ON A HYBRID VERSION OF THE VINOGRADOV MEAN VALUE THEOREM

C. CHEN and I. E. SHPARLINSKI[*]

Department of Pure Mathematics, University of New South Wales, Sydney, NSW 2052, Australia
e-mails: changhao.chenm@gmail.com, igor.shparlinski@unsw.edu.au

Abstract. Given a family $\boldsymbol{\varphi} = (\varphi_1, \ldots, \varphi_d) \in \mathbb{Z}[T]^d$ of d distinct nonconstant polynomials, a positive integer $k \le d$ and a real positive parameter ρ, we consider the mean value

$$M_{k,\rho}(\boldsymbol{\varphi}, N) = \int_{\mathbf{x} \in [0,1]^k} \sup_{\mathbf{y} \in [0,1]^{d-k}} |S_{\boldsymbol{\varphi}}(\mathbf{x}, \mathbf{y}; N)|^\rho \, d\mathbf{x}$$

of exponential sums

$$S_{\boldsymbol{\varphi}}(\mathbf{x}, \mathbf{y}; N) = \sum_{n=1}^{N} \exp\left(2\pi i \left(\sum_{j=1}^{k} x_j \varphi_j(n) + \sum_{j=1}^{d-k} y_j \varphi_{k+j}(n) \right) \right),$$

where $\mathbf{x} = (x_1, \ldots, x_k)$ and $\mathbf{y} = (y_1, \ldots, y_{d-k})$. The case of polynomials $\varphi_i(T) = T^i$, $i = 1, \ldots, d$ and $k = d$ corresponds to the classical *Vinaogradov mean value theorem*.

Here motivated by recent works of Wooley [14] and the authors [9] on bounds on $\sup_{\mathbf{y} \in [0,1]^{d-k}} |S_{\boldsymbol{\varphi}}(\mathbf{x}, \mathbf{y}; N)|$ for almost all $\mathbf{x} \in [0,1]^k$, we obtain nontrivial bounds on $M_{k,\rho}(\boldsymbol{\varphi}, N)$.

1. Introduction

For an integer $\nu \geqslant 2$, let $\mathsf{T}_\nu = (\mathbb{R}/\mathbb{Z})^\nu$ be the ν-dimensional unit torus. We denote

$$\mathbf{e}(x) = \exp(2\pi i x).$$

The exponential sums

$$S_d(\mathbf{u}; N) = \sum_{n=1}^{N} \mathbf{e}(u_1 n + \cdots + u_d n^d), \quad \mathbf{u} = (u_1, \ldots, u_d) \in \mathsf{T}_d,$$

introduced and estimated by Weyl [12], are commonly called the *Weyl sums*.

Weyl sums appear in a great variety of number theoretic problems starting with the problem of uniformity of distribution of fractional parts of real polynomials, see also [4]. They also play a crucial role in estimating the *zero-free region* of the *Riemann zeta-function* and thus in turn in bounding in the error term in the *prime number theorem*, see [11, Section 8.5], and the *Waring problem*, see [11, Section 20.2], in estimating short character sums modulo highly composite numbers [11, Section 12.6].

Thanks to recent striking results of Bourgain, Demeter and Guth [6] (for $d \geqslant 4$) and Wooley [13] (for $d = 3$) (see also [15]), for the mean value of $S_d(\mathbf{u}; N)$ we have

$$(1.1) \qquad \int_{\mathsf{T}_d} |S_d(\mathbf{u}; N)|^{2s(d)} \, d\mathbf{u} \leqslant N^{s(d)+o(1)}, \quad N \to \infty,$$

where

$$(1.2) \qquad s(d) = \frac{d(d+1)}{2},$$

which is the best possible form of the Vinogradov mean value theorem.

On the other hand, for individual sums it is known that their size depends on Diophantine properties of the coefficients $u_1, \ldots, u_d$, but generally the situation is not well understood, see [7,8].

The following best known bound is a direct implication of (1.1) and is given in [5, Theorem 5]. Let $\mathbf{u} = (u_1, \ldots, u_d) \in \mathsf{T}_d$ be such that for some ν with $2 \leq \nu \leq d$ and some positive integers a and q with $\gcd(a, q) = 1$ we have

$$\left| u_\nu - \frac{a}{q} \right| \leq \frac{1}{q^2}.$$

Then for any $\varepsilon > 0$ there exits a constant $C(\varepsilon)$ such that

$$|S_d(\mathbf{u}; N)| \leq C(\varepsilon) N^{1+\varepsilon} \big(q^{-1} + N^{-1} + q N^{-\nu} \big)^{\frac{1}{d(d-1)}}.$$

Recently, Wooley [14] has considered a hybrid scenario which interpolates between *individual* bounds and *mean value* estimates. In this setting one seeks results which hold for *all* values of the components of $\mathbf{u} = (u_1, \ldots, u_d) \in \mathsf{T}_d$ on some prescribed set of positions and *almost all* values of the components on the remaining positions.

Given a family $\boldsymbol{\varphi} = (\varphi_1, \ldots, \varphi_d) \in \mathbb{Z}[T]^d$ of d distinct nonconstant polynomials and a sequence of complex weights $\mathbf{a} = (a_n)_{n=1}^\infty$, for $\mathbf{u} = (u_1, \ldots, u_d) \in \mathsf{T}_d$ we define the trigonometric polynomials

$$(1.3) \qquad T_{\mathbf{a}, \boldsymbol{\varphi}}(\mathbf{u}; N) = \sum_{n=1}^N a_n \, \mathbf{e}(u_1 \varphi_1(n) + \cdots + u_d \varphi_d(n)).$$

Furthermore, for $k = 1, \ldots, d$, we decompose

$$\mathsf{T}_d = \mathsf{T}_k \times \mathsf{T}_{d-k}.$$

Given $\mathbf{x} \in \mathsf{T}_k$, $\mathbf{y} \in \mathsf{T}_{d-k}$ we refine the notation (1.3) and write

$$(1.4) \qquad T_{\mathbf{a},\boldsymbol{\varphi}}(\mathbf{x}, \mathbf{y}; N) = \sum_{n=1}^{N} a_n \, \mathbf{e} \left(\sum_{j=1}^{k} x_j \varphi_j(n) + \sum_{j=1}^{d-k} y_j \varphi_{k+j}(n) \right).$$

If $\mathbf{a} = \mathbf{e} = (1)_{n=1}^{\infty}$ (that is, $a_n = 1$ for each $n \in \mathbb{N}$) we just write

$$T_{\boldsymbol{\varphi}}(\mathbf{x}, \mathbf{y}; N) = T_{\mathbf{e},\boldsymbol{\varphi}}(\mathbf{x}, \mathbf{y}; N).$$

