

The asymptotic behavior of the average L^p –discrepancies and a randomized discrepancy

Stefan Steinerberger*

Department of Financial Mathematics, University of Linz
Altenbergstraße 69, A-4040 Linz, Austria

`stefan.steinerberger@gmail.com`

Submitted: June 7, 2010; Accepted: Jul 30, 2010; Published: Aug 9, 2010

Mathematics Subject Classification: 11K06, 11K38, 60D05

Keywords: discrepancy, average L^p discrepancy

Abstract

This paper gives the limit of the average L^p –star and the average L^p –extreme discrepancy for $[0, 1]^d$ and $0 < p < \infty$. This complements earlier results by Heinrich, Novak, Wasilkowski & Woźniakowski, Hinrichs & Novak and Gnewuch and proves that the hitherto best known upper bounds are optimal up to constants. We furthermore introduce a new discrepancy $D_N^{\mathbb{P}}$ by taking a probabilistic approach towards the extreme discrepancy D_N . We show that it can be interpreted as a centralized L^1 –discrepancy $D_N^{(1)}$, provide upper and lower bounds and prove a limit theorem.

1 Introduction.

This paper discusses two relatively separate problems in discrepancy theory, one being well-known and one being introduced. The reason for doing so is that our solution for the former was actually inspired by our investigating the average case for the latter. The paper is structured as follows: We introduce the L^p –discrepancies, known results and a motivation behind a probabilistic approach towards discrepancy in this section, give our results in the second section and provide proofs in the last part of the paper.

*The author is supported by the Austrian Science Foundation (FWF), Project S9609, part of the Austrian National Research Network “Analytic Combinatorics and Probabilistic Number Theory”.

L^p -discrepancies. In a seminal paper Heinrich, Novak, Wasilkowski and Woźniakowski [5] used probabilistic methods to estimate the inverse of the star-discrepancy, which is of great interest for Quasi Monte Carlo methods. Their approach relies on the notion of the average L^p -star discrepancy. Recall that the L^p -star discrepancy for a finite point set $\mathcal{P} \subset [0, 1]^d$ is defined as

$$D_N^{(p)*}(\mathcal{P}) = \left(\int_{[0,1]^d} \left| \frac{\#\{x_i \in \mathcal{P} : x_i \in [\mathbf{0}, \mathbf{x}]\}}{\#\mathcal{P}} - \lambda([\mathbf{0}, \mathbf{x}]) \right|^p d\mathbf{x} \right)^{1/p},$$

where

$$[\mathbf{0}, \mathbf{x}] := \{\mathbf{y} \in [0, 1]^d : 0 \leq y_1 \leq x_1 \wedge \dots \wedge 0 \leq y_d \leq x_d\},$$

$N = \#\mathcal{P}$ and λ is the usual Lebesgue measure. The average L^p -star discrepancy $av_p^*(N, d)$ is then defined as the expected value of the L^p -norm of the L^p -star discrepancy of N independently and uniformly distributed random variables over $[0, 1]^d$, i.e.

$$av_p^*(N, d) = \left(\int_{[0,1]^{Nd}} D_N^{(p)*}(\{\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_N\})^p d\mathbf{t} \right)^{1/p},$$

where $\mathbf{t} = (\mathbf{t}_1, \dots, \mathbf{t}_N)$ and $\mathbf{t}_i \in [0, 1]^d$. This averaging measure tells us something about the behaviour of this discrepancy measure as well as about the behaviour of random points in the unit cube and, in the words of Heinrich, Novak, Wasilkowski and Woźniakowski [5], “we believe that such an analysis is of interest *per se*”. Their original bound holds for even integers p and states

$$av_p^*(N, d) \leq 3^{2/3} 2^{5/2+d/p} p(p+2)^{-d/p} \frac{1}{\sqrt{N}}.$$

The derivation is rather complicated and depends on Stirling numbers of the first and second kind. This bound was then improved by Hinrichs and Novak [6] (again for even p). Their calculation, however, contained an error, which was later corrected by Gnewuch [4] and the result amounts to

$$av_p^*(N, d) \leq 2^{1/2+d/p} p^{1/2} (p+2)^{-d/p} \frac{1}{\sqrt{N}}.$$

Apparently, if one can consider the star-discrepancy, one can as well consider the discrepancy, thus giving rise to the L^p -extreme discrepancy. For the definition of the L^p -extreme discrepancy, we require

$$\Omega_d = \{(\mathbf{x}, \mathbf{y}) \in [0, 1]^d \otimes [0, 1]^d : x_1 \leq y_1 \wedge \dots \wedge x_d \leq y_d\},$$

and μ as the constant multiple of the Lebesgue measure, which turns (Ω_d, μ) into a probability space, i.e.

$$\mu = 2^d \lambda_{2d},$$

where λ_k is the k -dimensional Lebesgue measure. The L^p -extreme discrepancy for a point set $\mathcal{P} \subset [0, 1]^d$ is then defined as

$$D_N^{(p)}(\mathcal{P}) = \left(\int_{\Omega_d} \left| \frac{\#\{x_i \in \mathcal{P} : x_i \in [\mathbf{x}, \mathbf{y}]\}}{\#\mathcal{P}} - \lambda([\mathbf{x}, \mathbf{y}]) \right|^p d\mu \right)^{1/p}$$

and the average L^p -extreme discrepancy $av_p(N, d)$ is defined analogous to $av_p^*(N, d)$. The problem of finding bounds for this expression was tackled by Gnewuch.

Theorem (Gnewuch, [4]). *Let p be an even integer. If $p \geq 4d$, then*

$$av_p(N, d) \leq 2^{1/2+3d/p} p^{1/2} (p+2)^{-d/p} (p+4)^{-d/p} \frac{1}{\sqrt{N}}.$$

If $p < 4d$, then we have the estimate

$$av_p(N, d) \leq 2^{5/4} 3^{1/4-d} \frac{1}{\sqrt{N}}.$$

We study the general case $[0, 1]^d$ and $p > 0$ any real number: Our contribution is to find precise expressions for

$$\lim_{N \rightarrow \infty} av_p^*(N, d) \quad \text{and} \quad \lim_{N \rightarrow \infty} av_p(N, d).$$

Our results have four interesting aspects. First of all, they clearly constitute interesting results concerning L^p -discrepancies and are natural analogues to other well known results such as the law of iterated logarithm for the extreme discrepancy D_N . Secondly, they do imply all previous results for N large enough—it should be noted, however, that in applications definite bounds for fixed N are needed. However, our strategy for proving our limit theorems is quite flexible and we will sketch two possible ways to indeed get definite upper bounds further below. Thirdly, the precise form of the limits contains certain integrals, whose special form can be used to explain why in the previous derivation of the bounds unexpected things have appeared (i.e. Stirling numbers of the first and second kind). Finally, we can use our results to show that the already known results are effectively best possible and use a combination of them to show that the average L^p -discrepancies are stable in a certain way.

