Irregularities of Distributions with Respect to Polytopes

Michael Drmota

Dedicated to Prof. Edmund Hlawka on the occasion of his 80^{th} birthday

Abstract

In the first part of the paper we show that that the L^2-discrepancy with respect to squares is of the same order of magnitude as the usual L^2-discrepancy for point distributions in the k-dimensional torus.

In the second part we adapt this method to obtain a generalization of Roth's [7] lower bound $(\log N)^{(k-1)/2}$ (for the usual discrepancy) to the discrepancy with respect to homothetic simple convex polytopes.

1 Introduction

Let $\mathbf{x}_1, \ldots, \mathbf{x}_N \in \mathbf{R}^k/\mathbf{Z}^k$ be N points in the k-dimensional torus and for any measurable set $S \subseteq \mathbf{R}^k/\mathbf{Z}^k$ set

$$\Delta_N(\mathbf{x}_n, S) = \sum_{n=1}^{N} 1_S(\mathbf{x}_n) - N\lambda_k(S), \tag{1}$$

where 1_S denotes the characteristic function of the set S and λ_k the Lebesgue measure. (Obviously, $\Delta_N(\mathbf{x}_n, \cdot)$ is a signed measure.) For $\mathbf{x} = (x_1, \ldots, x_k)$, $x_j \geq 0$, let $[0, \mathbf{x})$ denote the k-dimensional box $[0, x_1) \times \cdots \times [0, x_k)$. Then the usual discrepancy D_N of $\mathbf{x}_1, \ldots, \mathbf{x}_N$ is defined by

$$D_N(\mathbf{x}_n) = \sup_{\mathbf{x} \in [0,1]^k} |\Delta_N(\mathbf{x}_n, [0, \mathbf{x}))| \tag{2}$$

and the L^2-discrepancy $D_N^{(2)}$ by

$$D_N^{(2)}(\mathbf{x}_n) = \left(\int_{[0,1]^k} (\Delta_N(\mathbf{x}_n, [0, \mathbf{x})))^2 \, d\mathbf{x} \right)^{\frac{1}{2}}. \tag{3}$$

(Sometimes $\frac{1}{N}\sup|\Delta_N(\mathbf{x}_n,[0,\mathbf{x}))|$ is called discrepancy and is also denoted by D_N [5, 3]. Furthermore observe that the value of D_N resp. that of $D_N^{(2)}$ will not change if we use closed boxes $[0,\mathbf{x}]$ instead of the half open boxes $[0,\mathbf{x})$.)

Roth [7] proved that there are constants $c_k>0$ such that

$$D_N(\mathbf{x}_n) \ge D_N^{(2)}(\mathbf{x}_n) \ge c_k(\log N)^{\frac{1}{2}(k-1)}. \tag{4}$$

In [1] these lower bounds are generalized in the following sense. The corresponding cube discrepancies

$$D_N(\mathbf{x}_n,\mathcal{Q}) = \sup_{\mathbf{y}\in[0,1)^k, 0\le\rho<1} |\Delta_N(\mathbf{x}_n, Q(\rho)+\mathbf{y})| \tag{5}$$

and

$$D_N^{(2)}(\mathbf{x}_n,\mathcal{Q}) = \left(\int_0^1\int_{[0,1]^k}(\Delta_N(\mathbf{x}_n,Q(\rho)+\mathbf{y}))^2\,d\mathbf{y}\,d\rho\right)^{\frac{1}{2}}, \tag{6}$$

where $Q(\rho)=[-\frac{1}{2}\rho,\frac{1}{2}\rho]^k$ denotes a cube of side length ρ, satisfy

$$D_N(\mathbf{x}_n,\mathcal{Q}) \ge D_N^{(2)}(\mathbf{x}_n,\mathcal{Q}) \ge c_k'(\log N)^{\frac{1}{2}(k-1)} \tag{7}$$

for some constants $c_k'>0$.

Therefore it is an interesting question to compare $D_N(\mathbf{x}_n)$ with $D_N(\mathbf{x}_n,\mathcal{Q})$ directly. Obviously, we always have $D_N(\mathbf{x}_n)\ge D_N(\mathbf{x}_n,\mathcal{Q})$. Furthermore, in the two dimensional case $k=2$ Ruzsa [8] obtained the remarkable relation

$$D_N(\mathbf{x}_n,\mathcal{Q}) \ge \frac{1}{25}D_N(\mathbf{x}_n). \tag{8}$$

However, his proof does not work in higher dimensions. In section 2 we will show that an L^2-analogue holds in any dimension.

Theorem 1 *Let $k\ge 2$ and $\mathbf{x}_1,\dots,\mathbf{x}_N\in\mathbf{R}^k/\mathbf{Z}^k$. Then*

$$D_N^{(2)}(\mathbf{x}_n,\mathcal{Q}) \gg\ll D_N^{(2)}(\mathbf{x}_n), \tag{9}$$

where the constants implied by $\gg\ll$ only depend on k.

As a corollary we immediately obtain by an application of an result by Niederreiter, Tichy, and Turnwald [6]:

Theorem 2 *Let $k \geq 2$ and $\mathbf{x}_1, \ldots, \mathbf{x}_N \in \mathbf{R}^k/\mathbf{Z}^k$. Then*

$$D_N(\mathbf{x}_n, \mathcal{Q}) \gg N^{-\frac{k}{2}} D_N(\mathbf{x}_n)^{1+\frac{k}{2}}, \tag{10}$$

where the constant implied by $\gg$ only depends on k.

This result is worse than Ruzsa's [8] in the case $k = 2$ but it seems to be the first quantitative relation between $D_N(\mathbf{x}_n, \mathcal{Q})$ and $D_N(\mathbf{x}_n)$ in arbitrary dimensions which is better than the quantitative relation between the isotropic discrepancy $J_N(\mathbf{x}_n)$ and $D_N(\mathbf{x}_n)$ (see [5]).

