

The curse of dimensionality for the L_p -discrepancy with finite p

Erich Novak and Friedrich Pillichshammer

June 13, 2023

Abstract

The L_p -discrepancy is a quantitative measure for the irregularity of distribution of an N -element point set in the d -dimensional unit cube, which is closely related to the worst-case error of quasi-Monte Carlo algorithms for numerical integration. Its inverse for dimension d and error threshold $\varepsilon \in (0, 1)$ is the minimal number of points in $[0, 1)^d$ such that the minimal normalized L_p -discrepancy is less or equal ε . It is well known, that the inverse of L_2 -discrepancy grows exponentially fast with the dimension d , i.e., we have the curse of dimensionality, whereas the inverse of L_∞ -discrepancy depends exactly linearly on d . The behavior of inverse of L_p -discrepancy for general $p \notin \{2, \infty\}$ has been an open problem for many years. In this paper we show that the L_p -discrepancy suffers from the curse of dimensionality for all p in $(1, 2]$ which are of the form $p = 2\ell/(2\ell - 1)$ with $\ell \in \mathbb{N}$.

This result follows from a more general result that we show for the worst-case error of numerical integration in an anchored Sobolev space with anchor 0 of once differentiable functions in each variable whose first derivative has finite L_q -norm, where q is an even positive integer satisfying $1/p + 1/q = 1$.

Keywords: Discrepancy, numerical integration, curse of dimensionality, tractability, quasi-Monte Carlo

MSC 2010: 11K38, 65C05, 65Y20

1 Introduction and main result

For a set \mathcal{P} consisting of N points $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$ in the d -dimensional unit-cube $[0, 1]^d$ the local discrepancy function $\Delta_{\mathcal{P}} : [0, 1]^d \rightarrow \mathbb{R}$ is defined as

$$\Delta_{\mathcal{P}}(\mathbf{t}) = \frac{|\{k \in \{1, 2, \dots, N\} : \mathbf{x}_k \in [\mathbf{0}, \mathbf{t}]\}|}{N} - \text{volume}([\mathbf{0}, \mathbf{t}]),$$

for $\mathbf{t} = (t_1, t_2, \dots, t_d)$ in $[0, 1]^d$, where $[\mathbf{0}, \mathbf{t}] = [0, t_1) \times [0, t_2) \times \dots \times [0, t_d)$. For a parameter $p \in [1, \infty]$ the L_p -discrepancy of the point set \mathcal{P} is defined as the L_p -norm of the local discrepancy function $\Delta_{\mathcal{P}}$, i.e.,

$$L_{p,N}(\mathcal{P}) := \left(\int_{[0,1]^d} |\Delta_{\mathcal{P}}(\mathbf{t})|^p d\mathbf{t} \right)^{1/p} \quad \text{for } p \in [1, \infty),$$

and

$$L_{\infty,N}(\mathcal{P}) := \sup_{\mathbf{t} \in [0,1]^d} |\Delta_{\mathcal{P}}(\mathbf{t})| \quad \text{for } p = \infty.$$

Traditionally, the L_{∞} -discrepancy is called star-discrepancy and is denoted by $D_N^*(\mathcal{P})$ rather than $L_{\infty,N}(\mathcal{P})$. The study of L_p -discrepancy has its roots in the theory of uniform distribution modulo one; see [2, 13, 21, 23] for detailed information. It is of particular importance because of its close relation to numerical integration. We will refer to this issue in Section 2.

Since one is interested in point sets with L_p -discrepancy as low as possible it is obvious to study for $d, N \in \mathbb{N}$ the quantity

$$\text{disc}_p(N, d) := \min_{\mathcal{P}} L_p(\mathcal{P}),$$

where the minimum is extended over all N -element point sets \mathcal{P} in $[0, 1]^d$. This quantity is called the N -th minimal L_p -discrepancy in dimension d .

Traditionally, the L_p -discrepancy is studied from the point of view of a fixed dimension d and one asks for the asymptotic behavior for increasing sample sizes N . The celebrated result of Roth [31] is the most famous result in this direction and can be seen as the initial point of discrepancy theory. Later, Schmidt [36] extended Roth's lower bound to arbitrary $p > 1$. For $p \in (1, \infty)$ it is known that for every dimension $d \in \mathbb{N}$ there exist positive reals $c_{d,p}, C_{d,p}$ such that for every $N \geq 2$ it holds true that

$$c_{d,p} \frac{(\log N)^{\frac{d-1}{2}}}{N} \leq \text{disc}_p(N, d) \leq C_{d,p} \frac{(\log N)^{\frac{d-1}{2}}}{N}.$$

The upper bound was proven by Davenport [7] for $p = 2$, $d = 2$, by Roth [32] for $p = 2$ and arbitrary d and finally by Chen [4] in the general case. For more details see [2]. Explicit constructions of point sets can be found in [6, 7, 8, 11, 22, 33].

Similar results, but less accurate, are available also for $p \in \{1, \infty\}$. See the above references for further information. The currently best asymptotical lower bound in the L_∞ -case can be found in [3].

All the classical bounds have a poor dependence on the dimension d . For large d these bounds are only meaningful in an asymptotic sense (for very large N) and do not give any information about the discrepancy in the pre-asymptotic regime (see, e.g., [27, 28] or [10, Section 1.7] for discussions). Nowadays, motivated by applications of point sets with low discrepancy for numerical integration, there is dire need of information about the dependence of discrepancy on the dimension.

This problem is studied with the help of the so-called inverse of L_p -discrepancy (or, in a more general context, the information complexity; see Section 2). This concept compares the minimal L_p -discrepancy with the initial discrepancy

$$\text{disc}_p(0, d) := \left(\int_{[0,1]^d} (\text{volume}([0, \mathbf{t}]))^p d\mathbf{t} \right)^{1/p},$$

which can be interpreted as the L_p -discrepancy of the empty point set, and asks for the minimal number N of nodes that is necessary in order to achieve that the N -th minimal L_p -discrepancy is smaller than ε times $\text{disc}_p(0, d)$ for a threshold $\varepsilon \in (0, 1)$. In other words, for $d \in \mathbb{N}$ and $\varepsilon \in (0, 1)$ the inverse of the minimal L_p -discrepancy is defined as

$$N_p^{\text{disc}}(\varepsilon, d) := \min\{N \in \mathbb{N} : \text{disc}_p(N, d) \leq \varepsilon \text{disc}_p(0, d)\}.$$

The question is now how fast $N_p^{\text{disc}}(\varepsilon, d)$ increases, when $d \rightarrow \infty$ and $\varepsilon \rightarrow 0$.

It is well known and easy to check that for the initial L_p -discrepancy we have

$$\text{disc}_p(0, d) = \begin{cases} \frac{1}{(p+1)^{d/p}} & \text{if } p \in [1, \infty), \\ 1 & \text{if } p = \infty. \end{cases} \quad (1)$$

Here we observe a difference in the cases of finite and infinite p . While for $p = \infty$ the initial discrepancy equals 1 for every dimension d , for finite values of p the initial discrepancy tends to zero exponentially fast with the dimension.

For $p \in \{2, \infty\}$ the behavior of the inverse of the minimal L_p -discrepancy is well understood. In the L_2 -case it is known that for all $\varepsilon \in (0, 1)$ we have

$$(1.125)^d(1 - \varepsilon^2) \leq N_2^{\text{disc}}(\varepsilon, d) \leq 1.5^d \varepsilon^{-2}.$$

Here the lower bound was first shown by Woźniakowski in [38] (see also [26, 35]) and the upper bound follows from an easy averaging argument, see, e.g., [28, Sec. 9.3.2].

In the L_∞ -case it was shown by Heinrich, Novak, Wasilkowski and Woźniakowski in [16] that there exists an absolute positive constant C such that for every $d \in \mathbb{N}$ and $\varepsilon \in (0, 1)$ we have

$$N_\infty^{\text{disc}}(\varepsilon, d) \leq Cd\varepsilon^{-2}.$$

The currently smallest known value of C is $6.23401\dots$ as shown in [15] (thereby improving on other results from [1, 12, 14, 29]). On the other hand, there exist numbers $c > 0$ and $\varepsilon_0 \in (0, 1)$ such that for all $d \in \mathbb{N}$ and all $\varepsilon \in (0, \varepsilon_0)$ we have

$$N_\infty^{\text{disc}}(\varepsilon, d) \geq cd\varepsilon^{-1},$$

as shown by Hinrichs in [17].

So while the inverse of L_2 -discrepancy grows exponentially fast with the dimension d , the inverse of star-discrepancy depends only linearly on the dimension d . One says that the L_2 -discrepancy suffers from the curse of dimensionality. In information based complexity theory the behavior of the inverse of star-discrepancy is called “polynomial tractability” (see, e.g., [27, 28]).