Probabilistic discrepancy. Now for something completely different. Assume we are given a finite set of points $\{x_1, x_2, \dots, x_N\} \subset [0, 1]$. The discrepancy is given by

$$D_N(\{x_1, x_2, \dots, x_N\}) = \sup_{0 \leq a \leq b \leq 1} \left| \frac{\#\{x_i : a \leq x_i \leq b\}}{N} - (b-a) \right|.$$

This immediately motivates another very natural measure by not looking for the largest value but for the average value: the deviation which is assumed by a “typically random” interval. Any such idea will be intimately tied to what makes an interval “typically

random”. Taking two random points in $[0, 1]$ and looking at the interval between them is doomed to fail: the point 0.5 will be an element in half of all cases, whereas the point 0 will never be part of an interval. It is thus only natural to go to the torus and consider sets of the type

$$\mathcal{I}[a, b] := \begin{cases} (a, b) & \text{if } 0 \leq a \leq b < 1 \\ [0, b] \cup (a, 1] & \text{if } 0 \leq b < a < 1, \end{cases}$$

with the usual generalization if we are in higher dimensions.

Definition. Let $\{x_1, x_2, \dots, x_N\} \subset [0, 1]^d$ and let X_1, X_2 be two independently and uniformly distributed random variables on $[0, 1]$. We define

$$D_N^{\mathbb{P}} := \mathbb{E} \left| \frac{\#\{x_i : x_i \in \mathcal{I}[X_1, X_2]\}}{N} - \lambda(\mathcal{I}[X_1, X_2]) \right|.$$

By definition of the extreme discrepancy D_N , we always have $D_N^{\mathbb{P}} \leq D_N$. Interestingly, even showing $D_N^{\mathbb{P}} < D_N$, which is, judging from the picture, obvious, is not completely trivial. The question is evident: what is the more precise relation between these two quantities? This entire concept is, of course, naturally related to toroidal discrepancies and can be viewed as an L^1 -analogue of a concept introduced by Lev [8] in 1995. We aim to present this probabilistic discrepancy as an object worthy of study, to present several initial results, discuss a possible application and motivate new lines of thought that might lead to new insight.

2 The results.

2.1 L^p -discrepancies.

Our main result gives the correct asymptotic behavior for the average L^p discrepancies for any dimension and any $p > 0$.

Theorem 1 (Limit case, average L^p -star discrepancy.). *Let $p > 0$, $d \in \mathbb{N}$. Then*

$$\begin{aligned} \lim_{N \rightarrow \infty} av_p^*(N, d) \sqrt{N} &= \frac{\sqrt{2}}{\pi^{\frac{1}{2d}}} \Gamma\left(\frac{1+p}{2}\right)^{1/p} \left(\int_{[0,1]^d} \left(\prod_{i=1}^d x_i \left(1 - \prod_{i=1}^d x_i\right) \right)^{p/2} dx_1 \dots dx_d \right)^{1/p} \\ &= \frac{\sqrt{2}}{\pi^{\frac{1}{2d}}} \Gamma\left(\frac{1+p}{2}\right)^{1/p} \left(\sum_{i=0}^{\infty} \binom{p/2}{i} (-1)^i \left(\frac{1}{\frac{p}{2} + i + 1}\right)^d \right)^{1/p}. \end{aligned}$$

As beautiful as these expressions might be, they are of little use if we have no idea how the integral behaves. Luckily, this is not the case and we can give several bounds for

it, the proofs of which are sketched within the proof of Theorem 1. We have the universal upper bound

$$\left(\int_{[0,1]^d} \left(\prod_{i=1}^d x_i \left(1 - \prod_{i=1}^d x_i \right) \right)^{p/2} dx_1 \dots dx_d \right)^{1/p} \leq \left(\frac{2}{p+2} \right)^{d/p}.$$

Regarding lower bounds, we have a universal lower bound

$$\left(\int_{[0,1]^d} \left(\prod_{i=1}^d x_i - \prod_{i=1}^d x_i^2 \right)^{p/2} dx_1 \dots dx_d \right)^{1/p} \geq \left(\left(\frac{2}{p+2} \right)^d - (2^{p/2} - 1) \left(\frac{2}{p+4} \right)^d \right)^{1/p},$$

where the term $(2^{p/2} - 1)$ gets very large very quickly, thus making the bound only useful for small values of p . For $p \geq 2$, we have the following better lower bound

$$\left(\int_{[0,1]^d} \left(\prod_{i=1}^d x_i \left(1 - \prod_{i=1}^d x_i \right) \right)^{p/2} dx_1 \dots dx_d \right)^{1/p} \geq \left(\left(\frac{2}{p+2} \right)^d - \frac{p}{2} \left(\frac{2}{p+4} \right)^d \right)^{1/p}.$$

Our proof of Theorem 1 can be transferred to the technically more demanding but not fundamentally different case of the average L^p -extreme discrepancy as well. Recall that we defined

$$\Omega_d = \{(\mathbf{x}, \mathbf{y}) \in [0, 1]^d \otimes [0, 1]^d : x_1 \leq y_1 \wedge \dots \wedge x_d \leq y_d\}$$

and μ as the normalized Lebesgue measure on Ω_d .

Theorem 2 (Limit case, average L^p -extreme discrepancy.). *Let $p > 0$, $d \in \mathbb{N}$. Then*

$$\lim_{N \rightarrow \infty} av_p(N, d) \sqrt{N} = \frac{\sqrt{2}}{\pi^{\frac{1}{2d}}} \Gamma\left(\frac{1+p}{2}\right)^{1/p} \left(\int_{\Omega_d} \left(\prod_{i=1}^d (y_i - x_i) - \prod_{i=1}^d (y_i - x_i)^2 \right)^{p/2} d\mu \right)^{1/p}.$$

Note, that the binomial theorem implies

$$\lim_{N \rightarrow \infty} \frac{av_p(N, d) \sqrt{N}}{\frac{\sqrt{2}}{\pi^{\frac{1}{2d}}} \Gamma\left(\frac{1+p}{2}\right)^{1/p}} = \left(\sum_{i=0}^{\infty} \binom{p/2}{i} (-1)^i \left(\frac{8}{(2+2i+p)(4+2i+p)} \right)^d \right)^{1/p}.$$

Furthermore, we have again a universal upper bound

$$\left(\int_{\Omega_d} \left(\prod_{i=1}^d (y_i - x_i) \left(1 - \prod_{i=1}^d (y_i - x_i) \right) \right)^{p/2} d\mu \right)^{1/p} \leq \left(\frac{8}{(p+2)(p+4)} \right)^{d/p}$$

and a derivation of lower bounds can be done precisely in the same way as above.

