Some Results on the Complexity of Numerical Integration

Erich Novak

We present some results on the complexity of numerical integration. We start with the seminal paper of Bakhvalov (1959) and end with new results on the curse of dimension and on the complexity of oscillatory integrals.

This survey paper consists of four parts:

1. Classical results till 1971
2. Randomized algorithms
3. Tensor product problems, tractability and weighted norms
4. Some recent results: $C^k$ functions and oscillatory integrals

1 Classical results till 1971

I start with a warning: We do not discuss the complexity of path integration and infinite-dimensional integration on $\mathbb{R}^N$ or other domains although there are exciting new results in that area, see [8] [14] [21] [22] [23] [40] [42] [43] [52] [67] [74] [87] [93] [117] [119]. For parametric integrals see [16] [17], for quantum computers, see [47] [48] [77] [112].

We mainly study the problem of numerical integration, i.e., of approximating the integral

$$S_d(f) = \int_{D_d} f(x) \, dx$$

over an open subset $D_d \subset \mathbb{R}^d$ of Lebesgue measure $\lambda^d(D_d) = 1$ for integrable functions $f: D_d \to \mathbb{R}$. The main interest is on the behavior of the minimal number of function values that are needed in the worst case setting to achieve an error at most $\varepsilon > 0$. Note that classical examples of domains $D_d$ are the unit cube $[0, 1]^d$ and the
normalized Euclidean ball (with volume 1), which are closed. However, we work with their interiors for definiteness of certain derivatives.

We state our problem. Let \( F_d \) be a class of integrable functions \( f : D_d \to \mathbb{R} \). For \( f \in F_d \), we approximate the integral \( S_d(f) \), see (1), by algorithms of the form

\[
A_n(f) = \phi_n(f(x_1), f(x_2), \ldots, f(x_n)),
\]

where \( x_j \in D_d \) can be chosen adaptively and \( \phi_n : \mathbb{R}^n \to \mathbb{R} \) is an arbitrary mapping. Adaption means that the selection of \( x_j \) may depend on the already computed values \( f(x_1), f(x_2), \ldots, f(x_{j-1}) \). We define \( N : F_d \to \mathbb{R}^n \) by \( N(f) = (f(x_1), \ldots, f(x_n)) \). The (worst case) error of the algorithm \( A_n \) is defined by

\[
e(A_n) = \sup_{f \in F_d} |S_d(f) - A_n(f)|,
\]

the optimal error bounds are given by

\[
e(n, F_d) = \inf_{A_n} e(A_n).
\]

The information complexity \( n(\varepsilon, F_d) \) is the minimal number of function values which is needed to guarantee that the error is at most \( \varepsilon \), i.e.,

\[
n(\varepsilon, F_d) = \min \{ n \mid \exists A_n \text{ such that } e(A_n) \leq \varepsilon \}.
\]

We minimize \( n \) over all choices of adaptive sample points \( x_j \) and mappings \( \phi_n \).

In this paper we give an overview on some of the basic results that are known about the numbers \( e(n, F_d) \) and \( n(\varepsilon, F_d) \). Hence we concentrate on complexity issues and leave aside other important questions such as implementation issues.

It was proved by Smolyak and Bakhvalov that as long as the class \( F_d \) is convex and symmetric we may restrict the minimization of \( e(A_n) \) by considering only non-adaptive choices of \( x_j \) and linear mappings \( \phi_n \), i.e., it is enough to consider \( A_n \) of the form

\[
A_n(f) = \sum_{i=1}^{n} a_i f(x_i).
\]

Theorem 0 (Bakhvalov [6]). Assume that the class \( F_d \) is convex and symmetric. Then linear and nonadaptive algorithms \( A_n \) of the form (2) are optimal and

\[
e(n, F_d) = \inf_{x_1, \ldots, x_n \in D_d, N(f) = 0} \sup_{f \in F_d} S_d(f).
\]

In this paper we only consider convex and symmetric \( F_d \) and then we can use the last formula for \( e(n, F_d) \).

Remark 0. a) For a proof of Theorem 0 see, for example, [86, Theorem 4.7]. This result is not really about complexity (hence it got its number), but it helps to prove complexity results.
b) A linear algorithm $A_n$ is called a quasi Monte Carlo (QMC) algorithm if $a_i = 1/n$ for all $i$ and is called a positive quadrature formula if $a_i > 0$ for all $i$. In general it may happen that optimal quadrature formulas have some negative weights and, in addition, we cannot say anything about the optimal points $x_i$.

c) More on the optimality of linear algorithms and on the power of adaption can be found in [13, 26, 86, 110, 111]. There are important classes of functions that are not symmetric and convex, and where Theorem 1 cannot be applied, see also [13, 92]. □

The optimal order of convergence plays an important role in numerical analysis. We start with a classical result of Bakhvalov (1959) for the class

$$ F^k_d = \{ f : [0,1]^d \rightarrow \mathbb{R} \mid \|D^\alpha f\|_\infty \leq 1, \ |\alpha| \leq k \}, $$

where $k \in \mathbb{N}$ and $|\alpha| = \sum_{i=1}^d \alpha_i$ for $\alpha \in \mathbb{N}_0^d$ and $D^\alpha f$ denotes the respective partial derivative. For two sequences $a_n$ and $b_n$ of positive numbers we write $a_n \asymp b_n$ if there are positive numbers $c$ and $C$ such that $c < a_n/b_n < C$ for all $n \in \mathbb{N}$.

**Theorem 1 (Bakhvalov [5]).**

$$ e(n, F^k_d) \asymp n^{-k/d}. \quad (4) $$

**Remark 1.** a) For such a complexity result one needs to prove an upper bound (for some algorithm) and a lower bound (for all algorithms). For the upper bound one can use tensor product methods based on a regular grid, i.e., one can use the $n = k^d$ points $x_i$ with coordinates from the set $\{1/(2k), 3/(2k), \ldots, (2k-1)/(2k)\}$.

The lower bound can be proved with the technique of “bump functions”: One can construct $2n$ functions $f_1, \ldots, f_{2n}$ with disjoint supports such that all $2^{2n}$ functions of the form $\sum_{i=1}^{2n} \delta_i f_i$ are contained in $F^k_d$, where $\delta_i = \pm 1$ and $S_d(f_i) \geq c_{d,k} n^{-k/d-1}$.

Since an algorithm $A_n$ can only compute $n$ function values, there are two functions $f^+ = \sum_{i=1}^{2n} f_i$ and $f^- = f^+ - 2 \sum_{i=1}^{2n} f_i$ such that $f^+, f^- \in F^k_d$ and $A_n(f^+) = A_n(f^-)$ but $|S_d(f^+) - S_d(f^-)| \geq 2 n c_{d,k} n^{-k/d-1}$. Hence the error of $A_n$ must be at least $c_{d,k} n^{-k/d}$. For the details see, for example, [75].

b) Observe that we can not conclude anything on $n(\varepsilon, F^k_d)$ if $\varepsilon$ is fixed and $d$ is large, since Theorem 1 contains hidden constants that depend on $k$ and $d$. Actually the lower bound is of the form

$$ e(n, F^k_d) \geq c_{d,k} n^{-k/d}, $$

where the $c_{d,k}$ are superexponentially small with $d \rightarrow \infty$.

c) The proof of the upper bound (using tensor product algorithms) is easy since we assumed that the domain is $D_d = (0,1)^d$. The optimal order of convergence is known for much more general spaces (such as Besov and Triebel-Lizorkin spaces) and arbitrary bounded Lipschitz domains, see [83, 113, 115]. Then the proof of the upper bounds is more difficult, however.

d) Integration on fractals was recently studied by Dereich and Müller-Gronbach [13]. These authors also obtain an optimal order of convergence $n^{-k/\alpha}$. The
definition of $S_d$ must be modified and $\alpha$ coincides, under suitable conditions, with the Hausdorff dimension of the fractal. □

By the curse of dimensionality we mean that $n(\varepsilon, F_d)$ is exponentially large in $d$. That is, there are positive numbers $c, \varepsilon_0$ and $\gamma$ such that

$$n(\varepsilon, F_d) \geq c(1 + \gamma)^d \quad \text{for all } \varepsilon \leq \varepsilon_0 \quad \text{and infinitely many } d \in \mathbb{N}. \quad (5)$$

If, on the other hand, $n(\varepsilon, F_d)$ is bounded by a polynomial in $d$ and $\varepsilon^{-1}$ then we say that the problem is polynomially tractable. If $n(\varepsilon, F_d)$ is bounded by a polynomial in $\varepsilon^{-1}$ alone, i.e., $n(\varepsilon, F_d) \leq C\varepsilon^{-\alpha}$, then we say that the problem is strongly polynomially tractable.

