NOTE ON IRREGULARITIES OF DISTRIBUTION

J. BECK AND W. W. L. CHEN

§1. *Introduction.* Let $U^K = [0, 1)^K$ be the K-dimensional unit cube, where $K \geq 2$. Suppose that we have a distribution $\mathcal{P}$ of N points in U^K. For $\mathbf{x} = (x_1, \ldots, x_K) \in U^K$, let $A(\mathbf{x})$ denote the box

$$A(\mathbf{x}) = [0, x_1] \times \ldots \times [0, x_K] \subset U^K,$$

and write

$$D[\mathcal{P}; A(\mathbf{x})] = \sum_{\mathbf{p} \in A(\mathbf{x}) \cap \mathcal{P}} 1 - N x_1 \ldots x_K. \tag{1}$$

Note that since N is the cardinality of $\mathcal{P}$ and $x_1 \ldots x_K$ is the K-dimensional volume of $A(\mathbf{x})$, the term $N x_1 \ldots x_K$ represents the "expected number" of points of $\mathcal{P}$ in $A(\mathbf{x})$.

The irregularity of the distribution $\mathcal{P}$ can be measured by the behaviour of the (discrepancy) function $D[\mathcal{P}; A(\mathbf{x})]$, where $\mathbf{x} \in U^K$. Here, we consider, for $0 < W < \infty$, the L^W-norm

$$\|D_A(\mathcal{P})\|_W = \left(\int_{U^K} |D[\mathcal{P}; A(\mathbf{x})]|^W d\mathbf{x} \right)^{1/W}.$$

More precisely, we study the behaviour of the function

$$\Delta_A(K, W, N) = \inf_{\mathcal{P}} \|D_A(\mathcal{P})\|_W,$$

where the infimum is taken over all distributions $\mathcal{P}$ of N points in U^K.

In 1954, K. F. Roth proved the following remarkable lower bound.

THEOREM 1 (Roth [6]). *There exists a postive constant $c_1(K)$, depending at most on K, such that for every $N \geq 2$, we have*

$$\Delta_A(K, 2, N) > c_1(K)(\log N)^{(K-1)/2}.$$

This was complemented by the following upper bound (the interested reader is also referred to the special cases studied by Davenport [4] and Roth [7]).

THEOREM 2 (Roth [8]). *There exists a positive constant $c_2(K)$, depending at most on K, such that for every $N \geq 2$, we have*

$$\Delta_A(K, 2, N) < c_2(K)(\log N)^{(K-1)/2}.$$

In other words, there exists a positive constant $c_3(K)$, depending at most on K, *such that for every $N \geq 2$, there exists a distribution $\mathcal{P}$ of N points in U^K such that*

$$\int_{U^K} |D[\mathcal{P}; A(\mathbf{x})]|^2 d\mathbf{x} < c_3(K)(\log N)^{K-1}. \tag{2}$$

Next, let $\mathbf{p}_1, \ldots, \mathbf{p}_N$ be a sequence of N points in U^K. For every natural number $M = 1, \ldots, N$, let

$$D[\mathbf{p}_1, \ldots, \mathbf{p}_M; A(\mathbf{x})] = \sum_{\substack{i=1 \\ \mathbf{p}_i \in A(\mathbf{x})}}^{M} 1 - Mx_1 \ldots x_K \tag{3}$$

(in other words, consider the discrepancy of the set $\{\mathbf{p}_1, \ldots, \mathbf{p}_M\}$ in the box $A(\mathbf{x})$; note also that (3) is equivalent to (1) when $M = N$), and consider the expression

$$\frac{1}{N} \sum_{M=1}^{N} \int_{U^K} |D[\mathbf{p}_1, \ldots, \mathbf{p}_M; A(\mathbf{x})]|^2 d\mathbf{x}.$$

For $i = 1, \ldots, N$, let $\mathbf{p}_i^* = (\mathbf{p}_i, (i-1)/N) \in U^{K+1}$. It is easy to see that

$$\frac{1}{N} \sum_{M=1}^{N} \int_{U^K} |D[\mathbf{p}_1, \ldots, \mathbf{p}_M; A(\mathbf{x})]|^2 d\mathbf{x} \gg_K \int_{U^{K+1}} |D[\mathbf{p}_1^*, \ldots, \mathbf{p}_N^*; A(\mathbf{y})]|^2 d\mathbf{y},$$

and so by the analogue of Theorem 1 in dimension $K+1$, we have

$$\frac{1}{N} \sum_{M=1}^{N} \int_{U^K} |D[\mathbf{p}_1, \ldots, \mathbf{p}_M; A(\mathbf{x})]|^2 d\mathbf{x} > c_4(K)(\log N)^K, \tag{4}$$

where $c_4(K)$ is a positive constant depending at most on K.

Comparing (2) and (4), we see that for a sequence, the square-integral of the discrepancy function is definitely greater than that for the "optimal set". In fact, there is an extra factor of order $(\log N)$. This is a nice quantitative distinction between sets and sequences. We call it *Roth phenomenon.*

Let us now consider the analogous question for balls (more precisely, for balls in the unit torus). For $0 \leq r \leq \frac{1}{2}$ and $\mathbf{x} \in U^K$, write

$$B(r, \mathbf{x}) = \left\{ \mathbf{y} \in \mathbb{R}^K : |\mathbf{x} - \mathbf{y}| = \left(\sum_{i=1}^{K} (x_i - y_i)^2 \right)^{1/2} \leq r \right\},$$

and

$$B^*(r, \mathbf{x}) = \bigcup_{\mathbf{l} \in \mathbf{Z}^K} (B(r, \mathbf{x}) + \mathbf{l}).$$

Given a distribution $\mathcal{P}$ of N points in U^K, write

$$Z(\mathcal{P}; r, \mathbf{x}) = \sum_{\mathbf{l} \in \mathbf{Z}^K} \sum_{\substack{\mathbf{p} \in \mathcal{P} \\ \mathbf{p} + \mathbf{l} \in B(r, \mathbf{x})}} 1 = \sum_{\mathbf{p} \in B^*(r, \mathbf{x}) \cap \mathcal{P}} 1.$$

If $\mathcal{P}$ is well distributed in U^K, then

$$Z(\mathcal{P}; r, \mathbf{x}) \sim N\mu(B(r, \mathbf{x})),$$

where μ denotes, as usual, the K-dimensional volume (Lebesgue measure).

In 1969, W. M. Schmidt proved the following pioneering result.

THEOREM 3A (Schmidt [9]). *For every $\varepsilon > 0$, there exists a positive constant $c_5(K, \varepsilon)$, depending at most on K and ε, such that for every distribution $\mathcal{P}$ of N points in U^K, there exists a ball $B(r, \mathbf{x})$ with $0 < r \leq \frac{1}{2}$ such that*

$$|Z(\mathcal{P}; r, \mathbf{x}) - N\mu(B(r, \mathbf{x}))| > c_5(K, \varepsilon)N^{\frac{1}{2}-(1/2K)-\varepsilon}.$$

Recently H. L. Montgomery and J. Beck slightly improved on Schmidt's lower bound by cancelling the ε from the exponent. In fact, both proved the following stronger L^2-norm estimate.

