The exponent of discrepancy is at most 1.4778...

Columbia University Computer Science Department Report CUCS-026-96

Grzegorz W. Wasilkowski and Henryk Woźniakowski *

December 1995

Abstract

We study discrepancy with arbitrary weights in the L_2 norm over the d dimensional unit cube. The exponent p^* of discrepancy is defined as the smallest p for which there exists a positive number K such that for all d and all $\varepsilon \leq 1$ there exist $K\,\varepsilon^{-p}$ points with discrepancy at most ε. It is well known that $p^* \in (1,2]$. We improve the upper bound by showing that

$$p^* \leq 1.478841.....$$

This is done by using relations between discrepancy and integration in the average case setting with the Wiener sheet measure. Our proof is *not* constructive. The known constructive bound on the exponent p^* is 2.454.

1 Introduction

We study discrepancy with arbitrary weights in the L_2 norm over the d dimensional unit cube $[0,1]^d$. This problem is defined as finding n points from $[0,1]^d$ which approximate the volumes of rectangles (starting from zero) with minimal error, see [8, 9] for the precise definition, history and basic properties. Discrepancy has been extensively studied in number theory and numerical analysis, see e.g., [1, 8, 9, 10, 11, 12]. Fast algorithms for computing

* *1991 Math. Subject Classification:* Primary 11K38, 41A55.

Key words: discrepancy, multivariate integration, average case.

The first author was partially supported by the the National Science Foundation under Grant CCR-9420543, and the second by the National Science Foundation and the Air Force Office of Scientific Research.

discrepancy of n given points are given in [6]. Discrepancy has been recently applied in computer science, see [3, 4, 7] and the references given there.

Discrepancy is related to multivariate integration in the worst case and average case settings. Indeed, discrepancy is an upper bound on the worst case integration error of functions whose variation in the sense of Hardy and Krause is at most one, see e.g., [2, 8, 9, 13]. It is also known, see [15], that discrepancy for points $x_1, x_2, \ldots, x_n$ is equal to the average case integration error for the points $\vec{1} - x_1, \vec{1} - x_2, \ldots, \vec{1} - x_n$, where $\vec{1} = [1, 1, \ldots, 1] \in \mathbb{R}^d$. The average error is defined for the class of continuous functions defined over $[0,1]^d$ and equipped with the Wiener sheet measure. Hence, bounds on discrepancy have immediate applications for multivariate integration.

Let $n(\varepsilon, d)$ be the minimal number of points from $[0,1]^d$ for which discrepancy is at most ε. It is well known that

$$n(\varepsilon, 1) = \left\lceil \left(1/\left(\varepsilon\sqrt{3}\right) - 1\right)/2 \right\rceil, \tag{1}$$

$$n(\varepsilon, d) = 0 \quad \text{for } 3^{-d/2} \le \varepsilon, \tag{2}$$

$$n(\varepsilon, d) \le 2^{-d}\,\varepsilon^{-2}. \tag{3}$$

The proof of the last inequality is not *constructive*, i.e., the construction of points satisfying (3) is unknown. The asymptotic behavior of $n(\varepsilon, d)$ for a fixed d and ε tending to zero is known,

$$n(\varepsilon, d) = \Theta\left(\varepsilon^{-1}(\log 1/\varepsilon)^{(d-1)/2}\right), \tag{4}$$

see [10] for the proof of a lower bound, and [5, 11] for the proof of an upper bound. However, the multiplicative factor in the Θ-notation is an unknown function of d. A constructive bound on $n(\varepsilon, d)$ with explicit dependence on d is given in Section 8.3 of [14],

$$n(\varepsilon, d) \le 3.304 \left(1.77959 + 2.714\,\frac{-1.12167 + \ln 1/\varepsilon}{d-1}\right)^{1.5(d-1)} \frac{1}{\varepsilon}, \tag{5}$$

and this is achieved by hyperbolic cross points.

We define the exponent p^* of discrepancy as the smallest (or the infimum of) p for which there exists a positive number K such that

$$n(\varepsilon, d) \le K\,\varepsilon^{-p}, \qquad \forall\, \varepsilon \le 1,\ \forall\, d = 1, 2 \ldots. \tag{6}$$

Obviously, (3) and (4) yield

$$p^* \in (1, 2]. \tag{7}$$

Since the bound in (3) is not constructive, so is the bound $p^* \leq 2$. From (2) and (5) it follows, see Section 8.3 of [14], that

$$n(\varepsilon, d) \leq 7.26\, \varepsilon^{-2.454}.$$

Hence, the known constructive bound[1] on p^* is

$$p^* \leq 2.454. \tag{8}$$

In this paper, we improve the upper bound by showing the following estimate.

