

THE INVERSE OF THE STAR-DISCREPANCY DEPENDS LINEARLY ON THE DIMENSION

STEFAN HEINRICH, ERICH NOVAK, GRZEGORZ W. WASILKOWSKI,
AND HENRYK WOŹNIAKOWSKI

ABSTRACT. We study bounds on the classical $*$ -discrepancy and on its inverse. Let $n_{\infty}^*(d, \varepsilon)$ be the inverse of the $*$ -discrepancy, i.e., the minimal number of points in dimension d with the $*$ -discrepancy at most ε . We prove that $n_{\infty}^*(d, \varepsilon)$ depends *linearly* on d and at most quadratically on ε^{-1} .

We present three upper bounds on $n_{\infty}^*(d, \varepsilon)$, all of them are based on probabilistic arguments and therefore they are non-constructive. The linear in d upper bound directly follows from deep results of the theory of empirical processes but it contains an unknown multiplicative factor. Two other upper bounds are without unknown factors but do not yield the linear (in d) upper bound. One upper bound is based on an average case analysis for the L_p -star discrepancy and our numerical results seem to indicate that it gives the best estimates for specific values of d and ε .

We also present two lower bounds on $n_{\infty}^*(d, \varepsilon)$. For lower bounds, we allow arbitrary coefficients in the discrepancy formula. We prove that $n_{\infty}^*(d, \varepsilon)$ must be of order $d \log \varepsilon^{-1}$ and, roughly, of order $d^{\lambda} \varepsilon^{-(1-\lambda)}$ for any $\lambda \in (0, 1)$.

1. INTRODUCTION

The classical L_p -star discrepancy is intimately related to the worst case error of multivariate integration for the Sobolev class of functions that are once differentiable in each variable with finite L_q -norm, $1/p + 1/q = 1$, see Drmota and Tichy (1997), and Niederreiter (1992). For $p = 2$, the L_2 -star discrepancy is also related to the average case error of multivariate integration for the class of continuous functions equipped with the Wiener sheet measure, see Woźniakowski (1991). Probably, the most commonly studied case corresponds to $p = \infty$, and the L_{∞} -star discrepancy is simply called the star discrepancy and denoted as the $*$ -discrepancy.

Let us recall that the L_p -star discrepancy of points $\mathbf{t}_1, \dots, \mathbf{t}_n \in [0, 1]^d$ is defined by

$$(1) \quad \text{disc}_p^*(\mathbf{t}_1, \dots, \mathbf{t}_n) = \left(\int_{[0,1]^d} \left| x_1 \cdots x_d - \frac{1}{n} \sum_{i=1}^n 1_{[0, \mathbf{x}]}(\mathbf{t}_i) \right|^p d\mathbf{x} \right)^{1/p},$$

Date: September 1999, this version: May 4, 2000.

The research reported in this paper has been initiated while the second and fourth authors were members of the FoCM semester at MSRI, Berkeley, in the fall of 1998. Research at MSRI was supported in part by NSF grant DMS-9701755. The second author was supported by a Heisenberg scholarship of the DFG. The third and fourth authors were partially supported by the National Science Foundation under Grants CCR-9729971 and CCR-9731858, respectively.

for $1 \leq p < \infty$, and

$$(2) \quad \text{disc}_\infty^*(\mathbf{t}_1, \dots, \mathbf{t}_n) = \sup_{\mathbf{x} \in [0,1]^d} \left| x_1 \cdots x_d - \frac{1}{n} \sum_{i=1}^n 1_{[\mathbf{0}, \mathbf{x}]}(\mathbf{t}_i) \right|$$

for $p = \infty$. Here, by $1_{[\mathbf{0}, \mathbf{x}]}$ we mean the function which is 1 on the box $\{\mathbf{y} : 0 \leq y_i < x_i, i = 1, 2, \dots, d\}$ and zero outside of this box. Let

$$\text{disc}_p^*(n, d) = \inf_{(\mathbf{t}_1, \dots, \mathbf{t}_n) \in [0,1]^{nd}} \text{disc}_p^*(\mathbf{t}_1, \dots, \mathbf{t}_n)$$

denote the minimal L_p -star discrepancy for n points in dimension d . For $n = 0$, the L_p -star discrepancy $\text{disc}_p^*(0, d)$ is the initial error of multivariate integration without sampling the function. We have $\text{disc}_p^*(0, d) = (p+1)^{-d/p}$ which goes exponentially fast to zero with d for all finite p . This may indicate that the multivariate integration problem is not properly scaled for finite p . For $p = \infty$, this problem disappears since we have $\text{disc}_\infty^*(0, d) = 1$.

The usual bounds on the L_p -star discrepancy are for a fixed dimension d and large n . It is well known that the asymptotic behavior of $\text{disc}_p^*(n, d)$ with respect to n is of order at most $n^{-1}(\log n)^{d-1}$, see once more Drmota and Tichy (1997), and Niederreiter (1992). Points which achieve such a bound are called low discrepancy points. There is a beautiful and still evolving theory how to construct such low discrepancy points. This theory is mostly due to Niederreiter and his collaborators.

For some applications such as in finance, the dimension d may be huge, see Chapter 4 in Traub and Werschulz (1998) for a thorough survey. Then the usual bounds on the L_p -star discrepancy are of no help since the function $n^{-1}(\log n)^{d-1}$ is increasing for $n \leq \exp(d-1)$. The latter number $\exp(d-1)$ is prohibitively large even for a modest d .

The problem how the L_p -star discrepancy depends on d is very important in practice. This is also a challenging theoretical problem which seems to require different proof techniques than used in the previous study of the L_p -star discrepancy.

For $p = 2$, there are a number of negative results. In particular, it is known that $\text{disc}_2^*(n, d) \leq \varepsilon \text{disc}_2^*(0, d)$ for $\varepsilon \in (0, 1)$ requires n to be exponential in d . More precisely, we must have

$$n \geq (1 - \varepsilon^2) (9/8)^d.$$

The same lower bound holds if we replace equal coefficients $a_i = 1/n$ in (1) by arbitrary positive coefficients, see Sloan and Woźniakowski (1998) and Woźniakowski (1999). A similar result holds also for all algorithms with arbitrary coefficients $a_i \in \mathbb{R}$, as recently shown in Novak and Woźniakowski (1999).¹ In particular, we have (with a little abuse of notation)

$$\lim_{d \rightarrow \infty} \frac{\text{disc}_2^*(\lfloor 1.0463^d \rfloor, d)}{\text{disc}_2^*(0, d)} = 1,$$

¹The exponential dependence on d means that multivariate integration is intractable since we cannot sample the function so many times if d is large. A possible way to break intractability is to switch to weighted Sobolev classes as was done in the papers cited above. We do not pursue this point here.

if we allow arbitrary coefficients.

The question of dependence on d for $p = \infty$ was raised by Larcher (1998) who asked whether there exists an $a > 1$ such that $\text{disc}_\infty^*(n, d)$ tends to 1 for $n = a^d$ as d goes to infinity, and also asked whether, in particular, $\text{disc}_\infty^*(2^d, d)$ goes to 1 as d goes to infinity. Based on the results for $p = 2$ one may be inclined to believe that the answer to at least one of these questions is affirmative.

It was surprising for us that this is *not* the case and that a positive result holds. We prove that $\text{disc}_\infty^*(n, d)$ depends only *polynomially* on d and n^{-1} (by this we mean that $\text{disc}_\infty^*(n, d)$ is bounded from above by a polynomial in d times a negative power of n). It will be done by showing polynomial bounds on the $*$ -discrepancy and on its inverse. The first bound has the best dependence on d and is of the form,

$$(3) \quad \text{disc}_\infty^*(n, d) \leq c d^{1/2} n^{-1/2} \quad \forall n, d = 1, 2, \dots,$$

with an unknown multiplicative factor c .

We shall see later that this dependence on d cannot be improved. We do not know whether the dependence on n in (3) is sharp. As already mentioned, the asymptotic behavior of the $*$ -discrepancy on n is much better but it does not necessarily mean that the uniform bound which is valid for all d and n cannot be of order $n^{-1/2}$. This problem is open and seems very difficult.

The proof of (3) follows directly from deep results of the theory of empirical processes. In particular, we use a result of Talagrand (1994) combined with a result of Haussler (1995), and the Vapnik-Červonenkis dimension of the family of cubes $[0, \mathbf{x}]$, see e.g., Dudley (1984). The proof is non-constructive in the sense that it is probabilistic. It shows that if the points are drawn at random from the uniform distribution, most choices satisfy the desired bound. It would be very useful to find such points analytically.

We also present another bound that does not contain unknown constants. It is of the form

$$(4) \quad \text{disc}_\infty^*(n, d) \leq 2\sqrt{2} n^{-1/2} \left(d \log \left(\left\lceil \frac{dn^{1/2}}{2(\log 2)^{1/2}} \right\rceil + 1 \right) + \log 2 \right)^{1/2}.$$

The dependence on d and n is now slightly worse and is of order $(d \log d)^{1/2}$ and $n^{-1/2}(\log n)^{1/2}$. (By \log we mean the natural logarithm.)

The proof of (4) is based on the use of Hoeffding's inequality and is quite elementary. Again, the proof is probabilistic and shows that most samples of uniformly distributed points are good.