From the proof of Theorem 1 we cannot conclude whether the curse of dimensionality holds for the classes $F_k^d$ or not; see Theorem 11. Possibly Maung Zho Newn and Sharygin [72] were the first who published (in 1971) a complexity result for arbitrary $d$ with explicit constants and so proved the curse of dimensionality for Lipschitz functions.

**Theorem 2 (Maung Zho Newn and Sharygin [72]).** Consider the class

$$F_d = \{ f : [0, 1]^d \to \mathbb{R} \mid |f(x) - f(y)| \leq \max_i |x_i - y_i| \}.$$ 

Then

$$e(n, F_d) = \frac{d}{2d + 2} n^{-1/d}$$

for $n = m^d$ with $m \in \mathbb{N}$.

**Remark 2.** One can show that for $n = k^d$ the regular grid (points $x_i$ with coordinates from the set $\{1/(2k), 3/(2k), \ldots, (2k - 1)/(2k)\}$) and the midpoint rule $A_\alpha(f) = n^{-1} \sum_{i=1}^n f(x_i)$ are optimal. See also [3, 4, 12, 105] for this result and for generalizations to similar function spaces. □

## 2 Randomized algorithms

The integration problem is difficult for all deterministic algorithms if the classes $F_d$ of inputs are too large, see Theorem 2. One may hope that randomized algorithms make this problem much easier.

Randomized algorithms can be formalized in slightly different ways. We do not explain the technical details of these models and only give a reason why it makes sense to study different models for upper and lower bounds, respectively; see [86] for more details.

- Assume that we want to construct and to analyze concrete algorithms that yield upper bounds for the (total) complexity of given problems including the arithmetic cost and the cost of the random number generator. Then it is reasonable to
consider a more restrictive model of computation where, for example, only the standard arithmetic operations are allowed. One may also restrict the use of random numbers and study so-called restricted Monte Carlo methods, where only random bits are allowed; see [50].

- For the proof of lower bounds we take the opposite view and allow general randomized mappings and a very general kind of randomness. This makes the lower bounds stronger.

It turns out that the results are often very robust with respect to changes of the computational model. For the purpose of this paper, it might be enough that a randomized algorithm \( A \) is a random variable \( (A^\omega)_{\omega \in \Omega} \) with a random element \( \omega \) where, for each fixed \( \omega \), the algorithm \( A^\omega \) is a (deterministic) algorithm as before. We denote by \( \mu \) the distribution of the \( \omega \). In addition one needs rather weak measurability assumptions, see also the textbook [71]. Let \( n(f, \omega) \) be the number of function values used for fixed \( \omega \) and \( f \).

The number

\[
n(A) = \sup_{f \in F} \int_{\Omega} n(f, \omega) \, d\mu(\omega)
\]

is called the **cardinality** of the randomized algorithm \( A \) and

\[
e^{\text{ran}}(A) = \sup_{f \in F} \left( \int_{\Omega} \|S(f) - \phi_\omega(N_\omega(f))\|^2 \, d\mu(\omega) \right)^{1/2}
\]

is the error of \( A \). By \( f^* \) we denote the upper integral. For \( n \in \mathbb{N} \), define

\[
e^{\text{ran}}(n, F_d) = \inf \{ e^{\text{ran}}(A) : n(A) \leq n \}.
\]

If \( A : F \to G \) is a (measurable) deterministic algorithm then \( A \) can also be treated as a randomized algorithm with respect to a Dirac (atomic) measure \( \mu \). In this sense we can say that deterministic algorithms are special randomized algorithms. Hence the inequality

\[
e^{\text{ran}}(n, F_d) \leq e(n, F_d)
\]

is trivial.

The number \( e^{\text{ran}}(0, F_d) \) is called the **initial error in the randomized setting**. For \( n = 0 \), we do not sample \( f \), and \( A^\omega(f) \) is independent of \( f \), but may depend on \( \omega \). It is easy to check that for a linear \( S \) and a balanced and convex set \( F \), the best we can do is to take \( A^\omega = 0 \) and then

\[
e^{\text{ran}}(0, F_d) = e(0, F_d).
\]

This means that for linear problems the initial errors are the same in the worst case and randomized settings.

The main advantage of randomized algorithms is that the curse of dimensionality is not present even for certain large classes of functions. With the standard Monte Carlo method we obtain
\[ e_{\text{ran}}(n, F_d) \leq \frac{1}{\sqrt{n}} \]

when \( F_d \) is the unit ball of \( L_p([0,1]^d) \) and \( 2 \leq p \leq \infty \). Mathé \[70\] proved that this is almost optimal and the optimal algorithm is

\[ A_n^{\omega}(f) = \frac{1}{n + \sqrt{n}} \sum_{i=1}^{n} f(X_i) \]

with i.i.d. random variables \( X_i \) that are uniformly distributed on \([0,1]^d\). It also follows that

\[ e_{\text{ran}}(n, F_d) = \frac{1}{1 + \sqrt{n}} \]

when \( F_d \) is the unit ball of \( L_p([0,1]^d) \) and \( 2 \leq p \leq \infty \). In the case \( 1 \leq p < 2 \) one can only achieve the rate \( n^{-1+1/p} \), for a discussion see \[49\].

Bakhvalov \[5\] found the optimal order of convergence already in 1959 for the class \( F_k^d = \{ f : [0,1]^d \to \mathbb{R} \mid \|D^\alpha f\|_{\infty} \leq 1, |\alpha| \leq k \} \),

where \( k \in \mathbb{N} \) and \( |\alpha| = \sum_{i=1}^{d} \alpha_i \) for \( \alpha \in \mathbb{N}_0^d \).

**Theorem 3 (Bakhvalov \[5\]).**

\[ e_{\text{ran}}(n, F_k^d) \asymp n^{-k/d - 1/2}. \] (7)

**Remark 3.** The proof of the upper bound can be given with a technique that is often called *separation of the main part* or also *control variates*. For \( n = 2m \) use \( m \) function values to construct a good \( L_2 \) approximation \( f_m \) of \( f \in F_k^d \) by a deterministic algorithm. The optimal order of convergence is

\[ \|f - f_m\|_2 \asymp m^{-k/d} \]

Then use the unbiased estimator

\[ A_n^{\omega}(f) = S_d(f_m) + \frac{1}{m} \sum_{i=1}^{m} (f - f_m)(X_i) \]

with i.i.d. random variables \( X_i \) that are uniformly distributed on \([0,1]^d\). See, for example, \[71, 75\] for more details. We add in passing that the optimal order of convergence can be obtained for many function spaces (Besov spaces, Triebel-Lizorkin spaces) and for arbitrary bounded Lipschitz domains \( \mathbb{D}_d \subset \mathbb{R}^d \); see \[83\], where the approximation problem is studied. To obtain an explicit randomized algorithm with the optimal rate of convergence one needs a random number generator for the set \( \mathbb{D}_d \). If it is not possible to obtain efficiently random samples from the uniform distribution on \( \mathbb{D}_d \) one can work with Markov chain Monte Carlo (MCMC) methods, see Theorem 5.
Some Results on the Complexity of Numerical Integration

All known proofs of lower bounds use the idea of Bakhvalov (also called Yao’s Minimax Principle): study the average case setting with respect to a probability measure on \( F \) and use the theorem of Fubini. For details see [44, 45, 46, 71, 75, 87].

We describe a problem that was studied by several colleagues and solved by Hinrichs [57] using deep results from functional analysis. Let \( H(K_d) \) be a reproducing kernel Hilbert space of real functions defined on a Borel measurable set \( D_d \subseteq \mathbb{R}^d \). Its reproducing kernel \( K_d : D_d \times D_d \to \mathbb{R} \) is assumed to be integrable,

\[
e_{\text{init}}^d := \left( \int_{D_d} \int_{D_d} K_d(x, y) \rho_d(x) \rho_d(y) \, dx \, dy \right)^{1/2} < \infty.
\]

Here, \( \rho_d \) is a probability density function on \( D_d \). Without loss of generality we assume that \( D_d \) and \( \rho_d \) are chosen such that there is no subset of \( D_d \) with positive measure such that all functions from \( H(K_d) \) vanish on it.

