



ANNALES  
HENRI LEBESGUE

---

KEVIN HUGHES

---

# $\ell^p$ -IMPROVING FOR DISCRETE SPHERICAL AVERAGES

GAIN D'INTÉGRABILITÉ  $\ell^p$  POUR LES  
MOYENNES SPHÉRIQUES DISCRÈTES

---

ABSTRACT. — We initiate the theory of  $\ell^p$ -improving inequalities for arithmetic averages over hypersurfaces and their maximal functions. In particular, we prove  $\ell^p$ -improving estimates for the discrete spherical averages and some of their generalizations. As an application of our  $\ell^p$ -improving inequalities for the dyadic discrete spherical maximal function, we give a new estimate for the full discrete spherical maximal function in four dimensions. Our proofs are analogous to Littman's result on Euclidean spherical averages. One key aspect of our proof is a Littlewood–Paley decomposition in both the arithmetic and analytic aspects. In the arithmetic aspect this is a major arc-minor arc decomposition of the circle method.

RÉSUMÉ. — Nous initions la théorie sur les inégalités de gain d'intégrabilité  $\ell^p$  pour des moyennes arithmétiques sur des hypersurfaces ainsi que pour leurs fonctions maximales. En particulier, nous prouvons un gain d'intégrabilité pour les moyennes sphériques discrètes et certaines généralisations. Comme application de notre résultat, nous obtenons une nouvelle estimation pour la fonction maximale sphérique discrète en dimension 4. Nos preuves sont analogues à celle de Littman sur les moyennes sphériques Euclidiennes. Un aspect important de la preuve est une décomposition de Littlewood–Paley reposant simultanément sur des aspects arithmétiques et analytiques. L'aspect arithmétique est encodé par une méthode de décomposition du cercle en arcs mineurs et majeurs.

---

*Keywords:*  $L^p$ -improving, discrete averages, discrete maximal functions, circle method, Littlewood–Paley theory.

2020 *Mathematics Subject Classification:* 37A45, 42B25, 11P05, 11P55, 11L07.

DOI: <https://doi.org/10.5802/ahl.50>

### 1. Introduction

The motivation for this paper is Littman’s  $L^p(\mathbb{R}^d)$ -improving result for spherical averages from [Lit73]. For dimensions  $d \geq 2$  and functions  $f : \mathbb{R}^d \rightarrow \mathbb{C}$  define the spherical average (over the unit sphere) by

$$\mathcal{A}f(\mathbf{x}) := \int_{\mathbb{S}^{d-1}} f(\mathbf{x} - \mathbf{y}) \, d\sigma(\mathbf{y})$$

where  $d\sigma$  is the Euclidean surface measure on the unit sphere  $\mathbb{S}^{d-1}$  in  $\mathbb{R}^d$ .

LITTMAN. — *If  $\mathcal{A}$  is the averaging operator over the unit sphere, then*

$$(1) \quad \|\mathcal{A}f\|_{L^{d+1}(\mathbb{R}^d)} \leq C \|f\|_{L^{\frac{d+1}{d}}(\mathbb{R}^d)}.$$

In this note we will be interested in estimates for the discrete spherical averages which are analogous to (1). Suppose that  $d \geq 2$ . For  $\lambda \in \mathbb{N}$  and functions  $f : \mathbb{Z}^d \rightarrow \mathbb{C}$ , define the discrete spherical averages

$$A_\lambda f(\mathbf{x}) := N_d(\lambda)^{-1} \sum_{\mathbf{y} \in \mathbb{Z}^d: |\mathbf{y}|^2 = \lambda} f(\mathbf{x} - \mathbf{y})$$

whenever  $N_d(\lambda) := \#\{\mathbf{y} \in \mathbb{Z}^d : |\mathbf{y}|^2 = \lambda\}$  is not zero. In other words,  $A_\lambda$  is the linear operator given by convolution with the discrete (or more appropriately named “arithmetic”) probability measure

$$\sigma_\lambda := N_d(\lambda)^{-1} \mathbf{1}_{\{\mathbf{y} \in \mathbb{Z}^d: |\mathbf{y}|^2 = \lambda\}}.$$

Our main result is the following.

THEOREM 1.1. — *If  $d \geq 4$  and  $\frac{d+1}{d-1} \leq p \leq 2$ , then for each  $\epsilon > 0$ , there exists constants  $C_{p,\epsilon}$  depending on  $p$  and  $\epsilon$  such that for all  $\lambda \in \mathbb{N}$  (further restrict  $\lambda$  to be odd when  $d = 4$ ), we have the  $\ell^p$ -improving inequality*

$$(1.1) \quad \|A_\lambda f\|_{\ell^{p'}(\mathbb{Z}^d)} \leq C_{p,\epsilon} \lambda^{\epsilon - \frac{d}{2}(\frac{2}{p} - 1)} \|f\|_{\ell^p(\mathbb{Z}^d)}.$$

The implicit constants are independent of  $\lambda$ .

We are also motivated by Lee’s work [Lee03] which proved that the dyadic spherical maximal function variant of Littman’s theorem holds. Moreover, our methods are flexible and we can use them to strengthen Theorem 1.1 when  $p$  is sufficiently large.

THEOREM 1.2. — *If  $d \geq 5$  and  $p := \frac{d}{d-2}$ , then there exists constants  $C_p$  depending on  $p$  such that for all  $\Lambda \in \mathbb{N}$ , we have the  $\ell^p$ -improving inequality for the  $\Lambda$ -dyadic discrete spherical maximal function*

$$(1.2) \quad \left\| \sup_{\Lambda \leq \lambda < 2\Lambda} |A_\lambda f| \right\|_{\ell^{p',\infty}(\mathbb{Z}^d)} \leq C_p \Lambda^{-\frac{d}{2}(\frac{2}{p} - 1)} \|f\|_{\ell^{p,1}(\mathbb{Z}^d)}.$$

The constant  $C_p$  is independent of  $\Lambda$ .

This clearly implies the same is true for a single average and hence we may remove the  $\epsilon$ -loss in Theorem 1.1 when  $p > \frac{d}{d-2}$ .

### 1.1. Motivation

When  $d \geq 5$  and  $\lambda \in \mathbb{N}$ , we have that  $100^{-1}\lambda^{\frac{d-2}{2}} \leq N_d(\lambda) \leq 100\lambda^{\frac{d-2}{2}}$ ; when  $d = 4$  and  $\lambda$  is restricted to be odd, we have that  $100^{-1}\lambda^{\frac{d-2}{2}} / \log \log \lambda \leq N_d(\lambda) \leq 100\lambda^{\frac{d-2}{2}} \cdot \log \log \lambda$ . Several years ago Jim Wright asked the author: *What is  $\ell^p$ -improving for the discrete spherical averages?* We interpret his question as the following.

QUESTION 1.3 (Jim Wright). — *When are there exponents  $1 \leq p, q \leq \infty$  and a constant  $C = C_{d,p,q}$ , possibly depending on  $d, p, q$  but independent of  $\lambda$  such that*

$$(1.3) \quad \|A_\lambda f\|_{\ell^q(\mathbb{Z}^d)} \leq C \|f\|_{\ell^p(\mathbb{Z}^d)}?$$

On the one hand, unlike the continuous case in  $\mathbb{R}^d$ , we do not have the dilational symmetry to exploit; this is why we want (1.3) to hold uniformly for  $\lambda \in \mathbb{N}$  instead of formulating the question for the unit sphere ( $\lambda = 1$ ) as stated in Littman’s result. On the other hand, we may quickly obtain some trivial off-diagonal results by using the contraction inequality

$$(1.4) \quad \|A_\lambda f\|_{\ell^p(\mathbb{Z}^d)} \leq \|f\|_{\ell^p(\mathbb{Z}^d)} \quad \text{for all } 1 \leq p \leq \infty$$

and the nesting property of  $\ell^p$ -spaces

$$(1.5) \quad \|f\|_{\ell^q} \leq \|f\|_{\ell^p} \quad \text{for all } 1 \leq p \leq q \leq \infty,$$

to see that (1.3) is true for all  $\lambda$  for  $1 \leq p \leq q \leq \infty$  with  $C = 1$ . This makes Question 1.3 trivial.

We consider the following two examples 1.4 and 1.5; assume for simplicity that  $d \geq 5$  and  $\lambda \in \mathbb{N}$  is large.

*Example 1.4 ( $\delta$ -function).* — Take  $f$  to be the delta function at the origin. Then  $A_\lambda f$  is supported on the sphere of radius  $\sqrt{\lambda}$  and has height  $N_d(\lambda)^{-1} \approx \lambda^{\frac{2-d}{2}}$ . Thus, for  $p, q \geq 1$ ,

$$\begin{aligned} \|A_\lambda f\|_{\ell^q(\mathbb{Z}^d)} &= \|A_\lambda f\|_{\ell^\infty(\mathbb{Z}^d)} N_d(\lambda)^{\frac{1}{q}} = N_d(\lambda)^{\frac{1}{q}-1} \\ &= N_d(\lambda)^{\frac{1}{q}-1} \|f\|_{\ell^p(\mathbb{Z}^d)} \approx \lambda^{-\left(\frac{d-2}{2}\right)\left(1-\frac{1}{q}\right)} \|f\|_{\ell^p(\mathbb{Z}^d)}. \end{aligned}$$

We have a similar bound for the dyadic maximal function

$$\left\| \sup_{\Lambda \leq \lambda < 2\Lambda} |A_\lambda f| \right\|_{\ell^q(\mathbb{Z}^d)}^q = \sum_{\Lambda \leq \lambda < 2\Lambda} \sum_{\mathbf{x} \in \mathbb{Z}^d: |\mathbf{x}|=\lambda} |A_\lambda f(\mathbf{x})|^q.$$

Consequently,

$$\left\| \sup_{\Lambda \leq \lambda < 2\Lambda} |A_\lambda f| \right\|_{\ell^q(\mathbb{Z}^d)}^q \lesssim \Lambda \cdot \Lambda^{\frac{d-2}{2}} \cdot \Lambda^{-q\left(\frac{d-2}{2}\right)} = \Lambda^{1+(1-q)\left(\frac{d-2}{2}\right)}.$$

Therefore, for  $p, q \geq 1$ , we have

$$\left\| \sup_{\Lambda \leq \lambda < 2\Lambda} |A_\lambda f| \right\|_{\ell^q(\mathbb{Z}^d)} \lesssim \Lambda^{\frac{1}{q}+\left(\frac{1}{q}-1\right)\left(\frac{d-2}{2}\right)} = \Lambda^{\frac{1}{q}+\left(\frac{1}{q}-1\right)\left(\frac{d-2}{2}\right)} \|f\|_{\ell^p(\mathbb{Z}^d)}.$$

*Example 1.5 (big ball-function).* — Take  $f$  to be the indicator function of a ball of radius  $R \approx \sqrt{\lambda}$ . Then  $A_\lambda f$  is supported on the ball of radius  $R + \sqrt{\lambda} \approx \sqrt{\lambda}$  and has height  $\gtrsim 1$  for a large chunk of it. Therefore,

$$\|A_\lambda f\|_{\ell^q(\mathbb{Z}^d)} \approx \|f\|_{\ell^q(\mathbb{Z}^d)} \approx R^{d/q}$$

for  $q \geq 1$  while

$$\|f\|_{\ell^p(\mathbb{Z}^d)} \approx R^{d/p}.$$

The same estimates hold with the dyadic maximal function  $\sup_{\Lambda \leq \lambda < 2\Lambda} |A_\lambda f|$  in place of  $A_\lambda f$ . Here we immediately see that we must have  $q \geq p$  to satisfy (1.3). Combining these two we see that for  $1 \leq p \leq 2$ :

$$\left\| \sup_{\Lambda \leq \lambda < 2\Lambda} |A_\lambda f| \right\|_{\ell^{p'}(\mathbb{Z}^d)} \approx \Lambda^{-\frac{d}{2}(\frac{2}{p}-1)} \|f\|_{\ell^p(\mathbb{Z}^d)}.$$

Here and throughout,  $p'$  is the dual exponent to  $p$  which means for  $1 \leq p \leq \infty$  that  $\frac{1}{p} + \frac{1}{p'} = 1$ .

