Proceedings of the Edinburgh Mathematical Society (2023) **66**, 923–939
doi:10.1017/S0013091523000366

DISCRETE RESTRICTION ESTIMATES FOR FORMS IN MANY VARIABLES

BRIAN COOK[1], KEVIN HUGHES[2,3] AND EYVINDUR PALSSON[1]

[1] *Department of Mathematics, Virginia Tech, Blacksburg, VA, USA*
(briancookmath@gmail.com; palsson@vt.edu)
[2] *School of Mathematics, The University of Bristol, Bristol, UK*
[3] *The Heilbronn Insitute for Mathematical Research, Bristol, UK*
(khughes.math@gmail.com)

(Received 16 December 2021)

Abstract We prove discrete restriction estimates for a broad class of hypersurfaces arising in seminal work of Birch. To do so, we use a variant of Bourgain's arithmetic version of the Tomas–Stein method and Magyar's decomposition of the Fourier transform of the indicator function of the integer points on a hypersurface.

Keywords: discrete restriction estimates; the circle method; Diophantine equations in many variables

2020 Mathematics subject classification: Primary 11D72; 11P55; 42B05; 42B25; 11L07

1. Introduction

In this paper, we consider discrete restriction estimates associated to integral, positive definite forms. Recall that a form is a homogeneous polynomial, integral means that the coefficients of this polynomial are integers and positive definite means that $\mathcal{Q}(\boldsymbol{x}) > 0$ for $\boldsymbol{x} \neq 0$. The positive definite criterion guarantees that the form is nondegenerate. Let $\mathcal{Q}(\boldsymbol{x}) \in \mathbb{Z}[\boldsymbol{x}]$ be such a form, where $\boldsymbol{x} = (x_1, x_2, \dots, x_d)$ with $d \geq 2$, and k denotes the degree of the form $\mathcal{Q}$. We always assume that $k \geq 2$. For each $\lambda \in \mathbb{R}$, the polynomial $\mathcal{Q}$ cuts out a real variety $V_{\mathcal{Q}=\lambda}(\mathbb{R}) := \{\boldsymbol{x} \in \mathbb{R}^d : \mathcal{Q}(\boldsymbol{x}) = \lambda\}$ containing a discrete set of integral points $V_{\mathcal{Q}=\lambda}(\mathbb{Z}) := \{\boldsymbol{x} \in \mathbb{Z}^d : \mathcal{Q}(\boldsymbol{x}) = \lambda\}$; either or both of these sets are possibly empty depending on the value of λ. For instance, $V_{\mathcal{Q}=\lambda}(\mathbb{R})$ is empty for negative λ since $\mathcal{Q}$ is positive definite, and $V_{\mathcal{Q}=\lambda}(\mathbb{Z})$ is empty for non-integral values of λ.

https://doi.org/10.1017/S0013091523000366 Published online by Cambridge University Press

In our discussion, we always consider a fixed form $\mathcal{Q}$. So, we suppress it from the notation below. For $\lambda \in \mathbb{N}$ and functions $a : \mathbb{Z}^d \to \mathbb{C}$, define the arithmetic extension operator

$$E_\lambda a(\boldsymbol{\xi}) := \sum_{\boldsymbol{x} \in V_{\mathcal{Q}=\lambda}(\mathbb{Z})} a(\boldsymbol{x}) e(\boldsymbol{x} \cdot \boldsymbol{\xi}).$$

Letting $\omega_\lambda := \mathbf{1}_{V_{\mathcal{Q}=\lambda}(\mathbb{Z})}$, we have $E_\lambda a(\boldsymbol{\xi}) = \mathcal{F}_{\mathbb{Z}^d}(a \cdot \omega_\lambda)(\boldsymbol{\xi})$, where $\mathcal{F}_{\mathbb{Z}^d}$ is the Fourier transform defined on complex-valued functions with domain $\mathbb{Z}^d$. In other words, E_λ is the adjoint to the restriction operator $R_\lambda f$ defined as $R_\lambda f := \mathcal{F}_{\mathbb{T}^d}(f) \cdot \omega_\lambda$ for functions $f : \mathbb{T}^d \to \mathbb{C}$. The extension operator is trivial when the variety has no integer points; that is, when $V_{\mathcal{Q}=\lambda}(\mathbb{Z})$ is the empty set. Consequently, we are interested in situations where the variety has many integer points. The prototypical examples here are spheres (centered at the origin) in five or more variables. Here the form is given by the sum of squares $x_1^2 + \cdots + x_d^2$, and the cardinality of $V_{\mathcal{Q}=\lambda}(\mathbb{Z})$ has order of magnitude $\lambda^{\frac{d}{2}-1}$ for $\lambda \in \mathbb{N}$. According to a theorem of Birch, there is a natural setting for these operators, which we review here.

Define the *Birch singular locus* of the form $\mathcal{Q}$ as the complex variety

$$V_{\mathcal{Q}}^{\dagger}(\mathbb{C}) := \{\boldsymbol{x} \in \mathbb{C}^d : \nabla \mathcal{Q}(\boldsymbol{x}) = \mathbf{0}\}.$$

Let $\dim_{\mathbb{C}}(V)$ denote the algebraic dimension of a complex variety V. We will say that an integral form is *regular* if it satisfies Birch's criterion:

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

We define the *Birch rank*, $B(\mathcal{Q})$ of a form $\mathcal{Q}$, to be the co-dimension $d - \dim_{\mathbb{C}}(V_{\mathcal{Q}}^{\dagger}(\mathbb{C}))$. The Birch rank is always non-negative since $V_{\mathcal{Q}}^{\dagger}(\mathbb{C})$ being a variety in $\mathbb{C}^d$ implies that $\dim_{\mathbb{C}}(V_{\mathcal{Q}}^{\dagger}(\mathbb{C})) \leq d$. To justify the term 'rank', one should note that this generalizes the notion of rank for quadratic forms. Indeed, for a quadratic form $\mathcal{Q}(\boldsymbol{x}) := \boldsymbol{x} M \boldsymbol{x}^{\mathrm{T}}$ defined by some $d \times d$-matrix M, a simple calculation gives $B(\mathcal{Q}) = \operatorname{rank}(M)$. Here, and in related examples, the point $\boldsymbol{x} \in \mathbb{Z}^d$ is regarded as a row vector of length d and $\boldsymbol{x}^{\mathrm{T}}$ is its transpose.

When Equation (1) is satisfied, Birch [2] tells us that there exists an infinite arithmetic progression $\Gamma_{\mathcal{Q}}$ in $\mathbb{N}$ depending on the form $\mathcal{Q}$ such that for each $\lambda \in \Gamma_{\mathcal{Q}}$, there exists a positive constant $C_{\mathcal{Q}}(\lambda)$ with the property that

$$N_{\mathcal{Q}}(\lambda) := \#\{\boldsymbol{n} \in \mathbb{Z}^d : \mathcal{Q}(\boldsymbol{n}) = \lambda\} = C_{\mathcal{Q}}(\lambda)\lambda^{\frac{d}{k}-1} + O_{\mathcal{Q}}(\lambda^{\frac{d}{k}-1-\delta}) > 0 \tag{2}$$

for some positive δ depending on the form $\mathcal{Q}$. Moreover, there exists constants $c_2 > c_1 > 0$ such that $c_1 \leq C_{\mathcal{Q}}(\lambda) \leq c_2$ for all $\lambda \in \Gamma_{\mathcal{Q}}$. Based on Birch's asymptotic Equation (2) and on the usual heuristics of the circle method, one expects the following estimates.

