The exponent of discrepancy is at least 1.0669

Jiří Matoušek*

Department of Applied Mathematics
Charles University
Malostranské nám. 25, 118 00 Praha 1
Czech Republic

Abstract

Let $P \subset [0,1]^d$ be an n-point set and let $w : P \to [0,\infty)$ be a weight function with $w(P) = \sum_{z \in P} w(z) = 1$. The L_2-discrepancy of the weighted set (P,w) is defined as the L_2-average of $D(x) = \mathrm{vol}(B_x) - w(P \cap B_x)$ over $x \in [0,1]^d$, where $\mathrm{vol}(B_x)$ is the volume of the d-dimensional interval $B_x = \prod_{k=1}^{d}[0,x_k)$. The exponent of discrepancy p^* is defined as the infimum of numbers p such that for all dimensions $d \geq 1$ and all $\varepsilon > 0$ there exists a weighted set of at most $K\varepsilon^{-p}$ points in $[0,1]^d$ with L_2-discrepancy at most ε, where $K = K(p)$ is a suitable number independent of ε and d. Wasilkowski and Woźniakowski [15] proved that $p^* \leq 1.4779$, by combining known bounds for the error of numerical integration and using their relation to L_2-discrepancy. In this note we observe that a careful treatment of a classical lower-bound proof of Roth yields $p^* \geq 1.04882$, and by a slight modification of the proof we get $p^* \geq 1.0669$. Determining p^* exactly seems to be quite a difficult problem.

1 Introduction

Let $P \subset [0,1]^d$ be an n-point set in the d-dimensional unit cube. The *discrepancy* of P quantifies the irregularity of distribution of P in some sense; the more "uniformly" distributed P is, the smaller discrepancy. Discrepancy is a classical object of study (see e.g., [2], [1], or [4] for history and background), and it is closely related to error estimates in numerical integration (see e.g., [10]). There are various notions of discrepancy; here we consider one of them, the so-called L_2-*discrepancy* (or, more verbosely, the L_2-discrepancy for anchored boxes). Among various definitions of discrepancy, the L_2-discrepancy seems to be particularly advantageous to deal with: For d fixed, the smallest possible L_2-discrepancy of an n-point set in $[0,1]^d$ is known up to a multiplicative constant depending on d, there are reasonably efficient algorithms for computing the L_2-discrepancy of a given set ([14],[6]), and there are results relating the L_2-discrepancy to the error of numerical integration in both the worst and average cases (see [10],[16]).

For a point $x = (x_1, x_2, \ldots, x_d) \in [0,1]^d$, we let B_x denote the Cartesian product $[0,x_1) \times [0,x_2) \times \ldots \times [0,x_d)$, and for an n-point set $P \subset [0,1]^d$ we put

$$D(x) = \mathrm{vol}(B_x) - \tfrac{1}{n}\left|P \cap B_x\right| = x_1 x_2 \cdots x_d - \tfrac{1}{n}\left|P \cap B_x\right| .$$

*Research supported by Czech Republic Grant GAČR 0194 and by Charles University grants No. 193,194.

The L_2-discrepancy of P, denoted by $D_2(P)$, equals $\|D\|_2 = \left(\int_{[0,1]^d} D(x)^2\,\mathrm{d}x\right)^{1/2}$. Further let us put

$$D_2(d,n) = \inf_{P\subset[0,1]^d:\,|P|=n} D_2(P)\,.$$

Analogously, we define $\bar{D}_2(d,n)$ as the smallest possible L_2-discrepancy of a weighted n-point set in $[0,1]^d$ (as was defined in the abstract).

As was mentioned above, the asymptotics of $D_2(d,n)$ is known for d fixed (and it is the same for $\bar{D}_2(d,n)$). But in many applications, such as high-dimensional numerical integration, the dimension is quite large (say 20, 50, or even few hundred in practice), while the number of points n cannot be made arbitrarily large, and so the dependence on d is of key importance. Here the knowledge is much less satisfactory. A natural question is, given d and a parameter $\varepsilon > 0$, what is the smallest n guaranteeing $\bar{D}_2(d,n) \le \varepsilon$. In order to quantify the answer by a single numerical parameter, Wasilkowski and Woźniakowski [15] defined the *exponent of discrepancy* as the smallest number p for which $\bar{D}_2(d, K\varepsilon^{-p}) \le \varepsilon$ holds for all $d \ge 1$ and $\varepsilon > 0$, with some constant $K = K(p)$. They proved the upper bound $p^* \le 1.4778\ldots$. One interesting question is whether their construction can be matched by point sets with all weights equal to $\frac{1}{n}$. Another one is of course how to improve the upper bound; there are several constructions of point sets with asymptotically optimal L_2-discrepancy for a fixed dimension which might be analyzed from this point of view ([12], [5], [3], [13], [7]).