Suggestions for improvement. These two results do not come with convergence estimates. Our method of proof could be used to obtain such bounds as well, if we were given (upper) bounds for the p -th central moment of the binomial distribution or, possibly, by using strong Berry-Esseen type results and suitable decompositions of the unit cube (i.e. bounds on the volume of the set \mathcal{A} from the proof). The second way seems to lead to a very technical path while the first way seems to be the more manageable one.

A short note on upper bounds. These two results allow us to estimate the quality of the already known bounds. The reader has probably noticed that if we use our universal upper bounds, we get almost precisely the same terms as the upper bounds in the results of Hinrichs and Novak [6] and Gnewuch [4], respectively. Our limit relation enables us thus to show that the previously known upper bounds are essentially best possible up to constants. We can even show a little bit more: any convergent sequence is bounded, the supremum of the sequence divided by the limit can thus serve as a measure of how well-behaved the sequence is.

Corollary 1 (Stability of the average L^2 -star discrepancy). *Let $d \in \mathbb{N}$ be arbitrary. Then*

$$\frac{\sup_{N \in \mathbb{N}} av_2^*(N, d)\sqrt{N}}{\lim_{N \rightarrow \infty} av_2^*(N, d)\sqrt{N}} \leq 2\sqrt{3}\pi^{1/4} \sim 4.611\dots$$

The implication of this corollary is the following: The limit case is already extremely typical, finitely many points behave at most a constant worse. It is clearly that by using the above results, this corollary can be extended to any other values of p as well. Clearly, a very similar result can be obtained for the average L^p -extreme discrepancy, where we would like to emphasize once more how good the previous results are. Let us compare Gnewuch's result (for even p and $p > 4d$) and a corollary of Theorem 2 (obtained by using the universal upper bound for the integral)

$$av_p(N, d)\sqrt{N} \leq \left[\sqrt{2} \cdot 8^{d/p} (p+2)^{-d/p} (p+4)^{-d/p} \right] p^{1/2}$$

$$\lim_{N \rightarrow \infty} av_p(N, d)\sqrt{N} \leq \left[\sqrt{2} \cdot 8^{d/p} (p+2)^{-d/p} (p+4)^{-d/p} \right] \Gamma\left(\frac{1+p}{2}\right)^{1/p} \pi^{\frac{1}{2d}}.$$

Furthermore,

$$\lim_{p \rightarrow \infty} \frac{\Gamma\left(\frac{1+p}{2}\right)^{1/p}}{\sqrt{p}} = \frac{1}{\sqrt{2e}},$$

i.e. the difference is indeed a matter of constants only. The reader will encounter a similar matching of terms when comparing the result of Hinrichs and Novak with Theorem 1. It would be certainly of interest to see whether upper bounds of similar quality can be proven when $p \notin 2\mathbb{N}$ - in such an attempt our result could serve as an orientation as to where the true answer lies.

2.2 Probabilistic discrepancy.

As usual, when a new piece of mathematics is defined, there are several different aspects that can be studied and one could focus on very detailed things. An example of a minor consideration would be the fact that the probabilistic discrepancy is more stable than the regular discrepancy in terms of removal of points and

$$D_{N-1}^{\mathbb{P}}(\{x_1, \dots, x_{n-1}\}) \leq D_N^{\mathbb{P}}(\{x_1, \dots, x_n\}) + \frac{1}{2N}$$

instead of the usual additive term $1/N$ for the extreme discrepancy. We are not going to undertake a detailed study but rather present two main points of interest.

Bounds for the probabilistic discrepancy. A natural question is how the probabilistic discrepancy is related to the extreme discrepancy. In a somewhat surprising fashion, our main result relies on a curious small fact concerning combinatorial aspects of Lebesgue integration (“what is the average oscillation of the graph of a bounded function?”). Recall that the essential supremum with respect to a measure μ is defined as

$$\operatorname{ess\,sup} |f(x)| = \|f\|_{L^\infty(\mu)} := \inf \{t \geq 0 : \mu(|f|^{-1}(t, \infty)) = 0\}.$$

Theorem 3. *Let (Ω, Σ, μ) be a probability space and let $f : \Omega \rightarrow \mathbb{R}$ be measurable. Then,*

$$\int_{\Omega} \int_{\Omega} |f(x) - f(y)| \, dx dy \leq \operatorname{ess\,sup} |f(x)|.$$

Note that the triangle inequality only gives the bound $\leq 2 \operatorname{ess\,sup}_{0 \leq x \leq 1} |f(x)|$, i.e. twice as large as our bound. Moreover, the function $f : [0, 1] \rightarrow [-1, 1]$ given by $f(x) = 2\chi_{[0, 0.5]} - 1$, where χ is the indicator function, shows that the inequality is sharp.

Theorem 4. *Let $\mathcal{P} = \{x_1, x_2, \dots, x_N\} \subset [0, 1]$. Then*

$$\frac{1}{8} D_N(\mathcal{P})^2 \leq D_N^{\mathbb{P}}(\mathcal{P}) \leq \inf_{0 \leq \alpha \leq 1} D_N^*(\{\mathcal{P} + \alpha\}).$$

Let us quickly illustrate this result by looking at the point set

$$\mathcal{P} = \left\{ 0, \frac{1}{N}, \frac{2}{N}, \dots, \frac{N-1}{N} \right\}$$

having extreme discrepancy $D_N(\mathcal{P}) = 1/N$. Its probabilistic discrepancy can be easily calculated to be $1/3N$, while our previous theorem tells us that

$$D_N^{\mathbb{P}}(\mathcal{P}) \leq \inf_{0 \leq \alpha \leq 1} D_N^*(\{\mathcal{P} + \alpha\}) = D_N^*\left(\left\{\mathcal{P} + \frac{1}{2N}\right\}\right) = \frac{1}{2N},$$

which is not that far off.

We also give the proof of another theorem, weaker than the previous one, because the proof is very interesting in itself and consists of many single components, whose improvements would lead the way to a better bound.