Conjecture 1. *Let $\mathcal{Q}$ be an integral, positive definite form of degree $k \geq 2$ in $d > 2k$ variables. For each $1 \leq p \leq \infty$ and $\epsilon > 0$, there exists a positive constant $C_{\mathcal{Q},p,\epsilon}$ such that*

$$\|E_\lambda a\|_{L^p(\mathbb{T}^d)} \leq C_{\mathcal{Q},p,\epsilon}\lambda^{\epsilon}(1 + \lambda^{\frac{d-k}{2k} - \frac{d}{kp}})\|a\|_{\ell^2(\mathbb{Z}^d)}. \tag{3}$$

https://doi.org/10.1017/S0013091523000366 Published online by Cambridge University Press

For $k \geq 3$, we further conjecture that one may remove the ϵ-loss; that is, for each $1 \leq p \leq \infty$, there exists a constant $C_{\mathcal{Q},p}$ such that

$$\|E_\lambda a\|_{L^p(\mathbb{T}^d)} \leq C_{\mathcal{Q},p}\left(1 + \lambda^{\frac{d-k}{2k} - \frac{d}{kp}}\right)\|a\|_{\ell^2(\mathbb{Z}^d)}. \tag{4}$$

There are two trivial estimates known for Conjecture 1. The first trivial estimate is the $\ell^2 \to L^2$ estimate, which is furnished by Plancherel's theorem. The second trivial estimate is the $\ell^2 \to L^\infty$ estimate, which is furnished by the Cauchy–Schwarz inequality and Equation (2) when the latter is known to hold. Conjecture 1 has been intensively studied in the quadratic case, especially for the spherical case $\mathcal{Q}(\boldsymbol{x}) := x_1^2 + \cdots + x_d^2$. Even for the sphere, this problem remains open despite major recent advances in the area. See [4–9] for more information regarding the spherical case and [10] for other quadratic hypersurfaces. In contrast, for forms of higher degree, there are no hitherto known non-trivial estimates towards this problem.

Our result is an affirmative answer to Conjecture 1 when the form is also assumed to be regular, and it yields Equation (4) when p and d are both sufficiently large. In particular, p will be much larger than the critical exponent $p_c = p_c(\mathcal{Q}) := \frac{2d}{d-k}$. (The critical exponent is defined as the exponent p where the two summands in Equation (3) or (4) balance. Supercritical p means that $p > p_c$, while subcritical p means that $p < p_c$.) To state our result, we introduce a relevant parameter. For a regular, integral form $\mathcal{Q}$ of degree k in d variables, define the parameter

$$\gamma_{\mathcal{Q}} := \frac{1}{6k}\left(\frac{d - \dim(V_{\mathcal{Q}}^\dagger(\mathbb{C}))}{(k-1)2^k} - 1\right). \tag{5}$$

Throughout we assume that d is sufficiently large with respect to k to satisfy the regularity criterion (1). This implies that $\gamma_{\mathcal{Q}} > 0$ and $d > 2k$. Our main result is the following.

Theorem 1. *Let $\mathcal{Q}$ be a regular, positive definite integral form in d variables of degree $k \geq 2$. If $p > 2 + \frac{2k}{\gamma_{\mathcal{Q}}}$, then Equation (4) holds for $\lambda \in \mathbb{N}$.*

We take a moment to orient ourselves with a few examples to record what Theorem 1 gives for these examples and to compare it with known bounds when applicable.

1.1. Spheres

For the form $\mathcal{Q}(\boldsymbol{x}) := |\boldsymbol{x}|^2$, its singular locus is $V_{\mathcal{Q}}^\dagger(\mathbb{C}) = \{\nabla\mathcal{Q}(\boldsymbol{x}) = 2\boldsymbol{x} = 0\}$ is $\{\boldsymbol{0}\}$. Therefore, the dimension of the singular locus is 0, and we require that $d - 0 > (2-1)2^2$. More simply, we require that $d \geq 5$ for spheres. Under this assumption on the dimension, $\gamma_{\mathcal{Q}} = (d-4)/48$ and Theorem 1 implies that the supercritical extension estimate Equation (4) holds for $p > 2 + 192/(d-4)$. This range of p is far away from the conjectured critical exponent of $2 + 4/(d-2)$. Fortunately, in this case, one may replace $\gamma_{\mathcal{Q}}$ (in Equation (13) below) by $(d-1)/4 + \epsilon$ for any $\epsilon > 0$ and $d \geq 4$ from [25]. In turn, this replacement improves the range of p in Theorem 1 to all $p > 2 + 8/(d-3)$ for $d \geq 4$. This recovers the bounds obtained in [4] but falls short of their subsequent improvements obtained in [6–9].

https://doi.org/10.1017/S0013091523000366 Published online by Cambridge University Press

1.2. Ellipsoids

Suppose that M is an invertible $d \times d$ matrix with integral coefficients such that the associated quadratic form $\mathcal{Q}(\boldsymbol{x}) := \boldsymbol{x}M\boldsymbol{x}^{\mathrm{T}}$ is positive definite. Spheres corresponding to M being the identity matrix. Since M is invertible, $V_{\mathcal{Q}}^{\dagger}(\mathbb{C}) = \{\mathbf{0}\}$, and Theorem 1 implies that the supercritical extension estimate Equation (4) holds for $p > 2+192/(d-4)$. While we are unaware of any results in this level of generality for quadratic forms, presumably, a more delicate approach following [4, 18] would yield a bound closer to the critical exponent.

1.3. *k*-Spheres

For the form $\mathcal{Q}(\boldsymbol{x}) := x_1^k + \cdots + x_d^k$, where k is an integer greater than 1, its singular locus is $V_{\mathcal{Q}}^{\dagger}(\mathbb{C}) = \{\nabla\mathcal{Q}(\boldsymbol{x}) = k(x_1^{k-1}, \ldots, x_d^{k-1}) = \mathbf{0}\} = \{\mathbf{0}\}$. Therefore, the dimension of the singular locus is 0, and we require that $d > (k-1)2^k$. Under this assumption on the dimension, $\gamma_{\mathcal{Q}} = (d-(k-1)2^k)/(6k(k-1)2^k)$, and Theorem 1 implies that the supercritical extension estimate Equation (4) holds for $p > 2 + (12k^2(k-1)2^k)/(d-(k-1)2^k)$. When $k = 3$, this becomes $d > 16$ and $p > 2 + 1728/(d-16)$.

A peculiar feature of Birch's method - and hence our results - is that the Birch rank is defined in terms of the *complex points* of the singular locus rather than its *real points*. Recall Euler's theorem: for any form $\mathcal{Q}$, we have the identity

$$\deg(\mathcal{Q})\mathcal{Q}(x_1, \ldots, x_d) = (x_1, \ldots, x_d) \cdot \nabla\mathcal{Q}(x_1, \ldots, x_d),$$

where the $\cdot$ on the right hand side denotes the inner product of two vectors. By Euler's theorem, a *real* singular point for a positive definite form is necessarily $\mathbf{0}$. In other words, $V_{\mathcal{Q}}^{\dagger}(\mathbb{R}) = \{\mathbf{0}\}$ for every positive definite form $\mathcal{Q}$. In contrast, the Birch singular locus can be huge as seen in the following 'non-example' of a positive definite form whose singular locus is too large for our theorem and methods to be applicable.

1.4. A non-example

Consider the form $\mathcal{Q}(\boldsymbol{x}) := (x_1^2 + \cdots + x_d^2)^2$; its Birch singular locus is

$$V_{\mathcal{Q}}^{\dagger}(\mathbb{C}) = \{\boldsymbol{x} \in \mathbb{C}^d : 4(x_1^2 + \cdots + x_d^2)\boldsymbol{x} = \mathbf{0}\} = \{\boldsymbol{x} \in \mathbb{C}^d : x_1^2 + \cdots + x_d^2 = \mathbf{0}\}.$$

This is a co-dimension 1 complex algebraic set. Consequently, this form fails to satisfy Birch's regularity criterion (1) regardless of how large d (the number of variables) is. Meanwhile, its real singular locus $V_{\mathcal{Q}}^{\dagger}(\mathbb{R})$ is the set $\{\mathbf{0}\}$. When λ is a square, the corresponding restriction operator is closely related to that for the form $x_1^2 + \cdots + x_d^2$. When λ is not a square, the behaviour of the corresponding restriction operator is subtle.

Remark 1.1. The expert can formulate conjectures analogous to Conjecture 1 for integral forms and their 0-level sets without difficulty. Practically, this presents only a technical difference from our hypothesis in Theorem 1 that forms are positive definite.

https://doi.org/10.1017/S0013091523000366 Published online by Cambridge University Press

Our methods also apply in this setting, but we do not pursue the analogous results in this paper.

Having considered a few examples, let us now discuss our motivations. One motivation is to extend discrete restriction theory for hypersurfaces beyond the setting of spheres and paraboloids. This is the first attempt to do so. This work fits into a broader program, initiated by Magyar in [24], which seeks to understand discrete (more appropriately termed 'arithmetic') harmonic analysis for hypersurfaces. Initial forays into this program have centered around Birch's theorem and have had applications to maximal functions and ergodic theorems [24] and [12], discrepancy estimates in [25], Szemeredi-type theorems in [26] and ℓ^p-improving estimates in [21].

