JEMS

Ben Krause · Joris Roos

Discrete analogues of maximally modulated singular integrals of Stein–Wainger type

Abstract. Consider the maximal operator

$$\mathscr{C} f(x) = \sup_{\lambda \in \mathbb{R}} \Big| \sum_{y \in \mathbb{Z}^n \setminus \{0\}} f(x - y) e(\lambda |y|^{2d}) K(y) \Big| \quad (x \in \mathbb{Z}^n),$$

where d is a positive integer, K a Calderón–Zygmund kernel and $n \geq 1$. This is a discrete analogue of a real-variable operator studied by Stein and Wainger. The nonlinearity of the phase introduces a variety of new difficulties that are not present in the real-variable setting. We prove $\ell^2(\mathbb{Z}^n)$-bounds for $\mathscr{C}$, answering a question posed by Lillian Pierce.

Keywords. Discrete analogues, Stein–Wainger type operators, circle method

1. Introduction

Let d and n be positive integers and K a homogeneous Calderón–Zygmund kernel on $\mathbb{R}^n$, taking the form

$$K(x) = \text{p.v.} \frac{\Omega(x)}{|x|^n},$$

where Ω is a smooth function on $\mathbb{R}^n \setminus \{0\}$ that is homogeneous of degree zero. We also assume that $\int_{\mathbb{S}^{n-1}} \Omega(x) \, d\sigma(x) = 0$, where σ denotes the surface measure on the sphere $\mathbb{S}^{n-1} \subset \mathbb{R}^n$. Consider the following operator acting on functions $f : \mathbb{Z}^n \to \mathbb{C}$:

$$\mathscr{C} f(x) = \sup_{\lambda \in \mathbb{R}} \Big| \sum_{y \in \mathbb{Z}^n \setminus \{0\}} f(x - y) e(\lambda |y|^{2d}) K(y) \Big| \quad (x \in \mathbb{Z}^n), \tag{1.1}$$

Ben Krause: Department of Mathematics, King's College London, London, WC2R 2LS, UK;
ben.krause@kcl.ac.uk

Joris Roos: Department of Mathematical Sciences, University of Massachusetts Lowell, Lowell, MA 01854, USA, and School of Mathematics, The University of Edinburgh, Edinburgh, Scotland, UK; jroos.math@gmail.com

where $|y| = (y_1^2 + \cdots + y_n^2)^{1/2}$ and $e(x) = e^{2\pi i x}$. This is a discrete analogue of a maximal operator studied by Stein and Wainger [11]. We also refer to $\mathscr{C}$ as a *discrete Carleson operator*. This is motivated by the formal resemblance to Carleson's operator given by the presence of a supremum over the modulation parameters λ. However, we stress that the (substantial) difficulties encountered in the analysis of the present operator are of a fundamentally different nature than those encountered in the analysis of Carleson's operator. The nonlinearity of the phase causes a number of new challenges arising from a curious fusion of number-theoretic and analytic phenomena which are not present in the real-variable case. We refer to the introduction of [3] for further discussions motivating the study of this operator and to [8], [7], [6], [9] for background and recent progress on some other related discrete analogues in harmonic analysis. The following is our main result.

Theorem 1.1. *There is a constant $C \in (0, \infty)$ such that*

$$\|\mathscr{C} f\|_{\ell^2(\mathbb{Z}^n)} \leq C \|f\|_{\ell^2(\mathbb{Z}^n)}. \tag{1.2}$$

The constant C only depends on d, n and K.

The case $n = d = 1$ was the subject of a question posed by Lillian Pierce during a 2015 workshop at the American Institute of Mathematics. We build on key partial progress previously obtained in [3], where a restricted supremum was considered.

The specific choice of the phase in (1.1) and the assumptions made on the kernel K are imposed primarily in favor of simplicity. Various extensions for other phase functions could be topics for further investigation. Another interesting problem is to prove ℓ^p bounds for $p \neq 2$, which is the topic of a forthcoming sequel to this paper, [4].

Structure of the paper. In §2 we introduce some basic facts and notations used throughout the proof. The most substantial of these are certain known exponential sum estimates from [10].

In §3 we give the proof of Theorem 1.1. The basic strategy follows that of [3], splitting the multiplier into a number-theoretic approximate ('major arcs') and an error term ('minor arcs'). This approach goes back to Bourgain [1] and can be viewed as an instance of the Hardy–Littlewood circle method. The proof involves four distinct components, which (with a slight abuse of terminology) we refer to as 'Minor arcs I/II' and 'Major arcs I/II'.

In §4 ('Minor arcs I') we perform a preliminary TT^* argument to reduce the set of modulation parameters λ.

In §5 ('Minor arcs II') we estimate the error terms from a number-theoretic approximation of the multipliers. This is a standard argument using the fundamental theorem of calculus (which only becomes possible after the crucial reduction from §4). This is already featured in [3].

In §6 ('Major arcs I') we handle the number-theoretic component of the main contribution to the multiplier by exploiting exponential sum estimates. A somewhat unanticipated dichotomy appears here between the cases $d = 1$ and $d \geq 2$.

In §7 ('Major arcs II') we handle the full supremum by combining the number-theoretic component with a delicate multi-frequency analysis similar to [3]. A new aspect

is that we make crucial use of a numerical inequality (see (2.3)) that also appeared in recent works of Mirek and Trojan and Mirek, Stein and Trojan [9], [8], [7].

This allows us to avoid the use of versions of Bourgain's logarithmic multi-frequency lemma [1] and variation-norm estimates from [2], which could be used to give an alternative argument (as suggested by a remark in [3]). Avoiding the use of Bourgain's lemma is desirable in view of extensions beyond ℓ^2.

2. Preliminaries

We write $A \lesssim B$ to denote existence of a constant C such that $A \leq C \cdot B$, where the admissible dependencies of the constant C will be specified, or clear from context. Throughout the text we allow constants to depend on the ambient dimension n, the degree d and the kernel K. Similarly, $A \approx B$ signifies that both $A \lesssim B$ and $B \lesssim A$. The notation $A = B + O(X)$ stands for $|A - B| \lesssim X$. Moreover, we write $A \asymp B$ to express that $\frac{1}{2}B \leq A \leq 2B$.

2.1. Fourier transforms on $\mathbb{Z}^n$, $\mathbb{T}^n$, $\mathbb{R}^n$ and transference

For Fourier transforms of functions $f : \mathbb{Z}^n \to \mathbb{C}$, $g : \mathbb{T}^n \to \mathbb{C}$ we use the notations

$$\widehat{f}(\xi) = \mathcal{F}_{\mathbb{Z}^n} f(\xi) = \sum_{x \in \mathbb{Z}^n} e(-\xi \cdot x) f(x),$$

$$\mathcal{F}^{-1}[g](x) = \mathcal{F}_{\mathbb{Z}^n}^{-1}[g](x) = \int_{\mathbb{T}^n} e(\xi \cdot x) \widehat{g}(\xi) \, d\xi.$$

Here $\mathbb{T}^n = (\mathbb{R}/\mathbb{Z})^n$. A function $g : \mathbb{R}^n \to \mathbb{C}$ that satisfies $g(x + z) = g(x)$ for all $z \in \mathbb{Z}^n$ will be called *periodic* and be silently identified with the corresponding function on $\mathbb{T}^n$. For a function $h : \mathbb{R}^n \to \mathbb{C}$ we write

$$\widehat{h}(\xi) = \mathcal{F}_{\mathbb{R}^n} h(\xi) = \int_{\mathbb{R}^n} e(-\xi \cdot x) h(x) \, dx,$$

$$\mathcal{F}^{-1}[h](x) = \mathcal{F}_{\mathbb{R}^n}^{-1}[h](x) = \widehat{h}(-x).$$

In particular, Fourier transforms on $\mathbb{Z}^n$ or $\mathbb{R}^n$ will be denoted by the same symbols unless the distinction is not clear from context, or is emphasized for other reasons.

For a bounded periodic function $m : \mathbb{R}^n \to \mathbb{C}$ we denote by $m(\mathrm{D})$ the associated Fourier multiplier acting on $\mathbb{Z}^n$, defined as

$$m(\mathrm{D}) f(x) = \mathcal{F}_{\mathbb{Z}^n}^{-1}[m \cdot \mathcal{F}_{\mathbb{Z}^n} f](x) \quad (x \in \mathbb{Z}^n).$$

We slightly abuse notation and also write $m(\mathrm{D})$ for the Fourier multiplier acting on $\mathbb{R}^n$, defined as

$$m(\mathrm{D}) h(x) = \mathcal{F}_{\mathbb{R}^n}^{-1}[m \cdot \mathcal{F}_{\mathbb{R}^n} f](x) \quad (x \in \mathbb{R}^n).$$

Let $(m_\lambda)_{\lambda \in \Lambda}$ be a family of bounded functions supported on a fundamental domain of $\mathbb{T}^n$ (such as a translate of the unit cube $[0, 1)^n$) and denote their periodizations by

$$\mathbf{m}_\lambda(\xi) = \sum_{z \in \mathbb{Z}^n} m_\lambda(\xi + z) \quad (\xi \in \mathbb{R}^n).$$

We will make use of the following transference principle.

Lemma 2.1. *Suppose that for some constant $A > 0$,*

$$\left\| \sup_{\lambda \in \Lambda} |m_\lambda(\mathrm{D}) f| \right\|_{L^2(\mathbb{R}^n)} \leq A \|f\|_{L^2(\mathbb{R}^n)}.$$

Then

$$\left\| \sup_{\lambda \in \Lambda} |\mathbf{m}_\lambda(\mathrm{D}) f| \right\|_{\ell^2(\mathbb{Z}^n)} \lesssim_n A \|f\|_{\ell^2(\mathbb{Z}^n)},$$

where the implicit constant only depends on n.

The proof of this fact is standard (see [1, Lemma 4.4]; there in the case $n = 1$, but the argument also works for $n \geq 2$).

2.2. Some notation and TT^*

For a function $\mathcal{K} : \mathbb{Z}^n \times \mathbb{Z}^n \to \mathbb{C}$ we denote by $T_{\mathcal{K}}$ the operator defined formally by

$$T_{\mathcal{K}} f(x) = \sum_{y \in \mathbb{Z}^n} \mathcal{K}(x, y) f(y). \tag{2.1}$$

Then the operator $T_{\mathcal{K}} T_{\mathcal{K}}^*$ is formally given by $T_{\mathcal{K}} T_{\mathcal{K}}^* = T_{\mathcal{K}^\sharp}$, where the kernel $\mathcal{K}^\sharp$ is

$$\mathcal{K}^\sharp(x, y) = \sum_{z \in \mathbb{Z}^n} \mathcal{K}(x, z) \overline{\mathcal{K}(y, z)}.$$

2.3. Kernel decomposition

Let ψ be a smooth function on $\mathbb{R}^n$ supported in $\{1/2 \leq |x| \leq 2\}$ with $0 \leq \psi \leq 1$ and $\sum_{j \in \mathbb{Z}} \psi_j(x) = 1$ for every $x \neq 0$, where $\psi_j(x) = \psi(2^{-j} x)$. Decompose

$$K(x) = \sum_{j \geq 1} K_j(x),$$

with $K_1(x) = \sum_{j \leq 1} \psi_j(x) K(x)$ and $K_j(x) = \psi_j(x) K(x)$ for $j \geq 2$. Then for all $j \geq 1$ and all $x \in \mathbb{R}^n \setminus \{0\}$,

$$|K_j(x)| \lesssim 2^{-jn}, \quad |\nabla K_j(x)| \lesssim 2^{-j(n+1)}, \quad \text{supp } K_j \subset \{x : |x| \leq 2^{j+1}\}. \tag{2.2}$$

2.4. A numerical inequality

We record a Rademacher–Menshov-type numerical inequality that was also crucially used in [8, Lemma 2.3]: for complex numbers $(a_j)_{j=0,\dots,2^s}$ we have

$$\max_{0 \le j \le 2^s} |a_j| \le |a_{j_0}| + \sqrt{2} \sum_{l=0}^{s} \Big(\sum_{0 \le \kappa < 2^{s-l}} |a_{(\kappa+1)2^l} - a_{\kappa 2^l}|^2 \Big)^{1/2}, \qquad (2.3)$$

for every integer j_0 with $0 \le j_0 \le 2^s$. This follows from an appropriate decomposition into dyadic intervals [5, 9].

2.5. Exponential sum estimates

Given integers $x_1, \dots, x_m$ at least one of which is non-zero we often use the notation $(x_1, \dots, x_m)$ for the greatest common divisor of $x_1, \dots, x_m$. It will be clear from context whether $(x_1, \dots, x_m)$ refers to the greatest common divisor, or the vector of the integers $x_1, \dots, x_m$. For a positive integer q we use the notation

$$[q] = \mathbb{Z} \cap [0, q).$$

The letter q always denotes a positive integer throughout the text. By a *reduced rational* we mean a fraction a/q with $a \in \mathbb{Z}$ and $(a, q) = 1$. For a positive integer $D \ge 2$, $x \in \mathbb{R}^n$ and real coefficients $\xi = (\xi_\alpha)_{1 \le |\alpha| \le D}$ we define the polynomial

$$P(\xi; x) = \sum_{1 \le |\alpha| \le D} \xi_\alpha x^\alpha,$$

where $\alpha \in \mathbb{N}_0^n$ denotes a multiindex. A key ingredient will be the following exponential sum estimate, due to Stein and Wainger [10, Proposition 3].

Proposition 2.2. *Let $R \ge 1$, φ a smooth function on $\mathbb{R}^n$ such that*

$$|\varphi(x)| \le 1 \quad and \quad |\nabla \varphi(x)| \le (1 + |x|)^{-1}$$

for all $x \in \mathbb{R}^n$, and ω a convex set contained in the ball of radius $100R$ centered at the origin. Then for every $\varepsilon > 0$ there exists $\delta > 0$ only depending on ε, n, D such that the following holds: for every ξ with the property that for some α_0 with $1 \le |\alpha_0| \le D$ there exists a reduced rational $a/q \in \mathbb{Q}$ such that

$$|\xi_{\alpha_0} - a/q| \le 1/q^2 \quad and \quad R^\varepsilon \le q \le R^{|\alpha_0| - \varepsilon},$$

we have

$$\Big| \sum_{x \in \mathbb{Z}^n \cap \omega} e(P(\xi; x)) \varphi(x) \Big| \le C R^{n - \delta},$$

where the constant C only depends on ε, n, D.