In section 3 we will generalize (4) to discrepancies with respect to arbitrary simple convex polytopes P. (A k-dimensional polytope is simple if every vertex of P has exactly k adjacent ones.) Similarly to (5) and (6) we can define discrepancies

$$D_N(\mathbf{x}_n, \mathcal{P}) = \sup_{\mathbf{y}\in[0,1)^k, 0\leq\rho<1} |\Delta_N(\mathbf{x}_n, \rho P + \mathbf{y})|$$

and

$$D_N^{(2)}(\mathbf{x}_n, \mathcal{P}) = \left(\int_0^1 \int_{[0,1]^k} (\Delta_N(\mathbf{x}_n, \rho P + \mathbf{y}))^2 \, d\mathbf{y} \, d\rho \right)^{\frac{1}{2}}.$$

Theorem 3 *Let $k \geq 2$ and $P \subseteq [0,1)^k$ a simple convex polytope. Then there exists a constant $c > 0$ such that for any finite point set $\mathbf{x}_1, \ldots, \mathbf{x}_N \in \mathbf{R}^k/\mathbf{Z}^k$*

$$D_N(\mathbf{x}_n, \mathcal{P}) \geq D_N^{(2)}(\mathbf{x}_n, \mathcal{P}) \geq c\,(\log N)^{\frac{1}{2}(k-1)}. \tag{11}$$

The method for proving Theorem 3 is a combination of the ideas used in the proof of Theorem 1 and a result by Károlyi [4].

It is also an interesing question to decide whether the lower bound (11) is optimal or not. In [4] it is proved that for any $N \geq 1$ there exists a point distribution $\mathbf{x}_1, \ldots, \mathbf{x}_N$ such that

$$D_N(\mathbf{x}_n, \mathcal{P}) \ll (\log N)^{\max\left(\frac{3}{2}k+\varepsilon, 2k-1\right)}.$$

This means that the order of magnitude of the smallest possible discrepancy is surely a power of $\log N$ which is indeed very small. Namely, the situation changes drastically if the underlying system of homothetic testing sets is generated by a convex body C with a sufficiently smooth boundary ∂C. In [2] it is shown that in this case

$$D_N(\mathbf{x}_n, \mathcal{C}) = \sup_{\mathbf{y}\in[0,1)^2, 0\leq\rho<1} |\Delta(\mathbf{x}_n, \rho C + \mathbf{y})| \gg N^{\frac{1}{2}-\frac{1}{2k}}.$$

This lower bound is optimal despite of a logarithmic factor $\sqrt{\log N}$

It should be further mentioned that in the 2-dimensional case $k = 2$ there exist much more precise results. Schmidt [9] showed that

$$D_N(\mathbf{x}_n) \gg \log N$$

and Halász (see [1]) improved this bound to

$$D_N(\mathbf{x}_n, Q) \gg \log N.$$

Both bounds are optimal. Moreover, Beck [1] provided a bound of the form

$$D_N(\mathbf{x}_n, \mathcal{C}) \gg (\log N)^{\frac{1}{2}}$$

for arbitrary convex sets $C \subseteq [0,1)^2$. Therefore Theorem 3 is only new for dimensions $k > 2$.

2 A Variation of Beck's Method

First, let us note that for technical reasons we will work with

$$\tilde{D}_N^{(2)}(\mathbf{x}_n) = \left(\int_{[0,1)^k} \int_{\mathbf{R}^k/\mathbf{Z}^k} (\Delta_N(\mathbf{x}_n, R(\mathbf{a}) + \mathbf{y}))^2 \, d\mathbf{y} \, d\mathbf{a} \right)^{\frac{1}{2}} \tag{12}$$

instead of $D_N^{(2)}(\mathbf{x}_n)$, where $R(\mathbf{a})$ denotes the rectangle $R(\mathbf{a}) = [-\frac{a_1}{2}, \frac{a_1}{2}] \times \cdots \times [-\frac{a_k}{2}, \frac{a_k}{2}]$, $\mathbf{a} = (a_1, \ldots, a_k) \in [0,1)^k$. By obvious reasoning they are related by the inequalities

$$D_N^{(2)}(\mathbf{x}_n) \le 4^k \tilde{D}_N^{(2)}(\mathbf{x}_n) \quad \text{and} \quad \tilde{D}_N^{(2)}(\mathbf{x}_n) \le 4^k D_N^{(2)}(\mathbf{x}_n).$$

Furthermore, we two sets A, B are called homothetic if there is a shift $\mathbf{y}$ and a factor $\rho > 0$ such that $A = \rho B + \mathbf{y}$.

We will indentify $\mathbf{R}^k/\mathbf{Z}^k$ with the unit cube $[0,1)^k \subseteq \mathbf{R}^k$ and consider the periodically expanded discrete set

$$\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\} + \mathbf{Z}^k \subseteq \mathbf{R}^k$$

instead of $\mathbf{x}_1, \ldots, \mathbf{x}_N \in [0,1)^k$. The distribution properties of $\mathbf{x}_1, \ldots, \mathbf{x}_N \in \mathbf{R}^k/\mathbf{Z}^k$ with respect to homothetic sets $S \subseteq \mathbf{R}^k/\mathbf{Z}^k$ are exactly the same as the distribution properties of $\mathbf{X}$ with respect to homothetic sets $S \subseteq \mathbf{R}^k$.

We only have to take care that S is not too large, i.e. $S \subseteq [0,1)^k + \mathbf{y}$ for some $\mathbf{y} \in \mathbf{R}^k$.