Our approach to Theorem 1 is motivated by a previously open question. This question, posed by the second author in May 2014 at the Workshop III: Kakeya problem, Restriction Problem and Sum-product Theory Workshop as part of IPAM's long program Algebraic Techniques for Combinatorial and Computational Geometry, asks: *Can one use Magyar–Stein–Wainger's decomposition of the surface measure for the integer points on a sphere from [27] to improve the discrete restriction estimates for the sphere?*

This question was natural since Magyar–Stein–Wainger's decomposition had been successfully used in the aforementioned works of Magyar, but at that time, it was unknown if Magyar–Stein–Wainger's decomposition could be used to prove non-trivial discrete restriction estimate for the sphere. Our proof of Theorem 1 reveals that Magyar–Stein–Wainger's decomposition can be used to prove non-trivial discrete restriction estimates for the sphere. Examining [4], the second author's question means: *What is the best way to control the error term in the decomposition?*

This latter question closely relates to another question, posed by Ákos Magyar at the *Georgia Discrete Analysis Conference* in May 2018, which asks: *how does one incorporate minor arc estimates for higher degree Diophantine equations in order to obtain discrete restriction estimates?* At that time, no discrete restriction estimates were known for a *single* degree 3 or higher multivariate form. Magyar's question was natural given the fact that for quadratic forms one does not need to use minor arcs but one must grapple with the minor arcs for hypersurfaces of degrees 3 or more. This relates to the first question because the minor arcs contribute the greatest error term in the decomposition formulas for hypersurfaces of degrees 3 and more. In the quadratic cases, there is no need for minor arcs, and they have not made an appearance in previous analyses.

It transpired that Magyar's question was partially answered in [17] where minor arc estimates were incorporated to prove discrete restriction estimates for 'k-paraboloids'. While [17] were predominately interested in ϵ-removal lemmas, the methods therein also used minor arc estimates to prove discrete restriction estimates. When one observes that the worst error term (in Magyar's generalizations of Magyar–Stein–Wainger's decomposition) arises from the minor arcs, the natural strategy becomes to adapt those methods to handle the other error terms. Our work answers these questions by successfully using this strategy.

We organized our argument to closely follow [17] so that Theorem 1 reduces to proving appropriate estimates for the main term and the error term, but we streamline the approach to fit our purposes. In particular, since we are interested in the sharp discrete restriction estimates, our approach 'bakes in' the ϵ-removal. The bounds for the error

https://doi.org/10.1017/S0013091523000366 Published online by Cambridge University Press

terms are taken from [24]. Meanwhile, the bulk of our work lies in handling the main term. This is done in Theorem 2 where we prove a dyadically refined decomposition of the main term, which is better suited to our purposes. Outlined in Section 4, this refinement is this paper's main technical contribution, which allows us to adapt the Tomas–Stein method in [4] to the main term.

Instead of striving to fully optimize every aspect of our argument, we have aimed to give a simplified version of the general method, which hopefully illuminates the main ideas. The main bottleneck in our argument is the poor state of knowledge for minor arcs bounds in Equation (13) that leads us to define $\gamma_{\mathcal{Q}}$ in Equation (5). Any better decay rate of Equation (13) (e.g., replacing $\gamma_{\mathcal{Q}}$ therein by a larger constant) immediately enlarges the range of p in Theorem 1. For example, one can improve the ranges of d and p in Theorem 1 for k-spheres by using superior minor arc estimates available in this case. Such estimates are possible by exploiting the diagonal structure of the underlying Diophantine equation; see [25] when $k=2$ and [1] when $k \geq 3$ for the best bounds presently known.

There has been much recent progress on decoupling estimates for affine-invariant systems of equations in many variables following [9, 11, 29]. (Affine invariance is also known as translation-dilation invariance or as parabolic rescaling.) For instance, see [14–16]. It is important to note the setting of Theorem 1 is *far* from affine invariant. By combining [28] and [23], there is a different way to use such decoupling results to prove Equation (3) for sufficiently large $p > p_c$. However, this procedure almost surely yields a smaller range of p than Theorem 1 provides, and it becomes increasingly worse as the degree or number of variables increases. Moreover, another method must be used to obtain the sharper estimate (4). The only known way to sharpen an estimate from Equation (3) to (4) is a circle method approach like the one used in the proof of Theorem 1.

1.5. Organization of the paper

The paper is organized as follows. In § 2, we set notation used throughout the paper. In § 3, we give an abstract formulation of Tomas's method for discrete L^2 restriction theorems dating to [3]; Lemma 1 therein reduces our problem to proving estimates related to the Fourier transform of the surface measure. In § 4, we recall a decomposition of the surface measure due to Magyar and related estimates from [24]; this is 'Magyar's Decomposition Theorem'. Combining Lemma 1 and Magyar's Decomposition Theorem, we reduce Theorem 1 to Theorem 2, which is an estimate for the major arcs. In § 5, we prove a bound for the major arc pieces by a further application of Tomas's methods.

2. Notation

We introduce here some notation that will streamline our exposition.

- For a positive integer, we let $\mathbb{Z}/q$ denote the group of integers modulo q and $U_q := \{1 \leq a < q : (a, q) = 1\}$ denote its unit group.
- We write $f(\lambda) \lesssim g(\lambda)$ if there exists a constant $C > 0$ independent of all λ under consideration (e.g., λ 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 $f(\lambda) \eqsim g(\lambda)$ if $f(\lambda) \lesssim$

$g(\lambda)$ and $f(\lambda) \gtrsim g(\lambda)$. Subscripts in the above notation 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 $\mathrm{e}^{-2\pi i t}$ for $t \in \mathbb{R}$ or $\mathbb{T}$.
- For a function $f : \mathbb{Z}^d \to \mathbb{C}$, its $\mathbb{Z}^d$-Fourier transform will be denoted $\mathcal{F}_{\mathbb{Z}^d} f(\boldsymbol{\xi})$ for $\boldsymbol{\xi} \in \mathbb{T}^d$. For a function $f : \mathbb{T}^d \to \mathbb{C}$, its $\mathbb{T}^d$-Fourier transform will be denoted $\mathcal{F}_{\mathbb{T}^d} f(\boldsymbol{x})$ for $\boldsymbol{x} \in \mathbb{Z}^d$. $\mathcal{F}_{\mathbb{Z}^d}$ and $\mathcal{F}_{\mathbb{T}^d}$ are defined so that they are inverses of one another. For a function $f : \mathbb{R}^d \to \mathbb{C}$, its $\mathbb{R}^d$-Fourier transform will be denoted $\mathcal{F}_{\mathbb{R}^d} f(\boldsymbol{x})$ for $\boldsymbol{x} \in \mathbb{R}^d$.
- For a function $f : \mathbb{R}^d \to \mathbb{C}$, we define dilation operator D_t by $\mathrm{D}_t f(\boldsymbol{x}) = f(\boldsymbol{x}/t)$.
- For a ring R, we will use the inner product notation $\boldsymbol{b} \cdot \boldsymbol{m}$ for vectors $\boldsymbol{b}, \boldsymbol{m} \in R^d$ 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 $q \in \mathbb{N}$.
- We also let $\mathbf{1}_X$ denote the indicator function of the set X.

3. The arithmetic Tomas–Stein method

Let ω_λ be the counting measure on $V_{\mathcal{Q}=\lambda}(\mathbb{Z})$ for a single integral, positive definite, homogenous form $\mathcal{Q}$ satisfying (1) and some $\lambda \in \mathbb{Z}$. Let $F = \mathcal{F}_{\mathbb{Z}^d}(\omega_\lambda)$ be the exponential sum corresponding to ω_λ. A common approach to problems involving ω_λ is to use the circle method so as to decompose the exponential sum F into a main piece $F_{\mathfrak{M}}$ and an error term $F_{\mathfrak{m}}$ corresponding, respectively, to major and minor arcs. (These are analogous to low and high frequency pieces, respectively.) To prove discrete restriction estimates, Bourgain in [4] combined this approach with Tomas's L^2 restriction argument in order to reduce matters to the following two estimates:

- Bounds for the operator given by convolution with the major arc operator $F_{\mathfrak{M}}$, and
- A uniform power saving bound on the minor arc piece $F_{\mathfrak{m}}$.