2.6. Approximation of the multipliers

For $j \geq 1$, $\lambda \in \mathbb{R}$ and $\xi \in \mathbb{R}^n$ we define the multipliers

$$m_{j,\lambda}(\xi) = \sum_{y \in \mathbb{Z}^n} e(\lambda |y|^{2d} + \xi \cdot y) K_j(y). \tag{2.4}$$

This defines a periodic function both in λ and ξ. Following Bourgain [1], the starting point for our arguments is an appropriate approximation for the value of $m_{j,\lambda}(\xi)$ when ξ and λ are close to rationals with small denominator. To formulate the result, we define the exponential sums

$$S\left(\frac{a}{q}, \frac{\mathbf{b}}{q}\right) = \frac{1}{q^n} \sum_{r \in [q]^n} e\left(\frac{a}{q} |r|^{2d} + \frac{\mathbf{b}}{q} \cdot r\right) \tag{2.5}$$

for rationals $a/q \in \mathbb{Q}$, $\mathbf{b}/q \in \mathbb{Q}^n$ with $(a, \mathbf{b}, q) = 1$ (note that this condition makes $S(a/q, \mathbf{b}/q)$ well-defined).

By Proposition 2.2 there exists $\delta > 0$ such that

$$|S(a/q, \mathbf{b}/q)| \lesssim_{d,n} q^{-\delta}. \tag{2.6}$$

The following observation will be crucial at various points in the proof of Theorem 1.1.

Lemma 2.3. *Suppose that $a/q \in \mathbb{Q}$, $\mathbf{b}/q \in \mathbb{Q}^n$, $(a, \mathbf{b}, q) = 1$ and $(a, q) > 1$. Then $S(a/q, \mathbf{b}/q) = 0$.*

We postpone the standard proof of this to the end of this section. Next, we define the real-variable versions of the multipliers $m_{j,\lambda}(\xi)$ by

$$\Phi_{j,\lambda}(\xi) = \int_{\mathbb{R}^n} e(\lambda |y|^{2d} + \xi \cdot y) K_j(y) \, dy. \tag{2.7}$$

At this point we record the following standard oscillatory integral decay estimate in the spirit of van der Corput's lemma:

$$|\Phi_{j,\lambda}(\xi)| \lesssim (1 + 2^{2dj} |\lambda| + 2^j |\xi|)^{-\frac{1}{2d}}. \tag{2.8}$$

For the proof we refer to [11, Proposition 2.1]. This estimate does not enter in the proof of the approximation result in this section, but will be important later on. Our basic approximation result for the multipliers $m_{j,\lambda}(\xi)$ now reads as follows.

Lemma 2.4. *Let j, q be positive integers with $q \leq 2^{j-2}$. Let $a \in \mathbb{Z}$ and $\mathbf{b} \in \mathbb{Z}^n$ with $(a, \mathbf{b}, q) = 1$. Further, assume that $\lambda \in \mathbb{R}$ and $\xi \in \mathbb{R}^n$ are such that*

$$|\lambda - a/q| \leq \delta 2^{-(2d-1)j} \quad \text{and} \quad |\xi - \mathbf{b}/q| \leq \delta, \tag{2.9}$$

where $\delta \in (2^{-j}, 1)$. Then

$$m_{j,\lambda}(\xi) = S(a/q, \mathbf{b}/q) \Phi_{j,\lambda-a/q}(\xi - \mathbf{b}/q) + O(q\delta), \tag{2.10}$$

where the implicit constant depends only on d, n, K.

Proof. The proof is similar to that of [1, Lemma 5.12]. Writing $y = uq + r$ with $u \in \mathbb{Z}^n$, $r \in [q]^n$, we can express $m_{j,\lambda}(\xi)$ as

$$q^{-n} \sum_{r \in [q]^n} e\left(\frac{a}{q}|r|^{2d} + \frac{\mathbf{b}}{q} \cdot r\right) I_{q,r}\left(\lambda - \frac{a}{q}, \xi - \frac{\mathbf{b}}{q}\right),$$

where

$$I_{q,r}(v, \eta) = q^n \sum_{u \in \mathbb{Z}^n} e(v|uq + r|^{2d} + \eta \cdot (uq + r)) K_j(uq + r).$$

It suffices to show that for every $r \in [q]^n$ and every $(v, \eta) \in \mathbb{R} \times \mathbb{R}^n$ with

$$|v| \leq \delta 2^{-(2d-1)j}, \quad |\eta| \leq \delta$$

we have the relation

$$I_{q,r}(v, \eta) = \int_{\mathbb{R}^n} e(v|t|^{2d} + \eta \cdot t) K_j(t) \, dt + O(\delta q). \tag{2.11}$$

The integral on the right-hand side of (2.11) equals

$$q^n \int_{\mathbb{R}^n} e(v|tq + r|^{2d} + \eta(tq + r)) K_j(tq + r) \, dt,$$

which in turn can be split as

$$q^n \sum_{u \in \mathbb{Z}^n} \int_{[0,1]^n} e\left(v|uq + r + tq|^{2d} + \eta \cdot (uq + r + tq)\right) K_j(uq + r + tq) \, dt. \tag{2.12}$$

In this display it holds that

$$\left|v|uq + r + tq|^{2d} - v|uq + r|^{2d}\right| \lesssim \delta q$$

since $|r| \leq q$, $|uq + r + qt| \approx |uq + r| \approx 2^j$ and $v \leq \delta 2^{-(2d-1)j}$. Similarly,

$$|\eta \cdot (uq + r + tq) - \eta \cdot (uq + r)| \lesssim \delta q.$$

Using also $\int_{\mathbb{R}^n} |K_j(t)| \, dt \approx 1$, this implies that (2.12) is

$$q^n \sum_{u \in \mathbb{Z}^n} \int_{[0,1]^n} e(v|uq + r|^{2d} + \eta \cdot (uq + r)) K_j(uq + r + tq) \, dt + O(\delta q). \tag{2.13}$$

Finally, note from (2.2) that

$$|K_j(uq + r + tq) - K_j(uq + r)| \lesssim 2^{-j(n+1)}q \leq 2^{-jn}\delta q.$$

Then we see that (2.13) can be written as

$$q^n \sum_{u \in \mathbb{Z}^n} e(v|uq + r|^{2d} + \eta \cdot (uq + r)) K_j(uq + r) + O(\delta q),$$

which establishes (2.11). ∎

Proof of Lemma 2.3. Let $(a, q) = v > 1$. Write $a = a'v$ and $q = q'v$. Then

$$q^n S(a/q, \mathbf{b}/q) = \sum_{u \in [v]^n} \sum_{r \in [q']^n} e\left(\frac{a'}{q'}|uq' + r|^{2d} + \frac{\mathbf{b}}{q} \cdot (uq' + r)\right)$$

$$= \left[\sum_{r \in [q']^n} e\left(\frac{a'}{q'}|r|^{2d} + \frac{\mathbf{b}}{q} \cdot r\right)\right] \prod_{i=1}^{n} \sum_{u_i \in [v]} e\left(\frac{\mathbf{b}_i}{v} \cdot u_i\right).$$

Since $(a, \mathbf{b}, q) = 1$ and $v > 1$, there must exist i_0 such that $\mathbf{b}_{i_0}$ is not divisible by v. But that implies $\sum_{\ell \in [v]} e\left(\frac{\mathbf{b}_{i_0}}{v}\ell\right) = 0$. ∎

3. Proof of Theorem 1.1

To prove the theorem, we need to obtain an $\ell^2(\mathbb{Z}^n)$ bound for the maximal operator

$$\sup_{\lambda \in \mathbb{R}} \left|\sum_{j \geq 1} m_{j,\lambda}(\mathrm{D}) f\right|,$$

where $m_{j,\lambda}$ is defined in (2.4). A first observation is that for each fixed j,

$$\left\|\sup_{\lambda \in \mathbb{R}} |m_{j,\lambda}(\mathrm{D}) f|\right\|_{\ell^2(\mathbb{Z}^n)} \lesssim \|f\|_{\ell^2(\mathbb{Z}^n)},$$

by the triangle inequality, Young's convolution inequality and (2.2). As a consequence, we may in the following assume that $j \geq j_0$, where j_0 is a sufficiently large constant depending on d and n.

Before we proceed, we give a rough description of what will be done. For this purpose, we will be deliberately vague when using the terms 'small' and 'close'. At this point, the reader should imagine these terms as being relative to appropriate fractional powers of 2^j, which might differ at each occurrence and will have to be chosen carefully in what follows. Roughly speaking, the approximation (2.10) tells us what $m_{j,\lambda}(\xi)$ is when λ and ξ are close to rationals with small denominator. On the other hand, Proposition 2.2 tells us that $|m_{j,\lambda}(\xi)|$ is small if any of $\lambda, \xi_1, \ldots, \xi_n$ is not close to a rational with small denominator. This naturally leads to a decomposition of $m_{j,\lambda}$ into two new functions. The first arises from summing the main contributions $S(a/q, \mathbf{b}/q)\Phi_{j,\lambda-a/q}(\xi - \mathbf{b}/q)$ over a suitable collection of rational $(a/q, \mathbf{b}/q)$ with small q. In the terminology of the Hardy–Littlewood circle method, these are the *major arcs*. The second function is an error term, which will subsume both the approximation error from (2.10) and the *minor arcs*, i.e. the cases when at least one of $\lambda, \xi_1, \ldots, \xi_n$ is not close to one of the chosen rationals. This decomposition is stated below as (3.6). Following this approach naively already leads to a fundamental problem: the error term crucially depends on λ, but we know only little more about it except that its absolute value is small. This leaves us with few strategies to handle the maximal operator corresponding to the error term. This was one of the reasons for the restriction on the parameters λ imposed in [3]. By a preliminary TT^*

argument on the multiplier $m_{j,\lambda}(\xi)$, we may discard 'most' parameters λ: as long as we discard λ sufficiently close to a rational with sufficiently small denominator, the TT^* argument yields summable decay in j (see Proposition 3.1 below). For each j, this only leaves λ contained in a union of a few small intervals (see (3.2) below). This allows us to bound the remaining maximal operator for the error term by a standard argument using the fundamental theorem of calculus, the crucial size information on the error and a crude λ-derivative estimate (see Proposition 3.2 below). We proceed with the precise estimates.

3.1. Decomposition of the multiplier and minor arcs

Define

$$\mathfrak{A}_j = \{a/q \in \mathbb{Q} : (a,q) = 1,\ q \in \mathbb{Z} \cap [1, 2^{\lfloor j\varepsilon_1 \rfloor})\}, \tag{3.1}$$

$$X_j = \bigcup_{\alpha \in \mathfrak{A}_j} \{\lambda \in \mathbb{R} : |\lambda - \alpha| \le 2^{-2dj + \varepsilon_1 j}\}, \tag{3.2}$$

where $\varepsilon_1 \in (0, 2^{-5})$ is a small fixed number that will be determined depending on d and n. Observe that the union in (3.2) is disjoint. The TT^* argument alluded to above yields the following result.

Proposition 3.1. *There exists $\gamma > 0$ only depending on d, n, ε_1 such that for all $j \ge 1$,*

$$\left\| \sup_{\lambda \notin X_j} |m_{j,\lambda}(\mathrm{D})f| \right\|_{\ell^2(\mathbb{Z}^n)} \lesssim 2^{-j\gamma} \|f\|_{\ell^2(\mathbb{Z}^n)}.$$

The proof can be seen as somewhat parallel to that of Stein–Wainger [11] and is given in §4. From now on we can restrict our attention to the multipliers $m_{j,\lambda}(\xi)\mathbf{1}_{X_j}(\lambda)$. In order to define the major arc approximations we need to set up some notation. For a positive integer s define

$$\mathcal{R}_s = \{(a/q, \mathbf{b}/q) \in \mathbb{Q} \times \mathbb{Q}^n : (a, \mathbf{b}, q) = 1,\ q \in \mathbb{Z} \cap [2^{s-1}, 2^s)\}.$$

Fix a smooth radial function χ on $\mathbb{R}^n$ with $0 \le \chi \le 1$ that is supported in $\{|\xi| \le 1/2\}$ and equal to 1 on $[-1/4, 1/4]^n$. For $s \ge 1$ and $\xi \in \mathbb{R}^n$ we write $\chi_s(\xi) = \chi(2^{10s}\xi)$. Further define, for s with $s \le \varepsilon_1 j$,

$$L^s_{j,\lambda}(\xi) = \sum_{(\alpha,\beta) \in \mathcal{R}_s} S(\alpha, \beta) \Phi^*_{j,\lambda-\alpha}(\xi - \beta) \chi_s(\xi - \beta), \tag{3.3}$$

where $\Phi^*_{j,\nu}$ is given by

$$\Phi^*_{j,\nu} = \Phi_{j,\nu} \cdot \mathbf{1}_{|\nu| \le 2^{-2dj + \varepsilon_1 j}}. \tag{3.4}$$

From the definition of $\mathcal{R}_s$ it is clear that $L^s_{j,\lambda}(\xi)$ is periodic in λ and ξ. Also note that if $L^s_{j,\lambda}(\xi) \ne 0$ (where $s \le \varepsilon_1 j$), then $\lambda \in X_j$. Define

$$L_{j,\lambda} = \sum_{1 \le s \le \varepsilon_1 j} L^s_{j,\lambda}. \tag{3.5}$$

Next, the function $E_{j,\lambda}$ is defined as the difference of $m_{j,\lambda}\mathbf{1}_{X_j}(\lambda)$ and $L_{j,\lambda}$ so that

$$m_{j,\lambda} \cdot \mathbf{1}_{X_j}(\lambda) = L_{j,\lambda} + E_{j,\lambda}. \qquad (3.6)$$

From the definitions, $L_{j,\lambda}(\xi)$ and $E_{j,\lambda}(\xi)$ are periodic in λ and ξ and vanish unless $\lambda \in X_j$.

