47], the quadrature error in this space can be estimated by standard functional analysis as

$$\left| \int_0^1 f(x)\,\mathrm{d}x - \sum_{i=1}^n w_i f(x_i) \right| \leq R_n(\mathbf{x},\mathbf{w}) \cdot \|f\|_{H^2},$$

where $R_n(\mathbf{x},\mathbf{w})$ is the so-called worst-case error which depends on the nodes $\mathbf{x}$ and weights $\mathbf{w}$. It is explicitly given by

$$\begin{aligned} R_n(\mathbf{x},\mathbf{w})^2 &= \int_0^1 \int_0^1 K(s,t)\,\mathrm{d}s\,\mathrm{d}t - 2\sum_{i=1}^n w_i \int_0^1 K(t,x_i)\,\mathrm{d}t + \sum_{i=1}^n \sum_{j=1}^n w_i w_j K(x_i,x_j) \\ &= \frac{\pi^2}{6} + 2\sum_{i=1}^n w_i \frac{\log(1-x_i)}{x_i} + \sum_{i=1}^n \sum_{j=1}^n w_i w_j \frac{1}{1-x_i x_j}. \end{aligned}$$

In Figure 3 one can see that all three new methods, i.e. the generalized Gaussian quadrature with respect to $\psi_{\log}$, ψ_{hyp} and ψ_{erf}, significantly outperform polynomial-

[6] 53 bit significant precision and 10 bit exponent precision.

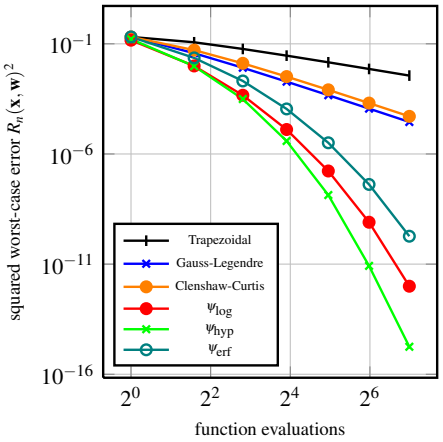


Fig. 3 Comparison of several methods for the worst-case error in the Hardy space H^2.

based methods like Gauss-Legendre as well as Clenshaw-Curtis and the trapezoidal rule in the Hardy space.

4.2 Multivariate results for standard test cases

Now we consider simple higher-dimensional model problems with a certain multiplicative and additive structure. To this end, let

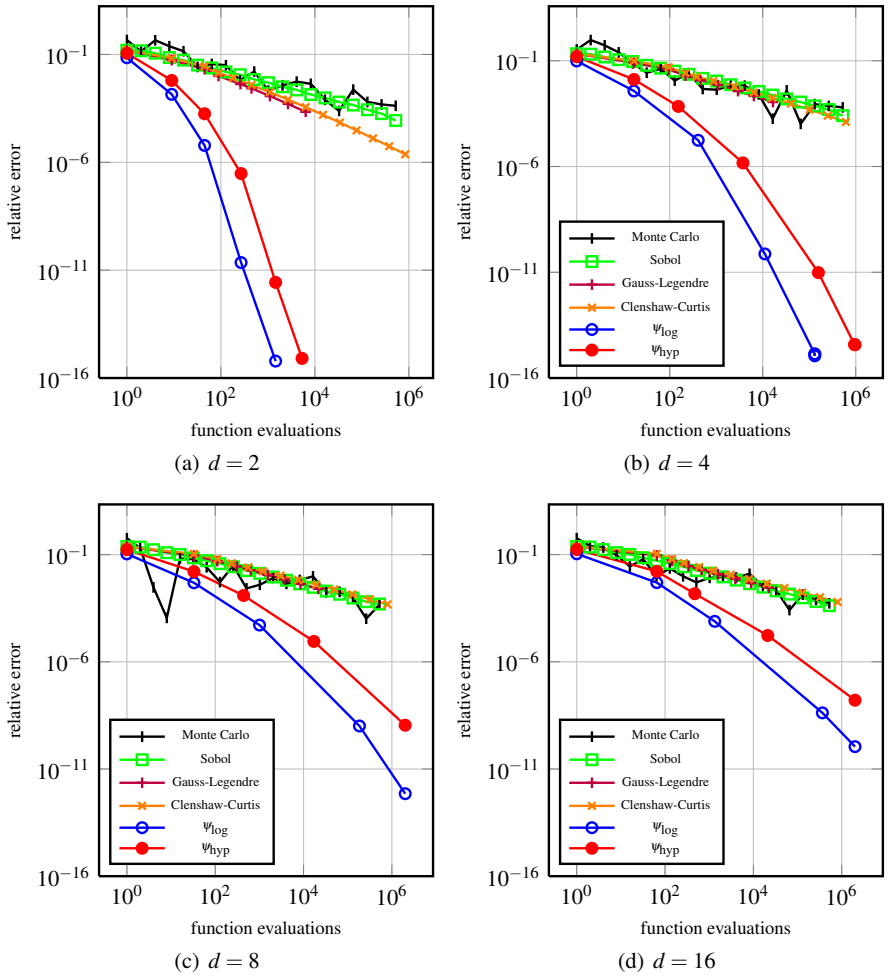
$$f(\mathbf{x}) = \sum_{\mathfrak{u}\subseteq\{1,\dots,d\}} \gamma_{\mathfrak{u}} \prod_{i\in\mathfrak{u}} \frac{1}{\sqrt[3]{1-x_i}}, \tag{23}$$