In fact Wooley [14] has studied only the classical case $a_n = 1$ for all $n \in \mathbb{N}$ and the polynomials

$$(1.5) \qquad \left\{ \varphi_1(T), \ldots, \varphi_d(T) \right\} = \left\{ T, \ldots, T^d \right\}$$

(we note that the order of $\varphi_1, \ldots, \varphi_d$ is not specified in (1.5)). Embedding in the argument of Wooley [14, Theorem 1.1] the modern form of the Vinogradov mean value theorem (1.1), one derives that for almost all $\mathbf{x} \in \mathsf{T}_k$ with respect to the k-dimensional Lebesgue measure on T_k, we have

$$(1.6) \qquad \sup_{\mathbf{y} \in \mathsf{T}_{d-k}} |T_{\boldsymbol{\varphi}}(\mathbf{x}, \mathbf{y}; N)| \leq N^{(1+\Delta_W(\boldsymbol{\varphi}, k))/2 + o(1)}, \quad N \to \infty,$$

where

$$\Delta_W(\boldsymbol{\varphi}, k) = \frac{2\sigma_k(\boldsymbol{\varphi}) + d - k + 1}{2s(d) + d - k + 1}$$

with $s(d)$ given by (1.2) and

$$(1.7) \qquad \sigma_k(\boldsymbol{\varphi}) = \sum_{j=k+1}^{d} \deg \varphi_j.$$

The authors [9] have extended and improved this and some other results of Wooley [14]. In particular, by [9, Theorems 2.1, 2.3 and Corollary 2.4], the bound (1.6) holds with

$$\Delta_{CS}(\boldsymbol{\varphi}, k) = \min \left\{ \frac{2\sigma_k(\boldsymbol{\varphi}) + d - k}{2s(d) + d - k}, \frac{\sigma_k(\boldsymbol{\varphi}) + 1}{s(d)} \right\} < \Delta_W(\boldsymbol{\varphi}, k)$$

instead of $\Delta_W(\boldsymbol{\varphi}, k)$ and also applies to sums $T_{\mathbf{a},\boldsymbol{\varphi}}(\mathbf{x}, \mathbf{y}; N)$ with complex weights $a_n = n^{o(1)}, n \in \mathbb{N}$, and a large family of polynomials.

2. Hybrid mean value theorems

2.1. Notation and conventions. Throughout the paper, the notation $U = O(V)$, $U \ll V$ and $V \gg U$ are equivalent to $|U| \leqslant cV$ for some positive constant c, which throughout the paper may depend on the degree d and occasionally on the small real positive parameter ε.

For any quantity $V > 1$ we write $U = V^{o(1)}$ (as $V \to \infty$) to indicate a function of V which satisfies $|U| \leq V^{\varepsilon}$ for any $\varepsilon > 0$, provided V is large enough. One additional advantage of using $V^{o(1)}$ is that it absorbs $\log V$ and other similar quantities without changing the whole expression.

We use $\#\mathcal{S}$ to denote the cardinality of a finite set $\mathcal{S}$.

We say that some property holds for almost all $\mathbf{x} \in \mathsf{T}_k$ if it holds for a set $\mathcal{X} \subseteq \mathsf{T}_k$ of k-dimensional Lebesgue measure $\lambda(\mathcal{X}) = 1$.

2.2. Main results. Here, instead of asking about bounds on

$$\sup_{\mathbf{y} \in \mathsf{T}_{d-k}} \left| T_{\mathbf{a},\boldsymbol{\varphi}}(\mathbf{x}, \mathbf{y}; N) \right|,$$

for a real $\rho > 0$ we consider the mean value

$$M_{k,\rho}(\mathbf{a}, \boldsymbol{\varphi}, N) = \int_{\mathbf{x} \in \mathsf{T}_k} \sup_{\mathbf{y} \in \mathsf{T}_{d-k}} |T_{\mathbf{a},\boldsymbol{\varphi}}(\mathbf{x}, \mathbf{y}; N)|^{\rho} \, d\mathbf{x}.$$

In some special cases of polynomials $\boldsymbol{\varphi}$ and the exponent ρ this quantity has been studied in [1, Lemma 5] and [2, Lemmas 3.2], see also [3, Lemma 12]. However our results and methods and those of [1,2] seem to apply to different regimes of parameters and do not overlap. More precisely, the argument of the proof of [1, Lemma 5] or [2, Lemmas 3.2] implies that for any integer $s \geq 1$ we have

$$(2.1) \qquad M_{k,2s}(\mathbf{a}, \boldsymbol{\varphi}, N) \ll N^{\sigma_k(\boldsymbol{\varphi})} M_{k,2s}(\mathbf{a}, \boldsymbol{\varphi}_k, N),$$

where $s(d)$ and $\sigma_k(\boldsymbol{\varphi})$ are given by (1.2) and (1.7), respectively, and $\boldsymbol{\varphi}_k = (\varphi_1, \ldots, \varphi_k)$, thus, recalling (1.3), we write

$$M_{k,2s}(\mathbf{a}, \boldsymbol{\varphi}_k, N) = \int_{\mathsf{T}_k} \left| T_{\mathbf{a},\boldsymbol{\varphi}_k}(\mathbf{x}; N) \right|^{2s} d\mathbf{x}.$$

If s is large enough, depending on $\max\{\deg \varphi_1, \ldots, \deg \varphi_k\}$, one can use (1.1) to estimate $M_{k,2s}(\mathbf{a}, \boldsymbol{\varphi}_k, N)$ (which is how it has been done in [1,2]). For smaller s one can apply the bound of Wooley [15, Theorem 1.1] (under some natural conditions on $\varphi_1, \ldots, \varphi_k$). However the presence of the factor $N^{\sigma_k(\boldsymbol{\varphi})}$ in (2.1) impedes the efficiency of this approach.

We now present a different approach, leading to different bounds which do not depend on the degrees of $\varphi_1, \ldots, \varphi_k$ and thus which is especially efficient when at least the degree is large compared to d.

Let

$$W(T; \boldsymbol{\varphi}) = \det\left(\varphi_i^{(j-1)}(T)\right)_{i,j=1}^d$$

denote the Wronskian of d polynomials $\boldsymbol{\varphi} = (\varphi_1, \ldots, \varphi_d) \in \mathbb{Z}[T]^d$.

It is convenient to formulate our results in terms of some real parameter ρ. However one can see that by convexity it is enough to establish them only for the largest admissible value of ρ.