The presented two upper bounds on the $*$ -discrepancy can be obviously used to estimate its inverse, i.e., the minimal number n of points with $*$ -discrepancy at most ε ,

$$(5) \quad n_\infty^*(d, \varepsilon) = \min\{n : \text{disc}_\infty^*(n, d) \leq \varepsilon\}.$$

In particular, (3) yields

$$(6) \quad n_\infty^*(d, \varepsilon) = O(d\varepsilon^{-2})$$

with an unknown factor in the big O -notation, whereas (4) yields

$$n_\infty^*(d, \varepsilon) = O(d\varepsilon^{-2}(\log d + \log \varepsilon^{-1}))$$

with a known factor in the big O -notation. We also present a third bound on $n_\infty^*(d, \varepsilon)$ which is not so explicit, see Theorem 6. In this bound we allow arbitrary equal weights c , instead of $c = 1/n$. We show, see Theorem 7, that this bound implies

$$(7) \quad n_\infty^*(d, \varepsilon) \leq C_k d^2 \varepsilon^{-2-1/k} \quad \forall k = 1, 2, \dots$$

with the specific factor C_k given there.

We decided to include the third bound in the paper for two reasons. The first one is that we wanted to check the quality of the bounds on the inverse of $*$ -discrepancy for some explicit values of d and ε . We could not use the first bound since the factor c in (3) is unknown. We therefore compared the second and third bound on $n_\infty^*(d, \varepsilon)$ and found out that the third bound was better, and sometimes much better, than the second bound as is reported at the end of Subsection 2.2.

The second reason for including the third bound is to present a different proof technique. It is based on the analysis of the average behavior of the L_p -star discrepancy for an even integer p . We believe that such an analysis is of interest per se. The average is taken with respect to points which are uniformly distributed in the unit d -dimensional cube. Clearly, $\text{disc}_p^*(n, d) \leq \text{disc}_\infty^*(n, d)$. It is therefore somehow surprising that an upper bound on the $*$ -discrepancy can be obtained by the L_p -star discrepancy. Inequalities of this form were already proved by Niederreiter et al. (1990). That paper considers a more general situation, but the results, applied to our case, are not as good as ours, see Remark 1.

It is well known that the average L_2 -star discrepancy is of order $n^{-1/2}$. We prove that the same is true for all even p and provide an explicit expression for the average L_p -star discrepancy. The analysis is, however, much harder and involves Stirling numbers of the first and second kind. To get this result we need to prove an identity, which seems to be new, between Stirling numbers. Although this identity is strictly combinatorial we could not find a direct proof of it. Instead, we use ‘‘Chernov-type’’ or ‘‘exponential’’ tail inequalities from probability theory.

We also study lower bounds on the inverse of the $*$ -discrepancy. For lower bounds, we extend the definition of the $*$ -discrepancy by allowing arbitrary coefficients a_i instead of n^{-1} in (2). We present two lower bounds. The first bound says that $n_\infty^*(d, \varepsilon)$ must be of order $d \log \varepsilon^{-1}$. In particular, we show that $n_\infty^*(d, 1/64) \geq 0.18 d$. This and (6) proves that the inverse of the $*$ -discrepancy depends linearly on d . Our second lower bound improves the dependence on ε^{-1} and states that $n_\infty^*(d, \varepsilon)$ must be of order, roughly, $d^\lambda \varepsilon^{-(1-\lambda)}$ for any $\lambda \in (0, 1)$, see Theorem 9.

2. UPPER BOUNDS

In this section we derive upper bounds on the $*$ -discrepancy and its inverse. The first two bounds use results from the theory of empirical processes and, in particular, use Hoeffding’s inequality. They are presented in Subsection 2.1. The third bound is based on the analysis of average L_p -discrepancy and is presented in Subsection 2.2. These two subsections are independent of each other and the reader can study them in any order.

2.1. Bounds based on empirical process theory. We first present an easy proof of (4). The proof essentially follows the line of probabilistic estimates in discrepancy theory, see, e.g., Drmota and Tichy (1997) or Matoušek (1999), with an explicit tracing of the dimension dependence. While the technical background and proofs of the results in empirical process theory needed to prove (3) are quite complicated, the proof of (4) easily shows the tractability of the $*$ -discrepancy, while at the same time it nevertheless reflects some of the basic features of empirical process theory such as exponential inequalities, see (9), and (bracketing) entropy, see (8).

Theorem 1. *For all $n, d \in \mathbb{N}$,*

$$\text{disc}_\infty^*(n, d) \leq 2\sqrt{2} n^{-1/2} \left(d \log \left(\left\lceil \frac{d n^{1/2}}{(2 \log 2)^{1/2}} \right\rceil + 1 \right) + \log 2 \right)^{1/2}.$$

Proof: We take a positive δ and define $m = \lceil d/\delta \rceil$. Let Γ_m be the equidistant grid on $[0, 1]^d$ with mesh-size $1/m$. The cardinality of Γ_m is obviously $(m+1)^d$.

We now prove that the supremum in (2) can be replaced by the maximum over the finite set Γ_m with a possible decrease of the $*$ -discrepancy by δ . More precisely, we have for all $\mathbf{t}_1, \dots, \mathbf{t}_n \in [0, 1]^d$,

$$(8) \quad \text{disc}_\infty^*(\mathbf{t}_1, \dots, \mathbf{t}_n) \leq \max_{\mathbf{x} \in \Gamma_m} \left| x_1 \dots x_d - \frac{1}{n} \sum_{i=1}^n 1_{[\mathbf{0}, \mathbf{x}]}(\mathbf{t}_i) \right| + \delta.$$

Indeed, for any positive η there exists $\mathbf{x}^* \in [0, 1]^d$ such that

$$\text{disc}_\infty^*(\mathbf{t}_1, \dots, \mathbf{t}_n) \leq \left| x_1^* \dots x_d^* - \frac{1}{n} \sum_{i=1}^n 1_{[\mathbf{0}, \mathbf{x}^*]}(\mathbf{t}_i) \right| + \eta.$$

Choose $\mathbf{x}, \mathbf{y} \in \Gamma_m$ with

$$x_j \leq x_j^* < x_j + \frac{1}{m} =: y_j \quad \forall j = 1, \dots, d.$$

Clearly,

$$\begin{aligned} & y_1 \dots y_d - x_1 \dots x_d \\ &= \sum_{k=1}^d (x_1 \dots x_{k-1} y_k y_{k+1} \dots y_d - x_1 \dots x_{k-1} x_k y_{k+1} \dots y_d) \leq \frac{d}{m} \leq \delta. \end{aligned}$$

Hence, $-\delta + \prod_{j=1}^d y_j \leq \prod_{j=1}^d x_j^* \leq \prod_{j=1}^d y_j \leq \delta + \prod_{j=1}^d x_j$. This yields

$$\begin{aligned} & y_1 \dots y_d - \frac{1}{n} \sum_{i=1}^n 1_{[\mathbf{0}, \mathbf{y}]}(\mathbf{t}_i) - \delta \leq x_1^* \dots x_d^* - \frac{1}{n} \sum_{i=1}^n 1_{[\mathbf{0}, \mathbf{x}^*]}(\mathbf{t}_i) \\ & \leq x_1 \dots x_d - \frac{1}{n} \sum_{i=1}^n 1_{[\mathbf{0}, \mathbf{x}]}(\mathbf{t}_i) + \delta. \end{aligned}$$

Since η can be arbitrarily small, this proves (8).

Let τ_1, \dots, τ_n be independent, uniformly distributed on $[0, 1]^d$ random variables. For $\mathbf{x} \in [0, 1]^d$, let

$$\xi_{\mathbf{x}}^{(i)} = x_1 \dots x_d - 1_{[0, \mathbf{x})}(\tau_i), \quad i = 1, \dots, n.$$

Letting \mathbf{E} to denote the expectation, we have $\mathbf{E}\xi_{\mathbf{x}}^{(i)} = 0$ and $|\xi_{\mathbf{x}}^{(i)}| \leq 1$.

By Hoeffding's inequality, see e.g. Pollard (1984), p. 191, for each $\mathbf{x} \in [0, 1]^d$, we have the following estimation for the corresponding probability \mathbf{P} :

$$(9) \quad \mathbf{P}\left\{\left|n^{-1} \sum_{i=1}^n \xi_{\mathbf{x}}^{(i)}\right| \geq \delta\right\} \leq 2 \exp(-\delta^2 n/2).$$

From (8) and (9) we have

$$\begin{aligned} \mathbf{P}\{\text{disc}_{\infty}^*(\tau_1, \dots, \tau_n) \leq 2\delta\} &\geq \mathbf{P}\left\{\max_{\mathbf{x} \in \Gamma_m} \left|x_1 \dots x_d - \frac{1}{n} \sum_{i=1}^n 1_{[0, \mathbf{x})}(\tau_i)\right| \leq \delta\right\} \\ &\geq 1 - 2(m+1)^d \exp(-\delta^2 n/2). \end{aligned}$$

The latter is strictly positive, if

$$(10) \quad \log 2 + d \log(\lceil d/\delta \rceil + 1) - \delta^2 n/2 < 0.$$

This holds for $\delta > \delta_0 = \delta_0(n, d)$, where δ_0 satisfies

$$(11) \quad \delta_0^2 = 2n^{-1} (d \log(\lceil d/\delta_0 \rceil + 1) + \log 2).$$

This implies that

$$\frac{1}{\delta_0} \leq \left(\frac{n}{4 \log 2}\right)^{1/2},$$

and, inserting this back into (11), we get

$$(12) \quad \delta_0^2 \leq 2n^{-1} \left(d \log\left(\left\lceil \frac{d n^{1/2}}{2(\log 2)^{1/2}} \right\rceil + 1\right) + \log 2\right).$$

This means that for any $\delta > \delta_0$ there exist points τ_1, \dots, τ_n such that

$$(13) \quad \text{disc}_{\infty}^*(\tau_1, \dots, \tau_n) \leq 2\delta.$$

Hence, $\text{disc}_{\infty}^*(n, d) \leq 2\delta_0$ which completes the proof. \square

Note that the proof above shows that if we replace the zero on the right hand side of (10) by $-\lambda$ for some constant $\lambda > 0$, then this means that the failure probability is bounded by $\exp(-\lambda)$. This implies that relation (13) holds with probability at least $1 - \exp(-\lambda)$, provided we replace $\log 2$ by $\log 2 + \lambda$ in (11).