No lower bound for p^* besides the trivial $p^* \ge 1$ seems to have been known. Here we show the lower bound stated in the title. Let us remark that the question of determining p^* exactly or giving at least a reasonably precise estimates appears fairly difficult (on the other hand, improving the current bounds somewhat may be much less challenging). One reason is perhaps that the definition of p^* is a bit unnatural, since one tries to approximate by a power function something which, at least for d fixed, has powers of logarithm in the asymptotics. Also, as was observed in [9], if n is not sufficiently large compared to d, the L_2-discrepancy of a point set need not be very much related to the "uniform distribution" of the set in the intuitive sense.

2 A lower bound directly from Roth's proof

The paper Roth [11] is a mathematical gem proving the lower bound $D_2(2,n) \ge \frac{c_2}{n}\sqrt{\ln n}$ for some $c_2 > 0$. A straightforward generalization of his proof to dimension d yields the asymptotically tight bound $D_2(d,n) \ge \frac{c_d}{n}(\ln n)^{(d-1)/2}$. This proof is presented, among others, in the books [8] and [1]. Kuipers and Niederreiter [8] give an explicit estimate for the value of c_d, namely $c_d \ge 2^{-4d}(\log_2 e/(d-1))^{(d-1)/2}$ (while Beck and Chen [1] do not write out an explicit bound). This already leads to a nontrivial lower bound $p^* \ge 1.00103$ (by a calculation analogous to the one indicated below and by noting that the same lower bound proof works for weighted point sets as well). Moreover, if one looks into Roth's proof, it is not difficult to find out that it actually yields the lower bound

$$\bar{D}_2(d,n) \ge \frac{c}{n} 2^{-2d} \binom{\lfloor \log_2 n \rfloor + d - 3}{d-1}^{1/2} \tag{1}$$

with an absolute constant $c > 0$. Later we will indicate how is this bound proved, but let us first see how it leads to a lower bound for the exponent of discrepancy. For a number $p > p^*$, we have $D_2(d, K\varepsilon^{-p}) \le \varepsilon$ for all d and $\varepsilon > 0$, and writing $n = K\varepsilon^{-p}$, we calculate

$$p \ge \frac{\ln(n/K)}{-\ln \bar{D}_2(d,n)}\,.$$

Further we set $n = \lfloor 2^{\alpha d} \rfloor$ for a constant $\alpha > 0$ (to be determined later) and let $d \to \infty$, getting

$$p^* \ge \limsup_{d\to\infty} \frac{\alpha d}{-\log_2 \bar{D}_2(d, \lfloor 2^{\alpha d} \rfloor)}\,. \tag{2}$$

By substituting for $\bar{D}_2(d,n)$ from (1) and using the well-known estimates of the factorial function $(k/e)^k \le k! \le k^{k+1}/e^{k-1}$ (a weak from of Stirling's formula, easily proved by induction) to estimate the binomial coefficient, we calculate

$$-\log_2 \bar{D}_2(d, \lfloor 2^{\alpha d} \rfloor) \ge (d + o(d)) \left(2 + \alpha - \tfrac{1}{2}\left[(\alpha+1)\log_2(\alpha+1) - \alpha \log_2 \alpha\right]\right)$$

where the $o(.)$ notation refers to $d \to \infty$. Consequently, from (2) we obtain

$$p^* \ge \sup_{\alpha > 0} \frac{\alpha}{2 + \alpha - \frac{1}{2}[(\alpha+1)\log_2(\alpha+1) + \alpha \log_2 \alpha]}\,,$$

and some simple calculus shows that the optimum value is attained for $\alpha = 15$, leading to the bound $p^* \ge 1.04882\ldots$.

For completeness and for later use, let us recall how (1) is proved; to this end, we need several definitions. For an integer $r \ge 0$, the Rademacher function $R_r(x) : \mathbf{R} \to \{-1, 1\}$ is given by $R_r(x) = (-1)^{\lfloor 2^{r+1} x \rfloor}$, and for a nonnegative integer vector $\mathbf{r} = (r_1, r_2, \ldots, r_d)$ and for $x \in \mathbf{R}^d$, put $R_{\mathbf{r}}(x) = \prod_{k=1}^d R_{r_k}(x_k)$. By an $\mathbf{r}$-box we mean a d-dimensional interval of the form $B = \prod_{k=1}^d [a_k 2^{-r_k}, (a_k+1)2^{-r_k})$, where each a_k is an integer between 0 and $2^{r_k} - 1$. An $\mathbf{r}$-function is any function $f : [0,1)^d \to \{-1,1\}$ such that for any $\mathbf{r}$-box B, f equals either $R_{\mathbf{r}}$ or $-R_{\mathbf{r}}$ on B.