As we see, the situation is clear (and quite different) for $p \in \{2, \infty\}$. But what happens for all other $p \notin \{2, \infty\}$? This question has been open for many years.

Just for completeness we remark that the problem has been considered also for other (semi) norms of the local discrepancy function rather than L_p -norms, that are in some sense “close” to the L_∞ -norm. Positive results (tractability) have been obtained for exponential Orlicz norms in [9]. On the other hand, for the BMO-seminorm the curse of dimensionality has been shown in [30].

All known proofs for lower bounds on the inverse of L_2 -discrepancy are based on Hilbert-space methods. A very powerful tool in this context developed in [26] (see also [28, Chapter 12]) is the method of decomposable

reproducing kernels. Unfortunately, it is not obvious how these L_2 -methods could be applied to the general L_p -case directly. So one has to find a new way or one has to figure out the essence of the Hilbert-space based proofs with the hope to get rid of all L_2 -specific factors in order to find a way to extend these proofs to the general L_p -case. We shall follow the latter path.

Our main result is formulated in the following theorem.

Theorem 1. *For every p of the form $p = \frac{2^\ell}{2^\ell - 1}$ with $\ell \in \mathbb{N}$ there exists a real C_p that is strictly larger than 1, such that for all $\varepsilon \in (0, 1)$ we have*

$$N_p^{\text{disc}}(\varepsilon, d) \geq C_p^d(1 + o(1)) \quad \text{for } d \rightarrow \infty.$$

In particular, for all these p the L_p -discrepancy suffers from the curse of dimensionality.

At first glance, one would think that the result could be easily extended to any number $p \in (1, 2]$ by means of monotonicity of the L_p -norm and squeezing any p between two values of the form given in the theorem above. However, notice that we have to take the normalized discrepancy into account, which destroys monotonicity. Having a more careful look at the problem shows that it might be not so easy to follow this first intuition and in fact, so far we did not succeed in showing the curse of dimensionality for all $p \in (1, 2]$.

Theorem 1 will follow from a more general result about the integration problem in the anchored Sobolev space with a q -norm that will be introduced and discussed in the following Section 2. This result will be stated as Theorem 3.

Sometimes a generalized notion of L_p -discrepancy is studied, where every point \mathbf{x}_k is equipped with an own weight a_k (rather than the weight $1/N$ for every point). The result from Theorem 1 even holds for this generalized L_p -discrepancy. This more general conclusion will be stated in Section 4 as Theorem 4.

We add a conjecture: We guess that the curse holds for all p with $1 < p < \infty$.

2 Relation to numerical integration

It is well known that the L_p -discrepancy is related to multivariate integration (see, e.g., [28, Chapter 9]). From now on let $p, q \geq 1$ be such that $1/p + 1/q =$

1. For $d = 1$ let $W_q^1([0, 1])$ be the space of absolutely continuous functions whose first derivatives belong to the space $L_q([0, 1])$. For $d > 1$ consider the d -fold tensor product space which is denoted by

$$W_q^{\mathbf{1}} := W_q^{(1,1,\dots,1)}([0, 1]^d)$$

and which is the Sobolev space of functions on $[0, 1]^d$ that are once differentiable in each variable and whose first derivative $\partial^d f / \partial \mathbf{x}$ has finite L_q -norm, where $\partial \mathbf{x} = \partial x_1 \partial x_2 \dots \partial x_d$. Now consider the subspace of functions that satisfy the boundary conditions $f(\mathbf{x}) = 0$ if at least one component of $\mathbf{x} = (x_1, \dots, x_d)$ equals 0 and equip this subspace with the norm

$$\|f\|_{d,q} := \left(\int_{[0,1]^d} \left| \frac{\partial^d}{\partial \mathbf{x}} f(\mathbf{x}) \right|^q d\mathbf{x} \right)^{1/q} \quad \text{for } q \in [1, \infty),$$

and

$$\|f\|_{d,\infty} := \sup_{\mathbf{x} \in [0,1]^d} \left| \frac{\partial^d}{\partial \mathbf{x}} f(\mathbf{x}) \right| \quad \text{for } q = \infty.$$

That is, consider the space

$$F_{d,q} := \{f \in W_q^{\mathbf{1}} : f(\mathbf{x}) = 0 \text{ if } x_j = 0 \text{ for some } j \in [d] \text{ and } \|f\|_{d,q} < \infty\},$$

where here and throughout this paper we write $[d] := \{1, 2, \dots, d\}$.

Now consider multivariate integration

$$I_d(f) := \int_{[0,1]^d} f(\mathbf{x}) d\mathbf{x} \quad \text{for } f \in F_{d,q}.$$

We approximate the integrals $I_d(f)$ by algorithms based on N function evaluations of the form

$$A_{d,N}(f) = \varphi(f(\mathbf{x}_1), f(\mathbf{x}_2), \dots, f(\mathbf{x}_N)), \quad (2)$$

where $\varphi : \mathbb{R}^N \rightarrow \mathbb{R}$ is an arbitrary function and where $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$ are points in $[0, 1]^d$, called integration nodes. Since $f \equiv 0$ belongs to $F_{d,q}$ we choose φ such that $\varphi(0, 0, \dots, 0) = 0$. Typical examples are linear algorithms of the form

$$A_{d,N}^{\text{lin}}(f) = \sum_{k=1}^N a_k f(\mathbf{x}_k), \quad (3)$$

where $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$ are in $[0, 1]^d$ and a_1, a_2, \dots, a_N are real weights that we call integration weights. If $a_1 = a_2 = \dots = a_N = 1/N$, then the linear algorithm (3) is a so-called quasi-Monte Carlo algorithm, for which we will write $A_{d,N}^{\text{QMC}}$.

Define the worst-case error of an algorithm (2) by

$$e(F_{d,q}, A_{d,N}) = \sup_{\substack{f \in F_{d,q} \\ \|f\|_{d,q} \leq 1}} |I_d(f) - A_{d,N}(f)|. \quad (4)$$

For a quasi-Monte Carlo algorithm $A_{d,N}^{\text{QMC}}$ it is well known (see, e.g., [35]) that

$$e(F_{d,q}, A_{d,N}^{\text{QMC}}) = L_p(\overline{\mathcal{P}}),$$

where $L_p(\overline{\mathcal{P}})$ is the L_p -discrepancy of the point set¹

$$\overline{\mathcal{P}} = \{\mathbf{1} - \mathbf{x}_k : k = 1, 2, \dots, N\}, \quad (5)$$

where $\mathbf{1} - \mathbf{x}_k$ is defined as the component-wise difference of the vector containing only ones and \mathbf{x}_k . From this point of view we now study the more general problem of numerical integration in $F_{d,q}$ rather than only the L_p -discrepancy (which corresponds to quasi-Monte Carlo algorithms – although with suitably “reflected” points).

We define the N -th minimal worst-case error as

$$e_q(N, d) := \min_{A_{d,N}} |e(F_{d,q}, A_{d,N})|$$

where the minimum is extended over all algorithms of the form (2) based on N function evaluations along points $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$ from $[0, 1]^d$. Note that for all $d, N \in \mathbb{N}$ we have

$$e_q(N, d) \leq \text{disc}_p(N, d). \quad (6)$$

The initial error is

$$e_q(0, d) = \sup_{\substack{f \in F_{d,q} \\ \|f\|_{d,q} \leq 1}} |I_d(f)|.$$

¹We remark that in [35] the anchored space with anchor $\mathbf{1}$ is considered which results in a worst case error of exactly $L_p(\mathcal{P})$, where \mathcal{P} is the node set of the QMC rule. Here we have chosen the anchor as $\mathbf{0}$, and therefore in the formula for the worst-case error the point set $\overline{\mathcal{P}}$ appears. For details see also [28, Section 9.5.1] for the case $p = 2$.

Lemma 2. *Let $d \in \mathbb{N}$ and let $q \in (1, \infty]$ and $p \in [1, \infty)$ with $1/p + 1/q = 1$. Then we have*

$$e_q(0, d) = \frac{1}{(p+1)^{d/p}}$$

and the worst-case function in $F_{d,q}$ is given by $h_d(\mathbf{x}) = h_1(x_1) \cdots h_1(x_d)$ for $\mathbf{x} = (x_1, \dots, x_d) \in [0, 1]^d$, where $h_1(x) = 1 - (1-x)^p$.

