

AVERAGES ALONG POLYNOMIAL SEQUENCES IN DISCRETE NILPOTENT GROUPS: SINGULAR RADON TRANSFORMS

ALEXANDRU D. IONESCU, AKOS MAGYAR, AND STEPHEN WAINGER

ABSTRACT. We consider a class of operators defined by taking averages along polynomial sequences in discrete nilpotent groups. As in the continuous case, one can consider discrete maximal Radon transforms, which have applications to pointwise ergodic theorems, and discrete singular Radon transforms. In this paper we prove L^2 boundedness of discrete singular Radon transforms along general polynomial sequences in discrete nilpotent groups of step 2.

CONTENTS

1. Introduction	1
2. A transference argument	6
3. The main kernels: identities and estimates	10
4. Proof of Theorem 2.3	16
5. Estimates on oscillatory sums and oscillatory integrals	26
6. An almost orthogonality lemma	35
References	41

1. INTRODUCTION

A class of interesting problems arises in studying averages of functions along polynomial sequences in discrete nilpotent groups. More precisely, assume $\mathbb{G}$ is a discrete nilpotent group of step $d \geq 1$ and $A : \mathbb{Z} \rightarrow \mathbb{G}$ is a polynomial sequence (see Definition 1.1 below), and consider the following problems:¹

Problem 1. (L^2 boundedness of maximal Radon transforms) Assume $f : \mathbb{G} \rightarrow \mathbb{C}$ is a function and let

$$\mathcal{M}f(g) = \sup_{N \geq 0} \frac{1}{2N+1} \sum_{|n| \leq N} |f(A^{-1}(n) \cdot g)|, \quad g \in \mathbb{G}.$$

The first author was partially supported by a Packard Fellowship and NSF grant DMS-1065710. The second author was partially supported by NSERC grant 22R44824.

¹One can also state similar problems in the case of L^q functions, $q > 1$, or for multi-dimensional polynomial sequences $A : \mathbb{Z}^k \rightarrow \mathbb{G}$, $k \geq 1$.

Then

$$\|\mathcal{M}f\|_{L^2(\mathbb{G})} \lesssim \|f\|_{L^2(\mathbb{G})}.$$

Problem 2. (L^2 pointwise ergodic theorems) Assume $\mathbb{G}$ acts by measure-preserving transformations on a probability space X , $f \in L^2(X)$, and let

$$A_N f(x) = \frac{1}{2N+1} \sum_{|n| \leq N} f(A^{-1}(n) \cdot x), \quad x \in X.$$

Then the sequence $A_N f$ converges almost everywhere in X as $N \rightarrow \infty$.

Problem 3. (L^2 boundedness of singular Radon transforms) Assume $K : \mathbb{R} \rightarrow \mathbb{R}$ is a Calderon–Zygmund kernel (see (1.1)), $f : \mathbb{G} \rightarrow \mathbb{C}$ is a (compactly supported) function, and let

$$Hf(g) = \sum_{n \in \mathbb{Z}} K(n) f(A^{-1}(n) \cdot g), \quad g \in \mathbb{G}.$$

Then

$$\|Hf\|_{L^2(\mathbb{G})} \lesssim \|f\|_{L^2(\mathbb{G})}.$$

The maximal Radon transform and the singular Radon transform can be thought of as discrete analogues of the continuous Radon transforms, which are averages along suitable curves or surfaces in Euclidean spaces. The theory of continuous Radon transforms has been extensively studied and is very well understood (including L^q , $q > 1$, estimates and multidimensional averages), see for example [8], [20], [9].

In the discrete setting, the three questions raised above have been answered in the affirmative in the commutative case $\mathbb{G} = \mathbb{Z}^d$.² The maximal function estimate and the pointwise ergodic theorem were proved by Bourgain [6], [4], [5], also in the case of L^q functions, $q > 1$. L^2 estimates for singular Radon transforms were obtained in [1], the L^q boundedness was established in [24] for $3/2 < q < 3$ and were extended for all $q > 1$ in [13]. Closely related fractional integral operators were treated in [17], [26], [18], [19].

Only partial results are available, however, in the case non-commutative discrete nilpotents groups, even in the case of step 2 nilpotent groups. A general feature of the partial results obtained in the non-commutative setting, see [12], [16], [25], is that the averages are taken over surfaces transversal to the center of the group, such that the "non-linear" part of the polynomial map is contained in the center. The point is that for such special polynomial sequences one can still use the Fourier transform in the central variables to analyze the operators.

However, it appears that one needs to proceed in an entirely different way in the case of general polynomial maps, when the Fourier transform method is not available. The present work is the first attempt to treat discrete Radon transforms along general

²The linear case $\mathbb{G} = \mathbb{Z}$, $A(n) = n$, is, of course, well-known.

polynomial sequences in the non-commutative nilpotent settings. More precisely, we will discuss the easier Problem 3 in the case of discrete nilpotent groups of step 2.

Finally let us remark that the L^2 ergodic theorems of Bergelson and Leibman [2] indicate that nilpotent groups provide the most general settings to which the results of Bourgain might extend. Indeed, they have shown that averages of measure preserving transformations generating a nilpotent group converge in the mean along any polynomial sequence, however this does not hold for transformations generating a solvable group.

To describe our settings in detail, recall that a polynomial sequence on a nilpotent group $\mathbb{G}$ is a map $A : \mathbb{Z} \rightarrow \mathbb{G}$, such that $D^k A(n) = 1$ for all n for some fixed k , where D^k is the k -fold iterate of the differencing operator D defined by $DA(n) = A(n)^{-1}A(n+1)$. It is known, see [14] that A is a polynomial sequence if and only if $A(n) = g_1^{p_1(n)} \dots g_t^{p_t(n)}$ for all n , where $g_1, \dots, g_t$ are elements of $\mathbb{G}$ and $p_1, \dots, p_t$ are integral polynomials. In particular the image of the map A is contained in a finitely generated subgroup of $\mathbb{G}$, thus without the loss of generality we will assume that $\mathbb{G}$ is finitely generated and hence countable. We will also assume that $\mathbb{G}$ is torsion free and then, by a result of Malcev [15], the group $\mathbb{G}$ can be embedded as a discrete, co-compact subgroup of a (connected and simply connected) nilpotent Lie group $\mathbb{G}^\sharp$. This motivates the following:

Definition 1.1. *Given $d \geq 1$, a group $\mathbb{G}$ will be called a discrete nilpotent group of step d if $\mathbb{G}$ is isomorphic to a discrete, co-compact subgroup of a (connected and simply connected) nilpotent Lie group $\mathbb{G}^\sharp$ of step d .*

Given a group $\mathbb{G}$, a sequence $A : \mathbb{Z} \rightarrow \mathbb{G}$ will be called a polynomial sequence if $A(0) = 1$ and $D^{k_0} A \equiv 1$ for some $k_0 \geq 1$, where, by definition,

$$D^0 A(n) = A(n), \quad D^{k+1} A(n) = D^k A(n)^{-1} D^k A(n+1), \quad n \in \mathbb{Z}.$$

In this paper we consider only the easier problem of L^2 boundedness of the discrete singular Radon transforms. To formulate our main result, let $K : \mathbb{R} \rightarrow \mathbb{R}$ be a Calderon–Zygmund kernel, i.e. a C^1 function satisfying

$$\sup_{t \in \mathbb{R}} [(1 + |t|)|K(t)| + (1 + |t|)^2 |K'(t)|] \leq 1, \quad \sup_{N \geq 0} \left| \int_{-N}^N K(t) dt \right| \leq 1. \quad (1.1)$$

The main theorem we prove in this paper is the following:

Theorem 1.2. *Assume $\mathbb{G}$ is a discrete nilpotent group of step 2, K is a Calderon–Zygmund kernel, and $A : \mathbb{Z} \rightarrow \mathbb{G}$ is a polynomial sequence. For any (compactly supported) function $f : \mathbb{G} \rightarrow \mathbb{C}$ let*

$$(Hf)(g) = \sum_{n \in \mathbb{Z}} K(n) f(A^{-1}(n) \cdot g), \quad g \in \mathbb{G}.$$

Then

$$\|Hf\|_{L^2(\mathbb{G})} \lesssim \|f\|_{L^2(\mathbb{G})}.$$

We describe now some of the main ideas in the proof of the theorem. We use first a transference principle to reduce matters to proving the theorem in a certain "universal"

case. More precisely, it will suffice to consider singular Radon transforms on the groups $\mathbb{G}_0 = \mathbb{G}_0(d)$ defined in section 2, and for explicit polynomial sequences $A_0 : \mathbb{Z} \rightarrow \mathbb{G}_0$, see Theorem 2.3. This reduction simplifies the overall picture and allows us to work in good systems of coordinates, which are well adapted to the natural homogeneities induced by the polynomial A_0 . However, the main problem, namely the lack of a good Fourier transform on the group $\mathbb{G}_0$ compatible with the structure of our convolution operators, remains even in this special setting.

A natural approach is to attempt to prove the theorem using the Cotlar–Stein lemma. More precisely, we may assume that

$$K = \sum_{j=1}^{\infty} K_j, \quad \int_{\mathbb{R}} K_j(t) dt = 0, \quad 2^j |K_j(t)| + 2^{2j} |K'_j(t)| \leq \mathbf{1}_{[-2^{j+3}, 2^{j+3}]}(t),$$

and consider the dyadic averages

$$H_j(f)(g) = \sum_{n \in \mathbb{Z}} K_j(n) f(A_0(n)^{-1} \cdot g), \quad g \in \mathbb{G}_0.$$

To apply the Cotlar–Stein lemma, we would have to prove an inequality of the form

$$\|H_k H_j^*\|_{L^2 \rightarrow L^2} + \|H_k^* H_j\|_{L^2 \rightarrow L^2} \lesssim 2^{-\delta'(j-k)} \quad (1.2)$$

for some $\delta' > 0$, and for any $k \leq j \in \{1, 2, \dots\}$. This is equivalent to proving that

$$\|H_k (H_j^* H_j)^r\|_{L^2 \rightarrow L^2} + \|H_k^* (H_j H_j^*)^r\|_{L^2 \rightarrow L^2} \lesssim 2^{-\delta(j-k)} \quad (1.3)$$

for some $\delta > 0$, $r \in \{1, 2, \dots\}$, and for any $k \leq j \in \{1, 2, \dots\}$.

