MATHEMATICS OF COMPUTATION
VOLUME 55, NUMBER 191
JULY 1990, PAGES 289-297

ERROR ESTIMATES ARISING FROM CERTAIN PSEUDORANDOM SEQUENCES IN A QUASI-RANDOM SEARCH METHOD

RICARDO A. MITCHELL

ABSTRACT. In this paper we apply number-theoretic results to estimate the dispersion, a measure of denseness for sequences in a bounded set, of the Halton and Hammersley sequences in the hypercube $I^s = [0, 1]^s$. It is seen that they attain the minimal order of magnitude for the dispersion.

1. INTRODUCTION

Random search methods are common in nonsmooth optimization. These methods are based on selecting random samples from the domain of the target function. The effectiveness of these methods depends on the distribution of the random sample selected. If the random sample is replaced by a deterministic point set, we have a type of quasi-random search method developed by Niederreiter [1].

To describe the method of Niederreiter, we consider the problem

$$M = \sup_{x\in A} f(x),$$

where $f: A \to \mathbb{R}$ is continuous on the bounded set $A \subseteq \mathbb{R}^s$, $s \geq 1$. Let $\{x_i\}_{i=1}^N$ be a deterministic point set in A.

We define the modulus of continuity of f on A by

$$\omega_f(t) = \sup_{\substack{x, y\in A \\ d(x, y)\leq t}} |f(x) - f(y)|, \qquad t \geq 0,$$

and the dispersion by $d_N = \sup_{x\in A} \min_{1\leq i\leq N} d(x, x_i)$, where $d(\cdot, \cdot)$ is a metric on A, normally taken to be the maximum or Euclidean metric. An approximation to M is $M_N = \max_{1\leq i\leq N} f(x_i)$ with the error bound $M - M_N \leq \omega_f(d_N)$. We note that as f is continuous, convergence to the global solution is assured, i.e., $M_N \to M$ as $N \to \infty$, if $d_N \to 0$ as $N \to \infty$.

Received August 30, 1988; revised May 9, 1989.

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 65C10; Secondary 11K38, 11K45.

Key words and phrases. Global optimization, dispersion, discrepancy, pseudorandom sequence, Chinese remainder theorem, denseness.

2. Bounds for d_N

It has been shown in [1] that for any N points in A, $d_N \geq C_A N^{-1/s}$, where C_A is a constant depending only on A. We assume from now on that A is a subset of $I^s = [0, 1]^s$. Then the dispersion is related to the most useful measure of uniform distribution for sequences, called the discrepancy. This is defined by

$$D_N = \sup_K \left| \frac{A(K, N)}{N} - V(K) \right| ,$$

where K runs through all subintervals of I^s and the counting function $A(K, N)$ is the number of i, $1 \leq i \leq N$, such that $x_i \in K$. The relation

$$d_N \leq \sqrt{s} D_N^{1/s} \tag{1}$$

is established in [1] for the Euclidean metric. For the maximum metric one obtains

$$d'_N \leq D_N^{1/s} \tag{2}$$

according to [6].

Consequently, bounds of the Erdös-Turán-Koksma type (see [2]) may be obtained for the dispersion, using (1) and (2). For the case $s = 1$, we have

$$d_N \leq C \left(\frac{1}{m+1} + \sum_{h=1}^{m} \left(\frac{1}{h} - \frac{1}{m+1} \right) \left| \frac{1}{N} \sum_{k=1}^{N} \exp(2\pi i h x_k) \right| \right)$$

for all $m \in \mathbb{N}$. However, employing a theorem of Niederreiter and Philipp [3], we obtain a different inequality. We start with