Proof. Since we are dealing with tensor products of one-dimensional spaces it suffices to prove the result for $d = 1$. For $f \in F_{1,q}$ we have

$$\int_0^1 f(x) dx = \int_0^1 \int_0^x f'(t) dt dx = \int_0^1 f'(t)g(t) dt,$$

where $g(t) = 1 - t$. Applying Hölder's inequality we obtain

$$\left| \int_0^1 f(x) dx \right| \leq \|f'\|_{L_q} \|g\|_{L_p} = \|f\|_{1,q} \|g\|_{L_p}$$

with equality if $f'(t) = c(g(t))^{p-1} = c(1-t)^{p-1}$ for some c . This holds for $f(t) = h_1(t) := 1 - (1-t)^p$.

We have

$$\int_0^1 h_1(t) dt = \frac{p}{p+1}$$

and

$$\|h_1\|_{1,q} = \left(\int_0^1 (px^{p-1})^q dx \right)^{1/q} = p \left(\frac{1}{(p-1)q+1} \right)^{1/q} = p(1+p)^{1/p-1}.$$

Hence,

$$e_q(0, 1) = \int_0^1 \frac{h_1(t)}{\|h_1\|_{1,q}} dt = \frac{1}{(p+1)^{1/p}}.$$

□

Note that for all $q \in (1, \infty]$ and $p \in [1, \infty)$ with $1/p + 1/q = 1$ and for all $d \in \mathbb{N}$ we have

$$e_q(0, d) = \text{disc}_p(0, d).$$

Now we define the information complexity as the minimal number of function evaluations necessary in order to reduce the initial error by a factor of ε . For $d \in \mathbb{N}$ and $\varepsilon \in (0, 1)$ put

$$N_q^{\text{int}}(\varepsilon, d) := \min\{N \in \mathbb{N} : e_q(N, d) \leq \varepsilon e_q(0, d)\}.$$

From (6), (1) and Lemma 2 it follows that for all $q \in (1, \infty]$ and $p \in [1, \infty)$ with $1/p + 1/q = 1$ and for all $d \in \mathbb{N}$ and $\varepsilon \in (0, 1)^d$ we have

$$N_q^{\text{int}}(\varepsilon, d) \leq N_p^{\text{disc}}(\varepsilon, d).$$

Hence, Theorem 1 follows from the following more general result.

Theorem 3. *For every positive even integer q there exists a real $C = C(q)$, which depends on q and which is strictly larger than 1, such that for all $\varepsilon \in (0, 1)$ we have*

$$N_q^{\text{int}}(\varepsilon, d) \geq C^d(1 + o(1)) \quad \text{for } d \rightarrow \infty. \quad (7)$$

In particular, for all even integer q the integration problem in $F_{d,q}$ suffers from the curse of dimensionality.

The proof of this result will be given in the following section.

Information about the base $C = C(q)$ of the exponentiation in Theorem 3 will be given in Section 4.

3 The proof of Theorem 3

The proof of Theorem 3 is based on a suitable decomposition of the worst-case function h_1 from Lemma 2. This decomposition depends on q and p , respectively, and will determine the base C of the exponentiation in (7).

Proof of Theorem 3. Let q be an even, positive integer and let p be such that $1/p + 1/q = 1$. Obviously, then $p = q/(q - 1)$ and $p \in (1, 2]$. According to Lemma 2 the univariate worst-case function equals

$$h_1(x) = 1 - (1 - x)^p$$

and

$$\|h_1\|_{1,q} = \frac{p}{(p+1)^{(p-1)/p}}.$$

Now we decompose h_1 in a suitable way into a sum of three functions. For real parameters c and a , with $a \in (0, 1)$ and $c \in (0, a/11)$ that will be determined

later put

$$h_{1,2,(0)}(x) = \begin{cases} \frac{1-(1-x)^p}{2} & \text{if } x \in [0, c), \\ \frac{1-(1-c)^p}{2} & \text{if } x \in [c, a-10c), \\ \frac{1-(1-c)^p}{20c}(a-x) & \text{if } x \in [a-10c, a), \\ 0 & \text{if } x \in [a, 1], \end{cases}$$

and

$$h_{1,2,(1)}(x) = \mathbf{1}_{[a,1]}(x) ((1-a)^p - (1-x)^p),$$

and

$$h_{1,1}(x) = \begin{cases} 1 - (1-x)^p - h_{1,2,(0)}(x) & \text{if } x \in [0, a), \\ 1 - (1-a)^p & \text{if } x \in [a, 1]. \end{cases}$$

Then we have

$$h_{1,1}(x) + h_{1,2,(0)}(x) + h_{1,2,(1)}(x) = 1 - (1-x)^p = h_1(x)$$

and $h_{1,2,(0)}(x)$ and $h_{1,2,(1)}(x)$ have disjoint support as well as the derivatives $h'_{1,1}$ and $h'_{1,2,(1)}$. See Figure 1 for a graphical illustration of the decomposition.

Let $d \in \mathbb{N}$. For $\mathbf{u} \subseteq [d]$ let $\mathbf{u}^c = [d] \setminus \mathbf{u}$. Now, for $\mathbf{x} = (x_1, \dots, x_d) \in [0, 1]^d$ let

$$\begin{aligned} h_d(\mathbf{x}) &= \prod_{j=1}^d h_1(x_j) \\ &= \prod_{j=1}^d (h_{1,1}(x_j) + h_{1,2,(0)}(x_j) + h_{1,2,(1)}(x_j)) \\ &= \sum_{\mathbf{u} \subseteq [d]} \prod_{j \in \mathbf{u}^c} h_{1,1}(x_j) \sum_{\mathbf{v} \subseteq \mathbf{u}} \prod_{j \in \mathbf{v}} h_{1,2,(0)}(x_j) \prod_{j \in \mathbf{u} \setminus \mathbf{v}} h_{1,2,(1)}(x_j). \end{aligned}$$

We have

$$\|h_d\|_{d,q} = \|h_1\|_{1,q}^d = \left(\frac{p}{(p+1)^{(p-1)/p}} \right)^d.$$

For a point $\mathbf{x} \in [0, 1]^d$ and for $\mathbf{u} \subseteq [d]$ let $\mathbf{x}(\mathbf{u})$ denote the projection of \mathbf{x} on the coordinates $j \in \mathbf{u}$.

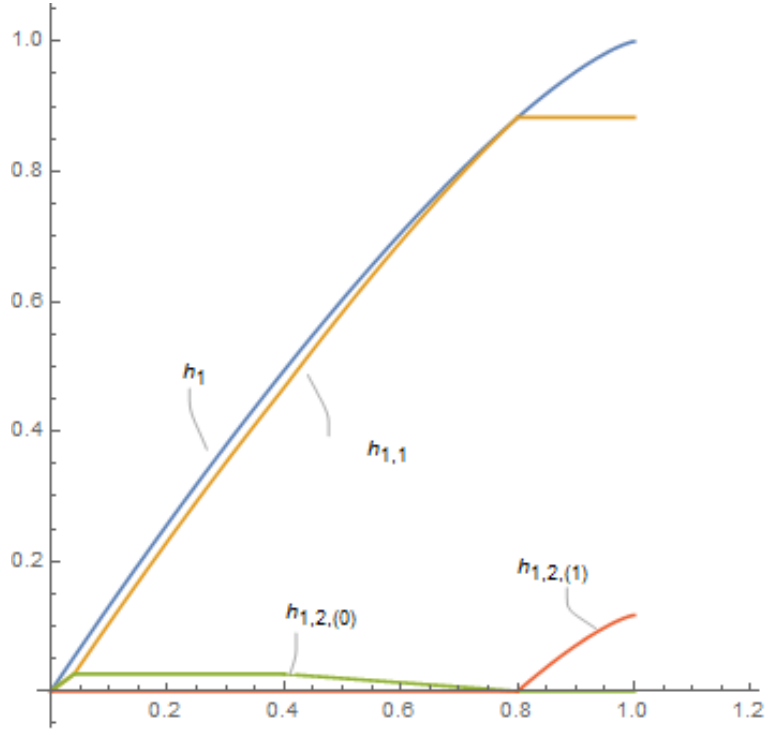


Figure 1: Plot of the functions $h_1, h_{1,1}, h_{1,2,(0)}$, and $h_{1,2,(1)}$ for $q = 4$ (hence $p = 4/3$) and $a = 0.8$ and $c = 0.04$ (later we will use a different choice of parameters; but with the ones chosen here one gets a better impression of the principle situation).