The advantage of proving (1.3) instead of (1.2) is that the operators $(H_j^* H_j)^r$ and $(H_j H_j^*)^r$ are more regular than the operators H_j , provided that $r \geq r(d)$ is sufficiently large. The kernels of these operators can be described precisely, see Proposition 3.2. Up to negligible errors, these operators are essentially sums of more standard oscillatory singular operators on the group $\mathbb{G}_0$, given by kernels of the form³

$$h \rightarrow \sum_{a/q} S^{(r)}(a/q) e^{2\pi i h \cdot a/q} K_j^{(r)}(h). \quad (1.4)$$

The sum is taken over suitable "irreducible fractions" a/q , the coefficients $S^{(r)}(a/q)$ have sufficiently fast decay as $q \rightarrow \infty$ (provided that r is sufficiently large), and $K_j^{(r)}$ is (almost) a standard singular integral kernel adapted to the canonical non-isotropic balls on the underlying Lie group $\mathbb{G}_0^\#$. This representation can be used to prove that

$$\|H_k (H_j^* H_j)^r\|_{L^2 \rightarrow L^2} + \|H_k^* (H_j H_j^*)^r\|_{L^2 \rightarrow L^2} \lesssim 2^{-\delta(j-k)} + 2^{-\delta k}, \quad \delta > 0, k \leq j \in \{1, 2, \dots\},$$

see Lemma 4.2, and, as a consequence,

$$\|H_k H_j^*\|_{L^2 \rightarrow L^2} + \|H_k^* H_j\|_{L^2 \rightarrow L^2} \lesssim 2^{-\delta'(j-k)} + 2^{-\delta' k}, \quad \delta' > 0, k \leq j \in \{1, 2, \dots\}. \quad (1.5)$$

³The proof of Proposition 3.2, which includes this description, relies on the complicated oscillatory sum estimates in Proposition 5.1. Having an elementary, essentially self-contained proof of these estimates is the main reason for working on step 2 groups, instead of the general case.

Unfortunately this last bound is weaker than the desired bound (1.2), and the additional factor $2^{-\delta'k}$ cannot be removed. As a consequence, the Cotlar–Stein lemma can be used to prove the weaker bound

$$\left\| \sum_{j \in [J, 2J]} H_j \right\|_{L^2 \rightarrow L^2} \lesssim 1, \quad \text{uniformly in } J,$$

but is not suitable to control the entire sum over j .

To estimate the entire sum we need an additional almost-orthogonality lemma, which we prove in section 6. This lemma appears to be new and might be of independent interest. In its simplest form, it says that if $S_1, \dots, S_K$ are bounded linear operators on a Hilbert space H satisfying, for any $m = 1, \dots, K$,

$$\begin{aligned} \sup_{m \in \{1, \dots, K\}} \|S_m\| &\leq 1, \\ \sup_{i_m, \dots, i_K \in \{0, 1\}} \|S_{m, i_m}^* [(S_{m+1, i_{m+1}} S_{m+1, i_{m+1}}^*)^{p_0} + \dots + (S_{K, i_K} S_{K, i_K}^*)^{p_0}]\| &\leq A 2^{-\delta_0 m}, \\ \sup_{i_m, \dots, i_K \in \{0, 1\}} \|S_{m, i_m} [(S_{m+1, i_{m+1}}^* S_{m+1, i_{m+1}})^{p_0} + \dots + (S_{K, i_K}^* S_{K, i_K})^{p_0}]\| &\leq A 2^{-\delta_0 m}, \end{aligned} \quad (1.6)$$

for some $\delta_0 > 0$, some dyadic number p_0 , and some constant A , then

$$\|S_1 + \dots + S_K\| \leq C(\delta_0, A, p_0).$$

The notation in (1.6) is $S_{m,0} = S_m$ and $S_{m,1} = 0$.

We apply this almost-orthogonality lemma with

$$S_m = \sum_{j \in [(1-\kappa)J_m, J_m]} H_j,$$

where $\kappa > 0$ is a sufficiently small constant and $J_1, J_2, \dots$ is a rapidly increasing sequence, $J_{m+1} \geq 2J_m$. The inequality in the first line of (1.6) is a consequence of the Cotlar–Stein lemma and (1.5). We prove the remaining inequalities in (1.6) in two steps: in Lemma 4.4 we prove the uniform bounds

$$\|(S_m^* S_m)^r + \dots + (S_n^* S_n)^r\|_{L^2 \rightarrow L^2} + \|(S_m S_m^*)^r + \dots + (S_n S_n^*)^r\|_{L^2 \rightarrow L^2} \lesssim 1,$$

for any $m \leq n \in \{1, 2, \dots\}$. For this we establish formulas similar to (1.4) for the kernels of the operators $(S_k^* S_k)^r$ and $(S_k S_k^*)^r$. Then we show in Lemma 4.5 that left composition with the operator S_{m-1} (or S_{m-1}^* respectively) contributes an additional factor of $2^{-\delta m}$, $\delta > 0$, thereby proving the desired bounds in (1.6).

The rest of the paper is organized as follows. In section 2 we use a transference argument to reduce the general case in Theorem 1.2 (corresponding to a general group $\mathbb{G}$ and a general sequence $A : \mathbb{Z} \rightarrow \mathbb{G}$) to a "universal" case (corresponding to a particular group $\mathbb{G}_0$ and a particular sequence $A_0 : \mathbb{Z} \rightarrow \mathbb{G}_0$).

In section 3 we define the operators H_j (the dyadic pieces of our singular Radon transforms), and describe the operators $H_{j_1}^* H_{k_1} \dots H_{j_r}^* H_{k_r}$ and $H_{j_1} H_{k_1}^* \dots H_{j_r} H_{k_r}^*$, for integers $j_1, k_1, \dots, j_r, k_r \in [J(1-\kappa), J]$. For $r \geq r(d)$ large enough we prove in Proposition 3.2

that the kernels of these operators are sums of more standard oscillatory singular integral kernels, similar to (1.4) (arising from "major arcs"), and negligible errors (arising from "minor arcs"). The bounds on these error terms rely on Proposition 5.1 and are delicate in our situation, due to the complicated structure of the polynomials that arise as a result of multiplication in the group $\mathbb{G}_0$.

Section 4 contains the proof of Theorem 2.3, i.e. the proof of the bounds in (1.6), along the line described above.

In section 5 we prove estimates for trigonometric sums and integrals, using a variant of the Weyl method developed by Davenport [10] and Birch [3]. These estimates are used at several places, for example to control the contributions of the "minor arcs" and to estimate the coefficients $S^{(r)}(a/q)$ in (1.4). For the sake of completeness we provide all the details needed in the proof.

Finally, in section 6 we state and prove a suitable version of the additional orthogonality lemma described in (1.6).

Acknowledgement: We would like to express our deep gratitude to Elias Stein, for his guidance and friendship throughout the years.

2. A TRANSFERENCE ARGUMENT

Let $\mathbb{G}^\#$ be a step 2 (connected and simply connected) nilpotent Lie group and let $\mathcal{G}$ denote its Lie algebra. Choose a basis $\mathcal{X} = \{X_1, \dots, X_{d_1}, Y_1, \dots, Y_{d_2}\}$ of the Lie algebra $\mathcal{G}$ such that $\mathbb{R}\text{-span}\{Y_1, \dots, Y_{d_2}\} = [\mathcal{G}, \mathcal{G}]$, the commutator subalgebra of $\mathcal{G}$. Note that this is a special case of a so-called strong Malcev basis passing through the lower central series $\mathcal{G} \geq [\mathcal{G}, \mathcal{G}] \geq \{0\}$ (see [7], Sec. 1.2). Associated to such a basis one defines coordinates on $\mathbb{G}^\#$ via the diffeomorphism $\phi : \mathbb{R}^d \rightarrow \mathbb{G}^\#$ defined by

$$\phi(x_1, \dots, x_{d_1}, y_1, \dots, y_{d_2}) = \exp(x_1 X_1) \dots \exp(x_{d_1} X_{d_1}) \exp(y_1 Y_1) \dots \exp(y_{d_2} Y_{d_2}).$$

Such coordinates associated to a Malcev basis are called exponential coordinates of the second kind. In these coordinates we have that

$$\mathbb{G}^\# = \{(x, y) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} : (x, y) \cdot (x', y') = (x + x', y + y' + R(x, x'))\}, \quad (2.1)$$

where $R : \mathbb{R}^{d_1} \times \mathbb{R}^{d_1} \rightarrow \mathbb{R}^{d_2}$ is a bilinear form. This follows easily from facts that $\exp(X) \cdot \exp(Y) = \exp(X + Y + \frac{1}{2}[X, Y])$ which implies that

$$\exp(x_i X_i) \exp(x'_j X_j) = \exp(x'_j X_j) \exp(x_i X_i) \exp(x_i x'_j [X_i, X_j]),$$

and $[X_i, X_j] = \sum_{l=1}^{d_2} c_{ij}^l Y_l$.

If $\mathbb{G} \leq \mathbb{G}^\#$ is a discrete co-compact subgroup, then one can choose such a basis $\mathcal{X} = \{X_1, \dots, Y_{d_2}\}$ so that

$$\mathbb{G} = \phi(\mathbb{Z}^d) = \exp(\mathbb{Z} X_1) \dots \exp(\mathbb{Z} X_{d_1}) \exp(\mathbb{Z} Y_1) \dots \exp(\mathbb{Z} Y_{d_2}),$$

see [7] Thm. 5.1.6 and Prop. 5.3.2. Thus the discrete subgroup $\mathbb{G}$ is identified with the integer lattice $\mathbb{Z}^d = \mathbb{Z}^{d_1} \times \mathbb{Z}^{d_2}$.

If $A : \mathbb{Z} \rightarrow \mathbb{G}$ is a polynomial sequence ($A(0) = 1$), then it is not hard to see that in these coordinates it takes the form

$$A(n) = (x_1(n), \dots, x_{d_1}(n), y_1(n), \dots, y_{d_2}(n)), \quad A(0) = 0,$$

where x_{l_1}, y_{l_2} are integral polynomials. Indeed, writing

$$DA(n) = (Dx_1(n), \dots, Dx_{d_1}(n), Dy_1(n), \dots, Dy_{d_2}(n)),$$

we have from (2.1) that $Dx_i(n) = x_i(n+1) - x_i(n)$ and $Dy_l(n) = y_l(n+1) - y_l(n) - R'_l(n)$ where $R'_l(n)$ is a polynomial expression of $x_1(n), \dots, x_{d_1}(n), x_1(n+1), \dots, x_{d_1}(n+1)$. Since $D^k x_i(n)$ is identically zero it follows that $x_i(n)$ is a polynomial of degree at most k , and then the vanishing of $D^k y_l(n)$ implies that $y_l(n)$ must be polynomial as well. Alternatively this fact can be easily derived from the characterization of polynomial sequences by Leibman [14] mentioned in the introduction. We will denote by d_3 the maximum of the degrees of the polynomials $x_i(n)$ and $y_l(n)$.

