ACTA ARITHMETICA
* (200*)

The inverse of the star-discrepancy depends linearly on the dimension

by

Stefan Heinrich (Kaiserslautern), Erich Novak (Erlangen),
Grzegorz W. Wasilkowski (Lexington, KY)
and Henryk Woźniakowski (New York, NY)

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

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

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

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

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

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

2000 *Mathematics Subject Classification*: 11K38, 11B73, 65D32, 68Q25.

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.

denote the minimal L_p-star discrepancy for n points in dimension d. For $n = 0$, the L_p-star discrepancy $\mathrm{disc}_p^*(0,d)$ is the initial error of multivariate integration without sampling the function. We have $\mathrm{disc}_p^*(0,d) = (p+1)^{-d/p}$, 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 $\mathrm{disc}_\infty^*(0,d) = 1$.

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 $\mathrm{disc}_p^*(n,d)$ 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 $n \le \exp(d-1)$. 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 $\mathrm{disc}_2^*(n,d) \le \varepsilon\,\mathrm{disc}_2^*(0,d)$ for $\varepsilon \in (0,1)$ requires n to be exponential in d. More precisely, we must have

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

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)

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

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 $\mathrm{disc}_\infty^*(n,d)$ tends to 1

([1]) The 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.

for $n = a^d$ as d goes to infinity, and also asked whether, in particular, $\mathrm{disc}^*_\infty(2^d, d)$ 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 $\mathrm{disc}^*_\infty(n, d)$ depends only *polynomially* on d and n^{-1} (by this we mean that $\mathrm{disc}^*_\infty(n, d)$ 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

$$\mathrm{disc}^*_\infty(n, d) \le c d^{1/2} n^{-1/2} \qquad \forall n, d = 1, 2, \ldots, \tag{3}$$

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 nonconstructive 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

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

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,

$$n^*_\infty(d, \varepsilon) = \min\{n : \mathrm{disc}^*_\infty(n, d) \le \varepsilon\}. \tag{5}$$

In particular, (3) yields

$$n^*_\infty(d, \varepsilon) = O(d \varepsilon^{-2}) \tag{6}$$

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 Lemma 2 and 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

$$n^*_\infty(d,\varepsilon) \le C_k d^2 \varepsilon^{-2-1/k} \quad \forall k = 1,2,\dots \tag{7}$$

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, $\mathrm{disc}^*_p(n,d) \le \mathrm{disc}^*_\infty(n,d)$. 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 $n^*_\infty(d,1/64) \ge 0.18d$. 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* $n, d \in \mathbb{N}$,

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

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 $\mathbf{t}_1, \dots, \mathbf{t}_n \in [0,1]^d$,

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

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

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

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

$$x_j \le x_j^* < x_j + 1/m =: y_j \qquad \forall j = 1,\dots,d.$$

Clearly,

$$y_1 \dots y_d - x_1 \dots x_d$$
$$= \sum_{k=1}^{d} (x_1 \dots x_{k-1} y_k y_{k+1} \dots y_d - x_1 \dots x_{k-1} x_k y_{k+1} \dots y_d) \le d/m \le \delta.$$

Hence,

$$-\delta + \prod_{i=1}^{d} y_i \le \prod_{i=1}^{d} x_i^* \le \prod_{i=1}^{d} y_i \le \delta + \prod_{i=1}^{d} x_i.$$

This yields

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

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

Let $\tau_1, \dots, \tau_n$ be independent random variables, uniformly distributed on $[0,1]^d$. For $\mathbf{x} \in [0,1]^d$, let

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

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

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 $\mathbb{P}$:

$$\mathbb{P}\left\{ \left| n^{-1} \sum_{i=1}^{n} \xi_{\mathbf{x}}^{(i)} \right| \ge \delta \right\} \le 2 \exp(-\delta^2 n/2). \tag{9}$$

From (8) and (9) we have

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

The latter is strictly positive if

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

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

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

This implies that

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

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

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

This means that for any $\delta > \delta_0$ there exist points $\tau_1, \dots, \tau_n$ such that

$$\operatorname{disc}^*_\infty(\tau_1, \dots, \tau_n) \le 2\delta. \tag{13}$$

Hence, $\operatorname{disc}^*_\infty(n,d) \le 2\delta_0$, which completes the proof. ∎

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 $(\Omega, \Sigma, \mathbb{P})$ 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 $s \ge Kv(\mathcal{C})^{1/2}$ and all natural n,*