where the so-called product-weights $\gamma_{\mathfrak{u}}$ [42, 46] are defined by

$$\gamma_{\mathfrak{u}} := \prod_{i\in\mathfrak{u}} \gamma_i.$$

Here, the sequence $\gamma_1 \geq \gamma_2 \geq \ldots \geq \gamma_d > 0$ moderates the contribution of the different coordinate directions to the value of the integral.

For the special case $\gamma_i = 2^{-i}$, the resulting convergence behaviour can be seen in Figure 4 for conventional integration methods like MC and QMC and the dimension-adaptive sparse grid approach based on both, the classical Clenshaw-Curtis and Gauss-Legendre formulae and our new quadrature formula with $\psi_{\log}$ chosen as (12) and ψ_{hyp} chosen as (16), respectively. We clearly see again that the conventional methods only possess an algebraic rate of convergence, whereas our new approach with $\psi_{\log}$ and ψ_{hyp} indeed shows exponential convergence. We furthermore see that, for this particular model problem, we obtain somewhat better results for $\psi_{\log}$ than for ψ_{hyp}. For higher dimensions this exponential behaviour gets less prominent. This



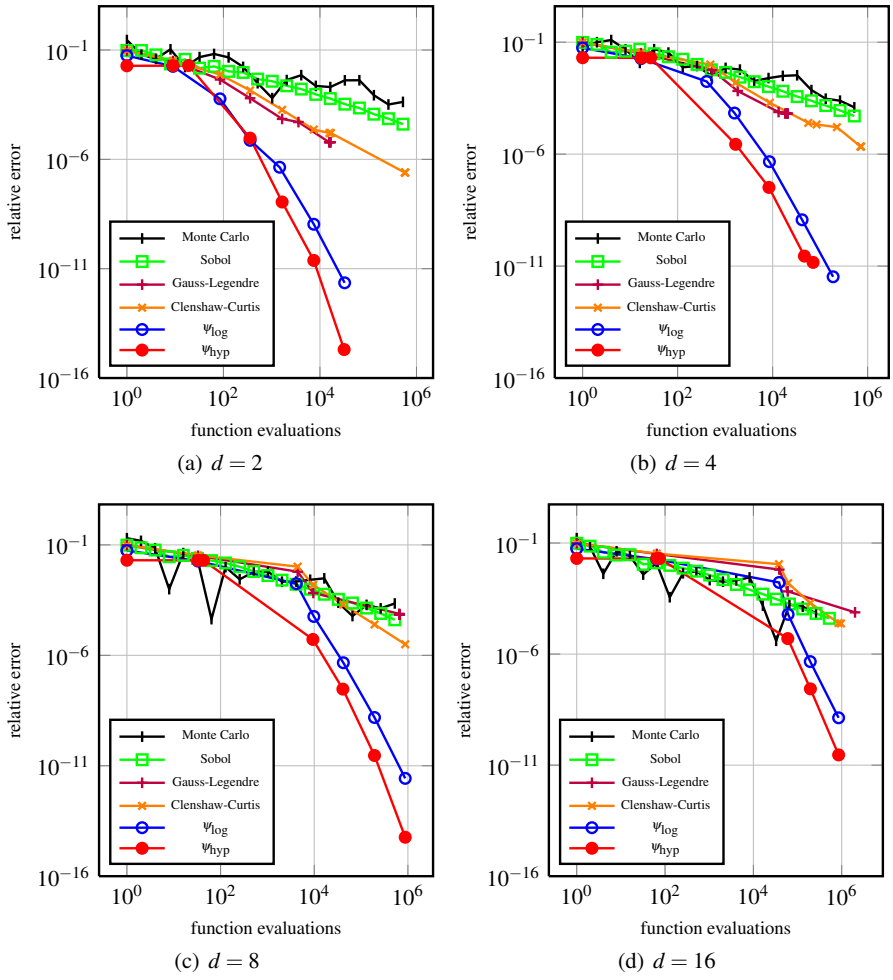
(a) $d = 2$ (b) $d = 4$

(c) $d = 8$ (d) $d = 16$

Fig. 4 Convergence behaviour for the test function with product-weights (23) with $\gamma_i = 2^{-i}$ for dimensions $d = 2, 4, 8, 16$.

stems from a delay in the onset of the asymptotic regime due to the dimension-dependent constants involved here [34]. But the superiority over the conventional methods still can be seen clearly.

As another example we will now deal with multivariate functions that can be represented as a superposition of q-dimensional functions. They are just a special case of the above framework where all $\gamma_{\mathfrak{u}} = 0$ for sets $\mathfrak{u}$ with cardinality bigger than q, i.e. $|\mathfrak{u}| > q$. This type of weights is often referred to as *finite-order weights*, see e.g. [41]. Here, we consider the test-function



(a) $d = 2$ (b) $d = 4$

(c) $d = 8$ (d) $d = 16$

Fig. 5 Convergence behaviour of the finite-order test function (24) with $q = 2$ and $d = 2, 4, 8, 16$.

$$f_{d,q}(\mathbf{x}) := \sum_{\substack{\mathfrak{u}\subset\mathcal{D}\\ |\mathfrak{u}|\le q}} \frac{1}{\sqrt{\sum_{j\in\mathfrak{u}}(1-x_j)}}. \tag{24}$$

