

# DISTANCES AND TREES IN DENSE SUBSETS OF $\mathbb{Z}^d$

NEIL LYALL      ÁKOS MAGYAR

ABSTRACT. In [4] Katznelson and Weiss establish that all sufficiently large distances can always be attained between pairs of points from any given measurable subset of  $\mathbb{R}^2$  of positive upper (Banach) density. A second proof of this result, as well as a stronger “pinned variant”, was given by Bourgain in [2] using Fourier analytic methods. In [9] the second author adapted Bourgain’s Fourier analytic approach to established a result analogous to that of Katznelson and Weiss for subsets  $\mathbb{Z}^d$  provided  $d \geq 5$ . We present a new direct proof of this discrete distance set result and generalize this to arbitrary trees. Using appropriate discrete spherical maximal function theorems we ultimately establish the natural “pinned variants” of these results.

## 1. INTRODUCTION

**1.1. Distance sets and existing results.** A result of Katznelson and Weiss [4] states that all sufficiently large distances can always be attained between pairs of points from any given measurable subset of  $\mathbb{R}^2$  of positive upper (Banach) density. Specifically, if  $A$  is a measurable subset of  $\mathbb{R}^2$  of positive upper Banach density, they established the existence of a threshold  $\lambda_0 = \lambda_0(A)$  such that the distance set

$$\text{dist}(A) = \{|x - y| : x, y \in A\} \supseteq [\lambda_0, \infty).$$

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

$$\delta^*(A) := \limsup_{N \rightarrow \infty} \sup_{t \in \mathbb{R}^d} \frac{|A \cap (t + Q_N)|}{|Q_N|},$$

where  $|\cdot|$  denotes Lebesgue measure on  $\mathbb{R}^d$  and  $Q_N$  denotes the cube  $[-N/2, N/2]^d$ .

This result was later established using Fourier analytic methods by Bourgain in [2]. Bourgain also established a “pinned variant”, namely that for any  $\lambda_1 \geq \lambda_0$  there is a fixed  $x \in A$  such that

$$\text{dist}(A; x) = \{|x - y| : y \in A\} \supseteq [\lambda_0, \lambda_1].$$

In [9] the second author adapted Bourgain’s Fourier analytic approach to established a result analogous to that of Katznelson and Weiss for subsets  $\mathbb{Z}^d$ , namely that if  $A \subseteq \mathbb{Z}^d$  of positive upper Banach density and  $d \geq 5$ , then there exists  $\lambda_0 = \lambda_0(A)$  and an integer  $q$ , depending on  $d$  and the density of  $A$ , such that

$$\text{dist}^2(A) = \{|x - y|^2 : x, y \in A\} \supseteq [\lambda_0, \infty) \cap q^2\mathbb{Z}.$$

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

$$\delta^*(A) := \limsup_{N \rightarrow \infty} \sup_{t \in \mathbb{Z}^d} \frac{|A \cap (t + Q_N)|}{|Q_N|},$$

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

Note that 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 result above must be divisible by the least common multiple of all integers  $1 \leq r \leq \delta^*(A)^{-1/d}$ .

---

2000 *Mathematics Subject Classification.* 11B30.

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

**1.2. New results.** We will denote, for any integer  $\lambda$ , the discrete sphere of radius  $\sqrt{\lambda}$  by  $S_\lambda$ , namely

$$S_\lambda := \{x \in \mathbb{R}^d : |x|^2 = \lambda\} \cap \mathbb{Z}^d.$$

In this paper we will present a new direct proof of the following discrete distance set result from [9].

**Theorem 1** (Unpinned Distances). *Let  $A \subseteq \mathbb{Z}^d$  with  $d \geq 5$  and  $\delta^*(A) > 0$ .*

*There exist  $q = q(\delta^*(A))$  and  $\lambda_0 = \lambda_0(A)$  such that for any integer  $\lambda \geq \lambda_0$  there exist a pair of points*

$$\{x, x + x_1\} \subseteq A \quad \text{with} \quad |x_1|^2 = q^2 \lambda.$$

*In fact, for any  $\varepsilon > 0$  there exist  $q = q(\varepsilon, d)$  and  $\Lambda_0 = \Lambda_0(A, \varepsilon)$  such that for any integer  $\lambda \geq \Lambda_0$  one has*

$$\frac{|A \cap (x + qS_\lambda)|}{|S_\lambda|} > \delta^*(A) - \varepsilon \quad \text{for some } x \in A.$$

It is impossible to take  $\lambda_0$  above to depend only on  $\delta^*(A)$  since 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  $\{x, x + x_1\}$  with  $|x_1| = qdM$ .

By considering sets  $A$  of the form  $\bigcup_{s \in \{1, \dots, q\}^d} A_s$  with each set  $A_s$  a “random” subset of the congruence class  $s + (q\mathbb{Z})^d$  one can further easily see that the second conclusion above is best possible or “ $\varepsilon$ -optimal”.

The first main new result of this paper is the following “pinned variant” of Theorem 1 above, in other words a discrete analogue of Bourgain’s pinned distances theorem in [2].

**Theorem 2** (Pinned Distances). *Let  $\varepsilon > 0$  and  $A \subseteq \mathbb{Z}^d$  with  $d \geq 5$ .*

*There exist  $q = q(\varepsilon, d)$  and  $\Lambda_0 = \Lambda_0(A, \varepsilon)$  such that for any  $\Lambda_1 \geq \Lambda_0$  there exists a fixed  $x \in A$  such that*

$$\frac{|A \cap (x + qS_\lambda)|}{|S_\lambda|} > \delta^*(A) - \varepsilon \quad \text{for all integers } \Lambda_0 \leq \lambda \leq \Lambda_1.$$

Our approach to Theorems 1 and 2 allows us to establish generalizations to arbitrary finite trees.

**Definition 1** (Trees and Down-Labeled Trees). A *tree*  $\Gamma = \Gamma(V, E)$  is a connected acyclic graph. It is easy to verify that the vertex set  $V$  and edge set  $E$  of any given finite tree  $\Gamma = \Gamma(V, E)$  must satisfy  $|E| = |V| - 1$  and that there exists an enumeration of  $V = \{v_0, v_1, \dots, v_n\}$  so that each edge  $e_j$  from  $E = \{e_1, \dots, e_n\}$  takes the form  $e_j = \{v_i, v_j\}$  for some unique  $0 \leq i < j$ . We shall refer to a tree  $\Gamma$  on  $n + 1$  vertices as *down-labeled* if its vertex set  $\{v_0, v_1, \dots, v_n\}$  has been enumerated as described above and use, for each  $1 \leq j \leq n$ , the convenient notation  $i_\Gamma(j)$  to denote the unique  $0 \leq i < j$  for which  $\{v_i, v_j\} \in E$ .



FIGURE 1. A tree

The following result gives the aforementioned generalizations of Theorem 1 to arbitrary finite trees.