Now let $N \in \mathbb{N}$ and let $\mathcal{P} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ be a set of N quadrature nodes in $[0, 1]^d$. For this specific choice define the “fooling function”

$$g_d(\mathbf{x}) = \sum_{\mathbf{u} \subseteq [d]} \prod_{j \notin \mathbf{u}} h_{1,1}(x_j) \sum_{\mathbf{v} \subseteq \mathbf{u}}^* \prod_{j \in \mathbf{v}} h_{1,2,(0)}(x_j) \prod_{j \in \mathbf{u} \setminus \mathbf{v}} h_{1,2,(1)}(x_j), \quad (8)$$

where the $*$ in $\sum_{\mathbf{v} \subseteq \mathbf{u}}^*$ indicates that the summation is restricted to all $\mathbf{v} \subseteq \mathbf{u}$ such that for all $k \in \{1, 2, \dots, N\}$ it holds that

$$(\mathbf{x}_k(\mathbf{v}), \mathbf{x}_k(\mathbf{u} \setminus \mathbf{v})) \notin [0, a]^{|\mathbf{v}|} \times [a, 1]^{|\mathbf{u}| - |\mathbf{v}|}.$$

This condition means that every point from the set \mathcal{P} is outside the support of $\prod_{j \notin \mathbf{u}} h_{1,1}(x_j) \prod_{j \in \mathbf{v}} h_{1,2,(0)}(x_j) \prod_{j \in \mathbf{u} \setminus \mathbf{v}} h_{1,2,(1)}(x_j)$ and this guarantees that

$$g_d(\mathbf{x}_k) = 0 \quad \text{for all } k \in \{1, 2, \dots, N\}.$$

Observe that the fooling function g_d is an element of the 3^d -dimensional (tensor product) subspace generated, for $d = 1$, by the three functions $h_{1,1}$, $h_{1,2,(0)}$ and $h_{1,2,(1)}$. Formally, we prove the lower bound for this finite-dimensional subspace. We use here notation from the theory of decomposable kernels (see, e.g., [28, Section 11.5]). This theory can only be applied for $q = p = 2$ and then one can choose the three functions orthogonal which simplifies the proof a lot (see also Section 5).

Now we show that $\|g_d\|_{d,q} \leq \|h_d\|_{d,q}$. This is the only part of the proof

where we need that q is an even integer. We have

$$\begin{aligned}
& \|h_d\|_{d,q}^q \\
&= \int_{[0,1]^d} \left| \sum_{\mathbf{u} \subseteq [d]} \prod_{j \notin \mathbf{u}} h'_{1,1}(x_j) \sum_{\mathbf{v} \subseteq \mathbf{u}} \prod_{j \in \mathbf{v}} h'_{1,2,(0)}(x_j) \prod_{j \in \mathbf{u} \setminus \mathbf{v}} h'_{1,2,(1)}(x_j) \right|^q d\mathbf{x} \\
&= \sum_{\mathbf{u}_1 \subseteq [d]} \sum_{\mathbf{v}_1 \subseteq \mathbf{u}_1} \sum_{\mathbf{u}_2 \subseteq [d]} \sum_{\mathbf{v}_2 \subseteq \mathbf{u}_2} \cdots \sum_{\mathbf{u}_q \subseteq [d]} \sum_{\mathbf{v}_q \subseteq \mathbf{u}_q} \int_{[0,1]^d} \\
&\quad \prod_{j \in \mathbf{u}_1^c} h'_{1,1}(x_j) \prod_{j \in \mathbf{u}_2^c} h'_{1,1}(x_j) \cdots \prod_{j \in \mathbf{u}_q^c} h'_{1,1}(x_j) \\
&\quad \prod_{j \in \mathbf{v}_1} h'_{1,2,(0)}(x_j) \prod_{j \in \mathbf{v}_2} h'_{1,2,(0)}(x_j) \cdots \prod_{j \in \mathbf{v}_q} h'_{1,2,(0)}(x_j) \\
&\quad \prod_{j \in \mathbf{u}_1 \setminus \mathbf{v}_1} h'_{1,2,(1)}(x_j) \prod_{j \in \mathbf{u}_2 \setminus \mathbf{v}_2} h'_{1,2,(0)}(x_j) \cdots \prod_{j \in \mathbf{u}_q \setminus \mathbf{v}_q} h'_{1,2,(1)}(x_j) d\mathbf{x} \\
&= \sum_{\mathbf{u}_1 \subseteq [d]} \sum_{\mathbf{v}_1 \subseteq \mathbf{u}_1} \sum_{\mathbf{u}_2 \subseteq [d]} \sum_{\mathbf{v}_2 \subseteq \mathbf{u}_2} \cdots \sum_{\mathbf{u}_q \subseteq [d]} \sum_{\mathbf{v}_q \subseteq \mathbf{u}_q} \prod_{j=1}^d \int_0^1 h'_{1,1}(x)^{\alpha_j} h'_{1,2,(0)}(x)^{\beta_j} h'_{1,2,(1)}(x)^{\gamma_j} dx,
\end{aligned}$$

where, for $j \in [d]$, we put

$$\begin{aligned}
\alpha_j &= |\{t \in \{1, 2, \dots, q\} : j \in \mathbf{u}_t^c\}|, \\
\beta_j &= |\{t \in \{1, 2, \dots, q\} : j \in \mathbf{v}_t\}|, \\
\gamma_j &= |\{t \in \{1, 2, \dots, q\} : j \in \mathbf{u}_t \setminus \mathbf{v}_t\}|.
\end{aligned}$$

Note that $\alpha_j, \beta_j, \gamma_j$ also depend on $\mathbf{u}_1, \mathbf{v}_1, \mathbf{u}_2, \mathbf{v}_2, \dots, \mathbf{u}_q, \mathbf{v}_q$. We have $\alpha_j, \beta_j, \gamma_j \in \{0, 1, \dots, q\}$ and $\alpha_j + \beta_j + \gamma_j = q$.

For given $\mathbf{u}_t \subseteq [d], \mathbf{v}_t \subseteq \mathbf{u}_t$ for $t \in \{1, 2, \dots, q\}$, put

$$I(\alpha_j, \beta_j, \gamma_j) = \int_0^1 h'_{1,1}(x)^{\alpha_j} h'_{1,2,(0)}(x)^{\beta_j} h'_{1,2,(1)}(x)^{\gamma_j} dx.$$

Now we show that $I(\alpha_j, \beta_j, \gamma_j) \geq 0$ for all $\alpha_j, \beta_j, \gamma_j \in \{0, 1, \dots, q\}$ with $\alpha_j + \beta_j + \gamma_j = q$. This will help us later on when we show that the norm of g_d is dominated by the norm of h_d .

If $\alpha_j > 0$ and $\gamma_j > 0$, then we have $I(\alpha_j, \beta_j, \gamma_j) = 0$, because $h'_{1,1}$ and $h'_{1,2,(1)}$ have disjoint support. Likewise, if $\beta_j > 0$ and $\gamma_j > 0$, then we have $I(\alpha_j, \beta_j, \gamma_j) = 0$, because $h'_{1,2,(0)}$ and $h'_{1,2,(1)}$ have disjoint support.

If $(\alpha_j, \beta_j, \gamma_j) \in \{(0, 0, q), (0, q, 0), (q, 0, 0)\}$, then obviously $I(\alpha_j, \beta_j, \alpha_j) > 0$.

It remains to consider the case $\alpha_j > 0, \beta_j > 0$ and $\gamma_j = 0$. In this case $\alpha_j + \beta_j = q$. If α_j (and then also β_j) is even, then $I(\alpha_j, \beta_j, \gamma_j) > 0$.

So it remains to deal with the cases where α_j is odd.

Let $a^* = 1 - 10^{-q}$ and $c^* = 10^{-q}$. Let $\alpha, \beta \in \{1, \dots, q-1\}$ be odd and assume that $\alpha + \beta = q$. We show that

$$I(\alpha, \beta, 0) \geq 0 \quad \text{for all } a \in [a^*, 1] \text{ and all } c \in [0, c^*]. \quad (9)$$

Let $a \in [a^*, 1]$ and $c \in [0, c^*]$. We have

$$\begin{aligned} I(\alpha, \beta, 0) &= \int_0^c \left(\frac{p}{2} (1-x)^{p-1} \right)^\alpha \left(\frac{p}{2} (1-x)^{p-1} \right)^\beta dx \\ &\quad + \int_c^{a-10c} 0 dx \\ &\quad + \int_{a-10c}^a \left(p(1-x)^{p-1} + \frac{1-(1-c)^p}{20c} \right)^\alpha \left(-\frac{1-(1-c)^p}{20c} \right)^\beta dx \\ &= \int_0^c \left(\frac{p}{2} (1-x)^{p-1} \right)^q dx \\ &\quad - \left(\frac{1-(1-c)^p}{20c} \right)^\beta \int_{a-10c}^a \left(p(1-x)^{p-1} + \frac{1-(1-c)^p}{20c} \right)^\alpha dx. \end{aligned}$$

Note that $(p-1)q = p$ and hence

$$\begin{aligned} I(\alpha, \beta, 0) &= \left(\frac{p}{2} \right)^q \frac{1-(1-c)^{p+1}}{p+1} \\ &\quad - \left(\frac{1-(1-c)^p}{20c} \right)^\beta \int_{a-10c}^a \left(p(1-x)^{p-1} + \frac{1-(1-c)^p}{20c} \right)^\alpha dx. \end{aligned}$$