Theorem 1 *The exponent p^* of discrepancy is bounded by*

$$p^* \leq q^* \approx 1.47788417318605565480, \tag{9}$$

where

$$q^* = \max_{x \geq 4/3} \frac{\ln\left(4^{x-1} x^x / (x-1)^{x-1}\right)}{\ln\left(4^x/\sqrt{6}\right)}.$$

We explain the idea behind the proof of (9); the complete proof is in Section 3. First, we switch to the equivalent problem of integration in the average case setting with the Wiener sheet measure. Let $I(f)$ denote the integral of a continuous function f defined over $[0,1]^d$. Then, instead of integrating f directly, we approximate f by a special function f_n whose computation requires n values of f. Obviously, $I(f) = I(f_n) + I(f - f_n)$. The next step is to apply the classical Monte Carlo algorithm to approximate $I(f - f_n)$ using function values at $n+1$ random points. The use of the Monte Carlo to the function $f - f_n$ may be regarded as variance reduction. The randomized error of the Monte Carlo is obviously bounded by the variance of $f - f_n$ divided by $\sqrt{n}$. We now take the expectation with respect to f. The expected variance of $f - f_n$ is equal to the average case error of approximation between f and f_n. The explicit formula for this is derived in Section 8.4 of [14]. Finally, we use a well-known fact that randomization does not help in the average case setting to conclude the existence of $2n+1$ points with the needed bound on discrepancy.

Since we use the Monte Carlo algorithm in one step, the proof of (9) is *not* constructive. Hence, the constructive bound (8) remains unimproved.

[1] In Section 8.2 of [14] we mentioned that the exponent $p = 2.454$ can be lowered by choosing different parameters of the basic algorithm. We tried there a number of such parameters but we always obtained exponents greater than two.

We regard the problem of finding the exact value of the exponent p^* as a very challenging and difficult one. Further improvements of the bounds (8) and (9) do not seem to be easy. From a practical point of view, the most challenging problem is to find n points such that $n \le K\varepsilon^{-p}$ with the exponent p less than two. Due to (9), such points exist.

We finally remark that the proof of (9) can be applied for integration of different classes of functions in the average case setting. For example, applying this proof for the class of periodic continuous functions equipped with the Brownian bridge, we obtain that the exponent of this integration problem is bounded from above by 1.29.

2 Discrepancy

In this section we recall the definition of discrepancy in the L_2 norm and briefly review bounds on discrepancy. A thorough discussion on discrepancy may be found in [8, 9].

Consider n points $z_1, z_2, \ldots, z_n$ from $[0,1]^d$. For a vector $t = [t_1, t_2, \ldots, t_d] \in [0,1]^d$, define the rectangle $[0,t) = [0,t_1) \times [0,t_2) \times \cdots \times [0,t_d)$. We wish to approximate the volume of the rectangle $[0,t)$, which is obviously $t_1 t_2 \cdots t_d$, by a weighted number of points z_i which lie in $[0,t)$. The error of such an approximation is called *discrepancy*. More precisely, let

$$\mathrm{DISC}_{n,d}(t; \{z_i\}, \{c_i\}) \;=\; t_1 t_2 \cdots t_d \;-\; \sum_{i=1}^{n} c_i \chi_{[0,t)}(z_i)$$

for some weights $c_i \in \mathbb{R}$. Here $\chi_{[0,t)}$ is the characteristic (indicator) function of $[0,t)$;

$$\chi_{[0,t)}(z) \;=\; 1 \;\text{ if } z \in [0,t), \quad \text{and} \quad \chi_{[0,t)}(z) \;=\; 0 \;\text{ if } z \notin [0,t).$$

The most popular choice of the weights c_i is to take $c_i = 1/n$.

The discrepancy of the n points z_i and weights c_i is given by

$$\|\mathrm{DISC}_{n,d}(\cdot\,; \{z_i\}, \{c_i\})\| \;=\; \left(\int_{[0,1]^d} \mathrm{DISC}_{n,d}^2(t; \{z_i\}, \{c_i\})\, dt \right)^{1/2}.$$

The discrepancy problem is to find the points $\{z_i\}$ and the weights $\{c_i\}$ that minimize the discrepancy,

$$\mathrm{DISC}_{n,d} \;=\; \inf \Big\{ \|\mathrm{DISC}_n(\cdot\,; \{z_i\}, \{c_i\})\| \;:\; z_i \in [0,1]^d,\, c_i \in \mathbb{R},\; i = 1, 2, \ldots, n \Big\}. \tag{10}$$