Theorem 5. Let $\mathcal{P} = \{x_1, x_2, \dots, x_N\} \subset [0, 1]$. Then

$$D_N(\mathcal{P}) \leq D_N^{\mathbb{P}}(\mathcal{P}) + \sqrt[3]{32D_N^{\mathbb{P}}(\mathcal{P})}.$$

Limit theorem. The classical result for the star discrepancy is very well known and, relying on the law of iterated logarithm, tells us that (independent of the dimension)

$$\limsup_{N \rightarrow \infty} \frac{\sqrt{2N}D_N^*}{\sqrt{\log \log N}} = 1 \quad \text{a.s.}$$

Since the entire definition of the probabilistic discrepancy rests on probabilistic principles, it would not be surprising if a similar result exists for $D_N^{\mathbb{P}}$. We will now compute a similar answer for the probabilistic discrepancy and show that the perhaps unexpectedly beautiful result suggests that the probabilistic discrepancy might indeed be worthy of study.

Theorem 6. Let $X_1, X_2, \dots, X_N, \dots$ be a sequence of independent random variables, which are uniformly distributed on $[0, 1]$. Then, almost surely,

$$\lim_{N \rightarrow \infty} \sqrt{N}D_N^{\mathbb{P}}(\{X_1, \dots, X_N\}) = \sqrt{\frac{\pi}{32}}.$$

Using the abbreviation

$$\{\mathcal{P} + \alpha\} = \{\{p + \alpha\} : p \in \mathcal{P}\},$$

where $\{\cdot\}$ denotes the fractional part, it is easily seen (and explained in the proof of Theorem 4) that

$$D_N^{\mathbb{P}}(\mathcal{P}) = \int_0^1 D_N^{(1)*}(\{\mathcal{P} + \alpha\})d\alpha.$$

The probabilistic discrepancy can hence be somehow thought of as a centralized L^1 -star discrepancy. It is noteworthy, that the relationship between L^1 -star discrepancy and probabilistic discrepancy seems to mirror the relationship between D_N^* and D_N , since both can be thought of as L^p -norms of associated functions, i.e. we have the relation

$$D_N^{\mathbb{P}}(\mathcal{P}) = \left\| D_N^{(1)*} \{\mathcal{P} + \cdot\} \right\|_{L^1([0,1])} \quad \text{and} \quad D_N = \left\| D_N^{(\infty)*} \{\mathcal{P} + \cdot\} \right\|_{L^\infty([0,1])}.$$

3 The proofs

3.1 The proof of Theorem 1

Proof. Recall that the L^p -star discrepancy $D_N^{(p)*}$ over a point set $\mathcal{P} \subset [0, 1]^d$ is defined as

$$D_N^{(p)*}(\mathcal{P}) = \left(\int_{[0,1]^d} \left| \frac{\#\{x_i \in \mathcal{P} : x_i \in [\mathbf{0}, \mathbf{x}]\}}{\#\mathcal{P}} - \lambda([\mathbf{0}, \mathbf{x}]) \right|^p d\mathbf{x} \right)^{1/p},$$

where N denotes again the cardinality of \mathcal{P} . The general approach would be now to consider the probability space (Ω_N, μ_N) consisting of N independently and uniformly distributed random variables over $[0, 1]^d$ and μ_N as the product Lebesgue measure

$$\mu_N = \underbrace{\lambda_d \times \lambda_d \times \dots \times \lambda_d}_{N \text{ times}}$$

to consider

$$(\text{av}_p^*(N, d))^p := \int_{\Omega_N} (D_N^{(p)*})^p(\mathcal{P}) d\mu_N$$

and to try to start proving bounds. We shall take another route by switching the order of integration and considering

$$\begin{aligned} (\text{av}_p^*(N, d))^p &= \int_{\Omega_N} \int_{[0,1]^d} \left| \frac{\#\{x_i \in \mathcal{P} : x_i \in [\mathbf{0}, \mathbf{x}]\}}{\#\mathcal{P}} - \lambda([\mathbf{0}, \mathbf{x}]) \right|^p d\mathbf{x} d\mu_N \\ &= \int_{[0,1]^d} \int_{\Omega_N} \left| \frac{\#\{x_i \in \mathcal{P} : x_i \in [\mathbf{0}, \mathbf{x}]\}}{\#\mathcal{P}} - \lambda([\mathbf{0}, \mathbf{x}]) \right|^p d\mu_N d\mathbf{x} \end{aligned}$$

instead. Fix any $\varepsilon > 0$ arbitrary. We shall restrict ourselves to not integrating over the entire set $[0, 1]^d$ but merely a subset $\mathcal{A} \subset [0, 1]^d$ given by

$$\mathcal{A} := \{\mathbf{x} \in [0, 1]^d : \varepsilon \leq \lambda([\mathbf{0}, \mathbf{x}]) \leq 1 - \varepsilon\}.$$

Since our integrand is nonnegative and at most 1, we especially have

$$\int_{[0,1]^d \setminus \mathcal{A}} \int_{\Omega_N} \left| \frac{\#\{x_i \in \mathcal{P} : x_i \in [\mathbf{0}, \mathbf{x}]\}}{\#\mathcal{P}} - \lambda([\mathbf{0}, \mathbf{x}]) \right|^p d\mu_N d\mathbf{x} \leq 1 - \lambda_d(\mathcal{A}).$$

Let us now keep a $\mathbf{x} \in [0, 1]^d \setminus \mathcal{A}$ fixed and only consider the expression

$$\frac{\#\{x_i \in \mathcal{P} : x_i \in [\mathbf{0}, \mathbf{x}]\}}{\#\mathcal{P}} - \lambda([\mathbf{0}, \mathbf{x}]).$$

Each single random variable either lands in $[\mathbf{0}, \mathbf{x}]$ or does not, which is just a Bernoulli trial with probability $\lambda([\mathbf{0}, \mathbf{x}])$ and thus the entire expression follows a Binomial distribution, i.e.

$$\#\{x_i \in \mathcal{P} : x_i \in [\mathbf{0}, \mathbf{x}]\} \sim \mathcal{B}(N, \lambda([\mathbf{0}, \mathbf{x}])).$$

The next step is simply the central limit theorem: as $n \rightarrow \infty$

$$\mathcal{B}(n, p) = \mathcal{N}(np, np(1-p))$$

and applying this to the above equation we get, after rescaling,

$$\frac{\sqrt{N}}{\sqrt{\lambda([\mathbf{0}, \mathbf{x}]) (1 - \lambda([\mathbf{0}, \mathbf{x}]))}} \left[\frac{\#\{x_i \in \mathcal{P} : x_i \in [\mathbf{0}, \mathbf{x}]\}}{\#\mathcal{P}} - \lambda([\mathbf{0}, \mathbf{x}]) \right] \sim \mathcal{N}(0, 1).$$