It will be useful to consider the polynomial map $A : \mathbb{Z} \rightarrow \mathbb{G}$ as a map $A : \mathbb{Z} \rightarrow \mathbb{G}^\#$, and the associated singular Radon transform acting on $L^2(\mathbb{G}^\#)$, defined by

$$(\tilde{H}f)(g) = \sum_{n \in \mathbb{Z}} K(n) f(A^{-1}(n) \cdot g), \quad g \in \mathbb{G}^\#.$$

In this settings our main result takes the form

Theorem 2.1. *Assume $\mathbb{G}^\#$ is a (connected and simply connected) nilpotent Lie group of step 2, K is a Calderon–Zygmund kernel, and $A : \mathbb{Z} \rightarrow \mathbb{G}^\#$ is a polynomial sequence. For any (continuous compactly supported) function $f : \mathbb{G}^\# \rightarrow \mathbb{C}$, we have*

$$\|\tilde{H}f\|_{L^2(\mathbb{G}^\#)} \lesssim \|f\|_{L^2(\mathbb{G}^\#)}.$$

We will show below that

$$\|\tilde{H}\|_{L^2(\mathbb{G}^\#) \rightarrow L^2(\mathbb{G}^\#)} = \|H\|_{L^2(\mathbb{G}) \rightarrow L^2(\mathbb{G})},$$

hence Theorem 2.1 and Theorem 1.2 are equivalent. To see this let $\mathbf{S}_d = \phi([0, 1]^d)$ where $\phi : \mathbb{R}^d \rightarrow \mathbb{G}^\#$ is the coordinate map defined above. From the multiplication structure given in (2.1) it is easy to see that $\mathbf{S}_d$ is a fundamental domain for $\mathbb{G}$, that is every element $g \in \mathbb{G}^\#$ can be written uniquely as $g = \gamma \cdot s$ with $\gamma \in \mathbb{G}$ and $s \in \mathbf{S}_d$. Moreover the map $\tilde{\phi} = \pi \circ \phi$ (π being the natural projection from $\mathbb{G}^\#$ to $\mathbb{G} \backslash \mathbb{G}^\#$) maps the Lebesgue measure on $[0, 1]^d$ to the normalized $\mathbb{G}^\#$ -invariant measure on $\mathbb{G} \backslash \mathbb{G}^\#$. For a given function $f : \mathbb{G} \rightarrow \mathbb{C}$ let $f^\# : \mathbb{G}^\# \rightarrow \mathbb{C}$ be such that $f^\#(\gamma \cdot s) = f(\gamma)$ for all $\gamma \in \mathbb{G}$ and $s \in \mathbf{S}_d$. Then

$$\|f^\#\|_{L^2(\mathbb{G}^\#)}^2 = \int_{\mathbb{G}^\#} |f^\#(g)|^2 dg = \sum_{\gamma \in \mathbb{G}} \int_{\mathbf{S}_d} |f^\#(\gamma \cdot s)|^2 ds = \sum_{\gamma \in \mathbb{G}} |f(\gamma)|^2 = \|f\|_{L^2(\mathbb{G})}^2.$$

Also

$$\tilde{H}f^\#(\gamma \cdot s) = \sum_{n \in \mathbb{Z}} K(n) f^\#(A(n)^{-1} \cdot \gamma \cdot s) = \sum_{n \in \mathbb{Z}} K(n) f(A(n)^{-1} \cdot \gamma) = Hf(\gamma),$$

thus $\tilde{H}f^\# = (Hf)^\#$ and hence the operators $\tilde{H}$ and H have the same norm.

The advantage of Theorem 2.1 is that it is easier to reduce it to a certain universal case. For integers $d \geq 1$ we define

$$Y_d = \{(l_1, l_2) \in \mathbb{Z} \times \mathbb{Z} : 0 \leq l_2 < l_1 \leq d\}$$

and the “universal” step-two nilpotent Lie groups $\mathbb{G}_0^\# = \mathbb{G}_0^\#(d)$

$$\mathbb{G}_0^\# = \{(x_{l_1 l_2})_{(l_1, l_2) \in Y_d} : x_{l_1 l_2} \in \mathbb{R}\},$$

with the group multiplication law

$$[x \cdot y]_{l_1 l_2} = \begin{cases} x_{l_1 0} + y_{l_1 0} & \text{if } l_1 \in \{1, \dots, d\} \text{ and } l_2 = 0, \\ x_{l_1 l_2} + y_{l_1 l_2} + x_{l_1 0} y_{l_2 0} & \text{if } l_1 \in \{1, \dots, d\} \text{ and } l_2 \in \{1, \dots, l_1 - 1\}. \end{cases}$$

Let $\mathbb{G}_0 = \mathbb{G}_0(d)$ denote the discrete subgroup $\mathbb{G}_0 = \mathbb{G}_0^\# \cap \mathbb{Z}^{|Y_d|}$. Let $A_0 : \mathbb{R} \rightarrow \mathbb{G}_0^\#$ denote the polynomial map

$$[A_0(x)]_{l_1 l_2} = \begin{cases} x^{l_1} & \text{if } l_2 = 0, \\ 0 & \text{if } l_2 \neq 0, \end{cases} \quad (2.2)$$

and notice that $A_0(\mathbb{Z}) \subseteq \mathbb{G}_0$.

Lemma 2.2. *Assuming $\mathbb{G}_0^\#$ and A are defined as before, there is d sufficiently large and a group morphism $T : \mathbb{G}_0 \rightarrow \mathbb{G}_0^\#$ such that*

$$A(n) = T(A_0(n)) \quad \text{for any } n \in \mathbb{Z}. \quad (2.3)$$

Proof of Lemma 2.2. Set

$$d = 2d_3$$

and let $g_1, \dots, g_d$ denote the generators of the group $\mathbb{G}_0$,

$$[g_m]_{l_1 l_2} = \begin{cases} 1 & \text{if } l_1 = m \text{ and } l_2 = 0, \\ 0 & \text{otherwise.} \end{cases}$$

We notice that any group morphism $T : \mathbb{G}_0 \rightarrow \mathbb{G}_0^\#$ is uniquely determined by the values $T(g_1), \dots, T(g_d)$. Indeed, any element

$$x = (x_{l_1 l_2})_{(l_1, l_2) \in Y_d} \in \mathbb{G}_0, \quad x_{l_1 l_2} \in \mathbb{Z},$$

can be written in the form

$$x = g_1^{x_{10}} \cdot \dots \cdot g_d^{x_{d0}} \cdot \prod_{1 \leq l_2 < l_1 \leq d} (g_{l_1} g_{l_2} g_{l_1}^{-1} g_{l_2}^{-1})^{x_{l_1 l_2}}.$$

Therefore, if $T(g_l) = h_l \in \mathbb{G}_0^\#$ then T is uniquely defined by

$$T(x) = h_1^{x_{10}} \cdot \dots \cdot h_d^{x_{d0}} \cdot \prod_{1 \leq l_2 < l_1 \leq d} (h_{l_1} h_{l_2} h_{l_1}^{-1} h_{l_2}^{-1})^{x_{l_1 l_2}}, \quad \text{if } x = (x_{l_1 l_2})_{(l_1, l_2) \in Y_d} \in \mathbb{G}_0.$$

It is easy to verify that this defines indeed a group morphism, using the fact that the elements $h_{l_1} h_{l_2} h_{l_1}^{-1} h_{l_2}^{-1}$ are in the center of the group $\mathbb{G}^\#$.

Assume that

$$A(n) = \left(\sum_{i=1}^{d_3} \alpha_i n^i, \sum_{i=1}^{d_3} \beta_i n^i \right), \quad \alpha_1, \dots, \alpha_{d_3} \in \mathbb{R}^{d_1}, \beta_1, \dots, \beta_{d_3} \in \mathbb{R}^{d_2}. \quad (2.4)$$

We define

$$T(g_l) = \begin{cases} (\alpha_l, \gamma_l) & \text{if } l \in \{1, \dots, d_3\}, \\ (0, \gamma_l) & \text{if } l \in \{d_3 + 1, \dots, d\}, \end{cases}$$

for some vectors $\gamma_1, \dots, \gamma_d \in \mathbb{R}^{d_2}$ to be fixed, and extend T as a group morphism from $\mathbb{G}_0 \rightarrow \mathbb{G}^\#$. Since

$$A_0(n) = g_1^n \cdot \dots \cdot g_d^{n^d},$$

it follows that

$$T(A_0(n)) = \left(\sum_{i=1}^{d_3} \alpha_i n^i, \sum_{i=1}^d \gamma_i n^i + \sum_{i=1}^{2d_3} \rho_i n^i \right),$$

for some coefficients $\rho_1, \dots, \rho_{2d_3}$ that depend only on $(\alpha_i)_{i \in \{1, \dots, d_3\}}$ and the bilinear form R . The desired identity $T(A_0(n)) = A(n)$ can be arranged by choosing the vectors $\gamma_1, \dots, \gamma_d$ appropriately. $\square$

Assume now that we could prove the following particular case of Theorem 1.2:

Theorem 2.3. *For any $d \geq 1$, $R \geq 1$, and $F : \mathbb{G}_0 \rightarrow \mathbb{C}$ let*

$$(H_0^R F)(g_0) = \sum_{|n| \leq R} K(n) F(A_0(n)^{-1} \cdot g_0),$$

where $A_0 : \mathbb{Z} \rightarrow \mathbb{G}_0$ is as in (2.2) and K is as in (1.1). Then

$$\|H_0^R F\|_{L^2(\mathbb{G}_0)} \lesssim_d \|F\|_{L^2(\mathbb{G}_0)} \quad \text{uniformly in } R.$$

