DISCRETE MULTILINEAR MAXIMAL OPERATORS AND PINNED SIMPLICES

NEIL LYALL ÁKOS MAGYAR ALEX NEWMAN PETER WOOLFITT

Abstract. Revisit [12] using theta functions and generalize to simplices. Provide dimensional improvements on ℓ^2 results from [1] and [7], and on the main result from [15] and strengthen to pinned.

1. Introduction

1.1. Simplices in dense subsets of $\mathbb{Z}^d$.

Recall that the *upper Banach density* of a set $A \subseteq \mathbb{Z}^d$ is defined by

$$\delta^*(A) = \lim_{N\to\infty} \sup_{t\in\mathbb{Z}^d} \frac{|A\cap(t+Q(N))|}{|Q(N)|},$$

where $|\cdot|$ denotes counting measure on $\mathbb{Z}^d$ and $Q(N)$ the discrete cube $[-N/2,N/2]^d\cap\mathbb{Z}^d$.

In light of the fact that the square of the distance between any two distinct points in $\mathbb{Z}^d$ is always a positive integer we also introduce the convenient notation $\sqrt{\mathbb{N}} := \{\lambda : \lambda>0 \text{ and } \lambda^2\in\mathbb{Z}\}$.

In [14] the second author established the following result on the existence of unpinned two point configurations (distances) in dense subsets of the integer lattice.

Theorem A (Magyar [14]). *Let $A\subseteq\mathbb{Z}^d$ with $d\geq 5$. If $\delta^*(A)>0$, then there exist an integer $q=q(\delta^*(A))$ and $\lambda_0=\lambda_0(A)$ such that for all $\lambda\in\sqrt{\mathbb{N}}$ with $\lambda\geq\lambda_0$ there exist a pair of points $\{x,x+y\}\subseteq A$ with $|y|=q\lambda$.*

The approach taken in [14] was an adaptation of Bourgain's in [2] to the analogous problem in the continuous setting of $\mathbb{R}^d$. In [15] the second author adapted this further to establish the following analogous result for non-degenerate k-simplices. Recall that for any $1\leq k\leq d$ we refer to a configuration $\Delta=\{v_0=0,v_1,\dots,v_k\}\subseteq\mathbb{Z}^d$ as a non-degenerate k-simplex if the vectors $v_1,\dots,v_k$ are linearly independent.

Theorem B (Magyar [15]). *Let $k\geq 2$, $A\subseteq\mathbb{Z}^d$ with $d\geq 2k+5$, and $\Delta=\{0,v_1,\dots,v_k\}\subseteq\mathbb{Z}^d$ be a non-degenerate k-simplex. If $\delta^*(A)>0$, there exists an integer $q=q(\delta^*(A))$ and $\lambda_0=\lambda_0(A,\Delta)$ such that for all $\lambda\in\sqrt{\mathbb{N}}$ with $\lambda\geq\lambda_0$ there exist $x\in A$ with $x+\Delta'\subseteq A$ for some $\Delta'=\{0,y_1,\dots,y_k\}\simeq\lambda q\Delta$.*

In the theorem above, and throughout this article, we say that two configurations $\lambda\Delta=\{0,\lambda v_1,\dots,\lambda v_k\}$ and $\Delta'=\{0,y_1,\dots,y_k\}$ in $\mathbb{Z}^d$ are *isometric*, and write $\Delta'\simeq\lambda\Delta$, if $|y_i-y_j|=\lambda|v_i-v_j|$ for all $0\leq i,j\leq k$.

In this article we establish an improvement on the dimension condition in Theorem B above from $d\geq 2k+5$ to $d\geq 2k+3$ and simultaneously establish a stronger *pinned* variant, namely

Theorem 1. *Let $k\geq 1$, $A\subseteq\mathbb{Z}^d$ with $d\geq 2k+3$, and $\Delta=\{0,v_1,\dots,v_k\}\subseteq\mathbb{Z}^d$ be a non-degenerate k-simplex. If $\delta^*(A)>0$, there exists an integer $q=q(\delta^*(A))$ and $\lambda_0=\lambda_0(A,\Delta)$ such that for any $\lambda_1\geq\lambda_0$ there exists a fixed $x\in A$ such that for all $\lambda\in[\lambda_0,\lambda_1]\cap\sqrt{\mathbb{N}}$ one has $x+\Delta'\subseteq A$ for some $\Delta'=\{0,y_1,\dots,y_k\}\simeq\lambda q\Delta$.*

Remark. The threshold λ_0 in the results above cannot be taken to depend on $\delta^*(A)$ only. Indeed, for any positive integers q and M the set $(Q_{qM}\cap\mathbb{Z}^d)+(4dqM\mathbb{Z})^d$ will have density $(4d)^{-d}$ but never contain pairs $\{x,x+y\}$ with $|y|=qdM$. Since A could fall entirely into a fixed congruence class of some integer $1\leq r\leq\delta^*(A)^{-1/d}$ the value of q in the results above must be divisible by the least common multiple of all integers $1\leq r\leq\delta^*(A)^{-1/d}$. Indeed if $A=(r\mathbb{Z})^d$ with $1\leq r\leq\delta^{-1/d}$ then A will have upper Banach density at least δ, but the distance between any two points $x,y\in A$ will always take the form $r\lambda$ for some $\lambda\in\sqrt{\mathbb{N}}$.

We note that the case $k=1$ of Theorem 1 was already established by the first two authors in [10]. To the best of our knowledge, there have been no previous results addressing *pinned* simplices in $\mathbb{Z}^d$ in any dimension when $k\geq 2$. Statements such as these are intimately connected with estimates for discrete maximal averages.

2010 *Mathematics Subject Classification.* 11B30.

The first and second authors were partially supported by grants NSF-DMS 1702411 and NSF-DMS 1600840, respectively.

1.2. **Discrete multilinear maximal averages associated to simplices.** The study of discrete analogues of central constructs of Euclidean harmonic analysis, initiated by Bourgain [3, 4, 5, 6], has grown into a vast, active area of research.

An important result in this development is the ℓ^p-boundedness of the so-called discrete spherical maximal function [13]. For any $\lambda \in \sqrt{\mathbb{N}}$ we let $S_\lambda = \{y \in \mathbb{Z}^d : |y| = \lambda\}$ denote the discrete sphere of radius λ centered at the origin. For $f : \mathbb{Z}^d \to \mathbb{R}$ we then define the discrete spherical averages

$$\mathcal{A}_\lambda f(x) = |S_\lambda|^{-1} \sum_{y \in S_\lambda} f(x+y).$$

noting that if $d \geq 5$, then $c_d \lambda^{d-2} \leq |S_\lambda| \leq C_d \lambda^{d-2}$ for some constants $0 < c_d < C_d < \infty$, see [18]. In [13] it was shown that for $p > d/(d-2)$ one has the following maximal function estimate

$$\big\| \sup_{\lambda \geq 1} |\mathcal{A}_\lambda f| \big\|_p \leq C_{p,d} \, \|f\|_p$$

where $\|f\|_p = (\sum_x |f(x)|^p)^{1/p}$ denotes the $\ell^p(\mathbb{Z}^d)$ norm of the function f.

In [12] the authors gave a new direct proof of ℓ^2-boundedness of the discrete spherical maximal function that neither relies on abstract transference theorems nor on delicate asymptotic for the Fourier transform of discrete spheres. Implicit in that paper is the fact that ℓ^2-boundedness follows as a consequence of stronger refined "mollified" estimates in which one obtains gains in ℓ^2 over suitably large scales when applied to functions whose Fourier transform is localized away from rational points with small denominators.

Recall that for $f \in \ell^1(\mathbb{Z}^d)$ we define its Fourier transform $\widehat{f} : \mathbb{T}^d \to \mathbb{C}$ by $\widehat{f}(\xi) = \sum_{x \in \mathbb{Z}^d} f(x) e^{-2\pi i x \cdot \xi}$. For each $\eta > 0$ we define

$$q_\eta := \operatorname{lcm}\{1 \leq q \leq \eta^{-2}\}$$

and for any $L \geq q_\eta$ we let

$$\Omega_{\eta,L} = \{\xi \in \mathbb{T}^d : \xi \in [-L^{-1}, L^{-1}]^d + (q_\eta^{-1}\mathbb{Z})^d\}.$$

Key to the proof of Theorem 1 is an extension of the approach from [12] to multilinear maximal operators associated to simplices. Given a non-degenerate k-simplex $\Delta = \{v_0 = 0, v_1, \ldots, v_k\} \subseteq \mathbb{Z}^d$ and $\lambda \in \sqrt{\mathbb{N}}$, we let

$$S_{\lambda\Delta} := \{(y_1, \ldots, y_k) \in \mathbb{Z}^{dk} : \ \Delta' = \{0, y_1, \ldots, y_k\} \simeq \lambda\Delta\}.$$

For functions $f_1, \ldots, f_k : \mathbb{Z}^d \to \mathbb{C}$ we then define the multilinear averaging operator

$$\mathcal{A}_{\lambda\Delta}(f_1, \ldots, f_k)(x) = |S_{\lambda\Delta}|^{-1} \sum_{(y_1,\ldots,y_k) \in S_{\lambda\Delta}} f_1(x+y_1) \cdots f_k(x+y_k)$$

noting that if $d \geq 2k+3$ and $\lambda \in \sqrt{\mathbb{N}}$, then

$$c_\Delta \, \lambda^{dk-k(k+1)} \leq |S_{\lambda\Delta}| \leq C_\Delta \, \lambda^{dk-k(k+1)} \tag{1}$$

for some constants $0 < c_\Delta < C_\Delta < \infty$, see [9] or [15].