Taking the p -th power, we get

$$\left(\frac{\sqrt{N}}{\sqrt{\lambda([\mathbf{0}, \mathbf{x}]) (1 - \lambda([\mathbf{0}, \mathbf{x}]))}} \right)^p \left| \frac{\#\{x_i \in \mathcal{P} : x_i \in [\mathbf{0}, \mathbf{x}]\}}{\#\mathcal{P}} - \lambda([\mathbf{0}, \mathbf{x}]) \right|^p \sim |X|^p,$$

where X is a random variable satisfying $X \sim \mathcal{N}(0, 1)$. This then implies for $N \rightarrow \infty$

$$\begin{aligned} \int_{\Omega_N} \left| \frac{\#\{x_i \in \mathcal{P} : x_i \in [\mathbf{0}, \mathbf{x}]\}}{\#\mathcal{P}} - \lambda([\mathbf{0}, \mathbf{x}]) \right|^p d\mu_N &= \\ \left(\frac{\sqrt{\lambda([\mathbf{0}, \mathbf{x}]) (1 - \lambda([\mathbf{0}, \mathbf{x}]))}}{\sqrt{N}} \right)^p \int_{-\infty}^{\infty} |X|^p d\mathcal{N}(0, 1) &= \\ = \left(\frac{\sqrt{\lambda([\mathbf{0}, \mathbf{x}]) (1 - \lambda([\mathbf{0}, \mathbf{x}]))}}{\sqrt{N}} \right)^p \frac{2^{p/2}}{\sqrt{\pi}} \Gamma\left(\frac{1+p}{2}\right). \end{aligned}$$

This is now only a pointwise estimate (\mathbf{x} is fixed) and for truly integrating over \mathbf{x} over the entire domain $[0, 1]^d$ would require uniform convergence, which is not given: the rule of thumb is that the binomial distribution is close to the normal distribution only if the success rate of a single Bernoulli trial (here $\lambda([\mathbf{0}, \mathbf{x}])$) is not close to 0 or 1 - this can be made precise by using a version of the central limit theorem that comes with error estimates, i.e. the Berry-Esseen theorem (see, for example, [1]). As it can be easily checked, the error estimates for a Bernoulli experiment with probability close to 0 or 1 diverge, this means that we have pointwise but not uniform convergence. However, integrating merely over \mathcal{A} works fine (this follows also from the Berry-Esseen theorem) and so, as $N \rightarrow \infty$

$$\begin{aligned} \frac{2^{p/2}}{\sqrt{\pi}} \Gamma\left(\frac{1+p}{2}\right) \left(\frac{1}{\sqrt{N}}\right)^p \int_{\mathcal{A}} \sqrt{\lambda([\mathbf{0}, \mathbf{x}]) (1 - \lambda([\mathbf{0}, \mathbf{x}]))}^p d\mathbf{x} &= \\ \int_{\mathcal{A}} \int_{\Omega_N} \left| \frac{\#\{x_i \in \mathcal{P} : x_i \in [\mathbf{0}, \mathbf{x}]\}}{\#\mathcal{P}} - \lambda([\mathbf{0}, \mathbf{x}]) \right|^p d\mu_N d\mathbf{x}. \end{aligned}$$

This, however, is a well-behaved integral and nothing prevents us from letting $\varepsilon \rightarrow 0$ and thus $\mathcal{A} \rightarrow [0, 1]^d$ in Hausdorff metric and so, as $N \rightarrow \infty$,

$$\begin{aligned} (\text{av}_p(N, d))^p &= \int_{[0,1]^d} \int_{\Omega_N} \left| \frac{\#\{x_i \in \mathcal{P} : x_i \in [\mathbf{0}, \mathbf{x}]\}}{\#\mathcal{P}} - \lambda([\mathbf{0}, \mathbf{x}]) \right|^p d\mu_N d\mathbf{x} \\ &= \frac{2^{p/2}}{\sqrt{\pi}} \Gamma\left(\frac{1+p}{2}\right) \left(\frac{1}{\sqrt{N}}\right)^p \int_{[0,1]^d} \sqrt{\lambda([\mathbf{0}, \mathbf{x}]) (1 - \lambda([\mathbf{0}, \mathbf{x}]))}^p d\mathbf{x}. \end{aligned}$$

Summarizing, one could say that the proof consists of switching the order of integration, using the central limit theorem and paying attention to small problem areas (which then, after evaluating the first integral, turn out to be no problem at all). Evaluating this last

integral precisely is probably very hard, however, an upper bound is easily obtained via $0 \leq \lambda([\mathbf{0}, \mathbf{x}]) \leq 1$ and

$$\begin{aligned} \int_{[0,1]^d} \sqrt{\lambda([\mathbf{0}, \mathbf{x}]) (1 - \lambda([\mathbf{0}, \mathbf{x}]))}^p d\mathbf{x} &\leq \int_{[0,1]^d} \sqrt{\lambda([\mathbf{0}, \mathbf{x}])}^p d\mathbf{x} = \int_{[0,1]^d} \prod_{i=1}^d x_i^{p/2} dx_1 \dots dx_d \\ &= \prod_{i=1}^d \int_0^1 x_i^{p/2} dx_i = \left(\frac{2}{2+p}\right)^d. \end{aligned}$$

For the lower bound, we distinguish between two cases: a general p and $p \geq 2$. In the first case, we can use the binomial expansion to get

$$\begin{aligned} (\lambda([\mathbf{0}, \mathbf{x}]) (1 - \lambda([\mathbf{0}, \mathbf{x}])))^{p/2} &= \lambda([\mathbf{0}, \mathbf{x}])^{p/2} \sum_{i=0}^{\infty} \binom{p/2}{i} (-1)^i \lambda([\mathbf{0}, \mathbf{x}])^i \\ &= \sum_{i=0}^{\infty} \binom{p/2}{i} (-1)^i \lambda([\mathbf{0}, \mathbf{x}])^{p/2+i} \\ &\geq \lambda([\mathbf{0}, \mathbf{x}])^{p/2} - \sum_{i=1}^{\infty} \binom{p/2}{i} \lambda([\mathbf{0}, \mathbf{x}])^{p/2+i} \\ &\geq \lambda([\mathbf{0}, \mathbf{x}])^{p/2} - \left(\sum_{i=1}^{\infty} \binom{p/2}{i}\right) \lambda([\mathbf{0}, \mathbf{x}])^{p/2+1} \\ &= \lambda([\mathbf{0}, \mathbf{x}])^{p/2} - (2^{p/2} - 1) \lambda([\mathbf{0}, \mathbf{x}])^{p/2+1}. \end{aligned}$$