It is not hard to see that Theorem 2.3 would imply Theorem 1.2. This follows from the standard transference principle, see [21, Proposition 5.1]. Indeed, given a polynomial map $A : \mathbb{R} \rightarrow \mathbb{G}^\#$ with $A(0) = 0$, we fix a group morphism $T : \mathbb{G}_0 \rightarrow \mathbb{G}^\#$ such as $A(n) = T(A_0(n))$, $n \in \mathbb{Z}$. Then we define the isometric representation π of $\mathbb{G}_0$ on $L^2(\mathbb{G}^\#)$,

$$\pi(g_0)(f)(g) = f(T(g_0^{-1}) \cdot g), \quad g_0 \in \mathbb{G}_0, f \in L^2(\mathbb{G}^\#), g \in \mathbb{G}^\#. \quad (2.5)$$

For $R \geq 1$ we define

$$\begin{aligned} K^R : \mathbb{Z} &\rightarrow \mathbb{C}, & K^R(n) &= K(n) \mathbf{1}_{[-R, R] \cap \mathbb{Z}}(n), \\ (H^R f)(g) &= \sum_{n \in \mathbb{Z}} K^R(n) f(A(n)^{-1} \cdot g), & f &\in C_0(\mathbb{G}^\#). \end{aligned}$$

Then, for any bounded open set $U \subseteq \mathbb{G}_0$, $R \geq 1$, and $f \in C_0(\mathbb{G}^\#)$

$$\|H^R f\|_{L^2(\mathbb{G}^\#)}^2 = \frac{1}{|U|} \int_U \int_{\mathbb{G}^\#} |\pi(g_0^{-1})(H^R f)(g)|^2 dg dg_0.$$

The definitions show that

$$\pi(g_0^{-1})(H^R f)(g) = (H^R f)(T(g_0) \cdot g) = \sum_{n \in \mathbb{Z}} K^R(n) f(T(A_0(n)^{-1} \cdot g_0) \cdot g) = H_0^R(F_g)(g_0),$$

where, by definition,

$$F_g(h_0) = f(T(h_0) \cdot g).$$

Notice that, for $g_0 \in U$,

$$H_0^R(F)(g_0) = H_0^R(F \cdot \mathbf{1}_{U'_R})(g_0), \quad U'_R = \{(\underline{u}, \underline{v}) \cdot h : h \in U, |\underline{u}|^2 + |\underline{v}| < C_d R^{2d}\}.$$

Therefore, using these identities and Theorem 2.3,

$$\begin{aligned} \|H^R f\|_{L^2(\mathbb{G}^\#)}^2 &= \frac{1}{|U|} \int_U \int_{\mathbb{G}^\#} |H_0^R(F_g \cdot \mathbf{1}_{U'_R})(g_0)|^2 dg dg_0 \\ &\lesssim_d \frac{1}{|U|} \int_{\mathbb{G}_0} \int_{\mathbb{G}^\#} |(F_g \cdot \mathbf{1}_{U'_R})(h_0)|^2 dg dh_0 \\ &\lesssim_d \frac{1}{|U|} \int_{\mathbb{G}_0} \int_{\mathbb{G}^\#} |f(T(h_0) \cdot g)|^2 \cdot \mathbf{1}_{U'_R}(h_0) dg dh_0 \\ &\lesssim_d \frac{|U'_R|}{|U|} \|f\|_{L^2(\mathbb{G}^\#)}^2. \end{aligned}$$

For R fixed we can fix U large enough such that $|U'_R|/|U| \leq 2$. Thus $\|H^R f\|_{L^2(\mathbb{G})} \lesssim_d \|f\|_{L^2(\mathbb{G})}$ uniformly in R , as desired.

The rest of the paper is concerned with the proof of Theorem 2.3. We will assume from now on that d is fixed, and all the implied constants are allowed to depend on d .

3. THE MAIN KERNELS: IDENTITIES AND ESTIMATES

We fix $\eta_0 : \mathbb{R} \rightarrow [0, 1]$ a smooth even function supported in the interval $[-2, 2]$ and equal to 1 in the interval $[-1, 1]$. We define

$$\eta_j(t) = \eta_0(2^{-j}t) - \eta_0(2^{-j+1}t), \quad t \in \mathbb{R}, j = 1, 2, \dots, \quad 1 = \sum_{j=0}^{\infty} \eta_j.$$

For $\lambda \geq 1$ let $\tilde{\eta}_{\leq \lambda} : \mathbb{R}^{|Y_d|} \rightarrow [0, 1]$,

$$\tilde{\eta}_{\leq \lambda}(x) = \prod_{(l_1, l_2) \in Y_d} \eta_0(x_{l_1 l_2} / 2^{\lambda(l_1 + l_2)}).$$

For $x = (x_{l_1 l_2})_{(l_1, l_2) \in Y_d} \in \mathbb{R}^{|Y_d|}$ and $\Lambda \in (0, \infty)$ let

$$\Lambda \circ x = (\Lambda^{l_1 + l_2} x_{l_1 l_2})_{(l_1, l_2) \in Y_d} \in \mathbb{R}^{|Y_d|}, \quad |x| = \sum_{(l_1, l_2) \in Y_d} |x_{l_1 l_2}|.$$

Let

$$\mathcal{D}_\Lambda^\# = \{x \in \mathbb{R}^{|Y_d|} : |(1/\Lambda) \circ x| < 1\}, \quad \mathcal{D}_\Lambda = \mathcal{D}_\Lambda^\# \cap \mathbb{Z}^{|Y_d|}.$$

For $j = 1, 2, \dots$ let

$$\begin{aligned} K_j(t) &= K(t)\eta_j(t) + c_j 2^{-j} \eta_j(t) - c_{j+1} 2^{-j-1} \eta_{j+1}(t), \\ \text{where } c_j &= 2 \left(\int_{\mathbb{R}} K(t) \left[\sum_{k=0}^{j-1} \eta_k(t) \right] dt \right) \left(\int_{\mathbb{R}} \eta_0(t) dt \right)^{-1}. \end{aligned} \quad (3.1)$$

Using this definition and the assumption (1.1), it follows that, for $j = 1, 2, \dots$

$$\begin{aligned} 2^j |K_j(t)| + 2^{2j} |K'_j(t)| &\lesssim \mathbf{1}_{[-2^{j+3}, 2^{j+3}]}(t), \quad \int_{\mathbb{R}} K_j(t) dt = 0, \quad \sup_{j=1,2,\dots} |c_j| \lesssim 1, \\ \sum_{j'=1}^j K_{j'}(t) &= K(t)\eta_0(2^{-j}t) - K(t)\eta_0(t) + c_1 2^{-1} \eta_1(t) - c_{j+1} 2^{-j-1} \eta_{j+1}(t). \end{aligned} \quad (3.2)$$

For $f \in L^2(\mathbb{G}_0)$ let

$$(H_j f)(g) = \sum_{n \in \mathbb{Z}} K_j(n) f(A_0(n)^{-1} \cdot g).$$

In this section we use the notation and the estimates in section 5, in particular Proposition 5.1 and Lemma 5.4. Any vector in $\mathbb{Q}^m$ has a unique representation in the form a/q , with $q \in \{1, 2, \dots\}$, $a \in \mathbb{Z}^m$, and $(a, q) = 1$. For $R \in [1, \infty]$ let $\mathcal{S}_R$ denote the set of irreducible fractions in $(\mathbb{Q} \cap (0, 1])^{|Y_d|}$ with denominators $\leq R$, i.e.

$$\mathcal{S}_R = \{a/q = (a_{l_1 l_2}/q)_{(l_1, l_2) \in Y_d} : 1 \leq q \leq R, a_{l_1 l_2} \in \mathbb{Z}_q, (a, q) = 1\}.$$

We fix once and for all three parameters ϵ, r, κ , $0 < \kappa \ll 1/r \ll \epsilon \ll 1$, $r \in 2^{\mathbb{Z}_+}$, depending only on d and satisfying

$$\epsilon = \overline{C}^{-1} (10d)^{-10}, \quad -2\overline{C} + r\epsilon/(2\overline{C}) \geq (10d)^{10}, \quad \kappa r^2 = 1, \quad (3.3)$$

where $\overline{C}$ is the constant in Proposition 5.1 and Lemma 5.4.

For $a/q \in \mathcal{S}_\infty$ let

$$S(a/q) = q^{-2r} \sum_{v, w \in \mathbb{Z}_q^r} e^{-2\pi i D(v, w) \cdot a/q}, \quad \tilde{S}(a/q) = q^{-2r} \sum_{v, w \in \mathbb{Z}_q^r} e^{-2\pi i \tilde{D}(v, w) \cdot a/q}. \quad (3.4)$$

where $D, \tilde{D}$ are defined in (5.2) and (5.3).

Lemma 3.1. *For any $a/q \in \mathcal{S}_\infty$*

$$|S(a/q)| + |\tilde{S}(a/q)| \lesssim q^{-(10d)^{10}}. \quad (3.5)$$

Proof of Lemma 3.1. For $(l_1, l_2) \in Y_d$ we write

$$\frac{a_{l_1 l_2}}{q} = \frac{a'_{l_1 l_2}}{q_{l_1 l_2}}, \quad (a'_{l_1 l_2}, q_{l_1 l_2}) = 1, \quad 1 \leq q_{l_1 l_2} \leq q.$$

The bound follows from Proposition 5.1 with $P = q$ if

$$\text{there is } (l_1, l_2) \in Y_d \text{ such that } q^\epsilon \leq q_{l_1 l_2} \leq q^{l_1 + l_2 - \epsilon}.$$

Otherwise, since

$$\sup_{(l_1, l_2) \in Y_d} q_{l_1 l_2} \leq q \leq \prod_{(l_1, l_2) \in Y_d} q_{l_1 l_2},$$

we necessarily have

$$q_{l_1 l_2} \leq q^\epsilon \text{ if } l_1 + l_2 \geq 2 \text{ and } q_{10} \geq q^{1-\epsilon}. \quad (3.6)$$

In this case we may assume $q \geq 2$, and let Q denote the smallest common multiple of $q_{l_1 l_2}$, $l_1 + l_2 \geq 2$, $q/Q \in \{2, 3, \dots\}$. Then we estimate, using the formula (5.2),

$$\begin{aligned} |S(a/q)| &\leq q^{-r} \sup_{v \in Z_q^r} \left| \sum_{w \in Z_q^r} e^{-2\pi i D(v, w) \cdot a/q} \right| \leq q^{-r} \sup_{v \in Z_q^r} \sum_{y \in Z_Q^r} \left| \sum_{x \in Z_{q/Q}^r} e^{-2\pi i D(v, Qx+y) \cdot a/q} \right| \\ &\leq q^{-r} \sup_{v \in Z_q^r} \sum_{y \in Z_Q^r} \left| \sum_{x \in Z_{q/Q}^r} e^{-2\pi i D(v, Qx+y)_{10} \cdot a_{10}/q_{10}} \right| = 0, \end{aligned}$$

which suffices. The bound on $|\tilde{S}(a/q)|$ is similar. $\square$

The main goal in this section is to describe the operators

$$H_{j_1} H_{j_2}^* \dots H_{j_{2r-1}} H_{j_{2r}}^* \quad \text{and} \quad H_{j_1}^* H_{j_2} \dots H_{j_{2r-1}}^* H_{j_{2r}},$$

for suitable values of $j_1, \dots, j_{2r}$. More precisely, we prove the following:

Proposition 3.2. *Assume $C(d)$ is a sufficiently large constant, $J \in [C(d), \infty)$, and $j_1, k_1, \dots, j_r, k_r \in [J(1 - \kappa), J] \cap \mathbb{Z}$. Then*

$$\begin{aligned} (H_{j_1}^* H_{k_1} \dots H_{j_r}^* H_{k_r} F)(g) &= \sum_{h \in \mathbb{G}_0} [K_{j_1, k_1, \dots, j_r, k_r}(h) + E_{j_1, k_1, \dots, j_r, k_r}(h)] F(h^{-1} \cdot g), \\ (H_{j_1} H_{k_1}^* \dots H_{j_r} H_{k_r}^* F)(g) &= \sum_{h \in \mathbb{G}_0} [\tilde{K}_{j_1, k_1, \dots, j_r, k_r}(h) + \tilde{E}_{j_1, k_1, \dots, j_r, k_r}(h)] F(h^{-1} \cdot g), \end{aligned}$$

for any $F \in L^2(\mathbb{G}_0)$ and $g \in \mathbb{G}_0$, where

$$\|E_{j_1, k_1, \dots, j_r, k_r}\|_{L^1(\mathbb{G}_0)} + \|\tilde{E}_{j_1, k_1, \dots, j_r, k_r}\|_{L^1(\mathbb{G}_0)} \lesssim 2^{-J/4}. \quad (3.7)$$

Moreover

$$\begin{aligned} K_{j_1, k_1, \dots, j_r, k_r}(h) &= \tilde{\eta}_{\leq J + \epsilon J}(h) \sum_{a/q \in \mathcal{S}_{2^{3d} 2^{\epsilon J}}} e^{2\pi i h \cdot a/q} S(a/q) \\ &\int_{\mathbb{R}^{|Y_d|}} \int_{\mathbb{R}^r \times \mathbb{R}^r} \prod_{(l_1, l_2) \in Y_d} \eta_0(2^{J(l_1 + l_2 - 2\epsilon)} \beta_{l_1 l_2}) G_{j_1, k_1, \dots, j_r, k_r}(x, y) e^{2\pi i (h - D(x, y)) \cdot \beta} dx dy d\beta, \end{aligned} \quad (3.8)$$

$$\begin{aligned} \tilde{K}_{j_1, k_1, \dots, j_r, k_r}(h) &= \tilde{\eta}_{\leq J+\epsilon J}(h) \sum_{a/q \in \mathcal{S}_{2^{3d^2\epsilon}J}} e^{2\pi i h \cdot a/q} \tilde{S}(a/q) \\ &\int_{\mathbb{R}^{|Y_d|}} \int_{\mathbb{R}^r \times \mathbb{R}^r} \prod_{(l_1, l_2) \in Y_d} \eta_0(2^{J(l_1+l_2-2\epsilon)} \beta_{l_1 l_2}) G_{j_1, k_1, \dots, j_r, k_r}(x, y) e^{2\pi i (h - \tilde{D}(x, y)) \cdot \beta} dx dy d\beta. \end{aligned} \quad (3.9)$$

The functions $G_{j_1, k_1, \dots, j_r, k_r}$ are defined by

$$G_{j_1, k_1, \dots, j_r, k_r}(x, y) = K_{j_1}(x_1) K_{k_1}(y_1) \dots K_{j_r}(x_r) K_{k_r}(y_r), \quad x, y \in \mathbb{R}^r.$$

The functions $D, \tilde{D} : \mathbb{R}^r \times \mathbb{R}^r \rightarrow \mathbb{R}^{|Y_d|}$ are defined in (5.2) and (5.3).

Proof of Proposition 3.2. We only prove the claims for the operators $H_{j_1}^* H_{k_1} \dots H_{j_r}^* H_{k_r}$ and the kernels $K_{j_1, k_1, \dots, j_r, k_r}, E_{j_1, k_1, \dots, j_r, k_r}$; the claims for the operators $H_{j_1} H_{k_1}^* \dots H_{j_r} H_{k_r}^*$ and the kernels $\tilde{K}_{j_1, k_1, \dots, j_r, k_r}, \tilde{E}_{j_1, k_1, \dots, j_r, k_r}$ follow by essentially identical arguments. Recall that ϵ, r, κ are fixed, depending only on d , so all the implicit constants are allowed to depend on ϵ, r, κ .

By definition,

$$\begin{aligned} (H_{j_1}^* H_{k_1} \dots H_{j_r}^* H_{k_r} F)(g) &= \sum_{n_1, m_1, \dots, n_r, m_r \in \mathbb{Z}} K_{j_1}(n_1) K_{k_1}(m_1) \dots K_{j_r}(n_r) K_{k_r}(m_r) \\ &F(A_0(m_r)^{-1} \cdot A_0(n_r) \cdot \dots \cdot A_0(m_1)^{-1} \cdot A_0(n_1) \cdot g). \end{aligned} \quad (3.10)$$

Recalling the definition (5.1) and letting

$$\begin{aligned} L_{j_1, k_1, \dots, j_r, k_r}(h) &= \tilde{\eta}_{\leq J+\epsilon J}(h) \int_{[0,1]^{|Y_d|}} \sum_{n, m \in \mathbb{Z}^r} G_{j_1, k_1, \dots, j_r, k_r}(n, m) \\ &e^{2\pi i \sum_{(l_1, l_2) \in Y_d} (h_{l_1 l_2} - D(n, m)_{l_1 l_2}) \theta_{l_1 l_2}} d\theta, \end{aligned} \quad (3.11)$$

this becomes

$$(H_{j_1}^* H_{k_1} \dots H_{j_r}^* H_{k_r} F)(g) = \sum_{h \in \mathbb{G}_0} L_{j_1, k_1, \dots, j_r, k_r}(h) F(h^{-1} \cdot g).$$

It remains to prove that we can decompose $L_{j_1, k_1, \dots, j_r, k_r} = K_{j_1, k_1, \dots, j_r, k_r} + E_{j_1, k_1, \dots, j_r, k_r}$ satisfying the claims in the proposition.

We decompose the integral over θ in (3.11) into the contribution of major and minor arcs. Let

$$\begin{aligned} L_{j_1, k_1, \dots, j_r, k_r}^1(h) &= \tilde{\eta}_{\leq J+\epsilon J}(h) \sum_{a/q \in \mathcal{S}_{2^{3d^2\epsilon}J}} \int_{\mathbb{R}^{|Y_d|}} \sum_{n, m \in \mathbb{Z}^r} G_{j_1, k_1, \dots, j_r, k_r}(n, m) \\ &e^{2\pi i \sum_{(l_1, l_2) \in Y_d} (h_{l_1 l_2} - D(n, m)_{l_1 l_2}) (a_{l_1 l_2}/q + \beta_{l_1 l_2})} \prod_{(l_1, l_2) \in Y_d} \eta_0(2^{J(l_1+l_2-2\epsilon)} \beta_{l_1 l_2}) d\beta. \end{aligned} \quad (3.12)$$

In view of the choice of ϵ, r, κ and the restriction $j_1, k_1, \dots, j_r, k_r \in [(1-\kappa)J, J]$, it follows from Proposition 5.1 and (3.2) that

$$|L_{j_1, k_1, \dots, j_r, k_r}(h) - L_{j_1, k_1, \dots, j_r, k_r}^1(h)| \lesssim 2^{-10d^2 J} \tilde{\eta}_{\leq J+\epsilon J}(h), \quad h \in \mathbb{G}_0,$$

which is consistent with the error estimate (3.7).

We consider now the sum over m, n in (3.12), and rewrite, for q fixed,

$$\begin{aligned} & \sum_{n, m \in \mathbb{Z}^r} G_{j_1, k_1, \dots, j_r, k_r}(n, m) e^{-2\pi i D(n, m) \cdot (a/q + \beta)} \\ &= \sum_{n, m \in \mathbb{Z}^r} \sum_{v, w \in \mathbb{Z}_q^r} G_{j_1, k_1, \dots, j_r, k_r}(qn + v, qm + w) e^{-2\pi i D(qn + v, qm + w) \cdot \beta} e^{-2\pi i D(v, w) \cdot a/q} \\ &= E'(a/q, \beta) + \sum_{n, m \in \mathbb{Z}^r} \sum_{v, w \in \mathbb{Z}_q^r} G_{j_1, k_1, \dots, j_r, k_r}(qn, qm) e^{-2\pi i D(qn, qm) \cdot \beta} e^{-2\pi i D(v, w) \cdot a/q}. \end{aligned}$$

For $q \leq 2^{3d^2\epsilon J}$ and $\beta = (\beta_{l_1 l_2})_{(l_1, l_2) \in Y_d}$, $|\beta_{l_1 l_2}| \leq 22^{-J(l_1 + l_2 - 2\epsilon)}$, we estimate, using (3.2), (5.2), and the assumption $j_1, k_1, \dots, j_r, k_r \in [(1 - \kappa)J, J]$,

$$|E'(a/q, \beta)| \lesssim 2^{-3J/4}.$$

Therefore, if we define

$$\begin{aligned} L_{j_1, k_1, \dots, j_r, k_r}^2(h) &= \tilde{\eta}_{\leq J + \epsilon J}(h) \sum_{a/q \in \mathcal{S}_{2^{3d^2\epsilon J}}} \int_{\mathbb{R}^{|Y_d|}} e^{2\pi i h \cdot (a/q + \beta)} \prod_{(l_1, l_2) \in Y_d} \eta_0(2^{J(l_1 + l_2 - 2\epsilon)} \beta_{l_1 l_2}) \\ &\quad \sum_{n, m \in \mathbb{Z}^r} \sum_{v, w \in \mathbb{Z}_q^r} G_{j_1, k_1, \dots, j_r, k_r}(qn, qm) e^{-2\pi i D(qn, qm) \cdot \beta} e^{-2\pi i D(v, w) \cdot a/q} d\beta, \end{aligned} \quad (3.13)$$

it follows that

$$|L_{j_1, k_1, \dots, j_r, k_r}^1(h) - L_{j_1, k_1, \dots, j_r, k_r}^2(h)| \lesssim 2^{-J/2} \prod_{(l_1, l_2) \in Y_d} 2^{-J(l_1 + l_2)} \tilde{\eta}_{\leq J + \epsilon J}(h).$$

This is consistent with the error estimate in (3.7).