THEOREM 3B (Montgomery [5], Beck [2]). *There exists a positive constant $c_6(K)$, depending at most on K, such that for every distribution $\mathcal{P}$ of N points in U^K, we have*

$$\int_0^{1/2} \int_{U^K} |Z(\mathcal{P}; r, \mathbf{x}) - N\mu(B(r, \mathbf{x}))|^2 d\mathbf{x}\, dr > c_6(K)N^{1-(1/K)}. \tag{5}$$

This lower bound is best possible, apart from the value of the constant $c_6(K)$ (see the random construction in Beck [1]).

Again, let $\mathbf{p}_1, \ldots, \mathbf{p}_N$ be a sequence of N points in U^K. For every natural number $M = 1, \ldots, N$, let

$$Z(\mathbf{p}_1, \ldots, \mathbf{p}_M; r, \mathbf{x}) = \sum_{\mathbf{l} \in \mathbb{Z}^K} \sum_{\substack{i=1 \\ \mathbf{p}_i + \mathbf{l} \in B(r, \mathbf{x})}}^{M} 1,$$

and consider the expression

$$\frac{1}{N} \sum_{M=1}^{N} \int_0^{1/2} \int_{U^K} |Z(\mathbf{p}_1, \ldots, \mathbf{p}_M; r, \mathbf{x}) - M\mu(B(r, \mathbf{x}))|^2 d\mathbf{x}\, dr. \tag{6}$$

One might suspect that (6) is greater than $f(N)N^{1-(1/K)}$ where $f(N) \to \infty$ as $N \to \infty$. The object of this paper, however, is to prove that this is not the case. In fact, we prove

THEOREM 4. *There exists a positive constant $c_7(K)$, depending at most on K, such that for every $N \geq 1$, there exists a sequence $\mathbf{p}_1, \ldots, \mathbf{p}_N$ of N points in U^K such that*

$$\frac{1}{N} \sum_{M=1}^{N} \int_0^{1/2} \int_{U^K} |Z(\mathbf{p}_1, \ldots, \mathbf{p}_M; r, \mathbf{x}) - M\mu(B(r, \mathbf{x}))|^2 d\mathbf{x}\, dr < c_7(K)N^{1-(1/K)}. \tag{7}$$

In other words, there is *no* Roth phenomenon for balls (compare (5) and (7)).

Theorem 4 can be reformulated in the following way. For $0 \leq r \leq \frac{1}{2}$, $\mathbf{x} \in U^K$ and $y \in [0, 1)$, let $C(r, \mathbf{x}, y) \subset \mathbb{R}^{K+1}$ be the cartesian product of the ball

$B(r, \mathbf{x}) \subset \mathbb{R}^K$ and the interval $[0, y]$, *i.e.*,

$$C(r, \mathbf{x}, y) = \left\{ \mathbf{z} = (z_1, \ldots, z_{K+1}) \in \mathbb{R}^{K+1} : \sum_{i=1}^{K} (x_i - z_i)^2 \leq r^2 \text{ and } 0 \leq z_{K+1} \leq y \right\}$$

(note that for $K = 2$, $C(r, \mathbf{x}, y)$ is an ordinary cylinder). Write

$$C^*(r, \mathbf{x}, y) = \bigcup_{\mathbf{l} \in \mathbb{Z}^{K+1}} (C(r, \mathbf{x}, y) + \mathbf{l}).$$

Let $\mathcal{P}$ be a distribution of N points in U^{K+1}, and write

$$Z(\mathcal{P}; r, \mathbf{x}, y) = \sum_{\mathbf{p} \in C^*(r, \mathbf{x}, y) \cap \mathcal{P}} 1.$$

Let

$$\|D_C(\mathcal{P})\|_2 = \left(\int_0^{1/2} \int_0^1 \int_{U^K} |Z(\mathcal{P}; r, \mathbf{x}, y) - N\mu(C(r, \mathbf{x}, y))|^2 d\mathbf{x}\, dy\, dr \right)^{1/2},$$

and

$$\Delta_C(K+1, 2, N) = \inf_{\mathcal{P}} \|D_C(\mathcal{P})\|_2,$$

where the infimum is taken over all distributions $\mathcal{P}$ of N points in U^{K+1}.

Let $\mathbf{p}_1, \ldots, \mathbf{p}_N$ be a sequence of N points in U^K satisfying Theorem 4. Then the $(K+1)$-dimensional set $\mathcal{P} = \{(\mathbf{p}_i, (i-1)/N) : 1 \leq i \leq N\} \subset U^{K+1}$ gives the following upper bound.

THEOREM 4′. *There exists a positive constant $c_8(K)$, depending at most on K, such that for every $N \geq 1$, we have*

$$\Delta_C(K+1, 2, N) < c_8(K) N^{\frac{1}{2} - (1/2K)}. \tag{8}$$

Note that the lower bound

$$\Delta_C(K+1, 2, N) > c_9(K) N^{\frac{1}{2} - (1/2K)} \tag{9}$$

follows from Theorem 3B.

Finally, we note that the analogous questions for L^1-norms are quite open. Also, one of the most important open problems is to determine the correct order of magnitude of the function

$$\Delta_A(K, \infty, N) = \inf_{\mathcal{P}} \sup_{\mathbf{x} \in U^K} |D[\mathcal{P}; A(\mathbf{x})]|,$$

where the infimum is taken over all distributions $\mathcal{P}$ of N points in U^K. It is not even known, for $K \geq 2$, whether

$$\frac{\Delta_A(K+1, \infty, N)}{\Delta_A(K, \infty, N)} \to \infty \quad \text{as } N \to \infty.$$

For other results and open problems in this field, we refer the reader to the excellent book of W. M. Schmidt [**10**].

The work in this paper was carried out while the first author was Visiting Fellow in the Department of Mathematics at Imperial College. We thank the Science and Engineering Research Council for the financial support.

§2. *The basic idea.* We shall only prove Theorem 4 in the special case $K=2$. The general case is similar but notationally more complicated.

Given any integer $n\geq 1$, we shall show that there exists a sequence $\mathbf{p}_1,\ldots,\mathbf{p}_N$ of $N=4^n$ points in the unit square U^2 such that

$$\frac{1}{N}\sum_{M=1}^{N}\int_0^{1/2}\int_{U^2}|Z(\mathbf{p}_1,\ldots,\mathbf{p}_M;r,\mathbf{x})-M\mu(B(r,\mathbf{x}))|^2d\mathbf{x}dr\leq c_{10}N^{1/2}, \quad (10)$$

where c_{10} is a positive absolute constant.

In §3, we first construct a sequence $\mathbf{q}_1,\ldots,\mathbf{q}_N$ of $N=4^n$ points in U^2 using ideas in combinatorics. This sequence plays a role (see Lemma 1) analogous to that of the Hammersley-Halton sequence in Roth [**8**], but, as in [**8**], is insufficient to give the desired result.