Integration over the whole domain and decomposition via Fubini then gives the result. For the second case, we study the function $(x - x^2)^{p/2}$ on $[0, 1]$. Since $p \geq 2$, the mean value theorem implies

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

and thus

$$(\lambda([\mathbf{0}, \mathbf{x}]) (1 - \lambda([\mathbf{0}, \mathbf{x}])))^{p/2} \geq \lambda([\mathbf{0}, \mathbf{x}])^{p/2} - \frac{p}{2} \lambda([\mathbf{0}, \mathbf{x}])^{p/2+1},$$

which, by integrating just as above, then yields the result. \square

3.2 The proof of Theorem 2

Proof. The proof is completely analogous to the proof of Theorem 1, due to the fact that the entire structure of the previous proof is based on the fact that the actual space is irrelevant right up to the end (which makes the entire proof technique very flexible). Regarding the proof of the upper bound, we proceed as in the previous proof via

$$\int_{\Omega_d} \left(\prod_{i=1}^d (y_i - x_i) \left(1 - \prod_{i=1}^d (y_i - x_i) \right) \right)^{p/2} d\mu \leq \int_{\Omega_d} \left(\prod_{i=1}^d (y_i - x_i) \right)^{p/2} d\mu.$$

Let us now take a closer look at $\lambda_{2d}(\Omega_d)$ (i.e. at the volume of Ω_d). This volume can be computed in many ways. A close look leads us to consider the first point fixed, to consider the choices for the second point and then integrate over the first point, which gives

$$\lambda_{2d}(\Omega_d) = \int_{[0,1]^d} \prod_{i=1}^d (1 - x_i) dx_1 \dots dx_d = \int_{[0,1]^d} \prod_{i=1}^d x_i dx_1 \dots dx_d = \frac{1}{2^d}.$$

Another argument would consist of taking a point $(\mathbf{x}, \mathbf{y}) \in [0, 1]^d \times [0, 1]^d$ completely at random. The inequality $x_1 \leq y_1$ is satisfied in half of all cases and likewise for all other components. Since all components are independent, we have $\lambda_{2d}(\Omega) = 2^{-d}$. This implies $\mu = 2^d \lambda_{2d}$ and therefore

$$\int_{\Omega_d} \left(\prod_{i=1}^d (y_i - x_i) \right)^{p/2} d\mu = \left(2 \int_0^1 \int_x^1 (y - x)^{p/2} dy dx \right)^d = \left(\frac{8}{(p+2)(p+4)} \right)^d.$$

□

3.3 The proof of Theorem 3

Proof. The statement is only non-trivial when $\text{ess sup}_{0 \leq x \leq 1} |f(x)| < \infty$. We can divide the entire function by $\text{ess sup}_{0 \leq x \leq 1} |f(x)|$ and are thus given a function which maps (ignoring sets of measure 0) $f : \Omega \rightarrow [-1, 1]$. By scaling and translation, it is thus only necessary to consider the case $f : \Omega \rightarrow [0, 1]$. We divide $[0, 1]$ into n intervals via

$$\mathcal{A}_i = \left[\frac{i}{n}, \frac{i+1}{n} \right) \quad \text{for } 0 \leq i \leq n-2 \quad \text{and} \quad \mathcal{A}_{n-1} = \left[\frac{n-1}{n}, 1 \right].$$

Then, we have

$$\int_{\Omega} \int_{\Omega} |f(x) - f(y)| dx dy \leq \frac{1}{n} + \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \frac{|i-j|}{n} \mu(f^{-1}(\mathcal{A}_i)) \mu(f^{-1}(\mathcal{A}_j)).$$

Rewriting the sum as a matrix expression and $\mu(f^{-1}(\mathcal{A}_i)) + \dots + \mu(f^{-1}(\mathcal{A}_{n-1})) = 1$ gives

$$\begin{aligned} \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \frac{|i-j|}{n} \mu(f^{-1}(\mathcal{A}_i)) \mu(f^{-1}(\mathcal{A}_j)) &\leq \max \{ \mathbf{x}^T A \mathbf{x} : \|\mathbf{x}\|_{\ell^1} \leq 1 \} \\ &= \max \{ \mathbf{x}^T A \mathbf{x} : \|\mathbf{x}\|_{\ell^1} = 1 \}, \end{aligned}$$

where the matrix A is given by

$$A = \left(\frac{|i-j|}{n} \right)_{i,j=1}^n.$$

This is a typical optimization under constraints problem and the usual Lagrangian approach yields that the maximum is assumed for $\mathbf{x} = (0.5, 0, 0, \dots, 0, 0.5)$, hence

$$\max \{ \mathbf{x}^T A \mathbf{x} : \|\mathbf{x}\|_{\ell^1} = 1 \} \leq \frac{1}{2} \frac{n-1}{n}.$$

Altogether,

$$\int_{\Omega} \int_{\Omega} |f(x) - f(y)| \, dx dy \leq \frac{1}{n} + \frac{1}{2} \frac{n-1}{n} = \frac{1}{2} \frac{n+1}{n}.$$

Letting $n \rightarrow \infty$ implies the result. \square

3.4 The proof of Theorem 4

Proof. We start with the lower bound. Taking a random interval could be simulated as follows: we start by taking any random point α in $[0, 1]$ and then take a random length β from $[0, 1]$ and move this length along the starting point (and maybe jumping from the right end of the interval to the left end of the interval, if necessary). This could also be simulated by substituting the point set \mathcal{P} by $\{\mathcal{P} + \alpha\}$ and just considering the interval $[0, \beta]$, i.e. the average L^1 -star discrepancy and so

$$D_N^{\mathbb{P}} = \int_0^1 \int_0^1 \left| \frac{A([0, x], N, \{\mathcal{P} + \alpha\})}{N} - \lambda([0, x]) \right| dx d\alpha = \int_0^1 D_N^{(1)}(\{\mathcal{P} + \alpha\}) d\alpha.$$