Let $Q(\rho)$, $\rho \in [0,1)$, and $R(\mathbf{a})$, $\mathbf{a} \in [0,1)^k$ be defined as above. For any positive integer $M \geq 1$ let $Q_M = [-M, M]^k$ and set

$$\Delta(M, S) = \sum_{\mathbf{x} \in \mathbf{X}} 1_{S \cap Q_M}(\mathbf{x}) - N\lambda_k \left(S \cap Q_M\right).$$

Furthermore define

$$\Delta^{(2)}(M, \mathcal{Q}) = (2M)^{-k} \int_0^1 \int_{\mathbf{R}^k} \left(\Delta(M, Q(\rho) + \mathbf{y})\right)^2 \, d\mathbf{y} \, d\rho$$

and

$$\Delta^{(2}(M, \mathcal{R}) = (2M)^{-k} \int_{[0,1)^k} \int_{\mathbf{R}^k} \left(\Delta(M, R(\mathbf{a}) + \mathbf{y})\right)^2 \, d\mathbf{y} \, d\mathbf{a}.$$

Using rather crude estimates around the boundary of Q_M we immediately get

$$\lim_{M \to \infty} \Delta^{(2)}(M, \mathcal{Q}) = \int\limits_0^1 \int\limits_{\mathbf{R}^k/\mathbf{Z}^k} \left(\Delta_N(\mathbf{x}_n, Q(\rho) + \mathbf{y})\right)^2 \, d\mathbf{y} \, d\rho = D_N^{(2)}(\mathbf{x}_n, \mathcal{Q})^2 \tag{13}$$

and

$$\lim_{M \to \infty} \Delta^{(2)}(M, \mathcal{R}) = \int\limits_{[0,1)^k} \int\limits_{\mathbf{R}^k/\mathbf{Z}^k} \left(\Delta_N(\mathbf{x}_n, R(\mathbf{a}) + \mathbf{y})\right)^2 \, d\mathbf{y} \, d\mathbf{a} = D_N^{(2)}(\mathbf{x}_n, \mathcal{R})^2. \tag{14}$$

Since $\Delta(M, S + \mathbf{y})$ can be rewritten as a convolution

$$\Delta(M, S + \mathbf{y}) = 1_{(-S)} * (d\Delta(M, \mathbf{y})),$$

we obtain by using Parselval's identity (and by observing that $-Q(\rho) = Q(\rho)$ and $-R(\mathbf{a}) = R(\mathbf{a})$)

$$\begin{aligned}
\Delta^{(2)}(M, \mathcal{Q}) &= (2M)^{-k} \int_0^1 \int_{\mathbf{R}^k} \left|(1_{Q(\rho)} * (d\Delta(M, \mathbf{y})\right|^2 \, d\mathbf{y} \, d\rho \\
&= (2M)^{-k} \int_0^1 \int_{\mathbf{R}^k} \left|\hat{1}_{Q(\rho)}(\mathbf{t})\right|^2 \cdot \left|(d\widehat{\Delta(M}, \cdot)(\mathbf{t})\right|^2 \, d\mathbf{t} \, d\rho \\
&= (2M)^{-k} \int_{\mathbf{R}^k} \left(\int_0^1 \left|\hat{1}_{Q(\rho)}(\mathbf{t})\right|^2 \, d\rho\right) \left|(d\widehat{\Delta(M}, \cdot))(\mathbf{t})\right|^2 \, d\mathbf{t}
\end{aligned}$$

and

$$\Delta^{(2)}(M,\mathcal{R}) = (2M)^{-k}\int_{\mathbf{R}^k}\left(\int_{[0,1)^k}\left|\hat{1}_{R(\mathbf{a})}(\mathbf{t})\right|^2\,d\mathbf{a}\right)\left|(d\widehat{\Delta(M},\cdot))(\mathbf{t})\right|^2\,d\mathbf{t},$$

where

$$\hat{f}(\mathbf{t}) = (2\pi)^{-\frac{k}{2}}\int_{\mathbf{R}^k} f(y)e^{-i\mathbf{t}\cdot\mathbf{y}}\,d\mathbf{y}$$

denotes the Fourier transform of a function $f(\mathbf{y})$ and

$$(d\widehat{\Delta(M},\cdot))(\mathbf{t}) = (2\pi)^{-\frac{k}{2}}\int_{\mathbf{R}^k} e^{-i\mathbf{t}\cdot\mathbf{y}}\,d\Delta(M,\mathbf{y})$$

denotes the Fourier transform of the (signed) measure $\Delta(M,\cdot)$.

The final step of the proof is to compare the quadratic mean values of

$$\hat{1}_{Q(\rho)}(\mathbf{t}) = \prod_{i=1}^{k}\frac{2\sin(\rho t_i)}{\sqrt{2\pi}t_i} \quad \text{and} \quad \hat{1}_{R(\mathbf{a})}(\mathbf{t}) = \prod_{i=1}^{k}\frac{2\sin(a_i t_i)}{\sqrt{2\pi}t_i}.$$

In [1, p. 140] it is mentioned that

$$\int_0^1\left|\hat{1}_{Q(\rho)}(\mathbf{t})\right|^2\,d\rho \gg \prod_{i=1}^{k}\frac{1}{(1+|t_i|)^2}$$

uniformly for $\mathbf{t}\in\mathbf{R}^k/\mathbf{Z}^k$ (compare also with Lemma 2). On the other hand, by using the inequality $|\sin x|\le\min\{1,|x|\}$, we get

$$\int_0^1\left|\hat{1}_{Q(\rho)}(\mathbf{t})\right|^2\,d\rho \ll \prod_{i=1}^{k}\frac{1}{(1+|t_i|)^2}$$

In the same fashion (ever more easily) we obtain

$$\int_{[0,1)^k}\left|\hat{1}_{R(\mathbf{a})}(\mathbf{t})\right|^2\,d\mathbf{a} \gg\ll \prod_{i=1}^{k}\frac{1}{(1+|t_i|)^2}$$

which implies

$$\int_0^1\left|\hat{1}_{Q(\rho)}(\mathbf{t})\right|^2\,d\rho \gg\ll \int_{[0,1)^k}\left|\hat{1}_{R(\mathbf{a})}(\mathbf{t})\right|^2\,d\mathbf{a}$$

and consequently

$$\Delta^{(2)}(M,\mathcal{Q}) \gg\ll \Delta^{(2)}(M,\mathcal{R}). \tag{15}$$

Hence (13), (14), and (15) prove Theorem 1.