Note that for $k = 1$ and $v_1 = (1, 0, \ldots, 0)$ we have that $S_{\lambda\Delta} = S_\lambda$ and hence $\mathcal{A}_{\lambda\Delta}(f) = \mathcal{A}_\lambda(f)$.

The ℓ^p mapping properties of the maximal operators corresponding to these averages were considered in [1] and [7]. Here we establish the following particular ℓ^2-estimates, the first non-trivial estimates of any type for such operators in dimensions lower that $d = 2k+5$ when $k \geq 2$. The stronger refined ℓ^2 estimate (3), in addition to implying (2), plays a crucial role in our proof of Theorem 1.

Theorem 2. *If $k \geq 1$, $d \geq 2k+3$, and $\Delta = \{0, v_1, \ldots, v_k\} \subseteq \mathbb{Z}^d$ be a non-degenerate k-simplex, then*

$$\big\| \sup_{\lambda \geq 1} |\mathcal{A}_{\lambda\Delta}(f_1, \ldots, f_k)| \big\|_2 \leq C_{d,\Delta} \|f_1\|_2 \cdots \|f_k\|_2. \tag{2}$$

In fact, for any $\eta > 0$, and $L \geq q_\eta^4$, we have

$$\Big\| \sup_{\lambda \geq \eta^{-2} L} |\mathcal{A}_{\lambda\Delta}(f_1, \ldots, f_k)| \Big\|_2 \leq C_{d,\Delta} \frac{\eta}{\log \eta^{-1}} \, \|f_1\|_2 \cdots \|f_k\|_2 \tag{3}$$

whenever $\operatorname{supp} \widehat{f_j} \subseteq \Omega^c_{\eta,L}$ *for some $1 \leq j \leq k$, where $\Omega^c_{\eta,L}$ denotes the complement of $\Omega_{\eta,L}$.*

Estimate (3) in the case $k = 1$ was originally established in joint work with Magyar [10] via an adaptation of the transference methods from [13].

Remark. The constants $C_{d,\Delta}$ above can in fact be taken independent of Δ.

2. Proof of Theorem 1

2.1. **Reduction to uniform distributed sets.** In light of the observation made after Theorem 1 above regarding the sensitivity of this problem to the local structure of A, it is natural to first consider the case when A is, in a suitable sense, well distributed in small congruence classes. In fact, this approach ultimately leads directly to a proof of Theorem 1.

Following [10] we define $A \subseteq \mathbb{Z}^d$ to be *η-uniformly distributed (modulo q_η)* if, for some $\eta > 0$, its *relative* upper Banach density on any residue class modulo q_η never exceeds $(1+\eta^4)$ times its density on $\mathbb{Z}^d$, namely if

$$\delta^*(A \,|\, s + (q_\eta\mathbb{Z})^d) \leq (1+\eta^4)\, \delta^*(A)$$

for all $s \in \{1, \ldots, q_\eta\}^d$. A straightforward density increment argument allows one to deduce Theorem 1 from the following analogue for η-uniformly distributed subsets of $\mathbb{Z}^d$.

Proposition 1. *Let $\varepsilon > 0$, $0 < \eta \ll \varepsilon^2$ and $k \geq 1$.*

If $A \subseteq \mathbb{Z}^d$ with $d \geq 2k+3$ is η-uniformly distributed, and $\Delta = \{0, v_1, \ldots, v_k\} \subseteq \mathbb{Z}^d$ is a non-degenerate k-simplex, then there exist $\lambda_0 = \lambda_0(A, \Delta, \eta)$ such that for any $\lambda_1 \geq \lambda_0$ there exists a fixed $x \in A$ such that

$$\mathcal{A}_{\lambda\Delta}(1_A, \ldots, 1_A)(x) > \delta^*(A)^k - \varepsilon$$

for all $\lambda \in [\lambda_0, \lambda_1] \cap \sqrt{\mathbb{N}}$, noting that

$$\mathcal{A}_{\lambda\Delta}(1_A, \ldots, 1_A)(x) = |S_{\lambda\Delta}|^{-1}\ |\{(y_1, \ldots, y_k) \in \mathbb{Z}^{dk} \,:\, x + \Delta' \subseteq A \text{ with } \Delta' = \{0, y_1, \ldots, y_k\} \simeq \lambda\Delta\}|.$$

In Proposition 1 above, and throughout this article, we use the notation $\alpha \ll \beta$ to denote that $\alpha \leq c\beta$ for some suitably small constant $c > 0$.

Proposition 1 in fact implies the following stronger optimal formulation of Theorem 1.

Corollary 1. *Let $k \geq 1$, $A \subseteq \mathbb{Z}^d$ with $d \geq 2k+3$, and $\Delta = \{0, v_1, \ldots, v_k\} \subseteq \mathbb{Z}^d$ be a non-degenerate k-simplex. For any $\varepsilon > 0$, there exists an integer $q = q(\varepsilon, d)$ and $\lambda_0(A, \Delta, \varepsilon)$ such that for any $\lambda_1 \geq \lambda_0$ there exists a fixed x such that*

$$|S_{\lambda\Delta}|^{-1}\ |\{(y_1, \ldots, y_k) \in (q\mathbb{Z})^{dk} \,:\, x + \Delta' \subseteq A \text{ with } \Delta' = \{0, y_1, \ldots, y_k\} \simeq \lambda q\Delta\}| > \delta^*(A)^k - \varepsilon \tag{4}$$

for all $\lambda \in [\lambda_0, \lambda_1] \cap \sqrt{\mathbb{N}}$.

Remark. By considering sets A of the form $\bigcup_{s \in \{1, \ldots, q\}^d} A_s$ with each set A_s a "random" subset of the congruence class $s + (q\mathbb{Z})^d$ one can further easily see that conclusion (4) above is in general best possible.

Proof that Proposition 1 implies Corollary 1. Let $0 < \varepsilon \leq \delta \leq 1$ and $A \subseteq \mathbb{Z}^d$ with $d \geq 2k+3$. To prove Corollary 1 it is enough to prove that if $\delta^*(A) \geq \delta$ then there exists $\lambda_0 = \lambda_0(A, \Delta, \varepsilon)$ and $q = q(\varepsilon, d)$ such that for any $\lambda_1 \geq \lambda_0$ there exists a fixed $x \in A$ such that (4) holds for all $\lambda \in \sqrt{\mathbb{N}}$ with $\lambda_0 \leq \lambda \leq \lambda_1$.

Let $0 < \eta \ll \varepsilon^2$. We prove the above for $\delta_m := (1+\eta^4)^{-m}$ inductively for all $m \geq 0$, using Proposition 1. For $m = 0$ the statement is trivial as $\delta^*(A) = \delta_0 = 1$ and hence A contains arbitrarily large cubic grids. Suppose it holds for $\delta = \delta_m$ and assume that $\delta^*(A) \geq \delta_{m+1}$. If A is η-uniformly distributed then the result holds for $\delta = \delta_{m+1}$ by Proposition 1. In the opposite case there is an $s \in \mathbb{Z}^d$ so that $\delta^*(A \,|\, s + (q_\eta\mathbb{Z})^d) > (1+\eta^4)\, \delta$. Let $\phi : s + (q_\eta\mathbb{Z})^d \to \mathbb{Z}^d$ be defined by $\phi(x) := q_\eta^{-1}(x - s)$ and let $A' := \phi(A)$. Then $\delta^*(A') \geq \delta_m$ thus (4) holds for A' and $\delta = \delta_m$, with some $q' = q'(\varepsilon, d)$ and $x' \in A'$. Note that

$$\begin{aligned}|\{(y_1, \ldots, y_k) \in (q'\mathbb{Z})^{dk} \,:\, x' + \Delta' \subseteq A' \text{ with } \Delta' = \{0, y_1, \ldots, y_k\} \simeq q'\lambda\Delta\}|\\ = |\{(y_1, \ldots, y_k) \in (q_\eta q'\mathbb{Z})^{dk} \,:\, q_\eta x' + \Delta' \subseteq A' \text{ with } \Delta' = \{0, y_1, \ldots, y_k\} \simeq q_\eta q'\lambda\Delta\}|\end{aligned}$$

which implies that (4) holds for A, $\delta = \delta_{m+1}$ with $q = q_\eta q'$ and $x = q_\eta x' + s$. $\square$

2.2. **Proof of Proposition 1.** Before proving proving Proposition 1 we need two preparatory lemmas.

We refer to a subset $Q \subseteq \mathbb{Z}^d$ as a cube of sidelength $\ell(Q) = N$ if

$$Q = t_0 + Q_N$$

for some $t_0 \in \mathbb{Z}^d$, where as usual $Q_N = [-N/2, N/2]^d$.