Proposition 3.2. *If the constant ε_1 is chosen small enough* (*depending only on d and n*), *there exists $\gamma > 0$ depending on d, n, ε_1 such that for all $j \geq 1$,*

$$\left\| \sup_{\lambda \in X_j} |E_{j,\lambda}(\mathrm{D})f| \right\|_{\ell^2(\mathbb{Z}^n)} \lesssim 2^{-j\gamma} \|f\|_{\ell^2(\mathbb{Z}^n)}.$$

The proof is given in §5. The basic idea is that the absolute value of $E_{j,\lambda}$ should be small (two reasons to believe this are Lemma 2.4 and Proposition 2.2) and its λ-derivatives are not too large. The structure of X_j then allows us to effectively deploy the fundamental theorem of calculus to deal with the supremum over λ.

3.2. Major arcs

It now remains to bound the maximal operator associated with the multiplier

$$\sum_{j \geq 1} L_{j,\lambda} = \sum_{j \geq 1} \sum_{1 \leq s \leq \varepsilon_1 j} L_{j,\lambda}^s = \sum_{s \geq 1} L_\lambda^s,$$

where we have set

$$L_\lambda^s = \sum_{j \geq \varepsilon_1^{-1} s} L_{j,\lambda}^s. \qquad (3.7)$$

The proof of Theorem 1.1 will be completed if we can exhibit $\gamma > 0$ such that for all $s \geq 1$,

$$\left\| \sup_{\lambda \in \mathbb{R}} |L_\lambda^s(\mathrm{D})f| \right\|_{\ell^2(\mathbb{Z}^n)} \lesssim_{d,n} 2^{-\gamma s} \|f\|_{\ell^2(\mathbb{Z}^n)}. \qquad (3.8)$$

We now begin with the definition of some auxiliary sets of rationals:

$$\begin{aligned}
\mathcal{A}_s &= \{\alpha \in \mathbb{Q} : (\alpha, \beta) \in \mathcal{R}_s \text{ for some } \beta\}, \\
\mathcal{B}_s(\alpha) &= \{\beta \in \mathbb{Q}^n : (\alpha, \beta) \in \mathcal{R}_s\}, \\
\mathcal{B}_s^\sharp &= \{\mathbf{b}/q : \mathbf{b} \in \mathbb{Z}^n, q \in \mathbb{Z} \cap [2^{s-1}, 2^s)\}.
\end{aligned} \qquad (3.9)$$

By definition,

$$(\alpha, \beta) \in \mathcal{R}_s \iff \alpha \in \mathcal{A}_s, \ \beta \in \mathcal{B}_s(\alpha)$$

and

$$\mathcal{B}_s(\alpha) \subset \mathcal{B}_s^\sharp \quad \text{for all } \alpha.$$

Also note that $\mathcal{B}_s(\alpha) = \emptyset$ if $\alpha \notin \mathcal{A}_s$. Fix a smooth radial function $\widetilde{\chi}$ with $0 \leq \widetilde{\chi} \leq 1$ that equals to one on $\{|\xi| \leq 1/2\}$ (and hence on the support of χ) and is supported in

$\{|\xi| \leq 1\}$. Set $\widetilde{\chi}_s(\xi) = \widetilde{\chi}(2^{10s}\xi)$. Given a bounded function m on $\mathbb{R}^n$ we define the periodic multipliers

$$\mathcal{L}_{s,\alpha}[m](\xi) = \sum_{\beta \in \mathcal{B}_s(\alpha)} S(\alpha, \beta) m(\xi - \beta) \chi_s(\xi - \beta), \tag{3.10}$$

$$\mathcal{L}_s^{\sharp}[m](\xi) = \sum_{\beta \in \mathcal{B}_s^{\sharp}} m(\xi - \beta) \widetilde{\chi}_s(\xi - \beta). \tag{3.11}$$

A crucial observation is the factorization

$$\mathcal{L}_{s,\alpha}[m] = \mathcal{L}_{s,\alpha}[1] \cdot \mathcal{L}_s^{\sharp}[m], \tag{3.12}$$

which holds because for each ξ, there is at most one $\beta \in \mathcal{B}_s^{\sharp}$ such that $\widetilde{\chi}_s(\xi - \beta) \neq 0$. The kernel associated with the multiplier (3.10) is given by

$$\mathcal{F}_{\mathbb{Z}^n}^{-1}[\mathcal{L}_{s,\alpha}[m]](y) = \sum_{\beta \in \mathcal{B}_s(\alpha)} S(\alpha, \beta) \int_{[0,1]^n} e(\xi \cdot y) m(\xi - \beta) \chi_s(\xi - \beta) \, d\xi$$

$$= \sum_{\beta \in \mathcal{B}_s(\alpha) \cap [0,1)^n} S(\alpha, \beta) e(\beta \cdot y) \mathcal{F}_{\mathbb{R}^n}^{-1}[m \cdot \chi_s](y), \tag{3.13}$$

where $y \in \mathbb{Z}^n$. With this notation in mind we write the multiplier in question as

$$L_\lambda^s = \mathcal{L}_{s,\alpha}[\Phi_{\lambda-\alpha}^s], \tag{3.14}$$

where we have set

$$\Phi_\lambda^s = \sum_{j \geq \varepsilon_1^{-1} s} \Phi_{j,\lambda}^*. \tag{3.15}$$

and α is the unique element of $\mathcal{A}_s$ such that $|\lambda - \alpha| \leq 2^{-2s-10}$ (say), or an arbitrary value from the complement of $\mathcal{A}_s$ if no such α exists (in this case, $L_\lambda^s(\xi) = 0$ anyways). Here, uniqueness of such α follows because two distinct rationals with denominators $\leq 2^s$ must be at least 2^{-2s} apart. In view of (3.14) and the factorization (3.12) it is reasonable to begin with the following number-theoretic estimate.

Proposition 3.3. *There exists $\gamma > 0$ depending on d and n such that for every $s \geq 1$,*

$$\left\| \sup_{\alpha \in \mathcal{A}_s} |\mathcal{L}_{s,\alpha}[1](\mathrm{D}) f| \right\|_{\ell^2(\mathbb{Z}^n)} \lesssim 2^{-\gamma s} \|f\|_{\ell^2(\mathbb{Z}^n)}. \tag{3.16}$$

This will be proved in §6 by making use of exponential sum estimates. The factorization (3.12) invites us to consider the companion maximal operator

$$f \mapsto \sup_{\mu \in \mathbb{R}} |\mathcal{L}_s^{\sharp}[\Phi_\mu^s](\mathrm{D}) f|.$$

Using Bourgain's multi-frequency lemma and the variational estimates from [2] it is possible to show that this maximal operator has $\ell^2 \to \ell^2$ operator norm $\lesssim s^2$ (the proof is

omitted in this paper, because this claim will not be needed). However, it is technically not straightforward to combine this result with Proposition 3.16 to treat the maximal operator associated with (3.14). Instead, we take a different route that relies on the numerical inequality (2.3) and a theorem of Stein and Wainger [11]. The following proposition is proved in §7.

Proposition 3.4. *The constant ε_1 can be chosen small enough depending on d and n so that there exists $\gamma > 0$ depending on d, n such that for every $s \geq 1$,*

$$\left\| \sup_{\lambda \in \mathbb{R}} |L_\lambda^s(\mathrm{D})f| \right\|_{\ell^2(\mathbb{Z}^n)} \lesssim 2^{-\gamma s} \|f\|_{\ell^2(\mathbb{Z}^n)}. \tag{3.17}$$

This establishes (3.8) and thereby Theorem 1.1.

4. Minor arcs I: Proof of Proposition 3.1

Since the output $m_{j,\lambda}(\mathrm{D})f(x)$ only depends on the values of f in a 2^{j+1}-neighborhood of the point x, a standard localization argument allows us to assume that f is supported in the set

$$B_j = \{y \in \mathbb{Z}^n : |y| \leq 2^j\}.$$

Fix an arbitrary function $\lambda : \mathbb{Z}^n \to \mathbb{R} \setminus X_j$ and write

$$T_{j,\lambda}f(x) = m_{j,\lambda(x)}(\mathrm{D})(f\mathbf{1}_{B_j})(x) = \sum_{y \in \mathbb{Z}^n} f(y)\mathcal{K}_{j,\lambda}(x, y),$$

where

$$\mathcal{K}_{j,\lambda}(x, y) = e(\lambda(x)|x - y|^{2d})K_j(x - y)\mathbf{1}_{B_j}(y).$$

Then the kernel of $T_{j,\lambda}T_{j,\lambda}^*$ is given by

$$\mathcal{K}_{j,\lambda}^\sharp(x, y) = \sum_{z \in \mathbb{Z}^n} e(\lambda(x)|z|^{2d} - \lambda(y)|y - x + z|^{2d})$$
$$\times K_j(z)\overline{K_j(y - x + z)}\mathbf{1}_{B_j}(x - z). \tag{4.1}$$

Note that $\mathcal{K}_{j,\lambda}^\sharp(x, y) = 0$ unless

$$|x| \leq 2^{j+2} \quad \text{and} \quad |y| \leq 2^{j+2}. \tag{4.2}$$

Let $\delta_0 > 0$ and $c_0 > 0$ be determined later and define

$$E_{j,\lambda} = \{(x, y) \in \mathbb{Z}^n \times \mathbb{Z}^n : |\mathcal{K}_{j,\lambda}^\sharp(x, y)| \geq c_0 2^{-j(n+\delta_0)}\}.$$

Lemma 4.1. *The constants c_0 and δ_0 can be chosen depending on d, n, ε_1 such that for every $j \geq 1$,*

$$|E_{j,\lambda}| \lesssim 2^{2nj - \frac{1}{10}\varepsilon_1 j}, \tag{4.3}$$

where ε_1 is as in (3.1) and (3.2).

Before proving this statement we show how it can be used to finish the proof of Proposition 3.1. By definition of $E_{j,\lambda}$,

$$|\mathcal{K}_{j,\lambda}^{\sharp}(x,y)| \lesssim 2^{-nj-\delta_0 j}\mathbf{1}_{B_{j+2}\times B_{j+2}}(x,y) + 2^{-nj}\mathbf{1}_{E_{j,\lambda}}(x,y).$$

With (4.3) this implies

$$\|\mathcal{K}_{j,\lambda}^{\sharp}\|_{\ell^2(\mathbb{Z}^n\times\mathbb{Z}^n)} \lesssim 2^{-\delta_0 j} + 2^{-\frac{1}{20}\varepsilon_1 j}. \tag{4.4}$$

By the Cauchy–Schwarz inequality we have

$$|\langle T_{\mathcal{K}_{j,\lambda}^{\sharp}}f,g\rangle| \leq \sum_{x\in\mathbb{Z}^n}\sum_{y\in\mathbb{Z}^n}|g(x)|\,|f(y)|\,|\mathcal{K}_{j,\lambda}^{\sharp}(x,y)|$$

$$\leq \|f\|_{\ell^2(\mathbb{Z}^n)}\|g\|_{\ell^2(\mathbb{Z}^n)}\|\mathcal{K}_{j,\lambda}^{\sharp}\|_{\ell^2(\mathbb{Z}^n\times\mathbb{Z}^n)},$$

which by (4.4) and ℓ^2 duality leads to

$$\|T_{j,\lambda}\|_{\ell^2\to\ell^2} = \|T_{\mathcal{K}_{j,\lambda}^{\sharp}}\|_{\ell^2\to\ell^2}^{1/2} \lesssim 2^{-\gamma j}$$

with $\gamma = \min(\frac{1}{2}\delta_0, \frac{1}{40}\varepsilon_1)$. It remains to prove Lemma 4.1.

In fact we will prove something stronger: the claim is that after choosing c_0 and δ_0 suitably we have, for every fixed $(x',y^*)\in\mathbb{Z}^{n-1}\times\mathbb{Z}^n$,

$$|\{x_1\in\mathbb{Z} : (x_1,x',y^*)\in E_{j,\lambda}\}| \lesssim 2^{j-\frac{1}{10}\varepsilon_1 j}. \tag{4.5}$$

In other words, each (x',y^*)-slice of $E_{j,\lambda}$ has small cardinality. By Fubini's theorem and (4.2) this implies the claimed inequality (4.3).

For future reference, we will be more careful with explicit constants than strictly necessary in this proof. The reader can safely ignore all constants only depending on d in the estimates that follow. Fixing $(x',y^*)\in\mathbb{Z}^{n-1}\times\mathbb{Z}^n$, we define

$$\mathcal{E} = \{x_1\in\mathbb{Z} : (x_1,x',y^*)\in E_{j,\lambda}\}.$$

Set $\varepsilon_0 = \frac{1}{10}\varepsilon_1$.