Let m be the smallest integer with $2^m \ge 2n$. The key step in the proof of (1) is the following lemma, whose proof we omit at this moment (see [1], Lemma 2.5): *For any n-point set $P \subset [0,1]^d$ and any nonnegative integer vector $\mathbf{r}$ with $|\mathbf{r}| = r_1 + r_2 + \cdots + r_k = m$, there exists an $\mathbf{r}$-function $f_{\mathbf{r}}$ with $\int f_{\mathbf{r}} \cdot D \ge 2^{-2d-4}/n$* (here and in the sequel, we omit the integration domain $[0,1]^d$ and the integration variable x). We fix such an $f_{\mathbf{r}}$ for each $\mathbf{r}$, and we let M denote the set of all nonnegative $\mathbf{r}$ with $|\mathbf{r}| = m$; note that $|M| = \binom{m+d-1}{d-1}$. Putting $F = \sum_{\mathbf{r} \in M} f_{\mathbf{r}}$ and applying the Cauchy-Schwarz inequality, we get

$$D_2(P) = \|D\|_2 \ge \frac{\int FD}{\|F\|_2}\,.$$

We have $\int FD \ge \frac{2^{-2d-4}}{n}|M|$ by the choice of the $f_{\mathbf{r}}$'s, and $\|F\|_2^2 = \sum_{\mathbf{r}\in M} \|f_{\mathbf{r}}\|_2^2 = |M|$ since $\int f_{\mathbf{r}} f_{\mathbf{s}} = 0$ for any $\mathbf{r}$-function $f_{\mathbf{r}}$ and any $\mathbf{s}$-function $f_{\mathbf{s}}$ with $\mathbf{r} \ne \mathbf{s}$. Consequently $D(P) \ge \frac{2^{-2d-4}}{n}|M|^{1/2}$ and (1) follows.

3 An improved lower bound

It turns out that a better lower bound for p^* is obtained if one considers "lower-dimensional" $\mathbf{r}$-functions in Roth's proof (although for d fixed and $n \to \infty$, these would only give an asymptotically suboptimal bound). Let $\mathbf{r}$ be a vector whose each component is either a nonnegative integer or $*$; write $\mathrm{supp}(\mathbf{r}) = \{k \in \{1, 2, \dots, d\} : r_k \neq *\}$ and $|\mathbf{r}| = \sum_{k \in \mathrm{supp}(\mathbf{r})} r_k$. Let $R_{\mathbf{r}}(x) = \prod_{k \in \mathrm{supp}(\mathbf{r})} R_{r_k}(x)$, define an $\mathbf{r}$-box as $\prod_{k=1}^d [a_k 2^{-r_k}, (a_k + 1)2^{-r_k})$ where we set $2^{-*} = 1$ (so $a_k = 0$ for $r_k = *$), and let an $\mathbf{r}$-function be defined as one equal to either $R_{\mathbf{r}}$ or $-R_{\mathbf{r}}$ on each $\mathbf{r}$-box.

Let P and m be as in the outline of Roth's proof in the previous section, and for a parameter $q \in \{1, 2, \dots, d\}$, let M_q denote the set of all $\mathbf{r}$ with $|\mathbf{r}| = m$ and with $|\mathrm{supp}(\mathbf{r})| = q$. This time we have $|M_q| = \binom{d}{q}\binom{m+q-1}{q-1}$. The analogue of the key lemma in the preceding proof is

Lemma. *For any $\mathbf{r} \in M_q$, there exists an $\mathbf{r}$-function $f_{\mathbf{r}}$ with $\int f_{\mathbf{r}} D \geq 2^{-d-q-4}/n$.*

Proof of the lemma (sketch). Let B be an $\mathbf{r}$-box. If $P \cap B \neq \emptyset$ we choose the sign of $f_{\mathbf{r}}$ on B in such a way that $\int_B f_{\mathbf{r}} D \geq 0$. If $P \cap B = \emptyset$, which happens for at least $2^m - n \geq 2^{m-1}$ $\mathbf{r}$-boxes B, we put $f_{\mathbf{r}} = R_{\mathbf{r}}$ on B. Similarly as in Roth's proof, we have

$$\int_B R_{\mathbf{r}}(x) D(x)\,\mathrm{d}x = \int_B R_{\mathbf{r}}(x) x_1 x_2 \cdots x_d\,\mathrm{d}x - \frac{1}{n}\int_B R_{\mathbf{r}}(x) |P \cap B_x|\,\mathrm{d}x\,.$$