Definition ($U^1_{q,L}(Q)$-norm)**.** For any cube $Q \subseteq \mathbb{Z}^d$, integers $1 \ll q \ll L \ll \ell(Q)$, and functions $f : Q \to \mathbb{R}$ we define

$$\|f\|_{U^1_{q,L}(Q)} = \Big(\frac{1}{|Q|}\sum_{t\in\mathbb{Z}^d} |f * \chi_{q,L}(t)|^2\Big)^{1/2} \tag{5}$$

where $\chi_{q,L}$ denotes the normalized characteristic function of the cubes $Q_{q,L} := Q_L \cap (q\mathbb{Z})^d$, namely

$$\chi_{q,L}(x) = \begin{cases} \left(\frac{q}{L}\right)^d & \text{if } \ x \in (q\mathbb{Z})^d \cap [-\frac{L}{2}, \frac{L}{2}]^d \\ 0 & \text{otherwise} \end{cases}. \tag{6}$$

In (5) above and in the sequel we denote the convolution $f * g$ of two functions f and g by

$$f * g(x) := \sum_{y\in\mathbb{Z}^d} f(x-y)g(y).$$

We note that the $U^1_{q,L}(Q)$-norm measures the mean square oscillation of a function with respect to cubic grids of size L and gap q. The first key ingredient in our proof of Proposition 1 is the simple, yet significant, observation from [10] that subsets of $\mathbb{Z}^d$ with positive upper Banach density that are η-uniformly distributed are also, in a precise sense, uniformly distributed at certain scales.

Lemma 1 (Consequence of Lemmas 1 and 2 in [10])**.** *Let $\eta > 0$ and $A \subseteq \mathbb{Z}^d$ be η-uniformly distributed with $\delta := \delta^*(A) > 0$. There exists a positive integer $L = L(A, \eta)$ and cubes $Q \subseteq \mathbb{Z}^d$ of arbitrarily large sidelength $\ell(Q)$ with $\ell(Q) \geq \eta^{-4}L$ such that*

$$|A \cap Q| \geq (\delta - O(\eta))|Q| \tag{7}$$

and

$$\|(1_A - \delta)1_Q\|_{U^1_{q_\eta,L}(Q)} = O(\eta). \tag{8}$$

The second key ingredient in our proof of Proposition 1 is the following *maximal* variant of a so-called generalized von-Neumann-type inequality, which follows in a straightforward manner from Theorem 2.

Lemma 2 (Corollary of Theorem 2)**.** *Let $k \geq 1$, $d \geq 2k + 3$, and $\Delta = \{0, v_1, \dots, v_k\} \subseteq \mathbb{Z}^d$ be a non-degenerate k-simplex. For any $\eta > 0$, positive integer L, cube $Q \subseteq \mathbb{Z}^d$ with sidelength $N \geq \eta^{-6}L$, and functions $f_1, \dots, f_k : Q \to [-1, 1]$ we have*

$$\frac{1}{|Q|}\sum_{x\in\mathbb{Z}^d} \sup_{\eta^{-3}L\leq\lambda\leq\eta^3 N} \big|\mathcal{A}_{\lambda\Delta}(f_1, \dots, f_k)(x)\big| \leq C_{d,\Delta}\Big(\min_{1\leq j\leq k} \|f_j\|_{U^1_{q_\eta,L}(Q)} + O(\eta)\Big).$$

Proof. By Cauchy-Schwarz, it suffices to prove the stronger estimate

$$\Big(\frac{1}{|Q|}\sum_{x\in\mathbb{Z}^d} \sup_{\eta^{-3}L\leq\lambda\leq\eta^3 N} \big|\mathcal{A}_{\lambda\Delta}(f_1, \dots, f_k)(x)\big|^2\Big)^{1/2} \leq C_{d,\Delta}\Big(\min_{1\leq j\leq k} \|f_j\|_{U^1_{q_\eta,L}(Q)} + O(\eta)\Big).$$

This follows from Theorem 2 by symmetry and sublinearity after decomposing $f_k = f_{k,1} + f_{k,2} + f_{k,3}$ with

$$f_{k,1} = f_k * \chi_{q_\eta,L}$$

where $f_{k,2}$ and $f_{k,3}$ satisfy

$$\widehat{f_{k,2}} = \widehat{f}_k\, 1_{\Omega_{\eta,\eta^{-1}L}}(1 - \widehat{\chi_{q_\eta,L}}) \quad \text{and} \quad \widehat{f_{k,3}} = \widehat{f}_k\, 1_{\Omega^c_{\eta,\eta^{-1}L}}(1 - \widehat{\chi_{q_\eta,L}}).$$

Indeed, estimate (2) implies that

$$\Big(\frac{1}{|Q|}\sum_{x\in\mathbb{Z}^d} \sup_{\lambda\geq 1} \big|\mathcal{A}_{\lambda\Delta}(f_1, \dots, f_{k-1}, g)(x)\big|^2\Big)^{1/2} \leq C_{d,\Delta}\Big(\frac{1}{|Q|}\sum_{x\in\mathbb{Z}^d} |g(x)|^2\Big)^{1/2}$$

for any $g : \mathbb{Z}^d \to \mathbb{C}$. Note that if $g = f_{k,1}$ then

$$\Big(\frac{1}{|Q|}\sum_{x\in\mathbb{Z}^d} |f_{k,1}(x)|^2\Big)^{1/2} = \|f_k\|_{U^1_{q_\eta,L}(Q)}.$$

In light of the fact that

$$\widehat{\chi}_{q_\eta,L}(\xi)=\frac{q_\eta^d}{L^d}\sum_{x\in[-\frac{L}{2},\frac{L}{2})^d,\ q_\eta|x} e^{-2\pi i x\cdot\xi}$$

it is easy to see that $\widehat{\chi}_{q,L}(\ell/q)=1$ for all $\ell\in\mathbb{Z}^d$ and that there exists some absolute constant $C>0$ such that

$$0\le 1-\widehat{\chi}_{q_\eta,L}(\xi)\le C\,L\,|\xi-\ell/q_\eta| \tag{9}$$

for all $\xi\in\mathbb{T}^d$ and $\ell\in\mathbb{Z}^d$, and hence that $1-\widehat{\chi_{q_\eta,L}}(\xi)=O(\eta)$ for all $\xi\in\Omega_{\eta,\eta^{-1}L}$. It thus follows, by Plancherel, that

$$\left(\frac{1}{|Q|}\sum_{x\in\mathbb{Z}^d}|f_{k,2}(x)|^2\right)^{1/2}=O(\eta).$$

Finally, since $\operatorname{supp}\widehat{f_{k,3}}\subseteq\Omega^c_{\eta,\eta^{-1}L}$, it follows from estimate (3) that

$$\left(\frac{1}{|Q|}\sum_{x\in\mathbb{Z}^d}\sup_{\eta^{-3}L\le\lambda\le\eta^3N}\left|\mathcal{A}_\lambda(f_1,\dots,f_{k-1},f_{k,3})(x)\right|^2\right)^{1/2}\le C_{d,\Delta}\,\eta. \qquad \square$$

Proof of Proposition 1. Let $0<\varepsilon\le\delta\le1$ and $0<\eta\ll\varepsilon^2$.

Suppose there exists a set $A\subseteq\mathbb{Z}^d$ with $d\ge 2k+3$ with $\delta=\delta^*(A)>0$ that is η-uniformly distributed but for which the conclusion of Proposition 1 fails, namely that there exists arbitrarily large pairs (λ_0,λ_1) such that for every $x\in A$ one has

$$\mathcal{A}_{\lambda\Delta}(1_A,\dots,1_A)(x)\le\delta^k-\varepsilon$$

for some $\lambda\in[\lambda_0,\lambda_1]\cap\sqrt{\mathbb{N}}$.

Combining this with Lemma 1 we can conclude that there exists a positive integer L and a cube $Q\in\mathbb{Z}^d$ with sidelength N sufficiently large so that in addition to the properties (7) and (8) we also have the property that

$$\mathcal{A}_{\lambda\Delta}(1_{A\cap Q},\dots,1_{A\cap Q})(x)\le\delta^k-\varepsilon$$

for every $x\in A$ for some $\lambda\in[\eta^{-3}L,\eta^3N]\cap\sqrt{\mathbb{N}}$.

We now let $A':=A\cap Q'$, where Q' denotes the cube of sidelength $(1-\eta^3)N$ with the same center as Q. It then follows, provided that N was chosen sufficiently large, that

$$\mathcal{A}_{\lambda\Delta}(1_Q,\delta1_Q,\dots,\delta1_Q)(x)=\delta^k$$

for every $x\in A'$ and hence that for each such x one has

$$\sum_{j=0}^{k-1}\mathcal{A}_{\lambda\Delta}(\underbrace{1_{A\cap Q},\dots 1_{A\cap Q}}_{j\text{ copies}},(1_A-\delta)1_Q,\delta1_Q,\dots,\delta1_Q)(x)\le-\varepsilon$$

for some $\lambda\in[\eta^{-3}L,\eta^3N]\cap\sqrt{\mathbb{N}}$. Consequently, we have that

$$\sum_{j=0}^{k-1}\sup_{\eta^{-3}L\le\lambda\le\eta^3N}\Big|\mathcal{A}_\lambda(\underbrace{1_{A\cap Q},\dots 1_{A\cap Q}}_{j\text{ copies}},(1_A-\delta)1_Q,\delta1_Q,\dots,\delta1_Q)(x)\Big|\ge\varepsilon \tag{10}$$

for every $x\in A'$.