THEOREM 2.1. *Suppose that $\boldsymbol{\varphi} \in \mathbb{Z}[T]^d$ is such that the Wronskian $W(T; \boldsymbol{\varphi})$ does not vanish identically. Let $\mathbf{a} = (a_n)_{n=1}^\infty$ be a sequence of complex weights with $a_n = n^{o(1)}$. Then for any real positive $\rho \leq 2s(d) + d - k$ we have*

$$M_{k,\rho}(\mathbf{a}, \boldsymbol{\varphi}, N) \leq N^{\mu(\boldsymbol{\varphi},k)\rho+o(1)}, \quad N \to \infty,$$

where

$$\mu(\boldsymbol{\varphi}, k) = \frac{s(d) + \sigma_k(\boldsymbol{\varphi}) + d - k}{2s(d) + d - k}.$$

Note that Theorem 2.1 gives a non-trivial bound, that is, $\mu(\boldsymbol{\varphi}, k) < 1$, provided that $\sigma_k(\boldsymbol{\varphi}) < s(d)$. Moreover for $k = d$, we have $\sigma_d(\boldsymbol{\varphi}) = 0$ and thus we recover the bound (1.1) in the Vinogradov mean value theorem.

From Theorem 2.1 we derive the following bounds for the mean values of short sums. For $K \in \mathbb{Z}$, we consider Weyl sums over short intervals

$$S_d(\mathbf{u}; K, N) = \sum_{n=K+1}^{K+N} \mathbf{e}(u_1 n + \cdots + u_d n^d), \quad \mathbf{u} = (u_1, \ldots, u_d) \in \mathsf{T}_d.$$

We now give an upper bound on the mean value of the largest value of all such sums, that is, for $\sup_{K \in \mathbb{Z}} |S_d(\mathbf{u}; K, N)|$. No result of this type seems to be known prior to this work.

COROLLARY 2.2. *For any real positive $\rho \leq d^2 + 2d - 1$ we have*

$$\int_{\mathsf{T}_d} \sup_{K \in \mathbb{Z}} |S_d(\mathbf{u}; K, N)|^\rho \, d\mathbf{u} \leq N^{\mu_d \rho + o(1)}, \quad N \to \infty,$$

where

$$\mu_d = 1 - \frac{d}{d^2 + 2d - 1}.$$

We note that while the bound

$$\sup_{K\in\mathbb{Z}}\int_{\mathsf{T}_d}\big|S_d(\mathbf{u};K,N)\big|^{2s(d)}\,d\mathbf{u}\le N^{s(d)+o(1)},\quad N\to\infty,$$

is trivial from (1.1) and the so called *translation invariance*, there is no obvious nontrivial bound on the integral of Corollary 2.2.

Recalling the definition (1.3) and (1.4), for $\mathbf{x}\in\mathsf{T}_k$ we can write

$$\sup_{\mathbf{y}\in T_{d-k}}|T_{\mathbf{a},\boldsymbol{\varphi}}(\mathbf{x},\mathbf{y};N)|=\sup_{\mathbf{u}\in T_d\cap\pi_{d,k}^{-1}(\mathbf{x})}|T_{\mathbf{a},\boldsymbol{\varphi}}(\mathbf{u};N)|,$$

where $\pi_{d,k}$ is the orthogonal projection of T_d onto T_k, that is,

$$\pi_{d,k}\colon (u_1,\ldots,u_d)\to(u_1,\ldots,u_k).$$

This suggests that one can try to extend Theorem 2.1 to more general settings by taking some other mappings instead of the orthogonal projection $\pi_{d,k}$.

Now, more generally, given a mapping $f\colon\mathbb{R}^d\to\mathbb{R}^k$, for $\mathbf{x}\in\mathbb{R}^k$ we denote

$$f^{-1}(\mathbf{x})=\big\{\mathbf{u}\in\mathbb{R}^d:f(\mathbf{u})=\mathbf{x}\big\}.$$

Then for $\rho>0$ we define

$$\mathsf{M}_{k,f,\rho}(\mathbf{a},\boldsymbol{\varphi},N)=\int_{\mathbb{R}^k}\sup_{\mathbf{u}\in T_d\cap f^{-1}(\mathbf{x})}|T_{\mathbf{a},\boldsymbol{\varphi}}(\mathbf{u};N)|^\rho\,d\mathbf{x},$$

where $T_{\mathbf{a},\boldsymbol{\varphi}}(\mathbf{u};N)$ is given by (1.3).

In the following we first take f to be an orthogonal projection onto some k-dimensional subspace, and second we take f to be some Hölder mapping.

Let $\mathcal{G}(d,k)$ denote the collection of all k-dimensional linear subspaces of $\mathbb{R}^d$. For $\mathcal{V}\in\mathcal{G}(d,k)$, let $\pi_{\mathcal{V}}\colon\mathbb{R}^d\to\mathcal{V}$ denote the orthogonal projection onto $\mathcal{V}$.

For the degree sequence $\deg\varphi_1,\ldots,\deg\varphi_d$ we denote them as

$$r_1\le\ldots\le r_d,$$

and define

(2.2)
$$\widetilde{\sigma}_k(\boldsymbol{\varphi})=\sum_{i=k+1}^d r_i.$$

THEOREM 2.3. *Suppose that $\boldsymbol{\varphi}\in\mathbb{Z}[T]^d$ is such that the Wronskian $W(T;\boldsymbol{\varphi})$ does not vanish identically. Let $\mathbf{a}=(a_n)_{n=1}^\infty$ be a sequence of*

complex weights with $a_n = n^{o(1)}$. *If* $\mathcal{V} \in \mathcal{G}(d,k)$, *then for any real positive* $\rho \leq 2s(d) + d - k$ *we have*

$$\mathsf{M}_{k,\pi_\mathcal{V},\rho}(\mathbf{a}, \boldsymbol{\varphi}, N) \leq N^{\mu_\mathcal{V}(\boldsymbol{\varphi},k)\rho + o(1)}, \quad N \to \infty,$$

where

$$\mu_\mathcal{V}(\boldsymbol{\varphi}, k) = \frac{s(d) + \widetilde{\sigma}_k(\boldsymbol{\varphi}) + d - k}{2s(d) + d - k}.$$

Now we turn to ϑ-*Hölder functions*, for some $0 < \vartheta \leq 1$, that is, functions $f \colon \mathbb{R}^d \to \mathbb{R}^k$ which satisfy

$$\|f(\mathbf{x}) - f(\mathbf{y})\| \ll \|\mathbf{x} - \mathbf{y}\|^\vartheta,$$

where $\|\mathbf{z}\|$ is the Euclidean norm of $\mathbf{z}$ (note that the left-hand side of this inequality is the Euclidean norm in $\mathbb{R}^k$, while the right-hand side is the Euclidean norm in $\mathbb{R}^d$). In particular, in the case $\vartheta = 1$ the function f is often called a *Lipschitz function*.