See [19, 20] for a variant. Bourgain's approach has been abstracted in [22] and [17]. We combine Lemmas 3.3 and 3.6 from [17] to form the following lemma.

Lemma 1. *For $\lambda \in \mathbb{N}$, let $F = \mathcal{F}_{\mathbb{Z}^d}(\omega_\lambda)$ be the $\mathbb{Z}^d$-Fourier transform of the arithmetic surface measure ω_λ defined on $V_{\mathcal{Q}=\lambda}(\mathbb{Z})$. Suppose that there exists a decomposition $F = F_{\mathfrak{M}} + F_{\mathfrak{m}}$ such that for each $f \in L^\infty(\mathbb{T}^d)$, we have the estimates*

$$\|F * f\|_{L^{p_0}(\mathbb{T}^d)} \lesssim \lambda^\epsilon \|f\|_{L^{p_0'}(\mathbb{T}^d)} \quad \text{for some } p_0 \le p_c, \tag{TS1}$$

$$\|F_{\mathfrak{M}} * f\|_{L^{p_1}(\mathbb{T}^d)} \lesssim \lambda^{\frac{d}{k}-1-\frac{2d}{kp_1}} \|f\|_{p_1'} \quad \text{for some } p_1 > p_c, \text{ and} \tag{TS2}$$

https://doi.org/10.1017/S0013091523000366 Published online by Cambridge University Press

$$\|F_{\mathfrak{m}}\|_{L^\infty(\mathbb{T}^d)} \lesssim \lambda^{\frac{d}{k}-1-\frac{\zeta}{k}} \quad \textit{for some } \zeta \in (0, d-k). \tag{TS3}$$

Then $\|F * f\|_{L^p(\mathbb{T}^d)} \lesssim \lambda^{\frac{d}{k}-1-\frac{2d}{kp}} \|f\|_{L^{p'}(\mathbb{T}^d)}$ *holds for* $p > \max\left[p_1, \frac{2d-(d-k)p_0}{\zeta} + p_0\right]$.

In our work, we only use Plancherel's theorem to exploit the subcritical estimate at $p_0 = 2$; this gives the exponent $p > \max\left[p_1, \frac{2d-(d-k)2}{\zeta} + 2\right] = \max\left[p_1, \frac{2k}{\zeta} + 2\right]$. We give the proof of Lemma 1 for completeness.

Proof of Lemma 1. Set $N = \lceil \lambda^{1/k} \rceil$. Fix $p > \max\left[p_1, \frac{2d-(d-k)p_0}{\zeta} + p_0\right]$ and let a be an element of ℓ^2. For notational convenience, we let E denote the extension operator defined on sequences $a : \mathbb{Z}^d \to \mathbb{C}$ by $Ea := \mathcal{F}_{\mathbb{Z}^d}(\omega_\lambda \cdot \mathcal{F}_{\mathbb{T}^d} a) = a * \mathcal{F}_{\mathbb{Z}^d}(\omega_\lambda)$. We may assume that a is not identically zero and by homogeneity normalize a so that $\|a\|_2 = 1$. We introduce a parameter $\alpha > 0$ in order to define the level sets and functions

$$S_\alpha = \{\boldsymbol{\xi} \in \mathbb{T}^d : |Ea(\boldsymbol{\xi})| \geq \alpha\} \quad \text{and} \quad f = 1_{S_\alpha} \frac{Ea}{|Ea|}.$$

By the Cauchy–Schwarz inequality and Birch's theorem in [2], we have

$$\|Ea\|_{L^\infty} \lesssim N^{\frac{d-k}{2}}. \tag{6}$$

Therefore, we may restrict α to lie in the interval $[0, CN^{\frac{d-k}{2}}]$ for some positive constant C. By Parseval's identity, we have

$$\alpha |S_\alpha| \leq \langle f, Ea \rangle = \langle \mathcal{F}_{\mathbb{T}^d} f, \omega_\lambda \cdot a \rangle = \langle \omega_\lambda \cdot \mathcal{F}_{\mathbb{T}^d} f, a \rangle.$$

By Cauchy–Schwarz and the assumption $\|a\|_2 = 1$, it follows that

$$\alpha^2 |S_\alpha|^2 \leq \|(\mathcal{F}_{\mathbb{T}^d} f)\omega_\lambda\|_{\ell^2}^2 = \langle (\mathcal{F}_{\mathbb{T}^d} f) \cdot \omega_\lambda, \mathcal{F}_{\mathbb{T}^d} f \rangle.$$

Another application of Parseval's identity implies that

$$\alpha^2 |S_\alpha|^2 \leq \langle f * F, f \rangle. \tag{7}$$

By Equation (7), Hölder's inequality and hypotheses (TS2) and (TS3) of the lemma, we have

$$\begin{aligned} \alpha^2 |S_\alpha|^2 &\leq \|f * F_{\mathfrak{M}}\|_{p_1} \|f\|_{p_1'} + \|f * F_{\mathfrak{m}}\|_\infty \|f\|_1 \\ &\lesssim N^{d-k-\frac{2d}{p_1}} \|f\|_{p_1'}^2 + \|F_{\mathfrak{m}}\|_\infty \|f\|_1^2 \\ &\lesssim N^{d-k-\frac{2d}{p_1}} |S_\alpha|^{\frac{2}{p_1'}} + N^{d-k-\zeta} |S_\alpha|^2. \end{aligned}$$

https://doi.org/10.1017/S0013091523000366 Published online by Cambridge University Press

Therefore, when $\alpha \gtrsim N^{\frac{d-k}{2}-\frac{\zeta}{2}}$, we have

$$\alpha^2 |S_\alpha|^2 \lesssim N^{d-k-\frac{2d}{p_1}} |S_\alpha|^{2-\frac{2}{p_1}}.$$

Rearranging implies that $|S_\alpha| \lesssim \alpha^{-p_1} N^{\frac{(d-k)p_1}{2}-d}$. Since $p > p_1$, we have

$$\begin{aligned}\int_{|Ea| \gtrsim N^{\frac{d-k}{2}-\frac{\zeta}{2}}} |Ea|^p \, \mathrm{d}m &= p \int_{CN^{\frac{d-k}{2}-\frac{\zeta}{2}}}^{CN^{\frac{d-k}{2}}} \alpha^{p-1} |S_\alpha| \, \mathrm{d}\alpha \\ &\lesssim N^{\frac{(d-k)p_1}{2}-d} \int_1^{CN^{\frac{d-k}{2}}} \alpha^{p-p_1-1} \, \mathrm{d}\alpha \\ &\lesssim N^{\frac{(d-k)p}{2}-d}.\end{aligned}$$

Altogether, we have

$$\int_{|Ea| \gtrsim N^{d/2-\zeta/2}} |Ea|^p \, \mathrm{d}m \lesssim N^{\frac{(d-k)p}{2}-d}. \tag{8}$$

We are left to consider that the regime where $|Ea| \lesssim N^{\frac{d-k}{2}-\frac{\zeta}{2}}$. We now make use of estimate (TS1) at the exponent p_0 to handle the regime where $|Ea| \lesssim N^{\frac{d-k}{2}-\frac{\zeta}{2}}$. This is possible by the trivial bound (6) as follows:

$$\int_{|Ea| \lesssim N^{\frac{d-k}{2}-\frac{\zeta}{2}}} |Ea|^p \, \mathrm{d}m \lesssim \left(N^{\frac{d-k}{2}-\frac{\zeta}{2}}\right)^{p-p_0} \int_{\mathbb{T}^r} |Ea|^{p_0} \, \mathrm{d}m \lesssim_\epsilon N^{\frac{(d-k-\zeta)(p-p_0)}{2}+\epsilon}.$$

Combining this estimate with Equation (8), we have that

$$\begin{aligned}\int |Ea|^p \, \mathrm{d}m &= \int_{|Ea| \lesssim N^{\frac{d-k}{2}-\frac{\zeta}{2}}} |Ea|^p \, \mathrm{d}m + \int_{|Ea| \gtrsim N^{\frac{d-k}{2}-\frac{\zeta}{2}}} |Ea|^p \, \mathrm{d}m \\ &\lesssim_\epsilon N^{\frac{(d-k)p}{2}-d} + N^{\frac{(d-k-\zeta)(p-p_0)}{2}+\epsilon}.\end{aligned}$$