We now turn to the proof of (3). As we shall see, the proof will be a direct consequence of deep results from the theory of empirical processes. We need to recall notions and some fundamental results of this theory, see Dudley (1978, 1984) and van der Vaart and Wellner (1996) for details. Let (M, \mathcal{M}) be a measurable space and let P be a probability on it. Let $X_i = X_i(\omega)$, $i = 1, \dots, n$, be independent, identically distributed random variables on some probability space $(\Omega, \Sigma, \mathbf{P})$ with values in M and distribution P .

Let $\mathcal{C} \subseteq \mathcal{M}$ be a countable family of measurable subsets of M . (Countability is often assumed in order to avoid measurability questions. For more general assumptions, see e.g., van der Vaart and Wellner (1996)).

The family \mathcal{C} is called a Vapnik-Červonenkis class, if there is a nonnegative integer m such that for each set $A \subset M$ with $m + 1$ elements we have

$$|\{A \cap C \in 2^A : C \in \mathcal{C}\}| < 2^{m+1}.$$

The smallest such m is called the VC-dimension of \mathcal{C} , and is denoted by $v(\mathcal{C})$. The following theorem is a result of Talagrand (1994), Theorem 6.6, combined with a result of Haussler (1995), Corollary 1.

Theorem 2. *There is a positive number K such that for each Vapnik-Červonenkis class \mathcal{C} and for each P as above the following holds: For all $s \geq K v(\mathcal{C})^{1/2}$ and all natural n ,*

$$(14) \quad \mathbb{P} \left\{ \omega : \sup_{C \in \mathcal{C}} \left| P(C) - \frac{1}{n} \sum_{i=1}^n 1_C(X_i(\omega)) \right| \geq sn^{-1/2} \right\} \leq \frac{1}{s} \left(\frac{Ks^2}{v(\mathcal{C})} \right)^{v(\mathcal{C})} e^{-2s^2}.$$

From this theorem, we easily obtain an upper bound on the *-discrepancy.

Theorem 3. *There is a positive number c such that for all $n, d \in \mathbb{N}$, $\varepsilon > 0$,*

$$(15) \quad \text{disc}_{\infty}^*(n, d) \leq cd^{1/2} n^{-1/2} \quad \text{and consequently} \quad n_{\infty}^*(d, \varepsilon) \leq \lceil c^2 d \varepsilon^{-2} \rceil.$$

Proof: We take $M = [0, 1]^d$. Let \mathcal{M} be the class of Borel sets of M and P the Lebesgue measure. We set

$$\mathcal{C} = \{[\mathbf{0}, \mathbf{x}] : \mathbf{x} = (x_1, \dots, x_d) \in ([0, 1] \cap \mathbb{Q})^d\},$$

where \mathbb{Q} denotes the set of rational numbers. Note that the supremum in (2) does not change if we restrict ourselves to rational points of $[0, 1]^d$. It is well-known, see, e.g., Dudley (1984), p. 91, that $v(\mathcal{C}) = d$. Take $s = \lambda v(\mathcal{C})^{1/2}$ and choose a positive λ_0 in such a way that $K\lambda^2 \leq e^{2\lambda^2}$

for all $\lambda \geq \lambda_0$. Then for $\lambda > \max(K, \lambda_0, 1)$ the right hand side of (14) is smaller than one, and we obtain Theorem 3 directly from Theorem 2. \square

It is obvious that Theorem 2 yields much more than just Theorem 3. To use the language of discrepancies, for $\mathbf{t}_1, \dots, \mathbf{t}_n \in M$ define

$$\text{disc}_{\infty}^{\mathcal{C}, P}(\mathbf{t}_1, \dots, \mathbf{t}_n) = \sup_{C \in \mathcal{C}} \left| P(C) - \frac{1}{n} \sum_{i=1}^n 1_C(\mathbf{t}_i) \right|.$$

Then the direct use of Theorem 2 yields

Theorem 4. *There exists an absolute positive constant c such that for each Vapnik-Červonenkis class \mathcal{C} and each P as above the following holds: For all $n \in \mathbb{N}$ there exist $\mathbf{t}_1, \dots, \mathbf{t}_n \in M$ with*

$$(16) \quad \text{disc}_{\infty}^{\mathcal{C}, P}(\mathbf{t}_1, \dots, \mathbf{t}_n) \leq cv(\mathcal{C})^{1/2} n^{-1/2}.$$

The proof above shows that a set of n points chosen independently, with distribution P , satisfies the estimate of Theorem 3 or 4, respectively, with probability at least

$$1 - \left(K \lambda^2 e^{-2\lambda^2} \right)^{v(\mathcal{C})},$$

where different choices of $\lambda \geq \max(K, \lambda_0, 1)$ lead to different constants in (15) or (16), respectively.

Many geometric classes of sets are Vapnik-Červonenkis. For example, the class of half-spaces in \mathbb{R}^d is of VC-dimension $d+1$, so is the class of all balls in \mathbb{R}^d . The class of all d -dimensional intervals $[x, y]$ is of dimension $2d$. For all classes with VC-dimension which polynomially depends on d , Theorem 4 shows that the corresponding discrepancies are tractable. Of course, not all geometric classes are Vapnik-Červonenkis. For instance, consider the class \mathcal{C} of convex sets in $[0, 1]^d$ for $d \geq 2$. This class is *not* Vapnik-Červonenkis, as is easily checked (see, e.g., Problem 2.6.15 of van der Vaart and Wellner, 1996). In this connection note also that the minimal discrepancy of convex sets is, up to logarithmic factors, of order $n^{-2/(d+1)}$, see Schmidt (1975), Stute (1977) and Beck (1988). Hence for $d \geq 4$ no behavior of $n^{-1/2}$ like in Theorem 4 can occur. For further results on VC-dimensions we refer the reader to e.g., Dudley (1984) and van der Vaart and Wellner (1996).

2.2. Bound based on the average L_p -star discrepancy. In this subsection we prove (7) by analyzing the L_p -star discrepancy for uniformly distributed points. We consider only even p and define the average L_p -star discrepancy as

$$\text{av}_p^*(n, d) = \left(\int_{[0,1]^{nd}} \text{disc}_p^*(\mathbf{t}_1, \dots, \mathbf{t}_n)^p d\mathbf{t} \right)^{1/p}, \quad \mathbf{t} = (\mathbf{t}_1, \dots, \mathbf{t}_n).$$

We have

$$\text{av}_p^*(n, d)^p = \int_{[0,1]^{nd}} \sum_{j=0}^p \binom{p}{j} \int_{[0,1]^d} (x_1 \cdots x_d)^{p-j} \cdot \left(-\frac{1}{n} \sum_{i=1}^n \mathbf{1}_{[\mathbf{0}, \mathbf{x}]}(\mathbf{t}_i) \right)^j d\mathbf{x} d\mathbf{t}.$$

This can be written as

$$\sum_{j=0}^p \binom{p}{j} (-n)^{-j} \sum_{(u_1, \dots, u_j) \in \{1, \dots, n\}^j} \int_{[0,1]^{n(d+1)}} (x_1 \cdots x_d)^{p-j} \mathbf{1}_{[\mathbf{0}, \mathbf{x}]}(\mathbf{t}_{u_1}) \cdots \mathbf{1}_{[\mathbf{0}, \mathbf{x}]}(\mathbf{t}_{u_j}) d\mathbf{x} d\mathbf{t}.$$

We compute the inner integral with respect to \mathbf{x} to get

$$\sum_{j=0}^p \binom{p}{j} (-n)^{-j} \sum_{(u_1, \dots, u_j) \in \{1, \dots, n\}^j} \int_{[0,1]^{nd}} (p-j+1)^{-d} \prod_{m=1}^d \min_{k=1, \dots, j} (1 - (t_{u_k, m})^{p-j+1}) d\mathbf{t}.$$

To compute the integral above, note that

$$\begin{aligned}
 & \int_{[0,1]^j} \min(1 - x_1^\alpha, \dots, 1 - x_j^\alpha) dx_1 \dots dx_j \\
 &= j! \int_{0 \leq x_1 \leq \dots \leq x_j \leq 1} (1 - x_j^\alpha) dx_1 \dots dx_j \\
 &= j! \int_0^1 (1 - x_j^\alpha) \left(\int_{0 \leq x_1 \leq \dots \leq x_{j-1} \leq x_j} dx_1 \dots dx_{j-1} \right) dx_j \\
 &= j! \int_0^1 (1 - x^\alpha) \frac{x^{j-1}}{(j-1)!} dx = \frac{\alpha}{j + \alpha}.
 \end{aligned}$$

Assume now that $u_1, \dots, u_j \in \{1, \dots, n\}$ and that they are not necessarily different. Then we obtain

$$\int_{[0,1]^n} \min(1 - x_{u_1}^\alpha, \dots, 1 - x_{u_j}^\alpha) dx_1 \dots dx_n = \frac{\alpha}{k(u_1, \dots, u_j) + \alpha},$$

where $k(u_1, \dots, u_j)$ is the number of different u_j 's. Hence, we have

$$\text{av}_p^*(n, d)^p = \sum_{j=0}^p \binom{p}{j} (-n)^{-j} \sum_{(u_1, \dots, u_j) \in \{1, \dots, n\}^j} (k(u_1, \dots, u_j) + p - j + 1)^{-d}.$$

Now we need to know the number of tuples $(u_1, \dots, u_j) \in \{1, \dots, n\}^j$ such that k different elements occur. We denote this number by $\#(j, k, n)$. It is the number of mappings from $\{1, \dots, j\}$ to $\{1, \dots, n\}$ such that the range has cardinality k . Using this notation we obtain

$$(17) \quad \text{av}_p^*(n, d)^p = \sum_{j=0}^p \binom{p}{j} (-n)^{-j} \sum_{k=0}^j (k + p - j + 1)^{-d} \cdot \#(j, k, n).$$

The numbers $\#(j, k, n)$ are well known in combinatorics, see, e.g., Riordan (1958), p. 92, and can be expressed by Stirling numbers $s(k, i)$ of the first kind and Stirling numbers $S(j, k)$ of the second kind. Indeed, we have

$$(18) \quad \#(j, k, n) = k! \binom{n}{k} S(j, k) = \sum_{i=0}^k s(k, i) S(j, k) n^i = \sum_{i=0}^k \binom{n}{k} \binom{k}{i} (-1)^i (k - i)^j.$$

Here we use $\#(0, 0, n) = 1$ and $\#(j, 0, n) = 0$ for $j > 0$. This is consistent with (18) if we use the usual definitions $s(0, 0) = S(0, 0) = 1$ and $S(j, 0) = 0$ for $j > 0$. For the third expression we use, as usual, the convention $0^0 = 1$ and $0^j = 0$ for $j > 0$.