**Theorem 3** (Unpinned Trees). *Let  $\Gamma$  be a down-labeled tree on  $n + 1$  vertices.*

*If  $A \subseteq \mathbb{Z}^d$  with  $d \geq 5$  and  $\delta^*(A) > 0$ , then there exist  $q = q(\delta^*(A))$  and  $\Lambda_0 = \Lambda_0(A, n)$  such that for any integers  $\lambda_1, \dots, \lambda_n \geq \Lambda_0$  there exists*

$$\{x + x_0, x + x_1, \dots, x + x_n\} \subseteq A \quad \text{with } x_0 = 0 \text{ and } |x_j - x_{i_\Gamma(j)}|^2 = q^2 \lambda_j \text{ for all } 1 \leq j \leq n.$$

*In fact, for any  $\varepsilon > 0$  there exist  $q = q(\varepsilon, d)$  and  $\Lambda_0 = \Lambda_0(A, n, \varepsilon)$  such that for any integers  $\lambda_1, \dots, \lambda_n \geq \Lambda_0$*

$$\frac{|\{\{x_1, \dots, x_n\} \subseteq A - x : x_j - x_{i_\Gamma(j)} \in qS_{\lambda_j} \text{ for } 1 \leq j \leq n\}|}{|S_{\lambda_1}| \cdots |S_{\lambda_n}|} > \delta^*(A)^n - \varepsilon$$

*for some  $x \in A$ , with the understanding that  $x_0 = 0$ .*

We remark that no result of this type can possibly hold with a threshold  $\Lambda_0$  independent of  $n$  and that a quantitatively weaker (and non-optimal) version of Theorem 3, in which the parameter  $q$  also depended on  $n$ , was previously established by Bulinski in [3] using methods from Ergodic theory. Related results on the embedding of specific distance graphs in subsets of  $\mathbb{R}^d$  of positive density have recently been obtained by the authors in [8], for subsets of  $\mathbb{R}^d$  of measure zero but large Hausdorff dimension by Bennett, Iosevich, and Taylor in [1], and in the setting of finite field geometries by Iosevich and Parshall in [6].

Our main result, which we stress does not follow from the arguments presented in [3], is the following ‘‘pinned variant’’ of Theorem 3 which generalizes Theorem 2 above.

**Theorem 4** (Pinned Trees). *Let  $\Gamma$  be a down-labeled tree on  $n + 1$  vertices,  $\varepsilon > 0$ , and  $A \subseteq \mathbb{Z}^d$  with  $d \geq 5$ .*

*There exist  $q = q(\varepsilon, d)$  and  $\Lambda_0 = \Lambda_0(A, n, \varepsilon)$  such that for any  $\Lambda_1 \geq \Lambda_0$  there exists a fixed  $x \in A$  such that*

$$\frac{|\{\{x_1, \dots, x_n\} \subseteq A - x : x_j - x_{i_\Gamma(j)} \in qS_{\lambda_j} \text{ for } 1 \leq j \leq n\}|}{|S_{\lambda_1}| \cdots |S_{\lambda_n}|} > \delta^*(A)^n - \varepsilon$$

*for any choice of integers  $\lambda_1, \dots, \lambda_n$  that satisfy  $\Lambda_0 \leq \lambda_1, \dots, \lambda_n \leq \Lambda_1$ , with the understanding that  $x_0 = 0$ .*

### 1.3. Outline of paper.

In Section 2 we state analogues of Theorems 1-4 for uniformly distributed subsets of  $\mathbb{Z}^d$  and reduce their proofs to that of analogous results for uniformly distributed compact subsets of  $\mathbb{Z}^d$ .

In Section 3 we complete the proofs of Theorems 1 and 3 by proving the analogous result for uniformly distributed compact subsets of  $\mathbb{Z}^d$ , namely Proposition 1. To do this we introduce a norm which measures the uniformity of distribution within residue classes modulo  $q$  with respect to a scale  $L$ . We then prove that this norm controls the frequency with which certain distances appear in compact subset of  $\mathbb{Z}^d$ , this is analogous to the so-called von-Neumann type inequalities in additive combinatorics. We ultimately demonstrate that this control on the frequency with which certain distances appear allows us to actually control the frequency with which trees with certain prescribed edge lengths appear in compact subset of  $\mathbb{Z}^d$ .

In Sections 4 and 5 we complete the proofs of Theorems 2 and 4 by proving the analogous result for uniformly distributed compact subsets of  $\mathbb{Z}^d$ , namely Proposition 2. In Section 4 we reduce matters to the *Discrete Spherical Maximal Function Theorem* of Magyar, Stein and Wainger [10] and a closely related ‘‘mollified variant’’ thereof, namely Proposition 5, whose statement and proof we presented in Section 5.

## 2. REDUCTION TO UNIFORMLY DISTRIBUTED COMPACT SUBSETS OF $\mathbb{Z}^d$

### 2.1. Distances and Trees in Uniformly Distributed Subsets of $\mathbb{Z}^d$ .

**Definition 2** (Definition of  $q_\eta$  and  $\eta$ -uniform distribution). For any  $\eta > 0$  we define

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

with  $C > 0$  a (sufficiently) large absolute constant and  $A \subseteq \mathbb{Z}^d$  to be  $\eta$ -uniformly distributed (modulo  $q_\eta$ ) if its relative upper Banach density on any residue class modulo  $q_\eta$  never exceeds  $(1 + \eta^4)$  times its density on  $\mathbb{Z}^d$ , namely if for all  $s \in \{1, \dots, q_\eta\}^d$  one has

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

Theorems 1 and 3 are immediate consequences, via an easy density increment argument which we sketch in Section 2.2 below, of the following analogous result for uniformly distributed sets.

**Theorem 5** (Theorems 1 and 3 for Uniformly Distributed Sets).

Let  $\varepsilon > 0$  and  $A \subseteq \mathbb{Z}^d$  with  $d \geq 5$  be  $\eta$ -uniformly distributed for some  $\eta > 0$ .

(i) If  $0 < \eta \ll \varepsilon^2$ , then there exists  $\Lambda_0 = \Lambda_0(A, \eta)$  such that for any integer  $\lambda \geq \Lambda_0$  one has

$$\frac{|A \cap (x + S_\lambda)|}{|S_\lambda|} > \delta^*(A) - \varepsilon \quad \text{for some } x \in A$$

(ii) If  $0 < \eta \ll \varepsilon^2/n$ , then for any down-labeled tree  $\Gamma$  on  $n+1$  vertices there exists  $\Lambda_0 = \Lambda_0(A, n, \eta)$  such that for any integers  $\lambda_1, \dots, \lambda_n \geq \Lambda_0$  there exist  $x \in A$  for which

$$\frac{|\{\{x_1, \dots, x_n\} \subseteq A - x : |x_j - x_{i_\Gamma(j)}|^2 = \lambda_j \text{ for } 1 \leq j \leq n\}|}{|S_{\lambda_1}| \cdots |S_{\lambda_n}|} > \delta^*(A)^n - \varepsilon$$

with the understanding that  $x_0 = 0$ .

In Theorem 5 above, and throughout the paper, we use the notation  $\alpha \ll \beta$  to denote that  $\alpha \leq c\beta$  for some suitably small constant  $c > 0$ .

Theorems 2 and 4 likewise reduce to the following analogous result for uniformly distributed sets.

**Theorem 6** (Theorems 2 and 4 for Uniformly Distributed Sets).

Let  $\varepsilon > 0$  and  $A \subseteq \mathbb{Z}^d$  with  $d \geq 5$  be  $\eta$ -uniformly distributed.

(i) If  $0 < \eta \ll \varepsilon^3$ , then there exist  $\Lambda_0 = \Lambda_0(A, \eta)$  such that for any  $\Lambda_1 \geq \Lambda_0$  there exists a fixed  $x \in A$  such that

$$\frac{|A \cap (x + S_\lambda)|}{|S_\lambda|} > \delta^*(A) - \varepsilon \quad \text{for all integers } \Lambda_0 \leq \lambda \leq \Lambda_1.$$

(ii) If  $0 < \eta \ll \varepsilon^3/n^3$ , then for any down-labeled tree  $\Gamma$  on  $n+1$  vertices there exist  $\Lambda_0 = \Lambda_0(A, n, \eta)$  such that for any  $\Lambda_1 \geq \Lambda_0$  there exists a fixed  $x \in A$  such that

$$\frac{|\{\{x_1, \dots, x_n\} \subseteq A - x : |x_j - x_{i_\Gamma(j)}|^2 = \lambda_j \text{ for } 1 \leq j \leq n\}|}{|S_{\lambda_1}| \cdots |S_{\lambda_n}|} > \delta^*(A)^n - \varepsilon$$

for all integers  $\lambda_1, \dots, \lambda_n$  that satisfy  $\Lambda_0 \leq \lambda_1, \dots, \lambda_n \leq \Lambda_1$ , with the understanding that  $x_0 = 0$ .