Finally, it remains to decompose the kernel $L_{j_1, k_1, \dots, j_r, k_r}^2$. For this we rewrite first

$$\begin{aligned} L_{j_1, k_1, \dots, j_r, k_r}^2(h) &= \tilde{\eta}_{\leq J + \epsilon J}(h) \sum_{a/q \in \mathcal{S}_{2^{3d^2\epsilon J}}} e^{2\pi i h \cdot a/q} S(a/q) \int_{\mathbb{R}^{|Y_d|}} e^{2\pi i h \cdot \beta} \prod_{(l_1, l_2) \in Y_d} \eta_0(2^{J(l_1 + l_2 - 2\epsilon)} \beta_{l_1 l_2}) \\ &\quad q^{2r} \sum_{n, m \in \mathbb{Z}^r} G_{j_1, k_1, \dots, j_r, k_r}(qn, qm) e^{-2\pi i D(qn, qm) \cdot \beta} d\beta, \end{aligned}$$

where $S(a/q)$ is defined in (3.4). Using the formula (5.2), we estimate for any $q \leq 2^{3d^2\epsilon J}$ and $\beta = (\beta_{l_1 l_2})_{(l_1, l_2) \in Y_d}$, $|\beta_{l_1 l_2}| \leq 22^{-J(l_1 + l_2 - 2\epsilon)}$,

$$\begin{aligned} & \left| \sum_{n, m \in \mathbb{Z}^r} G_{j_1, k_1, \dots, j_r, k_r}(qn, qm) e^{-2\pi i D(qn, qm) \cdot \beta} - \int_{\mathbb{R}^r \times \mathbb{R}^r} G_{j_1, k_1, \dots, j_r, k_r}(qx, qy) e^{-2\pi i D(qx, qy) \cdot \beta} dx dy \right| \\ & \lesssim q^{-2r} 2^{-3J/4}. \end{aligned}$$

Thus, with $K_{j_1, k_1, \dots, j_r, k_r}$ defined as in (3.8), we have the pointwise bound

$$|L_{j_1, k_1, \dots, j_r, k_r}^2(h) - K_{j_1, k_1, \dots, j_r, k_r}(h)| \lesssim 2^{-J/2} \prod_{(l_1, l_2) \in Y_d} 2^{-J(l_1 + l_2)} \tilde{\eta}_{\leq J + \epsilon J}(h),$$

which is consistent with the error estimate in (3.7). This completes the proof of the proposition. $\square$

Assume $J, j_1, k_1, \dots, j_r, k_r$ are as in Proposition 3.2 and define

$$\begin{aligned} P_{j_1, k_1, \dots, j_r, k_r}(v) &= \int_{\mathbb{R}^r \times \mathbb{R}^r} G_{j_1, k_1, \dots, j_r, k_r}(x, y) e^{-2\pi i D(x, y) \cdot (2^{-J} \circ v)} dx dy, \\ \tilde{P}_{j_1, k_1, \dots, j_r, k_r}(v) &= \int_{\mathbb{R}^r \times \mathbb{R}^r} G_{j_1, k_1, \dots, j_r, k_r}(x, y) e^{-2\pi i \tilde{D}(x, y) \cdot (2^{-J} \circ v)} dx dy. \end{aligned} \quad (3.14)$$

Notice that the formulas (3.8) and (3.9) become, after changes of variables

$$\begin{aligned} K_{j_1, k_1, \dots, j_r, k_r}(h) &= \tilde{\eta}_{\leq J + \epsilon J}(h) \prod_{(l_1, l_2) \in Y_d} 2^{-J(l_1 + l_2)} \sum_{a/q \in \mathcal{S}_{2^{3d} 2^{\epsilon J}}} e^{2\pi i h \cdot a/q} S(a/q) \\ &\quad \int_{\mathbb{R}^{|Y_d|}} \prod_{(l_1, l_2) \in Y_d} \eta_0(2^{-2\epsilon J} v_{l_1 l_2}) P_{j_1, k_1, \dots, j_r, k_r}(v) e^{2\pi i (2^{-J} \circ h) \cdot v} dv, \end{aligned} \quad (3.15)$$

and

$$\begin{aligned} \tilde{K}_{j_1, k_1, \dots, j_r, k_r}(h) &= \tilde{\eta}_{\leq J + \epsilon J}(h) \prod_{(l_1, l_2) \in Y_d} 2^{-J(l_1 + l_2)} \sum_{a/q \in \mathcal{S}_{2^{3d} 2^{\epsilon J}}} e^{2\pi i h \cdot a/q} \tilde{S}(a/q) \\ &\quad \int_{\mathbb{R}^{|Y_d|}} \prod_{(l_1, l_2) \in Y_d} \eta_0(2^{-2\epsilon J} v_{l_1 l_2}) \tilde{P}_{j_1, k_1, \dots, j_r, k_r}(v) e^{2\pi i (2^{-J} \circ h) \cdot v} dv. \end{aligned} \quad (3.16)$$

In view of the cancellation condition in the first line of (3.2),

$$P_{j_1, k_1, \dots, j_r, k_r}(0) = \tilde{P}_{j_1, k_1, \dots, j_r, k_r}(0) = 0. \quad (3.17)$$

We make the changes of variables $x = 2^J \mu$, $y = 2^J \nu$ to rewrite

$$\begin{aligned} P_{j_1, k_1, \dots, j_r, k_r}(v) &= \int_{\mathbb{R}^r \times \mathbb{R}^r} 2^{2rJ} G_{j_1, k_1, \dots, j_r, k_r}(2^J \mu, 2^J \nu) e^{-2\pi i D(\mu, \nu) \cdot v} d\mu d\nu, \\ \tilde{P}_{j_1, k_1, \dots, j_r, k_r}(v) &= \int_{\mathbb{R}^r \times \mathbb{R}^r} 2^{2rJ} G_{j_1, k_1, \dots, j_r, k_r}(2^J \mu, 2^J \nu) e^{-2\pi i \tilde{D}(\mu, \nu) \cdot v} d\mu d\nu. \end{aligned}$$

Using Lemma 5.4, for $m = 0, 1, \dots$

$$|\nabla_v^m P_{j_1, k_1, \dots, j_r, k_r}(v)| + |\nabla_v^m \tilde{P}_{j_1, k_1, \dots, j_r, k_r}(v)| \lesssim_m 2^{8r(J - \min(j_1, \dots, j_r, k_1, \dots, k_r))} (1 + |v|)^{-(10d)^{10}}. \quad (3.18)$$

4. PROOF OF THEOREM 2.3

In this section we complete the proof of Theorem 2.3. The main ingredients are Lemma 6.2 and the estimates and the identities proved in section 3. We use the notation introduced in section 3. In view of the identity in the second line of (3.2), it suffices to prove that for any integer $J \geq 1$

$$\left\| \sum_{j=1}^J H_j \right\|_{L^2(\mathbb{G}_0) \rightarrow L^2(\mathbb{G}_0)} \lesssim 1.$$

By further dividing into finitely many sums, it suffices to prove the following:

Proposition 4.1. *Assume $J_1, \dots, J_K \in [1, \infty)$ satisfy the separation condition*

$$J_{m+1} \geq 2J_m, \quad m = 1, \dots, K-1. \quad (4.1)$$

For $m = 1, \dots, K$ let

$$S_m = \sum_{j \in [J_m(1-\kappa), J_m] \cap \mathbb{Z}} H_j.$$

Then

$$\|S_1 + \dots + S_K\|_{L^2(\mathbb{G}_0) \rightarrow L^2(\mathbb{G}_0)} \lesssim 1.$$

The rest of the section is concerned with the proof of Proposition 4.1. We would like to apply Lemma 6.2, in the simplified form given in Remark 6.3. We will verify the conditions (6.21) in several steps.

Lemma 4.2. *We have*

$$\sup_{J \geq 1} \sup_{A \subseteq [J/2, J] \cap \mathbb{Z}} \left\| \sum_{j \in A} H_j \right\|_{L^2(\mathbb{G}_0) \rightarrow L^2(\mathbb{G}_0)} \lesssim 1.$$

Proof of Lemma 4.2. In view of the Cotlar–Stein lemma, it suffices to prove that, for some $\delta' > 0$,

$$\|H_k H_j^*\|_{L^2 \rightarrow L^2} + \|H_k^* H_j\|_{L^2 \rightarrow L^2} \lesssim 2^{-\delta'(j-k)} \quad \text{for any } k \leq j \in [J/2, J] \cap \mathbb{Z}.$$

Since $\|H_j\|_{L^2 \rightarrow L^2} \lesssim 1$ for any j , it follows that

$$\begin{aligned} \|H_k H_j^*\|_{L^2 \rightarrow L^2} &\lesssim \|H_k H_j^* H_j\|_{L^2 \rightarrow L^2}^{1/2} \lesssim \|H_k (H_j^* H_j)^2\|_{L^2 \rightarrow L^2}^{1/4} \lesssim \dots \lesssim \|H_k (H_j^* H_j)^r\|_{L^2 \rightarrow L^2}^{1/(2r)}, \\ \|H_k^* H_j\|_{L^2 \rightarrow L^2} &\lesssim \|H_k^* H_j H_j^*\|_{L^2 \rightarrow L^2}^{1/2} \lesssim \|H_k^* (H_j H_j^*)^2\|_{L^2 \rightarrow L^2}^{1/4} \lesssim \dots \lesssim \|H_k^* (H_j H_j^*)^r\|_{L^2 \rightarrow L^2}^{1/(2r)}. \end{aligned}$$

Therefore it suffices to prove that there is $\delta = \delta(d) > 0$ such that

$$\|H_k (H_j^* H_j)^r\|_{L^2 \rightarrow L^2} + \|H_k^* (H_j H_j^*)^r\|_{L^2 \rightarrow L^2} \lesssim 2^{-\delta(j-k)} \quad (4.2)$$

for any $k, j \in [C(d), \infty) \cap \mathbb{Z}$, $k \in [j/2, j]$.