The inner product and the norm of \( H(K_d) \) are denoted by \( \langle \cdot, \cdot \rangle_{H(K_d)} \) and \( \| \cdot \|_{H(K_d)} \). Consider multivariate integration

\[
S_d(f) = \int_{D_d} f(x) \rho_d(x) \, dx \quad \text{for all } f \in H(K_d),
\]

it is assumed that \( S_d : H(K_d) \to \mathbb{R} \) is continuous.

We approximate \( S_d(f) \) in the randomized setting using importance sampling. That is, for a positive probability density function \( \tau_d \) on \( D_d \) we choose \( n \) random sample points \( x_1, x_2, \ldots, x_n \) which are independent and distributed according to \( \tau_d \) and take the algorithm

\[
A_{n,d,\tau_d}(f) = \frac{1}{n} \sum_{j=1}^{n} \frac{f(x_j) \rho_d(x_j)}{\tau_d(x_j)}.
\]

The error of \( A_{n,d,\tau_d} \) is defined as

\[
e_{\text{ran}}(A_{n,d,\tau_d}) = \sup_{\|f\|_{H(K_d)} \leq 1} \left( \mathbb{E}_{\tau_d} \left( S_d(f) - A_{n,d,\tau_d}(f) \right)^2 \right)^{1/2},
\]

where the expectation is with respect to the random choice of the sample points \( x_j \).

For \( n = 0 \) we formally take \( A_{0,d,\tau_d} = 0 \) and then

\[
e_{\text{ran}}(0, H(K_d)) = e_{\text{init}}^d.
\]

Hinrichs [57] proved the following theorem.

**Theorem 4 (Hinrichs [57]).** Assume additionally that \( K_d(x, y) \geq 0 \) for all \( x, y \in D_d \). Then there exists a positive density function \( \tau_d \) such that

\[
e_{\text{ran}}(A_{n,d,\tau_d}) \leq \left( \frac{\pi}{2} \right)^{1/2} \frac{1}{\sqrt{n}} e_{\text{ran}}(0, H(K_d)).
\]
Hence, if we want to achieve $\epsilon^\text{ran}(A_{n,d,z_0}) \leq \epsilon \epsilon^\text{ran}(0,H(K_d))$ then it is enough to take
\[ n = \left\lceil \pi \left( \frac{1}{\epsilon^2} \right) \right\rceil. \]

**Remark 4.** In particular, such problems are strongly polynomially tractable (for the normalized error) if the reproducing kernels are pointwise nonnegative and integrable. In [88] we prove that the exponent 2 of $\epsilon^{-1}$ is sharp for tensor product Hilbert spaces whose univariate reproducing kernel is decomposable and univariate integration is not trivial for the two parts of the decomposition. More specifically we have
\[ n^\text{ran}(\epsilon,H(K_d)) \geq \left\lceil \frac{1}{8} \left( \frac{1}{\epsilon^2} \right) \right\rceil \quad \text{for all } \epsilon \in (0,1) \text{ and } d \geq \frac{2 \ln \epsilon^{-1} - \ln 2}{\ln \alpha^{-1}}, \]
where $\alpha \in [1/2,1)$ depends on the particular space.

We stress that these estimates hold independently of the smoothness of functions in a Hilbert space. Hence, even for spaces of very smooth functions the exponent of strong polynomial tractability must be 2. \(\square\)

Sometimes one cannot sample easily from the “target distribution” $\pi$ if one wants to compute an integral
\[ S(f) = \int_D f(x) \pi(dx). \]

Markov chain Monte Carlo (MCMC) methods are a very versatile and widely used tool to compute integrals and expectations.

We use an average of a finite Markov chain sample as approximation of the mean, i.e., we approximate $S(f)$ by
\[ S_{n,n_0}(f) = \frac{1}{n} \sum_{j=1}^{n} f(X_{j+n_0}), \]
where $(X_i)_{i \in \mathbb{N}_0}$ is a Markov chain with stationary distribution $\pi$. The number $n$ determines the number of function evaluations of $f$. The number $n_0$ is the burn-in or warm up time. Intuitively, it is the number of steps of the Markov chain to get close to the stationary distribution $\pi$.

We study the mean square error of $S_{n,n_0}$, given by
\[ e_{\nu}(S_{n,n_0},f) = \left( E_{\nu,K}|S_{n,n_0}(f) - S(f)\right)^{1/2}, \]
where $\nu$ and $K$ indicate the initial distribution and transition kernel of the chain; we work with the spaces $L_p = L_p(\pi)$. For the proof of the following error bound we refer to [96, Theorem 3.34 and Theorem 3.41].

**Theorem 5 (Rudolf [96]).** Let $(X_i)_{i \in \mathbb{N}}$ be a Markov chain with reversible transition kernel $K$, initial distribution $\nu$, and transition operator $P$. Further, let
\[ \Lambda = \sup \{ \alpha : \alpha \in \text{spec}(P - S) \}, \]

where \( \text{spec}(P - S) \) denotes the spectrum of the operator \((P - S) : L_2 \to L_2\), and assume that \( \Lambda < 1 \). Then

\[
\sup_{\|f\|_p \leq 1} e_\nu(S_{n,n_0},f)^2 \leq \frac{2}{n(1 - \Lambda)} + \frac{2C_\nu \gamma^{n_0}}{n^2(1 - \gamma)^2}
\]

holds for \( p = 2 \) and for \( p = 4 \) under the following conditions:

- for \( p = 2 \), \( \frac{d\nu}{dx} \in L_\infty \) and a transition kernel \( K \) which is \( L_1 \)-exponentially convergent with \( (\gamma,M) \) where \( \gamma < 1 \), i.e.,

\[
\|P^n - S\|_{L_1 \to L_1} \leq M\gamma^n
\]

for all \( n \in \mathbb{N} \) and \( C_\nu = \|\frac{d\nu}{dx} - 1\|_\infty \);

- for \( p = 4 \), \( \frac{d\nu}{dx} \in L_2 \) and \( \gamma = \|P - S\|_{L_2 \to L_2} < 1 \) where \( C_\nu = 64\|\frac{d\nu}{dx} - 1\|_2 \).

**Remark 5.** Let us discuss the results. If the transition kernel is \( L_1 \)-exponentially convergent, then we have an explicit error bound for integrands \( f \in L_2 \) whenever the initial distribution has a density \( \frac{d\nu}{dx} \in L_\infty \). However, in general it is difficult to provide explicit values \( \gamma \) and \( M \) such that the transition kernel is \( L_1 \)-exponentially convergent with \( (\gamma,M) \). This motivates to consider transition kernels which satisfy a weaker convergence property, such as the existence of an \( L_2 \)-spectral gap, i.e., \( \|P - S\|_{L_2 \to L_2} < 1 \). In this case we have an explicit error bound for integrands \( f \in L_4 \) whenever the initial distribution has a density \( \frac{d\nu}{dx} \in L_2 \). Thus, by assuming a weaker convergence property of the transition kernel we obtain a weaker result in the sense that \( f \) must be in \( L_4 \) rather than \( L_2 \).

If we want to have an error of \( \varepsilon \in (0,1) \) it is still not clear how to choose \( n \) and \( n_0 \) to minimize the total amount of steps \( n + n_0 \). How should we choose the burn-in \( n_0 \)? One can prove in this setting, see [96], that the choice \( n_{opt} = \lceil \log \frac{C_\nu}{\gamma} \rceil \) is a reasonable and almost optimal choice for the burn-in.

More details can be found in [81]. For a full discussion with all the proofs see [96].

### 3 Tensor product problems and weights

We know from the work of Bakhvalov already done in 1959 that the optimal order of convergence is \( n^{-k/d} \) for functions from the class \( C^k([0,1]^d) \). To obtain an order of convergence of roughly \( n^{-k} \) for every dimension \( d \), one needs stronger smoothness conditions. This is a major reason for the study of functions with **bounded mixed derivatives**, or **dominating mixed smoothness**, such as the classes

\[
W_p^{k,\text{max}}([0,1]^d) = \{ f : [0,1]^d \to \mathbb{R} \mid \|D^\alpha f\|_p \leq 1 \text{ for } \|\alpha\|_\infty \leq k \}.
\]
Observe that functions from this class have, in particular, the high order derivative $D^{(k_1,\ldots,k_d)} f \in L^p$ and one may hope that the curse of dimensionality can be avoided or at least moderated by this assumption. For $k = 1$ these spaces are closely related to various notions of discrepancy, see, for example, [23, 69, 87, 109].