The latter summand is dominated by the former summand when $\frac{(d-k-\zeta)(p-p_0)}{2} < \frac{(d-k)p}{2} - d$. This is equivalent to

$$\frac{(d-k-\zeta)(p-p_0)}{2} = \frac{(d-k)p}{2} - \frac{\zeta p}{2} - \frac{(d-k-\zeta)p_0}{2} < \frac{(d-k)p}{2} - d,$$

which is equivalent to

$$\frac{\zeta p}{2} + \frac{(d-k-\zeta)p_0}{2} > d.$$

https://doi.org/10.1017/S0013091523000366 Published online by Cambridge University Press

Rearranging this last expression, we find that we need

$$p > \frac{2}{\zeta}\left(d - \frac{(d-k-\zeta)p_0}{2}\right) = \zeta^{-1}(2d-(d-k-\zeta)p_0) = \frac{2d-(d-k)p_0}{\zeta} + p_0.$$

This is precisely the range of $p > \frac{2d-(d-k)p_0}{\zeta} + p_0$. □

4. Magyar's decomposition of the surface measure

Let $\mathcal{Q}(\boldsymbol{x}) \in \mathbb{Z}[\boldsymbol{x}]$ be an integral, positive definite form where $\boldsymbol{x} = (x_1, \ldots, x_d)$. The heavy lifting in our theorem lies in a decomposition of Magyar for the surface measure $\omega_\lambda := \mathbf{1}_{\{\boldsymbol{x}\in\mathbb{Z}^d:\mathcal{Q}(\boldsymbol{x})=\lambda\}}$, where $\lambda \in \mathbb{Z}$; this is the counting measure on the integer points $\boldsymbol{x}$ in $\mathbb{Z}^d$ such that $\mathcal{Q}(\boldsymbol{x}) = \lambda$. To state this theorem, we need to introduce a few objects.

For $q \in \mathbb{N}$, $a \in U_q$ and $\boldsymbol{m} \in \mathbb{Z}^d$, define the normalized Birch–Weyl sums

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

We have the bound

$$|G_{\mathcal{Q}}(a, q; \boldsymbol{m})| \lesssim_\epsilon q^{\epsilon-\kappa_{\mathcal{Q}}} \quad \text{for all } \epsilon > 0 \tag{9}$$

uniformly in $a \in U_q$ and $\boldsymbol{m} \in \mathbb{Z}^d$ with

$$\kappa_{\mathcal{Q}} := \frac{d - \dim V_{\mathcal{Q}}(\mathbb{C})}{2^{k-1}(k-1)}.$$

See [24] for a proof of this fact. The dimension d is sufficiently large so that $\kappa_{\mathcal{Q}} > 2$.

Let $d\sigma_{\mathcal{Q}}$ denote the singular measure on $\mathbb{R}^d$ defined as the Gelfand–Leray form whose $\mathbb{R}^d$-Fourier transform is defined distributionally by the oscillatory integral

$$\int_{\mathbb{R}} e(t(\mathcal{Q}(\boldsymbol{x}) - 1))\ \mathrm{d}t.$$

It is known that

$$\mathrm{d}\sigma_{\mathcal{Q}}(\boldsymbol{x}) = \mathrm{d}S_{\mathcal{Q}}(\boldsymbol{x})/|\nabla\mathcal{Q}(\boldsymbol{x})|, \tag{10}$$

where $\mathrm{d}S_{\mathcal{Q}}$ is the Euclidean surface area measure on the hypersurface $\{\boldsymbol{x} \in \mathbb{R}^d : \mathcal{Q}(\boldsymbol{x}) = 1\}$. These measures are compactly supported since $\mathcal{Q}$ is positive definite. We cite the following bound – see Lemma 6 on page 931 of [24] – for the $\mathbb{R}^d$-Fourier transform of the surface measure:

$$|\widetilde{d\sigma_{\mathcal{Q}}}(\boldsymbol{\xi})| \lesssim_\epsilon (1 + |\boldsymbol{\xi}|)^{1-\kappa_{\mathcal{Q}}+\epsilon} \quad \text{for each} \quad \boldsymbol{\xi} \in \mathbb{R}^d \quad \text{and for all } \epsilon > 0. \tag{11}$$

Let Ψ be a $C^\infty(\mathbb{R}^d)$ bump function supported in the cube $[-1/8, 1/8]^d$ and 1 on the cube $[-1/16, 1/16]^d$, where these cubes are regarded as subsets of the torus $\mathbb{T}^d$. For each

https://doi.org/10.1017/S0013091523000366 Published online by Cambridge University Press

$q \in \mathbb{N}$, let s be the integer such that $2^s \leq q < 2^{s+1}$. For such q and for $a \in U_q$, define the Fourier multipliers

$$\mu_\lambda^{a/q}(\boldsymbol{\xi}) := \sum_{\boldsymbol{m} \in \mathbb{Z}^d} G_{\mathcal{Q}}(a, q; \boldsymbol{m}) \Psi\left(2^s \left[\boldsymbol{\xi} - \frac{\boldsymbol{m}}{q}\right]\right) \widetilde{\mathrm{d}\sigma_{\mathcal{Q}}}\left(\lambda^{\frac{1}{k}}\left[\boldsymbol{\xi} - \frac{\boldsymbol{m}}{q}\right]\right)$$

for $\xi \in \mathbb{T}^d$. Generalizing work of [24], Magyar [27] obtained a flexible decomposition of the surface measure; we choose the following form.

Magyar's Decomposition Theorem ([24, 27]). *Let $\mathcal{Q}(\boldsymbol{x}) \in \mathbb{Z}[\boldsymbol{x}]$ be a regular, positive definite integral form. For each $\lambda \in \mathbb{N}$, the Fourier transform of the surface measure ω_λ decomposes as*

$$\lambda^{1-\frac{d}{k}} \cdot \mathcal{F}_{\mathbb{Z}^d}\omega_\lambda(\boldsymbol{\xi}) = \left(\sum_{s=0}^{\lceil \log_2 \lambda^{1/k} \rceil} \sum_{q=2^s}^{2^{s+1}-1} e\left(-\frac{a\lambda}{q}\right) \sum_{a \in U_q} \mu_\lambda^{a/q}(\boldsymbol{\xi})\right) + \varepsilon_\lambda(\boldsymbol{\xi}), \tag{12}$$

where

$$\|\varepsilon_\lambda\|_{L^\infty(\mathbb{T}^d)} \lesssim_{\mathcal{Q},\epsilon} \lambda^{\epsilon - \gamma_{\mathcal{Q}}} \quad \text{for all } \epsilon > 0. \tag{13}$$

Remark 4.1. Our form of the error term ε_λ and its estimate (13) do not explicitly appear in [24]. We outline the differences and how to prove this form of Magyar's Decomposition Theorem. Recall that Magyar's main term takes the shape as the Fourier multiplier

$$\sum_{q \in \mathbb{N}} \sum_{a \in U_q} e\left(\frac{-a\lambda}{q}\right) \sum_{\boldsymbol{m} \in \mathbb{Z}^d} G_{\mathcal{Q}}(a, q; \boldsymbol{m}) \Psi(q\boldsymbol{\xi} - \boldsymbol{m}) \widetilde{\mathrm{d}\sigma_{\mathcal{Q}}}\left(\lambda^{1/k}\left[\boldsymbol{\xi} - \frac{\boldsymbol{m}}{q}\right]\right). \tag{14}$$

The first notable difference is that we have dyadically refined the decomposition so that Equation (14) becomes

$$\sum_{s=0}^{\infty} \sum_{q=2^s}^{2^{s+1}-1} \sum_{a \in U_q} e\left(-\frac{a\lambda}{q}\right) \sum_{\boldsymbol{m} \in \mathbb{Z}^d} G_{\mathcal{Q}}(a, q; \boldsymbol{m}) \Psi\left(2^s\left[\boldsymbol{\xi} - \frac{\boldsymbol{m}}{q}\right]\right) \widetilde{\mathrm{d}\sigma_{\mathcal{Q}}}\left(\lambda^{1/k}\left[\boldsymbol{\xi} - \frac{\boldsymbol{m}}{q}\right]\right). \tag{15}$$

This modifies the analysis of Equations (2.15) and (2.16) of Proposition 4 in [24] in inconsequential ways since $2^s \leq q < 2^{s+1}$. In particular, this preserves the estimate (13). The second notable difference is that we truncated the sum over $q \in \mathbb{N}$. Following the analysis of Equation (2.17) of Proposition 4 in [24], we may truncate Equation (15) to

$$\sum_{s=0}^{\lfloor \log_2 \lambda^{1/k} \rfloor} \sum_{q=2^s}^{2^{s+1}-1} \sum_{a \in U_q} e\left(-\frac{a\lambda}{q}\right) \sum_{\boldsymbol{m} \in \mathbb{Z}^d} G_{\mathcal{Q}}(a, q; \boldsymbol{m}) \Psi\left(2^s\left[\boldsymbol{\xi} - \frac{\boldsymbol{m}}{q}\right]\right) \widetilde{\mathrm{d}\sigma_{\mathcal{Q}}}\left(\lambda^{1/k}\left[\boldsymbol{\xi} - \frac{\boldsymbol{m}}{q}\right]\right) \tag{16}$$

https://doi.org/10.1017/S0013091523000366 Published online by Cambridge University Press

and place the difference into the error term ε_λ while maintaining the estimate (13). The expert may immediately verify this by using the Magyar–Stein–Wainger transference principle (see Section 2 of [27]) and Birch's Weyl bound (9).