We now order the terms in (17) and obtain $\text{av}_p^*(n, d)^p$ as a polynomial in n^{-1} ,

$$(19) \quad \text{av}_p^*(n, d)^p = \sum_{r=0}^p C(r, p, d) \cdot n^{-r},$$

where

$$(20) \quad C(r, p, d) = (-1)^r \sum_{i=0}^{p-r} \binom{p}{r+i} (-1)^i \sum_{k=i}^{i+r} (p+1-r+k-i)^{-d} s(k, i) S(i+r, k).$$

Observe that $C(p, p, d)$ is zero, and therefore the sum in (19) is from $r = 0$ to $p - 1$. We will prove in a moment that also $C(r, p, d) = 0$ for $r = 0, \dots, p/2 - 1$.

The $C(r, p, d)$ can be written in the form

$$(21) \quad C(r, p, d) = \frac{(-1)^r}{(p+1-r)^d} \sum_{j=0}^r \left(\frac{p+1-r}{p+1-r+j} \right)^d \beta(r, p, j)$$

with

$$\beta(r, p, j) = (-1)^j \sum_{k=j}^{p-r+j} \binom{p}{r+k-j} (-1)^k s(k, k-j) S(k-j+r, k).$$

The following lemma about Stirling numbers is essential for our error bounds.

Lemma 1. *Assume that $p = 2m$ is an even integer. Then for $r = 0, \dots, m - 1$ and $j = 0, \dots, r$*

$$\sum_{k=j}^{p-r+j} \binom{p}{r+k-j} (-1)^k s(k, k-j) S(k-j+r, k) = 0.$$

Proof: We use a result of Kiefer (1961) who proved that for each d there is a positive $c(d)$ such that

$$\lambda\{\mathbf{t}_1, \dots, \mathbf{t}_n \in [0, 1]^{nd} : n^{1/2} \text{disc}_\infty^*(\mathbf{t}_1, \dots, \mathbf{t}_n) \geq r\} \leq c(d) \cdot \exp(-r^2), \quad \forall r > 0,$$

where λ denotes the Lebesgue measure. Let F denote the distribution function of disc_∞^* . Then $1 - F(x) \leq c(d) \exp(-x^2/n)$ and therefore

$$\begin{aligned} \text{av}_p^*(n, d)^p &\leq \int_{[0,1]^{nd}} \text{disc}_\infty^*(\mathbf{t}_1, \dots, \mathbf{t}_n)^p d\mathbf{t} = \int_0^1 x^p dF(x) = p \int_0^1 x^{p-1} (1 - F(x)) dx \\ &\leq c(d)p \int_0^1 x^{p-1} \exp(-nx^2) dx \leq c(d)pn^{-p/2} \int_0^{+\infty} y^{p-1} \exp(-y^2) dy = c_1(p, d)n^{-p/2}. \end{aligned}$$

Due to (19) this proves that

$$C(r, p, d) = 0 \quad \text{for } r = 0, 1, \dots, m - 1, \text{ and } d = 1, 2, \dots$$

Letting d to infinity in (21) we conclude that

$$\beta(r, p, j) = 0 \quad \text{for } j = 0, 1, \dots, r,$$

as claimed. \square

We are ready to present an upper bound on $\text{av}_p^*(n, d)$.

Theorem 5. *Assume that p is an even integer and $d \in \mathbb{N}$. Then*

$$(22) \quad \text{av}_p^*(n, d)^p = \sum_{r=p/2}^{p-1} C(r, p, d) \cdot n^{-r},$$

where the $C(r, p, d)$ are given by (20). Furthermore

$$|C(r, p, d)| \leq \frac{(r+1)(4p)^p}{(p+1-r)^d},$$

and

$$\text{av}_p^*(n, d) \leq 4\sqrt{2} p (1+p/2)^{-d/p} n^{-1/2} \left(\sum_{i=0}^{p/2-1} n^{-i} \left(\frac{1+p/2}{1+p/2-i} \right)^d \right)^{1/p}.$$

Proof: The formula (22) directly follows from Lemma 1. To get upper bounds on $C(r, p, d)$ we use (20) to obtain

$$|C(r, p, d)| \leq (p+1-r)^{-d} \sum_{i=0}^{p-r} \binom{p}{r+i} \sum_{k=i}^{i+r} |s(k, i) S(i+r, k)|.$$

The Stirling numbers of the first kind are defined by

$$\sum_{i=0}^k s(k, i) x^i = x(x-1)(x-2) \cdots (x-k+1).$$

As a consequence from the last equation, one can easily prove by induction that

$$\sum_{i=0}^k |s(k, i)| = k!.$$

Later we will use² only $|s(k, i)| \leq k!$. Since $|s(k+1, i)| = |s(k, i-1)| + k|s(k, i)|$, the numbers $|s(k, i)|$ are increasing in k . We also know that

$$S(n, m) = \frac{1}{m!} \sum_{k=0}^m (-1)^{m-k} \binom{m}{k} k^n$$

which leads to

$$S(n, m) \leq \frac{2^m m^n}{m!}.$$

Together we obtain

$$|C(r, p, d)| \leq (p+1-r)^{-d} \sum_{i=0}^{p-r} \sum_{k=i}^{i+r} \binom{p}{r+i} \cdot k! \cdot \frac{2^k k^{i+r}}{k!}.$$

²There are better asymptotic bounds on Stirling numbers of the first and second kind. Since we need estimates with *explicit* constants we choose explicit bounds and they are not always asymptotically sharp.

Since $2^k k^{i+r} \leq 2^{i+r} (i+r)^{i+r} \leq 2^p p^p$, we get

$$|C(r, p, d)| \leq (p+1-r)^{-d} 2^p p^p (r+1) \sum_{k=r}^p \binom{p}{k} \leq (p+1-r)^{-d} (r+1) (4p)^p,$$

as claimed. To get an estimate on $\text{av}_p^*(n, d)$, we simply notice that $|C(r, p, d)| \leq p(4p)^p / (1+p-r)^d$ and

$$\sum_{r=p/2}^{p-1} |C(r, p, d)| n^{-r} \leq \frac{p(4p)^p}{(1+p/2)^d} \frac{1}{n^{p/2}} \sum_{r=p/2}^{p-1} n^{-(r-p/2)} \left(\frac{1+p/2}{1+p-r} \right)^d.$$

Since $p^{1/p} \leq \sqrt{2}$ for even p , the estimate on $\text{av}_p^*(n, d)$ follows. \square

We are ready to prove the upper bound (7) for the inverse of $*$ -discrepancy. Our proof technique requires to consider a “scaled version” of the $*$ -discrepancy with an optimal weight. That is, for $h \in (0, 1]$ and $(\mathbf{t}_1, \dots, \mathbf{t}_n) \in [0, h]^{dn}$ we define

$$\text{disc}_\infty^h(\mathbf{t}_1, \dots, \mathbf{t}_n) = \inf_{c>0} \sup_{\mathbf{x} \in [0, h]^d} \left| x_1 \cdots x_d - c \cdot \sum_{i=1}^n 1_{[\mathbf{0}, \mathbf{x}]}(\mathbf{t}_i) \right|.$$

It is easy to check by rescaling that

$$(23) \quad \text{disc}_\infty^h(h\mathbf{t}_1, \dots, h\mathbf{t}_n) = h^d \cdot \text{disc}_\infty^1(\mathbf{t}_1, \dots, \mathbf{t}_n).$$

We derive upper bounds for the numbers

$$\text{disc}_\infty^1(n, d) = \inf_{(\mathbf{t}_1, \dots, \mathbf{t}_n) \in [0, 1]^{nd}} \text{disc}_\infty^1(\mathbf{t}_1, \dots, \mathbf{t}_n),$$

and its inverse function,

$$(24) \quad n_\infty^1(d, \varepsilon) = \min\{n : \text{disc}_\infty^1(n, d) \leq \varepsilon\}, \quad \varepsilon \in (0, 1).$$

We now show a close relation between the $n_\infty^*(d, \varepsilon)$ (where we use $c = 1/n$) with $n_\infty^1(d, \varepsilon)$ (where we use an optimal c). These quantities are related as shown below.