Let $n(\varepsilon, d)$ be the minimal number of points for which the discrepancy is at most ε,

$$n(\varepsilon, d) \;=\; \min \{ n \;:\; \mathrm{DISC}_{n,d} \;\le\; \varepsilon \}. \tag{11}$$

The exponent p^* of the discrepancy is defined as

$$p^* \;=\; \inf\Big\{p \;:\; \exists\, K \geq 0,\; \forall\, \varepsilon \leq 1,\; \forall\, d,\; n(\varepsilon, d) \;\leq\; K\, \varepsilon^{-p}\Big\}. \tag{12}$$

We now briefly indicate how the bounds (1), (2) and (3) are obtained. For $d = 1$, it is relatively easy to solve (10) and to get (1). Taking $n = 0$, we get $\mathrm{DISC}_{0,d} = 3^{-d/2}$ which yields (2). To show (3) it is enough to set $c_i = 1/n$ and integrate the function $\|\mathrm{DISC}_{n,d}(\cdot; \{z_i\}, \{1/n\})\|^2$ with respect to z_i. Then this integral is $(2^{-d} - 3^{-d})/n$. Applying the mean value theorem we get (3).

We will use a relation between discrepancy and integration in the average case setting, see [15]. This relation is given by the identity

$$\|\mathrm{DISC}_{n,d}(\cdot\,; \{z_i\}, \{c_i\})\|^2 \;=\; \int_F \left(\int_{[0,1]^d} f(t)\, dt \;-\; \sum_{i=1}^{n} c_i f(x_i) \right)^2 w(df), \tag{13}$$

where $F = C([0,1]^d)$ is the class of continuous functions defined over $[0,1]^d$, w is the classical Wiener sheet measure, and $x_i = \vec{1} - z_i$.

3 Proof

In this section we prove the bound (9) on the exponent p^*. Due to (1) and (2), we need only to consider $d \geq 2$ and $\varepsilon < 3^{-d/2}$. The proof is heavily based on a number of results from [14].

The first step of the proof is to consider the average case error e_n of a special algorithm $A_n(\cdot)$ for approximating $f \in F = C([0,1]^d)$;

$$e_n^2 \;=\; \int_F \left(\int_{[0,1]^d} (f(t) \;-\; A_n(f)(t))^2 \; dt \right) w(df).$$

The approximating function $f_n = A_n(f)$ is constructed as in Section 8.4 and Lemma 8 of [14] with the parameters $F_0 = 1/3$, $F = 4$, $D = 1/2$, and $C = B = 1/\sqrt{2}$. It is a linear combination of

$$n \;=\; \lfloor g_c^{d-1}(x)\, b(d) \rfloor$$

values of f, and its error satisfies

$$e_n \;\leq\; g_e^{d-1}(x)\, a(d),$$

where

$$g_e(x) \;=\; \left(\frac{1}{2}\right)^x \sqrt{\frac{3\, x^x}{2(x-1)^{x-1}}}, \qquad g_c(x) \;=\; \frac{4^{x-1}\, x^x}{(x-1)^{x-1}}$$

and

$$a(d) \;=\; \frac{1}{\sqrt{2}} \max\left\{ \left(\frac{d^2}{8\,\pi\,(d-1)} \right)^{1/4}, 1 \right\}, \qquad b(d) \;=\; \frac{128}{3\sqrt{2\pi(d-1)}}.$$

Here x is a parameter from $[4/3, \infty)$, to be chosen later.

We proceed to the next step of the proof. We wish to approximate

$$I(f) = \int_{[0,1]^d} f(t)\, dt.$$

Let

$$m \;=\; \lceil g_c^{d-1}(x)\, b(d) \rceil$$

and consider the following randomized (Monte Carlo) quadrature Q_m,

$$Q_m(f, \{t_i\}) = I(f_n) + \frac{1}{m} \sum_{i=1}^{m} (f(t_i) - f_n(t_i))$$

with independently and uniformly distributed points $t_1, \ldots t_m \in [0,1]^d$. Obviously, the quadrature Q_m uses $n + m \leq 2n + 1$ function values.