Example 1.4 actually reveals more since Young’s inequality implies

$$\|A_\lambda f\|_{\ell^\infty(\mathbb{Z}^d)} = \|\sigma_\lambda * f\|_{\ell^\infty(\mathbb{Z}^d)} \leq \|\sigma_\lambda\|_{\ell^\infty(\mathbb{Z}^d)} \|f\|_{\ell^1(\mathbb{Z}^d)} \leq 100\lambda^{-\frac{d-2}{2}} \|f\|_{\ell^1(\mathbb{Z}^d)}.$$

This bound extends to the dyadic maximal functions:

$$\left\| \sup_{\Lambda \leq \lambda < 2\Lambda} |A_\lambda f| \right\|_{\ell^\infty(\mathbb{Z}^d)} \leq 100\Lambda^{-\frac{d-2}{2}} \|f\|_{\ell^1(\mathbb{Z}^d)}.$$

Interpolating this with the contraction inequality we obtain the following estimate

$$(1.6) \quad \left\| \sup_{\Lambda \leq \lambda < 2\Lambda} |A_\lambda f| \right\|_{\ell^{p'}(\mathbb{Z}^d)} \leq 100\Lambda^{-\frac{d-2}{2}(\frac{2}{p}-1)} \|f\|_{\ell^p(\mathbb{Z}^d)} \quad \text{for } 1 \leq p \leq 2,$$

which we call the *trivial bound*. For large  $\lambda$  the trivial bound is larger than the bounds from our examples. Curiously, the trivial bound shows that there exist  $\ell^p(\mathbb{Z}^d)$ -improving estimates which decay with  $\lambda$ . Therefore a better interpretation of Jim Wright’s Question 1.3 is the following Question 1.6.

QUESTION 1.6. — For each  $1 < p < 2$ , what is the best exponent  $\eta_p$  so that

$$\|A_\lambda f\|_{\ell^{p',\infty}(\mathbb{Z}^d)} \leq C_p \lambda^{-\eta_p} \|f\|_{\ell^p(\mathbb{Z}^d)}$$

where the constant  $C_p$  is independent of  $\lambda \in \mathbb{N}$  and  $f \in \ell^p(\mathbb{Z}^d)$ ?

We also ask this question for the dyadic version.

QUESTION 1.7. — For each  $1 < p < 2$ , what is the best exponent  $\nu_p$  so that

$$\left\| \sup_{\Lambda \leq \lambda < 2\Lambda} |A_\lambda f| \right\|_{\ell^{p',\infty}(\mathbb{Z}^d)} \leq C'_p \Lambda^{-\nu_p} \|f\|_{\ell^p(\mathbb{Z}^d)}$$

where the constant  $C'_p$  is independent of  $\Lambda \in \mathbb{N}$  and  $f \in \ell^p(\mathbb{Z}^d)$ ?

Theorem 1.1 addresses Question 1.6 when we are close to  $p = 2$ . In particular it says that in Question 1.6 we may take any  $0 \leq \eta_p < \frac{d}{2}(\frac{2}{p} - 1)$  for  $\frac{d+1}{d-1} \leq p < 2$ . Comparing (1.1) to (1.6), we see that Theorem 1.1 improves upon the trivial bound. Recall that the trivial bound was given by a simple convexity estimate (in this case

applying Young’s inequality); consequently, Theorem 1.1 might be referred to as a *subconvexity estimate* in the argot of analytic number theorists.

One can also ask for  $\ell^p$ -improving estimates for the full discrete spherical maximal function. For  $f : \mathbb{Z}^d \rightarrow \mathbb{C}$ , let

$$A_*f(\mathbf{x}) := \sup_{\lambda \in \mathbb{N}: \lambda \text{ is odd}} |A_\lambda f(\mathbf{x})|$$

denote the discrete spherical maximal function when  $d = 4$ , and

$$A_*f(\mathbf{x}) := \sup_{\lambda \in \mathbb{N}} |A_\lambda f(\mathbf{x})|$$

denote the discrete spherical maximal function when  $d \geq 5$ .

QUESTION 1.8. — *When are there exponents  $1 \leq p, q \leq \infty$  so that*

$$(1.7) \quad \|A_*f\|_{\ell^q(\mathbb{Z}^d)} \lesssim \|f\|_{\ell^p(\mathbb{Z}^d)}?$$

When  $d \geq 5$ , there are some obvious ranges for which one may obtain bounds. In particular, by applying Magyar–Stein–Wainger’s theorem on the discrete spherical maximal function, see [MSW02, “Theorem”], we have that (1.7) holds for  $d \geq 5$  and  $\frac{d}{d-2} < q \leq \infty$  because

$$\|A_*f\|_{\ell^q} \lesssim \|f\|_{\ell^q}$$

for all  $q > \frac{d}{d-2}$  and the nesting property of  $\ell^p$ -spaces implies that

$$\|A_*f\|_{\ell^q} \lesssim \|f\|_{\ell^q} \leq \|f\|_{\ell^p}$$

for  $1 \leq p \leq q$ . Moreover this range is sharp; this can be seen by considering the delta-function example. In four dimensions, the analogous Magyar–Stein–Wainger estimate is expected to hold.

CONJECTURE 1.9. — *If  $p > 2$ , then*

$$(1.8) \quad \|A_*f\|_{\ell^p(\mathbb{Z}^4)} \lesssim_p \|f\|_{\ell^p(\mathbb{Z}^4)}.$$

Interestingly we can prove the following  $\ell^p$ -improving result when  $d = 4$  which would be a corollary of Conjecture 1 by the nesting properties of  $\ell^p$ -spaces.

THEOREM 1.10. — *Assume that  $d = 4$ . If  $q > 2$  and  $1 \leq p < q$ , then*

$$(1.9) \quad \|A_*f\|_{\ell^q(\mathbb{Z}^4)} \lesssim_{p,q} \|f\|_{\ell^p(\mathbb{Z}^4)}.$$

### 1.2. Comparison to recent works

Our examples show that (1.1) fails for  $p < \frac{d+2}{d}$ . So, one might expect that (1.1) would hold for all  $\frac{d+2}{d} \leq p \leq 2$  which would go beyond Theorem 1.1. Intriguingly, Kesler–Lacey [KL18] claimed that (1.1) fails for  $p < \frac{d+1}{d-1}$ . Moreover [KL18] removed the  $\epsilon$ -loss in (1.1) for  $\frac{d+1}{d-1} < p < \frac{d}{d-2}$ .

We encourage the reader to read Kesler–Lacey’s interesting work [KL18] which appeared independently of this paper. Kesler–Lacey also considered  $\ell^p(\mathbb{Z}^d)$ -improving inequalities for discrete spherical averages. In [KL18] their focus lied upon using  $\ell^p$ -improving inequalities to deduce “sparse bounds” for the discrete spherical maximal

function. The question of sparse bounds is not considered here. However, we prove  $\ell^p$ -improving inequalities for a broader class of averages which are not considered in [KL18].

As the reader may compare, the ideas in [KL18] (for proving  $\ell^p(\mathbb{Z}^d)$ -improving inequalities for discrete spherical averages) are very similar to those here. Their methods may appear more complicated due to the use of the Ramanujan bound from [Bou85, Hug19]. We point out that the Ramanujan bound first appeared in Bourgain's work on restriction of the parabola to the integer lattice to remove an  $\epsilon$ -loss there, and the Ramanujan bound was first adapted to the context of spherical averages in my work on discrete spherical maximal functions over sparse subsequences of radii; see [Hug19].

### 1.3. Overview of the proofs

For many discrete analogues in harmonic analysis not only are the statements of theorems analogous, but there is also an analogy between their proofs. We take a moment to describe this since this analogy does not appear to be explained in the literature.

For many problems in Fourier analysis such as Littman's theorem, one decomposes an operator into "low" and "high" frequencies and obtains bounds for these different pieces. For instance the Littlewood–Paley square function is one way to do this. There is a similar decomposition for analogous problems with an arithmetic flavor.

For problems over  $\mathbb{Z}^d$  instead of  $\mathbb{R}^d$ , all frequencies in the torus  $(\mathbb{R}/\mathbb{Z})^d$  are ostensibly low frequencies since the torus is compact. Unfortunately, this perspective is insufficiently nuanced to treat many problems. Instead one should recalibrate to the following: replace the sobriquets "low frequencies" and "high frequencies" of Fourier analysis with "major arcs" and "minor arcs" of the circle method respectively.

Recalibrating one's perspective to the above analogy allows us to import intuitions and paradigms from continuous Euclidean harmonic analysis to discrete Euclidean harmonic analysis via the circle method. By way of this analogy one sees that the circle method is akin to Littlewood–Paley theory. I encourage the reader to review the proofs of Littman's theorem before reading the proofs here to see this analogy in action.

### 1.4. Organization of the paper

The paper is organized as follows. Section 2 sets some notation used throughout the paper. Section 3 generalizes Theorem 1 to hypersurfaces defined by nice, positive definite homogeneous forms with integral coefficients. We prove Theorem 1.1 in Section 4 by making use of improved estimates for Kloosterman sums. Section 5 proves bounds for dyadic discrete  $k$ -spherical maximal functions. In Section 6 we prove Theorem 1.10. Section 7 concludes with a few questions. Finally, in the Appendix we record the range of  $\ell^p$ -boundedness of Magyar's maximal functions (which arise in Section 3) from [Mag02].

## 2. Notation

We introduce here some notation that will streamline our exposition.

- We write  $f(\lambda) \lesssim g(\lambda)$  if there exists a constant  $C > 0$  independent of all  $\lambda$  under consideration (e.g.  $\lambda$  in  $\mathbb{N}$  or in  $\Gamma_{\mathcal{Q}}$ ) such that

$$|f(\lambda)| \leq C|g(\lambda)|.$$

Furthermore, we will write  $f(\lambda) \gtrsim g(\lambda)$  if  $g(\lambda) \lesssim f(\lambda)$  while we will write  $f(\lambda) \approx g(\lambda)$  if  $f(\lambda) \lesssim g(\lambda)$  and  $f(\lambda) \gtrsim g(\lambda)$ .