The results are displayed in Figure 5. We basically observe a similar convergence behaviour as in the previous example. Now, however, ψ_{hyp} is slightly superior to $\psi_{\log}$. Moreover, the offset of the asymptotic convergence behaviour with respect to the dimension is more specific. The number of function evaluations, needed until the convergence rate for the dimension-adaptive sparse grid methods kicks in, depends quadratically on d. This is due to the fact that for $q = 2$ the sum in (24) consists of $\binom{d}{2}$ parts.

4.3 Computing multivariate normal probabilities by the Genz-algorithm

As a final example we consider the evaluation of multivariate normal probabilities which is an important part of many numerical methods in statistics [14], financial engineering [15], physics [31] and econometrics [1, 24]. It is defined by the integral

$$F(\mathbf{b}) := \frac{1}{\sqrt{\det(\Sigma)(2\pi)^d}} \int_{-\infty}^{b_1} \dots \int_{-\infty}^{b_d} \exp\left(-\frac{1}{2}\mathbf{x}^t \Sigma^{-1}\mathbf{x}\right) d\mathbf{x} \tag{25}$$

where $\Sigma \in \mathbb{R}^{d\times d}$ is a covariance matrix which depends on the specific problem under consideration.

For (25) it is common to use a sequence of variable transformations to obtain an integration problem that is defined on the open unit cube. This approach was independently developed by Genz [12], Geweke and Hajivassiliou [3, 18] and Keane [26]. In statistics, this method is often referred to as Genz-algorithm, while in econometrics it is called GHK-simulator. Regular sparse grids based on the Gauss-Legendre quadrature were utilized in [24] for the first time in this setting. In the following we will demonstrate that the sparse grid approach can benefit from our new univariate quadrature formulae. We remark that this appraoch can also be applied to the computation of other probabilities, e.g. the t-distribution [14].

The Genz-algorithm [12] consists of several transformations and finally leads to the integral

$$F(\mathbf{b}) = \hat{b}_1 \int_{(0,1)^{d-1}} \prod_{i=2}^{d} \hat{b}_i(w_1,\dots,w_{i-1})\, \mathrm{d}w \tag{26}$$