We will prove only the bound on the first term in the left-hand side of (4.2); the bound on the second term is very similar. We use Proposition 3.2 with $J = j_1 = k_1 = \dots = j_r = k_r = j$. With the notation in Proposition 3.2

$$[H_k(H_j^* H_j)^r](F)(g) = \sum_{h \in \mathbb{G}_0} F(h^{-1} \cdot g) \sum_{n \in \mathbb{Z}} K_k(n) (K_{j,j,\dots,j,j} + E_{j,j,\dots,j,j})(A_0(n)^{-1} \cdot h),$$

for any $F \in L^2(\mathbb{G}_0)$ and $g \in \mathbb{G}_0$. In view of (3.7), it suffices to prove that

$$\left\| \sum_{n \in \mathbb{Z}} K_k(n) K_{j,j,\dots,j,j}(A_0(n)^{-1} \cdot h) \right\|_{L_h^1(\mathbb{G}_0)} \lesssim 2^{-\delta(j-k)}.$$

We use now the formula (3.15). For $x \in \mathbb{R}^{|Y_d|}$ let

$$\begin{aligned} M_j(x) &= \tilde{\eta}_{\leq j+\epsilon_j}(x) \prod_{(l_1, l_2) \in Y_d} 2^{-j(l_1+l_2)} \\ &\quad \int_{\mathbb{R}^{|Y_d|}} \prod_{(l_1, l_2) \in Y_d} \eta_0(2^{-2\epsilon_j} \beta_{l_1 l_2}) P_{j,j,\dots,j,j}(\beta) e^{2\pi i(2^{-j} \circ x) \cdot \beta} d\beta. \end{aligned} \quad (4.3)$$

Recalling the rapid decay of the coefficients $S(a/q)$ (see Lemma 3.1), it suffices to prove that for any $a/q \in \mathcal{S}_{2^{3d^2}\epsilon_j}$

$$\left\| \sum_{n \in \mathbb{Z}} K_k(n) e^{2\pi i(A_0(n)^{-1} \cdot h) \cdot a/q} M_j(A_0(n)^{-1} \cdot h) \right\|_{L_h^1(\mathbb{G}_0)} \lesssim 2^{-\delta(j-k)} q^{(4d)^4}. \quad (4.4)$$

Using (3.18) and integration by parts

$$|M_j(x)| + \sum_{(l_1, l_2) \in Y_d} 2^{j(l_1+l_2)} |\partial_{x_{l_1, l_2}} M_j(x)| \lesssim (1 + |2^{-j} \circ x|)^{-(4d)^4} \prod_{(l_1, l_2) \in Y_d} 2^{-j(l_1+l_2)}. \quad (4.5)$$

Therefore, if $|n| \lesssim 2^k$ and $h \in \mathbb{G}_0$

$$|M_j(A_0(n)^{-1} \cdot h) - M_j(h)| \lesssim 2^{k-j} (1 + |2^{-j} \circ h|)^{-(4d)^4} \prod_{(l_1, l_2) \in Y_d} 2^{-j(l_1+l_2)}.$$

Thus

$$\left\| \sum_{n \in \mathbb{Z}} K_k(n) e^{2\pi i(A_0(n)^{-1} \cdot h) \cdot a/q} [M_j(A_0(n)^{-1} \cdot h) - M_j(h)] \right\|_{L_h^1(\mathbb{G}_0)} \lesssim 2^{k-j}. \quad (4.6)$$

On the other hand, using (3.2) and the assumption $k \geq j/2$, for any $h \in \mathbb{G}_0$ and $a/q \in \mathcal{S}_{2^{3d^2}\epsilon_j}$

$$\begin{aligned} \left| \sum_{n \in \mathbb{Z}} K_k(n) e^{2\pi i(A_0(n)^{-1} \cdot h) \cdot a/q} \right| &\leq \sum_{m \in \mathbb{Z}_q} \left| \sum_{n \in \mathbb{Z}} K_k(qn + m) e^{2\pi i(A_0(qn+m)^{-1} \cdot h) \cdot a/q} \right| \\ &\leq \sum_{m \in \mathbb{Z}_q} \left| \sum_{n \in \mathbb{Z}} K_k(qn + m) \right| \lesssim 2^{-j/4}. \end{aligned} \quad (4.7)$$

Thus, using also (4.5),

$$\left\| \sum_{n \in \mathbb{Z}} K_k(n) e^{2\pi i(A_0(n)^{-1} \cdot h) \cdot a/q} M_j(h) \right\|_{L_h^1(\mathbb{G}_0)} \lesssim 2^{-j/4}, \quad (4.8)$$

and the bound (4.4) follows from (4.6) and (4.8). This completes the proof. $\square$

Remark 4.3. We observe that it is important to assume that $j/k \lesssim 1$ in the proof of the bound (4.2). Otherwise one could only prove a weaker bound, of the form

$$\|H_k H_j^*\|_{L^2 \rightarrow L^2} + \|H_k^* H_j\|_{L^2 \rightarrow L^2} \lesssim 2^{-\delta'(j-k)} + 2^{-\delta'k} \quad \text{for any } k \leq j \in \{1, 2, \dots\}.$$

Such a bound does not suffice to apply the Cotlar–Stein lemma to prove the theorem directly. It is precisely to compensate for this failure that we need the additional orthogonality proposition in section 6.

We consider now long sums of operators $(S_m^* S_m)^r$ and $(S_m S_m^*)^r$.

Lemma 4.4. Assume $J_1, \dots, J_K \in [C(d), \infty)$ satisfy the separation condition

$$J_{m+1} \geq 2J_m, \quad m = 1, \dots, K-1. \quad (4.9)$$

For $m = 1, \dots, K$ let

$$S_m = \sum_{j \in [J_m(1-\kappa), J_m] \cap \mathbb{Z}} H_j.$$

Then

$$\|(S_1^* S_1)^r + \dots + (S_K^* S_K)^r\|_{L^2 \rightarrow L^2} + \|(S_1 S_1^*)^r + \dots + (S_K S_K^*)^r\|_{L^2 \rightarrow L^2} \lesssim 1. \quad (4.10)$$

Proof of Lemma 4.4. We prove only the bound on the first term in the left-hand side of (4.10). In view of Proposition 3.2, it suffices to prove that

$$\left\| \sum_{m=1}^K F * \left[\sum_{j_1, k_1, \dots, j_r, k_r \in [J_m(1-\kappa), J_m] \cap \mathbb{Z}} K_{j_1, k_1, \dots, j_r, k_r} \right] \right\|_{L^2(\mathbb{G}_0)} \lesssim \|F\|_{L^2(\mathbb{G}_0)}$$

for any $F \in L^2(\mathbb{G}_0)$. For $x \in \mathbb{R}^{|Y_d|}$ and $m = 1, \dots, K$ we define

$$\begin{aligned} N_m(x) &= \tilde{\eta}_{\leq J_m + \epsilon J_m}(x) \prod_{(l_1, l_2) \in Y_d} 2^{-J_m(l_1 + l_2)} \\ &\int_{\mathbb{R}^{|Y_d|}} \prod_{(l_1, l_2) \in Y_d} \eta_0(2^{-2\epsilon J_m} \beta_{l_1 l_2}) \sum_{j_1, k_1, \dots, j_r, k_r \in [J_m(1-\kappa), J_m] \cap \mathbb{Z}} P_{j_1, k_1, \dots, j_r, k_r}(\beta) e^{2\pi i(2^{-J_m} \circ x) \cdot \beta} d\beta. \end{aligned} \quad (4.11)$$

We use the formula (3.15) and the rapid decay of the coefficients $S(a/q)$ in Lemma 3.1. After rearranging the sum, it suffices to prove that for any $a/q \in \mathcal{S}_\infty$

$$\left\| \sum_{2^{J_m} \geq q^{(8d)^8}} \sum_{h \in \mathbb{G}_0} F(h^{-1} \cdot g) e^{2\pi i h \cdot a/q} N_m(h) \right\|_{L_g^2(\mathbb{G}_0)} \lesssim q^{(4d)^4} \|F\|_{L^2(\mathbb{G}_0)} \quad (4.12)$$

for any $F \in L^2(\mathbb{G}_0)$.

Using (3.18) and integration by parts

$$|N_m(x)| + \sum_{(l_1, l_2) \in Y_d} 2^{J_m(l_1+l_2)} |\partial_{x_{l_1 l_2}} N_m(x)| \lesssim_C 2^{4\epsilon J_m} (1 + |2^{-J_m} \circ x|)^{-C} \prod_{(l_1, l_2) \in Y_d} 2^{-J_m(l_1+l_2)}. \quad (4.13)$$

Using both (3.17) and (3.18), it follows that

$$\left| \int_{\mathbb{R}^{|Y_d|}} N_m(x) dx \right| \lesssim 2^{-J_m}. \quad (4.14)$$

We would like to prove (4.12) using the Cotlar–Stein lemma. For this we need to modify the kernels N_m to achieve a cancellation. More precisely, given a fixed fraction $a/q \in \mathcal{S}_\infty$ and $2^{J_m} \geq q^{(8d)^8}$ we would like to define kernels $N'_m : \mathbb{G}_0 \rightarrow \mathbb{C}$ with the properties

$$\begin{aligned} \sum_{h \in \mathbb{G}_0} N'_m(h) e^{2\pi i h \cdot a/q} e^{-2\pi i (v \cdot h) \cdot a/q} &= \sum_{h \in \mathbb{G}_0} N'_m(h) e^{2\pi i h \cdot a/q} e^{-2\pi i (h \cdot v) \cdot a/q} = 0, \quad v \in \mathbb{G}_0, \\ \|N_m - N'_m\|_{L^1(\mathbb{G}_0)} &\lesssim 2^{-J_m/4}, \\ N'_m(h) &= 0 \quad \text{if} \quad h \notin \mathcal{D}_{2^{J_m}(1+2\epsilon)}. \end{aligned} \quad (4.15)$$

To prove this, we introduce a decomposition of elements in the group $\mathbb{G}_0$, adapted to the denominator q . Let

$$\begin{aligned} \mathbb{H}_q &= \{h \in \mathbb{G}_0 : h = (qm_{l_1 l_2})_{(l_1, l_2) \in Y_d}, m_{l_1 l_2} \in \mathbb{Z}\}, \\ R_q &= \{b \in \mathbb{G}_0 : b_{l_1 l_2} \in [0, q-1] \cap \mathbb{Z}\}, \end{aligned} \quad (4.16)$$

and notice that

$$\text{the map } (h, b) \rightarrow h \cdot b \text{ defines a bijection from } \mathbb{H}_q \times R_q \text{ to } \mathbb{G}_0. \quad (4.17)$$

