PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 80, Number 1, September 1980

SHORTEST PATHS THROUGH PSEUDO-RANDOM POINTS IN THE d-CUBE

J. MICHAEL STEELE[1]

ABSTRACT. A lower bound for the length of the shortest path through n points in $[0, 1]^d$ is given in terms of the discrepancy function of the n points. This bound is applied to obtain an analogue for several pseudorandom sequences to the known limit behavior of the length of the shortest path through n independent uniformly distributed random observations from $[0, 1]^d$.

Introduction For a sequence $\{z_i: 1 \leqslant i < \infty\}$ of elements of $[0, 1]^d$, $d \geqslant 1$, to be considered as a candidate for a sequence of pseudorandom observations, it must at least be uniformly distributed in the sense that the empirical frequency of any subcube must tend to the volume of that subcube [5, pp. 127–157]. For such a minimally pseudorandom sequence, it seems of interest to determine the extent to which additional features of independent uniformly distributed random variables must also hold. The additional feature which is considered in this note is the asymptotic growth rate of the shortest path through the initial sample $\{z_1, z_2, \ldots, z_n\}$.

For a sequence $\{x_i, 1 \leqslant i < \infty\}$ of independent random variables uniformly distributed in $[0, 1]^d$, Beardwood, Halton and Hammersley [1] proved that $\lambda(n) = \lambda(x_1, x_2, \ldots, x_n)$, the length of the shortest path through $\{x_1, x_2, \ldots, x_n\}$, has the property that

$$\lim_{n\to\infty} \lambda(n)/n^{(d-1)/d} = k_d, \tag{1}$$

with probability one for a constant $k_d > 0$.

While a result of such precision is too much to expect of a general pseudorandom sequence, a somewhat weaker version of (1) can be obtained in sufficient generality to cover a variety of classical cases. To make this precise, first suppose $[0, 1]^d$ is partially ordered ($\ll$) by the usual coordinatewise ordering, and for each $x \in [0, 1]^d$, define the empirical distribution function

$$F_n(x) = (1/n)^{\#}\{1 \leqslant i \leqslant n: z_i \ll x\}$$

of the sequence $\{z_i: 1 \leqslant i < \infty\}$. The discrepancy function D_n is then defined by

$$D_n = \sup_{x\in[0,1]^d} |F_n(x) - F(x)|,$$

Received by the editors June 29, 1979 and, in revised form, November 8, 1979.
1980 *Mathematics Subject Classification*. Primary 65C10, 90B10, 10F40, 60D05.
Key words and phrases. Discrepancy, pseudorandom, shortest paths, uniform distribution.
[1]This research was supported in part by the National Science Foundation, Grant No. MCS78-07736.

0002-9939/80/0000-0423/$02.25

where $F(x)$ is the volume of the region $\{y \in [0, 1]^d: y \ll x\}$. For any pseudorandom sequence we may suppose that $\lim_{n\to\infty} D_n = 0$. It will be by relating the convergence rate of D_n to the growth rate of $\lambda(n) = \lambda(z_1, z_2, \ldots, z_n)$ that the analogues of (1) can be proved.

Since L. Few [2] has proved that there are constants c_d' such that for any $\{z_i: 1 \leqslant i \leqslant \infty, z_i \in [0, 1]^d\}$ and all $1 \leqslant n < \infty$

$$\lambda(n) = \lambda(z_1, z_2, \ldots, z_n) \leqslant c_d' n^{(d-1)/d},$$

the limit theory for $\lambda(n)$ will depend upon obtaining an appropriate lower bound. The following elementary bound suffices in many cases and is in a sense the best possible.

THEOREM 1. *There are constants $c_d > 0$ such that $\lambda(n) \geqslant c_d D_n^{-(d-1)/d}$ for any uniformly distributed sequence and any sufficiently large n.*

PROOF. Divide $[0, 1]^d$ into k^d cubic cells of edge $1/k$, and further divide each of these into 3^d cells of edge $1/3k$. The $1/3k$-edge cells which do not touch the boundary of the $1/k$-edge cells will be called center cells. We note that if each of the center cells is occupied by at least one element of $z_1, z_2, \ldots, z_n$, then $\lambda(n) \geqslant (2/3k)(k^d - 1)$, since there are k^d center cells, and two points in different center cells must be at least $2/3k$ apart.

Next note that if $D_n \leqslant 2^{-d}(3k)^{-d}$, then each of the $1/3k$-edge cells must be occupied by at least one z_i, $i = 1, 2, \ldots, n$. On choosing k so that $6^{-d}(k+1)^{-d} \leqslant D_n \leqslant 6^{-d}k^{-d}$, we finally have

$$\lambda(n) \geqslant (2/3)k^{d-1} \geqslant D_n^{-(d-1)/d}(k/6(k+1))^{d-1}(2/3),$$

which easily implies the theorem.

The main consequence of Theorem 1 is that together with Few's upper bound one obtains the following analogue of the Beardwood-Halton-Hammersley theorem.