In order to make use of this intermediate construction, we appeal to tools in probability theory in §4. However, the ideas here are quite different from those of Roth in [**8**]. Instead, we follow more closely the "randomization" arguments used in [**1**].

We complete the proof of Theorem 4 in §5.

§3. *A combinatorial approach.* Let $n\geq 1$ be given. We first partition U^2 into $N=4^n$ smaller squares as follows. For $0\leq i,j<2^n$, let

$$S(i,j)=[2^{-n}i,2^{-n}(i+1))\times[2^{-n}j,2^{-n}(j+1)).$$

For every integer l satisfying $0\leq l\leq n-1$, vectors $\boldsymbol{\tau}\in\{0,1,2,3\}^l$, $\mathbf{u}\in\{0,1\}^{n-l-1}$ and $\mathbf{v}\in\{0,1\}^{n-l-1}$, let $F[\boldsymbol{\tau};\mathbf{u},\mathbf{v}]$ be a bijective mapping from $\{0,1\}^2$ to $\{0,1,2,3\}$ (if $l=0$, write $\boldsymbol{\tau}=\varnothing$; if $l=n-1$, then write $\mathbf{u}=\mathbf{v}=\varnothing$).

Given these mappings, we can define a bijective mapping

$$G:\{0,1,\ldots,2^n-1\}^2\longrightarrow\{1,2,\ldots,N=4^n\}$$

as follows. Suppose that $0\leq i,j<2^n$. Write

$$\left.\begin{aligned} i&=a_12^{n-1}+a_22^{n-2}+\ldots+a_n\\ j&=b_12^{n-1}+b_22^{n-2}+\ldots+b_n\end{aligned}\right\}, \quad (11)$$

where, for $1\leq m\leq n$, a_m, $b_m\in\{0,1\}$. Let $\mathbf{a}=\mathbf{a}(i)=(a_1,\ldots,a_n)$ and $\mathbf{b}=\mathbf{b}(j)=(b_1,\ldots,b_n)$. Further, let $\tau_1,\tau_2,\ldots,\tau_n\in\{0,1,2,3\}$ be the solution of the following system of equations

$$\left.\begin{aligned}
&F[\varnothing;(a_1,\ldots,a_{n-1}),(b_1,\ldots,b_{n-1})](a_n,b_n)=\tau_1,\\
&F[\tau_1;(a_1,\ldots,a_{n-2}),(b_1,\ldots,b_{n-2})](a_{n-1},b_{n-1})=\tau_2,\\
&\quad\cdots\\
&F[(\tau_1,\ldots,\tau_l);(a_1,\ldots,a_{n-l-1}),(b_1,\ldots,b_{n-l-1})](a_{n-l},b_{n-l})=\tau_{l+1},\\
&\quad\cdots\\
&F[(\tau_1,\ldots,\tau_{n-1});\varnothing,\varnothing](a_1,b_1)=\tau_n.
\end{aligned}\right\} \quad (12)$$

Now let

$$G(i,j)=L=\tau_1 4^{n-1}+\tau_2 4^{n-2}+\ldots+\tau_n+1\in\{1,2,\ldots,N\}. \tag{13}$$

For $0\leq i,j<2^n$, let $\mathbf{q}(i,j)$ be a point in the square $S(i,j)$. Using G we can define a permutation $\mathbf{q}_1,\ldots,\mathbf{q}_N$ of the points $\mathbf{q}(i,j)$ $(0\leq i,j<2^n)$ as follows. For $L=1,\ldots,N$, write

$$\mathbf{q}_L=\mathbf{q}(i,j) \quad \text{if } G(i,j)=L. \tag{14}$$

For $0\leq l\leq n$ and $0\leq i,j<2^{n-l}$, let

$$S(l;\, i,j)=[2^{l-n}i,\, 2^{l-n}(i+1))\times[2^{l-n}j,\, 2^{l-n}(j+1)).$$

LEMMA 1. *Suppose that l and H are integers satisfying $0\leq l\leq n$ and $0\leq H<4^l$. Then every square $S(l;\, i,j)$ $(0\leq i,j<2^{n-l})$ contains exactly one element of the set*

$$\{\mathbf{q}_t\colon H4^{n-l}<t\leq(H+1)4^{n-l}\}.$$

Proof. Let

$$i=a_1 2^{n-l-1}+a_2 2^{n-l-2}+\ldots+a_{n-l}$$

and

$$j=b_1 2^{n-l-1}+b_2 2^{n-l-2}+\ldots+b_{n-l},$$

where, for $1\leq m\leq n-l$, $a_m, b_m\in\{0,1\}$; and let

$$H=\tau_1 4^{l-1}+\tau_2 4^{l-2}+\ldots+\tau_l,$$

where, for $1\leq r\leq l$, $\tau_r\in\{0,1,2,3\}$. Let $\tau_{l+1}, \tau_{l+2},\ldots,\tau_n\in\{0,1,2,3\}$ be the solution of the following system of equations

$$F[(\tau_1,\ldots,\tau_l);(a_1,\ldots,a_{n-l-1}),(b_1,\ldots,b_{n-l-1})](a_{n-l},b_{n-l})=\tau_{l+1},$$

$$F[(\tau_1,\ldots,\tau_{l+1});(a_1,\ldots,a_{n-l-2}),(b_1,\ldots,b_{n-l-2})](a_{n-l-1},b_{n-l-1})=\tau_{l+2},$$

$$\cdots$$

$$F[(\tau_1,\ldots,\tau_{n-1});\varnothing,\varnothing](a_1,b_1)=\tau_n.$$

Furthermore, let $(a_{n-l+1},b_{n-l+1}), (a_{n-l+2},b_{n-l+2}),\ldots,(a_n,b_n)\in\{0,1\}^2$ be the solution of the following system of equations

$$\left.\begin{aligned}
&F[(\tau_1,\ldots,\tau_{l-1});(a_1,\ldots,a_{n-l}),(b_1,\ldots,b_{n-l})](a_{n-l+1},b_{n-l+1})=\tau_l,\\
&F[(\tau_1,\ldots,\tau_{l-2});(a_1,\ldots,a_{n-l+1}),(b_1,\ldots,b_{n-l+1})](a_{n-l+2},b_{n-l+2})=\tau_{l-1},\\
&\quad\cdots\\
&F[\varnothing;(a_1,\ldots,a_{n-1}),(b_1,\ldots,b_{n-1})](a_n,b_n)=\tau_1.
\end{aligned}\right\} \tag{15}$$

Finally, let

$$i_0=a_1 2^{n-1}+a_2 2^{n-2}+\ldots+a_n,$$

$$j_0=b_1 2^{n-1}+b_2 2^{n-2}+\ldots+b_n,$$

and

$$H_0=\tau_1 4^{n-1}+\tau_2 4^{n-2}+\ldots+\tau_n+1.$$

Clearly

$$H_0 \in (H4^{n-l}, (H+1)4^{n-l}] \quad \text{and} \quad S(i_0, j_0) \subset S(l; i, j).$$

From (11)-(14), it follows that $\mathbf{q}_{H_0} = \mathbf{q}(i_0, j_0) \in S(i_0, j_0)$. This completes the proof of Lemma 1.