The next theorem establishes (TS2) of Lemma 1; that is, we treat the major arc terms.

Theorem 2. *Let $\mathcal{Q}(\boldsymbol{x}) \in \mathbb{Z}[\boldsymbol{x}]$ be a positive definite, regular, integral form satisfying Equation (1) and $\lambda \in \mathbb{N}$. If $p > 2 + \frac{4}{\kappa_{\mathcal{Q}}-2}$, we have*

$$\|F_{\mathfrak{M}} * f\|_{L^p(\mathbb{T}^d)} \lesssim_p \lambda^{\frac{d-k}{k} - \frac{2d}{kp}} \|f\|_{L^{p'}(\mathbb{T}^d)} \tag{17}$$

for each $\lambda \in \mathbb{N}$.

We may deduce Theorem 1 once Theorem 2 is proved as follows.

Proof of Theorem 1 assuming Theorem 2. Since $k \geq 2$, we have $2k/\gamma_{\mathcal{Q}} > 4/(\kappa_{\mathcal{Q}}-2)$, and Lemma 1 reduces Theorem 1 to applying the major arc bound in Theorem 2 and the (minor arc) bound for the error term (13). □

5. Proof of Theorem 2

Fix $\mathcal{Q}(\boldsymbol{x}) \in \mathbb{Z}[\boldsymbol{x}]$ a positive definite form of degree k satisfying Equation (1) and $\lambda \in \mathbb{N}$. Set $N = \lceil \lambda^{1/k} \rceil$. Define the functions

$$\Psi_j(\boldsymbol{\xi}) := \Psi(2^j \boldsymbol{\xi}) - \Psi(2^{j+1}\boldsymbol{\xi}) \quad \text{for} \quad 0 \leq j < \lfloor \log_2 N \rfloor \text{ and}$$
$$\Psi_j(\boldsymbol{\xi}) := \Psi(2^j \boldsymbol{\xi}) \quad \text{for} \quad j = \lfloor \log_2 N \rfloor.$$

Furthermore, for $q \in \mathbb{N}, a \in U_q$ and $0 \leq j < \lfloor \log_2 N \rfloor$, define the multipliers

$$\mu_\lambda^{a/q,j}(\boldsymbol{\xi}) := \lambda^{\frac{d}{k}-1} \sum_{\boldsymbol{m} \in \mathbb{Z}^d} G_{\mathcal{Q}}(a,q;\boldsymbol{m}) \Psi_j\left(2^{s+1}\left[\boldsymbol{\xi} - \frac{\boldsymbol{m}}{q}\right]\right) \cdot \mathcal{F}_{\mathbb{R}^d}\, \mathrm{d}\sigma_{\mathcal{Q}}\left(\lambda^{1/k}\left[\boldsymbol{\xi} - \frac{\boldsymbol{m}}{q}\right]\right).$$

We will collect these multipliers according to the scale of their moduli; to do so, define, for each $s \geq 0$, the set of fractions

$$\mathcal{R}_s := \{a/q \in \mathbb{Q} : 2^s \leq q < 2^{s+1} \text{ and } a \in U_q\}.$$

Let $K_\lambda^{a/q,j} := \mathcal{F}_{\mathbb{T}^d}(\mu_\lambda^{a/q,j})$ denote the inverse Fourier transform of $\mu_\lambda^{a/q,j}$. We start our proof by establishing an identity for these kernels.

Proposition 1. *Let $\mathcal{Q}(\boldsymbol{x}) \in \mathbb{Z}[\boldsymbol{x}]$ be a positive definite, non-singular, integral form satisfying Equation (1) and $\Gamma_{\mathcal{Q}}$ be a set of regular values for the form $\mathcal{Q}$. If $s \geq 0$, then for each $a/q \in \mathcal{R}_s$, we have*

$$K_\lambda^{a/q,j}(\boldsymbol{x}) = e\left(a\mathcal{Q}(\boldsymbol{x})/q\right) \lambda^{-1} [\mathcal{F}_{\mathbb{R}^d}(\mathrm{D}_{2^s\lambda^{-1/k}}\, \Psi_j) * \mathrm{d}\sigma_{\mathcal{Q}}](\lambda^{-1/k}\boldsymbol{x}) \tag{18}$$

for all $\boldsymbol{x} \in \mathbb{Z}^d$.

https://doi.org/10.1017/S0013091523000366 Published online by Cambridge University Press

The proof of this proposition follows the proof of Proposition 1 in [21]; in that proof, one replaces Ψ by Ψ_j and q by 2^s.

Now that we know the structure of our kernel, we will use a circle method decomposition and a further Littlewood–Paley decomposition to arbitrage $L^1(\mathbb{T}^d) \to L^\infty(\mathbb{T}^d)$ and $L^2(\mathbb{T}^d) \to L^2(\mathbb{T}^d)$ estimates and deduce Theorem 2. These bounds are the content of the two following lemmas.

Lemma 2. *Let $\mathcal{Q}(\boldsymbol{x}) \in \mathbb{Z}[\boldsymbol{x}]$ be a positive definite, non-singular, integral form satisfying Equation (1) and $\lambda \in \mathbb{N}$. If $0 \le s \le \lfloor \log_2 N \rfloor$ and $a/q \in \mathcal{R}_s$, then each major arc piece $\mu_\lambda^{a/q,j}$ satisfies*

$$\|\mu_\lambda^{a/q,j}\|_{L^\infty(\mathbb{T}^d)} \lesssim_\epsilon 2^{j-s} 2^{j(\epsilon-\kappa)} \lambda^{\frac{d}{k}-\kappa} \quad \textit{for} \quad 0 \le j \le \lfloor \log_2 N \rfloor - s \tag{19}$$

and

$$\|\mu_\lambda^{a/q,j}\|_{L^\infty(\mathbb{T}^d)} \lesssim_\epsilon 2^{s(\epsilon-\kappa)} \lambda^{\frac{d}{k}-1} \quad \textit{for} \quad j = \lfloor \log_2 N \rfloor - s \tag{20}$$

for all $\epsilon > 0$.

Lemma 3. *Let $\mathcal{Q}(\boldsymbol{x}) \in \mathbb{Z}[\boldsymbol{x}]$ be a positive definite, non-singular, integral form satisfying Equation (1) and $\lambda \in \mathbb{N}$. If $0 \le s \le \lfloor \log_2 N \rfloor$ and $a/q \in \mathcal{R}_s$, then each major arc piece $\mu_\lambda^{a/q,j}$ satisfies*

$$\|\mathcal{F}_{\mathbb{T}^d} \mu_\lambda^{a/q,j}\|_{\ell^\infty(\mathbb{Z}^d)} \lesssim 2^{j+s} \lambda^{-1-\frac{1}{k}} \tag{21}$$

for $0 \le j \le \lfloor \log_2 N \rfloor - s$.

Remark 5.1. Note that $j + s = \lfloor \log_2 N \rfloor$ is the natural cutoff because we do not capture any oscillation in $\mathcal{F}_{\mathbb{R}^d} d\sigma(\lambda^{1/k}\boldsymbol{\xi})$ when $|\boldsymbol{\xi}| \lesssim \lambda^{-1/k}$.

