# The $L_p$ -discrepancy for finite p > 1 suffers from the curse of dimensionality

Erich Novak and Friedrich Pillichshammer

#### Abstract

The  $L_p$ -discrepancy is a classical quantitative measure for the irregularity of distribution of an N-element point set in the d-dimensional unit cube. Its inverse for dimension d and error threshold  $\varepsilon \in (0, 1)$  is the number of points in  $[0, 1)^d$  that is required such that the minimal normalized  $L_p$ -discrepancy is less or equal  $\varepsilon$ . 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  $L_{\infty}$ discrepancy depends exactly linearly on d. The behavior of inverse of  $L_p$ -discrepancy for general  $p \notin \{2, \infty\}$  was an open problem since many years. Recently, the curse of dimensionality for the  $L_p$ -discrepancy was shown for an infinite sequence of values p in (1, 2], but the general result seemed to be out of reach.

In the present paper we show that the  $L_p$ -discrepancy suffers from the curse of dimensionality for all p in  $(1, \infty)$  and only the case p = 1 is still open.

This result follows from a more general result that we show for the worst-case error of positive quadrature formulas for an anchored Sobolev space of once differentiable functions in each variable whose first mixed derivative has finite  $L_q$ -norm, where q is the Hölder conjugate of p.

*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  $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N$  in the d-dimensional unitcube  $[0, 1)^d$  the local discrepancy function  $\Delta_{\mathcal{P}} : [0, 1]^d \to \mathbb{R}$  is defined as

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

for  $\boldsymbol{t} = (t_1, t_2, \dots, t_d)$  in  $[0, 1]^d$ , where  $[\boldsymbol{0}, \boldsymbol{t}) = [0, t_1) \times [0, t_2) \times \dots \times [0, t_d)$ . 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.,

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

and

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

Traditionally, the  $L_{\infty}$ -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 [1, 4, 9, 10] for detailed information. It has a close relation to numerical integration, see 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

$$\operatorname{disc}_p(N,d) := \min_{\mathcal{P}} L_{p,N}(\mathcal{P}),$$

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 [15] is the most famous result in this direction and can be seen as the initial point of discrepancy theory. 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

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

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  $L_{\infty}$ -case can be found in [2].

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 preasymptotic regime (see, e.g., [14] or [3, Section 1.7] for discussions). Nowadays, motivated from 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  $\operatorname{disc}_p(0,d)$ , which is 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  $\varepsilon$  times  $\operatorname{disc}_p(0, d)$  for a threshold  $\varepsilon \in (0, 1)$ . In other words, for  $d \in \mathbb{N}$  and  $\varepsilon \in (0, 1)$  the inverse of the N-th minimal  $L_p$ -discrepancy is defined as

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

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

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

$$\operatorname{disc}_{p}(0,d) = \begin{cases} \frac{1}{(p+1)^{d/p}} & \text{if } p \in [1,\infty), \\ 1 & \text{if } p = \infty. \end{cases}$$
(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 N-th minimal  $L_p$ -discrepancy is well understood. In the  $L_2$ -case it is known that for all  $\varepsilon \in (0, 1)$  we have

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

Here the lower bound was first shown by Woźniakowski in [16] (see also [13, 14]) and the upper bound follows from an easy averaging argument, see, e.g., [14, Sec. 9.3.2].

In the  $L_{\infty}$ -case it was shown by Heinrich, Novak, Wasilkowski and Woźniakowski in [6] that there exists an absolute positive constant C such that for every  $d \in \mathbb{N}$  and  $\varepsilon \in (0, 1)$  we have

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

The currently smallest known value of C is 6.23401... as shown in [5]. On the other hand, Hinrichs [7] proved that 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

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

So while the inverse of  $L_2$ -discrepancy grows exponentially fast with the dimension d, the inverse of the 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., [14]). Hence the situation is clear (and quite different) for  $p \in \{2, \infty\}$ . But what happens for all other  $p \notin \{2, \infty\}$ ? This question was open for many years.

Quite recently we proved in [12] that the  $L_p$ -discrepancy suffers from the curse for all values p of the form

$$p = \frac{2\ell}{2\ell - 1}$$
 with  $\ell \in \mathbb{N}$ .