THEOREM 2.4. *Let* $f \colon \mathbb{R}^d \to \mathbb{R}^k$ *be a* ϑ-*Hölder map for some* $0 < \vartheta \leq 1$. *Let* $\mathbf{a} = (a_n)_{n=1}^\infty$ *be a sequence of complex weights with* $a_n = n^{o(1)}$. *Suppose that* $\boldsymbol{\varphi} \in \mathbb{Z}[T]^d$ *is such that the Wronskian* $W(T; \boldsymbol{\varphi})$ *does not vanish identically. Denote*

$$\delta(\boldsymbol{\varphi}) = \min_{i=1,\ldots,d} \deg \varphi_i.$$

Then for any real positive $\rho \leq 2s(d) + d - k\vartheta$ *we have*

$$\mathsf{M}_{k,f,\rho}(\mathbf{a}, \boldsymbol{\varphi}, N) \leq N^{\mu_\vartheta(\boldsymbol{\varphi},k)\rho + o(1)}, \quad N \to \infty,$$

where

$$\mu_\vartheta(\boldsymbol{\varphi}, k) = \frac{s(d) + \sigma_0(\boldsymbol{\varphi}) + d - (\delta(\boldsymbol{\varphi}) + 1)\vartheta k}{2s(d) + d - k\vartheta}.$$

Note that for the classical choice of polynomials (1.5), we have $\delta(\boldsymbol{\varphi}) = 1$, $\sigma_0(\boldsymbol{\varphi}) = s(d)$ and thus

$$\mu_\vartheta(\boldsymbol{\varphi}, k) = \frac{2s(d) + d - 2k\vartheta}{2s(d) + d - k\vartheta} < 1$$

for any $k > 0$. Therefore Theorem 2.4 gives a non-trivial bound for $\mathsf{M}_{k,f,\rho}(\mathbf{a}, \boldsymbol{\varphi}, N)$ for any $k > 0$ and any ϑ-Hölder function f with $0 < \vartheta \leq 1$. Moreover if $\vartheta \to 0$ then $\mu_\vartheta(\boldsymbol{\varphi}, k) \to 1$. Indeed, it is expected that if the function f becomes "bad" (ϑ becomes "small") then the bounds for $\mathsf{M}_{k,f,\rho}(\mathbf{a}, \boldsymbol{\varphi}, N)$ also become "bad". However, we do not know whether there

exits a continuous function $f \colon \mathbb{R}^d \to \mathbb{R}^k$ such that we do not have non-trivial bounds for $\mathsf{M}_{k,f,\rho}(\mathbf{a}, \boldsymbol{\varphi}, N)$, that is, for any $\rho > 0$ and any $\varepsilon > 0$, one has

$$\mathsf{M}_{k,f,\rho}(\mathbf{a}, \boldsymbol{\varphi}, N) \gg N^{1-\varepsilon}$$

for infinitely many $N \in \mathbb{N}$.

2.3. Hybrid mean value theorems for discrepancy. Similar to works of Wooley [14] and the authors [9], we obtain similar results for the discrepancy.

Let ξ_n, $n \in \mathbb{N}$, be a sequence in $[0, 1)$. The *discrepancy* of this sequence at length N is defined as

$$(2.3) \qquad D_N = \sup_{0 \le a < b \le 1} \left| \#\{ 1 \le n \le N : \xi_n \in (a, b) \} - (b - a)N \right|.$$

We note that sometimes in the literature the scaled quantity $N^{-1}D_N$ is called the discrepancy, but since our argument looks cleaner with the definition (2.3), we adopt it here.

For $\mathbf{x} \in \mathsf{T}_k$, $\mathbf{y} \in \mathsf{T}_{d-k}$ we consider the sequence

$$\sum_{j=1}^{k} x_j \varphi_j(n) + \sum_{j=1}^{d-k} y_j \varphi_{k+j}(n), \quad n \in \mathbb{N},$$

and for each N we denote by $D_{\boldsymbol{\varphi}}(\mathbf{x}, \mathbf{y}; N)$ the corresponding discrepancy of its fractional parts.

For $\rho > 0$ let

$$\mathfrak{M}_{k,\rho}(\boldsymbol{\varphi}, N) = \int_{\mathsf{T}_k} \sup_{\mathbf{y} \in \mathsf{T}_{d-k}} D_{\boldsymbol{\varphi}}(\mathbf{x}, \mathbf{y}; N)^{\rho} \, d\mathbf{x}.$$

THEOREM 2.5. *Suppose that $\boldsymbol{\varphi} \in \mathbb{Z}[T]^d$ is such that the Wronskian $W(T; \boldsymbol{\varphi})$ does not vanish identically. Then for any $1 \le \rho \le 2s(d) + d - k$ we have*

$$\mathfrak{M}_{k,\rho}(\boldsymbol{\varphi}, N) \le N^{\mu(\boldsymbol{\varphi}, k)\rho + o(1)}, \quad N \to \infty,$$

where $\mu(\boldsymbol{\varphi}, k)$ is as in Theorem 2.1.

From Theorem 2.5 we derive a bound on the mean value of discrepancy over short intervals. More precisely, for each $K \in \mathbb{Z}$ denote by $D_d(\mathbf{u}; K, N)$ the discrepancy of the sequence of fractional parts

$$\{u_1 n + \cdots + u_d n^d\}, \quad n = K + 1, \ldots, K + N.$$

COROLLARY 2.6. *For any real positive $1 \leq \rho \leq d^2 + 2d - 1$ we have*

$$\int_{\mathsf{T}_d} \sup_{K \in \mathbb{Z}} D_d(\mathbf{u}; K, N)^\rho \, d\mathbf{u} \leq N^{\mu_d \rho + o(1)}, \quad N \to \infty,$$

where μ_d is given by Corollary 2.2.

3. Preliminaries

3.1. Packing large trigonometric polynomials in boxes. First we need a box-counting estimate which is essentially [9, Lemma 3.7]. Indeed, we certainly have $|T_{\mathbf{a},\boldsymbol{\varphi}}(\mathbf{u}; N)| \leq W_{\mathbf{a},\boldsymbol{\varphi}}(\mathbf{u}; N)$, where the sums $W_{\mathbf{a},\boldsymbol{\varphi}}(\mathbf{u}; N)$ are as in [9, Section 3.3].