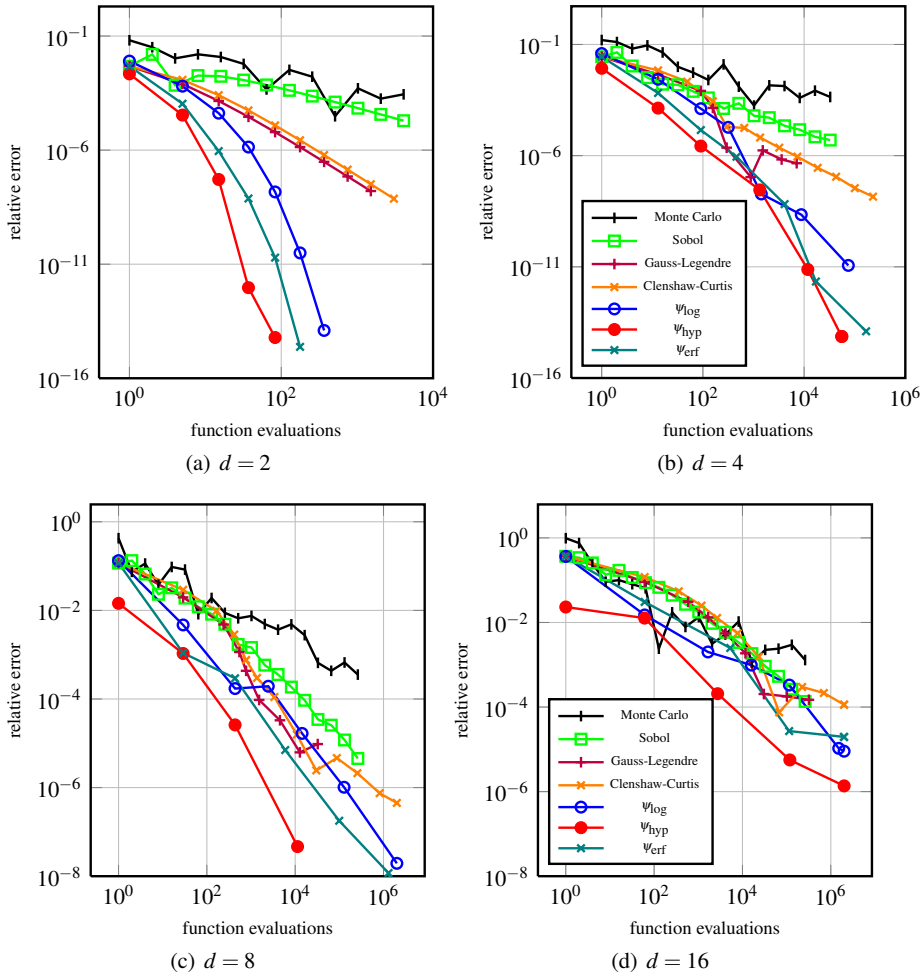
where the $\hat{b}_i$ are recursivly given by

$$\hat{b}_i(w_1,\dots,w_{i-1}) = \Phi\left(C_{i,i}^{-1}\cdot\left(b_i - \sum_{j=1}^{i-1} C_{i,j}\cdot\Phi^{-1}(w_j\cdot\hat{b}_j(w_1,\dots,w_{j-1})\right)\right).$$

Here, the matrix $\mathbf{C}\in\mathbb{R}^{d\times d}$ denotes a Cholesky factor[7] of the covariance-matrix, i.e. $\mathbf{C}\mathbf{C}^T = \Sigma$, and $\Phi : \mathbb{R}\to(0,1)$ is the cumulative Gaussian distribution function.

The main advantage of the Genz-algorithm in a dimension-adaptive sparse grid setting stems from the fact that it enforces a priority ordering onto the variables $w_1,\dots,w_{d-1}$, where w_1 contributes the most and w_{d-1} contributes the fewest to the value of $F(\mathbf{b})$. Furthermore, the dimensionality of the original integration problem is reduced by one. A disadvantge is of course the increased cost for the evaluation of the transformed integrand in formula (26). Moreover, while the original integrand was analytic in the whole complex plane, the new integrand is only analytic within the open disc $\{z\in\mathbb{C} : |z-\frac{1}{2}| < \frac{1}{2}\}$. This is due to the inverse cumulative distribution function Φ^{-1} that introduces a singularity at the origin and in some dimensions a

[7] Cholesky factorization is here only unique modulo row and column permutation.