- Subscripts in the above notations will denote parameters, such as the dimension  $d$  or degree  $k$  of a form  $\mathcal{Q}$ , on which the implicit constants may depend.
- $\mathbb{T}^d$  denotes the  $d$ -dimensional torus  $(\mathbb{R}/\mathbb{Z})^d$  identified with the unit cube  $[-1/2, 1/2]^d$ .
- $*$  denotes convolution on a group such as  $\mathbb{Z}^d$ ,  $\mathbb{T}^d$  or  $\mathbb{R}^d$ . It will be clear from context as to which group the convolution takes place.
- $e(t)$  will denote the character  $e^{-2\pi it}$  for  $t \in \mathbb{R}$  or  $\mathbb{T}$ .
- For a function  $f : \mathbb{Z}^d \rightarrow \mathbb{C}$ , its  $\mathbb{Z}^d$ -Fourier transform will be denoted  $\widehat{f}(\xi)$  for  $\xi \in \mathbb{T}^d$ . For a function  $f : \mathbb{R}^d \rightarrow \mathbb{C}$ , its  $\mathbb{R}^d$ -Fourier transform will be denoted  $\widetilde{f}(\xi)$  for  $\xi \in \mathbb{R}^d$ .
- For a function  $f : \mathbb{R}^d \rightarrow \mathbb{C}$ , we define dilation operator  $D_t$  by  $D_t f(\mathbf{x}) = f(\mathbf{x}/t)$ .
- For a ring  $R$ , we will use the inner product notation  $\mathbf{b} \cdot \mathbf{m}$  for vectors  $\mathbf{b}, \mathbf{m} \in R^d$  to mean the sum  $\sum_{i=1}^d b_i m_i$ . This is used for the rings  $\mathbb{R}, \mathbb{Z}, \mathbb{T}$  and  $\mathbb{Z}/q$  where  $q \in \mathbb{Z}$ .
- We also let  $\mathbf{1}_X$  denote the indicator function of the set  $X$ .

## 3. $\ell^p$ -improving for Magyar’s theorem

Our method is quite general, so we start by generalizing Theorem 1.2. Throughout this section  $\mathcal{Q}(\mathbf{x}) \in \mathbb{Z}[x_1, x_2, \dots, x_d]$ , where  $\mathbf{x} = (x_1, x_2, \dots, x_d)$ , will denote an integral, positive definite, homogeneous form, and  $k$  will denote the degree of the form  $\mathcal{Q}$ . We assume that  $k \geq 2$  and a natural number. Let  $V_{\mathcal{Q}}(\mathbb{C}) := \{\mathbf{x} \in \mathbb{C}^d : \nabla \mathcal{Q}(\mathbf{x}) = \mathbf{0}\}$  denote the (Birch) *singular locus* of the form  $\mathcal{Q}$ . We will say that a homogeneous, integral form is *non-singular* if it satisfies Birch’s criterion:

$$(3.1) \quad d - \dim_{\mathbb{C}}(V_{\mathcal{Q}}(\mathbb{C})) > (k - 1)2^k.$$

The notion of dimension “ $\dim_{\mathbb{C}}(V_{\mathcal{Q}}(\mathbb{C}))$ ” can be taken to be the algebraic dimension of the complex variety  $V_{\mathcal{Q}}(\mathbb{C})$ .

When (3.1) is satisfied, Birch [Bir62] tells us that there exists a positive constant  $C_{\mathcal{Q}}$  and an infinite arithmetic progression  $\Gamma_{\mathcal{Q}}$  in  $\mathbb{N}$  depending on the form  $\mathcal{Q}$  so that

$$N_{\mathcal{Q}}(\lambda) := \#\{\mathbf{n} \in \mathbb{Z}^d : \mathcal{Q}(\mathbf{n}) = \lambda\} \geq C_{\mathcal{Q}} \lambda^{\frac{d}{k}-1} > 0 \quad \text{for all } \lambda \in \Gamma_{\mathcal{Q}}.$$

Following Magyar [Mag02], we will call any such arithmetic progression  $\Gamma_{\mathcal{Q}}$  a *set of regular values for  $\mathcal{Q}$* . For each  $\lambda \in \mathbb{N}$ ,  $N_{\mathcal{Q}}(\lambda)$  is finite because the hypersurface

$\{\mathbf{n} \in \mathbb{R}^d : \mathcal{Q}(\mathbf{n}) = \lambda\}$  is defined by a positive definite form which implies that this hypersurface is compact. Consequently, the averages

$$A_\lambda^\mathcal{Q}(\mathbf{x}) := N_\mathcal{Q}(\lambda)^{-1} \sum_{\mathbf{n} \in \mathbb{Z}^d: \mathcal{Q}(\mathbf{n})=\lambda} f(\mathbf{x} - \mathbf{n})$$

make sense for all  $\lambda \in \Gamma_\mathcal{Q}$  and functions  $f : \mathbb{Z}^d \rightarrow \mathbb{C}$ . In this setting, our trivial bound (1.6) becomes

$$(3.2) \quad \left\| \sup_{\Lambda \leq \lambda < 2\Lambda} |A_\lambda^\mathcal{Q} f| \right\|_{\ell^{p'}(\mathbb{Z}^d)} \lesssim \Lambda^{-\left(\frac{d}{k}-1\right)\left(\frac{2}{p}-1\right)} \|f\|_{\ell^p(\mathbb{Z}^d)} \quad \text{for } 1 \leq p \leq 2 \text{ and } \Lambda \geq 1.$$

This follows from Young’s inequality and Birch’s estimate for  $N_\mathcal{Q}(\lambda)$ . In all of our dyadic maximal functions we will restrict to  $\lambda \in \Gamma_\mathcal{Q}$ .

For a non-singular, homogeneous, integral form define the parameters

$$(3.3) \quad \gamma_\mathcal{Q} := \frac{1}{6k} \left( \frac{d - \dim(V_\mathcal{Q}(\mathbb{C}))}{(k-1)2^k} - 1 \right) \quad \text{and} \quad \kappa_\mathcal{Q} := \frac{d - \dim V_\mathcal{Q}(\mathbb{C})}{2^{k-1}(k-1)}.$$

Throughout we assume that  $d > k \geq 2$  with  $d$  sufficiently large with respect to  $k$  to satisfy the Birch–Magyar non-singularity criterion (3.1) so that  $\gamma_\mathcal{Q} > 0$  and  $\kappa_\mathcal{Q} > 2$ . The following result gives an improvement over the trivial bound (3.2) when  $p$  is close to 2.

**THEOREM 3.1.** — *Let  $\mathcal{Q}$  be a positive definite, non-singular, homogeneous, integral form in  $d$  variables of degree  $k$  and  $\Gamma_\mathcal{Q}$  a set of regular values for  $\mathcal{Q}$ . If  $p := \frac{\kappa}{\kappa-1}$ , then for each  $\epsilon > 0$  there exists a constant  $C_{\epsilon, \mathcal{Q}, p}$  independent of  $\Lambda \geq 1$  so that*

$$(3.4) \quad \left\| \sup_{\Lambda \leq \lambda < 2\Lambda: \lambda \in \Gamma_\mathcal{Q}} |A_\lambda^\mathcal{Q} f| \right\|_{\ell^{p'}(\mathbb{Z}^d)} \leq C_{\mathcal{Q}, p, \epsilon} \Lambda^{\epsilon - \frac{d}{k}\left(\frac{2}{p}-1\right)} \left( 1 + \Lambda^{\frac{2}{p}-1-\gamma_\mathcal{Q}\left(2-\frac{2}{p}\right)} \right) \|f\|_{\ell^p(\mathbb{Z}^d)}$$

for all  $\Lambda \geq 1$ . Recall that the supremum is restricted so that each  $\lambda$  is in  $\Gamma_\mathcal{Q}$ .

Note that our assumption that  $\kappa > 2$  implies that  $\kappa/(\kappa - 1) < 2$ . Our theorem implies the following Corollary 3.2 which says that our dyadic maximal functions satisfy essentially sharp  $\ell^p(\mathbb{Z}^d) \rightarrow \ell^{p'}(\mathbb{Z}^d)$ -improving estimates for  $p \leq 2$  with  $p$  sufficiently close to 2.

**COROLLARY 3.2.** — *Let  $\mathcal{Q}$  be a positive definite, non-singular, homogeneous, integral form in  $d$  variables of degree  $k$  and  $\Gamma_\mathcal{Q}$  a set of regular values for  $\mathcal{Q}$ . If  $\frac{2(1+\gamma_\mathcal{Q})}{1+2\gamma_\mathcal{Q}} \leq p \leq 2$ , then for each  $\epsilon > 0$  there exists a constant  $C_{\epsilon, \mathcal{Q}, p}$  independent of  $\Lambda \geq 1$  so that*

$$\left\| \sup_{\Lambda \leq \lambda < 2\Lambda: \lambda \in \Gamma_\mathcal{Q}} |A_\lambda^\mathcal{Q} f| \right\|_{\ell^{p'}(\mathbb{Z}^d)} \leq C_{\mathcal{Q}, p, \epsilon} \Lambda^{\epsilon - \frac{d}{k}\left(\frac{2}{p}-1\right)} \|f\|_{\ell^p(\mathbb{Z}^d)}$$

for all  $\Lambda \geq 1$ .

Our corollary follows by determining when  $\frac{2}{p} - 1 - \gamma_\mathcal{Q} \left( 2 - \frac{2}{p} \right) \leq 0$  and noting that  $\gamma_\mathcal{Q} < \kappa_\mathcal{Q}$ . This is a computation that we leave to the reader.

The heavy lifting in our theorem lies in a decomposition of Magyar for the averages  $A_\lambda^\mathcal{Q}$ .



MAGYAR 1 ([Mag07]). — Let  $\mathcal{Q}(\mathbf{x}) \in \mathbb{Z}[\mathbf{x}]$  be a positive definite, non-singular, integral, homogeneous form, and  $\Gamma_{\mathcal{Q}}$  be a set of regular values for the form  $\mathcal{Q}$ . For each  $\lambda \in \Gamma_{\mathcal{Q}}$  the averaging operator  $A_{\lambda}^{\mathcal{Q}}$  decomposes into the sum of two convolution operators,  $A_{\lambda} = M_{\lambda}^{\mathcal{Q}} + E_{\lambda}^{\mathcal{Q}}$  such that for all  $\Lambda \geq 1$ , we have

$$(3.5) \quad \left\| \sup_{\Lambda \leq \lambda < 2\Lambda: \lambda \in \Gamma_{\mathcal{Q}}} |E_{\lambda}| \right\|_{\ell^2(\mathbb{Z}^d) \rightarrow \ell^2(\mathbb{Z}^d)} \lesssim_{\mathcal{Q}, \gamma} \Lambda^{\epsilon - \gamma \mathcal{Q}} \quad \text{for all } \epsilon > 0$$

The implicit constant is independent of  $\Lambda \geq 1$ . The main term,  $M_{\lambda}^{\mathcal{Q}}$ , is the sum of finitely many convolution operators,

$$M_{\lambda}^{\mathcal{Q}} = \sum_{q=1}^{\Lambda^{1/k}} \sum_{a \in (\mathbb{Z}/q)^*} e(a\lambda/q) M_{\lambda}^{\mathcal{Q}, a/q},$$

where  $\Lambda = 2^j$  satisfies  $\Lambda/2 \leq \lambda < \Lambda$ . Moreover,  $M_{\lambda}^{\mathcal{Q}}$  satisfies the estimate

$$(3.6) \quad \left\| \sup_{\lambda \in \Gamma_{\mathcal{Q}}} |M_{\lambda}^{\mathcal{Q}, a/q}| \right\|_{\ell^2(\mathbb{Z}^d) \rightarrow \ell^2(\mathbb{Z}^d)} \lesssim_{\epsilon} q^{\epsilon - \kappa \mathcal{Q}} \quad \text{for all } \epsilon > 0.$$

We remark that  $M_{\lambda}^{\mathcal{Q}, a/q}$  is the convolution operator corresponding to the Fourier multiplier

$$\widehat{M_{\lambda}^{\mathcal{Q}, a/q}}(\xi) := \sum_{\mathbf{m} \in \mathbb{Z}^d} G_{\mathcal{Q}}(a, q; \mathbf{m}) \Psi(q\xi - \mathbf{m}) \widetilde{d\sigma_{\mathcal{Q}}} \left( \lambda^{1/k} \left[ \xi - \frac{\mathbf{m}}{q} \right] \right).$$