Claim. *The numbers c_0 and δ_0 can be chosen such that the following holds: for every $u\in\mathcal{E}$ there exists a reduced rational a/q with $q\leq 2^{\varepsilon_0 j+1}d$ such that*

$$|(u-y_1^*)\lambda(y^*) - a/q| \leq 2^{-j(2d-1)+\varepsilon_0 j}. \tag{4.6}$$

Proof. Note that the coefficient of z_1^{2d-1} in the phase of (4.1) is equal to $2d(x_1-y_1)\lambda(y)$. By Dirichlet's approximation theorem, there exists a reduced rational a/q with $q\leq 2^{j(2d-1)-\varepsilon_0 j}$ such that

$$|2d(u-y_1^*)\lambda(y^*) - a/q| \leq q^{-1}2^{-j(2d-1)+\varepsilon_0 j} \leq 1/q^2.$$

Applying Proposition 2.2 (with $R=2^j$) we may choose c_0 and δ_0 (depending on the choice of ε_0) so that $q\leq 2^{\varepsilon_0 j}$ (because $|\mathcal{K}_{j,\lambda}^{\sharp}(u,x',y^*)|\geq c_0 2^{-j(n+\delta_0)}$). Dividing through by $2d$ yields the claim. ∎

From now on we fix c_0 and δ_0 to make the statement in the claim valid. We will also assume $j \geq j_0$, where j_0 is a large constant depending only on d that will be determined later. Our goal is now to show that $|\mathcal{E}| \leq 2^{j-\varepsilon_0 j}$. Arguing by contradiction, we assume that

$$|\mathcal{E}| > 2^{j-\varepsilon_0 j}. \tag{4.7}$$

It is clear that

$$\mathcal{E} \subset [-2^{j+2}, 2^{j+2}]. \tag{4.8}$$

We now exploit the three properties (4.6)–(4.8) to prove that $\lambda(y^*) \in X_j$, which establishes the required contradiction. First, we claim that there exist $u_1, u_2 \in \mathcal{E}$ such that

$$1 \leq u_2 - u_1 \leq 2^{\varepsilon_0 j+5}, \tag{4.9}$$

Indeed, suppose that all elements of $\mathcal{E}$ were pairwise separated by at least $2^{\varepsilon_0 j+5}$. Then, by (4.8) we would have $|\mathcal{E}| \leq 2^{j-\varepsilon_0 j-1}$, which contradicts (4.7). Consequently, there must exist $u_1, u_2 \in \mathcal{E}$ such that (4.9) holds. By (4.6) there exist reduced rationals $a/q, a'/q'$ with $\max(q, q') \leq 2^{\varepsilon_0 j+1}d$ and

$$|(u_1 - y_1^*)\lambda(y^*) - a/q| \leq 2^{-j(2d-1)+\varepsilon_0 j},$$

$$|(u_2 - y_1^*)\lambda(y^*) - a'/q'| \leq 2^{-j(2d-1)+\varepsilon_0 j}.$$

Then,

$$|\lambda(y^*) - a^*/q^*| \leq 2^{-j(2d-1)+\varepsilon_0 j+1}, \tag{4.10}$$

where $a^*/q^* = (u_2 - u_1)^{-1}(a'/q' - a/q)$ is a reduced rational with

$$q^* \leq qq'(u_2 - u_1) \leq 2^{3\varepsilon_0 j+7}d^2. \tag{4.11}$$

With (4.10) we have already obtained a somewhat decent rational approximation for $\lambda(y^*)$. However, to conclude $\lambda(y^*) \in X_j$, we need to show that the approximation is actually tighter by almost another factor of 2^{-j} on the right-hand side (see (3.2)). Denote the set of reduced rationals $a/q \in [0, 1)$ with $q \leq 2^{\varepsilon_0 j+1}d$ and $a \in [q]$ by $\mathcal{A}$. Then for each $\alpha \in \mathcal{A}$ we define

$$\mathcal{F}_\alpha = \{u \in \mathcal{E} : |(u - y_1^*)\lambda(y^*) - \alpha|_\mathbb{T} \leq 2^{-(2d-1)j+\varepsilon_0 j}\},$$

where $|\xi|_\mathbb{T} = \min_{z \in \mathbb{Z}} |\xi + z| \leq |\xi|$. By (4.6), we have $\mathcal{E} \subset \bigcup_{\alpha \in \mathcal{A}} \mathcal{F}_\alpha$. Since also $|\mathcal{A}| \leq d^2 2^{2\varepsilon_0 j+1}$, the pigeonhole principle and (4.7) imply that there exists $\alpha_0 = a_0/q_0 \in \mathcal{A}$ such that

$$|\mathcal{F}_{\alpha_0}| \geq 2^{j-3\varepsilon_0 j-1}d^{-2}.$$

Now we invoke the pigeonhole principle again in the following form (this step can be skipped if $d > 1$): for positive integers N, k with $(2N + 1)k^{-1} \geq 2$, cover a set $A \subset [-N, N] \cap \mathbb{Z}$ with k intervals, each of length $(2N + 1)k^{-1}$. One of the intervals, call it I, must satisfy $|A \cap I| \geq |A|k^{-1} - 1$. Writing $v_1 = \min A \cap I$ and $v_2 = \max A \cap I$ we

then have $|A|k^{-1} - 2 \le v_2 - v_1 \le (2N+1)k^{-1}$. Applying this fact to our situation with $N = 2^{j+2}$, $A = \mathscr{F}_{\alpha_0}$, and $k = \lceil 2^{5\varepsilon_0 j} \rceil$, we exhibit $v_1, v_2 \in \mathscr{F}_{\alpha_0}$ such that for $j \ge j_0$ large enough,

$$2^{j - 8\varepsilon_0 j - 3} d^{-2} \le v_2 - v_1 \le 2^{j - 5\varepsilon_0 j + 4}. \tag{4.12}$$

By definition of $\mathscr{F}_{\alpha_0}$ there exist integers ℓ_1, ℓ_2 such that

$$|(v_1 - y_1^*)\lambda(y^*) - (\alpha_0 + \ell_1)| \le 2^{-(2d-1)j + \varepsilon_0 j},$$
$$|(v_2 - y_1^*)\lambda(y^*) - (\alpha_0 + \ell_2)| \le 2^{-(2d-1)j + \varepsilon_0 j}.$$

This implies, using the lower bound in (4.12), that

$$\left| \lambda(y^*) - \frac{\ell_2 - \ell_1}{v_2 - v_1} \right| \le 2^{-2dj + 9\varepsilon_0 j + 3} d^2. \tag{4.13}$$

We claim that

$$\frac{\ell_2 - \ell_1}{v_2 - v_1} = \frac{a^*}{q^*}. \tag{4.14}$$

Indeed, suppose not. Then, from (4.12) and (4.11),

$$\left| \frac{\ell_2 - \ell_1}{v_2 - v_1} - \frac{a^*}{q^*} \right| \ge \frac{1}{(v_2 - v_1)q^*} \ge 2^{-j + 2\varepsilon_0 j - 11} d^{-2}.$$

On the other hand, from (4.10) and (4.13),

$$\left| \frac{\ell_2 - \ell_1}{v_2 - v_1} - \frac{a^*}{q^*} \right| \le 2^{-(2d-1)j + \varepsilon_0 j + 2},$$

for $j \ge j_0$ large enough. This yields a contradiction (again, for $j \ge j_0$ large enough). Thus, (4.14) holds. Summarizing, we have proved that

$$|\lambda(y^*) - a^*/q^*| \le 2^{-2dj + 10\varepsilon_0 j}$$

for $j \ge j_0$ large enough (from (4.14) and (4.13)). Further, $(a^*, q^*) = 1$ and $q^* \le d^2 2^{3\varepsilon_0 j + 7} \le 2^{\lfloor 10\varepsilon_0 j \rfloor}$ for large enough $j \ge j_0$. Recalling that we set $\varepsilon_0 = \frac{1}{10}\varepsilon_1$, this means precisely that $\lambda(y^*) \in X_j$.

Remarks. 1. The argument simplifies slightly in the case $d > 1$: in place of the upper bound in (4.12), the trivial upper bound 2^{j+3} would be sufficient.

2. From the proof it is clear that the factor $\frac{1}{10}$ appearing in (4.3) is not sharp. However, this is not relevant for our discussion.

5. Minor arcs II: Proof of Proposition 3.2

We will make use of the following fact.

Lemma 5.1. *Let $\Lambda \subset \mathbb{R}$ be a disjoint union of intervals $(I_j)_{1 \le j \le N}$ with $|I_j| \le \delta$, and $(m_\lambda)_{\lambda \in \Lambda}$ a family of bounded periodic functions on $\mathbb{R}^n$ such that*

$$\sup_{\lambda \in \Lambda} \|m_\lambda\|_{L^\infty(\mathbb{T}^n)} \le A, \tag{5.1}$$

the function $I_j \to \mathbb{C}$, $\lambda \mapsto m_\lambda(\xi)$, is absolutely continuous for a.e. $\xi \in \mathbb{R}^n$ and every $j = 1, \dots, N$, and

$$\sup_{\lambda \in \Lambda} \|\partial_\lambda m_\lambda\|_{L^\infty(\mathbb{T}^n)} \le B. \tag{5.2}$$

Then

$$\left\| \sup_{\lambda \in \Lambda} |m_\lambda(\mathrm{D}) f| \right\|_{\ell^2(\mathbb{Z}^n)} \le (N^{1/2} A + (2NAB\delta)^{1/2}) \|f\|_{\ell^2(\mathbb{Z}^n)}.$$

The proof is via a standard argument using the fundamental theorem of calculus which we postpone to the end of this section. In order to apply Lemma 5.1 to the multipliers $(E_{j,\lambda})_{\lambda \in X_j}$ we will prove that

$$|E_{j,\lambda}(\xi)| \lesssim 2^{-\gamma j} \tag{5.3}$$

for some $\gamma > 0$ only depending on d, n (in particular, not depending on the choice of ε_1) and all $\lambda \in X_j, \xi \in \mathbb{R}^n$, $j \ge 1$. Moreover, we see directly from the definitions (3.6), (3.3), (2.7), (2.4) that for a.e. $\lambda \in X_j, \xi \in \mathbb{R}^n$ and every $j \ge 1$,

$$|\partial_\lambda E_{j,\lambda}(\xi)| \lesssim 2^{2dj}. \tag{5.4}$$

Then Lemma 5.1 (with $\Lambda = X_j \cap [0,1), m_\lambda = E_{j,\lambda}, N = |\mathfrak{A}_j| \le 2^{2\varepsilon_1 j}, \delta \le 2^{-2dj+\varepsilon_1 j+1}$) gives

$$\left\| \sup_{\lambda \in X_j} |E_{j,\lambda}(\mathrm{D}) f| \right\|_{\ell^2(\mathbb{Z}^n)} \lesssim 2^{\frac{1}{2}(3\varepsilon_1 - \gamma)j} \|f\|_{\ell^2(\mathbb{Z}^n)}. \tag{5.5}$$

Thus we obtain the claimed decay in j as long as $\varepsilon_1 < \frac{1}{3}\gamma$. We turn our attention to proving (5.3).

Assume $\lambda \in X_j$ (otherwise $E_{j,\lambda}(\xi) = 0$). Fix $\varepsilon_2 = 2^{-5}$ (this can be replaced by any sufficiently small absolute constant with $\varepsilon_2 > \varepsilon_1$). We define the *major arcs*

$$\mathfrak{M}_j = \bigcup_{\substack{(\alpha,\beta) \in \mathcal{R}_s, \\ 1 \le s \le \varepsilon_2 j}} \mathfrak{M}_j(\alpha, \beta), \quad \text{where}$$

$$\mathfrak{M}_j(\alpha, \beta) = \{(\lambda, \xi) \in \mathbb{R} \times \mathbb{R}^n : |\lambda - \alpha| \le 2^{-2dj+\varepsilon_2 j}, |\xi - \beta| \le 2^{-j+\varepsilon_2 j}\}.$$

We need the following disjointness statement for the neighborhoods of the rationals involved in the sum defining $L_{j,\lambda}(\xi)$.

Lemma 5.2. *For each $(\lambda, \xi) \in \mathbb{R} \times \mathbb{R}^n$ there exists at most one (α, β) with $(\alpha, \beta) \in \mathcal{R}_s$ for some $1 \le s \le \varepsilon_2 j$ such that*

$$S(\alpha, \beta)\Phi_{j,\lambda-\alpha}^*(\xi - \beta)\chi_s(\xi - \beta) \ne 0. \tag{5.6}$$

If that is the case and also $s \le \varepsilon_1 j$, then

$$L_{j,\lambda}(\xi) = L_{j,\lambda}^s(\xi) = S(\alpha, \beta)\Phi_{j,\lambda-\alpha}^*(\xi - \beta)\chi_s(\xi - \beta).$$

(Otherwise, $L_{j,\lambda}(\xi) = 0$.)

Proof. Fix $(\lambda, \xi) \in \mathbb{R} \times \mathbb{R}^n$. Take $(\alpha, \beta) \in \mathcal{R}_s$ and $(\alpha', \beta') \in \mathcal{R}_{s'}$ such that (5.6) holds. Suppose that $\alpha \neq \alpha'$. Then

$$2^{-2\varepsilon_2 j} \leq 2^{-(s+s')} \leq |\alpha - \alpha'| \leq 2^{-2dj+\varepsilon_1 j+1}.$$

This is a contradiction. Thus, $\alpha = \alpha'$. Write $(\alpha, \beta) = (a/q, \mathbf{b}/q)$, $(\alpha', \beta') = (a'/q', \mathbf{b}'/q')$ with $(a, \mathbf{b}, q) = (a', \mathbf{b}', q') = 1$ and $2^{s-1} \leq q < 2^s$, $2^{s'-1} \leq q' < 2^{s'}$. By Lemma 2.3 and (5.6) we have $(a, q) = 1$ and $(a', q') = 1$. But since $\alpha = \alpha'$, this implies $q = q'$ and thus $s = s'$. Taking another look at (5.6) we see that $\beta = \beta'$ (by inspecting the support of $\chi_s = \chi_{s'}$). The claim about $L_{j,\lambda}(\xi)$ follows from the claim we just proved, and (3.5) and (3.3). ∎

The proof of (5.3) naturally splits into several cases.

Case 1: $(\lambda, \xi) \in \mathfrak{M}_j$. Then there exist $1 \leq s_0 \leq \varepsilon_2 j$ and $(\alpha_0, \beta_0) \in \mathcal{R}_{s_0}$ such that $(\lambda, \xi) \in \mathfrak{M}_j(\alpha_0, \beta_0)$. From Lemma 2.4 (with $\delta = 2^{-j+\varepsilon_2 j}$, $q \leq 2^{\varepsilon_2 j}$) we gather that

$$m_{j,\lambda}(\xi) = S(\alpha_0, \beta_0)\Phi_{j,\lambda-\alpha_0}(\xi - \beta_0) + O(2^{-j+2\varepsilon_2 j}). \tag{5.7}$$

We distinguish two further cases.

Case 1.1: $1 \leq s_0 \leq \varepsilon_1 j$. From Lemma 5.2 we deduce

$$L_{j,\lambda}(\xi) = L_{j,\lambda}^{s_0}(\xi) = S(\alpha_0, \beta_0)\Phi_{j,\lambda-\alpha_0}(\xi - \beta_0).$$

With (5.7) this gives

$$|E_{j,\lambda}(\xi)| = |m_{j,\lambda}(\xi) - L_{j,\lambda}(\xi)| \lesssim 2^{-j+2\varepsilon_2 j}.$$

Case 1.2: $\varepsilon_1 j < s_0 \leq \varepsilon_2 j$. We may write $\alpha_0 = a_0/q_0$, $\beta_0 = \mathbf{b}_0/q_0$ with $(a_0, \mathbf{b}_0, q_0) = 1$, $2^{s_0-1} \leq q_0 < 2^{s_0}$. In particular, $q_0 \geq 2^{\lfloor \varepsilon_1 j \rfloor}$.