It is easily seen, or follows directly from the general framework [9, Theorem 1] that for any point set $\mathcal{P} \subset [0, 1]$

$$D_N(\{\mathcal{P} + \alpha\}) \leq 2D_N(\mathcal{P})$$

(a proof of this statement is very similar to the usual proof for $D_N^* \leq D_N \leq 2D_N^*$). This inequality applied a second time yields

$$\frac{1}{2} D_N(\mathcal{P}) \leq D_N(\{\mathcal{P} + \alpha\}) \leq 2D_N(\mathcal{P}).$$

and in combination with [3, Theorem 1.8]

$$\frac{1}{2} D_N(\mathcal{P})^2 \leq D_N^{(1)*}(\mathcal{P}),$$

the lower bound follows via the inequality chain

$$D_N^{(1)}(\{\mathcal{P} + \alpha\}) \geq \frac{1}{2} D_N(\{\mathcal{P} + \alpha\})^2 \geq \frac{1}{8} D_N(\mathcal{P})^2.$$

Regarding the upper bound, by the symmetry of our torus-setting, it is of no importance whether we consider \mathcal{P} or the more general $\{\mathcal{P} + \alpha\}$ from the statement. Hence, it is sufficient to show the result for $\alpha = 0$. Let the discrepancy function $g : [0, 1] \rightarrow \mathbb{R}$ be defined as

$$g(x) = \frac{\#\{x_i : x_i \in [0, x]\}}{N} - x.$$

Let $a = X_1$ and $b = X_2$ be random values given by the uniform distributed random variables. If $a < b$ and $a, b \notin \mathcal{P}$, then

$$\left| \frac{\#\{x_i : x_i \in (a, b)\}}{N} - (b - a) \right| = |g(b) - g(a)|.$$

Similarly, if $b < a$ and $a, b \notin \mathcal{P}$, then

$$\begin{aligned} \left| \frac{\#\{x_i : x_i \in [0, 1] \setminus (b, a)\}}{N} - (1 - (a - b)) \right| &= \left| \frac{N - \#\{x_i : x_i \in [b, a]\}}{N} - (1 - (a - b)) \right| \\ &= \left| \frac{\#\{x_i : x_i \in [b, a]\}}{N} - (b - a) \right| \\ &= |g(a) - g(b)|. \end{aligned}$$

Thus we have, since $a, b \notin \mathcal{P}$ almost everywhere,

$$\begin{aligned} D_N^{\mathbb{P}}(\mathcal{P}) &= \int_0^1 \int_0^1 \left| \frac{\#\{x_i : x_i \in \mathcal{I}[a, b]\}}{N} - \lambda(\mathcal{I}[a, b]) \right| dadb \\ &= \int_0^1 \int_0^1 |g(a) - g(b)| dadb. \end{aligned}$$

The result then follows from Theorem 3 and $\text{ess sup}_{0 \leq x \leq 1} |g(x)| = D_N^*(\mathcal{P})$. \square

3.5 The proof of Theorem 5

Proof. We go back to our definition of $D_N^{\mathbb{P}}$ as the expected value of the random variable

$$\left| \frac{\#\{x_i : x_i \in \mathcal{I}[X_1, X_2]\}}{N} - \lambda(\mathcal{I}[X_1, X_2]) \right|,$$

where X_1, X_2 are uniformly distributed random variables. This random variable is obviously nonnegative and assumes, as largest value, D_N . We recall an inequality due to Bhatia and Davis.

Theorem (Bhatia & Davis, [2]). *Let a probability distribution f have compact support $\text{supp}(f) \subseteq [m, M]$ and let its expected value be $m \leq \mu \leq M$. Then the variance satisfies*

$$\mathbb{V}(f) \leq (\mu - m)(M - \mu).$$

This, applied to our problem, implies

$$\mathbb{V} \left| \frac{\#\{x_i : x_i \in \mathcal{I}[X_1, X_2]\}}{N} - \lambda(\mathcal{I}[X_1, X_2]) \right| \leq D_N^{\mathbb{P}}(D_N - D_N^{\mathbb{P}}).$$

If we combine this with Chebycheff's inequality, we get, for any real $k > 0$,

$$\mathbb{P} \left(\left| \frac{\#\{x_i : x_i \in \mathcal{I}[X_1, X_2]\}}{N} - \lambda(\mathcal{I}[X_1, X_2]) \right| - D_N^{\mathbb{P}} \geq k \sqrt{D_N^{\mathbb{P}}(D_N - D_N^{\mathbb{P}})} \right) \leq \frac{1}{k^2}.$$

There are two ways to see this: first, the usual interpretation that large deviations happen rarely. A second interpretation comes from the more measure-theoretic view of probability theory. Define the set $A \subset [0, 1]^2$ as

$$A = \left\{ (a, b) \in [0, 1]^2 : \left| \frac{\#\{x_i : x_i \in \mathcal{I}[a, b]\}}{N} - \lambda(\mathcal{I}[a, b]) \right| - D_N^{\mathbb{P}} > k\sqrt{D_N^{\mathbb{P}}(D_N - D_N^{\mathbb{P}})} \right\},$$

then Tchebycheff's inequality states that $\lambda_2(A) \leq k^{-2}$. We only consider the case $A \neq \emptyset$ (if A is empty, then the very last argument from the proof applies and we are done). Let us define the function $g : [0, 1]^2 \rightarrow \mathbb{R}$

$$g(a, b) = \left| \frac{\#\{x_i : x_i \in \mathcal{I}[a, b]\}}{N} - \lambda(\mathcal{I}[a, b]) \right|.$$

If this function *were* Lipschitz continuous (it is not!), the rest of the proof would be clear: Take a point $(a, b) \in [0, 1]^2$ where $g(a, b)$ assumes the maximal value $g(a, b) = D_N(\mathcal{P})$. Then, for k small enough, $(a, b) \in A$. Since A has small measure, it does not cover a large area and we can find another point $(c, d) \notin A$, which is close to the original point (a, b) . Since (c, d) is not in A , we have an upper bound for $g(c, d)$ and can use Lipschitz continuity to find an upper bound for $g(a, b)$ and thus $D_N(\mathcal{P})$. This strategy of proof is not feasible but can be slightly modified to yield a proof. A is not empty, it is hence possible to fix a $(a, b) \in A$ with $g(a, b) = D_N(\mathcal{P})$. We distinguish two cases.