3 An Extension to Convex Polytopes

The basic tools for the proof of Theorem 3 are the following two lemmata.

Lemma 1 *Let $k \geq 1$ and $P \subseteq \mathbf{R}^k$ a k-dimensional convex simple polytope. Then*

$$\hat{1}_P(\mathbf{t}) = (2\pi)^{-\frac{k}{2}} \sum_{\mathbf{x} \in V(P)} \frac{\left|\det(\mathbf{x}-\mathbf{y})_{\mathbf{y}\in\Gamma(\mathbf{x})}\right|}{\prod_{\mathbf{y}\in\Gamma(\mathbf{x})}(-i\mathbf{t}\cdot(\mathbf{x}-\mathbf{y}))} e^{-i\mathbf{t}\cdot\mathbf{x}}, \tag{16}$$

where $V(P)$ denotes the set of vertices of P and $\Gamma(\mathbf{x})$ the set of vertices adjacent to $\mathbf{x} \in V(P)$.

Proof. First, it is an easy exercise to verify (16) for k-dimensional simplices. Next, observe that $\hat{1}_P(\mathbf{t})$ surely depends only on the vertices of P and that it has the form

$$\hat{1}_P(\mathbf{t}) = \sum_{\mathbf{x}\in V(P)} c_{\mathbf{x}}(\mathbf{t}) e^{-i\mathbf{t}\mathbf{x}}.$$

Now choose any vertex $\mathbf{x}_0 \in V(P)$ and let $P = P' \cup S$, where S is the simplex obtained by taking the convex hull of $\mathbf{x}_0$ and all vertices $\Gamma(\mathbf{x}_0)$ adjacent to $\mathbf{x}_0$ and $P' = \overline{P \setminus S}$. It is clear that

$$\hat{1}_P(\mathbf{t}) = \hat{1}_{P'}(\mathbf{t}) + \hat{1}_S(\mathbf{t})$$

and that $\hat{1}_{P'}(\mathbf{t})$ has no influence on the coefficient $c_{\mathbf{x}_0}(\mathbf{t})$. Hence $c_{\mathbf{x}_0}(\mathbf{t})$ is given by

$$c_{\mathbf{x}_0}(\mathbf{t}) = (2\pi)^{-\frac{k}{2}} \frac{\left|\det(\mathbf{x}_0-\mathbf{y})_{\mathbf{y}\in\Gamma(\mathbf{x}_0)}\right|}{\prod_{\mathbf{y}\in\Gamma(\mathbf{x}_0)}(-i\mathbf{t}\cdot(\mathbf{x}_0-\mathbf{y}))}.$$

This completes the proof of Lemma 1. □

We can use (16) to obtain the following useful estimates.

Lemma 2 *Let $k \geq 1$ and $P \subseteq \mathbf{R}^k$ a k-dimensional convex simple polytope. Then*

$$|\hat{1}_P(\mathbf{t})|^2 \ll \sum_{\mathbf{x}\in V(P)} \prod_{\mathbf{y}\in\Gamma(\mathbf{x})} \frac{1}{(1+|\mathbf{t}\cdot(\mathbf{x}-\mathbf{y})|)^2} \tag{17}$$

and

$$\int_0^1 |\hat{1}_{\rho P}(\mathbf{t})|^2 \, d\rho \gg \sum_{\mathbf{x}\in V(P)} \prod_{\mathbf{y}\in\Gamma(\mathbf{x})} \frac{1}{(1+|\mathbf{t}\cdot(\mathbf{x}-\mathbf{y})|)^2} \tag{18}$$

uniformly for all $\mathbf{t} \in \mathbf{R}^k$, where the constants implied by $\ll$ and $\gg$ only depend on P.

We postpone the proof to the end of this section.

Let P be represented as the intersection $P = \bigcap_i H_i$ of closed halfspaces H_i and let $\tilde{P} = \bigcap_i \tilde{H}_i$ be a little bit pertubated parallel polytope which should mean that the supporting hyperplanes of H_i and $\tilde{H}_i$ are parallel (i.e. $H_i \cap \tilde{H}_i = H_i$ or $H_i \cap \tilde{H}_i = \tilde{H}_i$) and have small distance. Then the bounds for $|\hat{1}_P(\mathbf{t})|^2$ and for $|\hat{1}_{\tilde{P}}(\mathbf{t})|^2$ given by Lemma 2 do not really differ. Essentially the bounds (17) and (18) can be made uniform for a system $\mathcal{P}$ of polytopes $\tilde{P}$ parallel to P which do not differ too much. For example, let $\mathbf{x}_0 \in V(P)$ be a fixed vertex of P, $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\} = \Gamma(\mathbf{x}_0)$ the vertices adjacent to $\mathbf{x}_0$ and H_i, $1 \le i \le k$, the halfspaces supported by the hyperplanes

$$L_i = \left\{ \mathbf{x}_0 + \sum_{j \ne i} \lambda_j (\mathbf{x}_j - \mathbf{x}_0) \; : \; \lambda_j \in \mathbf{R} \right\}$$

which intersection $\bigcap\limits_{i=1}^{k} L_i$ locally describes P. Then the set

$$\tilde{\mathcal{P}} = \left\{ \tilde{P}(\rho_1, \ldots, \rho_k) = P \cap \bigcap_{i=1}^{k} (H_i + \rho_i(\mathbf{x}_i - \mathbf{x}_0)) \; : \; 0 \le \rho_i \le \varepsilon, \, 1 \le i \le k \right\} \tag{19}$$

is surely such a system for sufficiently small $\varepsilon > 0$. Using exactly the same methods as in the proof of Lemma 2 we obtain

$$\max_{\tilde{P} \in \tilde{\mathcal{P}}} |\hat{1}_{\tilde{P}}(\mathbf{t})|^2 \ll \sum_{\mathbf{x} \in V(P)} \prod_{\mathbf{y} \in \Gamma(\mathbf{x})} \frac{1}{(1 + |\mathbf{t} \cdot (\mathbf{x} - \mathbf{y})|)^2}$$

uniformly for all $\mathbf{t} \in \mathbf{R}^k$. The system $\tilde{\mathcal{P}}$ of parallel polytopes is interesting since it is proved by Károlyi [4] that

$$\int_{[0,\varepsilon)^k} \int_{[0,1)^k} \left(\Delta_N(\mathbf{x}_n, \tilde{P}(\rho_1, \ldots, \rho_k) + \mathbf{y}) \right)^2 \, d\mathbf{y} \, d\rho_1 \cdots d\rho_k \gg (\log N)^{k-1}.$$