(a) $d = 2$ (b) $d = 4$

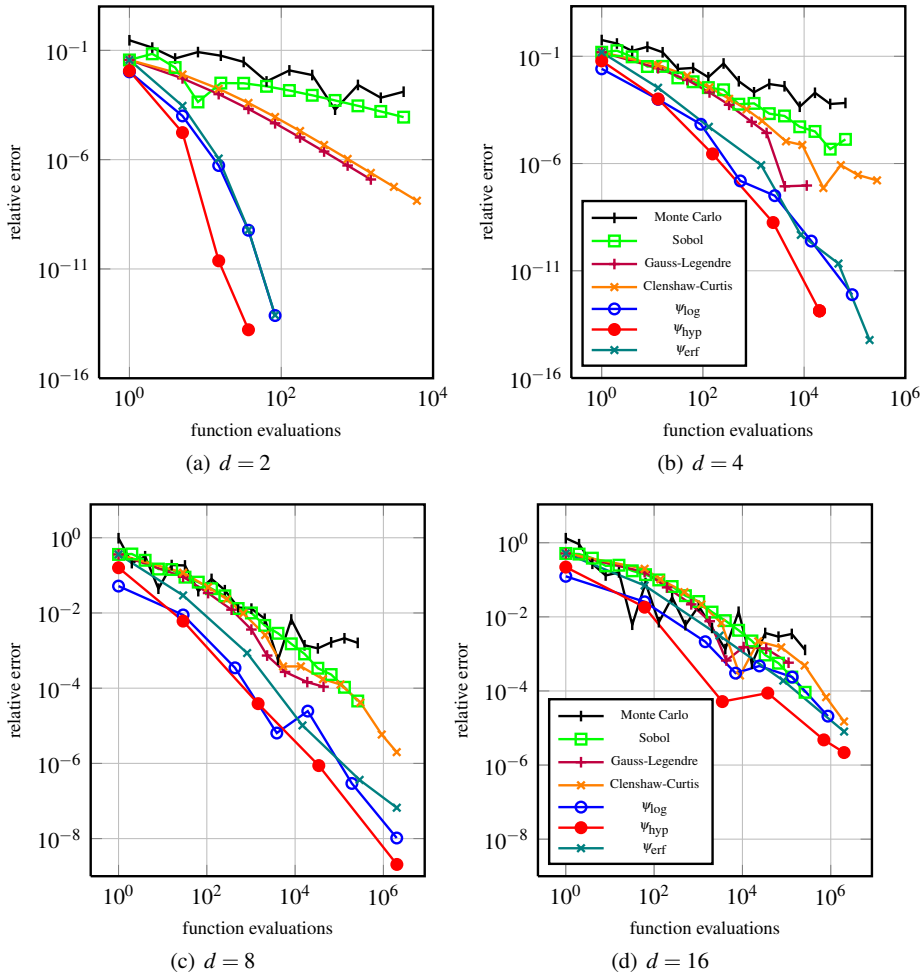
(c) $d = 8$ (d) $d = 16$

Fig. 6 Convergence behaviour for the Genz-integrand with $\rho_0 = 0.1$ and $b_i = \frac{1}{2}$.

fast growth of the integrand for arguments close to one. This is the reason why we now also included the ψ_{erf} method in our experiments for the Genz-integrand.

In our numerical experiments we will consider a special covariance structure for which the integral in (25) has a closed form solution [10, 12]. Namely we assume that the covariance matrix Σ has constant variance $\Sigma_{i,i} = 1$ and covariance $\Sigma_{i,j} = v_i \cdot v_j$ for $i \neq j$, where $v_i \in (-1,1), i = 1,\dots,d$. We remark that the normalization of the variance to one is not a restriction because it is always possible to shift the variance via a diagonal transformation to the boundaries of integration $b_1,\dots,b_d$.

In our first example we choose constant correlation $\Sigma_{i,j} = \rho_0 = 0.1$ and all $b_i = \frac{1}{2}$. In Figure 6 it can be observed that the dimension-adaptive sparse grid approach