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

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

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

$$\operatorname{disc}^*_\infty(n,d) \le cd^{1/2}n^{-1/2} \tag{15}$$

and consequently

$$n^*_\infty(d,\varepsilon) \le \lceil c^2 d\varepsilon^{-2}\rceil.$$

P r o o f. We take $M=[0,1]^d$. Let $\mathcal{M}$ be the class of Borel subsets 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 $K\lambda^2 \le e^{2\lambda^2}$ for all $\lambda\ge\lambda_0$. 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. ∎

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

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

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*

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

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

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

where different choices of $\lambda\ge\max(K,\lambda_0,1)$ 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 $d\ge 2$. 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 $d\ge 4$ 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

$$\mathrm{av}_p^*(n,d) = \Big(\int\limits_{[0,1]^{nd}} \mathrm{disc}_p^*(\mathbf{t}_1,\dots,\mathbf{t}_n)^p \, d\mathbf{t} \Big)^{1/p}, \qquad \mathbf{t} = (\mathbf{t}_1,\dots,\mathbf{t}_n).$$

We have

$$\mathrm{av}_p^*(n,d)^p = \int\limits_{[0,1]^{nd}} \sum_{j=0}^{p} \binom{p}{j} \int\limits_{[0,1]^d} (x_1 \dots x_d)^{p-j} \Big(-\frac{1}{n} \sum_{i=1}^{n} 1_{[\mathbf{0},\mathbf{x})}(\mathbf{t}_i) \Big)^j \, d\mathbf{x}\, d\mathbf{t}.$$

For $A = \{1,\dots,n\}^j$, this can be written as

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

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

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

To compute the integral above, note that

$$\begin{aligned} &\int\limits_{[0,1]^j} \min(1-x_1^\alpha,\dots,1-x_j^\alpha)\, dx_1 \dots dx_j \\ &\qquad = j! \int\limits_{0\le x_1\le\dots\le x_j\le 1} (1-x_j^\alpha)\, dx_1 \dots dx_j \\ &\qquad = j! \int_0^1 (1-x_j^\alpha) \Big(\int\limits_{0\le x_1\le\dots\le x_{j-1}\le x_j} dx_1 \dots dx_{j-1} \Big) dx_j \\ &\qquad = j! \int_0^1 (1-x^\alpha) \frac{x^{j-1}}{(j-1)!}\, dx = \frac{\alpha}{j+\alpha}. \end{aligned}$$

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

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

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

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

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

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

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

$$\begin{aligned}\#(j,k,n)=k!\binom{n}{k}S(j,k)&=\sum_{i=0}^{k}s(k,i)S(j,k)n^i\\&=\sum_{i=0}^{k}\binom{n}{k}\binom{k}{i}(-1)^i(k-i)^j.\end{aligned}$$

Here we use $\#(0,0,n)=1$ and $\#(j,0,n)=0$ for $j>0$. This is consistent with the last equality 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 $\mathrm{av}_p^*(n,d)^p$ as a polynomial in n^{-1},

$$\mathrm{av}_p^*(n,d)^p=\sum_{r=0}^{p}C(r,p,d)n^{-r}, \tag{18}$$

where

$$\begin{aligned}&C(r,p,d)\\&\quad=(-1)^r\sum_{i=0}^{p-r}\binom{p}{r+i}(-1)^i\sum_{k=i}^{i+r}(p+1-r+k-i)^{-d}s(k,i)S(i+r,k).\end{aligned} \tag{19}$$

Observe that $C(p,p,d)$ is zero, and therefore the sum in (18) 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

$$C(r,p,d)=\frac{(-1)^r}{(p+1-r)^d}\sum_{j=0}^{r}\left(\frac{p+1-r}{p+1-r+j}\right)^d\beta(r,p,j) \tag{20}$$

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 $r = 0, \ldots, m-1$ and $j = 0, \ldots, r$*

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

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

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

where λ denotes the Lebesgue measure. Let F denote the distribution function of $\operatorname{disc}^*_\infty$. Then $1 - F(x) \leq c(d) \exp(-x^2 n)$ and therefore

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

Due to (18) this proves that

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

Letting d tend to infinity in (20) we conclude that $\beta(r,p,j) = 0$ for $j = 0, 1, \ldots, r$, as claimed. ∎

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

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

$$\operatorname{av}^*_p(n,d)^p = \sum_{r=p/2}^{p-1} C(r,p,d) n^{-r}, \tag{21}$$

where the $C(r,p,d)$ are given by (19). Furthermore

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

and

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

P r o o f. The formula (21) directly follows from Lemma 1. To get upper bounds on $C(r,p,d)$ we use (19) to obtain

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

The Stirling numbers of the first kind are defined by

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

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

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

Later we will use ([2]) only $|s(k,i)| \le k!$. 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

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

Together we obtain

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

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

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

as claimed.