We shall denote by $\mathbf{q}(l; i, j; H)$ the one-element set in Lemma 1. In other words, for integers l, i, j, H satisfying the hypotheses of Lemma 1,

$$\mathbf{q}(l; i, j; H) = \{\mathbf{q}_t : H4^{n-l} < t \leq (H+1)4^{n-l}\} \cap S(l; i, j).$$

§4. *Some probabilistic lemmas.* We shall now use some elementary concepts and facts from probability theory (see, for example, Chung [3]), and define a "randomization" of the deterministic points $\mathbf{q}(i, j)$, mappings $F[\boldsymbol{\tau}; \mathbf{u}, \mathbf{v}]$ and G, and the sequence $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_N$ as follows.

(A) For $0 \leq i, j < 2^n$, let $\tilde{\mathbf{q}}(i, j)$ be a random variable (a random point, in fact) uniformly distributed in the square $S(i, j)$. More precisely,

$$\text{Prob}\,(\tilde{\mathbf{q}}(i, j) \in A) = \frac{\mu(A \cap S(i, j))}{\mu(S(i, j))} = N\mu(A \cap S(i, j))$$

for all Borel sets $A \subset \mathbb{R}^2$.

(B) For every integer l satisfying $0 \leq l \leq n-1$, vectors $\boldsymbol{\tau} \in \{0, 1, 2, 3\}^l$, $\mathbf{u} \in \{0, 1\}^{n-l-1}$ and $\mathbf{v} \in \{0, 1\}^{n-l-1}$, let $\tilde{F}[\boldsymbol{\tau}; \mathbf{u}, \mathbf{v}]$ be a uniformly distributed random bijective mapping from $\{0, 1\}^2$ to $\{0, 1, 2, 3\}$. More precisely, if $\pi: \{0, 1\}^2 \to \{0, 1, 2, 3\}$ is one of the $4! = 24$ different (deterministic) bijective mappings, then

$$\text{Prob}\,(\tilde{F}[\boldsymbol{\tau}; \mathbf{u}, \mathbf{v}] = \pi) = \tfrac{1}{24}.$$

(C) Let $\tilde{G}$ be the random bijective mapping from $\{0, 1, \ldots, 2^n - 1\}^2$ to $\{1, 2, \ldots, N = 4^n\}$ defined by (11), $(\widetilde{12})$ and (13), where $(\widetilde{12})$ denotes that in the system $(\widetilde{12})$ of equations, we replace each (deterministic) mapping $F[\boldsymbol{\tau}; \mathbf{u}, \mathbf{v}]$ by the random mapping $\tilde{F}[\boldsymbol{\tau}; \mathbf{u}, \mathbf{v}]$.

(D) Let $\tilde{\mathbf{q}}_1, \tilde{\mathbf{q}}_2, \ldots, \tilde{\mathbf{q}}_N$ denote the random sequence defined by $\tilde{G}$, *i.e.* for $L = 1, \ldots, N$,

$$\tilde{\mathbf{q}}_L = \tilde{\mathbf{q}}(i, j) \quad \text{if } \tilde{G}(i, j) = L; \tag{$\widetilde{14}$}$$

and let $\tilde{\mathbf{q}}(l; i, j; H)$ denote the randomization of $\mathbf{q}(l; i, j; H)$, *i.e.* for integers l, i, j, H satisfying the hypotheses of Lemma 1,

$$\tilde{\mathbf{q}}(l; i, j; H) = \{\tilde{\mathbf{q}}_t : H4^{n-l} < t \leq (H+1)4^{n-l}\} \cap S(l; i, j).$$

(E) Finally, assume that the random variables $\tilde{\mathbf{q}}(i, j)$ $(0 \leq i, j < 2^n)$ and

$$\tilde{F}[\boldsymbol{\tau}; \mathbf{u}, \mathbf{v}] \qquad (0 \leq l \leq n-1, \boldsymbol{\tau} \in \{0, 1, 2, 3\}^l, \mathbf{u} \in \{0, 1\}^{n-l-1}, \mathbf{v} \in \{0, 1\}^{n-l-1})$$

are independent of each other (the existence of such a set of random variables follows immediately from Kolmogorov's extension theorem in probability theory).

Let $(\Omega, \mathcal{F}, \text{Prob})$ denote the underlying probability measure space.

We shall prove (10) in the following way: There is a positive absolute constant c_{10} such that

$$\text{Prob}\left(\frac{1}{N}\sum_{M=1}^{N}\int_0^{1/2}\int_{U^2}|Z(\tilde{\mathbf{q}}_1, \dots, \tilde{\mathbf{q}}_M; r, \mathbf{x}) - M\mu(B(r, \mathbf{x}))|^2 d\mathbf{x}dr \leq c_{10}N^{1/2}\right) > 0. \tag{16}$$

Inequality (16) immediately implies the existence of a sequence $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N \in U^2$ satisfying (10), and so Theorem 4 follows in the case $K=2$. We need

LEMMA 2. *Suppose that l and H are integers satisfying $0 \leq l \leq n$ and $0 \leq H < 4^l$. Then the random point $\tilde{\mathbf{q}}(l; i, j; H)$ $(0 \leq i, j < 2^{n-l})$ is uniformly distributed in the square $S(l; i, j)$.*

Proof. Let

$$i = a_1 2^{n-l-1} + a_2 2^{n-l-2} + \dots + a_{n-l},$$

and

$$j = b_1 2^{n-l-1} + b_2 2^{n-l-2} + \dots + b_{n-l},$$

where, for $1 \leq m \leq n-l$, $a_m, b_m \in \{0, 1\}$; and let

$$H = \tau_1 4^{l-1} + \tau_2 4^{l-2} + \dots + \tau_l,$$

where, for $1 \leq r \leq l$, $\tau_r \in \{0, 1, 2, 3\}$. Since the random mapping

$$\tilde{F}[(\tau_1, \dots, \tau_{l-1}); (a_1, \dots, a_{n-l}), (b_1, \dots, b_{n-l})]$$

is uniformly distributed, it follows that the (random) solution $(\tilde{a}_{n-l+1}, \tilde{b}_{n-l+1})$ of the (random) equation (see (15))

$$\tilde{F}[(\tau_1, \dots, \tau_{l-1}); (a_1, \dots, a_{n-l}), (b_1, \dots, b_{n-l})](\tilde{a}_{n-l+1}, \tilde{b}_{n-l+1}) = \tau_l$$

has the following property. For $\delta_1, \delta_2 \in \{0, 1\}$,

$$\text{Prob}\,((\tilde{a}_{n-l+1}, \tilde{b}_{n-l+1}) = (\delta_1, \delta_2)) = \tfrac{1}{4}.$$