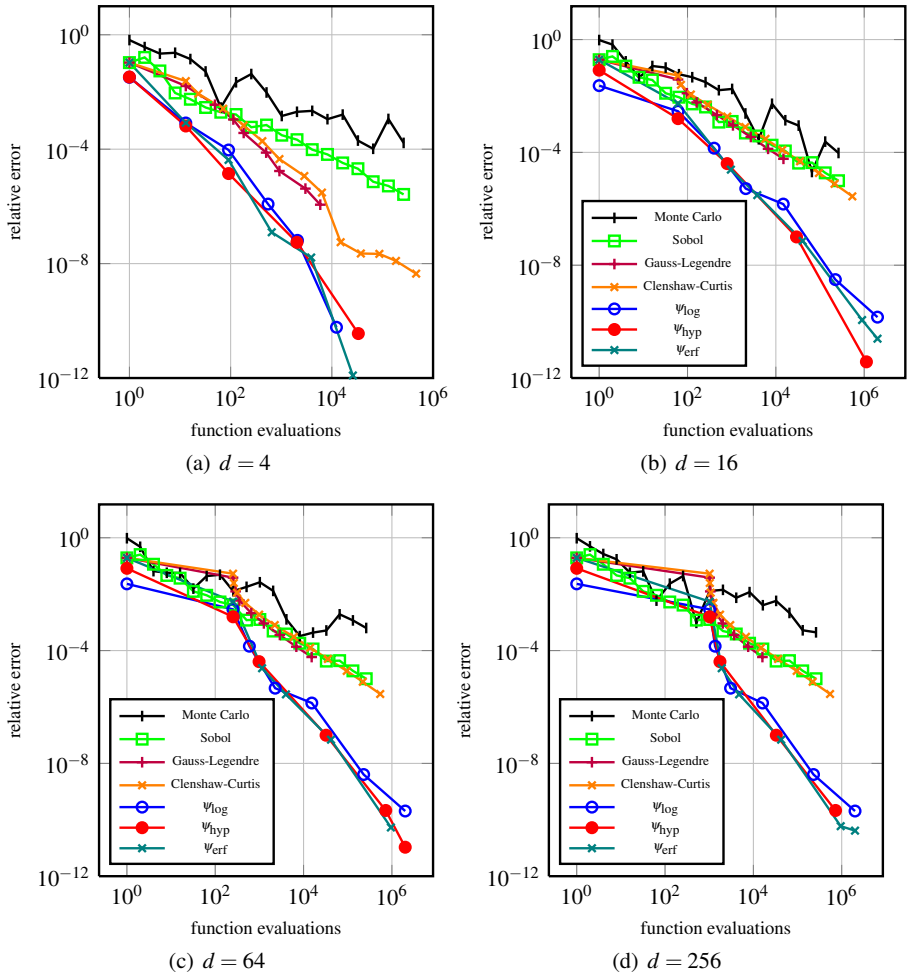
(a) $d = 2$ (b) $d = 4$

(c) $d = 8$ (d) $d = 16$

Fig. 7 Convergence behaviour of the Genz-integrand with $\rho_0 = 0.25$ and $b_i = -1 + \frac{i}{10}$.

is superior to (Q)MC for small values of d. Especially if it is based on the new generalized Gaussian formulae (11) with $\psi_{\log}$, ψ_{hyp} and ψ_{erf} , it performs very well and even achieves exponential convergence for small d. For higher dimensions the convergence should still be exponential, but it takes quite long until this asymptotic rate sets in. Thus exponential convergence is not visible in the case $d = 16$. This situation is common for sparse grids, as pointed out in [34, 35].

In our second example we use different values for the boundaries, namely $b_i = -1 + \frac{i}{10}$ and a bigger, but still constant correlation $\Sigma_{i,j} = \rho_0 = 0.25$. The convergence behaviour is similar to the first example, as can be seen from Figure 7.



(a) $d = 4$ (b) $d = 16$

(c) $d = 64$ (d) $d = 256$

Fig. 8 Convergence behaviour of the Genz-integrand with $\Sigma_{i,j} = 2^{-(i+j)}$ and $b_i = -1/2 + \frac{i}{10}$.

This demonstrates that our new approach indeed allows to deal with varying boundary values as it is needed in most pratical applications.

In the third example we look at a truely high-dimensional example with d up to 256. Here we set $\Sigma_{i,j} = 2^{-(i+j)}$ and $b_i = -1 + \frac{i}{10}$. This enforces an exponential decay of the correlation coefficients which weakens the interaction between the coordinates of the underlying integrand. Albeit involving a somewhat strong restriction, such situations appear often in practical problems with high nominal but low intrinsic dimensions, where measured data are involved, like for example in the panel probit models in econometrics [24], in density estimation in statistics [13] or in the pricing of various instruments in financial engineering [15].

It can be observed in Figure 8 that the dimension-adaptive sparse grid algorithm has now no trouble with nominally high dimensions and is able to correctly detect the important intrinsic coordinate directions. Moreover our new approach with $\psi_{\log}$, ψ_{hyp} and ψ_{erf} clearly outperforms all other methods in the case $d = 256$, even though exponential convergence, as in Figure 6, is not apparent yet. Still, to achieve a relative accuracy of 10^{-7} it needs less than 10^5 function evaluations whereas QMC and the Clenshaw-Curtis or Gauss-Legendre based sparse grids would need about 10^8 function evaluations and plain Monte Carlo would even involve about 10^{13} function values.