Case 1. We reach the discrepancy because there are too many points on too small a space, i.e. we have

$$\frac{\#\{x_i : x_i \in \mathcal{I}[a, b]\}}{N} - \lambda(\mathcal{I}[a, b]) = D_N(\mathcal{P}).$$

In order to rectify this imbalance, we make the interval slightly larger: We make a smaller and b larger (remember that we are on the torus, making a smaller if a is already 0 means reaching values slightly smaller than 1). We do this, until we reach the point $(c, d) \in \partial A$. A simple continuity argument shows, that there are still more points in (c, d) than its length and perfect distribution suggest (otherwise we have made the interval too large too quickly), this means that

$$\frac{\#\{x_i : x_i \in \mathcal{I}[c, d]\}}{N} - \lambda(\mathcal{I}[c, d]) = g(c, d).$$

As a consequence,

$$\begin{aligned} D_N(\mathcal{P}) - g(c, d) &= \frac{\#\{x_i : x_i \in \mathcal{I}[a, b]\}}{N} - \frac{\#\{x_i : x_i \in \mathcal{I}[c, d]\}}{N} + \lambda(\mathcal{I}[c, d]) - \lambda(\mathcal{I}[a, b]) \\ &\leq \lambda(\mathcal{I}[c, d]) - \lambda(\mathcal{I}[a, b]), \end{aligned}$$

where the difference between the cardinality of the sets is positive because $\mathcal{I}[a, b] \subseteq \mathcal{I}[c, d]$.

Case 2. The second case is somewhat dual to the first: We have a space with too few points in it, i.e.

$$\lambda(\mathcal{I}[a, b]) - \frac{\#\{x_i : x_i \in \mathcal{I}[a, b]\}}{N} = D_N(\mathcal{P}).$$

We rectify this by making the interval smaller, i.e. increasing a and decreasing b . We do so, until we reach a boundary point $(c, d) \in \partial A$. This boundary point satisfies

$$\lambda(\mathcal{I}[c, d]) - \frac{\#\{x_i : x_i \in \mathcal{I}[c, d]\}}{N} = g(c, d)$$

and so, as in the first case, as a consequence,

$$\begin{aligned} D_N(\mathcal{P}) - g(c, d) &= \lambda(\mathcal{I}[a, b]) - \lambda(\mathcal{I}[c, d]) + \frac{\#\{x_i : x_i \in \mathcal{I}[c, d]\}}{N} - \frac{\#\{x_i : x_i \in \mathcal{I}[a, b]\}}{N} \\ &\leq \lambda(\mathcal{I}[a, b]) - \lambda(\mathcal{I}[c, d]). \end{aligned}$$

It follows immediately from the definition of $\mathcal{I}[a, b]$ that generally

$$|\lambda(\mathcal{I}[x, y]) - \lambda(\mathcal{I}[z, w])| \leq |x - z| + |y - w|,$$

where in the case interesting to us, we even have equality. Now, let's go back to the two cases. We have a point $(a, b) \in A$ and are looking for the nearest point $(c, d) \notin A$, where in each case only a fourth (a quadrant) of all possible directions is feasible (because in the first case a has to get smaller and b has to get larger and conversely for the second case). Since the area of A is $\lambda(A) \leq 1/k^2$, the worst case in Euclidean distance would be a quarter of a circle with the midpoint in (a, b) and radius r . Since our above inequality is a statement in the ℓ_1 -norm, we have to substitute this with an isosceles triangle where the area condition implies for the length of the two legs ℓ of the triangle that

$$\ell \leq \sqrt{2} \frac{1}{k}$$

and so, in both cases,

$$D_N(\mathcal{P}) - g(c, d) \leq \sqrt{2} \frac{1}{k}.$$

Since $(c, d) \in \partial A$, we have

$$g(c, d) \leq k \sqrt{D_N^{\mathbb{P}}(D_N - D_N^{\mathbb{P}})} + D_N^{\mathbb{P}}$$

and hence altogether

$$D_N(\mathcal{P}) \leq k \sqrt{D_N^{\mathbb{P}}(D_N - D_N^{\mathbb{P}})} + D_N^{\mathbb{P}} + \sqrt{2} \frac{1}{k}.$$

Setting

$$k = \left(\frac{2}{D_N(\mathcal{P})D_N^{\mathbb{P}}(\mathcal{P}) - D_N^{\mathbb{P}}(\mathcal{P})^2} \right)^{1/4}$$

yields the result. □

The two components of the proof, which would be most suitable for improvement are clearly the bound on the variance, which is certainly far from best possible and knowledge about the structure the set A can possibly assume, which would also be very interesting in itself.

3.6 The proof of Theorem 6

Proof. As we have already seen in the proof of Theorem 4

$$D_N^{\mathbb{P}}(\mathcal{P}) = \int_0^1 D_N^{(1)}(\{\mathcal{P} + \alpha\}) da.$$

However, independently and uniformly distributed random variables are invariant under the translation $\mathcal{P} \rightarrow \{\mathcal{P} + \alpha\}$ and thus we are just dealing with the average L^1 -star discrepancy. Theorem 1 then gives

$$\lim_{N \rightarrow \infty} D_N^{\mathbb{P}}(\mathcal{P}) \sqrt{N} = \sqrt{\frac{\pi}{32}}.$$

□

Acknowledgements. The author is grateful to Friedrich Pillichshammer for valuable discussions and indebted to Michael Gnewuch for various remarks, which greatly improved the paper.

References

- [1] T. Arak and A. Zaitsev, Uniform Limit Theorems for Sums of Independent Random Variables (Proceedings of the Steklov Institute of Mathematics), American Mathematical Society (1988).
- [2] R. Bhatia and C. Davis, A Better Bound on the Variance, American Mathematical Monthly 107 (4): 353–357 (2000).
- [3] M. Drmota and R. Tichy, *Sequences, Discrepancies and Applications*. Lecture Notes in Mathematics 1651, Springer-Verlag, Berlin, 1997.
- [4] M. Gnewuch, Bounds for the average L^p -extreme and the L^∞ -extreme discrepancy, Electronic Journal of Combinatorics, Vol. 12, Research Paper 54, 11 pages, 2005.
- [5] S. Heinrich, E. Novak, G. W. Wasilkowski, H. Woźniakowski. The inverse of the star-discrepancy depends linearly on the dimension. Acta Arithmetica XCVI.3 (2001), 279–302.
- [6] A. Hinrichs and E. Novak, New bounds for the star discrepancy, Extended abstract of a talk at the Oberwolfach seminar Discrepancy Theory and its Applications, Report No. 13/2004, Mathematisches Forschungsinstitut Oberwolfach.

- [7] L. Kuipers and H. Niederreiter: *Uniform Distribution of Sequences*. John Wiley, New York, 1974.
- [8] V. Lev, On two versions of L^2 -discrepancy and geometrical interpretation of diaphony. *Acta Mathematica Hungarica* 69 (1995), no. 4, 281–300.
- [9] S. Steinerberger, Uniform distribution preserving mappings and variational problems, *Uniform Distribution Theory* 4, no. 1, 117–145 (2009).