We have $1 - (1 - c)^p \leq pc$ and $1 - a + 10c \leq 11c^*$. Hence,

$$\begin{aligned}
& \left(\frac{1 - (1 - c)^p}{20c} \right)^\beta \int_{a-10c}^a \left(p(1 - x)^{p-1} + \frac{1 - (1 - c)^p}{20c} \right)^\alpha dx \\
& \leq \left(\frac{p}{20} \right)^\beta \int_{a-10c}^a \left(p(1 - x)^{p-1} + \frac{p}{20} \right)^\alpha dx \\
& \leq \left(\frac{p}{20} \right)^\beta 10c \left(p(1 - a + 10c)^{p-1} + \frac{p}{20} \right)^\alpha \\
& = \left(\frac{p}{2} \right)^q \frac{10c}{10^\beta} \left(2(11c^*)^{p-1} + \frac{1}{10} \right)^\alpha \\
& \leq \left(\frac{p}{2} \right)^q \frac{10c}{10^q} \left(\frac{32}{10} \right)^{q-1},
\end{aligned}$$

where we used

$$2(11c^*)^{p-1} + \frac{1}{10} \leq 2 \frac{11^{p-1}}{10^{(p-1)q}} + \frac{1}{10} = 2 \frac{11^{p-1}}{10^p} + \frac{1}{10} = \frac{2}{11} \left(\frac{11}{10} \right)^p + \frac{1}{10} \leq \frac{1}{10} \frac{32}{10}.$$

Now we obtain

$$I(\alpha, \beta, 0) \geq \left(\frac{p}{2} \right)^q \left(\frac{1 - (1 - c)^{p+1}}{p+1} - \frac{c}{10^{q-1}} \left(\frac{32}{10} \right)^{q-1} \right).$$

We use $1 - (1 - c)^{p+1} \geq (p+1)(1 - c)^p c$. Then we get

$$\begin{aligned}
I(\alpha, \beta, 0) & \geq \left(\frac{p}{2} \right)^q c \left((1 - c)^p - \frac{1}{10^{q-1}} \left(\frac{32}{10} \right)^{q-1} \right) \\
& \geq \left(\frac{p}{2} \right)^q c \left(\left(1 - \frac{1}{10^q} \right)^p - \left(\frac{32}{100} \right)^{q-1} \right) \\
& = \left(\frac{p}{2} \right)^q c \left(\left(1 - \frac{1}{10^q} \right)^{q/(q-1)} - \left(\frac{32}{100} \right)^{q-1} \right) \\
& \geq \left(\frac{p}{2} \right)^q c \frac{6601}{10000} \geq 0.
\end{aligned}$$

Hence (9) is shown.

From now on let $a \in [a^*, 1)$ and $c \in (0, c^*]$. We summarize: for $\alpha_j, \beta_j, \gamma_j \in \{0, 1, \dots, q\}$ and $\alpha_j + \beta_j + \gamma_j = q$ we have the following situation:

α_j	β_j	γ_j	$I(\alpha_j, \beta_j, \gamma_j)$	reason
$= 0$	$= 0$	> 0	> 0	$\gamma_j = q$
$= 0$	> 0	$= 0$	> 0	$\beta_j = q$
0	> 0	> 0	$= 0$	disjoint support of $h_{1,2,(0)}$ and $h_{1,2,(1)}$
> 0	$= 0$	$= 0$	> 0	$\alpha_j = q$
> 0	$= 0$	> 0	$= 0$	disjoint support of $h'_{1,1}$ and $h'_{1,2,(1)}$
> 0	> 0	$= 0$	> 0	special choice $a \in [a^*, 1)$ and $c \in (0, c^*]$
> 0	> 0	> 0	$= 0$	disjoint support of $h_{1,2,(0)}$ and $h_{1,2,(1)}$

Thus, in any case we have $I(\alpha_j, \beta_j, \gamma_j) \geq 0$. Hence

$$\begin{aligned}
& \|h_d\|_{d,q}^q \\
&= \sum_{u_1 \subseteq [d]} \sum_{v_1 \subseteq u_1} \sum_{u_2 \subseteq [d]} \sum_{v_2 \subseteq u_2} \cdots \sum_{u_q \subseteq [d]} \sum_{v_q \subseteq u_q} \prod_{j=1}^d I(\alpha_j, \beta_j, \gamma_j) \\
&\geq \sum_{u_1 \subseteq [d]} \sum_{v_1 \subseteq u_1}^* \sum_{u_2 \subseteq [d]} \sum_{v_2 \subseteq u_2}^* \cdots \sum_{u_q \subseteq [d]} \sum_{v_q \subseteq u_q}^* \prod_{j=1}^d I(\alpha_j, \beta_j, \gamma_j) \\
&= \int_{[0,1]^d} \left| \sum_{u \subseteq [d]} \prod_{j \notin u} h'_{1,1}(x_j) \sum_{v \subseteq u}^* \prod_{j \in v} h'_{1,2,(0)}(x_j) \prod_{j \in u \setminus v} h'_{1,2,(1)}(x_j) \right|^q d\mathbf{x} \\
&= \|g_d\|_{d,q}^q.
\end{aligned}$$

Now set $\tilde{g}_d(\mathbf{x}) := (1/\|h_d\|_{d,q})g_d(\mathbf{x})$ such that $\|\tilde{g}_d\|_{d,q} \leq 1$.

Put, for $(a, c) \in [a^*, 1] \times [0, c^*]$,

$$\begin{aligned}
I_{1,2,(0)}(a, c) &:= \int_0^1 h_{1,2,(0)}(x) dx \\
&= \frac{(1 - (1 - c)^p)((p + 1)(a - 5c) - 1) + pc(1 - c)^p}{2(p + 1)}
\end{aligned}$$

and

$$I_{1,2,(1)}(a) := \int_0^1 h_{1,2,(1)}(x) dx = (1 - a)^{p+1} \frac{p}{p + 1}.$$

Next we show that we can choose $c \in (0, c^*)$ such that $I_{1,2,(0)}(a^*, c) = I_{1,2,(1)}(a^*)$.

We have

$$\begin{aligned}
\frac{I_{1,2,(0)}(a^*, c^*)}{I_{1,2,(1)}(a^*)} &= \frac{(1 - (1 - c^*)^p)((p + 1)(a^* - 5c^*) - 1) + pc^*(1 - c^*)^p}{2(1 - a^*)^{p+1}p} \\
&= \frac{(1 - (1 - c^*)^p)((p + 1)(1 - 6c^*) - 1) + pc^*(1 - c^*)^p}{2(c^*)^{p+1}p} \\
&\geq \frac{p(1 - c^*)^{p-1}((p + 1)(1 - 6c^*) - 1) + pc^*(1 - c^*)^p}{2(c^*)^{p+1}p} \\
&= \frac{(1 - c^*)^{p-1}}{2(c^*)^{p+1}}((p + 1)(1 - 6c^*) - 1 + c^*(1 - c^*)) \\
&\geq \frac{1 - 11c^* - (c^*)^2}{2(c^*)^2} \\
&= \frac{1}{2}10^{2q} - \frac{11}{2}10^q - \frac{1}{2} > 1.
\end{aligned}$$

Hence $I_{1,2,(0)}(a^*, c^*) > I_{1,2,(1)}(a^*)$. Now, since $I_{1,2,(0)}(a^*, 0) = 0 < I_{1,2,(1)}(a^*)$, we can find a $c \in (0, c^*)$ such that $I_{1,2,(0)}(a^*, c) = I_{1,2,(1)}(a^*)$, according to the mean value theorem.