Hence we get by exactly the same reasoning as in section 2

$$\begin{aligned} D_N^{(2)}(\mathbf{x}_n, \mathcal{P})^2 &= \int_0^1 \int_{[0,1)^k} \left(\Delta_N(\mathbf{x}_n, \rho P + \mathbf{y}) \right)^2 \, d\mathbf{y} \, d\rho \\ &= \lim_{M \to \infty} \frac{1}{(2M)^k} \int_{\mathbf{R}^k} \left(\int_0^1 |\hat{1}_{\rho P}(\mathbf{t})|^2 \, d\rho \right) \left| (d\Delta(\widehat{M, \cdot}))(\mathbf{t}) \right|^2 \, d\mathbf{t} \\ &\gg \lim_{M \to \infty} \frac{1}{(2M)^k} \int_{\mathbf{R}^k} \left(\int_{[0,\varepsilon)^k} |\hat{1}_{\tilde{P}(\rho_1, \ldots, \rho_k)}(\mathbf{t})|^2 \, d\rho \right) \left| (d\Delta(\widehat{M, \cdot}))(\mathbf{t}) \right|^2 \, d\mathbf{t} \end{aligned}$$

$$
\begin{aligned}
&= \int_{[0,\varepsilon)^k} \int_{[0,1)^k} \left(\Delta_N(\mathbf{x}_n, \tilde{P}(\rho_1, \ldots, \rho_k) + \mathbf{y}) \right)^2 \, d\mathbf{y} \, d\rho_1 \cdots d\rho_k \\
&\gg (\log N)^{k-1}.
\end{aligned}
$$

This completes the proof of Theorem 3.

Proof of Lemma 2. First, we will prove the upper bound (17). As usual, a facet F of P is the intersection of P with a supporting hyperplane, e.g. the 0-dimensional facets are the vertices $V(P)$ and the 1-dimensional facets are exactly the edges joining adjacent vertices.

(17) has two trivial cases. If $|\mathbf{t}| < C$ for some constant $C > 0$ then (17) is surely satisfied since

$$
|\hat{1}_P(\mathbf{t})| \leq (2\pi)^{-\frac{k}{2}} \lambda_k(P).
$$

The second one occurs if $|\mathbf{t} \cdot (\mathbf{x} - \mathbf{y})| \geq \varepsilon$ for all adjacent vertices $\mathbf{x}, \mathbf{y} \in V(P)$ and some fixed $\varepsilon > 0$. Here we have

$$
\begin{aligned}
|\hat{1}_P(\mathbf{t})| &\ll \sum_{\mathbf{x} \in V(P)} \prod_{\mathbf{y} \in \Gamma(\mathbf{x})} \frac{1}{|\mathbf{t} \cdot (\mathbf{x} - \mathbf{y})|} \\
&\leq \left(\frac{2}{\varepsilon} \right)^k \sum_{\mathbf{x} \in V(P)} \prod_{\mathbf{y} \in \Gamma(\mathbf{x})} \frac{1}{1 + |\mathbf{t} \cdot (\mathbf{x} - \mathbf{y})|}
\end{aligned}
$$

and consequently (17).

In what follows we will need a (fixed) sequence $\varepsilon = \varepsilon_1 < \varepsilon_2 < \cdots < \varepsilon_k$ such that ε_k and the fractions $\varepsilon_l / \varepsilon_{l+1}$, $1 \leq l < k$, are sufficiently small. We assume now that $|\mathbf{t}| \geq C$ for some sufficiently large constant C and that there exist two adjacent vertices $\mathbf{x}, \mathbf{y} \in V(P)$ with $|\mathbf{t} \cdot (\mathbf{x} - \mathbf{y})| < \varepsilon$. We partition $V(P)$ into vertex sets $V(P) = \bigcup_{j \in J} V_j$ such the following conditions hold. For any $j \in J$ there exists a facet F_j of P of dimension l_j such that $F_j \cap P = V_j$ and there exists $\mathbf{x}_j \in V_j$ such that $|\mathbf{t} \cdot (\mathbf{x}_j - \mathbf{y})| < \varepsilon_{l_j}$ for all $\mathbf{y} \in \Gamma(\mathbf{x}_j) \cap V_j$ and $|\mathbf{t} \cdot (\mathbf{x}_j - \mathbf{y})| \geq \varepsilon_{l_j + 1}$ for all $\mathbf{y} \in \Gamma(\mathbf{x}_j) \setminus V_j$. Note that $|\Gamma(\mathbf{x}_j) \cap V_j| = l_j$ and that there exists at least one set V_j with $|V_j| \geq 2$. Furthermore $|\mathbf{t}| \geq C$ ensures that $l_j < k$.