Acknowledgments This research was supported by the Collaborative Research Center 1060 of the DFG.

5 Conclusion

In the present paper we constructed a new univariate generalized Gaussian quadrature rule for bounded domains, which is related to classical Gaussian quadrature on unbounded domains with respect to a certain weight function. Special cases involve the Gauss-Laguerre and Gauss-Hermite rules and thus allow for an easy construction of the new generalized Gaussian quadrature by building onto existing implementations. Another example, which is related to the double exponential quadrature approach, was also presented.

Moreover, we used sparse tensor-products of this new univariate approach to cover multivariate problems. As application we considered a variant of the Genz-algorithm in which the multivariate integrals are evaluated by a dimension-adaptive sparse grid approach that was based on the new generalized Gaussian quadrature formulae. We demonstrated that our new method is able to significantly outperform dimension-adaptive sparse grid methods based on Gauss-Legendre or Clenshaw-Curtis quadrature as well as Monte Carlo and Quasi-Monte Carlo methods in moderate dimensions up to 16 and for special cases also in a truely high-dimensional setting with dimensions of $d = 256$.

References

1. I. Bertschek and M. Lechner. Convenient estimators for the panel probit model. *Journal of Econometrics*, 87(2):329–371, 1998.
2. T. Bonk. A new algorithm for multi-dimensional adaptive numerical quadrature. In W. Hackbusch and G. Wittum, editors, *Adaptive Methods – Algorithms, Theory, and Applications, NNFM 46*, pages 54–68. Vieweg, 1994.
3. A. Borsch-Supan and V. Hajivassiliou. Smooth unbiased multivariate probability simulators for maximum likelihood estimation of limited dependent variable models. *Journal of Econometrics*, 58(3):347–368, 1993.

4. H.-J. Bungartz and S. Dirnstorfer. Multivariate quadrature on adaptive sparse grids. *Computing*, 71(1):89–114, 2003.
5. H.-J. Bungartz and M. Griebel. A note on the complexity of solving Poisson's equation for spaces of bounded mixed derivatives. *Jornal of Complexity*, 15:167–199, 1999.
6. H.-J. Bungartz and M. Griebel. Sparse grids. *Acta Numerica*, 13:1–123, 2004.
7. P. Davis and P. Rabinowitz. *Methods of Numerical Integration*. Academic Press, 1984.
8. R. DeVore. Nonlinear approximation. *Acta Numerica*, 7:51–150, 1998.
9. R. DeVore and L. Scott. Error bounds for Gaussian quadrature and weighted L_1 polynomial approximation. *SIAM Journal on Numerical Analysis*, 21(2):400–412, 1984.
10. C. Dunnett and M. Sobel. Approximations to the probability integral and certain percentage points of a multivariate analogue of Student's t-distribution. *Biometrika*, 42:258–260, 1955.
11. W. Gautschi. Construction of Gauss-Christoffel quadrature formulas. *Mathematics of Computation*, 22(102):251–270, 1968.
12. A. Genz. Numerical computation of multivariate normal probabilities. *J. Comp. Graph Stat.*, 1:141–149, 1992.
13. A. Genz and F. Bretz. Comparison of methods for the computation of multivariate t probabilities. *Journal of Computational and Graphical Statistics*, 11(4):950–971, 2002.
14. A. Genz and F. Bretz. *Computation of Multivariate Normal and t Probabilities*. Springer, 2009.
15. T. Gerstner. Sparse grid quadrature methods for computational finance. Habilitation, Institut für Numerische Simulation, University of Bonn, 2007.
16. T. Gerstner and M. Griebel. Numerical integration using sparse grids. *Numerical Algorithms*, 18:209–232, 1998.
17. T. Gerstner and M. Griebel. Dimension–adaptive tensor–product quadrature. *Computing*, 71(1):65–87, 2003.
18. J. Geweke. Bayesian inference in econometric models using Monte Carlo integration. *Econometrica*, 57(6):1317–39, 1989.
19. A. Ghizzetti and A. Ossicini. Sull'esistenza e unicitat delle formule di quadrature Gaussiane. *Rend. Mat.*, 8:1–15, 1975.
20. M. Griebel. Sparse grids and related approximation schemes for higher dimensional problems. In *Proceedings of the Conference on Foundations of Computational Mathematics (FoCM05), Santander*, pages 106–161, 2005.
21. M. Griebel and M. Holtz. Dimension-wise integration of high-dimensional functions with applications to finance. *Journal of Complexity*, 26(5):455–489, 2010.
22. M. Griebel and S. Knapek. Optimized tensor-product approximation spaces. *Constructive Approximation*, 16(4):525–540, 2000.
23. M. Hegland. Adaptive sparse grids. In K. Burrage and R. B. Sidje, editors, *Proceedings of CTAC, Brisbane, 2001*, pages 335–353, 2001.
24. F. Heiss. The panel probit model: Adaptive integration on sparse grids. *Advances in Econometrics*, 26:41–64, 2010.
25. S. Karlin and W. J. Studden. *Tchebycheff systems: With applications in analysis and statistics (pure and applied mathematics)*. Interscience Publishers, 1966.
26. M. P. Keane. A computationally practical simulation estimator for panel data. *Econometrica*, 62(1):95–116, 1994.
27. F. Kuo, I. Sloan, and H. Woźniakowski. Periodization strategy may fail in high dimensions. *Numerical Algorithms*, 46(4):369–391, 2007.
28. D. Lubinsky and P. Rabinowitz. Rates of convergence of Gaussian quadrature. *Math. Comp.*, 43:219–242, 1984.
29. J. Ma, V. Rokhlin, and S. Wandzurra. Generalized Gaussian quadrature rules for systems of arbitrary functions. *SIAM Journal on Numerical Analysis*, 33(3):971–996, 1996.
30. G. Mastroianni and G. Monegato. Error estimates for Gauss-Laguerre and Gauss-Hermite quadrature formulas. In R. Zahar, editor, *Approximation and Computation: A Festschrift in Honor of Walter Gautschi*, volume 119 of *ISNM International Series of Numerical Mathematics*, pages 421–434. Birkhäuser Boston, 1994.
31. M. Masujima. *Path Integral Quantization and Stochastic Quantization*. Springer, 2008.