We claim that we must have $(a_0, q_0) > 1$. Indeed, suppose $(a_0, q_0) = 1$. Since $\lambda \in X_j$, there exists a reduced rational a_1/q_1 with $q_1 < 2^{\lfloor \varepsilon_1 j \rfloor}$ and

$$|a_1/q_1 - \lambda| \leq 2^{-2dj+\varepsilon_1 j}$$

Since $q_0 > q_1$, the reduced rationals a_1/q_1 and a_0/q_0 do not coincide. Therefore,

$$2^{-(\varepsilon_1+\varepsilon_2)j} \leq \frac{1}{q_0 q_1} \leq \left| \frac{a_1}{q_1} - \frac{a_0}{q_0} \right| \leq 2^{-2dj+\varepsilon_2 j+1}.$$

This is a contradiction. Thus we must have $(a_0, q_0) > 1$ and so $S(\alpha_0, \beta_0) = 0$ by Lemma 2.3. In particular, $|m_{j,\lambda}(\xi)| \lesssim 2^{-j+2\varepsilon_2 j}$ by (5.7). Also, from Lemma 5.2 we see that $L_{j,\lambda}(\xi) = 0$.

Case 2: $(\lambda, \xi) \notin \mathfrak{M}_j$. In this case we bound

$$|E_{j,\lambda}(\xi)| \leq |m_{j,\lambda}(\xi)| + |L_{j,\lambda}(\xi)|$$

and estimate the two terms on the right-hand side separately.

Fix $\epsilon < \frac{\varepsilon_2}{n+1}$ and set $N = 2^j$. By Dirichlet's approximation theorem there exist reduced fractions $a/q, b_1/r_1, \ldots, b_n/r_n$ with $q \leq N^{2d-\epsilon}$, $\max(r_1, \ldots, r_n) \leq N^{1-\epsilon}$ and

$$\left| \lambda - \frac{a}{q} \right| \leq \frac{1}{q} N^{-2d+\epsilon}, \qquad \left| \xi_k - \frac{b_k}{r_k} \right| \leq \frac{1}{r_k} N^{-1+\epsilon} \quad \text{for } k = 1, \ldots, n.$$

Setting $q_* = \mathrm{lcm}(q, r_1, \ldots, r_n)$, we must have $q_* \geq 2^{\lfloor \varepsilon_2 j \rfloor}$ because $(\lambda, \xi) \notin \mathfrak{M}_j$. Thus at least one of $q, r_1, \ldots, r_n$ must be $\geq 2^{\epsilon j}$ (otherwise $q_* \leq 2^{\epsilon(n+1)j}$, which is a contradiction because $\epsilon < \frac{\varepsilon_2}{n+1}$). By Proposition 2.2 we then obtain

$$|m_{j,\lambda}(\xi)| \lesssim 2^{-\delta j}.$$

It remains to estimate $|L_{j,\lambda}(\xi)|$. Suppose that $L_{j,\lambda}(\xi) \neq 0$. Then by Lemma 5.2 there exists $(\alpha, \beta) \in \mathcal{R}_s$ for some $1 \leq s \leq \varepsilon_1 j$ such that

$$L_{j,\lambda}(\xi) = S(\alpha, \beta) \Phi_{j,\lambda-\alpha}^*(\xi - \beta) \chi_s(\xi - \beta). \tag{5.8}$$

Then $|\lambda - \alpha| \leq 2^{-2dj+\varepsilon_2 j}$. Since $(\lambda, \xi) \notin \mathfrak{M}_j$,

$$|\xi - \beta| \geq 2^{-j+\varepsilon_2 j}.$$

With (2.8) and (5.8), this implies

$$|L_{j,\lambda}(\xi)| \leq |\Phi_{j,\lambda-\alpha}(\xi - \beta)| \lesssim 2^{-\frac{\varepsilon_2}{2d} j}.$$

Proof of Lemma 5.1. By the fundamental theorem of calculus we have, for absolutely continuous $g : [a, b] \to \mathbb{C}$,

$$\sup_{\lambda \in [a,b]} |g(\lambda)|^2 \leq |g(a)|^2 + 2 \int_a^b |g(t)| \, |g'(t)| \, dt.$$

Hence,

$$\left\| \sup_{\lambda \in \Lambda} |m_\lambda(\mathrm{D}) f| \right\|_{\ell^2(\mathbb{Z}^n)}^2 \leq \sum_{j=1}^N \|m_{\inf I_j}(\mathrm{D}) f\|_{\ell^2(\mathbb{Z}^n)}^2$$

$$+ 2 \sum_{j=1}^N \sum_{x \in \mathbb{Z}^n} \int_{I_j} |m_t(\mathrm{D}) f(x)| \, |\partial_\lambda m_t(\mathrm{D}) f(x)| \, dt. \tag{5.9}$$

By the Cauchy–Schwarz inequality and Fubini's theorem,

$$\sum_{x \in \mathbb{Z}^n} \int_{I_j} |m_t(\mathrm{D}) f(x)| \, |\partial_\lambda m_t(\mathrm{D}) f(x)| \, dt$$

$$\leq \left(\int_{I_j} \|m_t(\mathrm{D}) f\|_{\ell^2(\mathbb{Z}^n)}^2 \, dt \right)^{1/2} \left(\int_{I_j} \|\partial_\lambda m_t(\mathrm{D}) f\|_{\ell^2(\mathbb{Z}^n)}^2 \, dt \right)^{1/2}.$$

Combining this with (5.9) and using Plancherel's theorem with the assumptions (5.1) and (5.2) we obtain the claim. ∎

Remark. Observe that the same argument works for ℓ^p with $p \neq 2$ and more general families of operators.

6. Major arcs I: Proof of Proposition 3.3

Note that since $\mathscr{L}_{s,\alpha}[1] = \mathscr{L}_{s,\alpha+1}[1]$, we may restrict the supremum to $\alpha \in \mathcal{A}_s \cap [0, 1)$, without loss of generality. Also, from (3.13) we have

$$\mathscr{F}_{\mathbb{Z}^n}^{-1}[\mathscr{L}_{s,\alpha}[1]](y) = \sum_{\beta \in \mathcal{B}_s(\alpha) \cap [0,1)^n} S(\alpha, \beta) e(\beta \cdot y) \phi_s(y),$$

where $\phi_s = \mathscr{F}_{\mathbb{R}^n}^{-1}[\chi_s]$. Note that $\|\phi_s\|_{L^1(\mathbb{R}^n)} \approx 1$. For an arbitrary function $\alpha : \mathbb{Z}^n \to \mathcal{A}_s \cap [0, 1)$ we define

$$\mathscr{K}_{s,\alpha}(x, y) = \mathscr{F}_{\mathbb{Z}^n}^{-1}[\mathscr{L}_{s,\alpha(x)}[1]](x - y). \tag{6.1}$$

Then Proposition 3.3 is a consequence of the following.

Proposition 6.1. *There exists $\gamma > 0$ depending only on d, n such that*

$$\|T_{\mathscr{K}_{s,\alpha}} f\|_{\ell^2(\mathbb{Z}^n)} \lesssim 2^{-\gamma s} \|f\|_{\ell^2(\mathbb{Z}^n)},$$

with the implicit constant only depending on d and n, but not on the functions α, f. (The notation $T_{\mathscr{K}_{s,\alpha}}$ is defined in (2.1).)

Remark. The proof shows that the same result holds with ℓ^2 replaced by ℓ^p for every $p \in (1, \infty)$ (with decay rate depending on p).

Proof of Proposition 6.1. For every $x \in \mathbb{Z}^n$ there exist $q(x) \in \mathbb{Z} \cap [2^{s-1}, 2^s)$ and $a(x) \in [q(x)]$ with $(a(x), q(x)) = 1$ such that

$$\alpha(x) = a(x)/q(x).$$

For the proof we will employ a TT^*-argument. We begin by computing the kernel of TT^*. Note that $T_{\mathscr{K}_{s,\alpha}} T_{\mathscr{K}_{s,\alpha}}^* = T_{\mathscr{K}_{s,\alpha}^\sharp}$, where

$$\mathscr{K}_{s,\alpha}^\sharp(x, y) = \sum_{z \in \mathbb{Z}^n} \mathscr{K}_{s,\alpha}(x, z) \overline{\mathscr{K}_{s,\alpha}(y, z)}.$$

From (6.1),

$$\mathscr{K}_{s,\alpha}^\sharp(x, y) = \sum_{\substack{\beta \in \mathcal{B}_s(\alpha(x)) \cap [0,1)^n, \\ \beta' \in \mathcal{B}_s(\alpha(y)) \cap [0,1)^n}} S(\alpha(x), \beta) e(x \cdot \beta) \overline{S(\alpha(y), \beta')} e(-y \cdot \beta')$$
$$\times \left[\sum_{z \in \mathbb{Z}^n} \phi_s(x - z) \overline{\phi_s(y - z)} e(z \cdot (\beta' - \beta)) \right].$$

Next we claim that for every $\beta, \beta' \in \mathcal{B}_s^\sharp \cap [0, 1)^n$ with $\beta \neq \beta'$,

$$\sum_{z \in \mathbb{Z}^n} \phi_s(x - z) \overline{\phi_s(y - z)} e(z \cdot (\beta' - \beta)) = 0. \tag{6.2}$$

To see this, define a Schwartz function on $\mathbb{R}^n$ by

$$\Xi(t) = \phi_s(x - t) \overline{\phi_s(y - t)} e(t(\beta' - \beta)) \quad (t \in \mathbb{R}^n).$$

Then

$$\widehat{\Xi}(\xi) = [M_x \widetilde{\chi}_s * M_{-y} \widetilde{\chi}_s](\xi + \beta - \beta'),$$

where we have used the notation $M_u g(x) = e(u \cdot x) g(x)$. From the definitions of $\widetilde{\chi}_s$ and $\mathcal{B}_s^\sharp$ we then see for $\xi \in \mathbb{Z}^n$ that $\widehat{\Xi}(\xi) = 0$ unless $\xi + \beta - \beta' = 0$. However, $\beta, \beta' \in [0, 1)^n$ and $\beta \neq \beta'$ imply $\beta - \beta' \notin \mathbb{Z}^n$. Hence, by the Poisson summation formula the left-hand side of (6.2) is equal to

$$\sum_{z \in \mathbb{Z}^n} \Xi(z) = \sum_{\xi \in \mathbb{Z}^n} \widehat{\Xi}(\xi) = 0.$$

As a consequence,

$$\mathcal{R}_{s,\alpha}^\sharp(x, y) = \overline{\kappa_{s,\alpha}(x, y)} \cdot [\phi_s * \overline{\phi_s}](x - y), \tag{6.3}$$

where we set

$$\kappa_{s,\alpha}(x, y) = \sum_{\beta \in \mathcal{B}_s(\alpha(x)) \cap \mathcal{B}_s(\alpha(y)) \cap [0,1)^n} S(\alpha(y), \beta) \overline{S(\alpha(x), \beta)} e((y - x) \cdot \beta). \tag{6.4}$$

For the following computation we fix $(x, y) \in \mathbb{Z}^n \times \mathbb{Z}^n$ and write

$$a = a(y), \quad q = q(y), \quad a' = a(x), \quad q' = q(x)$$

for short. As a consequence of Lemma 2.3, we may assume $(a, q) = (a', q') = 1$ and read the sum over β in (6.4) as running over the set

$$\{\mathbf{b}/q : \mathbf{b} \in [q]^n\} \cap \{\mathbf{b}/q' : \mathbf{b} \in [q']^n\},$$

which is equal to

$$\{\mathbf{b}/q_\flat : \mathbf{b} \in [q_\flat]^n\},$$

where we have set $q_\flat = (q, q')$. Thus,

$$\kappa_{s,\alpha}(x, y) = \sum_{\mathbf{b} \in [q_\flat]^n} S(a/q, \mathbf{b}/q_\flat) \overline{S(a'/q', \mathbf{b}/q_\flat)} e((y - x) \cdot \mathbf{b}/q_\flat).$$

Expanding the exponential sums by (2.5), we can rewrite this as

$$(qq')^{-n} \sum_{r \in [q]^n, r' \in [q']^n} e\left(\frac{a}{q}|r|^{2d} - \frac{a'}{q'}|r'|^{2d}\right) \left[\sum_{\mathbf{b} \in [q_\flat]^n} e\left(\frac{\mathbf{b}}{q_\flat} \cdot (r - r' + y - x)\right)\right],$$

which, in view of the relation $N^{-1} \sum_{l \in [N]} e(lz/N) = \mathbf{1}_{z \equiv 0 \,(\text{mod } N)}$, is equal to

$$\left(\frac{q'}{q_\flat}\right)^{-n} \sum_{u \in [q'/q_\flat]^n} q^{-n} \sum_{r \in [q]^n} e\left(\frac{a}{q}|r|^{2d} - \frac{a'}{q'}|r + y - x + u \cdot q_\flat|^{2d}\right). \tag{6.5}$$

Inspection of this exponential sum reveals several scenarios in which no cancellation can be expected. For instance, a typical case where (6.5) exhibits no cancellation is when

$a = a', q = q'$ and $y - x$ is divisible by q (then $\kappa_{s,\alpha}(x, y) = 1$). Additional degeneracies arise in the case $d = 1$, requiring a more careful analysis. For $w \in \mathbb{Z}^n$ we define

$$S_{x,y}(w) = q^{-n} \sum_{r \in [q]^n} e\left(\frac{a}{q}|r|^{2d} - \frac{a'}{q'}|r + w|^{2d}\right). \tag{6.6}$$

In the case $d \geq 2$ it will suffice to exploit cancellation from the exponential sum (6.6), whereas in the case $d = 1$ we will sometimes need to make use of cancellation from the sum over u in (6.5).