Let $\mathbf{y} \in \Gamma(\mathbf{x}_j) \setminus V_j$. Then for any vertex $\mathbf{x}' \in V_j$ there exists $\mathbf{y}' \in \Gamma(\mathbf{x}')$ corresponding to $\mathbf{y}$. This means that $\{\mathbf{x} - \mathbf{x}'\} \cup \{\mathbf{y} - \mathbf{x}_j\}$ generate the same subspace as $\{\mathbf{x} - \mathbf{x}'\} \cup \{\mathbf{y}' - \mathbf{x}'\}$. For technical reasons we replace any $\mathbf{y}' \in \Gamma(\mathbf{x}')$ by $\mathbf{x}' + \lambda(\mathbf{y}' - \mathbf{x}')$, in which $\lambda > 0$, such that $(\mathbf{x}' + \lambda(\mathbf{y}' - \mathbf{x}')) - \mathbf{y}$ is parallel to $\{\mathbf{x} - \mathbf{x}'\}$. For brevity we will denote $\mathbf{x}' + \lambda(\mathbf{y}' - \mathbf{x}')$ again by $\mathbf{y}'$. Hence there exists $\mu_{\mathbf{y}}$ such that $\mathbf{x}' - \mathbf{y}' = \mathbf{x}_j - \mathbf{y} + \mu_{\mathbf{y}}(\mathbf{x}' - \mathbf{x}_j)$ for all $\mathbf{y}'$

corresponding to $\mathbf{y}$. Note that the coefficients $c_{\mathbf{x}'}(\mathbf{t})$ do not change during this procedure.

Our aim is to show that

$$\left| \sum_{\mathbf{x}\in V_j} c_{\mathbf{x}}(\mathbf{t}) e^{-i\mathbf{t}\cdot\mathbf{x}} \right| \ll \prod_{\mathbf{y}\in\Gamma(\mathbf{x}_j)\setminus V_j} \frac{1}{|\mathbf{t}\cdot(\mathbf{x}_j-\mathbf{y})|} \tag{20}$$

It is clear that (20) implies (17) since $|\mathbf{t}\cdot(\mathbf{x}'-\mathbf{x}'')| = \mathcal{O}(\varepsilon_{l_j})$ for any two adjacent vertices $\mathbf{x}',\mathbf{x}''\in V_j$ and

$$\prod_{\mathbf{y}'\in\Gamma(\mathbf{x}')\setminus V_j} (-i\mathbf{t}\cdot(\mathbf{x}'-\mathbf{y}')) = \prod_{\mathbf{y}\in\Gamma(\mathbf{x}_j)\setminus V_j} (-i\mathbf{t}\cdot(\mathbf{x}_j-\mathbf{y})) \left(1+\mathcal{O}\left(\frac{\varepsilon_{l_j}}{\varepsilon_{l_j+1}}\right)\right)$$

for any $\mathbf{x}'\in V_j$.

Next, consider a triangulation $\mathcal{T}$ of F_j such that for any simplex $S\in\mathcal{T}$ the vertices $V(S)$ of S are contained in V_j. For $\mathbf{x}\in V(S)$ let $\Gamma_S(\mathbf{x}) = (V(S)\setminus\{\mathbf{x}\})\cup(\Gamma(\mathbf{x})\setminus V_j)$ and set

$$c_S(\mathbf{t}) = (2\pi)^{-\frac{k}{2}} \sum_{\mathbf{x}\in V(S)} \frac{|\det(\mathbf{x}-\mathbf{y})_{\mathbf{y}\in\Gamma_S(\mathbf{x})}|}{\prod_{\mathbf{y}\in\Gamma_S(\mathbf{x})}(-i\mathbf{t}\cdot(\mathbf{x}-\mathbf{y}))} e^{-i\mathbf{t}\cdot\mathbf{x}}.$$

Then by considering the convex hull of $S\cup\bigcup_{\mathbf{x}\in V(S)}(\Gamma(\mathbf{x})\setminus V_j)$, and by using (16) it follows that

$$\sum_{S\in\mathcal{T}} c_S(\mathbf{t}) = \sum_{\mathbf{x}\in V_j} c_{\mathbf{x}}(\mathbf{t}) e^{-i\mathbf{t}\cdot\mathbf{x}}.$$

Therefore it remains to show that

$$|c_S(\mathbf{t})| \ll \prod_{\mathbf{y}\in\Gamma(\mathbf{x}_0)\setminus V_j} \frac{1}{|\mathbf{t}\cdot(\mathbf{x}_0-\mathbf{y})|} \tag{21}$$

where $\mathbf{x}_0$ is any fixed vertex of S.

Let $S\in\mathcal{T}$ and denote $\mathbf{x}_0,\mathbf{x}_1,\ldots,\mathbf{x}_{l_j}$ the vertices of S. Furthermore set

$$F(u) = \frac{e^{-iu}}{\prod_{\mathbf{y}\in\Gamma(\mathbf{x}_0)\setminus V_j}(-i\mathbf{t}\cdot(\mathbf{x}_0-\mathbf{y}) - \mu_{\mathbf{y}}(iu - i\mathbf{t}\cdot\mathbf{x}_0))},$$

$$f_1(u_0,u_1) = \frac{F(u_1)-F(u_0)}{u_1-u_0},$$

and inductively ($l \geq 1$)

$$f_{l+1}(u_0, u_1, \ldots, u_{l+1}) = \frac{f_l(u_0, u_1 \ldots, u_{l-1}, u_{l+1}) - f_l(u_0, u_1, \ldots, u_{l-1}, u_l)}{u_{l+1} - u_l}.$$

Then

$$f_l(u_0, \ldots, u_l) = \sum_{n=0}^{l} \frac{F(u_n)}{\prod_{m \neq n}(u_n - u_m)}$$

and consequently

$$c_S(\mathbf{t}) = (2\pi)^{-\frac{k}{2}}(-i)^{-l_j} f_{l_j}(\mathbf{t} \cdot \mathbf{x}_0, \ldots, \mathbf{t} \cdot \mathbf{x}_{l_j}).$$