**2.2. Proof that Theorem 5 (i) implies Theorem 1.** Let  $0 < \varepsilon \leq \alpha \leq 1$  and  $A \subseteq \mathbb{Z}^d$  with  $d \geq 5$ . To prove Theorem 1 it is enough to prove that if  $\delta^*(A) \geq \alpha$  then there exists  $\Lambda_0 = \Lambda_0(A, \varepsilon)$  and  $q = q(\varepsilon, d)$  such that for any integer  $\lambda \geq \Lambda_0$  one has

$$(1) \quad \frac{|A \cap (x + qS_\lambda)|}{|S_\lambda|} > \alpha - \varepsilon \quad \text{for some } x \in A.$$

Let  $0 < \eta \ll \varepsilon^2$ . We prove the above for  $\alpha_k := (1 + \eta^4)^{-k}$  inductively for all  $k \geq 0$ , using Theorem 5 (i). For  $k = 0$  the statement is trivial as  $\delta^*(A) = \alpha_0 = 1$  and hence  $A$  contains arbitrarily large cubic grids. Suppose it holds for  $\alpha = \alpha_k$  and assume that  $\delta^*(A) \geq \alpha_{k+1}$ . If  $A$  is  $\eta$ -uniformly distributed then the result holds for  $\alpha = \alpha_{k+1}$  by Theorem 5. In the opposite case there is an  $s \in \mathbb{Z}^d$  so that  $\delta^*(A | s + (q_\eta \mathbb{Z})^d) > (1 + \eta^4)\alpha$ . Let  $\phi : s + (q_\eta \mathbb{Z})^d \rightarrow \mathbb{Z}^d$  be defined by  $\phi(m) := q_\eta^{-1}(m - s)$  and let  $A' := \phi(A)$ . Then  $\delta^*(A') \geq \alpha_k$  thus (1) holds for  $A'$  and  $\alpha = \alpha_k$ , with some  $q' = q'(\varepsilon, d)$  and  $x' \in A'$ . Note that  $|A' \cap (x' + q'S_\lambda)| = |A \cap ((q_\eta x' + s) + q_\eta q'S_\lambda)|$  which implies that (1) holds for  $A$ ,  $\alpha = \alpha_{k+1}$  with  $q = q_\eta q'$  and  $x = q_\eta x' + s$ .  $\square$

**2.3. Compact variants of Theorems 5 and 6.** We shall now show that Theorems 5 and 6 can in turn be directly deduced from analogous *compact* variants, namely Corollary 1 and Proposition 2 below.

In what follows we shall use  $1_B$  to denote the characteristic function of any  $B \subseteq \mathbb{Z}^d$  and define

$$\sigma_\lambda = |S_\lambda|^{-1} 1_{S_\lambda}.$$

First we introduce a second related notion of uniformity.

**Definition 3** (Definition of  $(\eta, L)$ -uniform distribution). Let  $\eta > 0$  and  $q_\eta \ll \eta^2 L \ll \eta^6 N$ .

We define  $A \subseteq Q_N$  to be  $(\eta, L)$ -uniformly distributed if

$$\frac{1}{|Q_N|} \sum_{t \in Q_N} \left| \frac{|A \cap (t + Q_{q_\eta, L})|}{|Q_{q_\eta, L}|} - \frac{|A|}{|Q_N|} \right|^2 \leq \eta^2,$$

where as before  $Q_N$  denotes the discrete cube  $[-N/2, N/2]^d \cap \mathbb{Z}^d$  and now  $Q_{q, L} := Q_L \cap (q\mathbb{Z})^d$ .

**Proposition 1** (Average count of distances and trees in uniformly distributed subsets of  $Q_N$ ).

Let  $\Gamma$  be a down-labeled tree on  $n + 1$  vertices. If  $\eta > 0$  and  $A \subseteq Q_N \subseteq \mathbb{Z}^d$  with  $d \geq 5$  is  $(\eta, L)$ -uniformly distributed, then for all integers  $\lambda_1, \dots, \lambda_n$  that satisfy  $\eta^{-4} L^2 \leq \lambda_1, \dots, \lambda_n \leq \eta^4 N^2$  one has

$$\frac{1}{|Q_N|} \sum_{x \in \mathbb{Z}^d} 1_A(x) \sum_{x_1, \dots, x_n \in \mathbb{Z}^d} \prod_{j=1}^n 1_A(x - x_j) \sigma_{\lambda_j}(x_j - x_{i_\Gamma(j)}) = \left( \frac{|A|}{|Q_N|} \right)^{n+1} + O(n\eta)$$

with the understanding that  $x_0 = 0$ .

In particular, by taking  $n = 1$  above, one obtains that

$$\frac{1}{|Q_N|} \sum_{x \in \mathbb{Z}^d} 1_A(x) \sum_{x_1 \in \mathbb{Z}^d} 1_A(x - x_1) \sigma_\lambda(x_1) = \left( \frac{|A|}{|Q_N|} \right)^2 + O(\eta)$$

for all integers  $\lambda$  that satisfy  $\eta^{-4} L^2 \leq \lambda \leq \eta^4 N^2$ .

It is easy to see that Proposition 1 immediately implies the following

**Corollary 1** (Unpinned distances and trees in uniformly distributed subsets of  $Q_N$ ).

Let  $\varepsilon > 0$  and  $A \subseteq Q_N \subseteq \mathbb{Z}^d$  with  $d \geq 5$  be  $(\eta, L)$ -uniformly distributed for some  $\eta > 0$ .

(i) If  $0 < \eta \ll \varepsilon^2$ , then for all integers  $\lambda$  that satisfy  $\eta^{-4} L^2 \leq \lambda \leq \eta^4 N^2$  there exists  $x \in A$  such that

$$\frac{|A \cap (x + S_\lambda)|}{|S_\lambda|} = \sum_{x_1 \in \mathbb{Z}^d} 1_A(x - x_1) \sigma_\lambda(x_1) > \frac{|A|}{|Q_N|} - \varepsilon.$$

(ii) If  $0 < \eta \ll \varepsilon^2/n$ , then for any down-labeled tree  $\Gamma$  on  $n + 1$  vertices and integers  $\lambda_1, \dots, \lambda_n$  that satisfy  $\eta^{-4} L^2 \leq \lambda_1, \dots, \lambda_n \leq \eta^4 N^2$  there exist  $x \in A$  for which

$$\sum_{x_1, \dots, x_n \in \mathbb{Z}^d} \prod_{j=1}^n 1_A(x - x_j) \sigma_{\lambda_j}(x_j - x_{i_\Gamma(j)}) > \left( \frac{|A|}{|Q_N|} \right)^n - \varepsilon$$

with the understanding that  $x_0 = 0$ .

We will ultimately also establish the following ‘‘pinned’’ variant of Corollary 1.