Since $\eta\ll\delta$ and $|A'|\ge|A\cap Q|-\eta^3|Q|$ it follows from (7) that $|A'|/|Q|\ge\delta/2$. Combining this observation with (10) we obtain

$$\sum_{j=0}^{k-1}\frac{1}{|Q|}\sum_{x\in\mathbb{Z}^d}\sup_{\eta^{-3}L\le\lambda\le\eta^3N}\Big|\mathcal{A}_\lambda(\underbrace{1_{A\cap Q},\dots 1_{A\cap Q}}_{j\text{ copies}},(1_A-\delta)1_Q,\delta1_Q,\dots,\delta1_Q)(x)\Big|\ge\varepsilon\delta/2. \tag{11}$$

However, Lemma 2 and (8) clearly imply that for each $0\le j\le k-1$ one has

$$\frac{1}{|Q|}\sum_{x\in\mathbb{Z}^d}\sup_{\eta^{-3}L\le\lambda\le\eta^3N}\Big|\mathcal{A}_\lambda(\underbrace{1_{A\cap Q},\dots 1_{A\cap Q}}_{j\text{ copies}},(1_A-\delta)1_Q,\delta1_Q,\dots,\delta1_Q)(x)\Big|=O(\eta)$$

which leads to a contradiction if η is chosen sufficiently small with respect to ε^2. $\square$

3. Proof of Theorem 2

Following the approach in [12] we will deduce Theorem 2 from refined estimates for our maximal operators at a single dyadic scale, namely Proposition 2 below. We first need to introduce some notation closely related to that in Section 1.2. For any integer $j \geq 0$ we let

$$q_j = \operatorname{lcm}\{1, 2, \ldots, 2^j\}$$

noting that $q_j \asymp e^{2^j}$, and for any non-negative integers j and l that satisfy $2^j \leq l$, we let

$$\Omega_{j,l} := \{\xi \in \mathbb{T}^d : \xi \in [-2^{j-l}, 2^{j-l}]^d + (q_j^{-1}\mathbb{Z})^d\}.$$

Proposition 2. *If $k \geq 1$, $d \geq 2k+3$, and $\Delta = \{0, v_1, \ldots, v_k\} \subseteq \mathbb{Z}^d$ be a non-degenerate k-simplex, then*

$$\Big\| \sup_{2^l \leq \lambda \leq 2^{l+1}} |\mathcal{A}_{\lambda\Delta}(f_1, \ldots, f_k)| \Big\|_2 \leq C_{d,\Delta}\, 2^{-j/2} j^{-1} \, \|f_1\|_2 \cdots \|f_k\|_2 \tag{12}$$

whenever $\operatorname{supp} \widehat{f}_i \subseteq \Omega_{j,l}^c$ *for some* $1 \leq i \leq k$, *where* $\Omega_{j,l}^c$ *denotes the complement of* $\Omega_{j,l}$.

It is easy to see that Proposition 2 is equivalent to estimate (3) of Theorem 2. Indeed, note that in proving (3) one may restrict the sup to $\eta^{-2}L \leq \lambda \leq 2\eta^{-2}L$. Choosing $l, j \in \mathbb{N}$ such that $2^l \leq \eta^{-2}L \leq 2^{l+1}$ and $2^j \geq \eta^{-2}$ we have that $2^{l-j} \leq L$ and hence $\Omega_{j,l} \subseteq \Omega_{\eta,L}$. Applying Proposition 2 with j and l chosen as above implies estimate (3) of Theorem 2, while applying estimate (3) of Theorem 2 with $L = 2^{l-j}$ and $\eta = 2^{-j/2}$ immediately implies Proposition 2.

We are left with establishing that Proposition 2 implies estimate (2) of Theorem 2. Following the approach in [12] we start by introducing a smooth sampling function supported on $\Omega_{j,l}$.

3.1. A smooth sampling function supported on $\Omega_{j,l}$.

Let $\psi \in \mathcal{S}(\mathbb{R}^d)$ be a Schwartz function satisfying

$$1_Q(\xi) \leq \widetilde{\psi}(\xi) \leq 1_{2Q}(\xi)$$

where $Q = [-1/2, 1/2]^d$ and

$$\widetilde{\psi}(\xi) := \int_{\mathbb{R}^d} \psi(x) e^{-2\pi i x\cdot\xi} dx$$

denote the Fourier transform of ψ on $\mathbb{R}^d$. For a given $q \in \mathbb{N}$ and $L > q$ we define $\psi_{q,L} : \mathbb{Z}^d \to \mathbb{R}$ as

$$\psi_{q,L}(x) = \begin{cases} \left(\frac{q}{L}\right)^d \psi\left(\frac{m}{L}\right) & \text{if } x \in (q\mathbb{Z})^d \\ 0 & \text{otherwise} \end{cases}$$

Writing $x = qr + s$ with $r \in \mathbb{Z}^d$ and $s \in \mathbb{Z}^d/q\mathbb{Z}^d$, it follows from Poisson summation that

$$\widehat{\psi}_{q,L}(\xi) = \sum_{x\in\mathbb{Z}^d} \psi(x) e^{-2\pi i x\cdot\xi}$$

is a q^{-1}-periodic function on $\mathbb{T}^d$ that satisfies

$$\widehat{\psi}_{q,L}(\xi) = \sum_{\ell\in\mathbb{Z}^d} \widetilde{\psi}(L(\xi - \ell/q)).$$

For a given $l \in \mathbb{N}$ and $0 \leq j \leq J_l := [\log_2(l)] - 2$, we now define the sampling function

$$\Psi_{l,j} = \psi_{q_j, 2^{l-j}} \tag{13}$$

and note that $\operatorname{supp} \widehat{\Psi}_{l,j} \subseteq \Omega_{j,l}$.

Finally we define $\Delta\Psi_{l,j} = \Psi_{l,j+1} - \Psi_{l,j}$ and note the important almost orthogonality property they enjoy.

Lemma 3 (Lemma 1 in [12])**.** *There exists a constant $C = C_\Psi > 0$ such that*

$$\sum_{l \geq 2^j} |\widehat{\Delta\Psi}_{l,j}(\xi)|^2 \leq C$$

uniformly in $j \in \mathbb{N}$ and $\xi \in \mathbb{T}^d$.

3.2. **Proof that Proposition 2 implies estimate (2) of Theorem 2.** Let $k \geq 1$, $d \geq 2k+3$, and $\Delta = \{0, v_1, \ldots, v_k\} \subseteq \mathbb{Z}^d$ be a non-degenerate k-simplex. In [12] the authors gave a direct proof of estimate (2) of Theorem 2 when $k=1$, the ℓ^2-boundedness of the discrete spherical maximal function. We may thus, without loss in generality assume that $k \geq 2$, $\operatorname{supp} \widehat{f_k} \subseteq \Omega_{j,l}^c$, and that

$$\Big\| \sup_{\lambda \geq 1} |\mathcal{A}_{\lambda\widetilde{\Delta}}(f_1, \ldots, f_{k-1})| \Big\|_2 \leq C_{d,\widetilde{\Delta}} \|f_1\|_2 \cdots \|f_{k-1}\|_2 \tag{14}$$

where $\widetilde{\Delta} = \{0, v_1, \ldots, v_{k-1}\} \subseteq \mathbb{Z}^d$.

Let

$$\mathcal{M}_l(f_1, \ldots, f_k) := \sup_{2^l \leq \lambda \leq 2^{l+1}} |\mathcal{A}_{\lambda\Delta}(f_1, \ldots, f_k)|. \tag{15}$$

Writing

$$f_k = f_k * \Psi_{l,0} + \sum_{j=0}^{J_l - 1} f_k * \Delta\Psi_{l,j} + (f_k - f_k * \Psi_{l,J_l})$$

it follows by subadditivity that

$$\mathcal{M}_l(f_1, \ldots, f_k) \leq \mathcal{M}_l(f_1, \ldots, f_k * \Psi_{l,0}) + \sum_{j=0}^{J_l-1} \mathcal{M}_l(f_1, \ldots, f_k * \Delta\Psi_{l,j}) + \mathcal{M}_l(f_1, \ldots, f_k - f_k * \Psi_{l,J_l}). \tag{16}$$

Estimate (2) of Theorem 2 will now follow from a few observations and applications of Proposition 2, in light of the fact that

$$\sup_{\lambda \geq 1} |\mathcal{A}_{\lambda\Delta}(f_1, \ldots, f_k)| = \sup_l \mathcal{M}_l(f_1, \ldots, f_k).$$