For  $\mathbf{m} \in \mathbb{Z}^d$

$$G_{\mathcal{Q}}(a, q; \mathbf{m}) := q^{-d} \sum_{\mathbf{b} \in (\mathbb{Z}/q)^d} e \left( \frac{a\mathcal{Q}(\mathbf{b}) + \mathbf{b} \cdot \mathbf{m}}{q} \right)$$

is the normalized Weyl sum satisfying the bound

$$(3.7) \quad |G_{\mathcal{Q}}(a, q; \mathbf{m})| \lesssim_{\epsilon} q^{\epsilon - \kappa} \quad \text{for all } \epsilon > 0$$

uniformly in  $a \in (\mathbb{Z}/q)^*$  and  $\mathbf{m} \in \mathbb{Z}^d$ . The function  $\Psi$  is a  $C^{\infty}(\mathbb{R}^d)$  bump function supported in the cube  $[-1/4, 1/4]^d$  and 1 on the cube  $[-1/8, 1/8]^d$ , and the singular measure  $d\sigma_{\mathcal{Q}}$  is the Gelfand–Leray form defined distributionally by the oscillatory integral

$$\int_{\mathbb{R}} e(t(\mathcal{Q}(\mathbf{x}) - 1)) dt.$$

Alternatively,  $d\sigma_{\mathcal{Q}}(\mathbf{x}) = dS_{\mathcal{Q}}(\mathbf{x})/|\nabla \mathcal{Q}(\mathbf{x})|$  where  $dS_{\mathcal{Q}}$  is the Euclidean surface area measure on the hypersurface  $\{\mathbf{x} \in \mathbb{R}^d : \mathcal{Q}(\mathbf{x}) = 1\}$  which is compactly supported since  $\mathcal{Q}$  is positive definite. See [Mag02] for more information concerning Gelfand–Leray measures of hypersurfaces. We cite the following bound (see [Mag02, Lemma 6 p. 931]) for the  $\mathbb{R}^d$ -Fourier transform of the surface measure:

$$(3.8) \quad \left| \widetilde{d\sigma_{\mathcal{Q}}}(\xi) \right| \lesssim_{\epsilon} (1 + |\xi|)^{1 - \kappa + \epsilon} \quad \text{for each } \xi \in \mathbb{R}^d \quad \text{and for all } \epsilon > 0.$$

The estimate (3.5) does not explicitly appear in [Mag02]. Instead it appears for a slightly different definition of our error term, so we briefly indicate how one obtains it. Estimate (3.5) is encoded in the proofs of [Mag02, Propositions 3 and 4]. One key difference in this paper is that our main term  $M_{\lambda}^{\mathcal{Q}}$  is a finite sum depending on  $\lambda$ , and so we do not need [Mag02, (2.17), Proposition 4]. Meanwhile the estimates

(2.15) and (2.16) of [Mag02, Proposition 4] are superior to the minor arc estimate of [Mag02, Proposition 3]. Therefore the minimal exponent which defines  $\gamma_{\mathcal{Q}}$  comes from the minor arc estimate.

Magyar’s theorem gives us  $\ell^2$ -estimates, but we are interested in  $\ell^p \rightarrow \ell^{p'}$ -estimates. We will interpolate Magyar’s  $\ell^2$ -estimates with appropriate  $\ell^1 \rightarrow \ell^\infty$ -estimates to deduce the following Lemmas 3.3 and 3.4 which when added together immediately yield Theorem 3.1. Our lemma for the main term is the following.

LEMMA 3.3. — *Let  $\mathcal{Q}(\mathbf{x}) \in \mathbb{Z}[\mathbf{x}]$  be a positive definite, non-singular, integral, homogeneous form satisfying (3.1), and  $\Gamma_{\mathcal{Q}}$  be a set of regular values for the form  $\mathcal{Q}$ . If  $p := \frac{\kappa}{\kappa-1}$  (which is less than 2 by our assumption on  $\kappa$ ), then*

$$(3.9) \quad \left\| \sup_{\Lambda \leq \lambda < 2\Lambda} |M_\lambda| \right\|_{\ell^{p,1}(\mathbb{Z}^d) \rightarrow \ell^{p',\infty}(\mathbb{Z}^d)} \lesssim \Lambda^{\epsilon - \frac{d}{k}(\frac{2}{p}-1)} \quad \text{for all } \Lambda \geq 1.$$

Up to a factor of  $\Lambda^\epsilon$  the bound for our main term is the size we expect for our averages in the range  $\frac{\kappa}{\kappa-1} \leq p \leq 2$ . Unfortunately the bound for the error term is much worse.

LEMMA 3.4. — *Let  $\mathcal{Q}(\mathbf{x}) \in \mathbb{Z}[\mathbf{x}]$  be a positive definite, non-singular, integral, homogeneous form satisfying (3.1), and  $\Gamma_{\mathcal{Q}}$  be a set of regular values for the form  $\mathcal{Q}$ . If  $1 \leq p \leq 2$ , then for all  $\epsilon > 0$ ,*

$$(3.10) \quad \left\| \sup_{\Lambda \leq \lambda < 2\Lambda: \lambda \in \Gamma_{\mathcal{Q}}} |E_\lambda| \right\|_{\ell^p(\mathbb{Z}^d) \rightarrow \ell^{p'}(\mathbb{Z}^d)} \lesssim_\epsilon \Lambda^{(1-\frac{d}{k})(\frac{2}{p}-1) - \gamma_{\mathcal{Q}}(2-\frac{2}{p}) + \epsilon} \quad \text{for all } \Lambda \geq 1.$$

The proofs of these lemmas are motivated by proofs of Littman’s theorem and its variants. Proofs of Littman’s theorem often proceed by frequency decomposing the spherical average into pieces and finding  $L^1 \rightarrow L^\infty$  and  $L^2 \rightarrow L^2$  estimates with which to interpolate. In particular, we use a restricted weak-type argument that was used by Bourgain for the Euclidean spherical maximal function, by Ionescu [Ion04] for the discrete spherical maximal function and also by Hu–Li for discrete restriction to the sphere in [HL14]. This is the same strategy that we follow in our proofs of these lemmas. In the next Subsections 3.2 and 3.1, we will deduce our lemmas.

### 3.1. Proof of Lemma 3.3

Let  $K_\lambda^{a/q}$  denote the kernel (with domain  $\mathbb{Z}^d$ ) associated to the convolution operator  $M_\lambda^{\mathcal{Q},a/q}$ . We start our proof by establishing an identity for these kernels.

PROPOSITION 3.5. — *Let  $\mathcal{Q}(\mathbf{x}) \in \mathbb{Z}[\mathbf{x}]$  be a positive definite, non-singular, integral, homogeneous form satisfying (3.1), and  $\Gamma_{\mathcal{Q}}$  be a set of regular values for the form  $\mathcal{Q}$ . If  $1 \leq a < q < \infty$  with  $(a, q) = 1$ , then*

$$(3.11) \quad K_\lambda^{a/q}(\mathbf{x}) = e(a\mathcal{Q}(\mathbf{x})/q) \lambda^{-d/k} \widetilde{D_{q\lambda^{-1/k}} \Psi} * d\sigma_{\mathcal{Q}}(\lambda^{-1/k}\mathbf{x}) \quad \text{for } \mathbf{x} \in \mathbb{Z}^d.$$

*Proof.* — Fix a form  $\mathcal{Q}$  satisfying the hypotheses of the proposition. We drop the dependence on  $\mathcal{Q}$  in our notation in order to simplify it.

By Fourier inversion, our kernel is

$$\begin{aligned} K_\lambda^{a/q}(\mathbf{x}) &= \int_{\mathbb{T}^d} e(-\mathbf{x} \cdot \xi) \sum_{\mathbf{m} \in \mathbb{Z}^d} G(a, q; \mathbf{m}) \Psi(q\xi - \mathbf{m}) \widetilde{d\sigma} \left( \lambda^{1/k} \left[ \xi - \frac{\mathbf{m}}{q} \right] \right) d\xi \\ &= \sum_{\mathbf{m} \in \mathbb{Z}^d} G(a, q; \mathbf{m}) \int_{\mathbb{T}^d} e(-\mathbf{x} \cdot \xi) \Psi(q\xi - \mathbf{m}) \widetilde{d\sigma} \left( \lambda^{1/k} \left[ \xi - \frac{\mathbf{m}}{q} \right] \right) d\xi \\ &= \sum_{\mathbf{b} \in (\mathbb{Z}/q)^d} G(a, q; \mathbf{b}) \sum_{\mathbf{n} \in \mathbb{Z}^d} \int_{\mathbb{T}^d} e(-\mathbf{x} \cdot \xi) \Psi(q\xi - q\mathbf{n} - \mathbf{b}) \widetilde{d\sigma} \left( \lambda^{1/k} \left[ \xi - \mathbf{n} - \frac{\mathbf{b}}{q} \right] \right) d\xi \\ &= \sum_{\mathbf{b} \in (\mathbb{Z}/q)^d} G(a, q; \mathbf{b}) \int_{\mathbb{R}^d} e(-\mathbf{x} \cdot \xi) \Psi(q\xi - \mathbf{b}) \widetilde{d\sigma} \left( \lambda^{1/k} \left[ \xi - \frac{\mathbf{b}}{q} \right] \right) d\xi \end{aligned}$$

The second equality follows since there is only a single term in the sum for each  $\xi$  while the third follows from writing every  $\mathbf{m} \in \mathbb{Z}^d$  as  $q\mathbf{n} + \mathbf{b}$  for some  $\mathbf{n} \in \mathbb{Z}^d$  and a representative  $\mathbf{b} \in \{0, 1, \dots, q - 1\}^d$  which we identify with  $(\mathbb{Z}/q)^d$ . Using the well-known translation, modulation and dilation symmetries of the (inverse) Fourier transform, we have

$$\begin{aligned} K_\lambda^{a/q}(\mathbf{x}) &= \sum_{\mathbf{b} \in (\mathbb{Z}/q)^d} G(a, q; \mathbf{b}) \int_{\mathbb{R}^d} e(-\mathbf{x} \cdot \xi) D_{q^{-1}} \Psi \left( \xi - \frac{\mathbf{m}}{q} \right) \widetilde{D_{\lambda^{-1/k}} d\sigma} \left( \xi - \frac{\mathbf{m}}{q} \right) \\ &= \sum_{\mathbf{b} \in (\mathbb{Z}/q)^d} G(a, q; \mathbf{b}) e \left( -\frac{\mathbf{b} \cdot \mathbf{x}}{q} \right) \widetilde{D_{q^{-1}\lambda^{1/k}} \Psi} * d\sigma \left( \lambda^{-1/k} \mathbf{x} \right). \end{aligned}$$

Identity (3.11) immediately follows since

$$\sum_{\mathbf{b} \in (\mathbb{Z}/q)^d} G(a, q; \mathbf{b}) e \left( \frac{-\mathbf{b} \cdot \mathbf{x}}{q} \right) = e \left( \frac{a\mathcal{Q}(\mathbf{x})}{q} \right).$$

We leave this calculation to the reader since its just the inverse  $(\mathbb{Z}/q)^d$ -Fourier transform of the Gauss sum, and the Gauss sum is the  $(\mathbb{Z}/q)^d$ -Fourier transform of the function  $e \left( \frac{a\mathcal{Q}(\mathbf{x})}{q} \right)$ .  $\square$

Now that we know the structure of our kernel we will use a Littlewood–Paley decomposition and a circle method decomposition to arbitrage  $\ell^1 \rightarrow \ell^\infty$  and  $\ell^2 \rightarrow \ell^2$  estimates to deduce Lemma 3.3. In particular, the following Lemma 3.6 is motivated by the decompositions in [Ion04] when  $k = 2$  and [Hug17] when  $k \geq 3$ . With this in mind, we introduce a low-high frequency decomposition in the analytic aspect.