Lemma 2.

$$n_\infty^1(d, \varepsilon) \leq n_\infty^*(d, \varepsilon) \leq n_\infty^1(d, \varepsilon/2).$$

Proof: We only need to prove that $n_\infty^*(d, \varepsilon) \leq n := n_\infty^1(d, \varepsilon/2)$. Without loss of generality we may assume that there are optimal points $\mathbf{t}_i \in [0, 1]^d$, $i = 1, 2, \dots, n$, such that

$$\text{disc}_\infty^1(\mathbf{t}_1, \dots, \mathbf{t}_n) = \inf_{c>0} \sup_{\mathbf{x} \in [0, 1]^d} \left| x_1 \cdots x_d - c \cdot \sum_{i=1}^n 1_{[\mathbf{0}, \mathbf{x}]}(\mathbf{t}_i) \right| \leq \varepsilon/2.$$

Observe that all components of \mathbf{t}_i for $i = 1, 2, \dots, n$ must be less than 1. Indeed, if a component of some \mathbf{t}_i is one then $1_{[\mathbf{0}, \mathbf{x}]}(\mathbf{t}_i) = 0$ and we obtain the same result for at most $n-1$ points which contradicts the assumption that n is the minimal number for which the $*$ -discrepancy with an optimal weight is at most $\varepsilon/2$.

We now take $\mathbf{x} = \mathbf{1}$, and we see that for these \mathbf{t}_i the optimal c satisfies $|1 - cn| \leq \varepsilon/2$. Hence, the weight c differs from $1/n$ by at most $\varepsilon/(2n)$ and therefore, for all x ,

$$\left| c \cdot \sum_{i=1}^n 1_{[0, \mathbf{x})}(\mathbf{t}_i) - n^{-1} \cdot \sum_{i=1}^n 1_{[0, \mathbf{x})}(\mathbf{t}_i) \right| \leq |c - n^{-1}|n \leq \varepsilon/2.$$

Hence we obtain

$$\text{disc}_{\infty}^*(\mathbf{t}_1, \dots, \mathbf{t}_n) = \sup_{\mathbf{x} \in [0, 1]^d} \left| x_1 \cdots x_d - \frac{1}{n} \sum_{i=1}^n 1_{[0, \mathbf{x})}(\mathbf{t}_i) \right| \leq \varepsilon.$$

Thus, $n_{\infty}^*(d, \varepsilon) \leq n$, as claimed. \square

We now give upper bounds for $\text{disc}_{\infty}^1(n, d)$ or $n_{\infty}^1(d, \varepsilon)$ in terms of the $\text{av}_p^*(n, d)$. To prove such estimates, we assume that

$$(25) \quad \text{disc}_{\infty}^1(n, d) = \varepsilon.$$

From (23) we have

$$\text{disc}_{\infty}^h(\mathbf{t}_1, \dots, \mathbf{t}_n) \geq \varepsilon \cdot h^d$$

for arbitrary $(\mathbf{t}_1, \dots, \mathbf{t}_n) \in [0, h]^{nd}$ and $h \in (0, 1]$. In the following we assume that the points $(\mathbf{t}_1, \dots, \mathbf{t}_n)$ and h are fixed. Then there is an $\mathbf{x}^* \in [0, h]^d$ such that

$$\left| x_1^* \cdots x_d^* - \frac{1}{n} \sum_{i=1}^n 1_{[0, \mathbf{x}^*)}(\mathbf{t}_i) \right| \geq \varepsilon \cdot h^d.$$

There are two possible cases.

First case: We have

$$(26) \quad x_1^* \cdots x_d^* - \frac{1}{n} \sum_{i=1}^n 1_{[0, \mathbf{x}^*)}(\mathbf{t}_i) \geq \varepsilon \cdot h^d,$$

where $\mathbf{0} \leq \mathbf{x}^* \leq \mathbf{h}$ with $x_1^* \cdots x_d^* \geq \varepsilon h^d$, all inequalities are component-wise and by \mathbf{h} we mean the vector $\mathbf{h} = (h, h, \dots, h)$.

Then the box $[0, \mathbf{x}^*)$ does not contain enough of the sample points and this is also true for some smaller boxes. For all $\mathbf{y} \in [0, \mathbf{x}^*)$ we obtain

$$y_1 \cdots y_d - \frac{1}{n} \sum_{i=1}^n 1_{[0, \mathbf{y})}(\mathbf{t}_i) \geq \varepsilon \cdot h^d - x_1^* \cdots x_d^* + y_1 \cdots y_d$$

and therefore

$$\begin{aligned} \text{disc}_p^*(\mathbf{t}_1, \dots, \mathbf{t}_n)^p &\geq \int_{[0, \mathbf{x}^*)} (\varepsilon \cdot h^d - x_1^* \cdots x_d^* + y_1 \cdots y_d)_+^p d\mathbf{y} \\ &= \int_{[\mathbf{h} - \mathbf{x}^*, \mathbf{h})} (\varepsilon \cdot h^d - x_1^* \cdots x_d^* + (z_1 - \delta_1) \cdots (z_d - \delta_d))_+^p d\mathbf{z}, \end{aligned}$$

where $\delta_i = h - x_i^*$. Note that $-x_1^* \cdots x_d^* + (z_1 - \delta_1) \cdots (z_d - \delta_d) \geq -h^d + z_1 \cdots z_d$. Indeed, this follows from the fact that $f(\mathbf{z}) := z_1 \cdots z_d - (z_1 - \delta_1) \cdots (z_d - \delta_d)$ is increasing in each variable z_i and that $f(\mathbf{h}) = h^d - x_1^* \cdots x_d^*$. Therefore

$$\text{disc}_p^*(\mathbf{t}_1, \dots, \mathbf{t}_n)^p \geq \int_{[\mathbf{h}-\mathbf{x}^*, \mathbf{h}]} (\varepsilon \cdot h^d - h^d + y_1 \cdots y_d)_+^p d\mathbf{y}.$$

Note that $h^{d-1}x_i^* \geq x_1^* \cdots x_d^* \geq h^d \varepsilon$ implies $x_i^* \geq h\varepsilon$ and $h - x_i^* \leq h - h\varepsilon = h(1 - \varepsilon)$. Hence

$$\text{disc}_p^*(\mathbf{t}_1, \dots, \mathbf{t}_n)^p \geq \int_{[(1-\varepsilon)\mathbf{h}, \mathbf{h}]} (\varepsilon \cdot h^d - h^d + y_1 \cdots y_d)_+^p d\mathbf{y}.$$

If we further shrink the cube of integration to $[(1 - \varepsilon)^{1/d}\mathbf{h}, \mathbf{h}]$ then the integrand is nonnegative since $y_1 \cdots y_d \geq (1 - \varepsilon)h^d$. Therefore

$$(27) \quad \text{disc}_p^*(\mathbf{t}_1, \dots, \mathbf{t}_n)^p \geq \int_{[(1-\varepsilon)^{1/d}\mathbf{h}, \mathbf{h}]} (\varepsilon \cdot h^d - h^d + y_1 \cdots y_d)^p d\mathbf{y} =: A_p^d(\varepsilon).$$

Second case: We have

$$(28) \quad x_1^* \cdots x_d^* - \frac{1}{n} \sum_{i=1}^n 1_{[\mathbf{0}, \mathbf{x}^*]}(\mathbf{t}_i) \leq -\varepsilon \cdot h^d,$$

where, as before, $\mathbf{0} \leq \mathbf{x}^* \leq \mathbf{h}$. Then the box $[\mathbf{0}, \mathbf{x}^*]$ contains too many of the sample points and this is also true for some larger boxes. For all $\mathbf{y} \in [0, 1]^d$ with $\mathbf{x}^* \in [\mathbf{0}, \mathbf{y}]$ we obtain

$$y_1 \cdots y_d - \frac{1}{n} \sum_{i=1}^n 1_{[\mathbf{0}, \mathbf{y}]}(\mathbf{t}_i) \leq -\varepsilon \cdot h^d - x_1^* \cdots x_d^* + y_1 \cdots y_d$$

and therefore

$$\begin{aligned} \text{disc}_p^*(\mathbf{t}_1, \dots, \mathbf{t}_n)^p &\geq \int_{[\mathbf{x}^*, \mathbf{1}]} (\varepsilon \cdot h^d + x_1^* \cdots x_d^* - y_1 \cdots y_d)_+^p d\mathbf{y} \\ &\geq \int_{[\mathbf{h}, \mathbf{1}+\mathbf{h}-\mathbf{x}^*]} ((1 + \varepsilon) \cdot h^d - y_1 \cdots y_d)_+^p d\mathbf{y}, \end{aligned}$$

with the second inequality, as in the First case, due to a simple change of variables, $\mathbf{z} := \mathbf{y} + \mathbf{h} - \mathbf{x}^*$, and the monotonicity of the integrand.

From now on, we take $h = (1 + \varepsilon)^{-1/d}$. Then

$$(29) \quad \text{disc}_p^*(\mathbf{t}_1, \dots, \mathbf{t}_n)^p \geq \int_{[\mathbf{h}, \mathbf{1}]} (1 - y_1 \cdots y_d)^p d\mathbf{y} := B_p^d(\varepsilon).$$

This proves that

$$\text{disc}_p^*(\mathbf{t}_1, \dots, \mathbf{t}_n)^p \geq \min(A_p^d(\varepsilon), B_p^d(\varepsilon)).$$

As a consequence we obtain the following result by integration.

Theorem 6. *Assume that $d \in \mathbb{N}$, $\varepsilon = \text{disc}_\infty^1(n, d) \in (0, 1)$ and let $h = (1 + \varepsilon)^{-1/d}$. Then for all even p we have*

$$\text{av}_p^*(n, d) \geq \min(A_p^d(\varepsilon), B_p^d(\varepsilon))^{1/p}.$$

Therefore

$$n_\infty^1(d, \varepsilon) \leq \min \left\{ n : \exists \text{ even } p \text{ such that } \text{av}_p^*(n, d) \leq \min(A_p^d(\varepsilon), B_p^d(\varepsilon))^{1/p} \right\}.$$

To obtain a simpler estimate on $n_\infty^1(d, \varepsilon)$, we need lower bounds for the numbers $A_p^d(\varepsilon)$ and $B_p^d(\varepsilon)$.