LEMMA 3.1. *Let $0 < \alpha < 1$ and let ε be sufficiently small. For each $j = 1, \ldots, d$ let*

(3.1) $$\zeta_j = 1/\left\lceil N^{e_j + 1 + \varepsilon - \alpha} \right\rceil,$$

where $e_j = \deg \varphi_j$, $j = 1, \ldots, d$. We divide T_d into

$$U = \left(\prod_{j=1}^{d} \zeta_j \right)^{-1}$$

boxes of the form

$$[n_1 \zeta_1, (n_1 + 1)\zeta_1) \times \cdots \times [n_d \zeta_d, (n_d + 1)\zeta_d),$$

where $n_j = 1, \ldots, 1/\zeta_j$ for each $j = 1, \ldots, d$. Let $\mathfrak{R}$ be the collection of these boxes, and

$$\widetilde{\mathfrak{R}} = \left\{ \mathcal{R} \in \mathfrak{R} : \exists \, \mathbf{u} \in \mathcal{R} \text{ with } |T_{\mathbf{a},\boldsymbol{\varphi}}(\mathbf{u}; N)| \geq N^\alpha \right\}.$$

Then, uniformly over α, we have

$$\#\widetilde{\mathfrak{R}} \leq U N^{s(d)(1 - 2\alpha) + o(1)}.$$

3.2. The measure of the set of large Weyl sums. First, as before, we note that $|T_{\mathbf{a},\boldsymbol{\varphi}}(\mathbf{x}, \mathbf{y}; N)| \leq W_{\mathbf{a},\boldsymbol{\varphi}}(\mathbf{x}, \mathbf{y}; N)$, where $W_{\mathbf{a},\boldsymbol{\varphi}}(\mathbf{x}, \mathbf{y}; N)$ is as in [9, Section 3.3], thus the argument of [9] applies to the sums $T_{\mathbf{a},\boldsymbol{\varphi}}(\mathbf{x}, \mathbf{y}; N)$ as well. In particular, we have versions of [9, Corollaries 3.8 and 3.11] with $T_{\mathbf{a},\boldsymbol{\varphi}}(\mathbf{x}, \mathbf{y}; N)$ which we present below.

In fact we need a slightly more general version of [9, Corollary 3.8]. We recall that the result of [9, Corollary 3.8] is formulated with an $T = N^\alpha$ for

a fixed real α, however examining the argument one can easily see that as in Lemma 3.1, it is uniform with respect to α and thus works for an arbitrary parameter $T \geq 1$. More precisely, we have:

LEMMA 3.2. *Let $1 \leq T \leq N$. Then*

$$\lambda\big(\{ \mathbf{x} \in \mathsf{T}_k : \exists\, \mathbf{y} \in \mathsf{T}_{d-k} \text{ with } |T_{\mathbf{a},\boldsymbol{\varphi}}(\mathbf{x}, \mathbf{y}; N)| \geq T \} \big)$$
$$\leq N^{s(d)+\sigma_k(\boldsymbol{\varphi})+d-k+o(1)} T^{-2s(d)-d+k}.$$

The following result is similar to Lemma 3.2, with $\widetilde{\sigma}_k(\boldsymbol{\varphi})$ instead of $\sigma_k(\boldsymbol{\varphi})$. Let $\mathcal{V}$ be a k-dimensional subspace of $\mathbb{R}^d$. Also recall that $\widetilde{\sigma}_k(\boldsymbol{\varphi})$ is given by (2.2). For the orthogonal projection map $\pi_{\mathcal{V}}$, from [9, Corollary 3.11] we have the following.

LEMMA 3.3. *Let $1 \leq T \leq N$. Then*

$$\lambda\Big(\Big\{ \mathbf{x} \in \mathcal{V} : \sup_{\mathbf{u} \in \mathsf{T}_d \cap \pi_{\mathcal{V}}^{-1}(\mathbf{x})} |T_{\mathbf{a},\boldsymbol{\varphi}}(\mathbf{u}; N)| \geq T \Big\} \Big)$$
$$\leq N^{s(d)+\widetilde{\sigma}_k(\boldsymbol{\varphi})+d-k+o(1)} T^{-2s(d)-d+k}.$$

We now turn to ϑ-Hölder functions $f \colon \mathbb{R}^d \to \mathbb{R}^k$. Indeed, applying similar methods to the proofs of [9, Corollaries 3.8 and 3.11] we obtain Lemma 3.4 below. Since ϑ-Hölder functions do not appear in [9], for the sake of completeness, we outline the proof.

LEMMA 3.4. *Let $f \colon \mathbb{R}^d \to \mathbb{R}^k$ be a ϑ-Hölder function. Then for $1 \leq T \leq N$ we have*

$$\lambda\Big(\Big\{ \mathbf{x} \in \mathbb{R}^k : \sup_{\mathbf{u} \in \mathsf{T}_d \cap f^{-1}(\mathbf{x})} |T_{\mathbf{a},\boldsymbol{\varphi}}(\mathbf{u}; N)| \geq T \Big\} \Big)$$
$$\leq N^{s(d)+\sigma_0(\boldsymbol{\varphi})+d-(\delta(\boldsymbol{\varphi})+1)\vartheta k+o(1)} T^{-2s(d)-d+\vartheta k}.$$

PROOF. First of all suppose that $T = N^\alpha$ for some $0 < \alpha < 1$. We fix some sufficiently small $\varepsilon > 0$ and define the set

$$\mathfrak{U} = \bigcup_{\mathcal{R} \in \widetilde{\mathfrak{R}}} \mathcal{R}.$$

For $\mathcal{A} \subseteq \mathbb{R}^d$ denote $f(\mathcal{A}) = \{ f(\mathbf{x}) : \mathbf{x} \in \mathcal{A} \}$. Observe that

$$(3.2) \qquad \Big\{ \mathbf{x} \in \mathbb{R}^k : \sup_{\mathbf{u} \in \mathsf{T}_d \cap f^{-1}(\mathbf{x})} |T_{\mathbf{a},\boldsymbol{\varphi}}(\mathbf{u}; N)| \geq T \Big\} \subseteq f(\mathfrak{U}) \subseteq \bigcup_{\mathcal{R} \in \widetilde{\mathfrak{R}}} f(\mathcal{R}),$$

where $\widetilde{\mathfrak{R}}$ is as in Lemma 3.1.

Since f is ϑ-Hölder, for any $\mathcal{A} \subseteq \mathbb{R}^d$ we obtain

$$\lambda(f(\mathcal{A})) \ll (\operatorname{diam} \mathcal{A})^{\vartheta k},$$

where $\operatorname{diam} \mathcal{A} = \sup\{\|\mathbf{a} - \mathbf{b}\| : \mathbf{a}, \mathbf{b} \in \mathcal{A}\}$. For each $\mathcal{R} \in \widetilde{\mathfrak{R}}$, by (3.1) we have $\operatorname{diam} \mathcal{R} \ll N^{\alpha-1-\delta(\boldsymbol{\varphi})}$. Combining with Lemma 3.1 and the estimate (3.2), we derive

$$\lambda\left(\left\{\mathbf{x} \in \mathbb{R}^k : \sup_{\mathbf{u} \in \mathsf{T}_d \cap f^{-1}(\mathbf{x})} |T_{\mathbf{a},\boldsymbol{\varphi}}(\mathbf{u}; N)| \geq T\right\}\right)$$

$$\leq \#\widetilde{\mathfrak{R}}(\operatorname{diam} \mathcal{R})^{\vartheta k} \ll U N^{s(d)(1-2\alpha)} N^{(\alpha-\delta(\boldsymbol{\varphi})-1-\varepsilon)\vartheta k}.$$