We first observe that the first term on the right in (16) above satisfies

$$\mathcal{M}_l(f_1, \ldots, f_k * \Psi_{l,0}) \leq C_\Psi \mathcal{H}(f_k) \ \sup_{\lambda \geq 1} |\mathcal{A}_{\lambda\widetilde{\Delta}}(f_1, \ldots, f_{k-1})|$$

uniformly in l, where

$$\mathcal{H}(f)(x) = \sup_{N > 0} \frac{1}{|Q_N|} \Big| \sum_{y \in Q_N} f(x-y) \Big|$$

with $Q(N)$ the discrete cube $[-N/2, N/2]^d \cap \mathbb{Z}^d$ denotes the discrete Hardy-Littlewood maximal operator, which trivially satisfies $\|\mathcal{H}f\|_\infty \leq \|f\|_\infty \leq \|f\|_2$ by the nesting of discrete ℓ^p spaces. It therefore follows from the inductive hypothesis (14) that

$$\sup_l \|\mathcal{M}_l(f_1, \ldots, f_k * \Psi_{l,0})\|_2 \leq C\|f_1\|_2 \cdots \|f_k\|_2.$$

For the middle terms in (16) we first note that

$$\sup_l \sum_{j=0}^{J_l-1} \mathcal{M}_l(f_1, \ldots, f_k * \Delta\Psi_{l,j}) \ll \Big(\sum_{l=0}^{\infty} \Big| \sum_{j=0}^{J_l-1} \mathcal{M}_l(f_1, \ldots, f_k * \Delta\Psi_{l,j}) \Big|^2 \Big)^{1/2}$$

Taking ℓ^2 norms of both sides of the inequality above and applying Minkowski's inequality, followed by an application of Proposition 2, gives

$$\begin{aligned}
\Big\| \sup_l \sum_{0 \leq j \leq J_l} \mathcal{M}_l(f_1, \ldots, f_k * \Delta\Psi_{l,j}) \Big\|_2 &\leq \sum_j \Big(\sum_{l \geq 2^j} \|\mathcal{M}_l(f_1, \ldots, f_k * \Delta\Psi_{l,j})\|_2^2 \Big)^{1/2} \\
&\leq C\|f_1\|_2 \cdots \|f_{k-1}\|_2 \ \sum_j 2^{-j/2} \Big(\sum_{l \geq 2^j} \|f_k * \Delta\Psi_{l,j}\|_2^2 \Big)^{1/2} \\
&\leq C\|f_1\|_2 \cdots \|f_k\|_2
\end{aligned}$$

where the last inequality above follows from Lemma 3.

One more application of Proposition 2 with $j=[\log_2 l]-2$ to the last term in (16) gives

$$\begin{aligned}\Big\|\sup_l \mathcal{M}_l(f_1,\ldots,f_k-f_k*\Psi_{l,J_l})\Big\|_2 &\le \Big(\sum_{l=1}^{\infty}\|\mathcal{M}_l(f_1,\ldots,f_k-f_k*\Psi_{l,J_l})\|_2^2\Big)^{1/2}\\ &\le C\Big(\sum_{l=1}^{\infty} l^{-1}(\log_2 l)^{-2}\Big)^{1/2}\|f_1\|_2\cdots\|f_k\|_2\\ &\le C\|f_1\|_2\cdots\|f_k\|_2. \qquad\square\end{aligned}$$

4. Proof of Proposition 2

Given any simplex $\Delta=\{v_0=0,v_1,\ldots,v_k\}\subseteq\mathbb{R}^d$, we introduce the associated *inner product matrix* $T=T_\Delta=(t_{ij})_{1\le i,j\le k}$ with entries $t_{ij}:=v_i\cdot v_j$, where "$\cdot$" stands for the dot product in $\mathbb{R}^d$. Note that T is a positive semi-definite matrix with integer entries and T is positive definite if and only if Δ is non-degenerate. It is easy to see that $\Delta'\simeq\lambda\Delta$, with $\Delta'=\{y_0=0,y_1,\ldots,y_k\}$, if and only if

$$y_i\cdot y_j=\lambda^2 t_{ij}\quad\text{for all}\quad 1\le i,j\le k. \tag{17}$$

Let $M\in\mathbb{Z}^{d\times k}$ be a matrix with column vectors $y_1,\ldots,y_k\in\mathbb{Z}^d$. Then the system of equations given in (17) can be written as the matrix equation

$$M^tM=\lambda^2T, \tag{18}$$

where M^t is the transpose of the matrix M. We denote by $S_{\lambda^2T}(M)$ the indicator function of relation (18). Let $I_k=[0,2]^{k(k+1)/2}$ denote the space of symmetric $k\times k$ matrices with entries in the interval $[0,2]$. Using the fact that

$$\mathrm{tr}(X^tY)=\mathrm{tr}(YX^t)=\sum_{i=1}^{k}\sum_{j=1}^{k}x_{ij}y_{ij},$$

for any $k\times k$ matrices $X=(x_{ij})$, $Y=(y_{ij})$, one has

$$S_{\lambda^2T}(M)=2^{-k}\int_{I_k}e^{\pi i\,\mathrm{tr}[(M^tM-\lambda^2T)X]}\,dX \tag{19}$$

where $dX=\prod_{1\le i\le j\le k}dx_{ij}$. Moreover, if $M^tM=\lambda^2T$ then

$$\mathrm{tr}(T^{-1}M^tM)=\mathrm{tr}(MT^{-1}M^t)=\mathrm{tr}(\lambda^2I)=k\lambda^2.$$

Given $l\in\mathbb{N}$ write $\Lambda=2^l$ and $\varepsilon=2^{-2l}$. We have

$$S_{\lambda^2T}(M)=2^{-k}e^{k\varepsilon\lambda^2}\int_{I_k}e^{-\pi i\lambda^2\,\mathrm{tr}(TX)}\,e^{\pi i\,\mathrm{tr}(M(X+i\varepsilon T^{-1})M^t)}dX. \tag{20}$$

Let

$$G_{X,\varepsilon}(M)=G_{X,\varepsilon}(y_1,\ldots,y_k)=e^{\pi i\,\mathrm{tr}(M(X+i\varepsilon T^{-1})M^t)}$$

be the Gaussian function, where $y_1,\ldots,y_k\in\mathbb{Z}^d$ are the column vectors of the matrix M, and define the corresponding multi-linear operator

$$B_{X,\varepsilon}(f_1,\ldots,f_k)(x):=\sum_{y_1,\ldots,y_k\in\mathbb{Z}^d}f_1(x+y_1)\ldots f_k(x+y_k)\,G_{X,\varepsilon}(y_1,\ldots,y_k). \tag{21}$$

It follows that

$$A_\lambda(f_1,\ldots,f_k)(x)=2^{-k}e^{k\varepsilon\lambda^2}|S_{\lambda\Delta}|^{-1}\int_{I_k}e^{-\pi i\lambda^2\,\mathrm{tr}(TX)}\,B_{X,\varepsilon}(f_1,\ldots,f_k)(x)\,dX.$$

Thus for the maximal function

$$\mathcal{M}_l(f_1,\ldots,f_k):=\sup_{2^l\le\lambda\le2^{l+1}}|\mathcal{A}_{\lambda\Delta}(f_1,\ldots,f_k)|$$

we have the pointwise estimate

$$M_l(f_1,\ldots,f_k)(x)\le C\,\Lambda^{-k(d-k-1)}\int_{I_k}|B_{X,\varepsilon}(f_1,\ldots,f_k)(x)|\,dX, \tag{22}$$

as $\varepsilon = \Lambda^{-2} = 2^{-2l}$ and $\Lambda \le \lambda \le 2\Lambda$. Finally, by Minkowski's inequality

$$\|M_l(f_1,\ldots,f_k)\|_2 \le C\,\Lambda^{-k(d-k-1)} \int_{I_k} \|B_{X,\varepsilon}(f_1,\ldots,f_k)(x)\|_2\,dX. \tag{23}$$

Taking the Fourier transform of the expression in (21),

$$\widehat{B_{X,\varepsilon}}(f_1,\ldots,f_k)(\xi) = \tag{24}$$
$$= \int_{\Pi^{k-1}} \widehat{f}_1(\xi_1)\ldots\widehat{f}_{k-1}(\xi_{k-1})\widehat{f}_k(\xi-\xi_1-\ldots-\xi_{k-1})\,\widehat{G_{X,\varepsilon}}(\xi_1,\ldots,\xi_{k-1},\xi-\xi_1-\ldots-\xi_{k-1})\,d\xi_1\ldots d\xi_{k-1}.$$

Thus by the Cauchy-Schwarz inequality and Plancherel's indentity, one has

$$\|B_{X,\varepsilon}(f_1,\ldots,f_k)\|_2^2 \le \|\widehat{G_{X,\varepsilon}}\|_\infty^2 \prod_{i=1}^k \|f_i\|_2^2. \tag{25}$$

Thus, the $\ell^2\times\cdots\times\ell^2 \to \ell^2$ boundedness of the dyadic maximal operator $M_l(f_1,\ldots,f_k)$ follows from the estimate

$$\int_{I_k} \|\widehat{G_{X,\varepsilon}}\|_\infty\,dX \le C\,\Lambda^{k(d-k-1)} \tag{26}$$

with $\Lambda = 2^l$, for some constant $C_{d,k}$ which may depend on the k-simplex Δ.