The optimal order of convergence is known for all $k \in \mathbb{N}$ and $1 < p < \infty$ due to the work of Roth [94, 95], Frolov [38, 39], Bykovskii [10], Temlyakov [107] and Skriganov [99], see the survey Temlyakov [109]. The cases $p \in \{1, \infty\}$ are still unsolved.

The case $p = 1$ is strongly related to the star discrepancy, see also Theorem 10.

**Theorem 6.** Assume that $k \in \mathbb{N}$ and $1 < p < \infty$. Then

$$e(n, W_p^{k,\text{mix}}([0,1]^d)) \approx n^{-k}(\log n)^{(d-1)/2}.$$  

**Remark 6.** The upper bound was proved by Frolov 1976 for $p = 2$ and by Skriganov 1994 for all $p > 1$. The lower bound was proved by Roth 1954 and Bykovskii 1985 for $p = 2$ and by Temlyakov 1990 for all $p < \infty$.

For functions in $W_p^{k,\text{mix}}([0,1]^d)$ with compact support in $(0,1)^d$ one can take algorithms of the form

$$A_n(f) = \frac{|\det A|}{a^d} \sum_{m \in \mathbb{Z}^d} f \left(\frac{Am}{a}\right),$$

where $A$ is a suitable matrix that does not depend on $k$ or $n$, and $a > 0$. Of course the sum is finite since we use only the points $\frac{Am}{a}$ inside $(0,1)^d$.

This algorithm is similar to a lattice rule but is not quite a lattice rule since the points do not build an integration lattice. The sum of the weights is roughly 1, but not quite. Therefore this algorithm is not really a quasi-Monte Carlo algorithm. The algorithm $A_n$ can be modified to obtain the optimal order of convergence for the whole space $W_p^{k,\text{mix}}([0,1]^d)$. The modified algorithm uses different points $x_i$ but still positive weights $a_i$. For a tutorial on this algorithm see [114].

For the Besov-Nikolskii classes $S_r^{p,q}B(T_d)$ with $1 \leq p, q \leq \infty$ and $1/p < r < 2$, the optimal rate is

$$n^{-r}(\log n)^{(d-1)(1-1/q)}$$

and can be obtained constructively with QMC algorithms, see [59]. The lower bound was proved by Triebel [113].

A famous algorithm for tensor product problems is the Smolyak algorithm, also called sparse grids algorithm. We can mention just a few papers and books that deal with this topic: The algorithm was invented by Smolyak [104] and, independently, by several other colleagues and research groups. Several error bounds were proved by Temlyakov [106, 108]; explicit error bounds (without unknown constants) were obtained by Wasilkowski and Woźniakowski [118, 120]. Novak and Ritter [78, 79, 80] studied the particular Clenshaw-Curtis Smolyak algorithm. A survey is Bürgisser and Griebel [9] and another one is [87, Chapter 15]. For recent results on the order of convergence see Sickel and T. Ullrich [97, 98] and Dinh and T. Ullrich [30]. The
recent paper contains a tractability result for the Smolyak algorithm applied to very smooth functions. We display only one recent result on the Smolyak algorithm.

**Theorem 7 (Sickel and T. Ullrich [98]).** For the classes $W_{2,\text{mix}}^{k,2}([0,1]^d)$ one can construct a Smolyak algorithm with the order of the error

$$n^{-k}(\log n)^{(d-1)(k+1/2)}.$$  \hfill (9)

**Remark 7.**

a) The bound (9) is valid even for $L_2$ approximation instead of integration, but it is not known whether this upper bound is optimal for the approximation problem. Using the technique of control variates one can obtain the order

$$n^{-k-1/2}(\log n)^{(d-1)(k+1/2)}$$

for the integration problem in the randomized setting.

b) It is shown in Dinh and T. Ullrich [30] that the order (9) cannot be improved with the Smolyak method.

c) We give a short description of the Clenshaw-Curtis Smolyak algorithm for the computation of integrals $\int_{[-1,1]^d} f(x) \, dx$ that often leads to “almost optimal” error bounds, see [79].

We assume that for $d = 1$ a sequence of formulas

$$U^i(f) = \sum_{j=1}^{m_i} a^j_i f(x^j_i)$$

is given. In the case of numerical integration the $a^j_i$ are just numbers. The method $U^i$ uses $m_i$ function values and we assume that $U^{i+1}$ has smaller error than $U^i$ and $m_{i+1} > m_i$. Define then, for $d > 1$, the tensor product formulas

$$(U^{i_1} \otimes \cdots \otimes U^{i_d})(f) = \sum_{j_1=1}^{m_{i_1}} \cdots \sum_{j_d=1}^{m_{i_d}} a^{i_1}_{j_1} \cdots a^{i_d}_{j_d} f(x^{i_1}_{j_1}, \ldots, x^{i_d}_{j_d}).$$

A tensor product formula clearly needs

$$m_{i_1} \cdot m_{i_2} \cdots m_{i_d}$$

function values, sampled on a regular grid. The Smolyak formulas $A(q,d)$ are clever linear combinations of tensor product formulas such that

- only tensor products with a relatively small number of knots are used;
- the linear combination is chosen in such a way that an interpolation property for $d = 1$ is preserved for $d > 1$.

The Smolyak formulas are defined by

$$A(q,d) = \sum_{q-d+1 \leq |i| \leq q} (-1)^{|i|} \cdot \binom{d-1}{q-|i|} \cdot (U^{i_1} \otimes \cdots \otimes U^{i_d}),$$
where \( q \geq d \). Specifically, we use, for \( d > 1 \), the Smolyak construction and start, for \( d = 1 \), with the classical Clenshaw-Curtis formula with 

\[
m_1 = 1 \quad \text{and} \quad m_i = 2^{i-1} + 1 \quad \text{for} \ i > 1.
\]

The Clenshaw-Curtis formulas

\[
U^i(f) = \sum_{j=1}^{m_i} a^i_j f(x^i_j)
\]

use the knots

\[
x^i_j = -\cos\left(\pi \frac{j - 1}{m_i - 1}\right),
\]

(and \( x^1_1 = 0 \)). Hence we use nonequidistant knots. The weights \( a^i_j \) are defined in such a way that \( U^i \) is exact for all (univariate) polynomials of degree at most \( m_i \).

It turns out that many tensor product problems are still intractable and suffer from the curse of dimension, for a rather exhaustive presentation see [86, 87, 89]. Sloan and Woźniakowski [103] describe a very interesting idea that was further developed in hundreds of papers, the paper [103] is most important and influential. We can describe here only the very beginnings of a long ongoing story; we present just one example instead of the whole theory.

The rough idea is that \( f : [0, 1]^d \to \mathbb{R} \) may depend on many variables, \( d \) is large, but some variables or groups of variables are more important than others. Consider, for \( d = 1 \), the inner product

\[
\langle f, g \rangle_{1, \gamma} = \int_0^1 f(x) g(x) \, dx + \frac{1}{\gamma} \int_0^1 f'(x) g'(x) \, dx,
\]

where \( \gamma > 0 \). If \( \gamma \) is small then \( f \) must be “almost constant” if it has small norm. A large \( \gamma \) means that \( f \) may have a large variation and still the norm is relatively small. Now we take tensor products of such spaces and weights \( \gamma_1 \geq \gamma_2 \geq \ldots \) and consider the complexity of the integration problem for the unit ball \( F_d \) with respect to this weighted norm. The kernel \( K \) of the tensor product space \( H(K) \) is of the form

\[
K(x, y) = \prod_{i=1}^d K_{\gamma_i}(x_i, y_i),
\]

where \( K_{\gamma} \) is the kernel of the respective space \( H_{\gamma} \) of univariate functions.

**Theorem 8 (Sloan and Woźniakowski [103]).** Assume that \( \sum_{i=1}^{\infty} \gamma_i < \infty \). Then the problem is strongly polynomially tractable.