**Proposition 2** (Pinned distances and trees in uniformly distributed subsets of  $Q_N$ ).

Let  $\varepsilon > 0$  and  $A \subseteq Q_N \subseteq \mathbb{Z}^d$  with  $d \geq 5$  be  $(\eta, L)$ -uniformly distributed.

(i) If  $0 < \eta \ll \varepsilon^3$ , then there exists  $x \in A$  such that

$$\frac{|A \cap (x + S_\lambda)|}{|S_\lambda|} > \frac{|A|}{|Q_N|} - \varepsilon$$

for all integers  $\lambda$  that satisfy  $\eta^{-4} L^2 \leq \lambda \leq \eta^4 N^2$ .

(ii) If  $0 < \eta \ll \varepsilon^3/n^3$ , then for any down-labeled tree  $\Gamma$  on  $n + 1$  vertices there exist  $x \in A$  such that

$$\sum_{x_1, \dots, x_n \in \mathbb{Z}^d} \prod_{j=1}^n 1_A(x - x_j) \sigma_{\lambda_j}(x_j - x_{i_\Gamma(j)}) > \left( \frac{|A|}{|Q_N|} \right)^n - \varepsilon$$

for all integers  $\lambda_1, \dots, \lambda_n$  satisfying  $\eta^{-4} L^2 \leq \lambda_1, \dots, \lambda_n \leq \eta^4 N^2$ , with the understanding that  $x_0 = 0$ .

## 2.4. Reduction of Theorems 5 and 6 to Corollary 1 and Proposition 2.

The task of deducing Theorems 5 and 6 from Corollary 1 and Proposition 2 respectively simply amounts to establishing the following precise relationship between our two notions of uniform distribution.

**Lemma 1.** *Let  $\eta > 0$ . If  $A \subseteq \mathbb{Z}^d$  with  $\delta^*(A) > 0$  is  $\eta$ -uniformly distributed, then there exists a positive integer  $L = L(A, \eta)$  and an arbitrarily large integer  $N$  with  $N \geq \eta^{-4}L$  such that the set  $(A - t_0) \cap Q_N$  satisfies*

$$\frac{|(A - t_0) \cap Q_N|}{|Q_N|} > (1 - 2\eta^4) \delta^*(A)$$

for some  $t_0 \in \mathbb{Z}^d$  and simultaneously has the property that it is  $(C\eta, L)$ -uniformly distributed for some  $C > 0$ .

*Proof.* Since  $A \subseteq \mathbb{Z}^d$  is  $\eta$ -uniformly distributed we know there exists a positive integer  $L = L(A, \eta)$  such that

$$(2) \quad \frac{|A \cap (t + Q_{q_\eta, L})|}{|Q_{q_\eta, L}|} \leq (1 + 2\eta^4) \delta^*(A)$$

for all  $t \in \mathbb{Z}^d$ . Since  $\delta^*(A) > 0$  we further know that there exist arbitrarily large  $N \in \mathbb{N}$  such that

$$(3) \quad \frac{|A \cap (t_0 + Q_N)|}{|Q_N|} \geq (1 - 2\eta^4) \delta^*(A)$$

for some  $t_0 \in \mathbb{Z}^d$ . Combining (2) and (3) we see there exist  $N \in \mathbb{N}$  with  $N \geq \eta^{-4}L$  and  $t_0 \in \mathbb{Z}^d$  such that

$$\frac{|A \cap (t + Q_{q_\eta, L})|}{|Q_{q_\eta, L}|} \leq (1 + 5\eta^4) \frac{|A \cap (t_0 + Q_N)|}{|Q_N|}$$

for all  $t \in \mathbb{Z}^d$ . Setting  $A' := (A - t_0) \cap Q_N$  we further note that since  $A' \cap (t + Q_{q_\eta, L})$  is only supported in  $Q_N + Q_L$  it follows that

$$\frac{|A'|}{|Q_N|} = \frac{1}{|Q_N|} \sum_{t \in \mathbb{Z}^d} \frac{|A' \cap (t + Q_{q_\eta, L})|}{|Q_{q_\eta, L}|} = \frac{1}{|Q_N|} \sum_{t \in Q_N} \frac{|A' \cap (t + Q_{q_\eta, L})|}{|Q_{q_\eta, L}|} + O(L/N),$$

from which one can easily deduce that

$$\frac{1}{|Q_N|} \left| \left\{ t \in Q_N : \frac{|A' \cap (t + Q_{q_\eta, L})|}{|Q_{q_\eta, L}|} \leq (1 - \eta^2) \frac{|A'|}{|Q_N|} \right\} \right| = O(\eta^2)$$

provided  $L/N \ll \eta^2$  and hence that  $A'$  is  $(C\eta, L)$ -uniformly distributed for some  $C > 0$ .  $\square$

We are thus left with proving Propositions 1 and 2. These proofs are presented in Sections 3 and 4 below.

## 3. PROOF OF PROPOSITION 1

### 3.1. Reduction to a Generalized von-Neumann Inequality.

Let  $Q_N$  denote the discrete cube  $[-N/2, N/2]^d \cap \mathbb{Z}^d$  with  $d \geq 5$ .

**Definition 4** (Counting Function for Distances and Trees).

(i) For  $1 \ll \lambda \ll N^2$  and functions  $f_0, f_1 : Q_N \rightarrow [-1, 1]$  we define

$$T(f_0, f_1)(\lambda) = \frac{1}{|Q_N|} \sum_{x \in \mathbb{Z}^d} f_0(x) \sum_{x_1 \in \mathbb{Z}^d} f_1(x - x_1) \sigma_\lambda(x_1).$$

(ii) While, for any given down-labeled tree  $\Gamma$  on  $n + 1$  vertices,  $1 \leq m \leq n$ ,  $1 \ll \lambda_1, \dots, \lambda_m \ll N^2$  and functions  $f_0, f_1, \dots, f_m : Q_N \rightarrow [-1, 1]$ , we define

$$T_{\Gamma, m}(f_0, f_1, \dots, f_m)(\lambda_1, \dots, \lambda_m) = \frac{1}{|Q_N|} \sum_{x \in \mathbb{Z}^d} f_0(x) \sum_{x_1, \dots, x_m \in \mathbb{Z}^d} \prod_{j=1}^m f_j(x - x_j) \sigma_{\lambda_j}(x_j - x_{i_\Gamma(j)})$$

with the understanding that  $x_0 = 0$ .

**Definition 5** ( $U^1(q, L)$ -norm). For  $1 \ll q \ll L \ll N$  and functions  $f : Q_N \rightarrow \mathbb{R}$  we define

$$(4) \quad \|f\|_{U^1(q, L)} = \left( \frac{1}{|Q_N|} \sum_{t \in \mathbb{Z}^d} |f * \chi_{q, L}(t)|^2 \right)^{1/2}$$

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

$$(5) \quad \chi_{q, L}(x) = \begin{cases} \left(\frac{q}{L}\right)^d & \text{if } x \in (q\mathbb{Z})^d \cap [-\frac{L}{2}, \frac{L}{2}]^d \\ 0 & \text{otherwise} \end{cases}.$$

In (4) 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  $U^1(q, L)$ -norm measures the mean square oscillation of a function with respect to cubic grids of size  $L$  and gap  $q$ . It is a simple observation, that we record precisely below, that sets  $A \subseteq Q_N$  that are  $(\eta, L)$ -uniformly distributed have the property that their “balance functions” have small  $U^1(q_\eta, L)$ -norm.