In other words, the random variables $\tilde{a}_{n-l+1}$ and $\tilde{b}_{n-l+1}$ are independent of each other. Now let $\tilde{a}_{n-l+1} = a_{n-l+1}$, $\tilde{b}_{n-l+1} = b_{n-l+1}$ (*i.e.* fix the values of these random variables), and consider the next (random) equation in (15)

$$\tilde{F}[(\tau_1, \dots, \tau_{l-2}); (a_1, \dots, a_{n-l+1}), (b_1, \dots, b_{n-l+1})](\tilde{a}_{n-l+2}, \tilde{b}_{n-l+2}) = \tau_{l-1}.$$

Since this $\tilde{F}$ is also uniformly distributed, we have (conditional probability) that for $\delta_1, \delta_2 \in \{0, 1\}$,

$$\text{Prob}\,((\tilde{a}_{n-l+2}, \tilde{b}_{n-l+2}) = (\delta_1, \delta_2) | \tilde{a}_{n-l+1} = a_{n-l+1}, \tilde{b}_{n-l+1} = b_{n-l+1}) = \tfrac{1}{4}.$$

In other words, the random variables $\tilde{a}_{n-l+1}, \tilde{b}_{n-l+1}, \tilde{a}_{n-l+2}, \tilde{b}_{n-l+2}$ are independent. Repeating this argument, we conclude that $\tilde{a}_{n-l+1}, \dots, \tilde{a}_n, \tilde{b}_{n-l+1}, \dots, \tilde{b}_n$ are independent random variables with common distribution function

($n-l+1\leq m\leq n$ and $\delta\in\{0,1\}$)

$$\text{Prob}\,(\tilde{a}_m=\delta)=\text{Prob}\,(\tilde{b}_m=\delta)=\tfrac{1}{2}.$$

Let

$$\tilde{i}_0=\tilde{a}_{n-l+1}2^{l-1}+\tilde{a}_{n-l+2}2^{l-2}+\ldots+\tilde{a}_n,$$

and

$$\tilde{j}_0=\tilde{b}_{n-l+1}2^{l-1}+\tilde{b}_{n-l+2}2^{l-2}+\ldots+\tilde{b}_n.$$

Then both $\tilde{i}_0$ and $\tilde{j}_0$ are uniformly distributed on the set $\{0,1,2,\ldots,2^{l-1}\}$, and are independent of each other. Write

$$\tilde{i}_1=a_12^{n-1}+a_22^{n-2}+\ldots+a_{n-l}2^l+\tilde{i}_0,$$

and

$$\tilde{j}_1=b_12^{n-1}+b_22^{n-2}+\ldots+b_{n-l}2^l+\tilde{j}_0.$$

Then

$$\tilde{\mathbf{q}}(l;\,i,j;\,H)=\tilde{\mathbf{q}}(\tilde{i}_1,\tilde{j}_1).$$

Let $S(i_2,j_2)\subset S(l;\,i,j)$. Then clearly

$$\text{Prob}\,(\tilde{\mathbf{q}}(l;\,i,j;\,H)=\tilde{\mathbf{q}}(i_2,j_2))=\text{Prob}\,((\tilde{i}_1,\tilde{j}_1)=(i_2,j_2))=4^{-l}. \qquad (17)$$

Since $\tilde{\mathbf{q}}(i_2,j_2)$ is uniformly distributed in $S(i_2,j_2)$ and the random variables $\tilde{i}_1,\tilde{j}_1,\tilde{\mathbf{q}}(i_2,j_2)$ are independent of each other, Lemma 2 follows.

Let r $(0\leq r\leq\frac{1}{2})$ and $\mathbf{x}\in U^2$ be given. For integers l and H satisfying $0\leq l\leq n-1$ and $0\leq H<4^l$, consider the random set

$$\tilde{\mathcal{P}}(l,H)=\{\tilde{\mathbf{q}}(l;\,i,j;\,H):0\leq i,j<2^{n-l}\},$$

and write

$$\tilde{D}(l,H)=Z(\tilde{\mathcal{P}}(l,H);\,r,\mathbf{x})-4^{n-l}\mu(B(r,\mathbf{x})).$$

Let

$$T(l,H)=\{(i,j):0\leq i,j<2^{n-l}\quad\text{and}\quad S(l;\,i,j)\cap B^*(r,\mathbf{x})\neq\emptyset\quad\text{and}\quad S(l;\,i,j)\backslash B^*(r,\mathbf{x})\neq\emptyset\}.$$

It is easy to see that

$$\#\,T(l,H)\leq 4.2^{n-l}. \qquad (18)$$

Since every square $S(l;\,i,j)$ contains exactly one element (namely $\tilde{\mathbf{q}}(l;\,i,j;\,H)$) of the (random) set $\tilde{\mathcal{P}}(l,H)$, we have

$$\tilde{D}(l,H)=\sum_{\substack{(i,j)\in T(l,H)\\ \tilde{\mathbf{q}}(l;i,j;H)\in B^*(r,\mathbf{x})}}1-4^{n-l}\sum_{(i,j)\in T(l,H)}\mu(S(l;\,i,j)\cap B^*(r,\mathbf{x})).$$

For every $(i,j)\in T(l,H)$, let

$$\xi(l;\,i,j;\,H)=\begin{cases}1 & (\tilde{\mathbf{q}}(l;\,i,j;\,H)\in B^*(r,\mathbf{x})),\\ 0 & (\text{otherwise}).\end{cases}$$

From Lemma 2, $\tilde{\mathbf{q}}(l;\ i, j;\ H)$ is uniformly distributed in $S(l;\ i, j)$, and so, writing $\mathbb{E}$ to denote "expected value",

$$\mathbb{E}\xi(l;\ i, j;\ H) = \frac{\mu(S(l;\ i, j) \cap B^*(r, \mathbf{x}))}{\mu(S(l;\ i, j))} = 4^{n-l}\mu(S(l;\ i, j) \cap B^*(r, \mathbf{x})).$$

We can therefore write

$$\tilde{D}(l, H) = \sum_{(i,j) \in T(l,H)} \eta(l;\ i, j;\ H), \tag{19}$$

where

$$\eta(l;\ i, j;\ H) = \xi(l;\ i, j;\ H) - \mathbb{E}\xi(l;\ i, j;\ H).$$

Clearly $\mathbb{E}\eta = 0$ and $|\eta| \leqslant 1$.

We need

LEMMA 3. *Suppose that* $0 \leqslant l_1, l_2 \leqslant n$, $0 \leqslant H_1 < 4^{l_1}$, $0 \leqslant H_2 < 4^{l_2}$, $0 \leqslant i_1, j_1 < 2^{n-l_1}$ *and* $0 \leqslant i_2, j_2 < 2^{n-l_2}$. *Suppose further that either*

(α) $l_1 = l_2$ *and* $(i_1, j_1) \neq (i_2, j_2)$; *or*

(β) $l_1 < l_2$.