For  $\Delta \in (0, 1)$ , define the low frequency piece by the Fourier multiplier

$$\widehat{M_\lambda^{a/q, \text{low}}}(\xi) := \Psi(2q\Delta\lambda^{1/k}[\xi - \mathbf{m}/q]) \widehat{M_\lambda^{a/q}}(\xi).$$

The high frequency piece is defined as

$$M_\lambda^{a/q, \text{high}} := M_\lambda^{a/q} - M_\lambda^{a/q, \text{low}}.$$

By the Fourier localization of  $M_\lambda^{a/q}$  we have the restriction that  $\Delta \gtrsim \lambda^{-1/k}$ .

LEMMA 3.6. — Suppose that  $\Delta \in (0, 1)$ . Let  $\mathcal{Q}(\mathbf{x}) \in \mathbb{Z}[\mathbf{x}]$  be a positive definite, non-singular, integral, homogeneous form satisfying (3.1), and  $\Gamma_{\mathcal{Q}}$  be a set of regular values for the form  $\mathcal{Q}$ . If  $1 \leq a < q \lesssim \Lambda^{1/k}$  with  $(a, q) = 1$ , then each major arc piece  $M^{a/q, \mathcal{Q}} = M^{a/q, \mathcal{Q}, \text{low}} + M^{a/q, \mathcal{Q}, \text{high}}$  decomposes into a low frequency and high frequency piece such that

$$(3.12) \quad \left\| \sup_{\Lambda \leq \lambda < 2\Lambda} |M_{\lambda}^{a/q, \text{low}}| \right\|_{\ell^1(\mathbb{Z}^d) \rightarrow \ell^\infty(\mathbb{Z}^d)} \lesssim (q\Delta)^{-1} \Lambda^{-\frac{d}{k}}$$

and

$$(3.13) \quad \left\| \sup_{\Lambda \leq \lambda < 2\Lambda} |M_{\lambda}^{a/q, \text{high}}| \right\|_{\ell^2(\mathbb{Z}^d) \rightarrow \ell^2(\mathbb{Z}^d)} \lesssim_{\epsilon} \Lambda^{\epsilon} q^{-\frac{3}{2}} \Delta^{\kappa - \frac{3}{2}}.$$

*Proof.* — The  $\ell^2 \rightarrow \ell^2$ -estimate is proved using Bourgain’s  $L^2$ -estimates from [Bou85] as in [Ion04, Hug17]. We only sketch the proof of the  $\ell^1 \rightarrow \ell^\infty$ -estimate.

When  $\mathcal{Q}(\mathbf{x}) := \sum_{i=1}^d |x_i|^2$ , then we have the following known bound for the continuous spherical measure

$$(3.14) \quad \widetilde{D}_t \Psi * d\sigma_{\mathcal{Q}}(\mathbf{x}) \lesssim_d t^{-1} (1 + |\mathbf{x}/t|)^{-2d}.$$

See for instance [Ion04, p. 1415] where (3.14) is used to bound the discrete spherical maximal function, or [Gra08, (5.5.12)] for its derivation. The estimate (3.14) also holds for our varieties  $\{\mathbf{x} \in \mathbb{R}^d : \mathcal{Q}(\mathbf{x}) = 1\}$  since the proof only relies on the dimensionality of the measure  $d\sigma_{\mathcal{Q}}$ . To be precise all we require is that

$$(3.15) \quad d\sigma(B_r) \lesssim \min\{1, r^{d-1}\} \quad \text{for all balls of radius } r > 0 \text{ in } \mathbb{R}^d.$$

This implies that for each  $\mathbf{x} \in \mathbb{Z}^d$  we have

$$\begin{aligned} \sup_{\Lambda \leq \lambda < 2\Lambda: \lambda \in \Gamma_{\mathcal{Q}}} |M_{\lambda}^{a/q} f(\mathbf{x})| &\lesssim_d |f| * \left[ \sup_{\Lambda \leq \lambda < 2\Lambda: \lambda \in \Gamma_{\mathcal{Q}}} (q\Delta)^{-1} \lambda^{-\frac{d}{k}} \left(1 + |\lambda^{-1/k} \cdot |\right)^{-2d} \right] (\mathbf{x}) \\ &\lesssim_d |f| * \left[ (q\Delta)^{-1} \Lambda^{-\frac{d}{k}} \left(1 + |\Lambda^{-1/k} \cdot |\right)^{-2d} \right] (\mathbf{x}) \\ &\lesssim_d \|f\|_{\ell^1(\mathbb{Z}^d)} \cdot \left\| (q\Delta)^{-1} \Lambda^{-\frac{d}{k}} \left(1 + |\Lambda^{-1/k} \cdot |\right)^{-2d} \right\|_{\ell^\infty(\mathbb{Z}^d)} \\ &\lesssim_d (q\Delta)^{-1} \Lambda^{-\frac{d}{k}} \cdot \|f\|_{\ell^1(\mathbb{Z}^d)} \end{aligned}$$

□

Remark 3.7. — Estimates (3.8) and (3.15) and a Euclidean version of the low–high decomposition suffice to prove Euclidean versions of Theorem 3.1 generalizing Littman’s theorem.

We now return to the proof of Lemma 3.3. Fix  $\Lambda \geq 1$ . Let  $X$  be a fixed, finite subset of  $\mathbb{Z}^d$ . Our first observation is that estimate (3.12) implies that for  $\Lambda \leq \lambda < 2\Lambda$ .

$$(3.16) \quad \sup_{\Lambda \leq \lambda < 2\Lambda} |M_{\lambda} \mathbf{1}_X(\mathbf{x})| \lesssim \Lambda^{\frac{2-d}{k}} |X| \quad \text{for all } \mathbf{x} \in \mathbb{Z}^d$$

by taking  $\Delta \approx \Lambda^{-1/k}$  and summing over the moduli  $1 \leq q \lesssim \Lambda^{1/k}$ . The implicit bound is independent of the set  $X$ . Consequently we are reduced to proving

$$(3.17) \quad \left| \left\{ \mathbf{x} \in \mathbb{Z}^d : \sup_{\Lambda \leq \lambda < 2\Lambda} |M_\lambda \mathbf{1}_X(\mathbf{x})| > T \right\} \right| \lesssim_\epsilon \Lambda^{\frac{d(2-\kappa)}{k} + \epsilon} T^{-\kappa} |X|^{\kappa-1} \quad \text{for } 0 < T \lesssim \Lambda^{\frac{2-d}{k}} |X|.$$

Since

$$\begin{aligned} & \left| \left\{ \mathbf{x} \in \mathbb{Z}^d : \sup_{\Lambda \leq \lambda < 2\Lambda} |M_\lambda f(\mathbf{x})| > T \right\} \right| \\ & \leq \left| \left\{ \mathbf{x} \in \mathbb{Z}^d : \sup_{\Lambda \leq \lambda < 2\Lambda} \left| \sum_{q \leq Q} M_\Lambda^{q, \text{low}} f(\mathbf{x}) \right| > T/3 \right\} \right| \\ & \quad + \left| \left\{ \mathbf{x} \in \mathbb{Z}^d : \sup_{\Lambda \leq \lambda < 2\Lambda} \left| \sum_{q \leq Q} M_\Lambda^{q, \text{high}} f(\mathbf{x}) \right| > T/3 \right\} \right| \\ & \quad + \left| \left\{ \mathbf{x} \in \mathbb{Z}^d : \sup_{\Lambda \leq \lambda < 2\Lambda} \left| \sum_{q > Q} M_\Lambda^q f(\mathbf{x}) \right| > T/3 \right\} \right|, \end{aligned}$$

the inequalities of Lemma 3.6 combine to imply that

$$(3.18) \quad \left| \left\{ \mathbf{x} \in \mathbb{Z}^d : \sup_{\Lambda \leq \lambda < 2\Lambda} |M_\lambda f(\mathbf{x})| > T \right\} \right| \lesssim_\epsilon \Lambda^\epsilon \left[ Q^{1/2} \Delta^{\kappa - \frac{3}{2}} + Q^{2-\kappa} \right]^2 T^{-2} |X|$$

provided that we choose  $Q$  and  $\Delta$  such that  $Q\Delta^{-1}\Lambda^{-d/k}|X| \lesssim T$ .

The inequality (3.18) shows that we want to choose  $Q^{1/2}\Delta^{\kappa - \frac{3}{2}} \approx Q^{2-\kappa}$  which is  $\Delta \approx Q^{-1}$ ; we now make this assumption. Our restriction then takes the form  $Q^2\Lambda^{-d/k}|X| \lesssim T$  which is consistent with our reduction (3.17). Choosing  $Q^2 \approx \Lambda^{d/k}T|X|^{-1}$ , we deduce that

$$\begin{aligned} & \left| \left\{ \mathbf{x} \in \mathbb{Z}^d : \sup_{\Lambda \leq \lambda < 2\Lambda} |M_\lambda f(\mathbf{x})| > T \right\} \right| \lesssim_\epsilon \Lambda^\epsilon \left[ \Lambda^{d/k}T|X|^{-1} \right]^{2-\kappa} T^{-2} |X| \\ & \lesssim_\epsilon \Lambda^{\frac{d(2-\kappa)}{k} + \epsilon} T^{-\kappa} |X|^{\kappa-1} \end{aligned}$$

for all  $0 < T \leq 1$  as desired.

### 3.2. Proof of Lemma 3.4

The proof of Lemma 3.4 follows by interpolating the error term estimate (3.5) in Magyar’s theorem with estimate (3.19) below.

PROPOSITION 3.8. — *We have the following  $\ell^1 \rightarrow \ell^\infty$ -improving estimate for the error term:*

$$(3.19) \quad \left\| \sup_{\Lambda \leq \lambda < 2\Lambda : \lambda \in \Gamma_Q} |E_\lambda f| \right\|_{\ell^\infty(\mathbb{Z}^d)} \lesssim \Lambda^{1-\frac{d}{k}} \|f\|_{\ell^1(\mathbb{Z}^d)}.$$

*Proof.* — Since  $E_\lambda = A_\lambda - M_\lambda$ , the trivial bound (3.2) for the dyadic maximal function and our bound (3.16) for the main term imply (3.19).  $\square$

### 4. The discrete spherical averages

In this section we refine our main term analysis from Section 3 in order to prove Theorem 1.1. We now recall a decomposition of Magyar which, for the discrete spherical averages, improves upon the error term bound (3.5) in Magyar 1.