**Remark 8.** The paper [103] contains also a lower bound which is valid for all quasi-Monte Carlo methods. The proof of the upper bound is very interesting and an excellent example for the probabilistic method. Compute the mean of the quadratic worst case error of QMC algorithms over all \((x_1, \ldots, x_n) \in [0, 1]^m \) and obtain
Some Results on the Complexity of Numerical Integration

\[ \frac{1}{n} \left( \int_{[0,1]^d} K(x,x) \, dx - \int_{[0,1]^2d} K(x,y) \, dx \, dy \right). \]

This expectation is of the form \( C_d n^{-1} \) and the sequence \( C_d \) is bounded if and only if \( \sum \gamma_i < \infty \). The lower bound in [103] is based on the fact that the kernel \( K \) is always non-negative; this leads to lower bounds for QMC algorithms or, more general, algorithms with positive weights. ∎

As already indicated, Sloan and Woźniakowski [103] was continued in many directions. Much more general weights and many different Hilbert spaces were studied. By the probabilistic method one only obtains the existence of a good QMC algorithms but, in the meanwhile, there exist many results about the construction of good algorithms. In this paper the focus is on the basic complexity results and therefore I simply list a few of the most relevant papers: [7], [11], [24], [28], [29], [53], [54], [55], [64], [65], [66], [68], [90], [91], [100], [101] and [102]. See also the books [25, 69, 73, 87] and the excellent survey paper [23]. ∎

In complexity theory we want to study optimal algorithms and it is not clear whether QMC algorithms or quadrature formulas with positive weights are optimal. Observe that the Smolyak algorithm also uses negative weights and it is known that in certain cases positive quadrature formulas are far from optimal; for examples see [82] or [87, Section 10.6 and 11.3]. Therefore it is not clear whether the conditions on the weights can be relaxed if we allow arbitrary algorithms. The next result shows that this is not the case.

**Theorem 9** ([85]). The integration problem from Theorem 8 is strongly polynomially tractable if and only if \( \sum_{i=1}^{\infty} \gamma_i < \infty \).

**Remark 9.** Due to the known upper bound of Theorem 8 it is enough to prove a lower bound for arbitrary algorithms. This is done via the technique of decomposable kernels that was developed in [85]; see also [87, Chapter 11].

We do not describe this technique here and only remark that we need for this technique many non-zero functions \( f_i \) in the Hilbert space \( F_d \) with disjoint supports. Therefore this technique usually works for functions with finite smoothness, but not for analytic functions. ∎

Tractability of integration can be proved for many weighted spaces and one may ask whether there are also cases where tractability holds for certain unweighted spaces. A famous example for this are integration problems that are related to the star discrepancy.

For \( x_1, \ldots, x_n \in [0,1]^d \) define the star discrepancy by

\[ D^\infty(x_1, \ldots, x_n) = \sup_{t \in [0,1]^d} \left| t_1 \cdots t_d - \frac{1}{n} \sum_{i=1}^{n} \chi_{[0,t]}(x_i) \right|, \]

the respective QMC quadrature formula is \( Q_n(f) = \frac{1}{n} \sum_{i=1}^{n} f(x_i) \).

Consider the Sobolev space
\[ F_d = \{ f \in W_1^{\text{mix}} \mid \| f \| \leq 1, f(x) = 0, \text{if there exists an } i \text{ with } x_i = 1 \} \]

with the norm
\[ \| f \| := \left\| \frac{\partial^d f}{\partial x_1 \partial x_2 \ldots \partial x_d} \right\|_1. \]

Then the Hlawka-Zaremba-equality yields
\[ D^*_\infty(x_1, \ldots, x_n) = \sup_{f \in F_d} |S_d(f) - Q_n(f)|, \]

hence the star discrepancy is a worst case error bound for integration. As always we define
\[ n(\varepsilon, F_d) = \min\{n \mid \exists x_1, \ldots, x_n \text{ with } D^*_\infty(x_1, \ldots, x_n) \leq \varepsilon \}. \]

The following result shows that this integration problem is polynomially tractable and the complexity is linear in the dimension.

**Theorem 10** ([51]).
\[ n(\varepsilon, F_d) \leq C d \varepsilon^{-2} \quad (10) \]

and
\[ n(1/64, F_d) \geq 0.18d. \]

**Remark 10.** This result was modified and improved in various ways and we mention some important results. Hinrichs [56] proved the lower bound
\[ n(\varepsilon, F_d) \geq c d \varepsilon^{-1} \quad \text{for } \varepsilon \leq \varepsilon_0. \]

Aistleitner [11] proved that the constant \( C \) in (10) can be taken as 100. Aistleitner and Hofer [2] proved more on upper bounds. Already the proof in [51] showed that an upper bound \( D^*_\infty(x_1, \ldots, x_n) \leq C \sqrt{d/n} \) holds with high probability if the points \( x_1, \ldots, x_n \) are taken independently and uniformly distributed. Doerr [31] proved the respective lower bound, hence
\[ \mathbb{E}(D^*_\infty(x_1, \ldots, x_n)) \geq \sqrt{\frac{d}{n}} \quad \text{for } n \geq d. \]

Since the upper bounds are proved with the probabilistic method, we only know the *existence* of points with small star discrepancy. The existence results can be transformed into (more or less explicit) constructions and the problem is, of course, to minimize the computing time as well as the discrepancy. One of the obstacles is that already the computation of the star discrepancy of given points \( x_1, x_2, \ldots, x_n \) is very difficult. We refer the reader to [19, 26, 27, 32, 33, 34, 35, 41, 58]. □

Recently Dick [20] proved a tractability result for another unweighted space that is defined via an \( L_1 \)-norm and consists in periodic functions; we denote Fourier coefficients by \( \tilde{f}(k) \), where \( k \in \mathbb{Z}^d \). Let \( 0 < \alpha \leq 1 \) and \( 1 \leq p \leq \infty \) and
Some Results on the Complexity of Numerical Integration

Dick proved the upper bound

\[ e(n, F_{\alpha,p,d}) \leq \max \left( \frac{d - 1}{\sqrt{n}}, \frac{d^{\alpha/p}}{n^{\alpha}} \right) \]

for any prime number \( n \). Hence the complexity is at most quadratic in \( d \).

The proof is constructive, a suitable algorithm is the following. Use points \( x_k = \left\{ \frac{k}{n} \right\}, \left\{ \frac{k^2}{n} \right\}, \ldots, \left\{ \frac{k^d}{n} \right\} \), where \( k = 0, 1, \ldots, n - 1 \), and take the respective QMC algorithm. \( \square \)

### 4 Some recent results

We end this survey with two results that were still unpublished at the time of the conference, April 2014. First we return to the classes \( C_k([0,1]^d) \), see Theorem 1.

We want to be a little more general and consider the computation of

\[ S_d(f) = \int_{D_d} f(x) \, dx \]

up to some error \( \varepsilon > 0 \), where \( D_d \subset \mathbb{R}^d \) has Lebesgue measure 1. The results hold for arbitrary sets \( D_d \), the standard example of course is \( D_d = [0,1]^d \). We consider functions \( f: \mathbb{R}^d \to \mathbb{R} \) since our emphasis is on lower bounds.

It has not been known if the curse of dimensionality is present for probably the most natural class which is the unit ball of \( r \) times continuously differentiable functions,

\[ F^k_{d} = \{ f \in C^k(\mathbb{R}^d) \mid \|D^\alpha f\|_\infty \leq 1 \quad \text{for all} \quad |\alpha| \leq k \}, \]

where \( k \in \mathbb{N} \).

**Theorem 11** ([60]). The curse of dimensionality holds for the classes \( F^k_{d} \) with the super-exponential lower bound

\[ n(\varepsilon, F^k_{d}) \geq c_k (1 - \varepsilon) d^{d/(2k+3)} \quad \text{for all} \quad d \in \mathbb{N} \quad \text{and} \quad \varepsilon \in (0,1), \]

where \( c_k > 0 \) depends only on \( k \).

**Remark 11.** In [60, 61] we also prove that the curse of dimensionality holds for even smaller classes of functions \( F_d \) for which the norms of arbitrary directional derivatives are bounded proportionally to \( 1/\sqrt{d} \).

We start with the fooling function

\[ F_{\alpha,p,d} = \left\{ f : [0,1]^d \to \mathbb{R} \mid \sum_{k \in \mathbb{Z}^d} |\tilde{f}(k)| + \sup_{\|h\|_p^\alpha} \frac{|f(x+h) - f(x)|}{\|h\|_p^\alpha} \leq 1 \right\}. \]
\[ f_0(x) = \min \left\{ \frac{1}{\delta \sqrt{d}} \text{dist}(x, \mathcal{P}_\delta) \right\} \quad \text{for all} \quad x \in \mathbb{R}^d, \]

where

\[ \mathcal{P}_\delta = \bigcup_{i=1}^n B^d(x_i) \]

and \( B^d(x_i) \) is the ball with center \( x_i \) and radius \( \delta \sqrt{d} \). The function \( f_0 \) is Lipschitz.