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

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

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

([2]) There 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.

We now 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

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

It is easy to check by rescaling that

$$\operatorname{disc}^h_\infty(h\mathbf{t}_1,\dots,h\mathbf{t}_n) = h^d \cdot \operatorname{disc}^1_\infty(\mathbf{t}_1,\dots,\mathbf{t}_n). \tag{22}$$

We derive upper bounds for the numbers

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

and its inverse function,

$$n^1_\infty(d,\varepsilon) = \min\{n : \operatorname{disc}^1_\infty(n,d) \le \varepsilon\}, \quad \varepsilon \in (0,1). \tag{23}$$

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

LEMMA 2. *We have*

$$n^1_\infty(d,\varepsilon) \le n^*_\infty(d,\varepsilon) \le n^1_\infty(d,\varepsilon/2).$$

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

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

Observe that all components of $\mathbf{t}_i$ for $i = 1,\dots,n$ 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 $|1 - cn| \le \varepsilon/2$. Hence, the weight c differs from $1/n$ by at most $\varepsilon/(2n)$ and therefore, for all x,

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

Hence we obtain

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

Thus, $n^*_\infty(d,\varepsilon) \le n$, as claimed. ∎

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

$$\mathrm{disc}^1_\infty(n,d) = \varepsilon. \tag{24}$$

From (22) we have

$$\mathrm{disc}^h_\infty(\mathbf{t}_1,\dots,\mathbf{t}_n) \ge \varepsilon h^d$$

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

$$\left| x^*_1 \dots x^*_d - \frac{1}{n}\sum_{i=1}^n 1_{[\mathbf{0},\mathbf{x}^*)}(\mathbf{t}_i) \right| \ge \varepsilon h^d.$$

There are two possible cases.

Case 1: We have

$$x^*_1 \dots x^*_d - \frac{1}{n}\sum_{i=1}^n 1_{[\mathbf{0},\mathbf{x}^*)}(\mathbf{t}_i) \ge \varepsilon h^d, \tag{25}$$

where $\mathbf{0} \le \mathbf{x}^* \le \mathbf{h}$ with $x^*_1 \dots x^*_d \ge \varepsilon h^d$, all inequalities are component-wise and by $\mathbf{h}$ we mean the vector $\mathbf{h} = (h,h,\dots,h)$.

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

$$y_1 \dots y_d - \frac{1}{n}\sum_{i=1}^n 1_{[\mathbf{0},\mathbf{y})}(\mathbf{t}_i) \ge \varepsilon h^d - x^*_1 \dots x^*_d + y_1 \dots y_d$$

and therefore

$$\begin{aligned}\mathrm{disc}^*_p(\mathbf{t}_1,\dots,\mathbf{t}_n)^p &\ge \int\limits_{[\mathbf{0},\mathbf{x}^*)} (\varepsilon h^d - x^*_1 \dots x^*_d + y_1 \dots y_d)^p_+ \, d\mathbf{y} \\ &= \int\limits_{[\mathbf{h}-\mathbf{x}^*,\mathbf{h})} (\varepsilon h^d - x^*_1 \dots x^*_d + (z_1-\delta_1)\dots(z_d-\delta_d))^p_+ \, d\mathbf{z},\end{aligned}$$

where $\delta_i = h - x^*_i$. Note that $-x^*_1 \dots x^*_d + (z_1-\delta_1)\dots(z_d-\delta_d) \ge -h^d + z_1 \dots z_d$. Indeed, this follows from the fact that $f(\mathbf{z}) := z_1 \dots z_d - (z_1-\delta_1)\dots$ $\dots(z_d-\delta_d)$ is increasing in each variable z_i and that $f(\mathbf{h}) = h^d - x^*_1 \dots x^*_d$. Therefore

$$\mathrm{disc}^*_p(\mathbf{t}_1,\dots,\mathbf{t}_n)^p \ge \int\limits_{[\mathbf{h}-\mathbf{x}^*,\mathbf{h})} (\varepsilon h^d - h^d + y_1 \dots y_d)^p_+ \, d\mathbf{y}.$$