Remark 1. For any continuous function $f: A \subseteq \mathbb{R} \to \mathbb{R}$ and any compact interval $I \subseteq A$, if $\bigcup_{i \in J} I_i = I$, J a finite index set, we have

$$\sup_{x \in I} f(x) = \max_{i \in J} \sup_{x \in I_i} f(x).$$

Remark 2. Let $x_1 \leq x_2 \leq \cdots \leq x_N$ be N points in $I = [0, 1]$. The dispersion $d_N(I)$ of these points in I is given by

$$d_N(I) = \max_{1 \leq i \leq N} \sup_{x \in S(x_i)} |x_i - x| ,$$

where $S(x_i) = \{x \in I | a_{i-1} \leq x \leq a_i\}$, $1 \leq i \leq N$, with $a_0 = 0$, $a_i = (x_i + x_{i+1})/2$, $1 \leq i \leq N-1$, and $a_N = 1$.

Remark 3. For any sequence of N points $x_1 \leq x_2 \leq \cdots \leq x_N$ in $I = [0, 1]$, we have

$$d_N(I) = \max_{0 \leq i \leq N} \sup_{x \in [x_i, x_{i+1}]} |a_i - x| ,$$

where $x_0 = 0$ and $x_{N+1} = 1$.

Lemma 1. *The dispersion of the points* $x_1 \le x_2 \le \cdots \le x_N$ *in* $[0, 1]$ *satisfies*

$$d_N(I) \le \frac{4}{m+1} + \frac{4}{\pi}\sum_{h=1}^{m}\left(\frac{1}{h} - \frac{1}{m+1}\right)\left|\sum_{k=1}^{N}(a_k - a_{k-1})\exp(2\pi i h x_k)\right|$$

for all $m \in \mathbb{N}$.

Proof. Let $f\colon I = [0, 1] \to \mathbb{R}$ be defined by $f(0) = 0$ and $f(x) = a_i$ for $x_i < x \le x_{i+1}$, $0 \le i \le N$, where $x_0 = 0$ and $x_{N+1} = 1$. Then $f(0) = 0$ and $f(1) = 1$, and f is a nondecreasing function on $[0, 1]$. Also

$$\begin{aligned}\sup_{x\in I}|f(x) - x| &= \max_{0\le i\le N}\sup_{x\in(x_i, x_{i+1}]}|f(x) - x| \\ &= \max_{0\le i\le N}\sup_{x\in(x_i, x_{i+1}]}|a_i - x| = d_N(I).\end{aligned}$$

Invoking Theorem 1 in [3], we get

$$d_N(I) \le 4\left\{\frac{1}{m+1} + \frac{1}{\pi}\sum_{h=1}^{m}\left(\frac{1}{h} - \frac{1}{m+1}\right)|\hat{f}(h)|\right\}$$

for all $m \in \mathbb{N}$, where $\hat{f}(h) = \int_0^1 \exp(2\pi i h x)\,df(x)$. Clearly,

$$\hat{f}(h) = \sum_{k=1}^{N}(a_k - a_{k-1})\exp(2\pi i h x_k).$$

The result follows. □

H. Niederreiter [6] first proved a result of the type given in Lemma 1. By invoking the following theorem of Niederreiter [7], we obtain yet another inequality.

Lemma 2 (Niederreiter [7]). *Let* f *be a nondecreasing function on* $[0, 1] = I$ *with* $f(0) = 0$, $f(1) = 1$. *Suppose the function* g *on* I *satisfies a Lipschitz condition, i.e.,* $|g(u) - g(v)| \le L|u - v|$ *for all* $u, v \in I$, *as well as* $g(0) = 0$, $g(1) = 1$. *Then*

$$\begin{aligned}&\sup_{u, v\in I}|(f(u) - f(v)) - (g(u) - g(v))| \\ &\quad\le \left\{\frac{6L}{\pi^2}\sum_{h=1}^{\infty}\frac{1}{h^2}|\hat{f}(h) - \hat{g}(h)|^2\right\}^{1/3}.\end{aligned}$$