COROLLARY 1. *If $nD_n = O(n^\varepsilon)$ for all $\varepsilon > 0$, then*

$$\lim_{n\to\infty} \log \lambda(n)/\log n = (d-1)/d. \tag{2}$$

There is extensive literature devoted to the determination of the discrepancy of special sequences, and many of these satisfy the hypotheses of Corollary 1. Several such examples are considered below, but for a genuine survey of discrepancy bounds of pseudorandom sequences and their applications in numerical analysis one should consult Niederreiter [11]. In the first example we recall Halton's important generalization of the van der Corput sequence, but to avoid notational digression we omit its explicit description.

EXAMPLE 1 (HALTON [3]). For any $d > 1$ there is an infinite sequence in $[0, 1]^d$ such that $nD_n = O((\log n)^d)$.

The next two examples are of a more theoretical character.

EXAMPLE 2. Let $\{k\alpha\}$ denote $k\alpha$ reduced modulo 1, and set $z_k = (\{k\alpha_1\}, \{k\alpha_2\}, \ldots, \{k\alpha_d\})$ where $1, \alpha_1, \alpha_2, \ldots, \alpha_d$ are real algebraic numbers which are linearly independent over the rationals. Neiderreiter [8] has proved

$nD_n = O(n^\varepsilon)$ for all $\varepsilon > 0$ by using an extension of the Thue-Siegel-Roth theorem due to W. Schmidt [13].

EXAMPLE 3. (W. SCHMIDT [12]). For all $(\alpha_1\ \alpha_2, \ldots, \alpha_d) \in [0, 1]^d$, except a set of Lebesgue measure zero, the sequence $z_k = (\{k\alpha_1\}, \{k\alpha_2\}, \ldots, \{k\alpha_d\})$ satisfies $nD_n = O((\log n)^{d+1+\varepsilon})$.

Each of the preceding sequences has the benefit of being infinite, but the most widely used pseudorandom sequences do not have this property. The linear congruential pseudorandom numbers (LCPRN) are periodic with period m. This periodicity makes a limit result like Corollary 1 infeasible; but since bounds on D_n for LCPRN are known (Neiderreiter [9], [10]), the proof of Theorem 1 would yield a bound on $\lambda(n)$ for each $n < m$. Experience with LCPRN's indicates that in this case one should be able to obtain more precise information by direct analysis than by an application of discrepancy bounds, so this example has not been pursued.

One should similarly note that discrepancy estimates fall short of Corollary 1 for independent and uniformly distributed random variables. By Kiefer's [4] law of the iterated logarithm $nD_n = O((n \log\log n)^{1/2})$ with probability one, so by Theorem 1 and Few's bound one obtains only

$$(d-1)/d \geqslant \limsup \log \lambda(n)/\log n \geqslant \liminf \log \lambda(n)/\log n \geqslant (d-1)/2d.$$

Despite these two disappointments, Theorem 1 is essentially the best possible. In particular, the next result shows that if there is no estimate on the rate at which D_n tends to zero, the most one can say about $\lambda(n)$ is that $\lambda(n) \to \infty$.

THEOREM 2. *Given any positive $\phi(n)$ for which $\phi(n) \to \infty$, there is a uniformly distributed sequence for which, for all integers $n \geqslant 1$,*

$$\lambda(n) \leqslant \phi(n). \tag{3}$$

PROOF. For any finite sequences define their product by

$$(x_1, x_2, \ldots, x_k) * (x_1', x_2', \ldots, x_l') = (x_1, x_2, \ldots, x_k, x_1', x_2', \ldots, x_l')$$

and then define the powers $(x_1, x_2, \ldots, x_l)^{*k}$ correspondingly. Our basic lemma is the following:

For any uniformly distributed sequence (x_i) in $[0, 1]^d$ and any integers k_i for which ik_i increases, one has, for $(y_i)_{i=1}^\infty$ defined by

$$(y_1, y_2, \ldots) = (x_1)^{*k_1} * (x_1, x_2)^{*k_2} * (x_1, x_2, x_3)^{*k_3} * \ldots,$$

that $(y_i)_{i=1}^\infty$ is uniformly distributed.

To prove the lemma let D_n^x and D_n^y denote the respective discrepancies of $(x_i)_{i=1}^\infty$ and $(y_i)_{i=1}^\infty$. Now for any positive integer s there is a p such that

$$s = \sum_{i=1}^{p} ik_i + (p+1)q + r \quad \text{where } 0 \leqslant q < k_{p+1},\ 0 \leqslant r < p+1. \tag{4}$$

We therefore have for $(y_i)_{i=1}^\infty$ that

$$sD_s^y \leqslant \sum_{j=1}^{p} jk_j D_j^x + q(p+1)D_{p+1}^x + rD_r^x.$$

Letting $\psi(j) = \sup_{k \geqslant j} D_k^x$, we see $\psi(j)$ is nonincreasing and tends to zero by the uniform distribution of $(x_i)_{i=1}^{\infty}$. Since $D_r^x \leqslant 1$ and $q(p+1)/s \leqslant 1$,

$$D_s^y \leqslant \sum_{j=1}^{p} \psi(j)(jk_j/s) + \psi(p+1) + (p+1) \Big/ \sum_{i=1}^{p} i. \tag{5}$$

To show D_s^y tends to zero it thus suffices to prove the same for the sum $\Sigma_{j=1}^{p} \psi(j)(jk_j/s)$. For this recall Steffensen's inequality [7, p. 142].