Then

$$|\mathbb{E}(\eta(l_1;\ i_1, j_1;\ H_1)\eta(l_2;\ i_2, j_2;\ H_2))| \leqslant \frac{\mu(S(l_1;\ i_1, j_1) \cap S(l_2;\ i_2, j_2))}{\mu(S(l_2;\ i_2, j_2))}.$$

For conveience of notation, write

$$S_1 = S(l_1;\ i_1, j_1) \quad \text{and} \quad S_2 = S(l_2;\ i_2, j_2),$$

$$\xi_1 = \xi(l_1;\ i_1, j_1;\ H_1) \quad \text{and} \quad \xi_2 = \xi(l_2;\ i_2, j_2;\ H_2),$$

$$\eta_1 = \xi_1 - \mathbb{E}\xi_1 \quad \text{and} \quad \eta_2 = \xi_2 - \mathbb{E}\xi_2.$$

The proof of Lemma 3 depends on the following

LEMMA 4. *Let* $l_1, l_2, i_1, j_1, i_2, j_2$ *be defined as in Lemma* 3. *Suppose further that* $0 \leqslant i^*, j^* < 2^n$ *and*

$$S(i^*, j^*) \subset S_2 \backslash S_1.$$

Consider the event A defined by

$$A = A(l_2, i_2, j_2, H_2, i^*, j^*) \Leftrightarrow \tilde{\mathbf{q}}(l_2;\ i_2, j_2;\ H_2) \in S(i^*, j^*). \tag{20}$$

Then, writing

$$\xi^* = \xi(0;\ i^*, j^*;\ 0) = \begin{cases} 1 & (\tilde{\mathbf{q}}(i^*, j^*) \in B^*(r, \mathbf{x})), \\ 0 & (\textit{otherwise}), \end{cases}$$

we have

$$\mathbb{E}(\xi_1\xi_2|A) = \mathbb{E}(\xi_1)\mathbb{E}(\xi^*). \tag{21}$$

Proof of Lemma 3. let

$$V = \{(i, j):\ S(i, j) \subset S_2 \backslash S_1\},$$

and consider the event

$$A(i,j) \Leftrightarrow \tilde{\mathbf{q}}(l_2;\, i_2, j_2;\, H_2) \in S(i,j) \Leftrightarrow \tilde{\mathbf{q}}(l_2;\, i_2, j_2;\, H_2) = \tilde{\mathbf{q}}(i,j).$$

Let

$$B = \bigcup_{(i,j)\in V} A(i,j).$$

By (21),

$$\mathbb{E}(\xi_1\xi_2|B) = \sum_{(i,j)\in V} \frac{\text{Prob}\,(A(i,j))}{\text{Prob}\,(B)} \mathbb{E}(\xi_1)\mathbb{E}(\xi(i,j)), \tag{22}$$

where

$$\xi(i,j) = \xi(0;\, i,j;\, 0) = \begin{cases} 1 & (\tilde{\mathbf{q}}(i,j) \in B^*(r,\mathbf{x})), \\ 0 & (\text{otherwise}). \end{cases}$$

By (17),

$$\text{Prob}\,(A(i,j)) = 4^{-l_2},$$

and since the random variable $\tilde{\mathbf{q}}(l;\, i,j;\, H)$ is uniformly distributed in $S(l;\, i,j)$ (see Lemma 2), we have, for $t = 1, 2$, that

$$\mathbb{E}(\xi_t) = \frac{\mu(S_t \cap B^*(r,\mathbf{x}))}{\mu(S_t)} = 4^{n-l_t}\mu(S_t \cap B^*(r,\mathbf{x})) \tag{23}$$

and $(l = 0)$

$$\mathbb{E}\xi(i,j) = \frac{\mu(S(i,j) \cap B^*(r,\mathbf{x}))}{\mu(S(i,j))} = 4^n \mu(S(i,j) \cap B^*(r,\mathbf{x})). \tag{24}$$

Moreover,

$$\text{Prob}\,(B) = \frac{\mu(S_2 \backslash S_1)}{\mu(S_2)}. \tag{25}$$

Combining (22), (23) and (24), we have

$$\begin{aligned} \mathbb{E}(\xi_1\xi_2|B) &= \frac{4^{2n-l_1-l_2}}{\text{Prob}\,(B)} \sum_{(i,j)\in V} \mu(S_1 \cap B^*(r,\mathbf{x}))\mu(S(i,j) \cap B^*(r,\mathbf{x})) \\ &= \frac{4^{2n-l_1-l_2}}{\text{Prob}\,(B)} \mu(S_1 \cap B^*(r,\mathbf{x}))\mu((S_2 \backslash S_1) \cap B^*(r,\mathbf{x})). \end{aligned} \tag{26}$$

Clearly

$$\mathbb{E}(\eta_1\eta_2) = \mathbb{E}(\xi_1\xi_2) - \mathbb{E}(\xi_1)\mathbb{E}(\xi_2), \tag{27}$$

and

$$\mathbb{E}(\xi_1\xi_2) = \text{Prob}\,(B)\mathbb{E}(\xi_1\xi_2|B) + (1 - \text{Prob}\,(B))\mathbb{E}(\xi_1\xi_2|\Omega \backslash B). \tag{28}$$

By (23),

$$\mathbb{E}(\xi_1)\mathbb{E}(\xi_2) = 4^{2n-l_1-l_2}\mu(S_1 \cap B^*(r,\mathbf{x}))\mu(S_2 \cap B^*(r,\mathbf{x})); \tag{29}$$

and so, on combining (26) and (29), we have

$$\mathbb{E}(\xi_1)\mathbb{E}(\xi_2)-\text{Prob}\,(B)\mathbb{E}(\xi_1\xi_2|B)$$
$$=\frac{\mu(S_1\cap B^*(r,\mathbf{x}))\{\mu(S_2\cap B^*(r,\mathbf{x}))-\mu((S_2\backslash S_1)\cap B^*(r,\mathbf{x}))\}}{\mu(S_1)\mu(S_2)}$$
$$=\frac{\mu(S_1\cap B^*(r,\mathbf{x}))\mu((S_1\cap S_2)\cap B^*(r,\mathbf{x}))}{\mu(S_1)\mu(S_2)}. \tag{30}$$

It follows from (27), (28) and (30) that

$$\mathbb{E}(\eta_1\eta_2)=(1-\text{Prob}\,(B))\mathbb{E}(\xi_1\xi_2|\Omega\backslash B)$$
$$-\frac{\mu(S_1\cap B^*(r,\mathbf{x}))\mu((S_1\cap S_2)\cap B^*(r,\mathbf{x}))}{\mu(S_1)\mu(S_2)}. \tag{31}$$

Since $0\leqslant\xi_1,\ \xi_2\leqslant 1$, we have

$$0\leqslant\mathbb{E}(\xi_1\xi_2|\Omega\backslash B)\leqslant 1. \tag{32}$$

Furthermore,

$$0\leqslant\frac{\mu(S_1\cap B^*(r,\mathbf{x}))\mu((S_1\cap S_2)\cap B^*(r,\mathbf{x}))}{\mu(S_1)\mu(S_2)}\leqslant\frac{\mu(S_1\cap S_2)}{\mu(S_2)}. \tag{33}$$

On combining (25), (31), (32) and (33), we conclude that

$$-\frac{\mu(S_1\cap S_2)}{\mu(S_2)}\leqslant\mathbb{E}(\eta_1\eta_2)\leqslant 1-\text{Prob}\,(B)=\frac{\mu(S_1\cap S_2)}{\mu(S_2)},$$

and Lemma 3 follows.