On the other hand it follows by induction and by Taylor's theorem that

$$f_l(u_0, \ldots, u_l) = \frac{1}{l!}\left(F_1^{(l)}(v_1) + iF_2^{(l)}(v_2)\right),$$

where $F_1(u) = \Re(F(u))$, $F_2(u) = \Im(F(u))$, and $v_1, v_2 \in [\min u_n, \max u_n]$. Since $\mathbf{t}\cdot(\mathbf{x}_n - \mathbf{x}_m) = \mathcal{O}(\varepsilon_{l_j})$ and $|\mathbf{t}\cdot(\mathbf{x}_n - \mathbf{y})| > \varepsilon_{l_j+1} - \mathcal{O}(\varepsilon_{l_j})$ for $\mathbf{y} \in \Gamma(\mathbf{x}_n) \setminus V_j$ we obtain

$$c_S(\mathbf{t}) = (2\pi)^{-\frac{k}{2}}\lambda_k(\tilde{S})(-i)^{l_j}\left(F_1^{(l_j)}(\mathbf{t}\cdot\mathbf{x}_0 + \mathcal{O}(\varepsilon_{l_j})) + F_1^{(l_j)}(\mathbf{t}\cdot\mathbf{x}_0 + \mathcal{O}(\varepsilon_{l_j}))\right),$$

where $\tilde{S}$ denotes the convex hull of $S \cup (\Gamma(\mathbf{x}_0) \setminus V_j)$. Since

$$\begin{aligned} &(-i)^{l_j} F^{(l_j)}(u) \qquad (22)\\ &= e^{-iu}\sum_{n=0}^{l_j}\frac{l_j!}{(l_j-n)!}\sum_{\sum m_{\mathbf{y}}=n}\prod_{\mathbf{y}\in\Gamma(\mathbf{x}_0)\setminus V_j}\frac{\mu_{\mathbf{y}}^{m_{\mathbf{y}}}}{(-i\mathbf{t}\cdot(\mathbf{x}_0-\mathbf{y}) - \mu_{\mathbf{y}}(iu - i\mathbf{t}\cdot\mathbf{x}_0))^{m_{\mathbf{y}}+1}}\end{aligned}$$

and $|\mathbf{t}\cdot(\mathbf{x}_0 - \mathbf{y})| \gg \varepsilon_{l_j+1}$, $\mathbf{y} \in \Gamma(\mathbf{x}_0) \setminus V_j$, we finally get

$$\begin{aligned} c_S(\mathbf{t}) &= (2\pi)^{-\frac{k}{2}}\lambda_k(\tilde{S})e^{-i\mathbf{t}\cdot\mathbf{x}_0}\\ &\times \sum_{n=0}^{l_j}\frac{l_j!}{(l_j-n)!}\sum_{\sum m_{\mathbf{y}}=n}\prod_{\mathbf{y}\in\Gamma(\mathbf{x}_0)\setminus V_j}\frac{\mu_{\mathbf{y}}^{m_{\mathbf{y}}}}{(-i\mathbf{t}\cdot(\mathbf{x}_0-\mathbf{y}))^{m_{\mathbf{y}}+1}}\left(1 + \mathcal{O}\left(\frac{\varepsilon_{l_j}}{\varepsilon_{l_j+1}}\right)\right),\end{aligned}$$

which proves (21) and consequently (17).

The proof of the lower bound (18) uses a similar procedure but is more refined. First, the case $|\mathbf{t}| < C$ is trivial since

$$\min_{|\mathbf{t}|<C}\int_0^1 |\hat{1}_{\rho P}(\mathbf{t})|^2\, d\rho > 0.$$

Note that $\hat{1}_{\rho P}(\mathbf{t})$ has the representation

$$\hat{1}_{\rho P}(\mathbf{t}) = \rho^k \hat{1}_P(\rho \mathbf{t}) = \sum_{\mathbf{x}\in V(P)} c_{\mathbf{x}}(\mathbf{t}) e^{-i\rho\mathbf{t}\cdot\mathbf{x}}.$$

In order to give a flavour of the method let us consider the easiest case, in which $|\mathbf{t}\cdot(\mathbf{x}-\mathbf{y})| > \varepsilon_1$ for all adjacent vertices $\mathbf{x},\mathbf{y} \in V(P)$. Set $v = |V(P)|$. By elementary considerations it follows that there exists $\delta \in [\frac{1}{2v}, \frac{1}{v}]$ such that

$$\left| e^{-i\delta\mathbf{t}\cdot\mathbf{x}} - e^{-i\delta\mathbf{t}\cdot\mathbf{y}} \right| \geq \eta = \min\left(\frac{\delta\varepsilon_1}{2}, \frac{1}{2v^2}\right)$$

for any pair of adjacent vertices $\mathbf{x},\mathbf{y} \in V(P)$. Then the matrix

$$M(\tau) = \left(e^{-i(\delta m+\tau)\mathbf{t}\cdot\mathbf{x}}\right)_{0\leq m<v,\mathbf{x}\in V(P)}$$

is regular and its determinant can be estimated by

$$|\det M(\tau)| = \prod_{(\mathbf{x},\mathbf{y})\in E(P)} \left| e^{-i\delta\mathbf{t}\cdot\mathbf{x}} - e^{-i\delta\mathbf{t}\cdot\mathbf{y}} \right| \geq \eta^{\frac{1}{2}v(v-1)},$$

where $E(P)$ denotes the edges of P. Next, consider the vectors $\mathbf{c} = (c_{\mathbf{x}})_{\mathbf{x}\in V(P)}(\mathbf{t})$ and $\mathbf{v}(\tau) = \left(\hat{1}_{(\delta m+\tau)P}(\mathbf{t})\right)_{0\leq m<v}$. From $\mathbf{c} = M(\tau)^{-1}\mathbf{v}(\tau)$ we get

$$\begin{aligned}
\|\mathbf{v}(\tau)\|_\infty &= \max_{0\leq m<v} |\hat{1}_{(\delta m+\tau)P}(\mathbf{t})| \\
&\geq \frac{1}{\|M(\tau)^{-1}\|}\|\mathbf{c}\|_\infty \\
&\geq \frac{\eta^{\frac{1}{2}v(v-1)}}{v!}\|\mathbf{c}\|_\infty
\end{aligned}$$

and finally

$$\int_0^1 |\hat{1}_{\rho P}(\mathbf{t})|^2\, d\rho \gg \int_0^\delta \|\mathbf{v}(\tau)\|_\infty^2\, d\tau \gg \|\mathbf{c}\|_\infty^2$$

which proves (18).