Corollary 1. *Let* $x_1 \le x_2 \le \cdots \le x_N$ *be* N *points in* $I = [0, 1]$. *Then the dispersion of these points in* I *satisfies*

$$d_N(I) \le \left\{\frac{6}{\pi^2}\sum_{h=1}^{\infty}\frac{1}{h^2}\left|\sum_{k=1}^{N}(a_k - a_{k-1})\exp(2\pi i h x_k)\right|^2\right\}^{1/3}. \tag{3}$$

Proof. Same as that of Lemma 1. □

If we choose $x_i = 0$, $i = 1, 2, \ldots, N$, then $d_N(I) = 1$. We see that the right-hand side of (3) reduces to

$$\left(\frac{6}{\pi^2}\sum_{h=1}^{\infty}\frac{1}{h^2}\right)^{1/3} = 1.$$

Hence the constant $6/\pi^2$ in (3) is best possible.

The following results are easily obtained.

Remark 4. Let $x_1 \le x_2 \le \cdots \le x_N$ be N points in $I = [0, 1]$ with dispersion $d_N(I)$, and let f be a function of bounded variation $V(f)$ on I. Then

(i) $|\int_0^1 f(t)\,dt - \sum_{k=1}^N (a_k - a_{k-1}) f(x_k)| \le V(f) d_N(I)$ and

(ii) $|\sum_{k=1}^N (a_k - a_{k-1}) \exp(2\pi i x_k)| \le 4 d_N(I)$.

Similar inequalities were found by Kuipers and Niederreiter in [4] for the discrepancy D_N and the sum $\frac{1}{N}\sum_{k=1}^N f(x_k)$ in the case of (i), and D_N and the sum $\frac{1}{N}\sum_{k=1}^N \exp(2\pi i x_k)$ in the case of (ii).

3. The Halton and Hammersley sequences in I^s

Let $R \in \mathbb{N} - \{1\}$; then any nonnegative integer K may be uniquely represented as

$$K = \sum_{j=0}^{M} a_j R^j, \qquad 0 \le a_j \le R - 1. \tag{4}$$

Let $S_N = \{0, 1, \ldots, N-1\}$. Define the injective map $\phi_R : \mathbb{N} \cup \{0\} \to [0, 1]$, with radix R by

$$\phi_R(K) = \sum_{j=0}^{M} a_j R^{-j-1}, \tag{5}$$

where K, R, and a_j, $j = 0, 1, \ldots, M$, are as defined in (4).

Definition. The s-dimensional Halton sequences are defined by

$$(\phi_{R_1}(l), \phi_{R_2}(l), \ldots, \phi_{R_s}(l)), \qquad l = 0, 1, \ldots, \tag{6}$$

where R_i, $i = 1, \ldots, s$, are pairwise relatively prime and $\min_i R_i \ge 2$.

The s-dimensional Hammersley sequences are given by

$$\left(\frac{l}{N}, \phi_{R_1}(l), \ldots, \phi_{R_{s-1}}(l)\right), \qquad l \in S_N, \tag{7}$$

where R_i, $i = 1, \ldots, s-1$, are pairwise relatively prime (usually taken to be the first $s - 1$ primes) and $\min_i R_i \ge 2$.

Information on Halton and Hammersley sequences can be found in [2, 5].

Lemma 3. *Let $R_1, R_2, \ldots, R_s$ be pairwise relatively prime and $n_1, \ldots, n_s$ be nonnegative integers such that $N \geq \prod_{i=1}^{s} R_i^{n_i}$. Let $l_1, \ldots, l_s$ be integers with $0 \leq l_i \leq R_i^{n_i} - 1$ for $i = 1, \ldots, s$. Then there exists a number $L \in S_N$ such that*

$$(\phi_{R_1}(L), \ldots, \phi_{R_s}(L)) \in \mathop{\times}_{i=1}^{s} [l_i R_i^{-n_i}, (l_i + 1)R_i^{-n_i}).$$