The second integral can be shown to be 0 (using $P \cap B = \emptyset$). The first integral equals

$$\int_C \left(\mathrm{vol}(C) \prod_{k \notin \mathrm{supp}(\mathbf{r})} x_k\right) \mathrm{d}x\,, \tag{3}$$

where C denotes the "lower left part" of B: namely, for $B = \prod_{k=1}^d [a_k 2^{-r_k}, (a_k + 1)2^{-r_k})$, we have $C = \prod_{k=1}^d [a_k 2^{-r_k}, (a_k + \delta_k)2^{-r_k})$, with $\delta_k = \frac{1}{2}$ for $k \in \mathrm{supp}(\mathbf{r})$ and $\delta_k = 1$ otherwise. We have $\mathrm{vol}(C) = 2^{-m-q}$, and (3) equals $\mathrm{vol}(C)^2 2^{-(d-q)} = 2^{-2m-d-q}$. The lemma follows by summing over all the at least 2^{m-1} boxes B with $P \cap B = \emptyset$. $\square$

We now set $F = \sum_{\mathbf{r} \in M_q} f_{\mathbf{r}}$, with $f_{\mathbf{r}}$ as in the lemma. For two distinct vectors $\mathbf{r}, \mathbf{s} \in M_q$, we still get the orthogonality $\int f_{\mathbf{r}} f_{\mathbf{s}} = 0$, and so Cauchy-Schwarz gives

$$\bar{D}_2(P) \geq \frac{\int FD}{\|F\|_2} \geq \frac{2^{-d-q-4}}{n} |M_q|^{1/2} = \frac{2^{-d-q-4}}{n} \binom{d}{q}^{1/2} \binom{m+q-1}{q-1}^{1/2}.$$

To lower-bound p^*, we again set $n = \lfloor 2^{\alpha d} \rfloor$ for some constant α and use (2). Moreover, we put $q = \lfloor \beta d \rfloor$ for another suitable constant $\beta \in (0, 1]$. With the same estimates of the binomial coefficients and with some calculus and numerical calculation, one can find that the best bound for p^* is obtained for $\beta \approx 0.7473$ and $\alpha \approx 8.2205$, yielding $p^* \geq 1.0669\ldots$.

References

[1] J. Beck and W. Chen. *Irregularities of Distribution.* Cambridge University Press, 1987.

[2] J. Beck and V. Sós. Discrepancy theory. In *Handbook of Combinatorics*, pages 1405–1446. North-Holland, Amsterdam, 1995.

[3] W. L. Chen. On irregularities of distribution II. *Quart. J. Math. Oxford Ser. (2)*, 34:257–279, 1983.

[4] M. Drmota and R. F. Tichy. *Sequences, discrepancies and applications (Lecture Notes in Mathematics 1651).* Springer-Verlag, 1997.

[5] K. K. Frolov. An upper estimaste for the discrepancy in the L_p-metric, $2 \leq p \leq \infty$ (in Russian). *Dokl. Akad. Nauk SSSR*, 252:805–807, 1980. English translation in *Soviet. Math. Dokl.* 21(1980) 840–842.

[6] S. Heinrich. Efficient algorithms for computing the L_2 discrepancy. *Math. Comput.*, 65:1621–1633, 1996.

[7] F. J. Hickernell. The mean square discrepancy of randomized nets. *ACM Transactions on Modeling and Computer Simulation*, 6(4):274–296, 1996.

[8] L. Kuipers and H. Niederreiter. *Uniform Distribution of Sequences.* J. Wiley & Sons, New York, 1974.

[9] J. Matoušek. On the L_2-discrepancy for anchored boxes. Manuscript, 1997.

[10] H. Niederreiter. *Random number generation and quasi-Monte Carlo methods.* CBMS-NSF Regional Conference Series in Applied Mathematics. 63. Philadelphia, PA: SIAM, 1992.

[11] K. F. Roth. On irregularities of distribution. *Mathematika*, 1:73–79, 1954.

[12] K. F. Roth. On irregularities of distribution IV. *Acta Arith.*, 37:67–75, 1980.

[13] M. M. Skriganov. Constructions of uniform distributions in terms of geometry of numbers. *Algebra i Analiz*, 6:200–23, 1994. Also in *St. Petersburg Math. J.* 6:635–664,1995.

[14] T. T. Warnock. Computational investigations of low-discrepancy point sets. In *Applications of number theory to numerical analysis (S. K. Zaremba ed.)*, pages 319–343. Academic Press, New York, 1972.

[15] G. Wasilkowski and H. Woźniakowski. The exponent of discrepancy is at most 1.4778.... *Math. Comput*, 66:1125–1132, 1997.

[16] H. Woźniakowski. Average case complexity of multivariate integration. *Bull. Amer. Math. Soc.*, 24:185–194, 1991.