For the mollified estimate assume that, supp $\widehat{f_i} \subseteq \Omega_{j,l}^c$ i.e. $\widehat{f_i} = \mathbf{1}_{\Omega_{j,l}^c}\widehat{f_i}$ for some $1 \le i \le k$, using the notation $\mathbf{1}_A$ for the indicator function of a set A. By symmetry of the expression in (21) we may assume without loss of generality that $i=1$. In this case in equality (24) the function $\widehat{G_{X,\varepsilon}}(\xi_1,\ldots,\xi_{k-1},\xi-\xi_1-\ldots-\xi_{k-1})$ can be replaced by $\mathbf{1}_{\Omega_{j,l}^c}(\xi_1)\,\widehat{G_{X,\varepsilon}}(\xi_1,\ldots,\xi_{k-1},\xi-\xi_1-\ldots-\xi_{k-1})$, thus to prove Theorem 2, it is enough to show that for $j,l\in\mathbb{N}$ with $2^{j+2}\le l$, one has

$$\int_{I_k} \|\mathbf{1}_{\Omega_{j,l}^c}(\xi_1)\,\widehat{G_{X,\varepsilon}}(\xi_1,\ldots,\xi_k)\|_\infty\,dX \;\lesssim\; 2^{-j/2}j^{-1}\,\Lambda^{k(d-k-1)}, \tag{27}$$

with $\Lambda = 2^l$.

5. Estimates for theta functions on the Siegel upper half space.

To prove estimates (26)-(27) we will follow the approach given in [**?**], Sec. 5. For the sake of completeness we recall below some of the basic notions and constructs. If $M=[m_1,\ldots,m_k]\in\mathbb{Z}^{d\times k}$ and $\mathcal{X}=[\xi_1,\ldots,\xi_k]\in\mathbb{R}^{d\times k}$ are $d\times k$ matrices then one has that $tr(M^t\mathcal{X}) = m_1\cdot\xi+\ldots+m_k\cdot\xi_k$ where $\cdot$ denotes the usual dot product. Thus the Fourier transform of a function $f(m_1,\ldots,m_k)=f(M)$ may written as

$$\widehat{f}(\mathcal{X}) = \widehat{f}(\xi_1,\ldots,\xi_k) = \sum_{M\in\mathbb{Z}^{d\times k}} f(M)e^{-2\pi i tr(M^t\mathcal{X})}.$$

This implies that

$$\widehat{G}_{X,\varepsilon}(\mathcal{X}) = \sum_{M\in\mathbb{Z}^{d\times k}} e^{\pi i\,tr[(M(X+i\varepsilon T^{-1})M^t - 2M^t\mathcal{X}]} = \theta_{d,k}(X+i\varepsilon T^{-1},-\mathcal{X},0), \tag{28}$$

is the theta-function $\theta_{d,k}:\mathbb{H}_k\times\mathbb{R}^{d\times k}\times\mathbb{R}^{d\times k}\to\mathbb{C}$ defined by

$$\theta_{d,k}(Z,\mathcal{X},\mathcal{E}) = \sum_{M\in\mathbb{Z}^{d\times k}} e^{\pi i\,tr[(M-\mathcal{E})Z(M-\mathcal{E})^t+2M^t\mathcal{X}-\mathcal{E}^t\mathcal{X}]}, \tag{29}$$

for $Z=X+iY\in\mathbb{H}_k$, $\mathbb{H}_k$ being the Siegel upper space, see (5.1)-(5.3) in [**?**].

We partition the range of integration I_k and estimating the theta function separately on each part by exploiting its transformation properties. This may be viewed as the extension of the classical Farey arcs decomposition to $k>1$. Recall the integral symplectic group

$$\Gamma_k = \left\{\gamma = \begin{pmatrix} A & B \\ C & D \end{pmatrix};\ AB^t = BA^t,\ CD^t = DC^t,\ AD^t - BC^t = E_k,\right\} \tag{30}$$

which acts on the Siegel upper-half space $\mathcal{H}_k = \{Z = X + iY : \ X \in \mathcal{M}_k, Y \in \mathcal{P}_k\}$ as a group of analytic automorphisms; The action being defined by: $\gamma\langle Z\rangle = (AZ+B)(CZ+D)^{-1}$ for $\gamma \in \Gamma_k,\ Z \in \mathcal{H}_k$, see [?] and also[9]. Let us recall also the subgroup of integral modular substitutions:

$$\Gamma_{k,\infty} = \left\{\gamma = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix};\ AB^t = BA^t,\ AD^t = E_k\right\} \tag{31}$$

Writing $U = A^t$ and $S = AB^t$, it is easy to see that $D = U^{-1}$ and $B = SU^{-1}$, moreover S is symmetric and $U \in GL(k,\mathbb{Z})$, i.e. $\det(U) = \pm 1$. The action of such $\gamma \in \Gamma_{k,\infty}$ on $Z \in \mathcal{H}_k$ takes the form:

$$\gamma\langle Z\rangle = Z[U] + S, \tag{32}$$

using the notation $Z[U] = U^t ZU$. The general linear group $GL(k,\mathbb{Z})$ acts on the space $\mathcal{P}_k$ of positive $k \times k$ matrices, via the action: $Y \to Y[U],\ Y \in \mathcal{P}_k$, and let $\mathcal{R}_k$ denote the corresponding so-called Minkowksi domain, see [KL, Definition 1]. A matrix $Y = (y_{ij}) \in \mathcal{R}_k$ is called reduced. We recall that for a reduced matrix Y

$$Y \approx Y_D\ , \qquad y_{11} \le y_{22} \le \ldots \le y_{kk}, \tag{33}$$

where $Y_D = diag(y_{11},\ldots,y_{kk})$ denotes the diagonal part of Y, and $A \approx B$ means that $A - c_k B > 0$, $B - c_k A > 0$ for some constant $c_k > 0$. For a proof of these facts, see [KL, Lemma 2]. A fundamental domain $\mathcal{D}_k$ for the action of Γ_k on $\mathcal{H}_k$, called the Siegel domain, consists of all matrices $Z = X + iY$, $(X = (x_{ij}))$, satisfying

$$Y \in \mathcal{R}_k, \quad |x_{ij}| \le 1/2, \qquad |\det{(CZ+D)}| \ge 1, \quad \forall\ \gamma = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \Gamma_k. \tag{34}$$

The second rows of the matrices $\gamma \in \Gamma_k$ are parameterized by the so-called coprime symmetric pairs of integral matrices (C,D), which means that CD^t is symmetric and the matrices GC and GD with a matrix G of order k are both integral only if G is integral, see [A, Lemma 2.1.17]. It is clear from definition (5.6) that if $\gamma_2 = \gamma\gamma_1$ with second rows (C_2,D_2) and (C_1,D_1) for some $\gamma \in \Gamma_{k,\infty}$, then $(C_2,D_2) = (UC_1,UD_1)$ for some $U \in GL(k,\mathbb{Z})$. On the other hand, if both γ_1 and γ_2 have the same second row (C,D) then $\gamma_2\gamma_1^{-1} \in \Gamma_{k,\infty}$. This gives the parametrization of the group $\Gamma_{k,\infty}\backslash\Gamma_k$ by equivalence classes of coprime symmetric pairs (C,D) via the equivalence relation $(C_2,D_2) \sim (C_1,D_1)$ if $(C_2,D_2) = (UC_1,UD_1)$ for some $U \in GL(k,\mathbb{Z})$, see also [A, p.54]. We will use the notation $[\gamma] = [C,D] \in \Gamma_{k,\infty}\backslash\Gamma_k$.

If one defines the domain: $\mathbb{F}_k = \cup_{\gamma\in\Gamma_{k,\infty}}\gamma\mathcal{D}_k$, then $\mathcal{H}_k = \bigcup_{[\gamma]\in\Gamma_{k,\infty}\backslash\Gamma_k} \gamma^{-1}\mathbb{F}_k$ is a non-overlapping cover of the Siegel upper half-plane. Correspondingly, for a given matrix $T > 0$ of order k, define the Farey arc dissection of level T, as the cover

$$I_k = \bigcup_{[\gamma]\in\Gamma_{k,\infty}\backslash\Gamma_k} I_T[\gamma], \qquad I_T[\gamma] = \{X \in I_k :\ X + iT^{-1} \in \gamma^{-1}\mathbb{F}_k\} \tag{35}$$

We recall the basic estimates (5.14)-(5.16) in [?] whose proofs are based on the transformation property

$$|\theta_{d,k}(Z,\mathcal{X},0)| = |\det{(CZ+D)}|^{-\frac{d}{2}}\ |\theta_{d,k}(\gamma\langle Z\rangle,\ \mathcal{X}A^t - K_\gamma/2,\ \mathcal{X}C^t - N_\gamma/2)|$$

for some matrices $K_\gamma, N_\gamma \in \mathbb{Z}^{n\times k}$, see Proposition 5.2 in [?].
Namely, if (C,D) is a coprime symmetric pair, then for $Z \in I_T[C,D]$ one has

$$|\theta_{d,k}(Z,\mathcal{X},0)| \lesssim |\det{(CZ+D)}|^{-\frac{d}{2}}, \tag{36}$$

uniformly for $\mathcal{X} \in \mathcal{M}_k(\mathbb{R})$.