From now on let

$$c \in (0, c^*) \text{ be such that } I_{1,2,(0)}(a^*, c) = I_{1,2,(1)}(a^*). \quad (10)$$

In the following we use $(x)_+ = x$ if $x > 0$ and 0 if $x \leq 0$. Since $\tilde{g}_d(\mathbf{x}_k) = 0$ for all $k \in \{1, 2, \dots, N\}$ and since $\varphi(0, 0, \dots, 0) = 0$ (here φ is from (2)) we have for any algorithm (2) that is based on \mathcal{P} that

$$\begin{aligned}
&e(F_{d,q}, A_{d,N}) \\
&\geq \int_{[0,1]^d} \tilde{g}_d(\mathbf{x}) \, d\mathbf{x} \\
&= \frac{1}{\|h_d\|_{d,q}} \sum_{\mathbf{u} \subseteq [d]} \left(\int_0^1 h_{1,1}(x) \, dx \right)^{d-|\mathbf{u}|} \\
&\quad \times \sum_{\mathbf{v} \subseteq \mathbf{u}}^* \left(\int_0^1 h_{1,2,(0)}(x) \, dx \right)^{|\mathbf{v}|} \left(\int_0^1 h_{1,2,(1)}(x) \, dx \right)^{|\mathbf{u}|-|\mathbf{v}|} \quad (11)
\end{aligned}$$

Put

$$\alpha_1 := \int_0^1 h_{1,1}(x) \, dx = \frac{p}{p+1}(1 - (1 - a^*)^{p+1}) - I_{1,2,(0)}(a^*, c)$$

and remember that according to our choice (10) we have

$$\int_0^1 h_{1,2,(0)}(x) \, dx = \int_0^1 h_{1,2,(1)}(x) \, dx.$$

Now we continue from (11) and obtain

$$\begin{aligned} e(F_{d,q}, A_{d,N}) &\geq \frac{1}{\|h_d\|_{d,q}} \sum_{\mathbf{u} \subseteq [d]} \alpha_1^{d-|\mathbf{u}|} \left(\int_0^1 h_{1,2,(0)}(x) \, dx \right)^{|\mathbf{u}|} \sum_{\mathbf{v} \subseteq \mathbf{u}}^* 1 \\ &\geq \frac{1}{\|h_d\|_{d,q}} \sum_{\mathbf{u} \subseteq [d]} \alpha_1^{d-|\mathbf{u}|} \left(\int_0^1 h_{1,2,(0)}(x) \, dx \right)^{|\mathbf{u}|} (2^{|\mathbf{u}|} - N)_+ \\ &= \frac{1}{\|h_d\|_{d,q}} \sum_{\mathbf{u} \subseteq [d]} \alpha_1^{d-|\mathbf{u}|} \left(2 \int_0^1 h_{1,2,(0)}(x) \, dx \right)^{|\mathbf{u}|} \left(1 - \frac{N}{2^{|\mathbf{u}|}} \right)_+, \end{aligned}$$

where we used that for every $\mathbf{u} \subseteq [d]$ at most N of the intervals

$$\prod_{j \in \mathbf{v}} [0, a^*] \prod_{j \in \mathbf{u} \setminus \mathbf{v}} [a^*, 1] \prod_{j \notin \mathbf{u}} [0, 1]$$

with $\mathbf{v} \subseteq \mathbf{u}$ can contain a point from the node set \mathcal{P}_N .

Put

$$\alpha_2 := 2 \int_0^1 h_{1,2,(0)}(x) \, dx.$$

Then we have further

$$\begin{aligned} e(F_{d,q}, A_{d,N}) &\geq \frac{1}{\|h_d\|_{d,q}} \sum_{\mathbf{u} \subseteq [d]} \alpha_1^{d-|\mathbf{u}|} \alpha_2^{|\mathbf{u}|} \left(1 - \frac{N}{2^{|\mathbf{u}|}} \right)_+ \\ &= \frac{1}{\|h_d\|_{d,q}} \sum_{k=0}^d \binom{d}{k} \alpha_1^{d-k} \alpha_2^k \left(1 - \frac{N}{2^k} \right)_+ \\ &= \frac{1}{\|h_d\|_{d,q}} \alpha_1^d \sum_{k=0}^d \binom{d}{k} \alpha_3^k \left(1 - \frac{N}{2^k} \right)_+, \end{aligned}$$

where we put

$$\alpha_3 := \frac{\alpha_2}{\alpha_1} > 0.$$

Since the lower bound on the error is independent of the choice of \mathcal{P}_N it follows that

$$e_q(N, d) \geq \frac{1}{\|h_d\|_{d,q}} \alpha_1^d \sum_{k=0}^d \binom{d}{k} \alpha_3^k \left(1 - \frac{N}{2^k}\right)_+.$$

For the normalized n -th minimal error we therefore obtain

$$\frac{e_q(N, d)}{e_q(0, d)} \geq \frac{1}{\|h_d\|_{d,q} e_q(0, d)} \alpha_1^d \sum_{k=0}^d \binom{d}{k} \alpha_3^k \left(1 - \frac{N}{2^k}\right)_+.$$

We have

$$\|h_d\|_{d,q} e_q(0, d) = \left(\frac{p}{p+1}\right)^d$$

and hence

$$\frac{e_q(N, d)}{e_q(0, d)} \geq \left(\frac{p+1}{p}\right)^d \alpha_1^d \sum_{k=0}^d \binom{d}{k} \alpha_3^k \left(1 - \frac{N}{2^k}\right)_+. \quad (12)$$

Now we proceed using ideas from [28, p. 185]. Let $b \in (0, 1)$ and let $N = \lfloor C^d \rfloor$ with $C \in (1, 2^{\alpha_3/(1+\alpha_3)})$. This means that

$$\frac{C}{2^{\alpha_3/(1+\alpha_3)}} < 1.$$

Then there exists a positive c such that

$$c < \frac{\alpha_3}{1 + \alpha_3} \quad \text{and} \quad \frac{C}{2^c} < 1.$$

Put $k(d) := \lfloor cd \rfloor$. For sufficiently large d we have

$$\frac{N}{2^k} \leq \frac{C^d}{2^{cd-1}} = 2 \left(\frac{C}{2^c}\right)^d \leq b \quad \text{for all } k \in (k(d), d]. \quad (13)$$

Put further $C_{d,k} := \binom{d}{k} \alpha_3^k$. Then (12) and (13) imply that

$$\begin{aligned} \frac{e_q(\lfloor C^d \rfloor, d)}{e_q(0, d)} &\geq \left(\frac{p+1}{p}\right)^d \alpha_1^d \sum_{k=k(d)+1}^d C_{d,k} \left(1 - \frac{N}{2^k}\right)_+ \\ &\geq (1-b) \left(\frac{p+1}{p}\right)^d \alpha_1^d \sum_{k=k(d)+1}^d C_{d,k} \\ &= (1-b) \left(\frac{p+1}{p}\right)^d \alpha_1^d \left(\sum_{k=0}^d C_{d,k} - \sum_{k=0}^{k(d)} C_{d,k} \right). \end{aligned}$$

Note that

$$\sum_{k=0}^d C_{d,k} = (1 + \alpha_3)^d = \frac{(\alpha_1 + \alpha_2)^d}{\alpha_1^d}$$

and

$$\begin{aligned} \alpha_1 + \alpha_2 &= \int_0^1 h_{1,1}(x) \, dx + 2 \int_0^1 h_{1,2,(0)}(x) \, dx \\ &= \int_0^1 h_{1,1}(x) \, dx + \int_0^1 h_{1,2,(0)}(x) \, dx + \int_0^1 h_{1,2,(1)}(x) \, dx \\ &= \int_0^1 h_1(x) \, dx \\ &= \int_0^1 1 - (1 - x)^p \, dx \\ &= \frac{p}{p+1}. \end{aligned}$$

This yields

$$\begin{aligned} \frac{e_q(\lfloor C^d \rfloor, d)}{e_q(0, d)} &\geq (1 - b) \left(\frac{p+1}{p} (\alpha_1 + \alpha_2) \right)^d \left(1 - \frac{\sum_{k=0}^{k(d)} C_{d,k}}{(1 + \alpha_3)^d} \right) \\ &= (1 - b) \left(1 - \frac{\sum_{k=0}^{k(d)} C_{d,k}}{(1 + \alpha_3)^d} \right). \end{aligned}$$

Now put

$$\alpha(d) := \frac{\sum_{k=0}^{k(d)} C_{d,k}}{(1 + \alpha_3)^d}.$$

It is shown in [28, p. 185] that $\alpha(d)$ tends to zero when $d \rightarrow \infty$. Hence, for any positive δ we can find an integer $d(\delta)$ such that for all $d \geq d(\delta)$ we have

$$1 \geq \frac{e_q(\lfloor C^d \rfloor, d)}{e_q(0, d)} \geq (1 - b)(1 - \delta).$$

Since b and δ can be arbitrarily close to zero, it follows that

$$\lim_{d \rightarrow \infty} \frac{e_q(\lfloor C^d \rfloor, d)}{e_q(0, d)} = 1,$$

and this holds true for all $C \in (1, 2^{\alpha_3/(1+\alpha_3)})$.

Now we proceed like in [28, p. 186]. Take $\varepsilon \in (0, 1)$. For large d we have

$$\varepsilon e_q(0, d) < e_q(\lfloor C^d \rfloor, d) = e_q(0, d)(1 + o(1))$$

and therefore $N_q^{\text{int}}(\varepsilon, d) \geq C^d(1 + o(1))$. Since $C > 1$ this means the curse of dimensionality. \square

4 Some remarks

Some remarks are in order.