Proof. Let $(\alpha_1, \ldots, \alpha_s) \in \times_{i=1}^{s}[l_i R_i^{-n_i}, (l_i + 1)R_i^{-n_i})$. Then for any i, $1 \leq i \leq s$, α_i may be represented uniquely by

$$\begin{aligned}\alpha_i &= \sum_{j=0}^{n_i - 1} a_{ij} R_i^{j-n_i} + \sum_{j=n_i}^{\infty} a_{ij} R_i^{-j-1} \\ &= \sum_{j=0}^{n_i-1} a_{i,n_i-j-1} R_i^{-j-1} + \sum_{j=n_i}^{\infty} a_{ij} R_i^{-j-1}.\end{aligned}$$

By definition, there is a corresponding L, $0 \leq L \leq N - 1$, with $\phi_{R_i}(L) \in [l_i R_i^{-n_i}, (l_i + 1)R_i^{-n_i}]$ for $i = 1, \ldots, s$ if and only if

$$L = \sum_{j=0}^{n_i-1} a_{i,n_i-j-1} R_i^j + \sum_{j=n_i}^{K_i} a_{ij} R_i^j, \qquad K_i \in \mathbb{N}, \quad i = 1, \ldots, s,$$

or

$$L = \beta_i + \sum_{j=n_i}^{K_i} a_{ij} R_i^j, \qquad i = 1, \ldots, s,$$

where $0 \leq \beta_i = \sum_{j=0}^{n_i-1} a_{i,n_i-j-1} R_i^j \leq R_i^{n_i} - 1$. Hence, equivalently, L satisfies the congruences

$$L \equiv \beta_i \pmod{R_i^{n_i}}, \qquad i = 1, \ldots, s. \tag{8}$$

By the Chinese remainder theorem there exists an $L \in S_N$ satisfying (8). The lemma is established. □

With the aid of Lemma 3 we may now establish the following

Theorem 1. *Let the integers $R_1, \ldots, R_s \geq 2$ be pairwise relatively prime. Then the sequence (6) has dispersion $d_N(HALT)$ satisfying*

$$d_N(HALT) \leq C(R_i)N^{-1/s} \quad \textit{for } N \geq \prod_{i=1}^{s} R_i,$$

where $C(R_i)$ is a constant depending only on $R_1, \ldots, R_s$.

Proof. We can assume that $R_1 = \min(R_1, \ldots, R_s)$. For any $N \geq \prod_{i=1}^{s} R_i$ there exists a positive integer k_1 such that $R_1^{k_1} \leq N < R_1^{k_1+1}$. Now choose the integers $k_2, \ldots, k_s$ such that $R_i^{k_i} < R_1^{k_1+1} < R_i^{k_i+1}$ for $i = 2, \ldots, s$ and define the nonnegative integers $n_1, \dots, n_s$ by $n_i = [k_i/s]$, $i = 1, 2, \dots, s$, where $[x]$ is the greatest integer less than or equal to x. Then either

$$\text{(9)} \qquad \text{(i)} \quad R_1^{k_1+1} > N \geq \prod_{i=1}^{s} R_i^{n_i}$$

or

$$\text{(10)} \qquad \text{(ii)} \quad R_1^{k_1+1} \geq \prod_{i=1}^{s} R_i^{n_i} > N \geq R_1^{k_1}.$$

Case (i). We note that

$$[0, 1)^s = \bigcup_{l_i=0}^{R_i^{n_i}-1} \mathop{\times}_{i=1}^{s} [l_i R_i^{-n_i}, (l_i+1)R_i^{-n_i}).$$

Moreover, by Lemma 3 there is at least one point from the first N terms of (6) in each hyperrectangle

$$\mathop{\times}_{i=1}^{s} [l_i R_i^{-n_i}, (l_i+1)R_i^{-n_i}), \qquad 0 \leq l_i \leq R_i^{n_i} - 1.$$