MAGYAR 2 ([Mag07]). — Suppose that  $Q(\mathbf{x}) = x_1^2 + \dots + x_d^2$  for  $d \geq 4$ . For each  $\lambda \in \mathbb{N}$  (with  $\lambda$  assumed to be odd when  $d = 4$ ) and all  $\xi \in \mathbb{T}^d$ , we have  $A_\lambda^Q = M_\lambda^Q + E_\lambda^Q$  for  $\lambda \in \mathbb{N}$  such that

$$(4.1) \quad \|\widehat{E}_\lambda\|_{L^\infty(\mathbb{T}^d)} \lesssim_\epsilon \lambda^{\epsilon - \frac{d-3}{4}} \quad \text{for all } \epsilon > 0.$$

The main term  $M_\lambda^Q$  decomposes into a sum of pieces  $M_\lambda^Q = \sum_{q=1}^{\Lambda^{1/2}} M_\lambda^{Q,q}$  where  $\Lambda$  is the smallest dyadic integer (that is of the form  $2^j$  for some  $j \in \mathbb{N}$ ) such that  $\Lambda \leq \lambda < 2\Lambda$  and  $M_\lambda^Q$  has Fourier multiplier

$$\widehat{M}_\lambda^q(\xi) := \sum_{\mathbf{m} \in \mathbb{Z}^d} \text{Kl}(q, \lambda; \mathbf{m}) \Psi(2q\sqrt{\lambda}[\xi - \mathbf{m}/q]) \widetilde{d}\sigma(\sqrt{\lambda}[\xi - \mathbf{m}/q]).$$

Each piece of the main term satisfies the following “Weil bound”:

$$(4.2) \quad \|\widehat{M}_\lambda^q\|_{L^\infty(\mathbb{T}^d)} \lesssim_\epsilon q^{-\left(\frac{d-1}{2}\right) + \epsilon} (q, \lambda)^{1/2} \quad \text{for all } \epsilon > 0.$$

In (4.2) and below  $(a, b)$  represents the greatest common divisor of two integers  $a, b$ , and  $\text{Kl}(q, \lambda; \mathbf{m})$  is the Kloosterman/Salie sum defined as

$$\text{Kl}(q, \lambda; \mathbf{m}) := q^{-d} \sum_{\mathbf{a} \in (\mathbb{Z}/q)^\times} e\left(-\frac{a\lambda}{q}\right) \sum_{\mathbf{b} \in (\mathbb{Z}/q)^d} e\left(\frac{aQ(\mathbf{b}) + \mathbf{b} \cdot \mathbf{m}}{q}\right)$$

for  $q \in \mathbb{N}$  and  $\mathbf{m} \in \mathbb{Z}^d$ . For the estimate (4.1), see in [Mag07, (1.9) of Lemma 1]. When comparing, note that we have normalized our surface measure to be a probability measure. The estimate (4.2) follows from the famous Weil bounds for Kloosterman sums:

$$\text{Kl}(q, \lambda; \mathbf{m}) \lesssim_\epsilon q^{-\left(\frac{d-1}{2}\right) + \epsilon} (q, \lambda)^{1/2} \quad \text{for each } \mathbf{m} \in \mathbb{Z}^d \quad \text{and for all } \epsilon > 0.$$

Also note that  $M_\lambda^q = \sum_{a \in (\mathbb{Z}/q)^\times} M_\lambda^{a/q}$ .

Our strategy is the same as before; we need to prove bounds for the main term and the error term. We first improve our bounds for the main term.

LEMMA 4.1. — For  $Q(\mathbf{x}) := \sum_{i=1}^d x_i^2$ ,  $d \geq 4$  and  $p = \frac{d+1}{d-1}$ ,

$$(4.3) \quad \|M_\lambda\|_{\ell^{p,1}(\mathbb{Z}^d) \rightarrow \ell^{p',\infty}(\mathbb{Z}^d)} \lesssim_\epsilon \lambda^{\epsilon - \frac{d}{2} \left(\frac{2}{p} - 1\right)} \quad \text{for all } \epsilon > 0.$$

Subsequently, we improve our bound for the error term.

LEMMA 4.2. — When  $Q(\mathbf{x}) := \sum_{i=1}^d x_i^2$ ,  $d \geq 4$  and  $1 \leq p \leq 2$ .

$$(4.4) \quad \|E_\lambda\|_{\ell^p(\mathbb{Z}^d) \rightarrow \ell^{p'}(\mathbb{Z}^d)} \lesssim_\epsilon \lambda^{-\frac{d}{2} \left(\frac{2}{p} - 1\right) + \left(\frac{2}{p} - 1\right) - \left(\frac{d-3}{4}\right) \left(2 - \frac{2}{p}\right) + \epsilon} \quad \text{for all } \epsilon > 0.$$

A simple computation reveals that  $\left(\frac{2}{p} - 1\right) - \left(\frac{d-3}{4}\right) \left(2 - \frac{2}{p}\right) < 0$  for  $p > \frac{d+1}{d-1}$ . Theorem 1.1 follows immediately from combining Lemmas 4.1 and 4.2.

### 4.1. Proof of Lemma 4.1

We use the low-high decomposition of Section 3. First we have

$$\left\| \sum_{Q < q \lesssim \lambda^{1/2}} M_\lambda^q \right\|_{\ell^2(\mathbb{Z}^d) \rightarrow \ell^2(\mathbb{Z}^d)} \lesssim_\epsilon Q^{\epsilon - \frac{d-3}{2}} \lambda^\epsilon$$

because summing over  $(q, \lambda)^{1/2}$  only contributes a factor of  $\lambda^\epsilon$  on average. Fourier transform estimates for the spherical measure and the Weil bound for Kloosterman sums (4.2) implies the following bound on the high frequency pieces

$$\left\| M_\lambda^{q,\text{high}} \right\|_{\ell^2(\mathbb{Z}^d) \rightarrow \ell^2(\mathbb{Z}^d)} \lesssim_\epsilon q^{\epsilon - \frac{d-1}{2}} (q, \lambda)^{1/2} (q\Delta)^{-\frac{d-1}{2}} = q^\epsilon (q, \lambda)^{1/2} \Delta^{-\frac{d-1}{2}}.$$

Finally the low frequency bound is the same as before:

$$\left\| M_\lambda^{a/q,\text{low}} \right\|_{\ell^1(\mathbb{Z}^d) \rightarrow \ell^\infty(\mathbb{Z}^d)} \lesssim (q\Delta)^{-1} \lambda^{-\frac{d}{2}}.$$

The proof now proceeds by the restricted weak-type argument of Section 3. We will be brief in our description of this. Let  $X$  be a fixed subset of  $\mathbb{Z}^d$ . By the same reduction as in the proof of Lemma 3.3, we seek to prove

$$|\{ |M_\lambda \mathbf{1}_X(\mathbf{x})| > T \}| \lesssim_\epsilon \lambda^{-\frac{d(d-3)}{4} + \epsilon} T^{-\frac{d+1}{2}} |X|^{\frac{d-1}{2}} \quad \text{for } 0 < T \lesssim \lambda^{1-\frac{d}{2}} |X|.$$

Combining the above estimates for low pieces  $M_\lambda^{q,\text{low}}$  and high pieces  $M_\lambda^{q,\text{high}}$  implies that

$$|\{ \mathbf{x} \in \mathbb{Z}^d : M_\lambda f(\mathbf{x}) > T \}| \lesssim_\epsilon \lambda^\epsilon \left( [Q\Delta^{\frac{d-1}{2}}]^2 + Q^{3-d} \right) T^{-2} |X|$$

provided that we choose  $Q$  and  $\Delta$  such that  $Q\Delta^{-1}\lambda^{-d/2}|X| \lesssim T$ . We choose  $\Delta \approx Q^{-1}$  so that  $Q^2\lambda^{-d/2}|X| \approx T$ . Plugging this in we ascertain that

$$\begin{aligned} |\{ \mathbf{x} \in \mathbb{Z}^d : M_\lambda f(\mathbf{x}) > T \}| &\lesssim_\epsilon \lambda^\epsilon \left[ \lambda^{d/2} T |X|^{-1} \right]^{-\frac{d-3}{2}} \cdot T^{-2} |X| \\ &= \lambda^{-\frac{d(d-3)}{4} + \epsilon} T^{-\frac{d+1}{2}} |X|^{\frac{d-1}{2}} \end{aligned}$$

for all  $0 < T \leq 1$  as desired.

### 4.2. Proof of Lemma 4.2

We have Magyar’s Kloosterman bound (4.1) which says that we may take  $\gamma_Q := \frac{d-3}{4}$  for  $Q(\mathbf{x}) := \sum_{i=1}^d |x_i|^2$  with  $d \geq 4$ . Interpolating with the  $\ell^1(\mathbb{Z}^d) \rightarrow \ell^\infty(\mathbb{Z}^d)$ -bound (3.19), we obtain the Lemma 4.2.

### 5. The dyadic discrete $k$ -spherical maximal functions

Let  $\mathcal{Q}(\mathbf{x}) := \sum_{i=1}^d |x_i|^k$  for integers  $k \geq 2$ , and  $\Gamma_{\mathcal{Q}}$  a set of regular values for  $\mathcal{Q}$ . (For these forms,  $\Gamma_{\mathcal{Q}}$  contains all sufficiently large natural numbers when  $d \gtrsim k^2$ .) For degrees  $k \geq 3$ , define the dimensions

$$d_k := k^2 - \max_{2 \leq j \leq k-1} \left\{ \frac{kj - \min(2^j + 2, j^2 + j)}{j - j + 1} \right\},$$

and the exponents

$$p_{d,k} := \max \left\{ \frac{d}{d - k}, 1 + \frac{1}{2\gamma_{d,k}} \right\} \quad \text{for } k \geq 3 \quad \text{and } d > d_k$$

where  $\gamma_{d,k}$  is defined as

$$k \cdot \gamma_{d,k} := \begin{cases} (d - d_k)(k^2 + k - d_k)^{-1} & \text{if } d_k \leq d \leq k^2 + k, \\ 1 + (d - k^2 - k) \left( \max\{2^{1-k}, (k^2 - k)^{-1}\} \right) & \text{if } d > k^2 + k. \end{cases}$$

When  $k = 2$ , define  $d_2 := 4$ ,  $\gamma_{d,2} := 1 - \frac{d}{4}$  and  $p_{d,2} := \frac{d}{d-2}$ . These bounds relied on the works [BDG16, Bou17, BR15, Woo12].

For  $k$ -spheres we may generalize Theorem 1.2 to the following Theorem 5.1.

**THEOREM 5.1.** — *For  $k \geq 2$ ,  $d \geq d_k$  and  $p > p_{d,k}$ ,*

$$(5.1) \quad \left\| \sup_{\Lambda \leq \lambda < 2\Lambda} |A_\lambda| \right\|_{\ell^p(\mathbb{Z}^d) \rightarrow \ell^{p'}(\mathbb{Z}^d)} \lesssim \Lambda^{-\frac{d}{k}(\frac{2}{p}-1)}.$$

In the proof of this Theorem 5.1 we will see that we may easily deduce the restricted weak-type bound claimed in Theorem 1.2. As before we will break into two lemmas which handle the main term and error term respectively, and from which Theorem 5.1 follows immediately.

**LEMMA 5.2.** — *For  $k \geq 2$ ,  $d \geq 2k + 1$  and  $p := \frac{d}{d-k}$ ,*

$$(5.2) \quad \left\| \sup_{\Lambda \leq \lambda < 2\Lambda} |M_\lambda| \right\|_{\ell^{p,1}(\mathbb{Z}^d) \rightarrow \ell^{p',\infty}(\mathbb{Z}^d)} \lesssim \Lambda^{-\frac{d}{k}(\frac{2}{p}-1)}.$$

**LEMMA 5.3.** — *For  $k \geq 2$ ,  $d \geq d_k$  and  $p > p_{d,k}$ ,*

$$(5.3) \quad \left\| \sup_{\Lambda \leq \lambda < 2\Lambda} |E_\lambda| \right\|_{\ell^p(\mathbb{Z}^d) \rightarrow \ell^{p'}(\mathbb{Z}^d)} \lesssim_\epsilon \Lambda^{-\frac{d}{k}(\frac{2}{p}-1) + (\frac{2}{p}-1) - \gamma_{d,k}(2 - \frac{2}{p}) + \epsilon}.$$

When  $k = 2$  one may remove the  $\epsilon$ -loss.