By a suitable smoothing via convolution we construct a smooth fooling function \( f_k \in F_d \) with \( f_k \mid_{\mathcal{P}_0} = 0 \).

Important elements of the proof are volume estimates (in the spirit of Elekes [37] and Dyer, Füredi and McDiarmid [36]), since we need that the volume of a neighborhood of the convex hull of \( n \) arbitrary points is exponentially small in \( d \). \( \Box \)

Also classes of \( C^\infty \)-functions were studied recently. We still do not know whether the integration problem suffers from the curse of dimensionality for the classes

\[ F_d = \{ f : [0, 1]^d \to \mathbb{R} \mid \| D^\alpha f \|_\infty \leq 1 \text{ for all } \alpha \in \mathbb{N}_0^d \}, \]

this is Open Problem 2 from [86]. We know from Vybíral [116] and [61] that the curse is present for somewhat larger spaces and that a weak tractability holds for smaller classes; this can be proved with the Smolyak algorithm, see [62]. \( \Box \)

We now consider univariate oscillatory integrals for the standard Sobolev spaces \( H^s \) of periodic and non-periodic functions with an arbitrary integer \( s \geq 1 \). We study the approximate computation of Fourier coefficients

\[ I_k(f) = \int_0^1 f(x) e^{-2\pi ikx} \, dx, \quad i = \sqrt{-1}, \]

where \( k \in \mathbb{Z} \) and \( f \in H^s \).

There are several recent papers about the approximate computation of highly oscillatory univariate integrals with the weight \( \exp(2\pi ikx) \), where \( x \in [0, 1] \) and \( k \) is an integer (or \( k \in \mathbb{R} \)) which is assumed to be large in the absolute sense, see Huybrechs and Olver [63] for a survey.

We study the Sobolev space \( H^s \) for a finite \( s \in \mathbb{N} \), i.e.,

\[ H^s = \{ f : [0, 1] \to \mathbb{C} \mid f^{(s-1)} \text{ is abs. cont., } f^{(s)} \in L_2 \} \]

with the inner product

\[ \langle f, g \rangle_s = \sum_{l=0}^{s-1} \int_0^1 f^{(l)}(x) \, dx \int_0^1 \overline{g^{(l)}(x)} \, dx + \int_0^1 \overline{f^{(s)}(x) \, g^{(s)}(x)} \, dx \]

\[ = \sum_{l=0}^{s-1} \langle f^{(l)}, 1 \rangle_0 \langle g^{(l)}, 1 \rangle_0 + \langle f^{(s)}, g^{(s)} \rangle_0, \]

(13)

where \( \langle f, g \rangle_0 = \int_0^1 f(x) \overline{g(x)} \, dx \), and norm \( \| f \|_{H^s} = \langle f, f \rangle_s^{1/2} \).
For the periodic case, an algorithm that uses \( n \) function values at equally spaced points is nearly optimal, and its worst case error is bounded by \( C_s (n + |k|)^{-s} \) with an exponentially small \( C_s \) in \( s \). For the non-periodic case, we first compute successive derivatives up to order \( s - 1 \) at the end-points \( x = 0 \) and \( x = 1 \). These derivatives values are used to periodize the function and this allows us to obtain similar error bounds like for the periodic case. Asymptotically in \( n \), the worst case error of the algorithm is of order \( n^{-s} \) independently of \( k \) for both periodic and non-periodic cases.

**Theorem 12 ([84]).** Consider the integration problem \( I_k \) defined over the space \( H^s \) of non-periodic functions with \( s \in \mathbb{N} \). Then

\[
\frac{c_s}{(n + |k|)^s} \leq e(n,k,H^s) \leq \left( \frac{3}{2\pi} \right)^s \frac{2}{(n + |k| - 2s - 1)^s},
\]

for all \( k \in \mathbb{Z} \) and \( n \geq 2s \).

\[ \square \]

**Remark 12.** The minimal errors \( e(n,k,H^s) \) for the non-periodic case have a peculiar property for \( s \geq 2 \) and large \( k \). Namely, for \( n = 0 \) we obtain the initial error which is of order \( |k|^{-1} \), whereas for \( n \geq 2s \) it becomes of order \( |k|^{-s} \). Hence, the dependence on \( |k|^{-1} \) is short-lived and disappears quite quickly. For instance, take \( s = 2 \). Then \( e(n,k,s) \) is of order \( |k|^{-1} \) only for \( n = 0 \) and maybe for \( n = 1, 2, 3 \), and then becomes of order \( |k|^{-2} \).

\[ \square \]

**Acknowledgement.** I thank the following colleagues and friends for valuable remarks: Michael Gnewuch, Aicke Hinrichs, Robert Kunsch, Thomas Müller-Gronbach, Daniel Rudolf and Henryk Woźniakowski.