Hence, we have

$$d_N^2(HALT) \leq \sum_{i=1}^{s} (R_i^{-n_i})^2.$$

Now

$$k_i < sn_i + s - 1 \Leftrightarrow k_i + 1 \leq s(n_i + 1),$$

so

$$R_i^{s(n_i+1)} \geq R_i^{k_i+1} \geq R_1^{k_1+1} > N \quad \text{for } i = 1, \dots, s,$$

which implies that

$$R_i^{-n_i} < R_i N^{-1/s} \quad \text{for } i = 1, \dots, s.$$

It now follows that

$$d_N(HALT) < \left(\sum_{i=1}^{s} R_i^2 \right)^{1/2} N^{-1/s}.$$

Case (ii). If N does not satisfy (9), then we have (10). Hence,

$$N \geq R_1^{k_1} > \left(\prod_{i=2}^{s} R_i^{n_i} \right) R_1^{n_1 - 1}.$$

Note that $N \geq \prod_{i=1}^{s} R_i$ implies that $n_1 \geq 1$. Using arguments similar to those in Case (i), we find that

$$d_N^2(HALT) \leq R_1^{-2(n_1-1)} + \sum_{i=2}^{s} R_i^{-2n_i}.$$

It then follows that

$$d_N(HALT) < \left[R_1^4 + \sum_{i=2}^{s} R_i^2\right]^{1/2} N^{-1/s}.$$

The theorem is proved. □

Theorem 2. *Let* $R_1, \dots, R_{s-1}$ *be integers* ≥ 2 *that are pairwise relatively prime. Then the sequence* (7) *has dispersion* $d_N(HAMM)$ *satisfying*

$$d_N(HAMM) \leq C(R_i)N^{-1/s} \quad \textit{for } N \geq \prod_{i=1}^{s-1} R_i,$$

where $C(R_i)$ *is a constant depending only on* $R_1, \dots, R_{s-1}$.

Proof. In the first paragraph of the proof of Theorem 1, replace s by $s-1$ and N by $N^{1-1/s}$ and select $n_1, \dots, n_{s-1}$ in the same way. Then either

$$\text{(11)} \qquad \text{(i)} \quad R_1^{k_1+1} > N^{1-1/s} \geq \prod_{i=1}^{s-1} R_i^{n_i}$$

or

$$\text{(12)} \qquad \text{(ii)} \quad R_1^{k_1+1} \geq \prod_{i=1}^{s-1} R_i^{n_i} > N^{1-1/s} \geq R_1^{k_1}.$$

Choose $M = [N/\prod_{i=1}^{s-1} R_i^{n_i}]$, then divide the interval $[0, 1]$ into the M consecutive intervals $[pM^{-1}, (p+1)M^{-1}]$, where the integer p satisfies $0 \leq p \leq M-1$. We observe that for each p, $0 \leq p \leq M-1$, the interval $[pM^{-1}, (p+1)M^{-1}]$ contains, for some $l_1 \in \mathbb{Z}$, the Q numbers, $Q = \prod_{i=1}^{s-1} R_i^{n_i}$,

$$\frac{l_1}{N}, \frac{l_1+1}{N}, \dots, \frac{l_1+Q-1}{N}.$$

It follows that there exists an l satisfying the $(s-1)$ congruences

$$l \equiv \beta_i \pmod{R_i^{n_i}}, \quad i = 2, 3, \dots, s-1, \qquad l \equiv \beta_1 \pmod{R_1^p},$$

and $l/N \in [pM^{-1}, (p+1)M^{-1}]$, where $p = n_1$ or $n_1 - 1$ if (i) or (ii) holds, respectively. From this we deduce that

$$d_N^2(HAMM) \leq M^{-2} + R_1^{-2p} + \sum_{i=2}^{s-1} R_i^{-2n_i}.$$

Proceeding as in the proof of Theorem 1, we arrive at the conclusion

$$d_N(HAMM) \leq \left(R_1^{2a} + R_1^{2b} + \sum_{i=2}^{s-1} R_i^2\right)^{1/2} N^{-1/s},$$

where $a = 0$ if $M > N^{1/s}$, $a = 1$ if $M \leq N^{1/s} \leq M+1$, $b = 1$ if (11) holds, and $b = 2$ if (12) holds. This completes the proof. □

The estimates for $d_N(HALT)$ and $d_N(HAMM)$ obtained in the proofs of the above results are clearly not the best possible. If N is known and the radices R_i, $1 \le i \le s$, are chosen, then a direct calculation will yield better values for these estimates. Let us consider the following examples.