Possible values for C in Theorem 3. According to the proof of Theorem 3 we have

$$\alpha_3 = \frac{\alpha_2}{\alpha_1} = \frac{2I_{1,2,(0)}(a^*, c)}{\frac{p}{p+1}(1 - (1 - a^*)^{p+1}) - I_{1,2,(0)}(a^*, c)},$$

where c is chosen according to (10), which implies

$$I_{1,2,(0)}(a^*, c) = I_{1,2,(1)}(a^*) = \frac{p}{p+1}(1 - a^*)^{p+1}.$$

Hence,

$$\alpha_3 = \frac{2(1 - a^*)^{p+1}}{1 - 2(1 - a^*)^{p+1}}$$

and further

$$\frac{\alpha_3}{1 + \alpha_3} = 2(1 - a^*)^{p+1} = \frac{2}{10^{q(p+1)}} \quad \text{and} \quad 2^{\alpha_3/(1+\alpha_3)} = 4^{10^{-q(p+1)}}.$$

We compute some values for $2^{\alpha_3/(1+\alpha_3)}$:

q	$2^{\alpha_3/(1+\alpha_3)}$
2	1.01396...
4	1.000003482...
6	1.00000000051675...

Remark 1. Our choice for a^* was very arbitrary. One can try to find the smallest possible a^* such that all conditions are satisfied, i.e., $I(\alpha, 1 - \alpha, 0) \geq$ for all odd α in $\{1, \dots, q - 1\}$ and $I_{1,2,(0)} = I_{1,2,(1)}$. This way one can improve the value of $2^{\alpha_3/(1+\alpha_3)}$. For example for $q = 4$ we also found $a^* = 0.930338256\dots$ and $c = 0.00186068\dots$. With this choice we obtain

$$2^{\alpha_3/(1+\alpha_3)} = 1.00277\dots$$

Hence we can deduce that

$$N_4^{\text{int}}(\varepsilon, d) \geq (1.00277)^d(1 + o(1)).$$

History and related problems. During the proof of Theorem 3 we already observed that, formally, we prove a lower bound for the integration problem and a 3^d -dimensional tensor product space. We give some history of such problems, in particular (in)tractability results. The first results appeared in 1997. In [25] it is proved that some non-trivial problems are easy (tractable) and for other problems the curse of dimensionality is present. An important case is the space of trigonometric polynomials of degree one in each variable; Sloan and Woźniakowski [34] prove the curse for this space. The same space, but with different norms, was studied later in [24]. To prove the curse for this norm, a new technique was developed in [37] and further studied in [18] and [19].

As we explain in the next section, also the method of decomposable kernels (see, e.g., [28, Theorem 11.12]) can be seen as a contribution to this problem, even if the authors at that time did not present it that way. This technique works fine in the case of finite smoothness and is related to the technique of bump functions, see [20] for a discussion and new results. Bump functions do not work for analytic functions, see again [18, 19], and surprisingly they also do not work for functions of very low degree of smoothness, see [20].

Generalized L_p -discrepancy. Sometimes a more general version of L_p -discrepancy is considered. Here for points $\mathcal{P} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ and corresponding coefficients $\mathcal{A} = \{a_1, a_2, \dots, a_N\}$ the discrepancy function is

$$\overline{\Delta}_{\mathcal{P}, \mathcal{A}}(\mathbf{t}) = \sum_{k=1}^N a_k \mathbf{1}_{[0, \mathbf{t})}(\mathbf{x}_k) - t_1 t_2 \cdots t_d$$

for $\mathbf{t} = (t_1, t_2, \dots, t_d)$ in $[0, 1]^d$ and the generalized L_p -discrepancy is

$$\bar{L}_{p,N}(\mathcal{P}, \mathcal{A}) = \left(\int_{[0,1]^d} |\bar{\Delta}_{\mathcal{P},\mathcal{A}}(\mathbf{t})|^p d\mathbf{t} \right)^{1/p} \quad \text{for } p \in [1, \infty),$$

with the usual adaptations for $p = \infty$. If $a_1 = a_2 = \dots = a_N = 1/N$, then we are back to the classical definition of L_p -discrepancy from Section 1.

Now the N -th minimal generalized L_p -discrepancy in dimension d is defined as

$$\overline{\text{disc}}_p(N, d) := \min_{\mathcal{P}, \mathcal{A}} \bar{L}_p(\mathcal{P}, \mathcal{A})$$

where the minimum is extended over all N -element point sets \mathcal{P} in $[0, 1]^d$ and over all corresponding coefficients \mathcal{A} , and its inverse is

$$N_p^{\overline{\text{disc}}}(\varepsilon, d) := \min\{N \in \mathbb{N} : \overline{\text{disc}}_p(N, d) \leq \varepsilon \overline{\text{disc}}_p(0, d)\},$$

where, obviously, $\overline{\text{disc}}_p(0, d) = \text{disc}_p(0, d)$.

Turn to the integration problem in $F_{d,q}$, then the worst-case error (4) of a linear algorithm (3) is

$$e(F_{d,q}, A_{d,N}^{\text{lin}}) = \bar{L}_p(\bar{\mathcal{P}}, \mathcal{A}),$$

where $\bar{\mathcal{P}}$ is given in (5) and \mathcal{A} consists of exactly the coefficients from the given linear algorithm. Again, this is well known (see [28] or [35]).

With these preparations it follows from Theorem 3 that also the generalized L_p -discrepancy suffers for all p of the form $p = \frac{2\ell}{2\ell-1}$ with $\ell \in \mathbb{N}$ from the curse of dimensionality.

Theorem 4. *For every p of the form $p = \frac{2\ell}{2\ell-1}$ with $\ell \in \mathbb{N}$ there exists a real C_p that is strictly larger than 1, such that for all $\varepsilon \in (0, 1)$ we have*

$$N_p^{\overline{\text{disc}}}(\varepsilon, d) \geq C_p^d (1 + o(1)) \quad \text{for } d \rightarrow \infty.$$

In particular, for all these p the generalized L_p -discrepancy suffers from the curse of dimensionality.

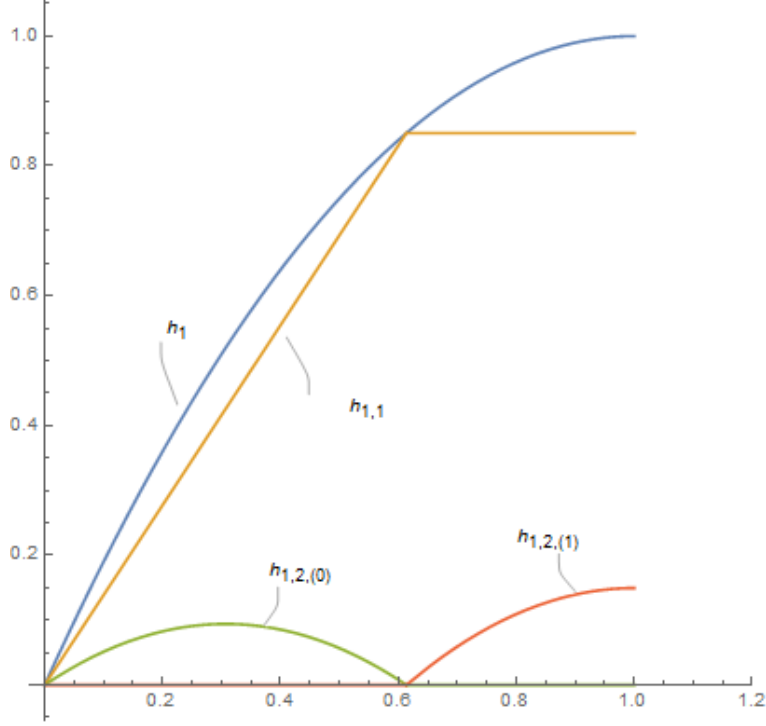


Figure 2: Plot of the functions $h_1, h_{1,1}, h_{1,2,(0)},$ and $h_{1,2,(1)}$ for $q = p = 2$ and $a = 4^{1/3}/(1 + 4^{1/3})$.