Note that $h^{d-1}x^*_i \ge x^*_1 \dots x^*_d \ge h^d\varepsilon$ implies $x^*_i \ge h\varepsilon$ and $h - x^*_i \le h - h\varepsilon = h(1-\varepsilon)$. Hence

$$\mathrm{disc}^*_p(\mathbf{t}_1,\dots,\mathbf{t}_n)^p \ge \int\limits_{[(1-\varepsilon)\mathbf{h},\mathbf{h})} (\varepsilon h^d - h^d + y_1 \dots y_d)^p_+ \, d\mathbf{y}.$$

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

$$\text{(26)} \qquad \operatorname{disc}_p^*(\mathbf{t}_1, \dots, \mathbf{t}_n)^p \geq \int\limits_{[(1-\varepsilon)^{1/d}\mathbf{h}, \mathbf{h})} (\varepsilon h^d - h^d + y_1 \dots y_d)^p \, d\mathbf{y} =: A_p^d(\varepsilon).$$

CASE 2: We have

$$\text{(27)} \qquad x_1^* \dots x_d^* - \frac{1}{n} \sum_{i=1}^{n} 1_{[\mathbf{0}, \mathbf{x}^*)}(\mathbf{t}_i) \leq -\varepsilon h^d,$$

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

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

and therefore

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

with the second inequality, as in Case 1, 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

$$\text{(28)} \qquad \operatorname{disc}_p^*(\mathbf{t}_1, \dots, \mathbf{t}_n)^p \geq \int\limits_{[\mathbf{h}, \mathbf{1})} (1 - y_1 \dots y_d)^p \, d\mathbf{y} := B_p^d(\varepsilon).$$

This proves that

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

As a consequence we obtain the following result by integration.

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

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

Therefore

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

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 $\varepsilon \le 0.5$. Then*

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

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

$$\begin{aligned} B_p^d(\varepsilon) &\ge \int\limits_{h \le y_i \le (1+\varepsilon/2)^{-1/d}} (1-(1+\varepsilon/2)^{-1})^p\,dx \\ &= ((1+\varepsilon/2)^{-1/d} - h)^d (1-(1+\varepsilon/2)^{-1})^p, \end{aligned}$$

with $h = (1+\varepsilon)^{-1/d}$. Using this and the inequality

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

which can be checked by an elementary technique, we obtain

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

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

$$\begin{aligned} A_p^d(\varepsilon) &\ge \int\limits_{h(1-\varepsilon/2)^{1/d} \le y_i \le h} (\varepsilon h^d - h^d + h^d(1-\varepsilon/2))^p\,dx \\ &= h^{d(1+p)}(1-(1-\varepsilon/2)^{1/d})^d(\varepsilon/2)^p. \end{aligned}$$

Thus

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

Since $1-(1-\varepsilon/2)^{1/d} \ge \varepsilon/(2d)$ can be checked in an elementary way, we finally have

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

and the lemma follows. ∎

We are ready to estimate $n_\infty^1(d,\varepsilon)$. We apply Theorem 6 with $\mathrm{av}_p^*(n,d)$ 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

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

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 $n \ge 2(1+kd)^2$ we have

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

Indeed, we have

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

The last inequality trivially holds for $i \ge d/2$, and for $i \le d/2$ 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*

$$n^*_\infty(d,\varepsilon) \le C_k d^2 \varepsilon^{-2-1/k},$$

where

$$C_k = \sup_{0<\varepsilon<0.5,\, d\in\mathbb{N}} \left(\frac{\min\{n : \mathrm{av}^*_{2kd}(n,d) \le \min(A^d_{2kd}(\varepsilon), B^d_{2kd}(\varepsilon))^{1/(2kd)}\}}{d^2\varepsilon^{-2-1/k}} \right)$$

is finite.

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 $\mathrm{av}^*_p(n,d)$ and the following result due to Niederreiter *et al.* (1990):

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

$$\mathrm{disc}^*_\infty(\mathbf{t}_1,\dots,\mathbf{t}_n) \le d(3d+4)c(d,p)^{-1/(p+d)}\,\mathrm{disc}^*_p(\mathbf{t}_1,\dots,\mathbf{t}_n)^{p/(p+d)}, \tag{29}$$

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

$$n^*_\infty(d,\varepsilon) \le C_\delta d^{4+2\delta}\varepsilon^{-2-\delta} \tag{30}$$

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

For completeness, we outline the proof of (30). Let $p = 2\alpha d$ for large integer α. We show that

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

To prove the lower bound, note that

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

Hence,

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

To prove the upper bound, note that

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

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

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

Hence,

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

Let $\varepsilon = \inf_{(\mathbf{t}_1,\dots,\mathbf{t}_n)\in[0,1]^{nd}} \mathrm{disc}^*_\infty(\mathbf{t}_1,\dots,\mathbf{t}_n)$. Then (29) implies that

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

Choosing α large enough, using the upper bound on $\mathrm{av}^*_p(n,d)$ and solving for $n = n^*_\infty(\varepsilon,d)$ yield (30). ∎

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 $\mathrm{av}_p^*(n,d)$, $A_p^d(\varepsilon)$, 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.