**Lemma 2.** *Let  $\eta > 0$  and  $1 \ll L \ll \eta^4 N$ .*

*If  $A \subseteq Q_N$  is  $(\eta, L)$ -uniformly distributed, then  $\|f_A\|_{U^1(q_\eta, L)} \leq 2\eta$  where  $f_A = 1_A - \frac{|A|}{|Q_N|} 1_{Q_N}$ .*

In light of Lemma 2 we see that the engine that drives our proof of Proposition 1 for  $n = 1$ , and thus our short proof of Theorem 1, via Part (i) of Corollary 1, is the following “generalized von-Neumann inequality”.

**Lemma 3** (Generalized von-Neumann). *Let  $\eta > 0$ , and  $\lambda, L$ , and  $N$  be integers with  $\eta^{-4} L^2 \leq \lambda \leq \eta^4 N^2$ .*

*Given any functions  $f_0, f_1 : Q_N \rightarrow [-1, 1]$  on  $Q_N \subseteq \mathbb{Z}^d$  with  $d \geq 5$  we have*

$$|T(f_0, f_1)(\lambda)| \leq \|f_1\|_{U^1(q_\eta, L)} + O(\eta).$$

In fact, as we shall see below, this result is also sufficient to establish Proposition 1 for any  $n$  and hence Theorem 3 via Part (ii) of Corollary 1.

*Proof of Proposition 1.* Let  $\eta > 0$  and  $A \subseteq Q_N \subseteq \mathbb{Z}^d$  with  $d \geq 5$  be  $(\eta, L)$ -uniformly distributed.

We let  $\alpha = |A|/|Q_N|$  and note that Lemma 2 ensures that  $\|f_A\|_{U^1(q_\eta, L)} \leq 2\eta$  where  $f_A = 1_A - \alpha 1_{Q_N}$ .

*Case  $n = 1$ :* In this case Proposition 1 follows immediately from Lemma 3 since

$$T(1_A, 1_A)(\lambda) = \alpha T(1_A, 1_{Q_N}) + T(1_A, f_A)(\lambda) = \alpha^2 + \|f_A\|_{U^1(q_\eta, L)} + O(\eta)$$

for all integers  $\lambda$  that satisfy  $\eta^{-4} L^2 \leq \lambda \leq \eta^4 N^2$ .

*Case  $n \geq 2$ :* We first fix  $2 \leq m \leq n$  and make two key observations.

Since  $0 \leq i_\Gamma(j) < j$  and  $f_j : Q_N \rightarrow [-1, 1]$  for all  $1 \leq j \leq m$ , and no function  $\sigma_{\lambda_j}$  involves the variable  $x$ , it follows that

$$(6) \quad |T_{\Gamma, m}(f_0, f_1, \dots, f_m)(\lambda_1, \dots, \lambda_m)| \leq \sum_{x_1, \dots, x_{m-1} \in \mathbb{Z}^d} \prod_{j=1}^{m-1} \sigma_{\lambda_j}(x_j - x_{i_\Gamma(j)}) \left| T(\tau_{x_{i_\Gamma(m)}} f_0, f_m)(\lambda_m) \right|$$

where  $\tau_{x_{i_\Gamma(m)}} f_0(x) = f_0(x + x_{i_\Gamma(m)})$ . We also note that

$$(7) \quad T_{\Gamma, m}(f_0, f_1, \dots, f_{m-1}, 1_{Q_N})(\lambda_1, \dots, \lambda_m) = T_{\Gamma, m-1}(f_0, f_1, \dots, f_{m-1})(\lambda_1, \dots, \lambda_{m-1}) + O(\sqrt{\lambda_m}/N).$$

It then follows from Lemma 3, together with observations (6) and (7) above, that

$$(8) \quad \begin{aligned} T_{\Gamma, n}(1_A, \dots, 1_A)(\lambda_1, \dots, \lambda_n) &= \alpha^n T_{\Gamma, 1}(1_A, 1_{Q_N})(\lambda_1) \\ &+ \sum_{m=1}^n \alpha^{n-m} T_{\Gamma, m}(1_A, \dots, 1_A, f_A)(\lambda_1, \dots, \lambda_m) + O(n\eta^2) \\ &= \alpha^{n+1} + O(n \|f_A\|_{U^1(q_\eta, L)}) + O(n\eta) \end{aligned}$$

for all integers  $\lambda_1, \dots, \lambda_n$  that satisfy  $\eta^{-4} L^2 \leq \lambda_1, \dots, \lambda_n \leq \eta^4 N^2$ .  $\square$

In order to prove Proposition 1 we thus left with the final task of establishing Lemma 3.

**3.2. Proof of Lemma 3.** For any  $f : Q_N \rightarrow [-1, 1]$  we define its *Fourier transform*  $\widehat{f} : \mathbb{T}^d \rightarrow \mathbb{C}$  by

$$\widehat{f}(\xi) = \sum_{x \in \mathbb{Z}^d} f(x) e^{-2\pi i x \cdot \xi}$$

noting that the support assumption on  $f$  ensures that the series defining  $\widehat{f}$  converges uniformly to a continuous function on the torus  $\mathbb{T}^d$ , which we will freely identify with the unit cube  $[0, 1)^d$  in  $\mathbb{R}^d$ .

It is easy to verify, using Cauchy-Schwarz and basic properties of the Fourier transform, that

$$|T(f_0, f_1)(\lambda)|^2 \leq \frac{1}{|Q_N|} \int |\widehat{f}_1(\xi)|^2 |\widehat{\sigma}_\lambda(\xi)|^2 d\xi$$

where

$$(9) \quad \widehat{\sigma}_\lambda(\xi) := \frac{1}{|S_\lambda|} \sum_{x \in S_\lambda} e^{-2\pi i x \cdot \xi}.$$

It is clear that whenever  $|\xi|^2 \ll \lambda^{-1}$  there can be no cancellation in the exponential sum (9), in fact it is easy to verify that the same is also true whenever  $\xi$  is *close* to a rational point with *small* denominator. The following proposition is a precise formulation of the fact that this is the only obstruction to cancellation.

**Proposition 3** (Key exponential sum estimates, Proposition 1 in [9]). *Let  $\eta > 0$ . If  $\lambda \geq C\eta^{-4}$  and*

$$\xi \notin (q_\eta^{-1}\mathbb{Z})^d + \{\xi \in \mathbb{R}^d : |\xi|^2 \leq \eta^{-1}\lambda^{-1}\},$$

*then*

$$\left| \frac{1}{|S_\lambda|} \sum_{x \in S_\lambda} e^{-2\pi i x \cdot \xi} \right| \leq \eta.$$

We now define  $\psi_{q_\eta, L}$  indirectly via the identity

$$\widehat{\psi}_{q_\eta, L}(\xi) := \widehat{\chi}_{q_\eta, L}(\xi)^2.$$

Since the definition of  $\chi_{q_\eta, L}$  in (5) above clearly implies that

$$(i) \quad 0 \leq \widehat{\psi}_{q_\eta, L}(\xi) \leq 1 \text{ for all } \xi \in \mathbb{T}^d \quad \text{and} \quad (ii) \quad \widehat{\psi}_{q_\eta, L}(\ell/q_\eta) = 1 \text{ for all } \ell \in \mathbb{Z}^d$$

it follows that

$$0 \leq 1 - \widehat{\psi}_{q_\eta, L}(\xi) \leq \sum_{x \in \mathbb{Z}^d} \psi_{q_\eta, L}(x) |1 - e^{2\pi i x \cdot (\xi - \ell/q_\eta)}| \ll L|\xi - \ell/q_\eta|$$

for all  $\xi \in \mathbb{T}^d$  and  $\ell \in \mathbb{Z}^d$ . In particular we note that

$$(10) \quad |1 - \widehat{\psi}_{q_\eta, L}(\xi)| \ll \eta \quad \text{if } |\xi - \ell/q_\eta| \leq \eta^{-1/2}\lambda^{-1/2} \text{ for some } \ell \in \mathbb{Z}^d.$$

while Proposition 3 ensures that

$$(11) \quad |\widehat{\sigma}_\lambda(\xi)| \leq \eta \quad \text{if } |\xi - \ell/q_\eta| > \eta^{-1/2}\lambda^{-1/2} \text{ for all } \ell \in \mathbb{Z}^d.$$