Lemma 3. *Assume that p is even and $\varepsilon \leq 0.5$. Then*

$$\min(A_p^d(\varepsilon), B_p^d(\varepsilon))^{1/p} \geq \frac{\sqrt{2}}{3\sqrt{3}} \varepsilon \left(\frac{\varepsilon}{4d} \right)^{d/p}.$$

Proof: From the definition of $B_p^d(\varepsilon)$ we have

$$B_p^d(\varepsilon) \geq \int_{h \leq y_i \leq (1+\varepsilon/2)^{-1/d}} (1 - (1 + \varepsilon/2)^{-1})^p dx = ((1 + \varepsilon/2)^{-1/d} - h)^d \cdot (1 - (1 + \varepsilon/2)^{-1})^p,$$

with $h = (1 + \varepsilon)^{-1/d}$. Using this and an inequality, which can be checked by an elementary technique,

$$(1 + \varepsilon/2)^{-1/d} - (1 + \varepsilon)^{-1/d} \geq \frac{\varepsilon}{4d}, \quad \varepsilon \in (0, 0.5),$$

we obtain

$$B_p^d(\varepsilon)^{1/p} \geq \frac{2}{5} \varepsilon \left(\frac{\varepsilon}{4d} \right)^{d/p}.$$

For $A_p^d(\varepsilon)$ we proceed similarly and obtain

$$A_p^d(\varepsilon) \geq \int_{h(1-\varepsilon/2)^{1/d} \leq y_i \leq h} (\varepsilon h^d - h^d + h^d(1 - \varepsilon/2))^p dx = h^{d(1+p)} (1 - (1 - \varepsilon/2)^{1/d})^d (\varepsilon/2)^p.$$

Thus

$$A_p^d(\varepsilon) \geq (2/3)^{p+1} \cdot 2^{-p} \cdot \varepsilon^p \cdot [1 - (1 - \varepsilon/2)^{1/d}]^d.$$

Since

$$1 - (1 - \varepsilon/2)^{1/d} \geq \frac{\varepsilon}{2d}$$

can be checked in an elementary way, we finally have

$$A_p^d(\varepsilon)^{1/p} \geq \frac{\sqrt{2}}{3^{3/2}} \varepsilon \left(\frac{\varepsilon}{2d} \right)^{d/p},$$

and the lemma follows. \square

We are ready to estimate $n_\infty^1(d, \varepsilon)$. We apply Theorem 6 with $\text{av}_p^*(n, d)$ replaced by its upper bound given by Theorem 5 and $\min(A_p^d(\varepsilon), B_p^d(\varepsilon))^{1/p}$ replaced by its

lower bound given by Lemma 3. Then $n_\infty^1(d, \varepsilon)$ is no larger than n which satisfies the inequality

$$\frac{4\sqrt{2}p}{(1+p/2)^{d/p}} \frac{1}{\sqrt{n}} A_n \leq \frac{\sqrt{2}}{3\sqrt{3}} \varepsilon \left(\frac{\varepsilon}{4d}\right)^{d/p},$$

where

$$A_n = \left(\sum_{i=0}^{p/2-1} n^{-i} \left(\frac{1+p/2}{1+p/2-i} \right)^d \right)^{1/p}.$$

Let $p = 2kd$ with a fixed integer k . We now show that for

$$n \geq 2(1+kd)^2$$

we have

$$A_n \leq (kd)^{1/(2kd)} \leq 3^{2/3}.$$

Indeed, we have

$$n^{-i} \left(\frac{1+p/2}{1+p/2-i} \right)^d \leq \frac{(1+kd)^{d-2i}}{2^i(1+kd-i)^d} \leq 1.$$

The last inequality trivially holds for $i \geq d/2$, and for $i \leq d/2$ it can be checked by induction. From this and Lemma 2 we easily conclude our final result.

Theorem 7. *For $\varepsilon < 0.5$ and $k \in \mathbb{N}$, we have*

$$n_\infty^*(d, \varepsilon) \leq C_k \cdot d^2 \cdot \varepsilon^{-2-1/k},$$

where

$$C_k = \sup_{0 < \varepsilon < 0.5, d \in \mathbb{N}} \left(\frac{\min\{n : av_{2kd}^*(n, d) \leq \min(A_{2kd}^d(\varepsilon), B_{2kd}^d(\varepsilon))^{1/(2kd)}\}}{d^2 \varepsilon^{-2-1/k}} \right) < \infty.$$

Remark 1. As already mentioned in the introduction, a polynomial upper bound on $n_\infty^*(d, \varepsilon)$ could be obtained using only the upper bound on $av_p^*(n, d)$ and the following result due to Niederreiter et al. (1990):

For every $n, p < \infty$ and $(\mathbf{t}_1, \dots, \mathbf{t}_n) \in [0, 1]^{nd}$,

$$(30) \quad \text{disc}_\infty^*(\mathbf{t}_1, \dots, \mathbf{t}_n) \leq d(3d+4) c(d, p)^{-1/(p+d)} \text{disc}_p^*(\mathbf{t}_1, \dots, \mathbf{t}_n)^{p/(p+d)},$$

where

$$c(d, p) = \int_{[0,1]^d} \|\mathbf{y}\|_1^p d\mathbf{y}.$$

However, the results obtained this way would exhibit stronger dependence on d . That is, we would only get

$$(31) \quad n_\infty^*(d, \varepsilon) \leq C_\delta d^{4+2\delta} \varepsilon^{-2-\delta}$$

instead of $n_\infty^*(d, \varepsilon) \leq C_\delta d^2 \varepsilon^{-2-\delta}$, as proven in Theorem 7.

For completeness, we outline the proof of (31). Let $p = 2\alpha d$ for large integer α . We show that

$$d^{-1+1/(2\alpha+1)} \leq c(d, p)^{-1/(p+d)} \leq 3^{1/3} d^{-1+2/(2\alpha+1)}.$$

To prove the lower bound, note that

$$c(d, p) = \int_{[0,1]^d} \left(\sum_{i=1}^d y_i \right)^p d\mathbf{y} \leq d^{p-1} \int_{[0,1]^d} \sum_{i=1}^d y_i^p d\mathbf{y} = d^p / (p + 1).$$

Hence,

$$c(d, p)^{-1/(p+d)} \geq \frac{(p + 1)^{1/(p+d)}}{d^{p/(p+d)}} \geq d^{-\frac{2\alpha}{2\alpha+1}}.$$

To prove the upper bound, note that

$$c(d, p) = \sum_{i_1, i_2, \dots, i_p \in \{1, 2, \dots, d\}} \int_{[0,1]^d} y_{i_1} \cdots y_{i_p} d\mathbf{y}.$$

Let $\beta_j = \#\{i_k : i_k = j\}$ for $j = 1, 2, \dots, d$. Then the last integral is

$$\prod_{j=1}^d \frac{1}{1 + \beta_j} \geq \frac{1}{(1 + p)^d}.$$

Hence,

$$c(d, p)^{-1/(p+d)} \leq \frac{(p + 1)^{d/(p+d)}}{d^{p/(p+d)}} \leq \frac{((2\alpha + 1)d)^{1/(2\alpha+1)}}{d^{2\alpha/(2\alpha+1)}} \leq 3^{1/3} d^{-\frac{2\alpha-1}{2\alpha+1}}.$$

Let $\varepsilon = \inf_{(\mathbf{t}_1, \dots, \mathbf{t}_n) \in [0,1]^{nd}} \text{disc}_\infty^*(\mathbf{t}_1, \dots, \mathbf{t}_n)$. Then (30) implies that

$$\begin{aligned} \varepsilon &\leq 7 \cdot 3^{1/3} d^{1+2/(2\alpha+1)} \left(\inf_{\mathbf{t}_1, \dots, \mathbf{t}_n \in [0,1]^{nd}} \text{disc}_p^*(\mathbf{t}_1, \dots, \mathbf{t}_n) \right)^{2\alpha/(2\alpha+1)} \\ &\leq 7 \cdot 3^{1/3} d^{1+2/(2\alpha+1)} (\text{av}_p^*(n, d))^{2\alpha/(2\alpha+1)}. \end{aligned}$$

Choosing α large enough, using the upper bound on $\text{av}_p^*(n, d)$ and solving for $n = n_\infty^*(\varepsilon, d)$ yield (31). \square

We end this section by comparing the upper bounds from Theorems 1 and 6. We computed the estimates from Theorem 1 or, more precisely, from

$$n_\infty^*(d, \varepsilon) \leq \lfloor 8\varepsilon^{-2}(d \log(\lceil 2d/\varepsilon \rceil + 1) + \log 2) + 1 \rfloor$$

which follows from relation (10) in the proof of it, and obtained:

$$\begin{aligned}
n_{\infty}^*(5, 0.45) &\leq 656, \\
n_{\infty}^*(10, 0.45) &\leq 1540, \\
n_{\infty}^*(20, 0.45) &\leq 3583, \\
n_{\infty}^*(40, 0.45) &\leq 8225, \\
n_{\infty}^*(60, 0.45) &\leq 13281, \\
n_{\infty}^*(80, 0.45) &\leq 18604, \\
n_{\infty}^*(5, 0.1) &\leq 19015, \\
n_{\infty}^*(10, 0.1) &\leq 42981, \\
n_{\infty}^*(20, 0.1) &\leq 96458, \\
n_{\infty}^*(40, 0.1) &\leq 214503.
\end{aligned}$$