Since ε is arbitrary, recalling the value of U, we now obtain

$$\lambda\left(\left\{\mathbf{x} \in \mathbb{R}^k : \sup_{\mathbf{u} \in \mathsf{T}_d \cap f^{-1}(\mathbf{x})} |T_{\mathbf{a},\boldsymbol{\varphi}}(\mathbf{u}; N)| \geq T\right\}\right)$$

$$\leq N^{s(d)+\sigma_0(\boldsymbol{\varphi})+d-(\delta(\boldsymbol{\varphi})+1)\vartheta k+o(1)} N^{-2s(d)\alpha-d\alpha+\vartheta k\alpha}.$$

Recalling that $T = N^\alpha$, we now obtain the desired result by taking $T = N^\alpha$. $\square$

3.3. Discrepancy and exponential sums.

We recall the classical *Erdős–Turán inequality* (see, for instance, [10, Theorem 1.21]).

LEMMA 3.5. *Let ξ_n, $n \in \mathbb{N}$, be a sequence in $[0,1)$. Then for the discrepancy D_N given by (2.3) and any $G \in \mathbb{N}$, we have*

$$D_N \leq 3\left(\frac{N}{G+1} + \sum_{g=1}^{G} \frac{1}{g}\left|\sum_{n=1}^{N} \mathbf{e}(g\xi_n)\right|\right).$$

4. Proofs of mean value theorems for exponential sums

4.1. Proof of Theorems 2.1, 2.3, 2.4.

Theorems 2.1, 2.3 and 2.4 follow by combining Lemmas 3.2, 3.3 and 3.4, respectively, with Lemma 4.1 below.

LEMMA 4.1. *Let N be a large positive number and $f \colon \mathsf{T}_k \to [0, N]$ be a function. Suppose that there exist positive constants $a < b$ such that for any $1 \leq T \leq N$,*

(4.1) $$\lambda\left(\left\{\mathbf{x} \in \mathsf{T}_k : F(\mathbf{x}) \geq T\right\}\right) \leq N^a T^{-b}.$$

Then for any positive $\rho \leq b$,

$$\int_{\mathsf{T}_k} F(\mathbf{x})^\rho \, d\mathbf{x} \ll N^{\rho a/b} \log N.$$

Proof. Let $R = N^{a/b}$. Note that for $T > R$ we have a nontrivial estimate in (4.1). We partition T_k into sets $\mathcal{X}_\ell$, $\ell \in \mathbb{N}$, where

$$\mathcal{X}_0 = \left\{ \mathbf{x} \in \mathsf{T}_k : F(\mathbf{x}) \leq R \right\},$$

and for $\ell \geq 1$,

$$\mathcal{X}_\ell = \left\{ \mathbf{x} \in \mathsf{T}_k : 2^{\ell-1} R < F(\mathbf{x}) \leq 2^\ell R \right\}.$$

By our assumption (4.1) we have $\lambda(\mathcal{X}_\ell) \ll N^a (2^\ell R)^{-b}$. Clearly for some $L = O(\log N)$ we have $\mathcal{X}_\ell = \emptyset$. Therefore,

$$\int_{\mathsf{T}_k} F(\mathbf{x})^\rho \, d\mathbf{x} = \sum_{\ell=0}^{L} \int_{\mathcal{X}_\ell} F(\mathbf{x})^\rho \, d\mathbf{x}$$

$$\ll R^\rho + \sum_{\ell=1}^{L} (2^\ell R)^\rho N^a (2^\ell R)^{-b} \ll R^\rho + N^a R^{\rho-b} \sum_{\ell=1}^{L} 2^{\ell(\rho-b)}.$$

By the choice of $R = N^{a/b}$ and the condition that $\rho \leq b$ we obtain the desired bound. $\square$

4.2. Proof of Corollary 2.2. For $K \in \mathbb{Z}$, recall that Weyl sums over short intervals are defined as follows

$$S_d(\mathbf{u}; K, N) = \sum_{n=K+1}^{K+N} \mathbf{e}(u_1 n + \cdots + u_d n^d).$$

We write

$$S_d(\mathbf{u}; K, N) = \sum_{n=1}^{N} \mathbf{e}(u_1(n+K) + \cdots + u_d(n+K)^d).$$

We observe that in the polynomial identity

$$u_1(T+K) + \cdots + u_d(T+K)^d = v_0 + v_1 T + \cdots + v_{d-1} T^{d-1} + u_d T^d \in \mathbb{R}[T],$$

where for $j = 0, 1, \ldots, d-1$, each v_j, depends only on $u_1, \ldots, u_d$ and K. It follows that

$$(4.2) \qquad \sup_{K \in \mathbb{Z}} |S_d(\mathbf{u}; K, N)| \leq \sup_{(v_1, \ldots, v_{d-1}) \in \mathsf{T}_{d-1}} \left| S_d((v_1, \ldots, v_{d-1}, u_d); N) \right|.$$

Note that for any fixed u_d for any $(u_1, \ldots, u_{d-1}) \in \mathsf{T}_{d-1}$ the estimate 4.2 holds for $\mathbf{u} = (u_1, \ldots, u_d)$. Thus we obtain

$$\int_{\mathsf{T}_d} \sup_{K \in \mathbb{Z}} |S_d(\mathbf{u}; K, N)|^{\rho} \, d\mathbf{u}$$

$$= \int_0^1 \left(\int_{\mathsf{T}_{d-1}} \sup_{K \in \mathbb{Z}} |S_d(\mathbf{u}; K, N)|^{\rho} du_1 \ldots du_{d-1} \right) du_d$$

$$\leq \int_0^1 \sup_{(v_1, \ldots, v_{d-1}) \in \mathsf{T}_{d-1}} |S_d((v_1, \ldots, v_{d-1}, u_d); N)|^{\rho} du_d.$$

Hence Theorem 2.1, applied $k = 1$, $\varphi_1(T) = T^d$, $\varphi_i(T) = T^{i-1}$ for $i = 2, \ldots, d$, and thus with $\sigma_1(\boldsymbol{\varphi}) = d(d-1)/2$, yields the desired bound.