#### 5.1. Proof of Lemma 5.2

We first handle the main term which relies on *Steckin’s estimate* (see [Hug17]):

$$(5.4) \quad G_{\mathcal{Q}}(a, q; \mathbf{m}) \lesssim_k q^{-d/k}$$



where the bound is uniform in  $a$  coprime to  $q$  and  $\mathbf{m} \in \mathbb{Z}^d$ , and Bruna–Nagel–Wainger’s estimates (see [BNW88]) imply that

$$|\widetilde{d\sigma_{\mathcal{Q}}}(\xi)| \lesssim (1 + |\xi|)^{-\frac{d-1}{k}}.$$

The effect of the following is that we may replace  $\kappa_{\mathcal{Q}}$  in Section 3 with  $d/k$ . An essential point below is that we have no  $\epsilon$ -loss; that is, no extraneous factors of  $\Lambda^\epsilon$  like before.

We first sharpen the bound for our main term. As shown in [MSW02] and [Hug17], the Gauss bound for Gauss sums and Steckin’s estimate (5.4) implies that

$$\left\| \sup_{\Lambda \leq \lambda < 2\Lambda: \lambda \in \Gamma_{\mathcal{Q}}} |M_{\lambda}^{a/q} f| \right\|_{\ell^2(\mathbb{Z}^d)} \lesssim q^{-\frac{d}{k}} \|f\|_{\ell^2(\mathbb{Z}^d)}.$$

Consequently,

$$\left\| \sup_{\Lambda \leq \lambda < 2\Lambda: \lambda \in \Gamma_{\mathcal{Q}}} \left| \sum_{Q < q \lesssim \Lambda^{1/2}} \sum_{a \in (\mathbb{Z}/q)^*} M_{\lambda}^{a/q} f \right| \right\|_{\ell^2(\mathbb{Z}^d)} \lesssim Q^{2-\frac{d}{k}} \|f\|_{\ell^2(\mathbb{Z}^d)}.$$

Using the low-high decomposition of Section 3, we have still have (3.12) for the low-frequency piece and for the high-frequency piece we now have

$$\left\| \sup_{\Lambda \leq \lambda < 2\Lambda} |M_{\lambda}^{a/q, \text{high}}| \right\|_{\ell^2(\mathbb{Z}^d) \rightarrow \ell^2(\mathbb{Z}^d)} \lesssim q^{-d/k} (q\Delta)^{\frac{d-1}{k} - \frac{1}{2}} = q^{-(\frac{1}{k} + \frac{1}{2})} \Delta^{\frac{d-1}{k} - \frac{1}{2}}.$$

Running the restricted weak-type argument as before we conclude the Lemma 5.2.

### 5.2. Proof of Lemma 5.3

We now handle the error term. From [MSW02] and [ACHK18], we have that

$$(5.5) \quad \left\| \sup_{\Lambda \leq \lambda < 2\Lambda} |E_{\lambda} f| \right\|_{\ell^2} \lesssim_{\epsilon} \Lambda^{\epsilon - \gamma_{d,k}} \|f\|_{\ell^2}$$

where one may remove the  $\epsilon$  in (5.5) when  $k = 2$ . Interpolating with the  $\ell^1(\mathbb{Z}^d)$ -bound from Section 3 (see the proof of Lemma 3.4), we obtain the Lemma 5.3.

## 6. The discrete spherical maximal function in four dimensions

In this section let  $\mathcal{Q}(\mathbf{x}) := x_1^2 + x_2^2 + x_3^2 + x_4^2$ . We have the following  $\ell^p(\mathbb{Z}^d)$ -improving estimates for Magyar’s discrete dyadic spherical maximal functions.

**THEOREM 6.1.** — *If  $d = 4$ , then for each  $q > 2$  and  $1 \leq p < q$  there exists  $\delta_{p,q} > 0$  such that*

$$(6.1) \quad \left\| \sup_{\Lambda \leq \lambda < \Lambda} |A_{\lambda} f| \right\|_{\ell^q(\mathbb{Z}^d)} \lesssim \Lambda^{-\delta_{p,q}} \|f\|_{\ell^p(\mathbb{Z}^d)}.$$

*We assume that the  $\lambda$  in the supremum are restricted to be odd.*

Theorem 1.10 follows from summing up Theorem 6.1 over  $\Lambda = 2^j$  for integers  $j \geq 0$ .

*Proof of Theorem 6.1.* — Our trivial bound (1.6) says that

$$\left\| \sup_{\Lambda \leq \lambda < 2\Lambda} |A_\lambda f| \right\|_{\ell^\infty(\mathbb{Z}^4)} \lesssim \Lambda^{-1} \|f\|_{\ell^1(\mathbb{Z}^4)}.$$

We need one additional ingredient which is from the author’s thesis.

**THEOREM 6.2** ([Hug12]). — *For all  $\lambda > 0$ , we have*

$$\left\| \sup_{\Lambda \leq \lambda < 2\Lambda} |A_\lambda f| \right\|_{\ell^2(\mathbb{Z}^4)} \lesssim (\log \Lambda)^2 \|f\|_{\ell^2(\mathbb{Z}^4)}.$$

*The implicit constant is independent of  $\Lambda$ .*

Interpolating our trivial bound and Theorem 6.2 we obtain that

$$(6.2) \quad \left\| \sup_{\Lambda \leq \lambda < \Lambda} |A_\lambda f| \right\|_{\ell^{p'}(\mathbb{Z}^4)} \lesssim \Lambda^{-\left(\frac{2}{p}-1\right)} (\log \Lambda)^2 \|f\|_{\ell^p(\mathbb{Z}^4)} \quad \text{for } 1 \leq p \leq 2.$$

The nesting property of  $\ell^p$ -spaces implies that

$$\left\| \sup_{\Lambda \leq \lambda < 2\Lambda} |A_\lambda f| \right\|_{\ell^q(\mathbb{Z}^4)} \lesssim \Lambda^{-\left(\frac{2}{q'}-1\right)} (\log \Lambda)^2 \|f\|_{\ell^p(\mathbb{Z}^4)} \quad \text{for all } q \geq 2 \text{ and } 1 \leq p \leq q'$$

which implies Theorem 6.1 in this range. Interpolating (6.2) with the trivial  $\ell^\infty$ -bound

$$\left\| \sup_{\Lambda \leq \lambda < 2\Lambda} |A_\lambda f| \right\|_{\ell^\infty(\mathbb{Z}^4)} \leq \|f\|_{\ell^\infty(\mathbb{Z}^4)},$$

we obtain for each  $q > 2$  that

$$\left\| \sup_{\Lambda \leq \lambda < 2\Lambda} |A_\lambda f| \right\|_{\ell^s(\mathbb{Z}^4)} \lesssim [\Lambda^{-\left(\frac{2}{p}-1\right)} (\log \Lambda)^2]^{q'/r} \|f\|_{\ell^r(\mathbb{Z}^4)} \quad \text{for all } r \geq q \text{ and } s = \frac{q'}{rq}.$$

Consequently, (6.1) holds for all  $q > 2$  and  $p < q$ . □

### 7. Further questions

Let  $X$  be a finite subset of  $\mathbb{Z}^d$  and  $\mu_X := |X|^{-1} \mathbf{1}_X$  be its normalized probability measure. By Young’s inequality, we have

$$(7.1) \quad \|f * \mu_X\|_{\ell^{p'}(\mathbb{Z}^d)} \lesssim |X|^{1-\frac{2}{p}} \|f\|_{\ell^p(\mathbb{Z}^d)}.$$

For instance if  $X = X_r$  is taken to be the integer points of a large ball  $\{\mathbf{x} \in \mathbb{Z}^d : |\mathbf{x}| < r\}$  for  $r > 0$ , then

$$\|f * \mu_X\|_{\ell^{p'}(\mathbb{Z}^d)} \lesssim r^{d\left(1-\frac{2}{p}\right)} \|f\|_{\ell^p(\mathbb{Z}^d)}$$

since  $\#\{\mathbf{x} \in \mathbb{Z}^d : |\mathbf{x}| < r\} \asymp r^d$  for all large  $r > 0$ . Moreover testing against Example 1.5 (big ball example) we see that this bound is sharp (aside possibly from the implicit constants). Therefore (7.1) is in general sharp for all  $1 \leq p \leq 2$ , and

another way to interpret Theorem 1.1 is that the discrete sphere smooths almost as well as the discrete ball which contains the same number of points as the sphere in the range  $\frac{d+1}{d-1} \leq p < 2$ . This begs with a few questions.

QUESTION 7.1 (Michael Fryers). — *For a random subset  $X$  what is the best constant in the inequality (7.1)?*

QUESTION 7.2 (Michael Fryers). — *Is (7.1) still true if we take  $X$  to be a perturbation of the discrete spherical measures  $\sigma_\lambda$ ?*

QUESTION 7.3. — *What are the extremizers satisfying (7.1) for all  $1 \leq p \leq 2$ ? Do they resemble the ball?*

### Appendix A. $\ell^p(\mathbb{Z}^d)$ -boundedness for Magyar’s theorem

[Mag02, Theorem 4] is only stated for  $\ell^2(\mathbb{Z}^d)$ . Since it has come up in conversation on multiple occasions, we record the range of  $\ell^p(\mathbb{Z}^d)$ -boundedness to which Magyar’s theorem extends and briefly indicate how. Standard density arguments extend the range of Magyar’s pointwise ergodic theorem ([Mag02, Theorem 3]) to the same range of  $L^p$ -spaces as  $\ell^p(\mathbb{Z}^d)$ -spaces below.

In this appendix assume that  $\mathcal{Q}(\mathbf{x}) \in \mathbb{Z}[\mathbf{x}]$  is a non-singular, integral, homogeneous form with a set of regular values  $\Gamma_{\mathcal{Q}}$  satisfying (3.1), define the maximal function for functions  $f : \mathbb{Z}^d \rightarrow \mathbb{C}$  by

$$A_*f(\mathbf{x}) := \sup_{\lambda \in \Gamma_{\mathcal{Q}}} |A_\lambda f(\mathbf{x})| \quad \text{for each } \mathbf{x} \in \mathbb{Z}^d.$$

MAGYAR 3. — *The maximal function  $A_*$  is bounded on  $\ell^p(\mathbb{Z}^d)$  for  $p > \max \left\{ \frac{\kappa}{\kappa-1}, \frac{2(\gamma+1)}{2\gamma+1} \right\}$  where  $\kappa = \kappa_{\mathcal{Q}}$  and  $\gamma = \gamma_{\mathcal{Q}}$  are defined by (3.3).*

As in Section 3, we have  $A_\lambda = M_\lambda + E_\lambda$ . This implies that for each function  $f : \mathbb{Z}^d \rightarrow \mathbb{C}$  and each  $\mathbf{x} \in \mathbb{Z}^d$

$$A_*f(\mathbf{x}) \leq \left( \sum_{q=1}^{\infty} \sum_{a \in (\mathbb{Z}/q)^*} |M_*^{a/q} f(\mathbf{x})| \right) + E_*f(\mathbf{x})$$

where

$$M_*^{a/q} f(\mathbf{x}) := \sup_{\lambda \in \Gamma_{\mathcal{Q}}} |M_\lambda^{a/q} f(\mathbf{x})|$$

and

$$E_*f(\mathbf{x}) := \sup_{\lambda \in \Gamma_{\mathcal{Q}}} |E_\lambda f(\mathbf{x})|.$$

Our strategy is then to give sufficiently decaying  $\ell^p$ -bounds on  $M_*^{a/q}$  to sum over  $a$  and  $q$  and to prove  $\ell^p$ -boundedness of  $E_*$  through its dyadic counterparts.