Figure 1: Plot of  $C_p$  for  $p \in [1, 50]$ . Note that  $C_1 = 1$  and  $\lim_{p\to\infty} C_p = 1$ . We have  $C_2 = 1.0022..., C_3 = 1.00248..., C_4 = 1.00238...$ 

We used ideas that only work if q (the Hölder conjugate of p) is an even integer and therefore could not handle other values of p. Now, with a different approach, we solve the question for all  $p \in (1, \infty)$ .

**Theorem 1.** For every p in  $(1, \infty)$  there exists a real  $C_p$  that is strictly larger than 1, such that for all  $d \in \mathbb{N}$  and all  $\varepsilon \in (0, 1/2)$  we have

$$N_p^{\text{disc}}(\varepsilon, d) \ge C_p^d (1 - 2\varepsilon).$$

We have

$$C_p = \left(\frac{1}{2} + \frac{p+1}{p} \frac{1+2^{p/(p+1)} - 2^{1/(p+1)}}{4}\right)^{-1} > 1.$$

In particular, for all p in  $(1, \infty)$  the  $L_p$ -discrepancy suffers from the curse of dimensionality.

Figure 1 shows the graph of  $C_p$  for  $p \in [1, 50]$ . An improvement will be given in Section 3.

The result 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. We end the paper with three open problems.

#### 2 Relation to numerical integration

It is well known that the  $L_p$ -discrepancy is related to multivariate integration (see, e.g., [14, Chapter 9]). Since this relation is essential for the present approach we repeat the brief summary from [12, Section 2]. From now on let

 $p,q \ge 1$  be Hölder conjugates, i.e., 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

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

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

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

and

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

That is, consider the space

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

Now consider multivariate integration

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

We approximate the integrals  $I_d(f)$  by linear algorithms of the form

$$A_{d,N}(f) = \sum_{k=1}^{N} a_j f(\boldsymbol{x}_k), \qquad (2)$$

where  $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N$  are in  $[0, 1)^d$  and  $a_1, a_2, \ldots, a_N$  are real weights that we call integration weights. If  $a_1 = a_2 = \ldots = a_N = 1/N$ , then the linear algorithm (2) is a so-called quasi-Monte Carlo algorithm, and we denote this by  $A_{d,N}^{\text{QMC}}$ .

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

$$e(F_{d,q}, A_{d,N}) = \sup_{\substack{f \in F_{d,q} \\ \|f\|_{d,q} \le 1}} |I_d(f) - A_{d,N}(f)|.$$
(3)

For a quasi-Monte Carlo algorithm  $A_{d,N}^{\text{QMC}}$  it is well known that

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

where  $L_{p,N}(\overline{\mathcal{P}})$  is the  $L_p$ -discrepancy of the point set

$$\overline{\mathcal{P}} = \{ \mathbf{1} - \boldsymbol{x}_k : k \in \{1, 2, \dots, N\} \},$$
(4)

where  $1 - x_k$  is defined as the component-wise difference of the vector containing only ones and  $x_k$ , see, e.g., [14, Section 9.5.1] for the case p = 2.

For general linear algorithms (2) the worst-case error is the so-called generalized  $L_p$ -discrepancy

$$e(F_{d,q}, A_{d,N}) = \overline{L}_{p,N}(\overline{\mathcal{P}}, \mathcal{A}),$$

where  $\overline{\mathcal{P}}$  is like in (4) and  $\mathcal{A}$  consists of exactly the coefficients from the given linear algorithm (see [14]). Here for points  $\mathcal{P} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N\}$  and corresponding coefficients  $\mathcal{A} = \{a_1, a_2, \ldots, a_N\}$  the discrepancy function is

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

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

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

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

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 consider linear algorithms where we restrict ourselves to nonnegative weights  $a_1, \ldots, a_N$  (thus QMC-algorithms are included in our setting).