As for Theorem 6 one can compute exactly $\text{av}_p^*(n, d)$, $A_p^d(\varepsilon)$, and $B_p^d(\varepsilon)$ and select the value of p in an optimal way to compute the following bounds.³ Observe that Theorem 6 gives upper bounds for the numbers $n_{\infty}^1(d, \varepsilon)$. Here we allow arbitrary constant weights c instead of $c = 1/n$. Of course one could use Lemma 2 to obtain also bounds for the $n_{\infty}^*(d, \varepsilon)$; these bounds would be larger – roughly by a factor of four.

$$\begin{aligned}
n_{\infty}^1(5, 0.45) &\leq 87 && \text{for } p = 24, \\
n_{\infty}^1(10, 0.45) &\leq 184 && \text{for } p = 50, \\
n_{\infty}^1(20, 0.45) &\leq 379 && \text{for } p = 106, \\
n_{\infty}^1(40, 0.45) &\leq 770 && \text{for } p = 216, \\
n_{\infty}^1(60, 0.45) &\leq 1160 && \text{for } p = 332, \\
n_{\infty}^1(80, 0.45) &\leq 1551 && \text{for } p = 440, \\
n_{\infty}^1(5, 0.1) &\leq 1541 && \text{for } p = 42, \\
n_{\infty}^1(10, 0.1) &\leq 3176 && \text{for } p = 86, \\
n_{\infty}^1(20, 0.1) &\leq 6464 && \text{for } p = 174, \\
n_{\infty}^1(40, 0.1) &\leq 13620 && \text{for } p = 356.
\end{aligned}$$

Hence, for these specific values of d and ε , the estimates based on the average case analysis of the L_p -star discrepancy are better.

3. LOWER BOUNDS

In this section we prove some lower bounds for the inverse of the *-discrepancy. In fact, these lower bounds are valid for arbitrary coefficients in the discrepancy formula.

³This computation is not trivial due to the great cancellation of positive and negative terms in the formula for $\text{av}_p^*(n, d)$. Norbert Müller (Trier) computed the values with his “real RAM implementation using C++”. The code of his program can be downloaded from <http://www.informatik.uni-trier.de/~mueller>.

More precisely, we extend the formula (2) by allowing arbitrary a_i 's instead of n^{-1} ,

$$d_{\infty}^*(\mathbf{t}_1, \dots, \mathbf{t}_n; a_1, \dots, a_n) = \sup_{\mathbf{x} \in [0,1]^d} \left| x_1 \cdots x_d - \sum_{i=1}^n a_i 1_{[0,\mathbf{x})}(\mathbf{t}_i) \right|$$

and

$$n_{\infty}(d, \varepsilon) = \min \left\{ n : \inf_{\mathbf{t}_i, a_i} d_{\infty}^*(\mathbf{t}_1, \dots, \mathbf{t}_n; a_1, \dots, a_n) \leq \varepsilon \right\}.$$

Of course,

$$n_{\infty}(d, \varepsilon) \leq n_{\infty}^*(d, \varepsilon).$$

We begin with the following theorem.

Theorem 8. *There exist positive numbers c and ε_0 such that for all d and all $\varepsilon \in (0, \varepsilon_0]$,*

$$n_{\infty}(d, \varepsilon) \geq c d \log \varepsilon^{-1}.$$

Proof: For a fixed d and $\varepsilon \in (0, 1/64]$, let $m = \lfloor (64\varepsilon)^{-1} \rfloor + 1$ and

$$X = \left\{ 1 - \frac{i}{2(m-1)d} \quad : \quad i = 0, 1, \dots, m-1 \right\}^d.$$

Let X_0 be a subset of X of maximal cardinality such that for all $\mathbf{x}, \mathbf{y} \in X_0$ with $\mathbf{x} \neq \mathbf{y}$,

$$(32) \quad |\{j : x_j \neq y_j\}| > \frac{d}{4}.$$

We now prove that

$$(33) \quad |X_0| \geq m^d e^{d/8} 2^{-d} (m-1)^{-d/4}.$$

Indeed, for any fixed $\mathbf{y} \in X_0$ we have

$$\left| \left\{ \mathbf{x} \in X : |\{j : x_j \neq y_j\}| \leq \frac{d}{4} \right\} \right| = \sum_{A \subset \{1, \dots, d\}, |A| \leq d/4} (m-1)^{|A|} \leq (m-1)^{d/4} f(d),$$

where $f(d)$ is the number of subsets of $\{1, \dots, d\}$ of cardinality not exceeding $d/4$,

$$f(d) = \binom{d}{0} + \cdots + \binom{d}{\lfloor d/4 \rfloor}.$$

We show that

$$f(d) \leq 2^d e^{-d/8}.$$

To do this, let ε_i , $i = 1, \dots, d$, be independent Bernoulli random variables with $\mathbb{P}\{\varepsilon_i = 1\} = \mathbb{P}\{\varepsilon_i = -1\} = 1/2$. Consider the sets

$$A_{\omega} = \{i : \varepsilon_i(\omega) = -1\}.$$

Then $k = |A_\omega| = \sum_{i=1}^d (1 - \varepsilon_i(\omega))/2$ iff $\sum_{i=1}^d \varepsilon_i(\omega) = d - 2k$. Clearly, $\mathbb{P}\{|A_\omega| = k\} = 2^{-d} \binom{d}{k}$ and

$$f(d) = 2^d \sum_{k=0}^{\lfloor d/4 \rfloor} \mathbb{P}\{|A_\omega| = k\} = 2^d \mathbb{P}\left\{\sum_{i=1}^d \varepsilon_i \geq d/2\right\} \leq 2^d e^{-d/8}$$

by Hoeffding's inequality, see Pollard, 1984, p. 191, as claimed. Hence,

$$\left|\left\{\mathbf{x} \in X : |\{j : x_j \neq y_j\}| \leq \frac{d}{4}\right\}\right| \leq 2^d e^{-d/8} (m-1)^{d/4}.$$

The maximality of X_0 implies

$$(34) \quad 2^d e^{-d/8} (m-1)^{d/4} |X_0| \geq |X| = m^d,$$

because otherwise we could find an $\mathbf{x} \in X$ such that for all $\mathbf{y} \in X_0$, (32) is satisfied. This proves (33).

Next we show that for any $\mathbf{x}, \mathbf{y} \in X_0$ with $\mathbf{x} \neq \mathbf{y}$,

$$(35) \quad \max(|[\mathbf{0}, \mathbf{x}]| - |[\mathbf{0}, \mathbf{z}]|, |[\mathbf{0}, \mathbf{y}]| - |[\mathbf{0}, \mathbf{z}]|) > \frac{1}{32(m-1)},$$

where $\mathbf{z} = \min(\mathbf{x}, \mathbf{y})$ denotes the vector of the component-wise minimum. Indeed, let

$$\ell_1 = |\{j : x_j > y_j\}| \quad \text{and} \quad \ell_2 = |\{j : x_j < y_j\}|.$$

From (32),

$$(36) \quad \max(\ell_1, \ell_2) \geq \frac{\ell_1 + \ell_2}{2} > \frac{d}{8}.$$

As in the proof of Theorem 1, we derive

$$\begin{aligned} |[\mathbf{0}, \mathbf{x}]| - |[\mathbf{0}, \mathbf{z}]| &= \prod_{k=1}^d x_k - \prod_{k=1}^d z_k \\ &= \sum_{k=1}^d (z_1 \dots z_{k-1} x_k x_{k+1} \dots x_d - z_1 \dots z_{k-1} z_k x_{k+1} \dots x_d) \\ &\geq \sum_{k=1}^d (x_k - z_k) \prod_{j=1}^d z_j \geq \frac{\ell_1}{2(m-1)d} \left(1 - \frac{1}{2d}\right)^d \geq \frac{\ell_1}{4(m-1)d}, \end{aligned}$$

since $(1 - \frac{1}{2d})^d \geq \frac{1}{2}$ holds due to Bernoulli's inequality. Arguing in the same way, with \mathbf{x} replaced by \mathbf{y} , and taking into account (36), we obtain (35).

Now assume that n is any natural number with

$$(37) \quad 2^n < |X_0|.$$

Let $\mathbf{t}_1, \dots, \mathbf{t}_n \in [0, 1]^d$ and $a_1, \dots, a_n \in \mathbb{R}$ be arbitrary. Since there are at most 2^n different subsets of $\{\mathbf{t}_1, \dots, \mathbf{t}_n\}$, (37) yields that there exist $\mathbf{x}, \mathbf{y} \in X_0$ such that $\mathbf{x} \neq \mathbf{y}$ and

$$[\mathbf{0}, \mathbf{x}] \cap \{\mathbf{t}_1, \dots, \mathbf{t}_n\} = [\mathbf{0}, \mathbf{y}] \cap \{\mathbf{t}_1, \dots, \mathbf{t}_n\}.$$

It follows that

$$\sum_{i=1}^n a_i 1_{[\mathbf{0}, \mathbf{x})}(\mathbf{t}_i) = \sum_{i=1}^n a_i 1_{[\mathbf{0}, \mathbf{y})}(\mathbf{t}_i) = \sum_{i=1}^n a_i 1_{[\mathbf{0}, \mathbf{z})}(\mathbf{t}_i),$$

where $\mathbf{z} = \min(\mathbf{x}, \mathbf{y})$. From (35) we conclude

$$\begin{aligned} & \sup_{\mathbf{u} \in [0,1]^d} \left| |[\mathbf{0}, \mathbf{u}]| - \sum_{i=1}^n a_i 1_{[\mathbf{0}, \mathbf{u})}(\mathbf{t}_i) \right| \\ & \geq \max_{\mathbf{u} \in \{\mathbf{x}, \mathbf{y}\}} \max \left(\left| |[\mathbf{0}, \mathbf{u}]| - \sum_{i=1}^n a_i 1_{[\mathbf{0}, \mathbf{u})}(\mathbf{t}_i) \right|, \left| |[\mathbf{0}, \mathbf{z}]| - \sum_{i=1}^n a_i 1_{[\mathbf{0}, \mathbf{z})}(\mathbf{t}_i) \right| \right) \\ & \geq \frac{1}{2} \max_{\mathbf{u} \in \{\mathbf{x}, \mathbf{y}\}} (|[\mathbf{0}, \mathbf{u}]| - |[\mathbf{0}, \mathbf{z}]|) > \frac{1}{64(m-1)} \geq \varepsilon. \end{aligned}$$