If a_i, b_i, c_i are positive, c_i nonincreasing and $\Sigma_{i=1}^{m} a_i \leqslant \Sigma_{i=1}^{m} b_i$ for $1 \leqslant m \leqslant p$, then $\Sigma_{i=1}^{p} c_i a_i \leqslant \Sigma_{i=1}^{p} c_i b_i$.

In our case $c_i = \psi(i)$, $a_i = ik_i/s$ and $b_i = 1/p$. The facts that $\Sigma_{i=1}^{p} ik_i \leqslant s$ and that ik_i is nondecreasing provide the hypothesis $\Sigma_{i=1}^{m} a_i \leqslant \Sigma_{i=1}^{m} b_i$, $1 \leqslant m \leqslant p$. Consequently we have $\Sigma_{i=1}^{p}(\psi(i)ik_i)/s \leqslant (1/p)\Sigma_{i=1}^{p} \psi(i)$, and since $\psi(i)$ tends to zero, so do both sums. Together with (5) this completes the proof of the lemma.

To prove the theorem we estimate $\lambda^y(n)$ in terms of $\lambda^x(n)$ and $(k_i)_{i=1}^{\infty}$. First we note the elementary relations

$$\lambda(x_1, x_2, \ldots, x_l, x_1', x_2', \ldots, x_m') \leqslant d^{1/2} + \lambda(x_1, x_2, \ldots, x_l) + \lambda(x_1', x_2', \ldots, x_m'),$$

$$\lambda\big((x_1)^{*k}\big) = 0 \quad \text{and} \quad \lambda\big((x_1, x_2, \ldots, x_n)^{*k}\big) = \lambda(x_1, x_2, \ldots, x_n).$$

For s as in (4), one thus obtains

$$\lambda^y(s) \leqslant (p+1)d^{1/2} + \sum_{i=1}^{p+1} \lambda(x_1, x_2, \ldots, x_i). \tag{6}$$

By Few's inequality, $\lambda(x_1, x_2, \ldots, x_i) \leqslant c_d'(i)^{(d-1)/d}$, so (6) certainly implies $\lambda^y(s) < p^2$ for p sufficiently large. Now for any $\phi(n)$ which is integer valued and nondecreasing, we can choose $k_i \uparrow \infty$ so that $k_{\phi(n)} \geqslant n$ for all n. Since $s \geqslant \Sigma_{i=1}^{p} ik_i \geqslant k_p$ and $s \leqslant k_{\phi(s)}$, we have $\phi(s) \geqslant p$ by the monotonicity of the k_i. By the bound $\lambda^y(s) < p^2$ we thus have $\lambda^y(s) \leqslant \phi(s)^2$ for s sufficiently large. Since $\phi(s)$ in this construction was arbitrary, this implies the theorem for all n sufficiently large. By special consideration of k_1, the result is easily shown to hold for all n.

References

1. J. Beardwood, J. H. Halton and J. M. Hammersley, *The shortest path through many points*, Proc. Cambridge Philos. Soc. **55** (1959), 299–327.

2. L. Few, *The shortest path and the shortest road through n points*, Mathematika **2** (1955), 141–144.

3. J. H. Halton, *On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals*, Numer. Math. **2** (1960), 84–90; correction, p. 196.

4. J. Kiefer, *On the large deviation of the empiric d.f. of vector chance variables and a law of the iterated logarithm*, Pacific J. Math. **11** (1961), 649–660.

5. D. E. Knuth, *The art of computer programming*, Vol. 2, *Seminumerical algorithms*, Addison-Wesley, Palo Alto, California, 1969.

6. L. Kuipers and H. Niederreiter, *Uniform distribution of sequences*, Wiley, Toronto, 1974.

7. D. S. Mitrinović, *Analytic inequalities*, Springer-Verlag, New York, 1970.

8. H. Niederreiter, *Application of diophantine approximations to numerical integrations*, Diophantine Approximation and its Application (C. F. Osgood, ed.), Academic Press, New York, 1973, pp. 129–199.

9. _____, *Statistical independence of linear congruential pseudo-random numbers*, Bull. Amer. Math. Soc. **82** (1976), 927–929.

10. _____, *Pseudo-random numbers and optimal coefficients*, Advances in Math. **26** (1977), 99–181.

11. _____, *Quasi-monte carlo methods and pseudo-random numbers*, Bull. Amer. Math. Soc. **84** (1978), 957–1041.

12. W. Schmidt, *Metrical theorems on fractional parts of sequences*, Trans. Amer. Math. Soc. **110** (1964), 493–518.

13. _____, *Simultaneous approximation to algebraic numbers by rationals*, Acta Math. **125** (1970), 189–201.

Department of Statistics, Stanford University, Stanford, California 94305