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

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

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

and

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

Of course,

$$n_\infty(d,\varepsilon) \le n_\infty^*(d,\varepsilon).$$

We begin with the following theorem.

([3]) This computation is not trivial due to the great cancellation of positive and negative terms in the formula for $\mathrm{av}_p^*(n,d)$. 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.

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

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

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

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

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

$$|\{j : x_j \neq y_j\}| > d/4. \tag{31}$$

We now prove that

$$|X_0| \geq m^d e^{d/8} 2^{-d} (m-1)^{-d/4}. \tag{32}$$

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

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

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

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

We show that

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

To do this, let ε_i, $i = 1, \ldots, d$, 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

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

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

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

The maximality of X_0 implies

$$2^d e^{-d/8} (m-1)^{d/4} |X_0| \geq |X| = m^d, \tag{33}$$

because otherwise we could find an $\mathbf{x} \in X$ such that for all $\mathbf{y} \in X_0$, (31) is satisfied. This proves (32).

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

$$\max(|[\mathbf{0},\mathbf{x})| - |[\mathbf{0},\mathbf{z})|, |[\mathbf{0},\mathbf{y})| - |[\mathbf{0},\mathbf{z})|) > \frac{1}{32(m-1)}, \tag{34}$$

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

$$l_1 = |\{j : x_j > y_j\}| \quad \text{and} \quad l_2 = |\{j : x_j < y_j\}|.$$

From (31),

$$\max(l_1, l_2) \geq (l_1 + l_2)/2 > d/8. \tag{35}$$

As in the proof of Theorem 1, we derive

$$\begin{aligned} |[\mathbf{0},\mathbf{x})| - |[\mathbf{0},\mathbf{z})| &= \prod_{k=1}^{d} x_k - \prod_{k=1}^{d} z_k \\ &= \sum_{k=1}^{d} (z_1 \dots z_{k-1} x_k x_{k+1} \dots x_d - z_1 \dots z_{k-1} z_k x_{k+1} \dots x_d) \\ &\geq \sum_{k=1}^{d} (x_k - z_k) \prod_{j=1}^{d} z_j \geq \frac{l_1}{2(m-1)d} \left(1 - \frac{1}{2d}\right)^d \geq \frac{l_1}{4(m-1)d}, \end{aligned}$$

since $(1 - 1/(2d))^d \geq 1/2$ due to Bernoulli's inequality. Arguing in the same way, with $\mathbf{x}$ replaced by $\mathbf{y}$, and taking into account (35), we obtain (34).

Now assume that n is any natural number with

$$2^n < |X_0|. \tag{36}$$

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 $\{\mathbf{t}_1, \dots, \mathbf{t}_n\}$, (36) 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 (34) we conclude

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

Consequently, (36) cannot hold with $n = n_\infty(d,\varepsilon)$, which together with (32) yields

$$2^{n_\infty(d,\varepsilon)} \ge m^d e^{d/8} 2^{-d} (m-1)^{-d/4} \tag{37}$$

and implies

$$n_\infty(d,\varepsilon) \ge \frac{3}{4} d \log \left\lfloor \frac{1}{64\varepsilon} \right\rfloor + \frac{d}{8} - d\log 2 \ge cd \log \varepsilon^{-1}$$

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. ∎

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 $d \in \mathbb{N}$,*

$$n_\infty(d, 1/64) \ge \frac{d}{8 \log 2} \ge 0.18d. \tag{38}$$

This follows directly from (37) 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,*