If there exist adjacent vertices $\mathbf{x},\mathbf{y} \in V(P)$ with $|\mathbf{t}\cdot(\mathbf{x}-\mathbf{y})| \leq \varepsilon_1$ then we use the same partition $V(P) = \bigcup_{j\in J} V_j$ as in the proof for the upper bound (17). Furthermore, for any $j \in J$ we get by exactly the same methods as above

$$\sum_{\mathbf{x}\in V_j} c_{\mathbf{x}}(\mathbf{t}) e^{-i\rho\mathbf{t}\cdot\mathbf{x}} = (2\pi)^{-\frac{k}{2}} \lambda_k(\tilde{V}_j) e^{-i\rho\mathbf{t}\cdot\mathbf{x}_j} \times \tag{23}$$

$$\sum_{n=0}^{l_j} \rho^{l_j-n} \frac{l_j!}{(l_j-n)!} \sum_{\sum m_{\mathbf{y}}=n} \prod_{\mathbf{y}\in\Gamma(\mathbf{x}_0)\setminus V_j} \frac{\mu_{\mathbf{y}}^{m_{\mathbf{y}}}}{(-i\mathbf{t}\cdot(\mathbf{x}_j-\mathbf{y}))^{m_{\mathbf{y}}+1}} \left(1+\mathcal{O}\left(\frac{\varepsilon_{l_j}}{\varepsilon_{l_j+1}}\right)\right),$$

where $\mathbf{x}_j \in V_j$. Now, let $v' = \sum_j (l_j + 1)$ and choose $\delta \in [\frac{1}{2v'}, \frac{1}{v'}]$ such that

$$\left| e^{-i\delta \mathbf{t}\cdot\mathbf{x}_j} - e^{-i\delta \mathbf{t}\cdot\mathbf{x}_{j'}} \right| \ge \eta' = \min\left(\frac{\delta\varepsilon_1}{2}, \frac{1}{2v'^2} \right)$$

for $j \ne j'$. Again the determinant of the matrix

$$M(\tau) = \left(e^{-i(\delta m+\tau)\mathbf{t}\cdot\mathbf{x}_j} (\delta m + \tau)^{l_j - n} \right)_{0\le m<v',(j\in J, 0\le n\le l_j)}$$

can be bounded below by

$$\begin{aligned} |\det M(\tau)| &= \delta^{\sum_{j\in J} l_j} \left(\prod_{j\in J} \prod_{n=0}^{l_j} \frac{1}{n!} \right) \prod_{j\ne j'\in J} \left| e^{-i\delta \mathbf{t}\cdot\mathbf{x}_j} - e^{-i\delta \mathbf{t}\cdot\mathbf{x}_{j'}} \right|^{(l_j+1)(l_{j'}+1)} \\ &\ge \delta^{\sum_{j\in J} l_j} \eta'^{\sum_{j\ne j'\in J} (l_j+1)(l_{j'}+1)} \prod_{j\in J} \prod_{n=0}^{l_j} \frac{1}{n!}. \end{aligned}$$

Finally consider the vectors $\mathbf{c} = (c_{jn})_{(j\in J, 0\le n\le l_j)} = M(\tau)^{-1}\mathbf{v}(\tau)$ and $\mathbf{v}(\tau) = \left(\hat{1}_{(\delta m+\tau)P}(\mathbf{t}) \right)_{0\le m<v}$. By (23) c_{jn} is given by

$$c_{jn} = \frac{l_j!}{(l_j - n)!} \sum_{\sum m_\mathbf{y} = n} \prod_{\mathbf{y}\in\Gamma(\mathbf{x}_0)\setminus V_j} \frac{\mu_\mathbf{y}^{m_\mathbf{y}}}{(-i\mathbf{t}\cdot(\mathbf{x}_j - \mathbf{y}))^{m_\mathbf{y}+1}} \left(1 + \mathcal{O}\left(\frac{\varepsilon_{l_j}}{\varepsilon_{l_j+1}} \right) \right).$$

As above they satisfy

$$\|\mathbf{v}(\tau)\|_\infty \gg \|\mathbf{c}\|_\infty \ge \max_{j\in J} |c_{j0}|.$$

Consequently we obtain (18) in any case. □

References

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

[2] M. Drmota, *Irregularities of distribution and convex sets*, Grazer Math. Berichte **318** (1992), 9–16.

[3] E. Hlawka, *Theorie der Gleichverteilung*, Bibl. Inst., Mannheim, 1979.

[4] G. Károlyi, *Geometric discrepancy theorems in higher dimensions*, Studia Scientiarum Mathematicorum Hungarica, to appear.

[5] L. Kuipers and H. Niederreiter, "Uniform Distribution of Sequences", John Wiley and Sons, London, 1974.

[6] H. Niederreiter, R. F. Tichy, and G. Turnwald, *An unequality for differences of distribtution functions*, Arch. Math. **54** (1990), 166–172.

[7] K. F. Roth, *On irregularities of distribution*, Mathematika **1** (1954), 73–79.

[8] I. Z. Ruzsa, *The discrepancy of rectangles and squares*, Grazer Math. Ber. **318** (1992), 135–140.

[9] W. M. Schmidt, *Irregularities of Distribution VII*, Acta Arith. **21** (1972), 63–74.

Michael Drmota
Department of Discrete Mathematics
Technical University of Vienna
Wiedner Hauptstrasse 8–10/118
A-1040 Vienna, Austria

E-mail: `mdrmota@rsmb.tuwien.ac.at`

AMS-classification: 11 K 38
Key words: irregularities of distribution, discrepancy