5. Proofs of mean value theorems for the discrepancy

5.1. Proof of Theorem 2.5. For any $\mathbf{x} \in \mathsf{T}_k, \mathbf{y} \in \mathsf{T}_{d-k}$, by Lemma 3.5 for any $G \in \mathbb{N}$ we obtain

$$D_{\boldsymbol{\varphi}}(\mathbf{x}, \mathbf{y}; N) \ll \frac{N}{G} + \sum_{g=1}^G \frac{1}{g} \left| T_{\boldsymbol{\varphi}}(g\mathbf{x}, g\mathbf{y}; N) \right|,$$

and therefore

$$(5.1) \qquad \sup_{\mathbf{y} \in \mathsf{T}_{d-k}} D_{\boldsymbol{\varphi}}(\mathbf{x}, \mathbf{y}; N) \ll \frac{N}{G} + \sum_{g=1}^G \frac{1}{g} \sup_{\mathbf{y} \in \mathsf{T}_{d-k}} \left| T_{\boldsymbol{\varphi}}(g\mathbf{x}, \mathbf{y}; N) \right|.$$

We now use the following invariant property of Lebesgue measure on a torus.

LEMMA 5.1. *Let $F \colon \mathsf{T}_k \to [0, N]$ be a continuous function. Then for any integer $g \neq 0$ we have*

$$(5.2) \qquad \int_{\mathsf{T}_k} F(g\mathbf{x}) \, d\mathbf{x} = \int_{\mathsf{T}_k} F(\mathbf{x}) \, d\mathbf{x}.$$

PROOF. For any Borel set $\mathcal{A} \subseteq \mathsf{T}_k$ and any integer $g \neq 0$, we have (for a proof see [14, Section 3])

$$\lambda\left(\left\{ \mathbf{x} \in \mathsf{T}_k : g\mathbf{x} \in \mathcal{A} \right\} \right) = \lambda(\mathcal{A}),$$

which is the same as the identity

$$\int_{\mathsf{T}_k} \mathbf{1}_{\mathcal{A}}(g\mathbf{x}) \, d\mathbf{x} = \int_{\mathsf{T}_k} \mathbf{1}_{\mathcal{A}}(\mathbf{x}) \, d\mathbf{x},$$

where $\mathbf{1}_{\mathcal{A}}$ is the characteristic function of $\mathcal{A}$. Thus (5.2) holds when $F = \mathbf{1}_{\mathcal{A}}$. It follows that the identity (5.2) still holds when F is a finite linear combination of characteristic functions, that is

$$F(\mathbf{x}) = \sum_{j=1}^{J} a_j \mathbf{1}_{\mathcal{A}_j}(\mathbf{x}),$$

for sets $\mathcal{A}_j \subseteq \mathsf{T}_k$, $j = 1, \ldots, J$. Since any continuous function can be arbitrarily approximated by a finite linear combination of such functions, the desired identity follows. $\square$

Here $\|f(\mathbf{x})\|_{\rho}$ denotes the $L^{\rho}(\mathsf{T}_k)$-norm of a function f on T_k. Then by (5.1) and the Minkowski inequality,

$$(5.3) \qquad \left\| \sup_{\mathbf{y} \in \mathsf{T}_{d-k}} D_{\boldsymbol{\varphi}}(\mathbf{x}, \mathbf{y}; N) \right\|_{\rho} \ll \frac{N}{G} + \sum_{g=1}^{G} \frac{1}{g} \left\| \sup_{\mathbf{y} \in \mathsf{T}_{d-k}} \left| T_{\boldsymbol{\varphi}}(g\mathbf{x}, \mathbf{y}; N) \right| \right\|_{\rho}.$$

For any positive integer $g \neq 0$, Lemma 5.1 implies

$$\left\| \sup_{\mathbf{y} \in \mathsf{T}_{d-k}} \left| T_{\boldsymbol{\varphi}}(g\mathbf{x}, \mathbf{y}; N) \right| \right\|_{\rho} = \left\| \sup_{\mathbf{y} \in \mathsf{T}_{d-k}} \left| T_{\boldsymbol{\varphi}}(\mathbf{x}, \mathbf{y}; N) \right| \right\|_{\rho},$$

which together with (5.3) and Theorem 2.1 yields

$$\left\| \sup_{\mathbf{y} \in \mathsf{T}_{d-k}} D_{\boldsymbol{\varphi}}(\mathbf{x}, \mathbf{y}; N) \right\|_{\rho} \ll \frac{N}{G} + \left\| \sup_{\mathbf{y} \in \mathsf{T}_{d-k}} \left| T_{\boldsymbol{\varphi}}(\mathbf{x}, \mathbf{y}; N) \right| \right\|_{\rho} \log G$$

$$\ll \frac{N}{G} + N^{\mu(\boldsymbol{\varphi}, k) + o(1)} \log G.$$

Choosing $G = N^{1 - \mu(\boldsymbol{\varphi}, k)}$, we derive the desired result.

5.2. Proof of Corollary 2.6. Recall that $D_d(\mathbf{u}; K, N)$ is the discrepancy of the sequence of fractional parts

$$\{u_1 n + \cdots + u_d n^d\}, \quad n = K + 1, \ldots, K + N.$$

Clearly this sequence is the same as

$$\{u_1(n + K) + \cdots + u_d(n + K)^d\}, \quad n = 1, \ldots, N,$$

and thus as before, see (4.2), we see that this sequence is the same as

$$\{v_0 + v_1 n + \cdots + v_{d-1} n^{d-1} + u_d n^d\}, \quad n = 1, \ldots, N,$$

where for $j = 0, 1, \ldots, d-1$, each v_j, depends only on $u_1, \ldots, u_d$ and K. Furthermore let $\mathbf{u}^* = (v_1, \ldots, v_{d-1}, u_d)$. It is not hard to see that the influence of the discarded constant term v_0 can be absorbed in a constant factor and does not change the order of magnitude of the discrepancy. More, precisely, we have

$$D_d(\mathbf{u}^*; N) \ll D_d(\mathbf{u}; K, N) \ll D_d(\mathbf{u}^*; N),$$

where the implied constant is absolute. It follows that

$$\sup_{K \in \mathbb{Z}} D_d(\mathbf{u}; K, N) \ll \sup_{(v_1, \ldots, v_{d-1}) \in \mathsf{T}_{d-1}} D_d((v_1, \ldots, v_{d-1}, u_d), N).$$

Using similar arguments as in the proof of Corollary 2.2 and applying Theorem 2.5, with $k = 1$, $\varphi_1(T) = T^d$, $\varphi_i(T) = T^{i-1}$ for $i = 2, \ldots, d$, and thus with $\sigma_1(\boldsymbol{\varphi}) = d(d-1)/2$, we obtain the desired bound.

6. Comments and open questions

6.1. Special cases of $d = 2$ and $d = 3$. We show that for special cases of $d = 2$ and $d = 3$ for the moments

(6.1)
$$\begin{cases} \mathcal{M}_{2,2}(N) = \int_0^1 \sup_{y \in [0,1]} \left| \sum_{n=1}^N \mathbf{e}(xn^2 + yn) \right|^2 dx, \\ \mathcal{M}_{3,4}(N) = \int_0^1 \sup_{y,z \in [0,1]} \left| \sum_{n=1}^N \mathbf{e}(xn^3 + yn^2 + zn) \right|^4 dx, \end{cases}$$

we have better bounds (nearly optimal) than the bounds in Theorem 2.1.