Hence, if we write

$$|\widehat{\sigma}_\lambda(\xi)|^2 = |\widehat{\sigma}_\lambda(\xi)|^2 \widehat{\psi}_{q_\eta, L}(\xi) + |\widehat{\sigma}_\lambda(\xi)|^2 (1 - \widehat{\psi}_{q_\eta, L}(\xi))$$

use the fact that  $|\widehat{\sigma}_\lambda(\xi)| \leq 1$  for all  $\xi \in \mathbb{T}^d$  and appeal to Plancherel we can deduce that

$$|T(f_0, f_1)(\lambda)|^2 \leq \frac{1}{|Q_N|} \int |\widehat{f}_1(\xi)|^2 \widehat{\chi}_{q_\eta, L}(\xi)^2 d\xi + O(\eta^2) = \|f_1\|_{U^1(q_\eta, L)}^2 + O(\eta^2)$$

which completes the proof of Lemma 3. □



## 4. PROOF OF PROPOSITION 2

Let  $Q_N$  denote the discrete cube  $[-N/2, N/2]^d \cap \mathbb{Z}^d$ .

**Definition 6** (Discrete Spherical Averages). Let  $f : Q_N \rightarrow \mathbb{R}$  be any function.

For any integer  $\lambda$  with  $1 \ll \lambda \ll N^2$  we define the *discrete spherical average*

$$\mathcal{A}_\lambda(f)(x) := f * \sigma_\lambda(x) = \frac{1}{|S_\lambda|} \sum_{y \in S_\lambda} f(x - y).$$

In Sections 4.1 and 4.2 below we reduce Proposition 2 to the *Discrete Spherical Maximal Function Theorem* of Magyar, Stein and Wainger [10], see Proposition 4, and a new ‘‘mollified variant’’ thereof, namely Proposition 5. The statement and proof of Proposition 5 is presented in Section 5.

## 4.1. Proof of Part (i) of Proposition 2.

Let  $\varepsilon > 0$  and  $0 < \eta \ll \varepsilon^3$ . Suppose, contrary to Part (i) of Proposition 2, that there exists a set  $A \subseteq Q_N \subseteq \mathbb{Z}^d$  with  $d \geq 5$  and  $\alpha = |A|/|Q_N| > 0$  that it is  $(\eta, L)$ -uniformly distributed, but has the property that for every  $x \in A$  there exists an integer  $\lambda$  with  $\eta^{-4}L^2 \leq \lambda \leq \eta^4N^2$  such that

$$\mathcal{A}_\lambda(1_A)(x) = \frac{|A \cap (x + S_\lambda)|}{|S_\lambda|} \leq \alpha - \varepsilon.$$

It easily follows that for every  $x \in A' := A \cap Q_{N-\sqrt{\lambda}}$  there exists an integer  $\lambda$  with  $\eta^{-4}L^2 \leq \lambda \leq \eta^4N^2$  such that

$$\mathcal{A}_\lambda(f_A)(x) \leq -\varepsilon$$

where  $f_A = 1_A - \alpha 1_{Q_N}$ . Hence for every  $x \in A'$  we may conclude that

$$(12) \quad \mathcal{A}_*(f_A)(x) \geq \varepsilon$$

where for any function  $f : \mathbb{Z}^d \rightarrow \mathbb{R}$ ,  $\mathcal{A}_*(f)$  denotes the *discrete spherical maximal function* defined by

$$\mathcal{A}_*(f)(x) := \sup_{\eta^{-4}L^2 \leq \lambda \leq \eta^4N^2} |\mathcal{A}_\lambda(f)(x)|.$$

**Proposition 4** ( $\ell^2$ -Boundedness of the Discrete Spherical Maximal Function [10]). *If  $d \geq 5$ , then*

$$\sum_{x \in \mathbb{Z}^d} |\mathcal{A}_*(f)(x)|^2 \leq C \sum_{x \in \mathbb{Z}^d} |f(x)|^2.$$

Since (12) implies, after an application of Cauchy-Schwarz, the inequality

$$(13) \quad \frac{\alpha \varepsilon}{2} \leq \frac{1}{|Q_N|} \sum_{x \in \mathbb{Z}^d} 1_A(x) \mathcal{A}_*(f_A)(x) \leq \alpha^{1/2} \left( \frac{1}{|Q_N|} \sum_{x \in \mathbb{Z}^d} |\mathcal{A}_*(f_A)(x)|^2 \right)^{1/2}$$

it follows that

$$(14) \quad \frac{\alpha^{1/2} \varepsilon}{2} \leq \left( \frac{1}{|Q_N|} \sum_{x \in \mathbb{Z}^d} |\mathcal{A}_*(f_A * \chi_{q_\eta, L})(x)|^2 \right)^{1/2} + \left( \frac{1}{|Q_N|} \sum_{x \in \mathbb{Z}^d} |\mathcal{A}_*(f_A - f_A * \chi_{q_\eta, L})(x)|^2 \right)^{1/2}.$$

In light of Proposition 4 it follows that the first sum above satisfies

$$\left( \frac{1}{|Q_N|} \sum_{x \in \mathbb{Z}^d} |\mathcal{A}_*(f_A * \chi_{q_\eta, L})(x)|^2 \right)^{1/2} \leq C \left( \frac{1}{|Q_N|} \sum_{t \in \mathbb{Z}^d} |f_A * \chi_{q_\eta, L}(t)|^2 \right)^{1/2} = C \|f_A\|_{U^1(q_\eta, L)} \leq 2C\eta.$$

Estimate (14) will therefore lead to a contradiction, if  $\eta$  is chosen sufficiently small with respect to  $\varepsilon^3$ , and hence complete the proof of Proposition 2 if we establish that the second sum in (14) satisfies

$$(15) \quad \left( \frac{1}{|Q_N|} \sum_{x \in \mathbb{Z}^d} |\mathcal{A}_*(f_A - f_A * \chi_{q_\eta, L})(x)|^2 \right)^{1/2} \leq C\eta^{1/3} \alpha^{1/2}$$

for some absolute constant  $C > 0$ . Estimate (15) follows immediately from Proposition 5 in Section 5 below.