### References

1. Ch. Aistleitner, Covering numbers, dyadic chaining and discrepancy, *J. Complexity* 27, 531–540, 2011.
2. Ch. Aistleitner and M. Hofer, Probabilistic discrepancy bounds for Monte Carlo point sets, *Math. Comp.* 83, 1373–1381, 2014.
3. V. F. Babenko, Asymptotically sharp bounds for the remainder for the best quadrature formulas for certain classes of functions, *Mat. Zametki* 19, no. 3, 313–322, 1976. English translation: *Math. Notes* 19, 187–193, 1976.
4. V. F. Babenko Exact asymptotics of the error of weighted cubature formulas optimal for certain classes of functions, *Mat. Zametki* 20, 589–595, 1976. English translation: *Math. Notes* 20.
5. N. S. Bakhvalov, On approximate computation of integrals, *Vestnik MGU, Ser. Math. Mech. Astron. Phys. Chem.* 4, 3–18, 1959, in Russian. A translation will soon appear in the J. Complexity.
6. N. S. Bakhvalov, On the optimality of linear methods for operator approximation in convex classes of functions, *USSR Comput. Maths. Math. Phys.* 11, 244–249, 1971.
7. J. Baldeaux, J. Dick, G. Leobacher, D. Nuyens and F. Pillichshammer, Efficient calculation of the worst-case error and (fast) component-by-component construction of higher order polynomial lattice rules, *Numer. Algorithms* 59, 403–431, 2012.
8. J. Baldeaux and M. Gnewuch, Optimal randomized multilevel algorithms for infinite-dimensional integration on function spaces with ANOVA-type decomposition, *SIAM J. Num. Anal.* 52, 1128–1155, 2014.
9. H.-J. Bungartz and M. Griebel, Sparse grids, Acta Numerica 13, 147–269, 2004.
10. V. A. Bykovskii, On the correct order of the error of optimal cubature formulas in spaces with dominant derivative, and on quadratic deviations of grids, Computing Center Far-Eastern Scientific Center, Akad. Sci. USSR, Vladivostok, 1985, preprint.
11. W. W. L. Chen and M. M. Skriganov, Explicit constructions in the classical mean squares problem in irregularities of point distribution, J. Reine Angew. Math. (Crelle) 545, 67–95, 2002.
12. E. V. Chernaya, Asymptotically exact estimation of the error of weighted cubature formulas optimal in some classes of continuous functions, Ukrainian Math. J. 47, 1606–1618.
13. N. Clancy, Y. Ding, C. Hamilton, F. J. Hickernell and Y. Zhang, The cost of deterministic, adaptive, automatic algorithms: Cones, not balls, J. Complexity 30, 21–45, 2014.
14. J. Creutzig, S. Dereich, Th. Müller-Gronbach and K. Ritter, Infinite-dimensional quadrature and approximation of distributions, Found. Comput. Math., 9, 391–429, 2009.
15. J. Creutzig and P. Wojtaszczyk, Linear vs. nonlinear algorithms for linear problems, J. Complexity 20, 807–820, 2004.
16. T. Daun and S. Heinrich, Complexity of Banach space valued and parametric integration, in: Monte Carlo and Quasi-Monte Carlo Methods 2012, J. Dick, F. Y. Kuo, G. W. Peters, I. H. Sloan (eds.), Springer Verlag, pp. 297–316.
17. T. Daun and S. Heinrich, Complexity of parametric integration in various smoothness classes, preprint.
18. S. Dereich and Th. Müller-Gronbach, Quadrature for self-affine distributions on $\mathbb{R}^d$, SPP-Preprint 160.
19. J. Dick, A note on the existence of sequences with small star discrepancy, J. Complexity 23, 649–652, 2007.
20. J. Dick, Numerical integration of Hölder continuous, absolutely convergent Fourier-, Fourier cosine-, and Walsh series, preprint.
21. J. Dick and M. Gnewuch, Optimal randomized changing dimension algorithms for infinite-dimensional integration on function spaces with ANOVA-type decomposition, J. Approx. Th. 184, 111–145, 2014.
22. J. Dick and M. Gnewuch, Infinite-dimensional integration in weighted Hilbert spaces: anchored decompositions, optimal deterministic algorithms, and higher order convergence, Found. Comput. Math., to appear.
23. J. Dick, F. Y. Kuo and I. H. Sloan, High-dimensional integration: The quasi-Monte Carlo way, Acta Numerica 22, 133–288, 2013.
24. J. Dick, G. Larcher, F. Pillichshammer and H. Woźniakowski, Exponential convergence and tractability of multivariate integration for Korobov spaces, Math. Comp. 80, 905–930, 2011.
25. J. Dick and F. Pillichshammer, Digital Nets and Sequences. Discrepancy Theory and Quasi-Monte Carlo Integration, Cambridge University Press, 2010.
26. J. Dick and F. Pillichshammer, Discrepancy theory and quasi-Monte Carlo integration, in: Panorama in Discrepancy Theory, W. Chen, A. Srivastav, G. Travaglini (eds.), Springer Verlag, to appear.
27. J. Dick and F. Pillichshammer, The weighted star discrepancy of Korobov’s $p$-sets, preprint.
28. J. Dick, I. H. Sloan, X. Wang and H. Woźniakowski, Liberating the weights, J. Complexity 20, 593–623, 2004.
29. J. Dick, I. H. Sloan, X. Wang and H. Woźniakowski, Good lattice rules in weighted Korobov spaces with general weights, Numer. Math. 103, 63–97, 2006.
30. Dün Dinh and T. Ullrich, Lower bounds for cubature errors of functions with dominating mixed smoothness in arbitrary dimension, preprint.
31. B. Doerr, A lower bound for the discrepancy of a random point set, J. Complexity 30, 16–20, 2014.
32. B. Doerr and M. Gnewuch, Construction of low-discrepancy point sets of small size by bracketing covers and dependent randomized rounding, in: Monte Carlo and Quasi-Monte Carlo Methods 2006, A. Keller, S. Heinrich, H. Niederreiter (eds.), 299–312, Springer, Berlin, 2008.
33. B. Doerr, M. Gnewuch, F. Kritzer and F. Pillichshammer, Component-by-component construction of low-discrepancy point sets of small size. Monte Carlo Meth. Appl. 14, 129–149, 2008.
34. B. Doerr, M. Gnewuch and M. Wahlström, Algorithmic construction of low-discrepancy point sets via dependent randomized rounding, J. Complexity 26, 490–507, 2010.
35. C. Doerr, M. Gnewuch and M. Wahlström, Calculation of discrepancy measures and applications, to appear in: W. W. L. Chen, A. Srivastav, G. Travaglini (eds.), Panorama of Discrepancy Theory, Springer-Verlag.
36. M. E. Dyer, Z. Füredi, C. McDiarmid, Random volumes in the $n$-cube. In: Polyhedral combinatorics (Morristown 1989), DIMACS Ser. Discrete Math. Theoret. Comput. Sci. vol. 1, 33–38, Amer. Math. Soc., Providence 1990.
37. G. Elekes, A geometric inequality and the complexity of computing volume. *Discrete Comput. Geom.* 1, 289–292, 1986.
38. K. K. Frolov, Upper bounds on the error of quadrature formulas on classes of functions, *Dokl. Akad. Nauk USSR* 231, 818–821, 1976. English translation: Soviet Math. Doklady 17, 1976.
39. K. K. Frolov, Upper bounds on the discrepancy in $L^p$, $2 \leq p < \infty$, *Dokl. Akad. Nauk USSR* 252, 805–807, 1980. English translation: Soviet Math. Doklady 18/1, 37–41, 1977.
40. M. Gnewuch, Infinite-dimensional integration on weighted Hilbert spaces, *Math. Comp.* 81, 2175–2205, 2012.
41. M. Gnewuch, Entropy, randomization, derandomization, and discrepancy, in: L. Plaskota, H. Woźniakowski (eds.), Monte Carlo and Quasi-Monte Carlo Methods 2010, Springer-Verlag, 2012, pp. 43–78.
42. M. Gnewuch, Lower error bounds for randomized multilevel and changing dimension algorithms, in: Monte Carlo and Quasi-Monte Carlo Methods 2012, J. Dick, F. Y. Kuo, G. W. Peters, I. H. Sloan (eds.), Springer Verlag, pp. 399–415.
43. M. Gnewuch, S. Mayer and K. Ritter, On weighted Hilbert spaces and integration of functions of infinitely many variables, *J. Complexity* 30, 29–47, 2014.
44. S. Heinrich, Lower bounds for the complexity of Monte Carlo function approximation, *J. Complexity* 8, 277–300, 1992.
45. S. Heinrich, Random approximation in numerical analysis, in: Functional Analysis, K. D. Bierstedt et al. (eds.), 123–171, Dekker, New York, 1994.
46. S. Heinrich, Complexity of Monte Carlo algorithms, in *The Mathematics of Numerical Analysis*, Lectures in Applied Mathematics 32, pp. 405–419, AMS-SIAM Summer Seminar, Park City, American Math. Society, Providence, RI, 1996.
47. S. Heinrich, Quantum Summation with an Application to Integration, *J. Complexity* 18, 1–50, 2001.
48. S. Heinrich, Quantum integration in Sobolev spaces, *J. Complexity* 19, 19–42, 2003.
49. S. Heinrich and E. Novak, Optimal summation and integration by deterministic, randomized, and quantum algorithms, in: Monte Carlo and Quasi-Monte Carlo Methods 2000, K.-T. Fang, F. J. Hickernell, H. Niederreiter (eds.), Springer Verlag 2002, pp. 50–62.
50. S. Heinrich, E. Novak and H. Pfeiffer, How many random bits do we need for Monte Carlo integration? In *Monte Carlo and Quasi-Monte Carlo Methods 2002*, H. Niederreiter (ed.), Springer 2004, pp. 27–49.
51. S. Heinrich, E. Novak, G. W. Wasilkowski and H. Woźniakowski, The inverse of the star-discrepancy depends linearly on the dimension, *Acta Arithmetica* 96, 279–302, 2001.
52. F. J. Hickernell, Th. Müller-Gronbach, B. Niu and Klaus Ritter, Multi-level Monte Carlo algorithms for infinite-dimensional integration on $\mathbb{R}^d$, *J. Complexity* 26, 229–254, 2010.
53. F. J. Hickernell, I. H. Sloan and G. W. Wasilkowski, On strong tractability of weighted multivariate integration, *Math. Comp.* 73, 1903–1911, 2004.
54. F. J. Hickernell and H. Woźniakowski, Integration and approximation in arbitrary dimension, *Adv. Comput. Math.* 12, 25–58, 2000.
55. F. J. Hickernell and H. Woźniakowski, Tractability of multivariate integration for periodic functions, *J. Complexity* 17, 660–682, 2001.
56. Á. Hinrichs, Covering numbers, Vapnik-Cervonenkis classes and bounds for the star discrepancy, *J. Complexity* 20, 477–483, 2004.
57. Á. Hinrichs, Optimal importance sampling for the approximation of integrals, *J. Complexity* 26, 125–134, 2010.
58. Á. Hinrichs, Discrepancy, integration and tractability, in: Monte Carlo and Quasi-Monte Carlo Methods 2012, Sydney, J. Dick, F. Kuo, G. Peters, I. Sloan (eds.), Springer Verlag, pp. 129–172, 2013.
59. Á. Hinrichs, L. Markhasin, J. Oettershagen and T. Ullrich, Optimal quasi-Monte Carlo rules on higher order digital nets for the numerical integration of multivariate periodic functions, in preparation.
60. Á. Hinrichs, E. Novak, M. Ullrich and H. Woźniakowski, The curse of dimensionality for numerical integration of smooth functions, *Math. Comp.*
61. A. Hinrichs, E. Novak, M. Ullrich and H. Woźniakowski, The curse of dimensionality for numerical integration of smooth functions II, *J. Complexity* **30**, 117–143, 2014.
62. A. Hinrichs, E. Novak and M. Ullrich, On weak tractability of the Clenshaw Curtis Smolyak algorithm, *J. Approx. Th.* **183**, 31–44, 2014.
63. D. Huybrechs and S. Olver, High oscillatory quadrature, *London Math. Soc. Lecture Note Ser.* **366**, pp. 25–50, Cambridge, 2009.
64. P. Kritzer, F. Pillichshammer and H. Woźniakowski, Multivariate integration of infinitely many times differentiable functions in weighted Korobov spaces, *Math. Comp.*
65. P. Kritzer, F. Pillichshammer and H. Woźniakowski, Tractability of multivariate analytic problems, to appear, De Gruyter, 2014.
66. F. Y. Kuo, Component-by-component constructions achieve the optimal rate of convergence for multivariate integration in weighted Korobov and Sobolev spaces, *J. Complexity* **19**, 301–320, 2003.
67. F. Y. Kuo, I. H. Sloan, G. W. Wasilkowski and H. Woźniakowski, Liberating the dimension, *J. Complexity* **26**, 422–454, 2010.
68. F. Y. Kuo, G. W. Wasilkowski and B. J. Waterhouse, Randomly shifted lattice rules for unbounded integrands, *J. Complexity* **22**, 630–651, 2006.
69. G. Leobacher and F. Pillichshammer, *Introduction to Quasi-Monte Carlo Integration and Applications*, Springer Verlag, 2014.
70. Maung Zho Newn and I. F. Sharygin, Optimal cubature formulas in the classes $D_{k}^{1,c}$ and $D_{k}^{1,l}$, in: *Problems of Numerical and Applied Mathematics*, Institute of Cybernetics, Uzbek Academy of Sciences, Tashkent (1991), pp. 22–27, in Russian.
71. H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods*, SIAM, Philadelphia, 1992.
72. E. Novak, On the power of adaption, *J. Complexity* **12**, 199–237, 1996.
73. E. Novak, Quantum complexity of integration, *J. Complexity* **17**, 2–16, 2001.
74. E. Novak and K. Ritter, High dimensional integration of smooth functions over cubes, *Numer. Math.* **75**, 79–97, 1996.
75. E. Novak and K. Ritter, The curse of dimension and a universal method for numerical integration, in: *Multivariate Approximation and Splines*, G. Nürnberger, J. W. Schmidt and G. Walz, eds., ISNM **125**, Birkhäuser, Basel, 177–188, 1997.
76. E. Novak and K. Ritter, Simple cubature formulas with high polynomial exactness, *Constr. Approx.* **15**, 499–522, 1999.
77. E. Novak and H. Woźniakowski, Tractability of Multivariate Problems, Volume I: Linear Information, European Math. Soc., Zürich, 2008.
78. E. Novak and H. Woźniakowski, Tractability of Multivariate Problems, Volume II: Standard Information for Functionals, European Math. Soc. Publ. House, Zürich, 2010.
79. E. Novak and H. Woźniakowski, Complexity of oscillatory integration for univariate Sobolev spaces, *J. Complexity* **17**, 388–441, 2001.
80. E. Novak and H. Woźniakowski, *Tractability of Multivariate Problems*, Volume I: Linear Information, European Math. Soc., Zürich, 2008.
81. E. Novak and H. Woźniakowski, *Tractability of Multivariate Problems*, Volume II: Standard Information for Operators, European Math. Soc. Publ. House, Zürich, 2012.
90. D. Nuyens and R. Cools, Fast algorithms for component-by-component construction of rank-1 lattice rules in shift invariant reproducing kernel Hilbert spaces, *Math. Comp.* 75, 903–920, 2006.