Consequently, (37) cannot hold with $n = n_\infty(d, \varepsilon)$, which together with (33) yields

$$(38) \quad 2^{n_\infty(d, \varepsilon)} \geq m^d e^{d/8} 2^{-d} (m-1)^{-d/4}$$

and implies

$$n_\infty(d, \varepsilon) \geq \frac{3}{4} d \log \left\lceil \frac{1}{64\varepsilon} \right\rceil + \frac{d}{8} - d \log 2 \geq c d \log \varepsilon^{-1}$$

for all $\varepsilon \in (0, \varepsilon_0]$, where $\varepsilon_0 \in (0, 1/64]$ and a positive c are appropriately chosen. This completes the proof. \square

Theorems 3 and 8 show that the inverse of the $*$ -discrepancy depends linearly on d . However, the dependence on ε^{-1} in these theorems is quite different. We now show a lower bound with an improved dependence on ε^{-1} at the expense of the behavior on d . To do this we first notice what happens for $\varepsilon = 1/64$.

Corollary 1. *For all $d \in \mathbb{N}$,*

$$(39) \quad n_\infty(d, 1/64) \geq \frac{d}{8 \log 2} \geq 0.18 d.$$

This follows directly from (38) since $m = 2$ for $\varepsilon = 1/64$.

By combining known results on the $*$ -discrepancy and Corollary 1, we are ready to show the following theorem.

Theorem 9. *For any $\lambda \in (0, 1)$ and any natural number m there exists a positive number $c_{\lambda, m}$ such that for all $\varepsilon \in (0, 1/64]$ and all d ,*

$$n_\infty(d, \varepsilon) \geq c_{\lambda, m} d^\lambda \left(\frac{1}{\varepsilon} \right)^{1-\lambda} \left(\log \frac{1}{\varepsilon} \right)^m.$$

Proof: Denote

$$d_0 = 2 \left\lceil \frac{m}{1-\lambda} \right\rceil + 1.$$

Observe that, obviously, $n_\infty(d_1, \varepsilon) \geq n_\infty(d_2, \varepsilon)$ whenever $d_1 \geq d_2$. By the results of Roth (1954) for the L_2 -star discrepancy⁴, which is a lower bound on the $*$ -discrepancy, we have for all $\varepsilon \in (0, 1)$ and $d \geq d_0$,

$$(40) \quad n_\infty(d, \varepsilon) \geq n_\infty(d_0, \varepsilon) \geq c_{d_0} \frac{1}{\varepsilon} \left(\log \frac{1}{\varepsilon} \right)^{\frac{d_0-1}{2}} \geq c_{d_0} \frac{1}{\varepsilon} \left(\log \frac{1}{\varepsilon} \right)^{\frac{m}{1-\lambda}}$$

for some positive c_{d_0} . Combining (39) and (40) through the trivial relation

$$\max(a, b) \geq a^\lambda b^{1-\lambda} \quad (a, b > 0),$$

we get

$$(41) \quad \begin{aligned} n_\infty(d, \varepsilon) &\geq c_{d_0}^{1-\lambda} (8 \log 2)^{-\lambda} d^\lambda \left(\frac{1}{\varepsilon} \right)^{1-\lambda} \left(\log \frac{1}{\varepsilon} \right)^m \\ &= c_{\lambda, m} d^\lambda \left(\frac{1}{\varepsilon} \right)^{1-\lambda} \left(\log \frac{1}{\varepsilon} \right)^m, \end{aligned}$$

for all $\varepsilon \in (0, 1/64]$ and $d \geq d_0$.

On the other hand, for $d \leq d_0$ and $0 < \varepsilon \leq 1/64$, there exists a positive c such that

$$\begin{aligned} n_\infty(d, \varepsilon) &\geq n_\infty(1, \varepsilon) \geq \frac{c}{\varepsilon} \\ &\geq c d_0^{-\lambda} \left(\frac{1}{\varepsilon} \right)^\lambda \left(\log \frac{1}{\varepsilon} \right)^{-m} d^\lambda \left(\frac{1}{\varepsilon} \right)^{1-\lambda} \left(\log \frac{1}{\varepsilon} \right)^m \\ &\geq c'_{\lambda, m} d^\lambda \left(\frac{1}{\varepsilon} \right)^{1-\lambda} \left(\log \frac{1}{\varepsilon} \right)^m, \end{aligned}$$

with

$$c'_{\lambda, m} = c d_0^{-\lambda} \inf_{\varepsilon \in (0, 1/64]} \left(\frac{1}{\varepsilon} \right)^\lambda \left(\log \frac{1}{\varepsilon} \right)^{-m} > 0,$$

which completes the proof. \square

Acknowledgments. The authors would like to thank Norbert Müller (Trier) for various computations related to the estimates in Section 2.2 (see also the footnote at the end of that section). The first named author would like to thank Silvelyn Zwanzig (Hamburg) for valuable information on empirical process theory.

REFERENCES

1. J. Beck (1988): On the discrepancy of convex plane sets. *Monatsh. Math.* **105**, 91-106.
2. W. W. L. Chen (1985): On irregularities of distribution and approximate evaluation of certain functions. *Quarterly J. Math. Oxford Ser.* **36**, 173-182.

⁴Roth (1954) proved this result for $a_i = n^{-1}$ but the same proof technique applies for arbitrary a_i as shown by Chen (1985).

3. M. Drmota and R. F. Tichy (1997): Sequences, Discrepancies and Applications. Lecture Notes in Mathematics 1651, Springer-Verlag, Berlin.
4. R. M. Dudley (1978): Central limit theorems for empirical measures. *Annals of Probability*, **6**, 899–929.
5. R. M. Dudley (1984): A course on empirical processes (École d’Été de Probabilités de Saint-Flour XII-1982). Lecture Notes in Mathematics 1097, 2–141, Springer-Verlag, New York.
6. D. Haussler (1995): Sphere packing numbers for subsets of the Boolean n -cube with bounded Vapnik-Červonenkis dimension. *Journal of Combinatorial Theory*, **A 69**, 217–232.
7. J. Kiefer (1961): On large deviations of the empiric d.f. of vector chance variables and a law of the iterated logarithm. *Pacific J. Math.* **11**, 649–660.
8. G. Larcher (1998): Talk at Dagstuhl-Seminar 98201.
9. J. Matoušek (1999): Geometric Discrepancy. Springer-Verlag, Berlin.
10. H. Niederreiter (1992): Random Number Generation and Quasi-Monte Carlo Methods. SIAM, Philadelphia.
11. H. Niederreiter, R. F. Tichy and G. Turnwald (1990): An inequality for differences of distribution functions. *Arch. Math.* **54**, 166–172.
12. E. Novak and H. Woźniakowski (1999): Intractability results for integration and discrepancy. Submitted to *J. Complexity*.
13. D. Pollard (1984): Convergence of Stochastic Processes. Springer-Verlag, New York.
14. J. Riordan (1958): An Introduction to Combinatorial Analysis. Wiley and Sons, New York.
15. K. F. Roth (1954): On irregularities of distribution. *Mathematika* **1**, 73–79.
16. W. M. Schmidt (1975): On irregularities of distribution IX. *Acta Arith.* **27**, 385–396.
17. I. H. Sloan and H. Woźniakowski (1998): When are quasi-Monte Carlo algorithms efficient for high dimensional integrals? *J. Complexity* **14**, 1–33.
18. W. Stute (1977): Convergence rates for the isotropic discrepancy. *Ann. Prob.* **5**, 707–723.
19. M. Talagrand (1994): Sharper bounds for Gaussian and empirical processes. *Annals of Probability*, **22**, 28–76.
20. J. F. Traub and A. G. Werschulz (1998): Complexity and Information. Cambridge University Press.
21. A. W. van der Vaart and J. A. Wellner (1996): Weak Convergence and Empirical Processes, with Applications to Statistics. Springer-Verlag, New York.
22. H. Woźniakowski (1991): Average case complexity of multivariate integration. *Bull of AMS (new series)* **24**, 185–191.
23. H. Woźniakowski (1999): 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, 114–136.

S. HEINRICH, FB INFORMATIK, UNIVERSITÄT KAISERSLAUTERN, POSTFACH 3049, D-67653
KAISERSLAUTERN, GERMANY

E-mail address: heinrich@informatik.uni-kl.de

E. NOVAK, MATHEMATISCHES INSTITUT, UNIVERSITÄT ERLANGEN-NÜRNBERG, BISMARCKSTR.
1 1/2, D-91054 ERLANGEN, GERMANY

E-mail address: novak@mi.uni-erlangen.de

G. W. WASILKOWSKI, DEPARTMENT OF COMPUTER SCIENCE, UNIVERSITY OF KENTUCKY, LEX-
INGTON, KY 40506, USA

E-mail address: greg@cs.uky.edu

H. WOŹNIAKOWSKI, DEPARTMENT OF COMPUTER SCIENCE, COLUMBIA UNIVERSITY, NEW YORK
10027, NY, USA

E-mail address: henryk@cs.columbia.edu