We conclude this section by proving Lemma 4. Let

$$i^*=a_12^{n-1}+a_22^{n-2}+\ldots+a_n,$$

and

$$j^*=b_12^{n-1}+b_22^{n-2}+\ldots+b_n,$$

where, for $1\leqslant m\leqslant n$, $a_m,\ b_m\in\{0,1\}$, and let

$$H_2=\tau_14^{l_2-1}+\tau_24^{l_2-2}+\ldots+\tau_{l_2},$$

where, for $1\leqslant r\leqslant l_2$, $\tau_r\in\{0,1,2,3\}$. Event A is equivalent to the fulfillment of the following system of (random) equations (see the proof of Lemma 2).

$$\left.\begin{aligned}&(\tilde{F}[(\tau_1,\ldots,\tau_{l_2-1});(a_1,\ldots,a_{n-l_2}),(b_1,\ldots,b_{n-l_2})])^{-1}(\tau_{l_2})\\&\qquad=(a_{n-l_2+1},b_{n-l_2+1}),\\&(\tilde{F}[(\tau_1,\ldots,\tau_{l_2-2});(a_1,\ldots,a_{n-l_2+1}),(b_1,\ldots,b_{n-l_2+1})])^{-1}(\tau_{l_2-1})\\&\qquad=(a_{n-l_2+2},b_{n-l_2+2}),\\&\quad\cdots\\&\quad(\tilde{F}[\varnothing;(a_1,\ldots,a_{n-1}),(b_1,\ldots,b_{n-1})])^{-1}(\tau_1)=(a_n,b_n).\end{aligned}\right\} \tag{34}$$

Write

$$i_1 = c_1 2^{n-l_1-1} + c_2 2^{n-l_1-2} + \ldots + c_{n-l_1},$$

and

$$j_1 = d_1 2^{n-l_1-1} + d_2 2^{n-l_1-2} + \ldots + d_{n-l_1},$$

where, for $1 \leq m \leq n - l_1$, $c_m, d_m \in \{0, 1\}$, and let

$$H_1 = \lambda_1 4^{l_1-1} + \lambda_2 4^{l_1-2} + \ldots + \lambda_{l_1},$$

where, for $1 \leq r \leq l_1$, $\lambda_r \in \{0, 1, 2, 3\}$. The random variable $\tilde{\mathbf{q}}(l_1; i_1, j_1; H_1)$ depends only on the following random variables (and is independent of all the others): the random mappings

$$\left.\begin{aligned} &\tilde{F}[(\lambda_1, \ldots, \lambda_{l_1-1}); (c_1, \ldots, c_{n-l_1}), (d_1, \ldots, d_{n-l_1})], \\ &\tilde{F}[(\lambda_1, \ldots, \lambda_{l_1-2}); (c_1, \ldots, c_{n-l_1}, x_1), (d_1, \ldots, d_{n-l_1}, y_1)], \\ &\quad \cdot \quad \cdot \quad \cdot \\ &\tilde{F}[\varnothing; (c_1, \ldots, c_{n-l_1}, x_1, \ldots, x_{l_1-1}), (d_1, \ldots, d_{n-l_1}, y_1, \ldots, y_{l_1-1})], \end{aligned}\right\} \tag{35a}$$

where, for $1 \leq m \leq l_1 - 1$, $x_m, y_m \in \{0, 1\}$; and the random points

$$\{\tilde{\mathbf{q}}(i, j) : S(i, j) \subset S(l_1; i_1, j_1)\}. \tag{35b}$$

Consider now the following three sets of random variables:

$$\left.\begin{aligned} &\{\tilde{\mathbf{q}}(i^*, j^*)\}, \\ &\{\text{random mappings occurring in (34)}\}, \\ &\{\text{random mappings and points occurring in (35)}\}. \end{aligned}\right\} \tag{36}$$

Observe that these three sets are pairwise disjoint. Indeed, if $l_1 = l_2$ and $(i_1, j_1) \neq (i_2, j_2)$, then it is trivial; on the other hand, if $l_1 < l_2$, then the assumption $S(i^*, j^*) \subset S_2 \backslash S_1$ yields

$$(a_1, \ldots, a_{n-l_1}, b_1, \ldots, b_{n-l_1}) \neq (c_1, \ldots, c_{n-l_1}, d_1, \ldots, d_{n-l_1}),$$

and again the result is obvious. Lemma 4 follows immediately from

LEMMA 5. *Let* $(\Omega, \mathscr{F}, \text{Prob})$ *be a probability measure space. Let* $A \in \mathscr{F}$, *and let* ξ_1, ξ_2, ξ^* *be random variables on this space. Suppose that*

$$A \Leftrightarrow \xi_2 = \xi^*.$$

Furthermore, suppose that ξ_1, ξ^* *and* χ_A *(characteristic function of the event* A*) are independent. Then*

$$\mathbb{E}(\xi_1 \xi_2 | A) = \mathbb{E}(\xi_1) \mathbb{E}(\xi^*). \tag{37}$$

Proof. By definition

$$\mathbb{E}(\xi_1\xi_2|A) = \frac{1}{\text{Prob}(A)} \int_A \xi_1(w)\xi_2(w)d\ \text{Prob}(w)$$

$$= \frac{1}{\text{Prob}(A)} \int_A \xi_1(w)\xi^*(w)d\ \text{Prob}(w)$$

$$= \frac{1}{\text{Prob}(A)} \int_\Omega \xi_1(w)\xi^*(w)\chi_A(w)d\ \text{Prob}(w). \tag{38}$$

By the independence of ξ_1, ξ^* and χ_A, we have

$$\int_\Omega \xi_1(w)\xi^*(w)\chi_A(w)d\ \text{Prob}(w) = \mathbb{E}(\xi_1)\mathbb{E}(\xi^*)\mathbb{E}(\chi_A) = \mathbb{E}(\xi_1)\mathbb{E}(\xi^*)\ \text{Prob}(A). \tag{39}$$

(37) follows immediately on combining (38) and (39).