#### 4.2. Proof of Part (ii) of Proposition 2.

Let  $\varepsilon > 0$  and  $0 < \eta \ll \varepsilon^3/n^3$ . Suppose, contrary to Part (ii) of Proposition 2, that there exists a set  $A \subseteq Q_N \subseteq \mathbb{Z}^d$  with  $d \geq 5$  and  $\alpha = |A|/|Q_N| > 0$  that it is  $(\eta, L)$ -uniformly distributed, but has the property that for every  $x \in A$  there exist integers  $\lambda_1, \dots, \lambda_n$  satisfying  $\eta^{-4}L^2 \leq \lambda_1, \dots, \lambda_n \leq \eta^4N^2$  such that

$$(16) \quad \sum_{x_1, \dots, x_n \in \mathbb{Z}^d} \prod_{j=1}^n 1_A(x - x_j) \sigma_{\lambda_j}(x_j - x_{i_\Gamma(j)}) \leq \alpha^n - \varepsilon$$

where  $x_0 = 0$ . As in (8) we may write the left hand side of (16) for every  $x \in A' := A \cap Q_{(1-n\eta^2)N}$  as

$$\alpha^n + \sum_{m=1}^n \alpha^{n-m} \left[ \sum_{x_1, \dots, x_{m-1} \in \mathbb{Z}^d} \prod_{j=1}^{m-1} 1_A(x - x_j) \sigma_{\lambda_j}(x_j - x_{i_\Gamma(j)}) \sum_{x_m \in \mathbb{Z}^d} f_A(x - x_m) \sigma_{\lambda_m}(x_m - x_{i_\Gamma(m)}) \right]$$

where  $f_A = 1_A - \alpha 1_{Q_N}$ . It therefore follows, from the pigeonhole principle, that for every  $x \in A'$  there must exist  $2 \leq m \leq n$  and integers  $\lambda_1, \dots, \lambda_{m-1}$  satisfying  $\eta^{-4}L^2 \leq \lambda_1, \dots, \lambda_{m-1} \leq \eta^4N^2$  such that

$$\sum_{x_1, \dots, x_{m-1} \in \mathbb{Z}^d} \prod_{j=1}^{m-1} \sigma_{\lambda_j}(x_j - x_{i_\Gamma(j)}) |\mathcal{A}_*(f_A)(x - x_{i_\Gamma(m)})| \geq \varepsilon/n.$$

Since

$$\sum_{x \in \mathbb{Z}^d} \sum_{x_1, \dots, x_{m-1} \in \mathbb{Z}^d} \prod_{j=1}^{m-1} \sigma_{\lambda_j}(x_j - x_{i_\Gamma(j)}) |\mathcal{A}_*(f_A)(x - x_{i_\Gamma(m)})|^2 = \sum_{x \in \mathbb{Z}^d} |\mathcal{A}_*(f_A)(x)|^2$$

for any choice of integers  $\lambda_1, \dots, \lambda_{m-1}$ , it follows by Cauchy-Schwarz that

$$\frac{1}{|Q_N|} \sum_{x \in \mathbb{Z}^d} 1_A(x) \sum_{x_1, \dots, x_{m-1} \in \mathbb{Z}^d} \prod_{j=1}^{m-1} \sigma_{\lambda_j}(x_j - x_{i_\Gamma(j)}) |\mathcal{A}_*(f_A)(x - x_{i_\Gamma(m)})| \leq \alpha^{1/2} \left( \frac{1}{|Q_N|} \sum_{x \in \mathbb{Z}^d} |\mathcal{A}_*(f_A)(x)|^2 \right)^{1/2}.$$

We can therefore conclude that

$$(17) \quad \frac{1}{|Q_N|} \sum_{x \in \mathbb{Z}^d} |\mathcal{A}_*(f_A)(x)|^2 \geq \frac{\alpha \varepsilon^2}{n^2}$$

an estimate which, in light of Proposition 5 below, leads to a contradiction if  $\eta$  is chosen sufficiently small with respect to  $\varepsilon^3/n^3$  (as described after estimate (13) in Section 4.1 above).

### 5. A “MOLLIFIED” DISCRETE SPHERICAL MAXIMAL FUNCTION THEOREM

Let  $\eta > 0$  and  $\lambda, L$ , and  $N$  be integers that satisfy  $\eta^{-4}L^2 \leq \lambda \leq \eta^4N^2$ . For functions  $f : Q_N \rightarrow [-1, 1]$  we now define

$$\mathcal{A}_{\lambda, \eta}(f)(x) := \mathcal{A}_\lambda(f - f * \chi_{q_\eta, L})(x) = f * (\sigma_\lambda - \sigma_\lambda * \chi_{q_\eta, L})(x)$$

where  $\sigma_\lambda = \frac{1}{|S_\lambda|} 1_{S_\lambda}$ , and introduce the corresponding “mollified” discrete spherical maximal function

$$(18) \quad \mathcal{A}_{*, \eta}(f)(x) := \sup_{\eta^{-4}L^2 \leq \lambda \leq \eta^4N^2} |\mathcal{A}_{\lambda, \eta}(f)(x)|.$$

We note that the convolution operator  $\mathcal{A}_{\lambda, \eta}$  corresponds to the Fourier multiplier  $\widehat{\sigma_{\lambda, \eta}} := \widehat{\sigma_\lambda}(1 - \widehat{\chi_{q_\eta, L}})$ .

**Proposition 5** ( $\ell^2$ -Decay of the “Mollified” Discrete Spherical Maximal Function). *If  $d \geq 5$ , then for any  $\eta > 0$  we have*

$$(19) \quad \sum_{x \in \mathbb{Z}^d} |\mathcal{A}_{*, \eta}(f)(x)|^2 \leq C\eta^{2/3} \sum_{x \in \mathbb{Z}^d} |f(x)|^2.$$

*Proof of Proposition 5.* We follow the proof of Proposition 4 as given in [10]. For each  $x \in \mathbb{Z}^d$  we now define

$$\mathcal{B}_\lambda(f)(x) = \mathcal{A}_{\lambda^2}(f)(x)$$

noting that when considering  $\mathcal{B}_\lambda$  we are now allowing all values of  $\lambda$  for which  $\lambda^2$  is an integer, and that

$$\mathcal{B}_*(f)(x) := \sup_{\eta^{-2}L \leq \lambda \leq \eta^2 N} \mathcal{B}_\lambda(f)(x) = \mathcal{A}_*(f)(x) \quad \text{and} \quad \mathcal{B}_{*,\eta}(f)(x) = \mathcal{B}_*(f - f * \chi_{q_\eta, L})(x).$$

We now recall the approximation to  $\mathcal{B}_\lambda$  given in Section 3 of [10] as a convolution operator  $\mathcal{M}_\lambda$  acting on functions on  $\mathbb{Z}^d$  of the form

$$\mathcal{M}_\lambda = c_d \sum_{q=1}^{\infty} \sum_{\substack{1 \leq a \leq q \\ (a,q)=1}} e^{-2\pi i \lambda a/q} \mathcal{M}_\lambda^{a/q}$$

where for each reduced fraction  $a/q$  the corresponding convolution operator  $\mathcal{M}_\lambda^{a/q}$  has Fourier multiplier

$$m_\lambda^{a/q}(\xi) := \sum_{\ell \in \mathbb{Z}^k} G(a/q, \ell) \varphi_q(\xi - \ell/q) \tilde{\sigma}_\lambda(\xi - \ell/q)$$

with  $\varphi_q(\xi) = \varphi(q\xi)$  a standard smooth cut-off function,  $G(a/q, l)$  a normalized Gauss sum, and  $\tilde{\sigma}_\lambda(\xi) = \tilde{\sigma}(\lambda\xi)$  where  $\tilde{\sigma}(\xi)$  is the Fourier transform (on  $\mathbb{R}^d$ ) of the measure on the unit sphere in  $\mathbb{R}^d$  induced by Lebesgue measure and normalized to have total mass 1. By Proposition 4.1 in [10] we have

$$\left\| \sup_{\Lambda \leq \lambda \leq 2\Lambda} |\mathcal{B}_\lambda(f) - \mathcal{M}_\lambda(f)| \right\|_{\ell^2(\mathbb{Z}^d)} \leq C\Lambda^{-1/2} \|f\|_{\ell^2(\mathbb{Z}^d)}$$

provided  $d \geq 5$ . Writing

$$\mathcal{M}_*(f) := \sup_{\eta^{-2}L \leq \lambda \leq \eta^2 N} |\mathcal{M}_\lambda(f)| \quad \text{and} \quad \mathcal{M}_{*,\eta}(f) := \mathcal{M}_*(f - f * \chi_{q_\eta, L})$$

this implies

$$\|\mathcal{B}_{*,\eta}(f) - \mathcal{M}_{*,\eta}(f)\|_{\ell^2} \leq C\eta L^{-1/2} \|f - f * \chi_{q_\eta, L}\|_{\ell^2} \leq C\eta L^{-1/2} \|f\|_{\ell^2}$$

thus matters reduce to showing (19) for the operator  $\mathcal{M}_{*,\eta}$ .