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 linear algorithms of the form (2) based on N function evaluations along points  $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N$  from  $[0, 1)^d$  and with non-negative weights  $a_1, \ldots, a_N \geq 0$ . Note that for all  $d, N \in \mathbb{N}$  we have

$$e_q(N,d) \le \operatorname{disc}_p(N,d). \tag{5}$$

The initial error is

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

We call  $f \in F_{d,q}$  a worst-case function, if  $I_d(f/||f||_{d,q}) = e_q(0,d)$ .

**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  $h_d(\boldsymbol{x}) = h_1(x_1) \cdots h_1(x_d)$  for  $\boldsymbol{x} = (x_1, \ldots, x_d) \in [0, 1]^d$ , where  $h_1(x) = 1 - (1 - x)^p$ . Furthermore, we have

$$\int_0^1 h_1(t) \, \mathrm{d}t = \frac{p}{p+1} \quad and \quad \|h_1\|_{1,q} = \frac{p}{(p+1)^{1/q}}$$

For a proof of Lemma 2 see [12].

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

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

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  $\varepsilon$ . For  $d \in \mathbb{N}$  and  $\varepsilon \in (0, 1)$  put

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

We stress that  $N_q^{\text{int}}(\varepsilon, d)$  is a kind of restricted complexity since we only allow positive quadrature formulas.

From (5), (1) and Lemma 2 it follows that for all Hölder conjugates  $q \in (1, \infty]$  and  $p \in [1, \infty)$  and for all  $d \in \mathbb{N}$  and  $\varepsilon \in (0, 1)^d$  we have

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

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

**Theorem 3.** For every q in  $(1, \infty)$  put

$$C_p = \left(\frac{1}{2} + \frac{p+1}{p} \frac{1+2^{p/(p+1)}-2^{1/(p+1)}}{4}\right)^{-1}$$

where p is the Hölder conjugate of q. Then  $C_p > 1$  and for all  $d \in \mathbb{N}$  and  $\varepsilon \in (0, 1/2)$  we have

$$N_q^{\text{int}}(\varepsilon, d) \ge C_p^d \left(1 - 2\varepsilon\right). \tag{6}$$

In particular, for all q in  $(1, \infty)$  the integration problem in  $F_{d,q}$  suffers from the curse of dimensionality for positive quadrature formulas.

*Proof.* 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 value of  $C_p$  in (6).



Figure 2: Decomposition functions  $h_{1,1}$ ,  $h_{1,2,(0)}$  and  $h_{1,2,(1)}$  for p = 3 and decomposition point a = 1/2.

For a decomposition point a in (0, 1) that will be determined in a moment define the functions

$$h_{1,1}(x) = \frac{1 - (1 - a)^p}{a} \min(x, a),$$
  

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

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

Then we have  $h_1(x) = h_{1,1}(x) + h_{1,2,(0)}(x) + h_{1,2,(1)}(x)$ . See Figure 2 for an illustration.

We choose the decomposition point a such that

$$||h_{1,1} + h_{1,2,(0)}||_{1,q} = ||h_{1,1} + h_{1,2,(1)}||_{1,q} =: \alpha.$$

We have

$$\|h_{1,1} + h_{1,2,(0)}\|_{1,q}^q = \int_0^a |p(1-x)^{p-1}|^q \, \mathrm{d}x = p^q \frac{1 - (1-a)^{p+1}}{p+1}$$

and

$$\|h_{1,1} + h_{1,2,(1)}\|_{1,q}^q = \int_a^1 |p(1-x)^{p-1}|^q \, \mathrm{d}x = p^q \frac{(1-a)^{p+1}}{p+1}.$$

Thus we have to choose

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

and then

$$\alpha = \frac{1}{2^{1/q}} \frac{p}{(p+1)^{1/q}} = \frac{\|h_1\|_{1,q}}{2^{1/q}}$$

Obviously,  $\alpha < ||h_1||_{1,q}$ . Put

$$\beta := \max\left(\int_0^1 h_{1,1}(x) + h_{1,2,(0)}(x) \,\mathrm{d}x, \int_0^1 h_{1,1}(x) + h_{1,2,(1)}(x) \,\mathrm{d}x\right)$$

Then we have  $\beta < \int_0^1 h_1(x) \, \mathrm{d}x = \frac{p}{p+1}$ .