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

Proof. Denote

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

Observe that, obviously, $n_\infty(d_1,\varepsilon) \ge n_\infty(d_2,\varepsilon)$ whenever $d_1 \ge d_2$. By the results of Roth (1954) for the L_2-star discrepancy ([4]), which is a lower bound on the $*$-discrepancy, we have for all $\varepsilon \in (0,1)$ and $d \ge d_0$,

$$n_\infty(d,\varepsilon) \ge n_\infty(d_0,\varepsilon) \ge c_{d_0} \frac{1}{\varepsilon} \left(\log \frac{1}{\varepsilon}\right)^{(d_0-1)/2} \ge c_{d_0} \frac{1}{\varepsilon} \left(\log \frac{1}{\varepsilon}\right)^{m/(1-\lambda)} \tag{39}$$

for some positive c_{d_0}. Combining (38) and (39) through the trivial relation

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

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

we get

$$n_\infty(d,\varepsilon) \geq c_{d_0}^{1-\lambda}(8\log 2)^{-\lambda} d^\lambda \left(\frac{1}{\varepsilon}\right)^{1-\lambda} \left(\log \frac{1}{\varepsilon}\right)^m \tag{40}$$
$$= c_{\lambda,m} d^\lambda \left(\frac{1}{\varepsilon}\right)^{1-\lambda} \left(\log \frac{1}{\varepsilon}\right)^m,$$

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 exits a positive c such that

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

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. ∎

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

References

J. Beck (1988), *On the discrepancy of convex plane sets*, Monatsh. Math. 105, 91–106.

W. W. L. Chen (1985), *On irregularities of distribution and approximate evaluation of certain functions*, Quart. J. Math. Oxford Ser. 36, 173–182.

M. Drmota and R. F. Tichy (1997), *Sequences, Discrepancies and Applications*, Lecture Notes in Math. 1651, Springer, Berlin.

R. M. Dudley (1978), *Central limit theorems for empirical measures*, Ann. Probab. 6, 899–929.

R. M. Dudley (1984), *A course on empirical processes*, Lecture Notes in Math. 1097, Springer, New York, 2–141.

D. Haussler (1995), *Sphere packing numbers for subsets of the Boolean n-cube with bounded Vapnik–Červonenkis dimension*, J. Combin. Theory A 69, 217–232.

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.

G. Larcher (1998), Talk at Dagstuhl-Seminar 98201.

J. Matoušek (1999), *Geometric Discrepancy*, Springer, Berlin.
H. Niederreiter (1992), *Random Number Generation and Quasi-Monte Carlo Methods*, SIAM, Philadelphia.
H. Niederreiter, R. F. Tichy and G. Turnwald (1990), *An inequality for differences of distribution functions*, Arch. Math. (Basel) 54, 166–172.
E. Novak and H. Woźniakowski (1999), *Intractability results for integration and discrepancy*, J. Complexity, to appear.
D. Pollard (1984), *Convergence of Stochastic Processes*, Springer, New York.
J. Riordan (1958), *An Introduction to Combinatorial Analysis*, Wiley, New York.
K. F. Roth (1954), *On irregularities of distribution*, Mathematika 1, 73–79.
W. M. Schmidt (1975), *On irregularities of distribution IX*, Acta Arith. 27, 385–396.
I. H. Sloan and H. Woźniakowski (1998), *When are quasi-Monte Carlo algorithms efficient for high dimensional integrals?* J. Complexity 14, 1–33.
W. Stute (1977), *Convergence rates for the isotropic discrepancy*, Ann. Probab. 5, 707–723.
M. Talagrand (1994), *Sharper bounds for Gaussian and empirical processes*, ibid. 22, 28–76.
J. F. Traub and A. G. Werschulz (1998), *Complexity and Information*, Cambridge Univ. Press.
A. W. van der Vaart and J. A. Wellner (1996), *Weak Convergence and Empirical Processes, with Applications to Statistics*, Springer, New York.
H. Woźniakowski (1991), *Average case complexity of multivariate integration*, Bull. Amer. Math. Soc. (N.S.) 24, 185–191.
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, Berlin, 114–136.

FB Informatik
Universität Kaiserslautern
Postfach 3049
D-67653 Kaiserslautern, Germany
E-mail: heinrich@informatik.uni-kl.de

Mathematisches Institut
Universität Erlangen-Nürnberg
Bismarckstr. 1 1/2
D-91054 Erlangen, Germany
E-mail: novak@mi.uni-erlangen.de

Department of Computer Science
University of Kentucky
Lexington, KY 40506, U.S.A.
E-mail: greg@cs.uky.edu

Department of Computer Science
Columbia University
New York 10027, NY, U.S.A.
E-mail: henryk@cs.columbia.edu

Received on 12.10.1999
and in revised form on 4.5.2000 (3701)