Proof of Lemma 2. Fix $0 \le s \le \lfloor \log_2 N \rfloor$ and $a/q \in \mathcal{R}_s$. For $0 \le j < \lfloor \log_2 N \rfloor - s$, Equation (11) implies that

$$\|\mu^{a/q,j}\|_{L^\infty(\mathbb{T}^d)} \lesssim_\epsilon \lambda^{\frac{d}{k}-1} (2^s)^{\epsilon-\kappa} (\lambda^{1/k}/2^{s+j})^{1-\kappa+\epsilon} \lesssim_\epsilon 2^{j-s} 2^{j(\epsilon-\kappa)} \lambda^{\frac{d}{k}-\kappa}$$

for all $\epsilon > 0$ since $\kappa > 2$. For $j = \lfloor \log_2 N \rfloor - s$, Equation (11) implies that

$$\|\mu^{a/q,j}\|_{L^\infty(\mathbb{T}^d)} \lesssim_\epsilon 2^{s(\epsilon-\kappa)} \lambda^{\frac{d}{k}-1}.$$

□

Before proving Lemma 3, we need a geometric property of our measures $d\sigma_{\mathcal{Q}}$. The estimate below is best known for $\mathcal{Q}(\boldsymbol{x}) = |\boldsymbol{x}|^2$; see [13] for this estimate. However, we are unaware of a reference for more general hypersurfaces aside from estimate (23) in [21]. For completeness, we include the statement and its proof below.

Proposition 2. *Let ϕ be a Schwartz function on $\mathbb{R}^d$. If $t > 0$, then*

$$\|t^{-d}(\mathrm{D}_t\,\phi) * \mathrm{d}\sigma_Q\|_{L^\infty(\mathbb{R}^d)} \lesssim t^{-1}. \tag{22}$$

Proof. Since $\mathcal{Q}$ is positive definite, the variety $V_{\mathcal{Q}=1}(\mathbb{R})$ is compact. Moreover, Equation (10) implies that for every ball B of radius $r > 0$, we have

$$\sigma(B) \lesssim r^{d-1}. \tag{23}$$

For each point $\boldsymbol{x} \in \mathbb{R}^d$, define the sets $S_0(\boldsymbol{x}) := \{\boldsymbol{y} \in \mathbb{R}^d : |\boldsymbol{x} - \boldsymbol{y}| < t\}$ and $S_j(\boldsymbol{x}) := \{\boldsymbol{y} \in \mathbb{R}^d : 2^j t \leq |\boldsymbol{x} - \boldsymbol{y}| < 2^{j+1} t\}$ for $j \in \mathbb{N}$. By Equation (23), we have that

$$\sigma(S_j(\boldsymbol{x})) \lesssim (2^j t)^{d-1} \tag{24}$$

for each $\boldsymbol{x} \in \mathbb{R}^d$.

Since ϕ is Schwartz, we have

$$\mathrm{D}_t\,\phi(\boldsymbol{x}) \lesssim_\phi (1 + |\boldsymbol{x}/t|)^{-M}$$

for all $M \in \mathbb{N}$. Therefore,

$$\mathrm{D}_t\,\phi * \mathrm{d}\sigma_{\mathcal{Q}}(\boldsymbol{x}) \lesssim (1 + |\cdot/t|)^{-M} * \mathrm{d}\sigma_{\mathcal{Q}}(\boldsymbol{x})$$

for all $\boldsymbol{x} \in \mathbb{R}^d$. Decomposing $\mathbb{R}^d$ into the sets $S_j(\boldsymbol{x})$, we have

$$\begin{aligned}
\mathrm{D}_t\,\phi * \mathrm{d}\sigma_{\mathcal{Q}}(\boldsymbol{x}) &\lesssim_{\phi,M} \int_{\mathbb{R}^d} (1 + |\boldsymbol{x} - \boldsymbol{y}|/t)^{-M}\,\mathrm{d}\sigma(\boldsymbol{y}) \\
&\lesssim \sum_{j=0}^{\infty} \int_{S_j(\boldsymbol{x})} (1 + |\boldsymbol{y}|/t)^{-M}\,\mathrm{d}\sigma(\boldsymbol{y}) \\
&\lesssim \sum_{j=0}^{\infty} \int_{S_j(\boldsymbol{x})} 2^{-jM}\,\mathrm{d}\sigma(\boldsymbol{y})
\end{aligned}$$

Using estimate (24), we obtain that

$$\mathrm{D}_t\,\phi * \mathrm{d}\sigma_{\mathcal{Q}}(\boldsymbol{x}) \lesssim_{\phi,M} \lesssim \sum_{j=0}^{\infty} \sigma_{\mathcal{Q}}(S_j(\boldsymbol{x})) 2^{-jM} \lesssim \sum_{j=0}^{\infty} (2^j t)^{d-1} 2^{-jM} \lesssim t^{d-1}.$$

Normalizing by t^{-d}, we obtain the desired estimate. □

Proof of Lemma 3. Fix $0 \leq s \leq \lfloor \log_2 N \rfloor$ and $a/q \in \mathcal{R}_s$. For each $0 \leq j \leq \lfloor \log_2 N \rfloor - s$, identity (18) and estimate (22) imply that for each $\boldsymbol{x} \in \mathbb{Z}^d$, we have

$$\mu_\lambda^{a/q,j}(\boldsymbol{x}) \lesssim_d 2^j (\lambda^{1/k}/2^s)^{-1} \lambda^{-1} \lesssim_d 2^{j+s} \lambda^{-1-\frac{1}{k}}$$

https://doi.org/10.1017/S0013091523000366 Published online by Cambridge University Press

by taking $\phi = \mathcal{F}_{\mathbb{R}^d}(\mathrm{D}_{2^s}\ \lambda^{-1/k}\psi_j)$ and $t = \lambda^{1/k}2^{-s}$ in Proposition 2. $\square$

Proof of Theorem 2. Let $1 \leq p \leq 2$ and $f \in L^{p'}(\mathbb{T}^d)$ be normalized so that $\|f\|_{L^{p'}(\mathbb{T}^d)} = 1$. Interpolating the bounds (19) and (21) for $\mu_\lambda^{a/q,j}$ when $0 \leq j+s < \lfloor \log_2 N \rfloor$, we obtain

$$\begin{aligned}\|\mu_\lambda^{a/q,j} * f\|_p &\lesssim_\epsilon \left(2^{j+s}\lambda^{-1-\frac{1}{k}}\right)^{\frac{2}{p}} \cdot \left(2^{j-s}2^{j(\epsilon-\kappa)}\lambda^{\frac{d}{k}-\kappa}\right)^{1-\frac{2}{p}} \\ &= 2^{j(\frac{2}{p}+(1+\epsilon-\kappa)(1-\frac{2}{p}))} \cdot 2^{s(\frac{2}{p}-1+\frac{2}{p})} \cdot \lambda^{(\frac{d}{k}-\kappa)(1-\frac{2}{p})-\frac{2}{p}(1+\frac{1}{k})} \\ &= 2^{j(1+(\epsilon-\kappa)(1-\frac{2}{p}))} \cdot 2^{s(\frac{4}{p}-1)} \cdot \lambda^{(\frac{d}{k}-\kappa)(1-\frac{2}{p})-\frac{2}{p}(1+\frac{1}{k})}.\end{aligned}$$

Summing over fractions $a/q \in R_s$ for $j \leq s < \lfloor \log_2 N \rfloor$, we find that

$$\left\| \left(\sum_{a/q\in R_s} \mu_\lambda^{a/q,j}(\boldsymbol{x}) \right) * f \right\|_{L^p(\mathbb{T}^d)} \lesssim_{\mathcal{Q},\epsilon} 2^{j(1+(\epsilon-\kappa)(1-\frac{2}{p}))} \cdot 2^{s(\frac{4}{p}+1)} \cdot \lambda^{(\frac{d}{k}-\kappa)(1-\frac{2}{p})-\frac{2}{p}(1+\frac{1}{k})}.$$

Provided $1-\kappa(1-\frac{2}{p}) < 0$, which is equivalent to the range $p > 2 + \frac{2}{\kappa-1}$, we have

$$\left\| \left(\sum_{j=0}^{\lfloor \log_2 N \rfloor - s - 1} \sum_{a/q\in R_s} \mu_\lambda^{a/q,j}(\boldsymbol{x}) \right) * f \right\|_{L^p(\mathbb{T}^d)} \lesssim_{\mathcal{Q},\epsilon} 2^{s(\frac{4}{p}+1)} \cdot \lambda^{(\frac{d}{k}-\kappa)(1-\frac{2}{p})-\frac{2}{p}(1+\frac{1}{k})}.$$