Consider a linear algorithm  $A_{d,N}$  based in nodes  $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$  and nonnegative weights  $a_1, \ldots, a_N \geq 0$ . Then for  $i \in \{1, \ldots, N\}$  we define functions

$$P_i(\boldsymbol{x}) = \prod_{k=1}^d (h_{1,1}(x_k) + h_{1,2,(z_k)}(x_k)), \quad \boldsymbol{x} = (x_1, \dots, x_d) \in [0, 1]^d,$$

where  $z_k \in \{0, 1\}, k \in \{1, \ldots, d\}$  are chosen in a way, such that

$$P_i(\boldsymbol{x}_i) = h_d(\boldsymbol{x}_i).$$

The functions  $P_i$  are tensor products and therefore we have the simple formulas

$$\|P_i\|_{d,q} = lpha^d \quad ext{and} \quad \int_{[0,1]^d} P_i(oldsymbol{x}) \, \mathrm{d}oldsymbol{x} \leq eta^d$$

for every  $i \in \{1, \ldots, N\}$ .

In order to estimate the error of  $A_{d,N}$  we consider the two functions  $h_d$  and

$$f^* = \sum_{i=1}^N P_i.$$

For every  $i \in \{1, \ldots, N\}$  we have

$$h_d(\boldsymbol{x}_i) \le f^*(\boldsymbol{x}_i),$$

because  $P_k \ge 0$  for all k and  $P_i(\boldsymbol{x}_i) = h_d(\boldsymbol{x}_i)$ . This and the use of non-negative quadrature weights implies that

$$A_{d,N}(h_d) \le A_{d,N}(f^*).$$

Now for real y we use the notation  $(y)_+ := \max(y, 0)$ . Then we have the error estimate

$$e(F_{d,q}, A_{d,N}) \ge \frac{\left(\int_{[0,1]^d} h_d(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} - \int_{[0,1]^d} f^*(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}\right)_+}{2\max(\|h_d\|_{d,q}, \|f^*\|_{d,q})},\tag{7}$$

which is trivially true if  $\int_{[0,1]^d} h_d(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \leq \int_{[0,1]^d} f^*(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}$  and which is easily shown if  $\int_{[0,1]^d} h_d(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} > \int_{[0,1]^d} f^*(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}$ , because then

$$\left( \int_{[0,1]^d} h_d(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} - \int_{[0,1]^d} f^*(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \right)_+ \\ \leq \int_{[0,1]^d} h_d(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} - A_{d,N}(h_d) + A_{d,N}(f^*) - \int_{[0,1]^d} f^*(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \\ \leq \|h_d\|_{d,q} \, e(F_{d,q}, A_{d,N}) + \|f^*\|_{d,q} \, e(F_{d,q}, A_{d,N}) \\ \leq 2 \max(\|h_d\|_{d,q}, \|f^*\|_{d,q}) \, e(F_{d,q}, A_{d,N}).$$

From the triangle inequality we obtain

$$||f^*||_{d,q} \le N\alpha^d = N \frac{||h_d||_{d,q}}{2^{d/q}},$$

and we also have

$$\int_{[0,1]^d} f^*(oldsymbol{x}) \,\mathrm{d}oldsymbol{x} \leq Neta^d$$

Inserting into (7) yields

$$e(F_{d,q}, A_{d,N}) \ge \frac{\left(\int_{[0,1]^d} h_d(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} - N\beta^d\right)_+}{2\|h_d\|_{d,q} \max(1, N/2^{d/q})}.$$

Since the right hand side is independent of  $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$  and  $a_1, \ldots, a_N$  we obtain

$$e_q(N,d) \ge \frac{\left(\int_{[0,1]^d} h_d(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} - N\beta^d\right)_+}{2\|h_d\|_{d,q} \max(1, N/2^{d/q})}.$$
(8)

Put  $\gamma := \frac{1}{\beta} \int_0^1 h_1(x) \, dx = \frac{1}{\beta} \frac{p}{p+1}$ . Then we have  $\gamma > 1$ . Now let  $\varepsilon \in (0, 1/2)$  and assume that  $e_q(N, d) \leq \varepsilon e_q(0, d)$ .  $e_q(0, d) \|h_d\|_{d,q} = (\frac{p}{p+1})^d$  implies that This and

$$2\varepsilon \left(\frac{p}{p+1}\right)^d \max\left(1, \frac{N}{2^{d/q}}\right) \ge \left(\int_{[0,1]^d} h_d(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} - N\beta^d\right)_+.$$