Next we describe the "mollified" estimate (5.16) in [?] in slightly different form. For $q \in \mathbb{N}$ and $\tau > 0$ define the region

$$\Omega_{q,\tau} = \{\mathcal{X} \in \mathbb{R}^{d\times k} :\ |\mathcal{X} - P/2q| \le \tau \ \ for\ some\ \ P \in \mathbb{Z}^{d\times k}\}. \tag{37}$$

If $[\gamma] = [C,D]$ coprime symmetric pair, $q := |det(C)| > 0$, then for $Z \in I_T[C,D]$

$$|\theta_{d,k}(Z,\mathcal{X},0)| \lesssim |\det{(CZ+D)}|^{-\frac{d}{2}}\ \left(e^{-c\min(Y)} + e^{-c\,\tau^2\mu(C^tYC)}\right), \tag{38}$$

uniformly for $\mathcal{X} \in \Omega^c_{q,\tau}$. Here $Y = Im\gamma\langle Z\rangle$, $\min(Y) = \min_{x\in\mathbb{Z}^d, x\neq 0} |Yx\cdot x|$ and $\mu(Y) = \min_{x\in\mathbb{R}^d,\, |x|=1} |Yx\cdot x|$.

Define, similarly as in (5.20) in [?]

$$J_T[C,D] = \int_{I_T[C,D]} \sup_{\mathcal{X}} |\theta_{d,k}(X+iT^{-1}, -\mathcal{X}, 0)|\, dX. \tag{39}$$

By (36) we have that

$$J_T[C,D] \lesssim J^0_T[C,D], \tag{40}$$

where

$$J^0_T[C,D] = \int_{X\in I_T[C,D]} |\det(CZ+D)|^{-\frac{d}{2}}\, dX. \tag{41}$$

If $q := |\det(C)| > 0$, then for $\tau > 0$ let,

$$J_{T,\tau}[C,D] := \int_{I_T[C,D]} \sup_{\mathcal{X}} \mathbf{1}_{\Omega^c_{\tau,q}}(\mathcal{X})\, |\theta_{d,k}(X+iT^{-1}, -\mathcal{X}, 0)|\, dX. \tag{42}$$

By estimate (38) one has

$$J_{T,\tau}[C,D] \;\lesssim\; J^1_T[C,D] + J^2_{T,\tau}[C,D], \tag{43}$$

where

$$J^1_T[C,D] = \int_{I_T[C,D]} |\det(CZ+D)|^{-\frac{d}{2}}\, e^{-c\, min(Y)}\, dX \tag{44}$$

$$J^2_{T,\tau}[C,D] = \int_{I_T[C,D]} |\det(CZ+D)|^{-\frac{d}{2}}\, e^{-c\tau^2\,\mu(C^tYC)}\, dX. \tag{45}$$

where $Y = Im\,\gamma\langle Z\rangle$ and $\gamma \in \Gamma_k$ such that $[\gamma] = [C,D] \in \Gamma_{k,\infty}\backslash\Gamma$.

Then by inequalities (5.24)-(5.26) given in Propositions 5.3-5.4 in [?], we have the estimates

$$\sum_{S^t=S} J_T[C,D+CS] \lesssim det(T)^{\frac{d-k-1}{2}} |det(C)|^{-\frac{d}{2}}, \tag{46}$$

$$\sum_{S^t=S} J_{T,\tau}[C,D+CS] \lesssim det(T)^{\frac{d-k-1}{2}} \left(|det(C)|^{-k} min(T)^{-\frac{d-2k}{4}} + |det(C)|^{-\frac{d}{2}} (\tau^2\mu(T))^{-\frac{d-2k}{4}}\right), \tag{47}$$

where the summation is over all symmetric integral matrices $S \in \mathcal{M}_k(\mathbb{Z})$.

Recall that the map $[C,D] \to C^{-1}D$ provides a one-one and onto correspondence between the classes of coprime symmetric pairs $[C,D] \in \Gamma_{k,\infty}\backslash\Gamma_k$, with $det(C) \neq 0$, and symmetric rational matrices R of order k, and the pairs $[C,D+CS]$ correspond to the matrices $R+S$ with symmetric $S \in \mathbb{Z}^{k\times k}$. Let us write $\mathbb{Q}(1)^{k\times k}$ for the space of modulo 1 incongruent symmetric rational matrices, where $\mathbb{Q}(1) = \mathbb{Q}/\mathbb{Z}$, $\mathbb{Q}$ being the set of rational numbers. If $R = C^{-1}D$, for a coprime symmetric pair $[C,D]$ then will write

$$J_T[R] := \sum_{S^t=S} J_T[C,D+CS], \tag{48}$$

$$J_{T,\tau}[R] := \sum_{S^t=S} J_{T,\tau}[C,D+CS], \tag{49}$$

which is well-defined as it only depends on the equivalence class $[R] \in \mathbb{Q}(1)^{k\times k}$. Finally write $d(R) = |\det(C)|$ for $R = C^{-1}D$. Then by (29)-4.3, we have with $\varepsilon = \Lambda^{-2}$,

$$\int_{I_k} \sup_{\mathcal{X}} |\theta_{d,k}(X+i\varepsilon T^{-1}, -\mathcal{X}, 0)|\, dX = \sum_{[C,D], det(C)\neq 0} J_{\Lambda^2 T}[C,D] + \sum_{[C,D], det(C)=0} J_{\Lambda^2 T}[C,D] =: \sum_1 + \sum_2. \tag{50}$$

An estimate for the second sum is given in Corollary 5.1 in [?], namely it is shown that

$$\sum_2 \;\lesssim\; |\Lambda^2 T|^{(k-1)(d-k)/2} \;\lesssim\; \Lambda^{(d-k)(k-1)}, \tag{51}$$

where $|T| = (\sum_{ij} t_{ij}^2)^{1/2}$ is the Euclidean norm of the matrix T. For the first sum we use estimate (46) for the matrix $\Lambda^2 T$, which implies

$$\sum_1 = \sum_{[R]\in\mathbb{Q}(1)^{k\times k}} J_{\Lambda^2 T}[R] \;\lesssim\; \Lambda^{k(d-k-1)} \sum_{[R]\in\mathbb{Q}(1)^{k\times k}} d(R)^{-d/2}. \tag{52}$$

Recall the following estimate, proved in Lemma 1.4.9 in [K]; for $u \geq 1$ and $s > 1$ one has

$$u^{-s} \sum_{1\leq d(R)\leq u} d(R)^{-k} + \sum_{d(R)\geq u} d(R)^{-k-s} \lesssim (2 + \frac{1}{s-1})\, u^{1-s}, \tag{53}$$

where the summation is taken over $[R] \in \mathbb{Q}(1)^{k\times k}$. In particular $\sum_R d(R)^{-d/2} \lesssim 1$ in dimensions $d > 2k+2$, thus estimate (26) follows from (28), (50) and estimates (51)-(52).

For the mollified estimate (27), we set $\tau = 2^{j-l}$ besides $\Lambda = 2^l$ and $\varepsilon = 2^{-2l}$. Again, we note that if $q = |det(C)| > 0$ and if $q \mid q_j$ i.e. if q divides q_j then $\xi_1 \in \Omega_{j,l}^c$ implies that $\mathcal{X} \in \Omega_{\tau,q}$ for $\mathcal{X} = (\xi_1, \ldots, \xi_d)$, for the sets $\Omega_{j,l}$ and $\Omega_{\tau,q}$ defined in (**??**) and (**??**). Using this observation, we have

$$\int_{I_k} \sup_{\mathcal{X}} \mathbf{1}_{\Omega_{j,k}^c}(\xi_1) |\theta_{d,k}(X + i\varepsilon T^{-1}, -\mathcal{X}, 0)|\, dX \;\lesssim\; \sum_{d(R)|q_j} J_{\Lambda^2 T,\tau}[R] + \sum_{d(R)\nmid q_j} J_{\Lambda^2 T}[R] + \sum_2 . \tag{54}$$

In dimensions $d \geq 2k+3$, using (47) and (53), the first sum on the right side of (54) is crudely estimated by

$$\begin{aligned} \sum_{d(R)|q_j} J_{\Lambda^2 T,\tau}[R] \;&\lesssim\; \Lambda^{k(d-k-1)} \sum_{1\leq d(R)\leq q_j} \left(d(R)^{-k} \Lambda^{-\frac{d-2k}{2}} + d(R)^{-\frac{d}{2}} (\tau\Lambda)^{-\frac{d-2k}{2}} \right) \\ &\lesssim\; \Lambda^{k(d-k-1)} \left(q_j 2^{-\frac{3l}{2}} + 2^{-\frac{3j}{2}} \right) \;\lesssim\; \Lambda^{k(d-k-1)} 2^{-\frac{3j}{2}}. \end{aligned} \tag{55}$$

Indeed, $q_j = l.c.m.\{1 \leq q \leq 2^j\} \approx e^{2^j} \leq 2^l$ as $2^{j+2} \leq l$ by our assumptions. To estimate the second term on the right side of (54), we need the following.