It is well known that the expected error with respect to t_i of this quadrature is bounded as

$$\int_{[0,1]^{dm}} (I(f) \;-\; Q_m(f, \{t_i\}))^2 \; dt_1\, dt_2 \cdots dt_m \;\leq\; \frac{1}{m} \int_{[0,1]^d} (f(t) \;-\; f_n(t))^2 \; dt.$$

We now integrate this with respect to f's that are distributed according to the Wiener sheet measure w. Let

$$\mathrm{E}_m^2 \;=\; \int_F \int_{[0,1]^{dm}} (I(f) \;-\; Q_m(f, \{t_i\}))^2 \; dt_1\, dt_2 \cdots dt_m\, w(df).$$

Then

$$\mathrm{E}_m \;\leq\; e_n/\sqrt{m}. \tag{14}$$

It is known that randomization does not help in the average case setting. This simply follows from the mean value theorem since there exist points t_i^*, $i = 1, 2 \ldots, m$, such that

$$\int_F (I(f) \;-\; Q_m(f, \{t_i^*\}))^2 \; w(df) \;\leq\; \mathrm{E}_m^2.$$

From (13) and (14) we thus conclude

$$\mathrm{DISC}_{2n+1,d} \;\leq\; e_n/\sqrt{m}.$$

From the definition of n and the bound on e_n, we get

$$\mathrm{DISC}_{2n+1,d} \;\le\; c(d)\, h(x)^{d-1}$$

with

$$c(d) \;=\; \frac{a(d)}{\sqrt{b(d)}} \quad \text{and} \quad h(x) \;=\; \frac{\sqrt{6}}{4^x}.$$

To guarantee that $\mathrm{DISC}_{2n+1,d} \le \varepsilon$, we define x as the smallest $x \in [4/3, \infty)$ for which

$$c(d)\, h(x)^{d-1} \;\le\; \varepsilon.$$

Note that such an x exists. Let

$$q^* \;=\; \max_{x \ge 4/3} \frac{\ln(g_c(x))}{\ln(1/h(x))}.$$

Then

$$(g_c(x))^{d-1} \;\le\; (1/h(x))^{q^*(d-1)} \;\le\; (c(d)/\varepsilon)^{q^*},$$

and, therefore,

$$2n \;\le\; 2\, b(d) \left(\frac{c(d)}{\varepsilon}\right)^{q^*}. \tag{15}$$

Observe that $b(d)\, c^{q^*}(d) \;=\; \Theta\left(d^{(q^*-1)/2}\right)$ and, as we shall see, q^* is greater than 1. Since $\varepsilon < 3^{-d/2}$, we have that $b(d)\, c^{q^*}(d)\, \varepsilon^{\eta}$ is uniformly bounded for any positive η. Therefore, from (15) we get that, for any positive η,

$$n(\varepsilon, d) \;\le\; 2n+1 \;\le\; K_\eta\, \varepsilon^{-q^*-\eta} \qquad \text{with} \qquad K_\eta = \sup_{d \ge 2} \left(2\, b(d)\, c^{q^*}(d)\, 3^{-d\eta/2} \;+\; 3^{-(q^*+\eta)d/2}\right).$$

This proves that $p^* \le q^*$, as claimed.

To complete the proof of the theorem, we need to estimate q^*. By differentiating the function

$$\rho(x) \;=\; \frac{\ln g_c(x)}{\ln 1/h(x)} \;=\; \frac{(x-1)\ln(4) \;+\; x\,\ln(x) \;-\; (x-1)\ln(x-1)}{\ln(\sqrt{1/6}) \;+\; x\,\ln(4)}$$

we get

$$\rho'(x) \;=\; \frac{z(x)}{\left(\ln(4^x/\sqrt{6})\right)^2}$$

with

$$z(x) \;=\; \ln\left(\frac{4\,x}{x-1}\right)\,\ln\left(\sqrt{8/3}\right) \;-\; \ln(4)\,\ln(x)\,.$$

Hence, to find the extreme points of ρ we need to find the roots of z. To this end, observe that

$$z'(x) = -(x(x-1))^{-1}\,\ln\left(\sqrt{8/3}\right) - x^{-1}\,\ln(4),$$

so that z' is negative. Since $z(2.0) \approx 0.0588...$ and $z(2.5) \approx -0.3398...$, the unique root x^* is in $(2, 2.5)$. Since $z''(x)$ is positive, the root can efficiently be approximated by Newton's method starting with $x_0 = 2$. After a few Newton's steps we received

$$x^* \;\approx\; \tilde{x} \;=\; 2.06426868257207727735.$$

The value of z at that point is less that 10^{-20}, and $\tilde{x}$ equals x^* to within at least 18 most significant digits.

Finally, by computing $\rho(\tilde{x})$, we get

$$q^* \;=\; \rho(x^*) \;\approx\; \rho(\tilde{x}) \;\approx\; 1.47788417318605565480,$$

which is correct to at least 18 most significant digits. This completes the proof.

We end the paper by the following remarks.