For a given reduced fraction  $a/q$  we now define the maximal operator

$$\mathcal{M}_*^{a/q}(f) := \sup_{\eta^{-2}L \leq \lambda \leq \eta^2 N} |\mathcal{M}_\lambda^{a/q}(f)|$$

where  $\mathcal{M}_\lambda^{a/q}$  is the convolution operator with multiplier  $m_\lambda^{a/q}(\xi)$ . It is proved in Lemma 3.1 of [10] that

$$(20) \quad \|\mathcal{M}_*^{a/q}(f)\|_{\ell^2} \leq Cq^{-d/2} \|f\|_{\ell^2}.$$

We will show here that if  $q \leq C\eta^{-2/3}$ , then

$$(21) \quad \|\mathcal{M}_*^{a/q}(f - f * \chi_{q_\eta, L})\|_{\ell^2} \leq C\eta^{1/3} q^{-d/2} \|f\|_{\ell^2}.$$

Taking estimates (20) and (21) for granted, one obtains

$$\|\mathcal{M}_*(f - f * \chi_{q_\eta, L})\|_{\ell^2} \ll \left( \eta^{1/3} \sum_{1 \leq q \leq C\eta^{-2/3}} q^{-d/2+1} + \sum_{q \geq C\eta^{-2/3}} q^{-d/2+1} \right) \|f\|_{\ell^2} \ll \eta^{1/3} \|f\|_{\ell^2}$$

as required. It thus remains to prove (21).

Writing  $\varphi_q(\xi) = \varphi'_q(\xi)\varphi_q(\xi)$ , with a suitable smooth cut-off function  $\varphi'$ , we can introduce the decomposition

$$m_\lambda^{a/q}(\xi) = \left( \sum_{\ell \in \mathbb{Z}^k} G(a/q, \ell) \varphi'_q(\xi - \ell/q) \right) \left( \sum_{\ell \in \mathbb{Z}^k} \varphi_q(\xi - \ell/q) \tilde{\sigma}(\xi - \ell/q) \right) =: g^{a/q}(\xi) n_\lambda^q(\xi),$$

since for each  $\xi$  at most one term in each of the above sums is non-vanishing. Accordingly

$$\mathcal{M}_*^{a/q}(f - f * \chi_{q_\eta, L}) = G^{a/q} \mathcal{N}_*^q(f - f * \chi_{q_\eta, L})$$

where the maximal operator  $\mathcal{N}_*^q$  and the convolution operator  $G_{a/q}$  correspond to the multipliers  $n_\lambda^q$  and  $g^{a/q}$  respectively. Now by the standard Gauss sum estimate we have  $|g^{a/q}(\xi)| \ll q^{-d/2}$  uniformly in  $\xi$ , hence

$$\|G^{a/q} \mathcal{N}_*^q(f - f * \chi_{q_\eta, L})\|_{\ell^2} \ll q^{-d/2} \|\mathcal{N}_*^q(f - f * \chi_{q_\eta, L})\|_{\ell^2}.$$

Thus by our choice  $q_\eta := \text{lcm}\{1 \leq q \leq C\eta^{-2}\}$  it remains to show that if  $q$  divides  $q_\eta$  then

$$(22) \quad \|\mathcal{N}_*^q(f - f * \chi_{q_\eta, L})\|_{\ell^2} \ll \eta^{1/3} \|f\|_{\ell^2}.$$

As before we write  $\mathcal{N}_{*,\eta}^q(f) = \mathcal{N}_*^q(f - f * \chi_{q_\eta, L})$ , and note that this is a maximal operator with multiplier

$$n_\lambda^q(\xi)(1 - \widehat{\chi_{q_\eta, L}})(\xi) = \sum_{\ell \in \mathbb{Z}^d} \varphi_q(\xi - \ell/q)(1 - \widehat{\chi_{q_\eta, L}})(\xi - \ell/q) \tilde{\sigma}_\lambda(\xi - \ell/q).$$

For a fixed  $q$ , the multiplier  $\varphi_q(1 - \widehat{\chi_{q_\eta, L}}) \tilde{\sigma}_\lambda$  is supported on the cube  $[-\frac{1}{2q}, \frac{1}{2q}]^d$  thus by Corollary 2.1 in [10]

$$\|\mathcal{N}_{*,\eta}^q\|_{\ell^2 \rightarrow \ell^2} \leq C \|\tilde{\mathcal{N}}_{*,\eta}^q\|_{L^2 \rightarrow L^2}$$

where  $\tilde{\mathcal{N}}_{*,\eta}^q$  is the maximal operator corresponding to the multipliers  $\varphi_q(1 - \widehat{\chi_{q_\eta, L}}) \tilde{\sigma}_\lambda$ , for  $\eta^{-2}L \leq \lambda \leq \eta^2N$ , acting on  $L^2(\mathbb{R}^d)$ . By the definition of the function  $\chi_{q_\eta, L}$

$$|1 - \widehat{\chi_{q_\eta, L}}(\xi)| \ll \min\{1, L|\xi|\},$$

thus from Theorem 6.1 (with  $j = 1$ ) in [5] we obtain

$$\|\tilde{\mathcal{N}}_{*,\eta}^q\|_{L^2 \rightarrow L^2} \ll \left( \frac{L}{\eta^{-2}L} \right)^{1/6} = \eta^{1/3}$$

which establishes (22) and completes the proof.  $\square$

#### REFERENCES

- [1] M. BENNETT, A. IOSEVICH, K. TAYLOR *Finite chains inside thin subsets of  $\mathbb{R}^d$* , Anal. PDE. (2015); 9:597-614.
- [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] K. BULINSKI, *Spherical Recurrence and locally isometric embeddings of trees into positive density subsets of  $\mathbb{Z}^d$* , Math. Proc. Cambridge Philos. Soc. 165 (2018), no. 2, 267-278
- [4] H. FURSTENBERG, Y. KATZNELSON AND B. WEISS, *Ergodic theory and configurations in sets of positive density*, Israel J. Math. 54 (1986), no. 3, 307–316.
- [5] 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
- [6] A. IOSEVICH AND H. PARSHALL, *Embedding distance graphs in finite field vector spaces*, arXiv:1802.06460
- [7] N. LYALL AND Á. MAGYAR, *Optimal polynomial recurrence*, Canad. J. Math. 65 (2013), no. 1, 171-194
- [8] N. LYALL AND Á. MAGYAR, *Distance Graphs and sets of positive upper density in  $\mathbb{R}^d$* , arXiv:1803.08916
- [9] Á. MAGYAR, *On distance sets of large sets of integer points*, Israel J. Math. 164 (2008), 251–263.
- [10] Á. MAGYAR, E.M. STEIN, S. WAINGER, *Discrete analogues in harmonic analysis: spherical averages*, Annals of Math., 155 (2002), 189-208

DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF GEORGIA, ATHENS, GA 30602, USA

*E-mail address:* lyall@math.uga.edu

*E-mail address:* magyar@math.uga.edu