Lemma 4. *Let $j \in \mathbb{N}$ and $s > 1$. Then*

$$\sum_{d(R)\nmid q_j} d(R)^{-k-s} \;\lesssim\; 2^{j(1-s)}\, j^{-1}, \tag{56}$$

where the implicit constant may depend on d, k and s.

Proof. Let

$$\Psi(s) := \sum_{[R]\in\mathbb{Q}(1)^{k\times k}} d(R)^{-k-s} = \sum_{n\geq 1} a_k(n) n^{-s}, \tag{57}$$

with $a_k(n) = \sum_{d(R)=n} d(R)^{-k}$. For two Dirichlet series $\Psi(s) = \sum_{n\geq 1} a(n) n^{-s}$ and $\Phi(s) = \sum_{n\geq 1} b(n) n^{-s}$ we will write $\Psi(s) \preceq \Phi(s)$ if $|a(n)| \leq b(n)$ for all $n \geq 1$.

It is proved in [9], see Lemma 1.4.9 there, that

$$\Psi(s) \preceq \zeta(s+1)^K \zeta(s) =: \sum_{n\geq 1} b_K(n) n^{-s}, \tag{58}$$

with $K = 2^k + k - 3$. Clearly the coefficients of the Dirichlet series $\zeta(s+1)^K\zeta(s)$ are multiplicative i.e. $b_K(nm) = b_K(n) b_K(m)$ if $(n,m) = 1$, moreover are easy to show that,

$$b_K(n) = \sum_{m|n} \frac{d_K(m)}{m}, \tag{59}$$

where $d_K(m) = |\{m_1, \ldots, m_k \in \mathbb{N} : m_1 m_2 \cdots m_K = m\}|$. Since $q_j = l.c.m.\{1 \leq q \leq 2^j\}$, if $n \nmid q_j$ the either there is a prime $p > 2^j$ such that $p \mid n$ or there is a prime $p < 2^j$ such that $p^{\gamma_p} > 2^j$ but $p^{\gamma_p} \mid n$. Accordingly, we have the estimate

$$\sum_{d(R)\nmid q_j} d(R)^{-k-s} = \sum_{n\nmid q_j} a_k(n) n^{-s} \leq \sum_{p>2^j} \sum_{n\geq 1} b_K(pn) p^{-s} n^{-s} + \sum_{p<2^j} \sum_{n\geq 1} b_K(p^{\gamma_p} n) p^{-\gamma_p s} n^{-s}. \tag{60}$$

Writing $n = p^r m$, the first sum on the right side of (60) is estimated by

$$\sum_{p>2^j}\sum_{n\geq 1} b_K(pn)p^{-s}n^{-s} = \sum_{p>2^j}\sum_{r=1}^{\infty}\sum_{m\geq 1, p\nmid m} b_K(p^r)b_K(m)p^{-rs}m^{-s}. \tag{61}$$

Using the fact that $b_K(p^r m) = b_K(p^r)b_k(m)$ and by (59)

$$b_K(p^r) = 1 + \sum_{s=1}^{r}\frac{d_K(p^s)}{p^s} \leq 1 + \sum_{s=1}^{\infty}\frac{(s+1)^K}{2^s} \lesssim 1, \tag{62}$$

uniformly in $r \geq 1$. Thus, for $s > 1$,

$$\sum_{p>2^j}\sum_{r=1}^{\infty}\sum_{m\geq 1, p\nmid m} b_K(p^r)b_K(m)p^{-rs}m^{-s} \;\lesssim\; \sum_{p>2^j} p^{-s} \;\lesssim\; 2^{j(1-s)}j^{-1}, \tag{63}$$

using the fact that the number of primes $2^J \leq p < 2^{J+1}$ is bounded by $2^J\, J^{-1}$ for all $J \geq j$.

The second term on the right side of (60) is estimated similarly, except that here we use the fact that $p^{\gamma_p} > 2^j$ for $p < 2^j$. We have

$$\begin{aligned}\sum_{p<2^j}\sum_{n\geq 1} b_K(p^{\gamma_p}n)p^{-\gamma_p s}n^{-s} &= \sum_{p<2^j}\sum_{r=\gamma_p}^{\infty}\sum_{m\geq 1, p\nmid m} b_K(p^r)b_K(m)p^{-rs}m^{-s} \\ &\lesssim \sum_{p<2^j}\sum_{r=\gamma_p}^{\infty} p^{-rs} \;\lesssim\; \sum_{p<2^j} p^{-\gamma_p s} \;\lesssim\; 2^{j(1-s)}j^{-1},\end{aligned} \tag{64}$$

as the number of primes $p < 2^j$ is bounded by $2^j j^{-1}$. Estimate (56) follows immediately from (63)-(64). $\square$

In dimensions $d > 2k+2$, Lemma 4 with $s = d/2 - k \geq 3/2$ implies that

$$\sum_{d(R)\nmid q_j} J_{\Lambda^2 T}[R] \;\lesssim\; \Lambda^{k(n-k-1)}d(R)^{-d/2} \;\lesssim\; \Lambda^{k(n-k-1)}2^{-j/2}j^{-1}, \tag{65}$$

with $\Lambda = 2^l$. Finally, by (51) (54)-(55) and (65) one obtains, in dimensions $d > 2k+2$,

$$\int_{I_k}\sup_{\mathcal{X}} \mathbf{1}_{\Omega^c_{j,k}}(\xi_1)|\theta_{d,k}(X+i\varepsilon T^{-1}, -\mathcal{X}, 0)|\,dX \;\lesssim \Lambda^{k(d-k-1)}\big(2^{-j/2}j^{-1} + 2^{-3j/2} + 2^{-3l}\big) \;\lesssim\; \Lambda^{k(d-k-1)}2^{-j/2}j^{-1}. \tag{66}$$

Estimate (27) follows immediately from (28) and (66).

References

[1] T. C. Anderson, A. V. Kumchev, and E. A. Palsson, *Discrete maximal operators over surfaces of higher codimension*, arXiv:2006.09968

[2] J. Bourgain, *A Szemerédi type theorem for sets of positive density in* $\mathbb{R}^k$, Israel J. Math. 54 (1986), no. 3, 307–316.

[3] J. Bourgain, *On the maximal ergodic theorem for certain subsets of the integers*, Israel J. Math. 61 (1988), no. 1, 39-72.

[4] J. Bourgain, *On the pointwise ergodic theorem on* L^p *for arithmetic sets*, Israel J. Math. 61 (1988), 73-84.

[5] J. Bourgain, *Pointwise ergodic theorems for arithmetic sets. With an appendix by the author, Harry Furstenberg, Yitzhak Katznelson and Donald S. Ornstein*, Publ. Math.-Paris 69 (1989), no. 1, pp. 5-45.

[6] J. Bourgain, *Eigenfunction bounds for the Laplacian on the n-torus*, Internat. Math. Res. Notices (1993), no. 3, 61-66.

[7] B. Cook, N. Lyall, Á. Magyar, *Multilinear maximal operators associated to simplices*, to appear in Journal of the LMS

[8] L. Huckaba, N. Lyall and Á. Magyar, *Simplices and sets of positive upper density in* $\mathbb{R}^d$, Proc. Amer. Math. Soc. 145 (2017), no. 6, 2335-2347

[9] Y.Kitaoka, *Siegel modular forms and representation by quadratic forms* Lectures on Mathe- matics and Physics, Tata Institute of Fundamental Research, Springer-Verlag, (1986).

[10] N. Lyall and Á. Magyar, *Distances and trees in dense subsets of* $\mathbb{Z}^d$, arXiv 1509.09298

[11] N. Lyall and Á. Magyar, *Weak hypergraph regularity and applications to geometric Ramsey theory*, to appear in Trans. Amer. Math. Soc.

[12] N. Lyall, Á. Magyar, A. Newman, P. Woolfitt, *The discrete spherical maximal function: A new proof of* ℓ^2*-boundedness*, to appear in the Proc. Amer. Math. Soc.

[13] Á. Magyar, E. M. Stein, and S. Wainger, *Discrete analogues in harmonic analysis: Spherical averages*, Ann. Math. (2) 155 (2002), no. 1, 189-208.
[14] Á. Magyar, *Distance sets of large sets of integer points*, Israel J. Math., v (2008) pp.
[15] Á. Magyar, *k-point configurations in sets of positive density of Z^n*, Duke Math. J., v 146/1, (2009) pp. 1-34.
[16] S. Raghavan, *Modular forms of degree n and representation by quadratic forms*, Ann. Math. (2) 70 (1959), no. 3, 446-477.
[17] C. L. Siegel, *On the theory of indefinite quadratic forms*, Ann. of Math. (2) 45 (1944), 577-622.
[18] R. C. Vaughan, *The Hardy-Littlewood Method*, Second ed., Cambridge University Press, Cambridge, 1997.

Department of Mathematics, The University of Georgia, Athens, GA 30602, USA

Email address: lyall@math.uga.edu

Email address: magyar@math.uga.edu

Email address: alxjames@uga.edu

Email address: pwoolfitt@uga.edu