LEMMA A.1. — *If  $\mathcal{Q}(\mathbf{x}) \in \mathbb{Z}[\mathbf{x}]$  is a non-singular, integral, homogeneous form satisfying (3.1), then for  $1 \leq p \leq 2$  we have the bound*

$$(A.1) \quad \left\| M_*^{a/q} \right\|_{\ell^p(\mathbb{Z}^d) \rightarrow \ell^p(\mathbb{Z}^d)} \lesssim_\epsilon q^{\epsilon - \kappa(2 - \frac{2}{p})} \quad \text{for each } \epsilon > 0.$$

Summing over  $a$  and  $q$  immediately yields the following corollary. We see that we only need  $2 - \kappa(2 - \frac{2}{p}) < 0$  which is when  $p > \frac{\kappa}{\kappa-1}$ .

**COROLLARY A.2.** — *Let  $\mathcal{Q}(\mathbf{x}) \in \mathbb{Z}[\mathbf{x}]$  be a non-singular, integral, homogeneous form satisfying (3.1), and  $\Gamma_{\mathcal{Q}}$  be a set of regular values for the form  $\mathcal{Q}$ . If  $\frac{\kappa}{\kappa-1} < p \leq 2$ , then*

$$(A.2) \quad \left\| \sup_{\lambda \in \Gamma_{\mathcal{Q}}} |M_{\lambda} f| \right\|_{\ell^p(\mathbb{Z}^d)} \lesssim_{\mathcal{Q}, p} \|f\|_{\ell^p(\mathbb{Z}^d)}.$$

Our lemma for the error term is the following Lemma A.3 .

**LEMMA A.3.** — *Let  $\mathcal{Q}(\mathbf{x}) \in \mathbb{Z}[\mathbf{x}]$  be a non-singular, integral, homogeneous form satisfying (3.1), and  $\Gamma_{\mathcal{Q}}$  be a set of regular values for the form  $\mathcal{Q}$ . If  $1 \leq p \leq 2$ , then for all  $\epsilon > 0$ ,*

$$(A.3) \quad \left\| \sup_{\Lambda \leq \lambda < 2\Lambda: \lambda \in \Gamma_{\mathcal{Q}}} |E_{\lambda} f| \right\|_{\ell^p(\mathbb{Z}^d)} \lesssim_{\epsilon} \Lambda^{(\frac{2}{p}-1)-\gamma_{\mathcal{Q}}(2-\frac{2}{p})+\epsilon} \|f\|_{\ell^p(\mathbb{Z}^d)} \quad \text{for all } \Lambda \geq 1.$$

Summing (A.3) over  $\Lambda = 2^j$  for integers  $j \geq 0$  immediately yields the following corollary.

**COROLLARY A.4.** — *The error term maximal function  $E_*$  is bounded on  $\ell^p(\mathbb{Z}^d)$  for  $p > \frac{2(\gamma+1)}{2\gamma+1} = 1 + \frac{1}{2\gamma+1}$ .*

### A.1. Proof of Lemma A.1

We use the Magyar–Stein–Wainger approach of [MSW02]. For this we will need the following estimate for analogous Euclidean maximal functions.

**PROPOSITION A.5.** — *If  $\mathcal{Q}(\mathbf{x}) \in \mathbb{Z}[\mathbf{x}]$  is a non-singular, integral, homogeneous form satisfying (3.1), then the maximal function  $\sup_{t>0} |F * d\sigma_{\mathcal{Q}}|$ , defined a priori for smooth compactly supported functions  $F : \mathbb{R}^d \rightarrow \mathbb{C}$ , is bounded on  $L^p(\mathbb{R}^d)$  for  $p > 1 + (2\kappa - 2)^{-1} = \frac{2\kappa-1}{2(\kappa-1)}$ .*

*Proof.* — The proposition follows from Magyar’s estimate (3.8) by applying [RdF86, Theorem A] since  $\kappa > 2$ . □

We resume the proof of Lemma A.1. By (3.11) of Proposition 3.5 we have that

$$\left\| M_*^{a/q} \right\|_{\ell^1(\mathbb{Z}^d) \rightarrow \ell^1(\mathbb{Z}^d)} \lesssim 1.$$

The Magyar–Stein–Wainger transference principle ([MSW02, Proposition 2.1]) and the estimate (3.7) imply that

$$\left\| M_*^{a/q} \right\|_{\ell^2(\mathbb{Z}^d) \rightarrow \ell^2(\mathbb{Z}^d)} \lesssim q^{\epsilon-\kappa}.$$

Interpolation handles the remaining  $1 < p < 2$ .

## A.2. Proof of Lemma A.3

The  $\ell^2 \rightarrow \ell^2$ -estimate is (3.5). Interpolation with the  $\ell^1 \rightarrow \ell^1$ -estimate

$$\left\| \sup_{\Lambda \leq \lambda < 2\Lambda} |E_\lambda f| \right\|_{\ell^1(\mathbb{Z}^d)} \lesssim \Lambda \|f\|_{\ell^1(\mathbb{Z}^d)}$$

concludes the Lemma A.3.

To prove this  $\ell^1 \rightarrow \ell^1$ -estimate we note that the union bound implies that

$$\left\| \sup_{\Lambda \leq \lambda < 2\Lambda} |A_\lambda f| \right\|_{\ell^1(\mathbb{Z}^d)} \leq \Lambda \|f\|_{\ell^1(\mathbb{Z}^d)}$$

since each average is a contraction on  $\ell^1$ . Meanwhile, Proposition 3.5 implies that

$$(A.4) \quad \left\| \sup_{\Lambda \leq \lambda < 2\Lambda} |M_\lambda| \right\|_{\ell^1(\mathbb{Z}^d) \rightarrow \ell^1(\mathbb{Z}^d)} \lesssim_\epsilon \Lambda^{2/k}.$$

Since  $k \geq 2$ , we have that

$$\left\| \sup_{\Lambda \leq \lambda < 2\Lambda} |E_\lambda f| \right\|_{\ell^1(\mathbb{Z}^d)} = \left\| \sup_{\Lambda \leq \lambda < 2\Lambda} |A_\lambda f - M_\lambda f| \right\|_{\ell^1(\mathbb{Z}^d)} \lesssim \Lambda \|f\|_{\ell^1(\mathbb{Z}^d)}.$$

## BIBLIOGRAPHY

- [ACHK18] Theresa Anderson, Brian Cook, Kevin Hughes, and Angel Kumchev, *Improved  $\ell^p$  boundedness for integral  $k$ -spherical maximal functions*, Discrete Anal. **2018** (2018), 10 (18 pages). ↑975
- [BDG16] Jean Bourgain, Ciprian Demeter, and Larry Guth, *Proof of the main conjecture in Vinogradov's mean value theorem for degrees higher than three*, Ann. Math. **184** (2016), no. 2, 633–682. ↑974
- [Bir62] Bryan J. Birch, *Forms in many variables*, Proc. R. Soc. Lond. Ser. A **265** (1961/62), 245–263. ↑965
- [BNW88] Joaquim Bruna, Alexander Nagel, and Stephen Wainger, *Convex hypersurfaces and Fourier transforms*, Ann. Math. **127** (1988), no. 2, 333–365. ↑975
- [Bou85] Jean Bourgain, *Estimations de certaines fonctions maximales*, C. R. Acad. Sci. Paris Sér. I Math. **301** (1985), no. 10, 499–502. ↑964, 970
- [Bou17] ———, *On the Vinogradov mean value*, Proc. Steklov Inst. Math. **296** (2017), 30–40. ↑974
- [BR15] Jörg Brüdern and Olivier Robert, *Rational points on linear slices of diagonal hypersurfaces*, Nagoya Math. J. **218** (2015), 51–100. ↑974
- [Gra08] Loukas Grafakos, *Classical Fourier Analysis*, second ed., Graduate Texts in Mathematics, vol. 249, Springer, 2008. ↑970
- [HL14] Yi Hu and Xiaochun Li, *Discrete Fourier restriction associated with Schrödinger equations*, Rev. Mat. Iberoam. **30** (2014), no. 4, 1281–1300. ↑968
- [Hug12] Kevin Hughes, *Problems and results related to Waring's problem: Maximal functions and ergodic averages*, Ph.D. thesis, Princeton University, USA, 2012. ↑976
- [Hug17] ———, *Restricted weak-type endpoint estimates for  $k$ -spherical maximal functions*, Math. Z. **286** (2017), no. 3-4, 1303–1321. ↑969, 970, 974, 975
- [Hug19] ———, *The discrete spherical averages over a family of sparse sequences*, J. Anal. Math. **138** (2019), no. 1, 1–21. ↑964

- [Ion04] Alexandru D. Ionescu, *An endpoint estimate for the discrete spherical maximal function*, Proc. Am. Math. Soc. **132** (2004), no. 5, 1411–1417. ↑968, 969, 970
- [KL18] Robert Kesler and Michael T. Lacey,  *$\ell^p$ -improving inequalities for discrete spherical averages*, <https://arxiv.org/abs/1804.09845>, 2018. ↑963, 964
- [Lee03] Sanghyuk Lee, *Endpoint estimates for the circular maximal function*, Proc. Am. Math. Soc. **131** (2003), no. 5, 1433–1442. ↑960
- [Lit73] Walter Littman,  *$L^p - L^q$ -estimates for singular integral operators arising from hyperbolic equations*, Partial differential Equations, Berkeley 1971, Proceedings of Symposia in Pure Mathematics, vol. 23, American Mathematical Society, 1973, pp. 479–481. ↑960
- [Mag02] Ákos Magyar, *Diophantine equations and ergodic theorems*, Am. J. Math. **124** (2002), no. 5, 921–953. ↑964, 965, 967, 968, 977
- [Mag07] ———, *On the distribution of lattice points on spheres and level surfaces of polynomials*, J. Number Theory **122** (2007), no. 1, 69–83. ↑967, 972
- [MSW02] Ákos Magyar, Elias M. Stein, and Stephen Wainger, *Discrete analogues in harmonic analysis: spherical averages*, Ann. Math. **155** (2002), no. 1, 189–208. ↑963, 975, 978
- [RdF86] José L. Rubio de Francia, *Maximal functions and Fourier transforms*, Duke Math. J. **53** (1986), no. 2, 395–404. ↑978
- [Woo12] Trevor D. Wooley, *The asymptotic formula in Waring’s problem*, Int. Math. Res. Not. **2012** (2012), no. 7, 1485–1504. ↑974

Manuscript received on 5th October 2018,  
revised on 25th November 2019,  
accepted on 11th February 2020.

Recommended by Editor P. Gérard.  
Published under license CC BY 4.0.



This journal is a member of Centre Mersenne.



Kevin HUGHES  
School of Mathematics,  
The University of Bristol,  
Howard House, Queens Avenue,  
Bristol, BS8 1TW, (UK);  
and the Heilbronn Institute  
for Mathematical Research,  
Bristol, (UK)  
[khughes.math@gmail.com](mailto:khughes.math@gmail.com)