Example 1. Take $N = 72$ and $s = 2$. Good choices for the radices are $R_1 = 2$ and $R_2 = 3$ in the case of the Halton sequence. Clearly, $N = 72 = 2^3 3^2$, i.e., $n_1 = 3$, $n_2 = 2$, then $d_N(HALT) < \sqrt{(1/8)^2 + (1/9)^2}$. Since it is more economical to convert the integers $\{l: 0 \le l \le N-1\}$ to bases 8 and 9 than to the bases 2 and 3, the radices $R_1 = 8$ and $R_2 = 9$ would be preferred as the estimate for $d_N(HALT)$ remains the same.

Example 2. If $s = 2$ and $N = 675$ for the Halton sequence, and $R_1 = 2$, $R_2 = 3$, then

$$2^5 3^3 = 864 > N > 2^{5-1} 3^3.$$

Hence, $d_N(HALT) \le \sqrt{(1/16)^2 + (1/27)^2}$. However, if we take $R_1 = 25$, $R_2 = 27$, then $N = R_1 R_2$ and $d_N(HALT) < \sqrt{(1/25)^2 + (1/27)^2}$. Thus the pair of radices $R_1 = 25$, $R_2 = 27$ gives a better distribution. In both cases the estimates are better than those implied by Theorem 1.

4. Conclusion

We have already seen that for any N points in a bounded set A, $d_N(A) \ge C_A N^{-1/s}$, where C_A is a constant depending only on A. Hence both the Halton and Hammersley sequences possess the minimal order of magnitude for the dispersion. This justifies their importance in the search method mentioned in the introduction. When using the Halton sequences, the following is suggested: let $N_m = m \prod_{i=1}^{s} R_i^{n_i}$ be the number of function evaluations available, where radices R_i, $i = 1, \ldots, s$, and the numbers n_i, $i = 1, \ldots, s$, are selected to minimize the estimate for d_{N_m}. We start the search with N points where $N = N_m/m$. We then keep adding N points.

From the proof of Theorem 1, we observe that the first N points divide the cube I^s into N rectangles, and after N_m points have been used, each rectangle, by Lemma 3, has m points. Thus the Halton sequences, when used in this manner, behave like a "stratified" sample.

Bibliography

1. H. Niederreiter, *A quasi-Monte Carlo method for the approximate computation of the extreme values of a function*, Studies in Pure Mathematics (To the Memory of Paul Turán), Birkhäuser Verlag, Basel, 1983, pp. 523–529.

2. ——, *Quasi-Monte Carlo methods and pseudo-random numbers*, Bull. Amer. Math. Soc. **84** (1978), 957–1041.

3. H. Niederreiter and W. Philipp, *Berry-Esseen bounds and a theorem of Erdös and Turán on uniform distribution mod* 1, Duke Math. J. **40** (1973), 633–649.

4. L. Kuipers and H. Niederreiter, *Uniform distribution of sequences*, Wiley, New York, 1974.

5. L. K. Hua and Y. Wang, *Applications of number theory to numerical analysis*, Springer-Verlag, Berlin, 1981.

6. H. Niederreiter, *On a measure of denseness for sequences*, Topics in Classical Number Theory (Budapest, 1981), Colloq. Math. Soc. János Bolyai, vol 34, North-Holland, Amsterdam, 1984, pp. 1163–1208.

7. ——, *Quantitative versions of a result of Hecke in the theory of uniform distribution mod* 1, Acta Arith. **28** (1975), 321–339.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF THE WEST INDIES, MONA, KINGSTON 7, JAMAICA, WEST INDIES