The case $d \geq 2$. Viewing the phase in (6.6) as a polynomial in r, the coefficient of r_1^{2d-1} is equal to $-2da'w_1/q'$. This leads us to define

$$\mathcal{E}_x = \{w \in \mathbb{Z}^n : (2dw_1, \mathsf{q}(x)) \geq 2^{s/2}\}.$$

By sorting modulo $\mathsf{q}(x)$ and counting divisors of $\mathsf{q}(x)$ we see that for $z \in \mathbb{Z}^n$, $N \geq 2^s$ and every $\varepsilon > 0$,

$$N^{-n}|\mathcal{E}_x \cap (z + [N]^n)| \lesssim_\varepsilon 2^{-s/2+\varepsilon s}. \tag{6.7}$$

If $w \notin \mathcal{E}_x$, then Proposition 2.2 (with $R = q \leq R^{(2d-1)-\varepsilon}$, crucially using $d \geq 2$) yields

$$|S_{x,y}(w)| \lesssim 2^{-\gamma s}$$

for some sufficiently small $\gamma \in (0, 1/2)$ depending on d and n. Using the triangle inequality on the sum over u in (6.5) leads to the estimation

$$|\kappa_{s,\alpha}(x, y)| \lesssim 2^{-\gamma s} + \sum_{v|\mathsf{q}(x)} (\mathsf{q}(x)/v)^{-n} \sum_{u \in [\mathsf{q}(x)/v]^n} \mathbf{1}_{\mathcal{E}_x}(y - x + u \cdot v), \tag{6.8}$$

where we have removed the (x, y)-dependence of $q_\flat = (\mathsf{q}(x), \mathsf{q}(y))$ by summing over all divisors of $\mathsf{q}(x)$. Hence, recalling (6.3), we see for every $x \in \mathbb{Z}^n$ that

$$\sum_{y \in \mathbb{Z}^n} |\mathcal{R}_{s,\alpha}^\sharp(x, y)| \lesssim 2^{-\gamma s} + \tau(\mathsf{q}(x)) \sup_{u \in \mathbb{Z}^n} \sum_{y \in \mathbb{Z}^n} \mathbf{1}_{\mathcal{E}_x}(y - x + u)|\phi_s * \phi_s|(x - y),$$

where $\tau(q)$ denotes the number of divisors of q. Using the standard divisor bound $\tau(q) \lesssim_\varepsilon q^\varepsilon$, (6.7) and rapid decay of $\phi_s * \phi_s$, we obtain

$$\sum_{y \in \mathbb{Z}^n} |\mathcal{R}_{s,\alpha}^\sharp(x, y)| \lesssim 2^{-\gamma s}$$

for every $x \in \mathbb{Z}^n$. Since also $\mathcal{R}_{s,\alpha}^\sharp(x, y) = \overline{\mathcal{R}_{s,\alpha}^\sharp(y, x)}$, we infer from Schur's test that

$$\|T_{\mathcal{R}_{s,\alpha}^\sharp}\|_{\ell^2(\mathbb{Z}^n)\to\ell^2(\mathbb{Z}^n)} \lesssim 2^{-\gamma s}.$$

This concludes the proof of Proposition 6.1 in this case.

The case $d = 1$. First assume that $q_\flat = (\mathsf{q}(x), \mathsf{q}(y)) \leq 2^{s/3}$. Then $q'/q_\flat \geq 2^{2s/3-1}$. Viewing the phase in (6.5) as a polynomial in u, the coefficient of u_1^2 is $-a'q_\flat^2/q'$, which equals

a reduced rational with denominator in $[2^{s/3}, 2^s] \cap \mathbb{Z}$. Thus, applying Proposition 2.2 to the exponential sum over u yields

$$|\kappa_{s,\alpha}(x, y)| \lesssim 2^{-\gamma s}$$

for a small enough $\gamma > 0$. Next we handle the case that $q_\flat \geq 2^{s/3}$. We will exploit cancellation from the summation over r in (6.5). The exponential sum on the right-hand side of (6.6) factors into n one-dimensional sums. It will be enough to estimate the first factor, which is given by

$$I = q^{-1} \sum_{r_1 \in [q]} e\left(\frac{A}{Q} r_1^2 - \frac{2a'w_1}{q'} r_1\right),$$

where $A/Q = a/q - a'/q'$ with $(A, Q) = 1$. We are led to distinguish two cases. Suppose that $Q \geq 2^{s/3}$. Then, since also $Q \leq qq'/q_\flat \leq 2^{5s/3}$, we may apply Proposition 2.2 to obtain

$$|I| \lesssim 2^{-\delta s} \tag{6.9}$$

for some small enough $\delta > 0$. On the other hand, assume $Q \leq 2^{s/3}$. Then, by reorganizing the summation modulo Q,

$$I = q^{-1} \left[\sum_{s \in [Q]} e\left(\frac{A}{Q} s^2 - \frac{2a'w_1}{q'} s\right)\right] \cdot \left[\sum_{u \in [M]} e\left(-\frac{2a'w_1 Q}{q'} u\right)\right] + O(2^{-2s/3}),$$

where $M = \lfloor q/Q \rfloor$. Summing the geometric sum over u and using the triangle inequality on the sum over s we get

$$|I| \lesssim 2^{-2s/3} |1 - e(2a'w_1 Q/q')|^{-1} \lesssim 2^{-2s/3} |2a'w_1 Q/q'|_{\mathbb{T}}^{-1}, \tag{6.10}$$

where $|\xi|_{\mathbb{T}} = \min_{z \in \mathbb{Z}} |\xi + z|$. Note that Q depends on both x and y. To remove the dependence on y we define for a positive integer $\upsilon \leq 2^{s/3}$ the set

$$E_x^{(\upsilon)} = \{w_1 \in \mathbb{Z} : |2a'w_1 \upsilon/q'|_{\mathbb{T}} \leq 2^{-s/2}\}. \tag{6.11}$$

Let $\iota = (2\upsilon, q') \lesssim 2^{s/3}$ and $\mathfrak{q} = q'/\iota$. Let $\mathcal{R} \subset \mathbb{Z}$ be a complete residue system modulo $\mathfrak{q}$. Then $a' \frac{2\upsilon}{\iota} \mathcal{R}$ is also a complete residue system modulo $\mathfrak{q}$. Thus

$$|E_x^{(\upsilon)} \cap \mathcal{R}| = |\{\ell \in [\mathfrak{q}] : |\ell/\mathfrak{q}|_{\mathbb{T}} \leq 2^{-s/2}\}| \lesssim \mathfrak{q} 2^{-s/2}.$$

Since $\mathfrak{q} \leq q' < 2^s$, we then have, for every $N \geq 2^s$ and $z \in \mathbb{Z}$,

$$N^{-1} |E_x^{(\upsilon)} \cap (z + [N])| \lesssim 2^{-s/2}.$$

Define

$$\mathcal{E}_x = \bigcup_{\upsilon \leq 2^{s/3}} \{w \in \mathbb{Z}^n : w_1 \in E_x^{(\upsilon)}\}. \tag{6.12}$$

Then if $w \notin \mathcal{E}_x$, we gather from (6.6) and (6.9)–(6.11) that

$$|S_{x,y}(w)| \leq |I| \lesssim \max(2^{-s/6}, 2^{-\delta s})$$

and for every $z \in \mathbb{Z}^n$ and $N \geq 2^s$,

$$N^{-n}|\mathcal{E}_x \cap (z + [N]^n)| \lesssim 2^{-s/6}.$$

The fact that we have chosen the exceptional set $\mathcal{E}_x$ only depending on x (as opposed to both x and y) allows us to recycle the crude argument using Schur's test seen in the case $d \geq 2$. Indeed, summarizing the above we have shown that (6.8) again holds for all $(x, y) \in \mathbb{Z}^n \times \mathbb{Z}^n$ with $\mathcal{E}_x$ defined as in (6.12) (and $\gamma > 0$ small enough, possibly different from above). This completes the proof of Proposition 6.1.

7. Major arcs II: Proof of Proposition 3.4

Before we begin with the proof we collect some preliminary results. First note from (3.13) that for every bounded function m on $\mathbb{R}^n$,

$$\mathscr{L}_{s,\alpha}[m](\mathrm{D})f(x) = \sum_{\beta \in \mathcal{B}_s(\alpha) \cap [0,1)^n} S(\alpha, \beta)e(x \cdot \beta)\mathscr{F}_{\mathbb{R}^n}^{-1}[m\,\chi_s] * M_{-\beta}f(x), \quad (7.1)$$

where $M_a f(x) = e(x \cdot a)f(x)$ denotes modulation by a (and $*$ must denote convolution on $\mathbb{Z}^n$, since f is only defined on $\mathbb{Z}^n$). The factorization (3.12) will allow us to prove the following consequence of Proposition 3.3.

Lemma 7.1. *Let $\mathcal{I}$ be a countable set and $(m_\nu)_{\nu \in \mathcal{I}}$ a family of bounded functions on $\mathbb{R}^n$. Then there exists $\gamma > 0$ such that for every $s \geq 1$,*

$$\left\| \sup_{\alpha \in \mathcal{A}_s} |\mathscr{L}_{s,\alpha}[m_\nu](\mathrm{D})f(x)| \right\|_{\ell^2_{x,\nu}(\mathbb{Z}^n \times \mathcal{I})} \lesssim_{d,n} 2^{-\gamma s}\left(\sup_{|\xi| \leq 1} \|m_\nu(\xi)\|_{\ell^2_\nu(\mathcal{I})} \right)\|f\|_{\ell^2(\mathbb{Z}^n)}.$$

Observe that Proposition 3.3 is the special case $\mathcal{I} = \{\cdot\}$, $m. \equiv 1$. We also need the following result which can be seen as a variant of Lemma 7.1 (in the case $\mathcal{I} = \{\cdot\}$) for maximally truncated singular integrals.

Lemma 7.2. *For $j \geq 1$ let $\mathcal{K}_j$ be a mean zero C^1 function supported on $\{|x| \asymp 2^j\}$ such that there exists a constant $A > 0$ with*

$$2^{jn}|\mathcal{K}_j(x)| + 2^{j(n+1)}|\nabla \mathcal{K}_j(x)| \leq A \quad (7.2)$$

for all $j \geq 1$ and $x \in \mathbb{R}^n$. Write

$$\mathcal{K}^{a,b}(x) = \sum_{a \leq j < b} \mathcal{K}_j(x).$$

Then there exists $\gamma > 0$ such that for all $s \geq 1$,

$$\left\| \sup_{J \geq 1} \sup_{\alpha \in \mathcal{A}_s} |\mathscr{L}_{s,\alpha}[\widehat{\mathcal{K}^{0,J}}](\mathrm{D})f| \right\|_{\ell^2(\mathbb{Z}^n)} \lesssim 2^{-\gamma s}\|f\|_{\ell^2(\mathbb{Z}^n)}. \quad (7.3)$$

Here the implicit constant depends only on A, d, n, and $\widehat{}$ denotes the Fourier transform on $\mathbb{R}^n$.

The proofs of Lemmas 7.1 and 7.2 are postponed to §7.4 and §7.5, respectively. We now begin with the proof of Proposition 3.4. The maximal operator in question can be bounded by

$$\sup_{\alpha \in \mathscr{A}_s} \sup_{|\mu| \leq 2^{-C_0 s}} \Big| \sum_{j \geq \varepsilon_1^{-1} s} \mathscr{L}_{s,\alpha}[\Phi_{j,\mu}](\mathrm{D}) f(x) \Big|, \tag{7.4}$$

where $C_0 = (2d - 1)\varepsilon_1^{-1}$ is a large constant.

Motivated by phase considerations, we introduce a frequency scale parameter $\ell \in \mathbb{Z}$ to partition the scales j for every fixed μ as

$$\mathscr{J}_{\ell,\mu} = \{j \geq \varepsilon_1^{-1} s : |\mu| 2^{2dj} \asymp 2^{\ell}\},$$

where $A \asymp B$ means $\frac{1}{2} B \leq A \leq 2B$. Note that $\mathscr{J}_{\ell,\mu}$ has at most one element and is often empty. Define

$$\widetilde{\Phi}_{\ell,\mu}(\xi) = \sum_{j \in \mathscr{J}_{\ell,\mu}} \Phi_{j,\mu}(\xi).$$

Let C_1 be a large positive constant that is to be determined later. We distinguish three cases:

$$\begin{aligned} \mathscr{L}_1 &= \{\ell \in \mathbb{Z} : \ell \geq C_1 s\}, \\ \mathscr{L}_2 &= \{\ell \in \mathbb{Z} : -C_1 s < \ell < C_1 s\}, \\ \mathscr{L}_3 &= \{\ell \in \mathbb{Z} : \ell \leq -C_1 s\}, \end{aligned}$$

and bound the ℓ^2 norm of (7.4) accordingly by

$$\sum_{i=1}^{3} \Big\| \sup_{\alpha \in \mathscr{A}_s} \sup_{|\mu| \leq 2^{-C_0 s}} \Big| \sum_{\ell \in \mathscr{L}_i} \mathscr{L}_{s,\alpha}[\widetilde{\Phi}_{\ell,\mu}](\mathrm{D}) f \Big| \Big\|_{\ell^2(\mathbb{Z}^n)}. \tag{7.5}$$

It remains to bound these three summands separately, which is the content of the next three subsections.

7.1. The high frequency case: $\ell \in \mathscr{L}_1$

This is the easiest case. Here $\ell \geq C_1 s$. From Stein and Wainger's theorem [11] we see

$$\Big\| \sup_{\mu \in \mathbb{R}} |\widetilde{\Phi}_{\ell,\mu}(\mathrm{D})g| \Big\|_{L^2(\mathbb{R}^n)} \lesssim 2^{-\gamma \ell} \|g\|_{L^2(\mathbb{R}^n)}$$

for some $\gamma > 0$. Since

$$\#\mathscr{R}_s \lesssim 2^{Cs} \quad \text{for some } C > 0,$$

we may choose C_1 large enough so that the trivial bound $|S(\alpha, \beta)| \leq 1$, the triangle inequality on the summation over β and a transference argument using Lemma 2.1 imply that the first summand in (7.5) is $\lesssim 2^{-\gamma' s} \|f\|_{\ell^2(\mathbb{Z}^n)}$ for some $\gamma' > 0$.