Consequently, when $p > 2 + \frac{2}{\kappa-1}$, we have

$$\left\| \left(\sum_{s=0}^{\lfloor \log_2 N \rfloor} \sum_{j=0}^{\lfloor \log_2 N \rfloor - s - 1} \sum_{a/q\in R_s} \mu_\lambda^{a/q,j}(\boldsymbol{x}) \right) * f \right\|_{L^p(\mathbb{T}^d)} \lesssim_{\mathcal{Q},p} \lambda^{(\frac{4}{p}+1)/k} \cdot \lambda^{(\frac{d}{k}-\kappa)(1-\frac{2}{p})-\frac{2}{p}(1+\frac{1}{k})}.$$

Comparing the exponent of λ with the desired one of $\frac{d}{k} - 1 - \frac{2d}{kp}$, we find that we have Equation (2) for $p > 2 + \frac{4}{k\kappa-\kappa-1}$. This is better than the range of $p > 2 + \frac{4}{\kappa-2}$ claimed in the theorem.

When $0 \leq j + s = \lfloor \log_2 N \rfloor$, we have

$$\|\mu_\lambda^{a/q,j} * f\|_p \lesssim_\epsilon \left(\lambda^{-1}\right)^{\frac{2}{p}} \cdot \left(2^{s(\epsilon-\kappa)}\lambda^{\frac{d}{k}-1}\right)^{1-\frac{2}{p}} = 2^{s(\epsilon-\kappa)(1-\frac{2}{p})} \cdot \lambda^{\frac{d}{k}-1-\frac{2d}{kp}}.$$

Summing over $0 \leq s \leq \lfloor \log_2 N \rfloor$, we find that

$$\left\| \sum_{s=0}^{\lfloor \log_2 N \rfloor} \mu_\lambda^{a/q,j} * f \right\|_p \lesssim \lambda^{\frac{d}{k}-1-\frac{2d}{kp}}$$

https://doi.org/10.1017/S0013091523000366 Published online by Cambridge University Press

provided that $(\epsilon - \kappa)(1 - \frac{2}{p}) < 0$ for arbitrarily small, positive ϵ. For each $0 < \epsilon < \kappa - 2$, this is equivalent to the range of $p > \frac{2(\kappa-\epsilon)}{\kappa-2-\epsilon}$. Thereby, taking ϵ to 0, we arrive at the range of $p > \frac{2\kappa}{\kappa-2} = 2 + \frac{4}{\kappa-2}$, as claimed. □

Acknowledgements. The authors would like to thank the Heilbronn Institute for Mathematical Research for enabling this collaboration through their Focused Research Workshop 'Efficient Congruencing and Decoupling' in June 2019. KH thanks Virginia Tech for their hospitality whilst part of this paper was written.

KH thanks Dr. Efthalia Tzitzili for a discussion on positive definite forms. We thank the anonymous referees for their feedback on this paper.

Funding Statement. EP was supported in part by Simons Foundation grant no. 360560.

Competing Interests. The authors declare no competing interests pertaining to the undertaken research.

References

(1) T. Anderson, B. Cook, K. Hughes and A. Kumchev, Improved l^p boundedness for integral k-spherical maximal functions, *Discrete Anal.* (2018), 1–18. doi: 10.19086/da.3675.

(2) B. J. Birch, Forms in many variables, *Proc. Roy. Soc. Ser. A* **265** (1961), 245–263.

(3) J. Bourgain, On $\Lambda(p)$-subsets of squares, *Israel J. Math.* **67**(3) (1989), 291–311.

(4) J. Bourgain, Eigenfunction bounds for the Laplacian on the n-torus, *Int. Math. Res. Not.* **1993**(3) (1993), 61–66. doi: 10.1155/S1073792893000066.

(5) J. Bourgain, Analysis results and problems related to lattice points on surfaces, in *Harmonic Analysis and Nonlinear Differential Equations (Riverside, CA, 1995)*, Contemporary Mathematics, Volume 208, pp. 85–109 (American Mathematical Society, Providence, RI, 1997).

(6) J. Bourgain, Moment inequalities for trigonometric polynomials with spectrum in curved hypersurfaces, *Israel J. Math.* **193**(1) (2013), 441–458.

(7) J. Bourgain and C. Demeter, Improved estimates for the discrete Fourier restriction to the higher dimensional sphere, *Illinois J. Math.* **57**(1) (2013), 213–227.

(8) J. Bourgain and C. Demeter, New bounds for the discrete Fourier restriction to the sphere in 4D and 5D, *Int. Math. Res. Not.* **2015**(11) (2015), 3150–3184. doi: 10.1093/imrn/rnu036.

(9) J. Bourgain and C. Demeter, The proof of the ℓ^2 decoupling conjecture, *Ann. of Math. (2)* **182.1** (2015), 351–389.

(10) J. Bourgain and C. Demeter, Decouplings for curves and hypersurfaces with nonzero Gaussian curvature, *J. Anal. Math.* **133** (2017), 279–311.

(11) J. Bourgain, C. Demeter and L. Guth, Proof of the main conjecture in Vinogradov's mean value theorem for degrees higher than three, *Ann. of Math. (2)* **184**(2) (2016), 633–682.

(12) B. Cook and K. Hughes, Bounds for lacunary maximal functions given by Birch–Magyar averages, *Trans. Amer. Math. Soc.* **374** (2021), 3859–3879. doi: 10.1090/tran/8152.

(13) L. Grafakos, *Classical Fourier Analysis*, 2nd ed. (Springer, 2008).

(14) S. Guo and R. Zhang, On integer solutions of Parsell-Vinogradov systems, *Invent. Math.* **218**(1) (2019), 1–81.

(15) S. Guo and P. Zorin-Kranich, Decoupling for certain quadratic surfaces of low codimensions, Preprint arXiv:1902.03450 (2019).

(16) S. Guo and P. Zorin-Kranich, Decoupling for moment manifolds associated to Arkhipov–Chubarikov–Karatsuba systems, *Adv. Math.* **360** (2020), 1–56.

(17) K. Henriot and K. Hughes, Discrete restriction estimates of epsilon-removal type for kth-powers and k-paraboloids, *Math. Ann.* **372**(3) (2018), 963–998.

(18) K. Henriot and K. Hughes, On restriction estimates for discrete quadratic surfaces, *Int. Math. Res. Not.* **2019**(23) (2019), 7139–7159.

(19) Y. Hu and X. Li, Discrete Fourier restriction associated with KdV equations, *Anal. PDE* **6**(4) (2013), 859–892.

(20) Y. Hu and X. Li, Discrete Fourier restriction associated with Schrödinger equations, *Rev. Mat. Iberoam.* **30**(4) (2014), 1281–1300.

(21) K. Hughes, ℓ^p-improving for discrete spherical averages, *Ann. H. Lebesgue* **3** (2020), 959–980.

(22) E. Keil, On a diagonal quadric in dense variables, *Glasg. Math. J.* **56**(3) (2014), 601–28.

(23) X. Lai and Y. Ding, A note on the discrete Fourier restriction problem, *Proc. Amer. Math. Soc.* **146**(9) (2018), 3839–3846.

(24) A. Magyar, Diophantine equations and ergodic theorems, *Amer. J. Math.* **124**(5) (2002), 921–953.

(25) A. Magyar, On the distribution of lattice points on spheres and level surfaces of polynomials, *J. Number Theory* **122**(1) (2007), 69–83.

(26) A. Magyar, On distance sets of large sets of integer points, *Israel J. Math.* **164** (2008), 251–263.

(27) A. Magyar, E. M. Stein and S. Wainger, Discrete analogues in harmonic analysis: spherical averages, *Ann. of Math. (2)* **155**(1) (2002), 189–208.

(28) S. T. Parsell, S. M. Prendiville and T. D. Wooley, Near-optimal mean value estimates for multidimensional Weyl sums, *Geom. Funct. Anal.* **23**(6) (2013), 1962–2024.

(29) T. Wooley, Nested efficient congruencing and relatives of Vinogradov's mean value theorem, *Proc. Lond. Math. Soc.* **118**(4) (2019), 942–1016.