If  $N \leq \min(2^{d/q}, \gamma^d)$ , then we obtain

$$\left(\frac{p}{p+1}\right)^d - N\beta^d = \int_{[0,1]^d} h_d(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} - N\beta^d$$
$$= \left(\int_{[0,1]^d} h_d(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} - N\beta^d\right)_+$$
$$\leq 2\varepsilon \left(\frac{p}{p+1}\right)^d \max\left(1, \frac{N}{2^{d/q}}\right)$$
$$= 2\varepsilon \left(\frac{p}{p+1}\right)^d.$$

Hence

$$N \ge \left(\frac{1}{\beta}\frac{p}{p+1}\right)^d (1-2\varepsilon) \ge (\min(2^{1/q},\gamma))^d (1-2\varepsilon).$$

If  $N \ge \min(2^{d/q}, \gamma^d)$ , then we trivially have

$$N \ge (\min(2^{1/q}, \gamma))^d (1 - 2\varepsilon).$$

This yields

$$N_q^{\text{int}}(\varepsilon, d) \ge (\min(2^{1/q}, \gamma))^d (1 - 2\varepsilon),$$

and we are done.

It remains to compute the values for  $C_p := \min(2^{1/q}, \gamma)$ . Obviously,  $C_p > 1$ , since  $\gamma > 1$ . For  $a = 1 - \frac{1}{2^{1/(p+1)}}$  we have

$$\int_0^1 h_{1,1}(x) + h_{1,2,(0)}(x) \, \mathrm{d}x = \int_0^a 1 - (1-x)^p \, \mathrm{d}x + \int_a^1 1 - (1-a)^p \, \mathrm{d}x$$
$$= a - \frac{1 - (1-a)^{p+1}}{p+1} + (1-a)(1 - (1-a)^p)$$
$$= \frac{1}{2}\frac{p}{p+1}$$

and

$$\int_0^1 h_{1,1}(x) + h_{1,2,(1)}(x) \, \mathrm{d}x = \int_0^a \frac{1 - (1 - a)^p}{a} x \, \mathrm{d}x + \int_a^1 1 - (1 - x)^p \, \mathrm{d}x$$
$$= \frac{a(1 - (1 - a)^p)}{2} + 1 - a - \frac{(1 - a)^{p+1}}{p+1}$$
$$= \frac{1}{2} \frac{p}{p+1} + \frac{1 + 2^{p/(p+1)} - 2^{1/(p+1)}}{4}.$$

Hence

$$\beta = \frac{1}{2} \frac{p}{p+1} + \frac{1 + 2^{p/(p+1)} - 2^{1/(p+1)}}{4}.$$

Therefore

$$\gamma = \frac{1}{\frac{1}{\frac{1}{2} + \frac{p+1}{p}\frac{1+2^{p/(p+1)} - 2^{1/(p+1)}}{4}}}$$

and

$$C_p = \min(2^{1/q}, \gamma) = \min(2^{1-1/p}, \gamma) = \gamma = \frac{1}{\frac{1}{\frac{1}{2} + \frac{p+1}{p} \frac{1+2^{p/(p+1)} - 2^{1/(p+1)}}{4}}}.$$

### **3** Remark and open problems

There are many spaces where the curse of dimensionality is present for high dimensional integration, see the recent survey paper [11]. For  $p = \infty$  the discrepancy is tractable and this property does not hold for many unweighted spaces; for a recent example see Krieg [8].