Remark 1 The proof of (9) also yields the following estimate

$$n(\varepsilon, d) \;\leq\; 2.51\,d^{\,0.24}\,\varepsilon^{-1.4778842}\,(1+o(1))\,, \quad \text{as } d \;\to\; \infty.$$

This follows easily from the formulas for $a(d)$, $b(d)$ and $c(d)$.

Remark 2. One could try to improve the bound q^* on the exponent p^* by using significantly different values of n and m in the proof of the theorem. We have checked this approach and, unfortunately, it does not lead to any improvement; $n \approx m$ is optimal.

Lemma 8 of [14] can be also used for different values of the parameters F_0, F, D, C and B. We have checked that the choice reported in the proof leads to the smallest q^*.

Remark 3. As mentioned in the introduction, we have also analyzed the exponent for the integration problem in the average case setting with respect to the Brownian bridge.

For this problem, the results of [14] apply with parameters $C = B = 1/\sqrt{6}$, $F_0 = 1$, $F = 2$, and $D = 1/\sqrt{2}$. Hence, the functions g_e, g_e, a, b from the beginning of Section 3 become:

$$g_e(x) = \sqrt{\frac{x^{x-1}}{6(x-1)^{x-1}2^x}}, \qquad g_c(x) = 2^{x-1}\frac{x^x}{(x-1)^{x-1}},$$

$$a(d) = \frac{1}{\sqrt{6}} \max\left\{\left(\frac{d^2}{4\pi(d-1)}\right)^{1/4}, 1\right\}, \qquad b(d) = \frac{16}{\sqrt{\pi(d-1)}}.$$

Thus, repeating the proof of Section 3, we find

$$q^* \approx 1.28898137370363213616.$$

Acknowledgment

We thank A. G. Werschulz for useful comments on this paper.

References

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

[2] V. A. Bykovskij, On exact order of optimal quadrature formulas for spaces of functions with bounded mixed derivatives, *Report of Dalnevostochnoi Center of Academy of Sciences*, Vladivostok, USSR, (in Russian), 1985.

[3] B. Chazelle, Geometric Discrepancy Revisited, *34th Annual Symposium on Foundations of Computer Science*, p. 392-399, 1993.

[4] D. P. Dobkin and D. P. Mitchell, Random-Edge Discrepancy of Supersampling Patterns, *Graphics Interface '93*, York, Ontario, 1993.

[5] K. K. Frolov, Upper bounds on discrepancy in L_p, $2 \le p \le \infty$, *Dokl. Akad. Nauk SSSR*, 252, p. 805-807, 1980.

[6] S. Heinrich, Efficient algorithms for computing the L_2 discrepancy, *Math. of Computation*, 1995, to appear.

[7] S. Heinrich and A. Keller, Quasi-Monte Carlo methods in computer graphics, Part I: The QMC-Buffer, Report, Univ. of Kaiserslautern, Dept. Computer Science, Report No.242/1994; Part II: The radiance equation, Report, Univ. of Kaiserslautern, Dept. Computer Science, Report No.243/1994.

[8] H. Niederreiter, Quasi-Monte Carlo methods and pseudo-random numbers, *Bull, Amer. Math. Soc.*, 84, p. 957-1041, 1978.

[9] H. Niederreiter, Random Number Generation and Quasi-Monte Carlo Methods, *CBMS-NSF Reg. Conf. Series Appl. Math.*, 63, *SIAM*, Philadelphia, 1992.

[10] K. F. Roth, On irregularities of distribution, *Mathematika*, 1, p. 73-79, 1954.

[11] K. F. Roth, On irregularities of distribution, IV, *Acta Arith.*, 37, p. 67-75, 1980.

[12] I. H. Sloan and S. Joe Lattice methods for multiple integration, *Oxford Science Publication*, Oxford, 1994.

[13] V. N. Temlyakov, Approximation of Functions with Bounded Mixed Derivatives, *Proc. Steklov Inst. Math*, Moscow, 1989.

[14] G. W. Wasilkowski and H. Woźniakowski, Explicit cost bounds of algorithms for multivariate tensor product problems, *J. Complexity*, 11, p. 1-56, 1995

[15] H. Woźniakowski, Average Case Complexity of Multivariate Integration, *Bull. Amer. Math. Soc. (N.S)*, 24, p. 185-194, 1991.

Author's Addresses:

Grzegorz W. Wasilkowski, Department of Computer Science, University of Kentucky, Lexington, KY 40506, USA, email: greg@cs.engr.uky.edu

Henryk Woźniakowski, Department of Computer Science, Columbia University, New York, NY 10027, USA, and Institute of Applied Mathematics, University of Warsaw, ul. Banacha 2, 02-097 Warszawa, Poland, email: henryk@cs.columbia.edu