5 The case $p = q = 2$ revisited and an outlook on the general problem

As already mentioned, the case $p = q = 2$ is well established and can be tackled with the method of decomposable kernels (see [26] and also [28, Chapter 12]). This method can be mimicked with the method presented here by taking the decomposition

$$\begin{aligned} h_{1,1}(x) &= (2 - a) \min(x, a), \\ h_{1,2,(0)}(x) &= \mathbf{1}_{[0,a]}(x) x(a - x), \\ h_{1,2,(1)}(x) &= \mathbf{1}_{[a,1]}(x) (x(2 - x) - a(2 - a)) \end{aligned}$$

for some $a \in (0, 1)$ (compare this with [28, p. 192]). See Figure 2. Define the fooling function g_a like in (8), but with the present decomposition, it follows

almost immediately from orthogonality arguments that $\|g_d\|_{d,2} \leq \|h_d\|_{d,2}$. Choosing $a = 4^{1/3}/(1 + 4^{1/3})$ (compare again with [28, p. 192]) gives in addition that $\int_0^1 h_{1,2,(0)}(x) dx = \int_0^1 h_{1,2,(1)}(x) dx$ such that the whole proof of Theorem 3 goes through with the present decomposition, but now one obtains the improved quantity

$$2^{\alpha_3/(1+\alpha_3)} = 1.08332 \dots,$$

i.e.,

$$N_2^{\text{int}}(\varepsilon, d) \geq (1.08332)^d (1 + o(1)),$$

which is exactly the value obtained in [28, p. 193]. This reproves the result for $p = q = 2$, but without using the method of decomposable kernels.

From this point of view, the following decomposition for general $q \in (1, \infty)$ would be somehow obvious:

$$\begin{aligned} h_{1,1}(x) &= \frac{1 - (1 - a)^p}{a} \min(x, a), \\ h_{1,2,(0)}(x) &= \mathbf{1}_{[0,a]}(x) \left((1 - (1 - x)^p) - \frac{x}{a} (1 - (1 - a)^p) \right), \\ h_{1,2,(1)}(x) &= \mathbf{1}_{[a,1]}(x) ((1 - (1 - x)^p) - (1 - (1 - a)^p)). \end{aligned}$$

Choosing $a = 2^{\frac{p}{p+1}}/(1 + 2^{\frac{p}{p+1}})$ one obtains again that $\int_0^1 h_{1,2,(0)}(x) dx = \int_0^1 h_{1,2,(1)}(x) dx$. However, choosing the fooling function g_d like in (8) the problem is to show the estimate $\|g_d\|_{d,q} \leq \|h_d\|_{d,q}$. As soon as this could be done, one can tackle the general q case with the proof method from Theorem 3.

References

- [1] C. Aistleitner: Covering numbers, dyadic chaining and discrepancy. J. Complexity 27 (6): 531–540, 2011.
- [2] J. Beck and W.W.L. Chen: Irregularities of Distribution. Cambridge University Press, Cambridge, 1987.
- [3] D. Bilyk, M.T. Lacey, and A. Vagharshakyan: On the small ball inequality in all dimensions. J. Funct. Anal. 254 (9): 2470–2502, 2008.

- [4] W.W.L. Chen: On irregularities of distribution. *Mathematika* 27: 153–170, 1981.
- [5] W.W.L. Chen: On irregularities of distribution II. *Quart. J. Math. Oxford* 34: 257–279, 1983.
- [6] 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.* 545: 67–95, 2002.
- [7] H. Davenport: Note on irregularities of distribution. *Mathematika* 3: 131–135, 1956.
- [8] J. Dick: Discrepancy bounds for infinite-dimensional order two digital sequences over \mathbb{F}_2 . *J. Number Theory* 136: 204–232, 2014.
- [9] J. Dick, A. Hinrichs, F. Pillichshammer, and J. Prochno: Tractability properties of the discrepancy in Orlicz norms. *J. Complexity* 61, paper ref. 101468, 9 pp., 2020.
- [10] J. Dick, P. Kritzer, and F. Pillichshammer: *Lattice Rules – Numerical Integration, Approximation, and Discrepancy*. Springer Series in Computational Mathematics 58, Springer, Cham, 2022.
- [11] J. Dick and F. Pillichshammer: Optimal \mathcal{L}_2 discrepancy bounds for higher order digital sequences over the finite field \mathbb{F}_2 . *Acta Arith.* 162: 65–99, 2014.
- [12] B. Doerr: A sharp discrepancy bound for jittered sampling. *Math. Comp.* 91(336): 1871–1892, 2022.
- [13] M. Drmota and R.F. Tichy: *Sequences, Discrepancies and Applications*. Lecture Notes in Mathematics 1651, Springer Verlag, Berlin, 1997.
- [14] M. Gnewuch and N. Hebbinghaus: Discrepancy bounds for a class of negatively dependent random points including Latin hypercube samples. *Ann. Appl. Probab.* 31(4): 1944–1965, 2021.
- [15] M. Gnewuch, H. Pasing, and Ch. Weiss: A generalized Faulhaber inequality, improved bracketing covers, and applications to discrepancy. *Math. Comp.* 90 (332): 2873–2898, 2021.

- [16] S. Heinrich, E. Novak, G. Wasilkowski, and H. Woźniakowski: The inverse of the star-discrepancy depends linearly on the dimension. *Acta Arith.* 96(3): 279–302, 2001.
- [17] A. Hinrichs: Covering numbers, Vapnik-Červonenkis classes and bounds for the star-discrepancy. *J. Complexity* 20(4): 477–483, 2004.
- [18] A. Hinrichs, D. Krieg, E. Novak and J. Vybíral: Lower bounds for the error of quadrature formulas for Hilbert spaces, *J. Complexity* 65, paper ref. 101544, 20 pp., 2021.
- [19] A. Hinrichs, D. Krieg, E. Novak and J. Vybíral: Lower bounds for integration and recovery in L_2 , *J. Complexity* 72, paper ref. 101662, 15 pp., 2022.
- [20] D. Krieg and J. Vybíral: New lower bounds for the integration of periodic functions. Preprint arXiv 2302.02639.
- [21] L. Kuipers and H. Niederreiter: *Uniform Distribution of Sequences*. John Wiley, New York, 1974.
- [22] L. Markhasin: L_p - and $S_{p,q}^r B$ -discrepancy of (order 2) digital nets. *Acta Arith.* 168: 139–159, 2015.
- [23] J. Matoušek: *Geometric Discrepancy – An Illustrated Guide, Algorithms and Combinatorics*, 18, Springer-Verlag, Berlin, 1999.
- [24] E. Novak: Intractability results for positive quadrature formulas and extremal problems for trigonometric polynomials. *J. Complexity* 15: 299–316, 1999.
- [25] E. Novak, I.H. Sloan and H. Woźniakowski: Tractability of tensor product linear operators, *J. Complexity* 13: 387–418, 1997.
- [26] E. Novak and H. Woźniakowski: Intractability results for integration and discrepancy. *J. Complexity* 17(2): 388–441, 2001.
- [27] E. Novak and H. Woźniakowski: *Tractability of Multivariate Problems. Volume I: Linear Information*. EMS Tracts in Mathematics 16, Zürich, 2008.

- [28] E. Novak and H. Woźniakowski: Tractability of Multivariate Problems. Volume II: Standard Information for Functionals. EMS Tracts in Mathematics 12, Zürich, 2010.
- [29] H. Pasing and C. Weiss: Improving a constant in high-dimensional discrepancy estimates. Publ. Inst. Math. (Beograd) (N.S.) 107(121): 67–74, 2020.
- [30] F. Pillichshammer: The BMO-discrepancy suffers from the curse of dimensionality. J. Complexity 76, paper ref. 101739, 7 pp., 2023.
- [31] K.F. Roth: On irregularities of distribution. Mathematika 1: 73–79, 1954.
- [32] K.F. Roth: On irregularities of distribution. IV. Acta Arith. 37: 67–75, 1980.
- [33] M.M. Skriganov: Harmonic analysis on totally disconnected groups and irregularities of point distributions. J. Reine Angew. Math. 600: 25–49, 2006.
- [34] I.H. Sloan and H. Woźniakowski: An intractability result for multiple integration. Math. Comp. 66: 1119–1124, 1997.
- [35] I.H. Sloan and H. Woźniakowski: When are quasi-Monte Carlo algorithms efficient for high-dimensional integrals? J. Complexity 14: 1–33, 1998.
- [36] W.M. Schmidt: Irregularities of distribution. X. In: Number Theory and Algebra, pp. 311–329, Academic Press, New York, 1977.
- [37] J. Vybíral: A variant of Schur’s product theorem and its applications. Adv. Math. 368, paper ref. 107140, 9 pp., 2020.
- [38] H. Woźniakowski: Efficiency of quasi-Monte Carlo algorithms for high dimensional integrals. In: Monte Carlo and Quasi-Monte Carlo Methods 1998 (H. Niederreiter and J. Spanier, eds.), pp. 114–136, Springer Verlag, Berlin, 2000.

Author’s Address:

Erich Novak, Mathematisches Institut, FSU Jena, Ernst-Abbe-Platz 2, 07740 Jena, Germany. Email: erich.novak@uni-jena.de

Friedrich Pillichshammer, Institut für Finanzmathematik und Angewandte Zahlentheorie, JKU Linz, Altenbergerstraße 69, A-4040 Linz, Austria. Email: friedrich.pillichshammer@jku.at