7.2. The intermediate frequency case: $\ell \in \mathcal{L}_2$

First note that $\#\mathcal{L}_2 \lesssim s$, so we may consider the terms for each fixed $\ell \in \mathcal{L}_2$ separately. The arguments differ slightly depending on the sign of ℓ with the terms for $\ell \geq 0$ being the more problematic ones. Set $\ell^+ = \max(0, \ell)$. We begin estimating

$$\left\| \sup_{\alpha \in \mathcal{A}_s} \sup_{|\mu| \leq 2^{-C_0 s}} |\mathscr{L}_{s,\alpha}[\widetilde{\Phi}_{\ell,\mu}](\mathrm{D})f| \right\|_{\ell^2(\mathbb{Z}^n)}$$

by

$$\left\| \sup_{j \geq \varepsilon_1^{-1} s} \sup_{\alpha \in \mathcal{A}_s} \sup_{|\mu| \asymp 2^{\ell-2dj}} |\mathscr{L}_{s,\alpha}[\Phi_{j,\mu}](\mathrm{D})f| \right\|_{\ell^2(\mathbb{Z}^n)}. \tag{7.6}$$

Let ϑ denote an appropriately compactly supported, smooth and nonnegative function such that $\int \vartheta = 1$ and let $\vartheta_j(x) = 2^{-jn}\vartheta(2^{-j}x)$. With μ such that $|\mu| \asymp 2^{\ell-2dj}$ we define

$$\Phi_{j,\mu}^\flat(\xi) = \Phi_{j,\mu}(\xi) - \Phi_{j,\mu}(0)\widehat{\vartheta_{j-\ell^+}}(\xi).$$

The definition is made so that $\Phi_{j,\mu}^\flat(0) = 0$ and $\Phi_{j,\mu}^\flat$ satisfies favorable estimates that will be stated below (see also [3, Lemma 3.22]). Observe that $|\Phi_{j,\mu}(0)| \lesssim 1$ (there is also decay in $|\ell|$ but we will only need that later on). We now estimate (7.6) by the sum of

$$\left(\sum_{j \geq \varepsilon_1^{-1} s} \left\| \sup_{\alpha \in \mathcal{A}_s} \sup_{|\mu| \asymp 2^{\ell-2dj}} |\mathscr{L}_{s,\alpha}[\Phi_{j,\mu}^\flat](\mathrm{D})f| \right\|_{\ell^2(\mathbb{Z}^n)}^2 \right)^{1/2} \tag{7.7}$$

and

$$\left\| \sup_{j \geq \varepsilon_1^{-1} s} \sup_{\alpha \in \mathcal{A}_s} |\mathscr{L}_{s,\alpha}[\widehat{\vartheta_{j-\ell^+}}](\mathrm{D})f| \right\|_{\ell^2(\mathbb{Z}^n)}. \tag{7.8}$$

We begin by handling the latter term. Since $\ell^+ \leq C_1 s$ we can bound (7.8) by

$$\left\| \sup_{J \geq 1} \sup_{\alpha \in \mathcal{A}_s} |\mathscr{L}_{s,\alpha}[\widehat{\vartheta_J}](\mathrm{D})f| \right\|_{\ell^2(\mathbb{Z}^n)} \tag{7.9}$$

as long as ε_1 is chosen small enough so that $\varepsilon_1^{-1} \geq C_1$ (which is possible, since the choice of C_1 is independent of that of ε_1). Expanding into a telescoping sum

$$\vartheta_J = \vartheta_0 + \sum_{0 \leq j < J} (\vartheta_{j+1} - \vartheta_j),$$

we bound (7.9) by

$$\left\| \sup_{\alpha \in \mathcal{A}_s} |\mathscr{L}_{s,\alpha}[\widehat{\vartheta_0}](\mathrm{D})f| \right\|_{\ell^2(\mathbb{Z}^n)} + \left\| \sup_{J \geq 1} \sup_{\alpha \in \mathcal{A}_s} \left| \mathscr{L}_{s,\alpha}\left[\sum_{0 \leq j < J} (\widehat{\vartheta_{j+1}} - \widehat{\vartheta_j}) \right](\mathrm{D})f \right| \right\|_{\ell^2(\mathbb{Z}^n)}. \tag{7.10}$$

Using Lemma 7.1 for the first term and Lemma 7.2 for the second term we see that the previous display is $\lesssim 2^{-\gamma s}\|f\|_{\ell^2(\mathbb{Z}^n)}$ as required.

It remains to estimate the main term (7.7). The uncertainty principle suggests that the value of $|\Phi^\flat_{j,\mu}(D)g(x)|$ stays approximately constant as μ varies over an interval of length $\lesssim 2^{-2dj}$. This motivates the following standard argument (similar to §5). Define $I = \{|\mu| \asymp 2^{\ell-2dj}\} = [-b, -a] \cup [a, b]$ with $0 < a < b$, set $\delta = 2^{-2dj}$ and define

$$\mathfrak{F} = \mathfrak{F}_{\ell,j} = \bigcup_{0 \le k < \lceil (b-a)/\delta \rceil} \{a + k\delta, -b + k\delta\}.$$

The set $\mathfrak{F}_{\ell,j}$ has cardinality $\lceil 3 \cdot 2^\ell \rceil \approx 2^{\ell^+}$. For every differentiable function $G : I \to \mathbb{C}$ we have, by the fundamental theorem of calculus,

$$\sup_{\mu \in I} |G(\mu)| \le \|G\|_{\ell^2(\mathfrak{F})} + \delta \int_0^1 \|G'(\mu + t\delta)\|_{\ell^2_\mu(\mathfrak{F})} \, dt.$$

Using this with

$$G(\mu) = \mathscr{L}_{s,\alpha}[\Phi^\flat_{j,\mu}](D)f(x)$$

we bound (7.7) by

$$\left(\sum_{j \ge \varepsilon_1^{-1}s} \sum_{\mu \in \mathfrak{F}_{\ell,j}} \left\| \sup_{\alpha \in \mathscr{A}_s} |\mathscr{L}_{s,\alpha}[\Phi^\flat_{j,\mu}](D)f| \right\|^2_{\ell^2(\mathbb{Z}^n)} \right)^{1/2}$$

$$+ \int_0^1 \left(\sum_{j \ge \varepsilon_1^{-1}s} \sum_{\mu \in \mathfrak{F}_{\ell,j}} 2^{-4dj} \left\| \sup_{\alpha \in \mathscr{A}_s} |\mathscr{L}_{s,\alpha}[\partial_\mu \Phi^\flat_{j,\mu+2^{-2dj}t}](D)f| \right\|^2_{\ell^2(\mathbb{Z}^n)} \right)^{1/2} dt.$$

Applying Lemma 7.1, we see that (7.7) is $\le A \, 2^{-\gamma s} \|f\|_{\ell^2(\mathbb{Z}^n)}$ with

$$A^2 \lesssim \sup_{|\xi| \le 1} \sum_{j \ge \varepsilon_1^{-1}s} \sum_{\mu \in \mathfrak{F}_{\ell,j}} |\Phi^\flat_{j,\mu}(\xi)|^2$$

$$+ \sup_{|\xi| \le 1} \sup_{t \in [0,1]} \sum_{j \ge \varepsilon_1^{-1}s} \sum_{\mu \in \mathfrak{F}_{\ell,j}} 2^{-4dj} |\partial_\mu \Phi^\flat_{j,\mu+2^{-2dj}t}(\xi)|^2.$$

To show that $A \lesssim 1$ we will use the following estimates. Let $j \ge 1$ and $|\mu| \asymp 2^{\ell-2dj}$. If $\ell \ge 0$, then for all $N \ge 0$,

$$|\Phi^\flat_{j,\mu}(\xi)| + 2^{-2dj}|\partial_\mu \Phi^\flat_{j,\mu}(\xi)| \lesssim_N 2^{-\ell n/2}\mathbf{1}_{|\xi| \approx 2^{\ell-j}} + 2^{-\ell N}2^j|\xi|\mathbf{1}_{|\xi| \lesssim 2^{\ell-j}}$$
$$+ (2^j|\xi|)^{-N}\mathbf{1}_{|\xi| \gtrsim 2^{\ell-j}}. \tag{7.11}$$

If $\ell < 0$, then for all $N \ge 0$,

$$|\Phi^\flat_{j,\mu}(\xi)| + 2^{-2dj}|\partial_\mu \Phi^\flat_{j,\mu}(\xi)| \lesssim_N \min(2^j|\xi|, (2^j|\xi|)^{-N}). \tag{7.12}$$

These estimates imply that $A \lesssim 1$. Both estimates follow from direct computation using the definitions. We indicate some of the details for the term $\Phi^\flat_{j,\mu}(\xi)$; the term $2^{-2dj}\partial_\mu \Phi^\flat_{j,\mu}(\xi)$ satisfies the same estimates and is handled in the same way. First suppose

that $\ell \geq 0$. If $|\xi| \approx 2^{\ell-j}$, we use van der Corput's lemma, which gives $|\Phi_{j,\mu}(\xi)| \lesssim 2^{-\ell n/2}$. Moreover, $|\Phi_{j,\mu}(0)| \lesssim_N 2^{-\ell N}$ for all $N \geq 0$ using integration by parts. If $|\xi| \gtrsim 2^{\ell-j}$, then integration by parts gives $|\Phi_{j,\mu}(\xi)| \lesssim_N (2^j|\xi|)^{-N}$ for all $N \geq 0$. Also using rapid decay of the Schwartz function $\widehat{\vartheta_{j-\ell}}$ we obtain the claim. If $|\xi| \lesssim 2^{\ell-j}$, we use $\Phi^{\flat}_{j,\mu}(0) = 0$, the mean value theorem and the estimate $|\nabla \Phi^{\flat}_{j,\mu}(\xi)| \lesssim_N 2^j 2^{-\ell N}$. It remains to consider (7.12). Suppose that $\ell < 0$. If $|\xi| \lesssim 2^{-j}$, we argue in the same way as before using the mean value theorem. If $|\xi| \gtrsim 2^{-j}$, integration by parts and rapid decay of $\widehat{\vartheta_j}$ give $|\Phi^{\flat}_{j,\mu}(\xi)| \lesssim_N (2^j|\xi|)^{-N}$ as required.

7.3. The low frequency case: $\ell \in \mathcal{L}_3$

The third summand in (7.5) can be written as

$$\left\| \sup_{\alpha \in \mathcal{A}_s} \sup_{|\mu| \leq 2^{-C_0 s}} \left| \sum_{J_- \leq j \leq J_{+,\mu}} \sum_{\beta \in \mathcal{B}_s(\alpha) \cap [0,1)^n} S(\alpha, \beta) e(x \cdot \beta) \mathcal{F}^{-1}_{\mathbb{R}^n}[\Phi_{j,\mu} \cdot \chi_s] \right. \right.$$
$$\left. \left. * \, M_{-\beta} f(x) \right| \right\|_{\ell^2_x(\mathbb{Z}^n)}, \qquad (7.13)$$

where $J_- = \varepsilon_1^{-1} s$ and $J_{+,\mu}$ is the largest integer j such that $\mu 2^{2dj} < 2^{-C_1 s+1}$. By the triangle inequality, the previous display is

$$\lesssim 2^{(n+1)s} \left\| \sup_{|\mu| \leq 2^{-C_0 s}} \tau_\mu * |f| \right\|_{\ell^2_x(\mathbb{Z}^n)}$$
$$+ \left\| \sup_{\alpha \in \mathcal{A}_s} \sup_{J \geq J_-} \left| \mathcal{L}_{s,\alpha} \left[\sum_{J_- \leq j \leq J} \widehat{K_j} \right](\mathrm{D}) f \right| \right\|_{\ell^2(\mathbb{Z}^n)}, \qquad (7.14)$$

where $\tau_\mu = |\widetilde{\tau}_\mu *_{\mathbb{R}^n} \mathcal{F}^{-1}_{\mathbb{R}^n}(\chi_s)|$ and

$$\widetilde{\tau}_\mu(y) = \sum_{J_- \leq j \leq J_{+,\mu}} \left(\mathcal{F}^{-1}_{\mathbb{R}^n}[\Phi_{j,\mu}](y) - K_j(-y) \right) = \sum_{J_- \leq j \leq J_{+,\mu}} (e(\mu|y|^{2d}) - 1) K_j(-y).$$

Using the definition of $J_{+,\mu}$ we obtain $\tau_\mu * |f| \lesssim 2^{-C_1 s} M_{\mathrm{HL}} f$, where M_{HL} denotes the discrete Hardy–Littlewood maximal function. If we choose $C_1 > n + 1$, this takes care of the first term in (7.14), while the second term is handled by an application of Lemma 7.2 (setting $\mathcal{K}_j = 0$ for $0 \leq j < J_-$).

7.4. Proof of Lemma 7.1

By (3.12),
$$\mathcal{L}_{s,\alpha}[m_\nu](\mathrm{D}) f = \mathcal{L}_{s,\alpha}[1](\mathrm{D}) \left(\mathcal{L}^{\sharp}_s[m_\nu](\mathrm{D}) f \right).$$

From Proposition 3.3 we obtain $\gamma > 0$ such that for every $\nu \in \mathcal{J}$ and $s \geq 1$,

$$\left\| \sup_{\alpha \in \mathbb{R}} |\mathcal{L}_{s,\alpha}[m_\nu](\mathrm{D}) f(x)| \right\|_{\ell^2_x(\mathbb{Z}^n)} \lesssim_{d,n} 2^{-\gamma s} \| \mathcal{L}^{\sharp}_s[m_\nu](\mathrm{D}) f \|_{\ell^2(\mathbb{Z}^n)}.$$

Using (3.11) and Parseval's identity, we get

$$\|\mathscr{L}_s^\sharp[m_\nu](\mathrm{D})f\|_{\ell^2(\mathbb{Z}^n)}^2 = \int_{[0,1]^n}\Big|\sum_{\beta\in\mathcal{B}_s^\sharp} m_\nu(\xi-\beta)\widetilde{\chi}_s(\xi-\beta)\widehat{f}(\xi)\Big|^2 d\xi.$$

Summing over $\nu\in\mathcal{J}$ and using disjointness of the supports of the functions $\widetilde{\chi}_s(\cdot-\beta)$ for different β we have

$$\|\mathscr{L}_s^\sharp[m_\nu](\mathrm{D})f(x)\|_{\ell_{x,\nu}^2(\mathbb{Z}^n\times\mathcal{J})}^2 = \int_{[0,1]^n}\sum_{\nu\in\mathcal{J}}\sum_{\beta\in\mathcal{B}_s^\sharp}|m_\nu(\xi-\beta)\widetilde{\chi}_s(\xi-\beta)\widehat{f}(\xi)|^2 d\xi$$

$$\leq\Big(\sup_{\xi\in[0,1]^n}\sum_{\nu\in\mathcal{J}}\sum_{\beta\in\mathcal{B}_s^\sharp}|m_\nu(\xi-\beta)\widetilde{\chi}_s(\xi-\beta)|^2\Big)\|f\|_{\ell^2(\mathbb{Z}^n)}^2,$$

which by definition of $\widetilde{\chi}_s$ (and disjointness again) is

$$\leq\Big(\sup_{|\xi|\leq 1}\sum_{\nu\in\mathcal{J}}|m_\nu(\xi)|^2\Big)\|f\|_{\ell^2(\mathbb{Z}^n)}^2.$$

7.5. Proof of Lemma 7.2

From the assumptions we have the standard estimate

$$|\widehat{\mathcal{K}_j}(\xi)|\lesssim\min(2^j|\xi|,(2^j|\xi|)^{-1}) \tag{7.15}$$

for all $\xi\in\mathbb{R}^n$ and $j\geq 1$. We distinguish two cases: either $J\leq 2^{Cs}$ or $J>2^{Cs}$. Here C is a large constant (to be determined).