32. M. Mori. Discovery of the double exponential transformation and its developments. *Publications of the Research Institute for Mathematical Sciences*, 41(4):897–935, 2005.
33. T. Nahm. Error estimation and index refinement for dimension adaptive sparse grid quadrature with applications to the computation of path integrals. Master's thesis, Institut für Numerische Simulation, Universität Bonn, 2005.
34. F. Nobile, R. Tempone, and C. Webster. An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 46(5):2411–2442, 2008.
35. F. Nobile, R. Tempone, and C. Webster. A sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 46(5):2309–2345, 2008.
36. E. Novak and K. Ritter. High dimensional integration of smooth functions over cubes. *Numer. Math.*, 75:79–97, 1996.
37. D. Nuyens and R. Cools. Higher order quasi-Monte Carlo methods: A comparison. In T. E. Simos, G. Psihoyios, and C. Tsitouras, editors, *AIP Conference Proceedings*, pages 553–557. Springer New York LLC, 2010.
38. K. Petras. Smolyak cubature of given polynomial degree with few nodes for increasing dimension. *Numer. Math.*, 93(4):729–753, 2003.
39. A. Sidi. Extension of a class of periodizing variable transformations for numerical integration. *Mathematics of Computation*, 75(253):327–343, 2005.
40. I. Sloan and S. Joe. *Lattice Methods for Multiple Integration*. Oxford science publications. Clarendon Press, 1994.
41. I. Sloan, X. Wang, and H. Woźniakowski. Finite-order weights imply tractability of multivariate integration. *J. Complexity.*, 20(1):46–74, 2004.
42. I. Sloan and H. Wozniakowski. When are Quasi-Monte Carlo algorithms efficient for high-dimensional integrals? *J. Complexity*, 14:1–33, 1998.
43. S. Smolyak. Quadrature and interpolation formulas for tensor products of certain classes of functions. *Dokl. Akad. Nauk SSSR*, 4:240–243, 1963.
44. F. Stenger. Integration formulae based on the trapezoidal formula. *IMA Journal of Applied Mathematics*, 12(1):103–114, 1973.
45. H. Takahasi and M. Mori. Double exponential formulas for numerical integration. *Publications of the Research Institute for Mathematical Sciences*, 9(3):721–741, 1973.
46. G. Wasilkowski and H. Wozniakowski. Weighted tensor product algorithms for linear multivariate problems. *J. Complexity*, 15:402–447, 1999.
47. B. Zwicknagl and R. Schaback. Interpolation and approximation in Taylor spaces. *Journal of Approximation Theory*, 171:65–83, 2013.