The value of  $C_p$  in Theorem 1 and 3 (see also Figure 1) can be improved with the following "spline" method:

For  $y \in [0, 1]$  define the linear splines

$$s_y(x) = \begin{cases} \frac{1 - (1 - y)^p}{y} x & \text{if } x < y, \\ 1 - (1 - y)^p & \text{if } x \ge y. \end{cases}$$

Then  $s_y \in F_{1,q}$ ,  $s_y \ge 0$  and  $s_y(y) = h_1(y)$ , the worst-case function from Lemma 2. Put

$$\alpha := \max_{y \in [0,1]} \|s_y\|_{1,q}$$
 and  $\beta := \max_{y \in [0,1]} \int_0^1 s_y(x) \, \mathrm{d}x.$ 

It is easily shown that

$$\alpha = \max_{y \in [0,1]} \frac{1 - (1-y)^p}{y^{1/p}} \quad \text{and} \quad \beta = \max_{y \in [0,1]} (1 - (1-y)^p) \left(1 - \frac{y}{2}\right).$$

Furthermore, it is elementary but tedious to show that  $\alpha < ||h_1||_{1,q}$ . On the other hand, since for every  $y \in [0,1]$  we have  $0 \leq s_y(x) < h_1(x)$ , it is clear that  $\beta < \int_0^1 h_1(x) \, dx$ . This yields that

$$\widetilde{C}_p := \min\left(\frac{\|h_1\|_{1,q}}{\alpha}, \frac{1}{\beta}\int_0^1 h_1(x)\,\mathrm{d}x\right) > 1.$$
(9)

Now we modify the approach in the proof of Theorem 3 in the following way: As before, consider a linear algorithm  $A_{d,N}$  based in nodes  $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$  in  $[0, 1]^d$ and with non-negative weights  $a_1, \ldots, a_N \geq 0$ . Let  $x_{i,j}$  be the *j*-th coordinate of the point  $\boldsymbol{x}_i, i \in \{1, \ldots, N\}$  and  $j \in \{1, \ldots, d\}$ . For  $i \in \{1, \ldots, N\}$  we define functions

$$P_i(\boldsymbol{x}) := s_{x_{i,1}}(x_1) s_{x_{i,2}}(x_2) \cdots s_{x_{i,d}}(x_d), \quad \boldsymbol{x} = (x_1, \dots, x_d) \in [0, 1]^d.$$

Consider the two functions  $h_d$  and  $f^* := \sum_{i=1}^N P_i$ . Since  $A_{d,N}$  uses only non-negative weights we have

$$A_{d,N}(f^*) = \sum_{i=1}^{N} a_i \sum_{j=1}^{N} P_j(\boldsymbol{x}_i) \ge \sum_{i=1}^{N} a_i P_i(\boldsymbol{x}_i)$$
$$= \sum_{i=1}^{N} a_i s_{x_{i,1}}(x_{i,1}) s_{x_{i,2}}(x_{i,2}) \cdots s_{x_{i,d}}(x_{i,d}) = \sum_{i=1}^{N} a_i h_d(\boldsymbol{x}_i) = A_{d,N}(h_d).$$

In the same way as in the proof of Theorem 3 we obtain

$$e(F_{d,q}, A_{d,N}) \ge \frac{\left(\int_{[0,1]^d} h_d(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} - \int_{[0,1]^d} f^*(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}\right)_+}{2 \max(\|h_d\|_{d,q}, \|f^*\|_{d,q})}$$
$$\ge \frac{\left(\int_{[0,1]^d} h_d(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} - N\beta^d\right)_+}{2 \max(\|h_d\|_{d,q}, N\alpha^d)},$$

where we used that  $||f^*||_{d,q} \leq N\alpha^d$  and  $\int_{[0,1]^d} f^*(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \leq N\beta^d$ . This yields

$$e_q(N,d) \ge \frac{\left(\int_{[0,1]^d} h_d(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} - N\beta^d\right)_+}{2\max(\|h_d\|_{d,q}, N\alpha^d)}.$$

From here it follows in the same way as in the proof of Theorem 3 that

$$N_q^{\text{int}}(\varepsilon, d) \ge \widetilde{C}_p^d (1 - 2\varepsilon) \text{ where } \widetilde{C}_p > 1 \text{ is given in (9).}$$