Applying Theorem 2.1 with $d = 2$, $k = 1$, $\sigma_k(\boldsymbol{\varphi}) = 1$, $\rho = 2$ and with $d = 3$, $k = 1$, $\sigma_k(\boldsymbol{\varphi}) = 3$, $\rho = 4$ we derive that

$$\mathcal{M}_{2,2}(N) \le N^{10/7+o(1)} \quad \text{and} \quad \mathcal{M}_{3,4}(N) \le N^{22/7+o(1)},$$

respectively. On the other hand, we have the lower bounds

$$\mathcal{M}_{2,2}(N) \ge \int_0^1 \int_0^1 \left| \sum_{n=1}^N \mathbf{e}(xn^2 + yn) \right|^2 dy\, dx = N,$$

and

$$\mathcal{M}_{3,4}(N) \ge \int_0^1 \int_0^1 \int_0^1 \left| \sum_{n=1}^N \mathbf{e}(xn^3 + yn^2 + zn) \right|^4 dz\, dy\, dx \ge 2N^2 + O(N),$$

see, for example, [13, Section 1].

Now we use a different way to bound the square mean values (6.1) and reduce the gap between the above lower and upper bounds.

For $d = 2$ we need the following well known inequality, for a proof see [11, Inequality (8.8)]:

$$\sup_{y \in [0,1]} \left| \sum_{n=1}^{N} \mathbf{e}(xn^2 + yn) \right|^2 \ll N + \sum_{h=1}^{N} \min\left\{ \frac{1}{\|2hx\|}, N \right\},$$

where $\|x\| = \min\{|x - n| : n \in \mathbb{Z}\}$ denotes the distance of x from the nearest integer. Moreover, for any positive integer h, applying Lemma 5.1 we obtain that uniformly over h, we have

$$\int_0^1 \min\left\{ \frac{1}{\|2hx\|}, N \right\} dx = \int_0^1 \min\left\{ \frac{1}{\|x\|}, N \right\} dx \ll \log N,$$

hence

$$\mathcal{M}_{2,2}(N) \ll N \log N.$$

Note that this is tight except, possibly, for the logarithm factor.

For $d = 3$ we use that by [11, Proposition 8.2]:

$$\sup_{y,z \in [0,1]} \left| \sum_{n=1}^{N} \mathbf{e}(xn^3 + yn^2 + zn) \right|^4$$

$$\ll N \sum_{g=-N}^{N} \sum_{h=-N}^{N} \min\left\{ \frac{1}{\|6ghx\|}, N \right\} \ll N^3 + N \sum_{\substack{g,h=-N \\ gh \neq 0}}^{N} \min\left\{ \frac{1}{\|6ghx\|}, N \right\},$$

and arguing as before we obtain

$$\mathcal{M}_{3,4}(N) \ll N^3 \log N.$$

It is natural to ask whether similar arguments can improve Theorem 2.1 for higher degrees. Note that typically methods based on the Vinogradov mean value theorem yield better bounds for higher degrees. Indeed this also happens here, which means that Theorem 2.1 gives stronger bounds than using [11, Proposition 8.2] for $d \geq 4$ and applying the above arguments.

6.2. Open questions. It is interesting to try to use the "self-improving" idea of [9] and in particular [9, Corollary 3.9] to obtain stronger results in the case when one of the polynomials $\varphi_{k+1}, \ldots, \varphi_d$ is linear. There are however some obstacles which the authors have not been able to overcome.

It is also natural to study the mixed mean values

$$M_{k,\rho,\tau}(\mathbf{a}, \boldsymbol{\varphi}, N) = \int_{\mathbf{x} \in [0,1]^k} \left(\int_{\mathbf{y} \in [0,1]^{d-k}} \left| T_{\mathbf{a}, \boldsymbol{\varphi}}(\mathbf{x}, \mathbf{y}; N) \right|^\tau d\mathbf{y} \right)^\rho d\mathbf{x},$$

and obtain bounds which are stronger than those following from the trivial inequality

$$M_{k,\rho,\tau}(\mathbf{a}, \boldsymbol{\varphi}, N) \le M_{k,\rho\tau}(\mathbf{a}, \boldsymbol{\varphi}, N)$$

and Theorem 2.1. The case of the classical Weyl sums is of special interest.

References

[1] T. C. Anderson, B. Cook, K. Hughes and A. Kumchev, On the ergodic Waring–Goldbach problem, arxiv.org/abs/1703.02713 (2017).

[2] T. C. Anderson, B. Cook, K. Hughes and A. Kumchev, Improved ℓ^p-boundedness for integral k-spherical maximal functions, *Discrete Anal.* (2018), Paper No. 10, 18 pp.

[3] T. C. Anderson, A. Kumchev and E. A. Palsson, Discrete maximal operators over surfaces of higher codimension, arxiv.org/abs/2006.09968 (2020).

[4] R. C. Baker, Metric number theory and the large sieve, *J. London Math. Soc.*, **24** (1981), 34–40.

[5] J. Bourgain, On the Vinogradov mean value, *Proc. Steklov Math. Inst.*, **296** (2017), 30–40.

[6] J. Bourgain, C. Demeter and L. Guth, Proof of the main conjecture in Vinogradov's mean value theorem for degrees higher than three, *Ann. Math.*, **184** (2016), 633–682.

[7] J. Brüdern, Approximations to Weyl sums, *Acta Arith.*, **184** (2018), 287–296.

[8] J. Brüdern and D. Daemen, Imperfect mimesis of Weyl sums, *Internat. Math. Res. Notices*, **2009** (2009), 3112–3126.

[9] C. Chen and I. E. Shparlinski, New bounds of Weyl sums, *Internat. Math. Res. Notices* (to appear).

[10] M. Drmota and R. Tichy, *Sequences, Discrepancies and Applications*, Springer-Verlag (Berlin, 1997).

[11] H. Iwaniec and E. Kowalski, *Analytic Number Theory*, Amer. Math. Soc. (Providence, RI, 2004).

[12] H. Weyl, Über die Gleichverteilung von Zahlen mod Eins, *Math. Ann.*, **77** (1916), 313–352.

[13] T. D. Wooley, The cubic case of the main conjecture in Vinogradov's mean value theorem, *Adv. in Math.*, **294** (2016), 532–561.

[14] T. D. Wooley, Perturbations of Weyl sums, *Internat. Math. Res. Notices*, **2016** (2016), 2632–2646.

[15] T. D. Wooley, Nested efficient congruencing and relatives of Vinogradov's mean value theorem, *Proc. London Math. Soc.*, **118** (2019), 942–1016.