Case I: $J\leq 2^{Cs}$. Using the numerical inequality (2.3), we bound $|\mathscr{L}_{s,\alpha}[\widehat{\mathcal{K}^{0,J}}](\mathrm{D})f(x)|$ by

$$\sqrt{2}\sum_{l\leq Cs}\Big(\sum_{\kappa\leq 2^{Cs-l}}|\mathscr{L}_{s,\alpha}[\widehat{\mathcal{K}^{\kappa 2^l,(\kappa+1)2^l}}](\mathrm{D})f(x)|^2\Big)^{1/2}$$

plus

$$|\mathscr{L}_{s,\alpha}[\widehat{\mathcal{K}_1}](\mathrm{D})f(x)|.$$

By Lemma 7.1 the maximal operator associated with the second term has $\ell^2\to\ell^2$ operator norm $\lesssim 2^{-\gamma s}$ as required. Hence, it remains to estimate

$$\Big(\sum_{\kappa\leq 2^{Cs-l}}\Big\|\sup_{\alpha\in\mathcal{A}_s}|\mathscr{L}_{s,\alpha}[\widehat{\mathcal{K}^{\kappa 2^l,(\kappa+1)2^l}}](\mathrm{D})f(x)|\Big\|_{\ell_x^2(\mathbb{Z}^n)}^2\Big)^{1/2}$$

for every fixed $l\leq Cs$. By Lemma 7.1 this quantity is

$$\lesssim 2^{-\gamma s}\sup_{|\xi|\leq 1}\Big(\sum_{\kappa\leq 2^{Cs-l}}|\widehat{\mathcal{K}^{\kappa 2^l,(\kappa+1)2^l}}(\xi)|^2\Big)^{1/2}\|f\|_{\ell^2(\mathbb{Z}^n)},$$

which in turn is $\lesssim 2^{-\gamma s}\|f\|_{\ell^2(\mathbb{Z}^n)}$ by (7.15). This concludes the proof for Case I.

Case II: $J > 2^{Cs}$. Here we will exploit the enormous magnitude of the spatial scale J. The argument is vaguely similar to the proof of Cotlar's inequality for maximally truncated Calderón–Zygmund operators. By subtracting a Calderón–Zygmund operator (leading to a term that no longer involves a supremum over J and is handled directly by Lemma 7.1) it suffices to consider

$$\left\| \sup_{\alpha \in \mathcal{A}_s} \sup_{J \geq 2^{Cs}} |\mathscr{L}_{s,\alpha}[\widehat{\mathcal{K}^{J,\infty}}](\mathrm{D})f| \right\|_{\ell^2(\mathbb{Z}^n)}. \tag{7.16}$$

Let φ denote a nonnegative Schwartz function on $\mathbb{R}^n$ with $\mathrm{supp}\,\widehat{\varphi} \subset \{|\xi| \leq 1/2\}$ and $\int \varphi = 1$. Write

$$\varphi_j(x) = 2^{-jn}\varphi(2^{-j}x).$$

We claim that it suffices to estimate

$$\left\| \sup_{\alpha \in \mathcal{A}_s} \sup_{J \geq 2^{Cs}} |\mathscr{L}_{s,\alpha}[\widehat{\varphi_J}\,\widehat{\mathcal{K}}](\mathrm{D})f| \right\|_{\ell^2(\mathbb{Z}^n)}. \tag{7.17}$$

where $\mathcal{K} = \mathcal{K}^{0,\infty}$. To see this we consider

$$\left\| \sup_{\alpha \in \mathcal{A}_s} \sup_{J \geq 2^{Cs}} |\mathscr{L}_{s,\alpha}[\widehat{\mathcal{K}^{J,\infty}} - \widehat{\varphi_J}\,\widehat{\mathcal{K}}](\mathrm{D})f| \right\|_{\ell^2(\mathbb{Z}^n)}. \tag{7.18}$$

This is majorized by

$$\left\| \sup_{\alpha \in \mathcal{A}_s} |\mathscr{L}_{s,\alpha}[\widehat{\mathcal{K}^{J,\infty}} - \widehat{\varphi_J}\,\widehat{\mathcal{K}}](\mathrm{D})f| \right\|_{\ell^2_{x,J}(\mathbb{Z}^n \times \mathbb{N}_0)},$$

which by Lemma 7.1 is

$$\lesssim 2^{-\gamma s} \sup_{|\xi| \leq 1} \left(\sum_{J \geq 1} |\widehat{\mathcal{K}^{J,\infty}}(\xi) - \widehat{\varphi_J}(\xi)\,\widehat{\mathcal{K}}(\xi)|^2 \right)^{1/2} \|f\|_{\ell^2(\mathbb{Z}^n)}.$$

From (7.15) and the definition of φ one derives

$$|\widehat{\mathcal{K}^{J,\infty}}(\xi) - \widehat{\varphi_J}(\xi)\widehat{\mathcal{K}}(\xi)| \lesssim \min(2^J|\xi|, (2^J|\xi|)^{-1}),$$

which concludes the proof of the claim. Thus it remains to bound (7.17). Using (7.1) and band-limitedness of φ we write

$$\mathscr{L}_{s,\alpha}[\widehat{\varphi_J}\,\widehat{\mathcal{K}}](\mathrm{D})f(x) = \sum_{\beta \in \mathcal{B}_s(\alpha) \cap [0,1)^n} S(\alpha,\beta)e(x \cdot \beta)(\varphi_J * \mathcal{K}_s * M_{-\beta}f)(x),$$

where the convolutions are in $\mathbb{Z}^n$ and $\mathcal{K}_s = \mathcal{F}_{\mathbb{R}^n}^{-1}(\widehat{\mathcal{K}} \cdot \chi_s)$.

Denote the least common multiple of all integers in the interval $[2^{s-1}, 2^s]$ by Q_s. By choosing C large enough we may achieve that if $J \geq 2^{Cs}$, then

$$2^J \geq Q_s^{100n}.$$

Then Q_s is small compared with the spatial scale of φ_J and this allows us to compare (7.17) favorably with the averaged version

$$\left(\frac{1}{Q_s^n}\sum_{u\in[Q_s]^n}\left\|\sup_{\alpha\in\mathcal{A}_s}\sup_{j\geq1}\left|\sum_{\beta\in\mathcal{B}_s(\alpha)\cap[0,1)^n}S(\alpha,\beta)e(x\cdot\beta)(\varphi_J*\mathcal{K}_s\right.\right.\right.$$
$$\left.\left.\left.*M_{-\beta}f)(x-u)\right|\right\|_{\ell_x^2(\mathbb{Z}^n)}^2\right)^{1/2}. \quad (7.19)$$

The resulting difference term can be handled by the mean value theorem and using the inequality $2^J\geq Q_s^{100n}$. More specifically, given $u\in[Q_s]^n$ we brutally estimate

$$\left\|\sup_{\alpha\in\mathcal{A}_s}\sup_{J\geq2^{Cs}}\left|\sum_{\beta\in\mathcal{B}_s(\alpha)\cap[0,1)^n}S(\alpha,\beta)e(x\cdot\beta)\big((\varphi_J*\mathcal{K}_s*M_{-\beta}f)(x)\right.\right.$$
$$\left.\left.-(\varphi_J*\mathcal{K}_s*M_{-\beta}f)(x-u)\big)\right|\right\|_{\ell_x^2(\mathbb{Z}^n)}.$$
$$\lesssim 2^{(n+1)s}\sup_{\beta\in[0,1)^n}\left\|\sup_{J\geq2^{Cs}}|(\varphi_J-\varphi_J(\cdot-u))*\mathcal{K}_s*M_{-\beta}f|\right\|_{\ell^2(\mathbb{Z}^n)}.$$

Since $|u|2^{-J}\leq Q_s^{-100n+1}\leq2^{-(100n+1)s}$, the previous display is (say)

$$\lesssim2^{-10s}\sup_{\beta\in[0,1)^n}\|M_{\mathrm{HL}}(\mathcal{K}_s*M_{-\beta}f)\|_{\ell^2(\mathbb{Z}^n)}\lesssim2^{-10s}\|f\|_{\ell^2(\mathbb{Z}^n)},$$

where M_{HL} denotes the discrete Hardy–Littlewood maximal function.

It now remains to estimate (7.19). Expanding the ℓ^2 norm and changing variables $x\mapsto x+u$ we write the square of (7.19) as

$$\frac{1}{Q_s^n}\sum_{u\in[Q_s]^n}\sum_{x\in\mathbb{Z}^n}\sup_{\alpha\in\mathcal{A}_s}\sup_{j\geq1}\left|\sum_{\beta\in\mathcal{B}_s(\alpha)\cap[0,1)^n}S(\alpha,\beta)e((x+u)\cdot\beta)(\varphi_J*\mathcal{K}_s*M_{-\beta}f)(x)\right|^2.$$

By changing variables $u\mapsto v-x$ and using periodicity (Q_s is divisible by the denominator of β for all $\beta\in\mathcal{B}_s^\sharp$) this becomes

$$\frac{1}{Q_s^n}\sum_{v\in[Q_s]^n}\sum_{x\in\mathbb{Z}^n}\sup_{\alpha\in\mathcal{A}_s}\sup_{j\geq1}\left|\sum_{\beta\in\mathcal{B}_s(\alpha)\cap[0,1)^n}S(\alpha,\beta)e(v\cdot\beta)(\varphi_J*\mathcal{K}_s*M_{-\beta}f)(x)\right|^2.$$

After having decoupled v and x we arrive at the pointwise estimate

$$\sup_{\alpha\in\mathcal{A}_s}\sup_{j\geq1}\left|\varphi_J*\sum_{\beta\in\mathcal{B}_s(\alpha)\cap[0,1)^n}S(\alpha,\beta)e(v\cdot\beta)(\mathcal{K}_s*M_{-\beta}f)\right|$$
$$\lesssim M_{\mathrm{HL}}\left(\sup_{\alpha\in\mathcal{A}_s}\left|\sum_{\beta\in\mathcal{B}_s(\alpha)\cap[0,1)^n}S(\alpha,\beta)e(v\cdot\beta)(\mathcal{K}_s*M_{-\beta}f)\right|\right).$$

Applying the ℓ^2 bound for M_{HL}, changing variables back $v\mapsto u+x$ and using Lemma 7.1 finishes the estimate.

References

[1] Bourgain, J.: Pointwise ergodic theorems for arithmetic sets. Inst. Hautes Études Sci. Publ. Math. **69**, 5–45 (1989) Zbl 0705.28008 MR 1019960

[2] Guo, S., Roos, J., Yung, P.-L.: Sharp variation-norm estimates for oscillatory integrals related to Carleson's theorem. Anal. PDE **13**, 1457–1500 (2020) Zbl 1452.42010 MR 4149067

[3] Krause, B., Lacey, M. T.: A discrete quadratic Carleson theorem on ℓ^2 with a restricted supremum. Int. Math. Res. Notices **2017**, 3180–3208 Zbl 1405.42034 MR 3658135

[4] Krause, B., Roos, J.: Discrete analogues of maximally modulated singular integrals of Stein–Wainger type, II. arXiv:2107.14616 (2021)

[5] Lewko, A., Lewko, M.: Estimates for the square variation of partial sums of Fourier series and their rearrangements. J. Funct. Anal. **262**, 2561–2607 (2012) Zbl 1239.42027 MR 2885959

[6] Mirek, M.: Square function estimates for discrete Radon transforms. Anal. PDE **11**, 583–608 (2018) Zbl 1383.42014 MR 3738256

[7] Mirek, M., Stein, E. M., Trojan, B.: $\ell^p(\mathbb{Z}^d)$-estimates for discrete operators of Radon type: variational estimates. Invent. Math. **209**, 665–748 (2017) Zbl 1397.42006 MR 3681393

[8] Mirek, M., Stein, E. M., Trojan, B.: $\ell^p(\mathbb{Z}^d)$-estimates for discrete operators of Radon type: maximal functions and vector-valued estimates. J. Funct. Anal. **277**, 2471–2521 (2019) Zbl 1444.42026 MR 3990725

[9] Mirek, M., Trojan, B.: Discrete maximal functions in higher dimensions and applications to ergodic theory. Amer. J. Math. **138**, 1495–1532 (2016) Zbl 1404.37006 MR 3595493

[10] Stein, E. M., Wainger, S.: Discrete analogues in harmonic analysis. I. l^2 estimates for singular Radon transforms. Amer. J. Math. **121**, 1291–1336 (1999) Zbl 0945.42009 MR 1719802

[11] Stein, E. M., Wainger, S.: Oscillatory integrals related to Carleson's theorem. Math. Res. Lett. **8**, 789–800 (2001) Zbl 0998.42007 MR 1879821