This re-proves Theorem 3 (and Theorem 1). The advantage of  $C_p$  in Theorem 1 and 3 is that it is stated explicitly for any  $p \in (1, \infty)$ . The value of  $\widetilde{C}_p$  can be computed numerically for every  $p \in (1, \infty)$ . Experiments show a strong improvement of  $\widetilde{C}_p$  over  $C_p$ . See the following table and Figure 3:

p	2	3	4	5
$C_p$	1.0022	1.00248	1.00238	1.0022
$\widetilde{C}_p$	1.06066	1.07231	1.07276	1.07005
p	10	20	30	100
$C_p$	1.00148	1.00086	1.0006	1.00019
$\widetilde{\alpha}$	1 05005	1.025	1 09697	1 0101

Figure 3 shows the strong improvement of  $\widetilde{C}_p$  over  $C_p$ . The blue line is the graph from Figure 1.

We end the paper with three open problems.

- 1. In order to estimate the error of quadrature formulas, we only considered two functions  $h_d$  and  $f^* = \sum_{i=1}^{N} P_i$ . The reason is that we have an exact formula for the norm of  $P_i$ , while the norms of "better" fooling functions are difficult to estimate. Our first Open Problem is to improve the lower bounds by finding bigger values for the constant  $C_p$  in the main result and  $\tilde{C}_p$ .
- 2. The proof in [12] only works for even q. However, for even q, it is more general since we prove the lower bound for all quadrature formulas, the



Figure 3: Plot of  $\widetilde{C}_p$  compared to  $C_p$  for  $p \in [1, 50]$ . Note that  $\widetilde{C}_1 = C_1 = 1$ .

weights  $a_i$  do not have to be positive. Hence we ask, this is Open Problem 2, whether the curse also holds for all  $p \in (1, \infty)$  for arbitrary quadrature formulas. We guess that the answer is yes, but our attempts to prove it failed.

3. We already mentioned that the problem is still open for p = 1. Our technique does not work in this case and we even do not guess an answer to this third Open Problem.

## References

- J. Beck and W.W.L. Chen: Irregularities of Distribution. Cambridge University Press, Cambridge, 1987.
- [2] D. Bilyk, M.T. Lacey, and A. Vagharshakyan: On the small ball inequality in all dimensions. J. Funct. Anal. 254 (9): 2470–2502, 2008.
- [3] J. Dick, P. Kritzer, and F. Pillichshammer: Lattice Rules Numerical Integration, Approximation, and Discrepancy. Springer Series in Computational Mathematics 58, Springer, Cham, 2022.
- [4] M. Drmota and R.F. Tichy: Sequences, Discrepancies and Applications. Lecture Notes in Mathematics 1651, Springer Verlag, Berlin, 1997.
- [5] 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.

- [6] 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.
- [7] A. Hinrichs: Covering numbers, Vapnik-Červonenkis classes and bounds for the star-discrepancy. J. Complexity 20(4): 477–483, 2004.
- [8] D. Krieg: Tractability of sampling recovery on unweighted function classes, arxiv 2304.14169.
- [9] L. Kuipers and H. Niederreiter: Uniform Distribution of Sequences. John Wiley, New York, 1974.
- [10] J. Matoušek: Geometric Discrepancy An Illustrated Guide, Algorithms and Combinatorics, 18, Springer-Verlag, Berlin, 1999.
- [11] E. Novak: Optimal algorithms for numerical integration: recent results and open problems. To appear in: A. Hinrichs, P. Kritzer, F. Pillichshammer (eds.). Monte Carlo and Quasi-Monte Carlo Methods 2022. Springer Verlag, 2024.
- [12] E. Novak and F. Pillichshammer: The curse of dimensionality for the  $L_{p}$ -discrepancy with finite p. J. Complexity 79: paper ref. 101769, 19 pp., 2023.
- [13] E. Novak and H. Woźniakowski: Intractability results for integration and discrepancy. J. Complexity 17(2): 388–441, 2001.
- [14] E. Novak and H. Woźniakowski: Tractability of Multivariate Problems. Volume II: Standard Information for Functionals. EMS Tracts in Mathematics 12, Zürich, 2010.
- [15] K.F. Roth: On irregularities of distribution. Mathematika 1: 73–79, 1954.
- [16] 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.), Springer Verlag, Berlin, 1999.

#### 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