§5. *Completion of the proof.* In this section, we complete the proof of (16). Let M, satisfying $1 \leqslant M < N = 4^n$, be an arbitrary but fixed integer. Write

$$M = \tau_1 4^{n-1} + \tau_2 4^{n-2} + \ldots + \tau_n$$
$$= \tau_{l_1} 4^{n-l_1} + \tau_{l_2} 4^{n-l_2} + \ldots + \tau_{l_m} 4^{n-l_m},$$

where $\tau_{l_1}, \ldots, \tau_{l_m}$ are the non-vanishing coefficients; in other words, $\tau_{l_1}, \ldots, \tau_{l_m} \in \{1, 2, 3\}$, where $1 \leqslant l_1 < \ldots < l_m \leqslant n$. We can write

$$\{\tilde{\mathbf{q}}_1, \ldots, \tilde{\mathbf{q}}_M\} = \bigcup_{t=1}^{m} \bigcup_{h_t=0}^{\tau_{l_t}-1} \tilde{\mathscr{P}}(l_t, H(t, h_t)), \tag{40}$$

where, for $1 \leqslant t \leqslant m$,

$$H(t, h_t) = \sum_{s=1}^{t-1} \tau_{l_s} 4^{l_t - l_s} + h_t.$$

It follows from (19) that

$$Z(\tilde{\mathbf{q}}_1, \ldots, \tilde{\mathbf{q}}_M; r, \mathbf{x}) - M\mu(B(r, \mathbf{x})) = \sum_{t=1}^{m} \sum_{h_t=0}^{\tau_{l_t}-1} \tilde{D}(l_t, H(t, h_t))$$

$$= \sum_{t=1}^{m} \sum_{h_t=0}^{\tau_{l_t}-1} \sum_{(i,j) \in T(l_t, H(t,h_t))} \eta(l_t; i, j; H(t, h_t)). \tag{41}$$

For $t = 1, \ldots, m$, let

$$X_t = \{\eta(l_t; i, j; H(t, h_t)) : 0 \leqslant h_t < \tau_{l_t} \quad \text{and} \quad (i, j) \in T(l_t, H(t, h_t))\},$$

and let

$$X=\bigcup_{t=1}^{m} X_t.$$

By (41), we have

$$\mathbb{E}(Z(\tilde{\mathbf{q}}_1,\ldots,\tilde{\mathbf{q}}_M;\, r,\mathbf{x})-M\mu(B(r,\mathbf{x})))^2=\sum_{\eta_1\in X}\sum_{\eta_2\in X}\mathbb{E}(\eta_1\eta_2)=I_1+2I_2, \tag{42}$$

where

$$I_1=\sum_{t=1}^{m}\sum_{\eta_1,\eta_2\in X_t}\mathbb{E}(\eta_1\eta_2),$$

and

$$I_2=\sum_{1\leqslant s<t\leqslant m}\sum_{\eta_1\in X_s}\sum_{\eta_2\in X_t}\mathbb{E}(\eta_1\eta_2).$$

Consider first I_1. By Lemma 3 and (18),

$$\begin{aligned}|I_1\|&\leqslant\sum_{t=1}^{m}\sum_{h_t=0}^{\tau_{l_t}-1}\sum_{h_t'=0}^{\tau_{l_t}-1}\#(T(l_t;\,H(t,h_t))\cap T(l_t;\,H(t,h_t')))\\&\leqslant\sum_{t=1}^{m}4\tau_{l_t}^2 2^{n-l_t}\leqslant 36\sum_{t=1}^{m}2^{n-l_t}\leqslant 36\sum_{l=1}^{\infty}2^{n-l}\\&=36.2^n=36N^{1/2}.\end{aligned} \tag{43}$$

Suppose now that $1\leqslant l<l'\leqslant n$ and $0\leqslant i,j<2^{n-l}$. Then there is exactly one index pair (i',j'), satisfying $0\leqslant i',j'<2^{n-l'}$, such that $S(l';\,i',j')\cap S(l;\,i,j)=\varnothing$. In fact, we have $S(l;\,i,j)\subset S(l';\,i',j')$, and so

$$\frac{\mu(S(l';\,i',j')\cap S(l;\,i,j))}{\mu(S(l';\,i',j'))}=4^{l-l'}.$$

It follows from Lemma 3 that

$$\begin{aligned}|I_2|&\leqslant\sum_{s=1}^{m}\sum_{\eta\in X_s}\sum_{t=s+1}^{m}\sum_{h_t=0}^{\tau_{l_t}-1}4^{l_s-l_t}\leqslant\sum_{s=1}^{m}\sum_{\eta\in X_s}\sum_{l=1}^{\infty}3.4^{-l}\\&=\sum_{s=1}^{m}\sum_{\eta\in X_s}1=\sum_{s=1}^{m}\sum_{h_s=0}^{\tau_{l_s}-1}\#\,T(l_s;\,H(s,h_s)).\end{aligned} \tag{44}$$

On combining (18) and (44), we conclude that

$$|I_2|\leqslant\sum_{s=1}^{m}\sum_{h_s=0}^{\tau_{l_s}-1}4.2^{n-l_s}\leqslant 12\sum_{s=1}^{m}2^{n-l_s}\leqslant 12\sum_{l=1}^{\infty}2^{n-l}=12.2^n=12N^{1/2}. \tag{45}$$

On combining (42), (43) and (45), we have, for $1\leqslant M<N$,

$$\mathbb{E}(z(\tilde{\mathbf{q}}_1,\ldots,\tilde{\mathbf{q}}_M;\, r,\mathbf{x})-M\mu(B(r,\mathbf{x})))^2\leqslant 60N^{1/2}.$$

Also, a simpler argument applies when $M=N$. Consequently,

$$\mathbb{E}\left(\frac{1}{N}\sum_{M=1}^{N}\int_0^{1/2}\int_{U^2}|Z(\tilde{\mathbf{q}}_1,\ldots,\tilde{\mathbf{q}}_M;\, r,\mathbf{x})-M\mu(B(r,\mathbf{x}))|^2 d\mathbf{x}dr\right)\leqslant 30N^{1/2}.$$

(16) follows immediately with $c_{10}=30$.

References

1. J. Beck. Some upper bounds in the theory of irregularities of distribution. *Acta Arith.*, 43 (1984), 115-130.
2. J. Beck. Irregularities of distribution I. To appear in *Acta Math.*
3. K. L. Chung. *A course in probability theory* (Academic Press, New York, 1974).
4. H. Davenport. Note on irregularities of distribution. *Mathematika*, 3 (1956), 131-135.
5. H. L. Montgomery. Irregularities of distribution by means of power sums. Manuscript submitted to the *Proceedings of the Congress de Teoria de los Numeros, Bilbao*, 1984.
6. K. F. Roth. On irregularities of distribution. *Mathematika*, 1 (1954), 73-79.
7. K. F. Roth. On irregularities of distribution III. *Acta Arith.*, 35 (1979), 373-384.
8. K. F. Roth. On irregularities of distribution IV. *Acta Arith.*, 37 (1980), 67-75.
9. W. M. Schmidt. Irregularities of distribution IV. *Invent. Math.*, 7 (1969), 55-82.
10. W. M. Schmidt. *Lectures on irregularities of distribution* (Tata Institute of Fundamental Research, Bombay, 1977).

Dr. J. Beck,
Eötvös Loránd University,
Budapest, Hungary.

Dr. W. W. L. Chen,
Imperial College,
London, SW7 2BZ

10F40: *NUMBER THEORY; Diophantine Approximation; Distribution modulo one.*

Received on the 12*th of November*, 1985.