91. D. Nuyens and R. Cools, Fast algorithms for component-by-component construction of rank-1 lattice rules with a non-prime number of points, *J. Complexity* 22, 4–28, 2006.

92. L. Plaskota and G. W. Wasilkowski, The power of adaptive algorithms for functions with singularities, *J. Fixed Point Theory Appl.* 6, 227–248, 2009.

93. L. Plaskota and G. W. Wasilkowski, The power of adaptive algorithms for functions with singularities, *J. Complexity* 22, 4–28, 2006.

94. L. Plaskota and G. W. Wasilkowski, Tractability of infinite-dimensional integration in the worst case and randomized settings, *J. Complexity* 27, 505–518, 2011.

95. K. F. Roth, On irregularities of distributions, *Mathematika* 1, 73–79, 1954.

96. K. F. Roth, On irregularities of distributions IV, *Acta Arith.* 37, 67–75, 1980.

97. K. F. Roth, Explicit error bounds for Markov chain Monte Carlo, *Dissertationes Mathematicae* 485, Warszawa, 2012.

98. W. Sickel and T. Ullrich, Smolyak’s algorithm, sampling on sparse grids and function spaces of dominating mixed smoothness, *East J. on Approx.* 13, 387–425, 2007.

99. M. M. Skriganov, Constructions of uniform distributions in terms of geometry of numbers, *St. Petersburg Math. J.* 6, 635–664, 1995.

100. I. H. Sloan, F. Y. Kuo and S. Joe, On the step-by-step construction of quasi-Monte Carlo integration rules that achieves strong tractability error bounds in weighted Sobolev spaces, *Math. Comp.* 71, 1609–1640, 2002.

101. I. H. Sloan and A. V. Reztsov, Component-by-component construction of good lattice rules, *Math. Comp.* 71, 263–273, 2002.

102. I. H. Sloan, X. Wang and H. Woźniakowski, Finite-order weights imply tractability of multivariate integration, *J. Complexity* 20, 46–74, 2004.

103. I. H. Sloan and H. Woźniakowski, When are quasi-Monte Carlo algorithms efficient for high dimensional integrals? *J. Complexity* 14, 1–33, 1998.

104. A. G. Sukharev, Optimal numerical integration formulas for some classes of functions, *Dokl. Akad. Nauk SSSR* 4, 240–243, 1963.

105. A. G. Sukharev, Optimal numerical integration formulas for some classes of functions, *Sov. Math. Dokl.* 20, 472–475, 1979.

106. V. N. Temlyakov, Approximate recovery of periodic functions of several variables, *Math. USSR Sb.* 56, 249–261, 1987.

107. V. N. Temlyakov, On a way of obtaining lower estimates for the error of quadrature formulas, *Math. USSR Sb.* 181, 1403–1413, 1990, [in Russian]; *Math. USSR Sb.* 71, 247–257, 1992 [in English].

108. V. N. Temlyakov, On approximate recovery of functions with bounded mixed derivative, *J. Complexity* 9, 41–59, 1993.

109. V. N. Temlyakov, Cubature formulas, discrepancy, and nonlinear approximation, *J. Complexity* 19, 352–391, 2003.

110. J. F. Traub, G. W. Wasilkowski and H. Woźniakowski, *Information-Based Complexity*, Academic Press, New York, 1988.

111. J. F. Traub and H. Woźniakowski, A General Theory of Optimal Algorithms, Academic Press, New York, 1980.

112. J. F. Traub and H. Woźniakowski, Path integration on a quantum computer, *Quantum Inf. Process.* 1, 365–388, 2003.

113. H. Triebel, *Bases in Function Spaces, Sampling, Discrepancy, Numerical Integration*, European Mathematical Society (EMS), Zürich, 2010.

114. M. Ullrich, On “Upper error bounds for quadrature formulas on function classes” by K. K. Frolov, preprint, 2014.

115. J. Vybiral, Sampling numbers and function spaces, *J. Complexity* 23, 773–792, 2007.

116. J. Vybiral, Weak and quasi-polynomial tractability of approximation of infinitely differentiable functions, *J. Complexity* 30, 48–55, 2014.

117. G. W. Wasilkowski, Average case tractability of approximating $\infty$-vari ate functions, *Math. Comp.* 83, 1319–1336, 2014.

118. G. W. Wasilkowski and H. Woźniakowski, Explicit cost bounds of algorithms for multivariate tensor product problems, *J. Complexity* 11, 1–56, 1995.

119. G. W. Wasilkowski and H. Woźniakowski, On tractability of path integration, *J. Math. Phys.* 37, 2071–2088, 1996.

120. G. W. Wasilkowski and H. Woźniakowski, Weighted tensor-product algorithms for linear multivariate problems, *J. Complexity* 15, 402–447, 1999.