The cancellation condition in the first line of (4.15) holds provided that

$$\sum_{h \in \mathbb{H}_q} N'_m(h \cdot b) = 0 \text{ for any } b \in R_q. \quad (4.18)$$

Therefore we set, for any $h \in \mathbb{G}_0$

$$\begin{aligned} N'_m(h) &= N_m(h) - \tilde{\eta}_{\leq J_m}(h) \prod_{(l_1, l_2) \in Y_d} 2^{-J_m(l_1+l_2)} \sum_{b \in R_q} \gamma_b \mathbf{1}_{\mathbb{H}_q \cdot b}(h), \\ \gamma_b &= \left[\sum_{g \in \mathbb{H}_q} N_m(g \cdot b) \right] \left[\sum_{g \in \mathbb{H}_q} \tilde{\eta}_{\leq J_m}(g \cdot b) \prod_{(l_1, l_2) \in Y_d} 2^{-J_m(l_1+l_2)} \right]^{-1} \quad \text{for any } b \in R_q. \end{aligned} \quad (4.19)$$

The support assertion in (4.15) follows from the definition. The cancellation assertion in (4.15) follows from (4.18). Finally, to prove that $\|N_m - N'_m\|_{L^1(\mathbb{G}_0)} \lesssim 2^{-J_m/4}$ it suffices to prove that

$$|\gamma_b| \lesssim 2^{-J_m/4} \quad \text{for any } b \in R_q.$$

Recalling that $2^{J_m} \geq q^{(8d)^8}$ and using the definition of γ_b , it remains to prove that

$$\left| \sum_{g \in \mathbb{H}_q} N_m(g \cdot b) \right| \lesssim 2^{-J_m/3} \quad \text{for any } b \in R_q. \quad (4.20)$$

Using (4.11),

$$\sup_{b \in R_q} |N_m(g \cdot b) - N_m(g)| \lesssim 2^{-J_m/2} (1 + |2^{-J_m} \circ x|)^{-(4d)^4} \prod_{(l_1, l_2) \in Y_d} 2^{-J_m(l_1 + l_2)}.$$

Moreover, using (4.13), (4.14),

$$\left| \sum_{g \in \mathbb{H}_q} N_m(g) \right| \lesssim 2^{-J_m/2}.$$

The bound (4.20) follows from the last two bounds, which completes the proof of (4.15).

We turn now to the proof of (4.12). Let

$$T_m F(g) = \sum_{h \in \mathbb{G}_0} F(h^{-1} \cdot g) e^{2\pi i h \cdot a/q} N'_m(h).$$

For (4.12) it suffices to prove that

$$\left\| \sum_{2^{J_m} \geq q^{(8d)^8}} T_m \right\|_{L^2 \rightarrow L^2} \lesssim q^{(4d)^4}.$$

In view of the Cotlar–Stein lemma, it suffices to prove that, for some $\delta = \delta(d) > 0$

$$\|T_m T_{m'}^*\|_{L^2 \rightarrow L^2} + \|T_m^* T_{m'}\|_{L^2 \rightarrow L^2} \lesssim 2^{-\delta(m' - m)} q^{2(4d)^4}, \quad 2^{J_{m'}} \geq 2^{J_m} \geq q^{(8d)^8}. \quad (4.21)$$

We first prove (4.21) when $m' \geq m + 1$. Using the cancellation conditions in (4.15),

$$\begin{aligned} T_m T_{m'}^* F(g) &= \sum_{v \in \mathbb{G}_0} F(v \cdot g) \left[\sum_{h \in \mathbb{G}_0} N'_m(h) e^{2\pi i h \cdot a/q} \overline{N'_{m'}(v \cdot h)} e^{-2\pi i (v \cdot h) \cdot a/q} \right] \\ &= \sum_{v \in \mathbb{G}_0} F(v \cdot g) \left[\sum_{h \in \mathbb{G}_0} N'_m(h) e^{2\pi i h \cdot a/q} [\overline{N'_{m'}(v \cdot h)} - \overline{N'_{m'}(v)}] e^{-2\pi i (v \cdot h) \cdot a/q} \right], \end{aligned}$$

and

$$\begin{aligned} T_m^* T_{m'} F(g) &= \sum_{v \in \mathbb{G}_0} F(v^{-1} \cdot g) \left[\sum_{h \in \mathbb{G}_0} \overline{N'_m(h)} e^{-2\pi i h \cdot a/q} N'_{m'}(h \cdot v) e^{2\pi i (h \cdot v) \cdot a/q} \right] \\ &= \sum_{v \in \mathbb{G}_0} F(v^{-1} \cdot g) \left[\sum_{h \in \mathbb{G}_0} \overline{N'_m(h)} e^{-2\pi i h \cdot a/q} [N'_{m'}(h \cdot v) - N'_{m'}(v)] e^{2\pi i (h \cdot v) \cdot a/q} \right]. \end{aligned}$$

Thus

$$\begin{aligned}
& \|T_m T_{m'}^*\|_{L^2 \rightarrow L^2} + \|T_m^* T_{m'}\|_{L^2 \rightarrow L^2} \\
& \lesssim \left\| \sum_{h \in \mathbb{G}_0} |N'_m(h)| [|N'_{m'}(h \cdot v) - N'_{m'}(v)| + |N'_{m'}(v \cdot h) - N'_{m'}(v)|] \right\|_{L_v^1} \\
& \lesssim \|N'_m\|_{L^1} \|N_{m'} - N'_{m'}\|_{L^1} \\
& + \|N'_m\|_{L^1} \sup_{h \in \mathcal{D}_{2^{J_m}(1+2\epsilon)}^\#} (\|N_{m'}(h \cdot v) - N_{m'}(v)\|_{L_v^1} + \|N_{m'}(v \cdot h) - N_{m'}(v)\|_{L_v^1}).
\end{aligned} \tag{4.22}$$

Using the bounds (4.13) and (4.15) and the separation assumption $J_{m'} \geq 2J_m$,

$$\begin{aligned}
& \|N_{m'} - N'_{m'}\|_{L^1} \lesssim 2^{-J_{m'}/4}, \quad \|N_m\|_{L^1} + \|N'_m\|_{L^1} \lesssim 2^{J_m/20}, \\
& \sup_{h \in \mathcal{D}_{2^{J_m}(1+2\epsilon)}^\#} |N_{m'}(h \cdot v) - N_{m'}(v)| \lesssim \mathbf{1}_{\mathcal{D}_{2^{J_{m'}(1+8\epsilon)}}(v)} 2^{J_m - J_{m'}} 2^{8\epsilon J_{m'}} \prod_{(l_1, l_2) \in Y_d} 2^{-J_{m'}(l_1 + l_2)}, \\
& \sup_{h \in \mathcal{D}_{2^{J_m}(1+2\epsilon)}^\#} |N_{m'}(v \cdot h) - N_{m'}(v)| \lesssim \mathbf{1}_{\mathcal{D}_{2^{J_{m'}(1+8\epsilon)}}(v)} 2^{J_m - J_{m'}} 2^{8\epsilon J_{m'}} \prod_{(l_1, l_2) \in Y_d} 2^{-J_{m'}(l_1 + l_2)}.
\end{aligned} \tag{4.23}$$

Using (4.22) it follows that

$$\|T_m T_{m'}^*\|_{L^2 \rightarrow L^2} + \|T_m^* T_{m'}\|_{L^2 \rightarrow L^2} \lesssim 2^{-J_{m'}/10}, \tag{4.24}$$

which clearly suffices to prove (4.21) in this case.

Finally, we prove (4.21) when $m' = m$, which is equivalent to

$$\left\| \sum_{h \in \mathbb{G}_0} F(h) e^{2\pi i(g \cdot h) \cdot a/q} N_m(g \cdot h) \right\|_{L_g^2} \lesssim q^{(4d)^4} \|F\|_{L^2}. \tag{4.25}$$

Using the decomposition (4.16)-(4.17), it suffices to prove that

$$\sum_{b \in R_q} \left\| \sum_{h \in \mathbb{H}_q} F(h \cdot b) e^{2\pi i(g \cdot h \cdot b) \cdot a/q} N_m(g \cdot h \cdot b) \right\|_{L_g^2} \lesssim q^{(4d)^4} \|F\|_{L^2}.$$

Since $e^{2\pi i(g \cdot h \cdot b) \cdot a/q}$ does not depend on $h \in \mathbb{H}_q$, this is equivalent to proving that

$$\sum_{b \in R_q} \left\| \sum_{h \in \mathbb{H}_q} F_b(h) N_m(g^{-1} \cdot h \cdot b) \right\|_{L_g^2} \lesssim q^{(4d)^4} \left[\sum_{b \in R_q} \|F_b\|_{L^2(\mathbb{H}_q)}^2 \right]^{1/2}. \tag{4.26}$$

We notice that R_q has $q^{|Y_d|}$ elements. Therefore, it suffices to prove that for any $F, G \in L^2(\mathbb{G}_0)$

$$\left| \sum_{h, g \in \mathbb{G}_0} F(h) N_m(h^{-1} \cdot g) G(g) \right| \lesssim \|F\|_{L^2} \|G\|_{L^2}. \tag{4.27}$$

We derive (4.27) as a consequence of L^2 boundedness of a singular Radon transform on the nilpotent Lie group $\mathbb{G}_0^\#$. Let

$$\mathcal{C} = [0, 1)^{|Y_d|} \subseteq \mathbb{G}_0^\#$$

and notice that

the map $(g, \mu) \rightarrow g \cdot \mu$ defines a measure-preserving bijection from $\mathbb{G}_0 \times \mathcal{C}$ to $\mathbb{G}_0^\#$. (4.28)

For any function $f \in L^2(\mathbb{G}_0)$ let

$$f^\#(g \cdot \mu) = f(g) \text{ for any } (g, \mu) \in \mathbb{G}_0 \times \mathcal{C}, \quad f^\# \in L^2(\mathbb{G}_0^\#), \quad \|f^\#\|_{L^2(\mathbb{G}_0^\#)} = \|f\|_{L^2(\mathbb{G}_0)}.$$

Then we write, for any $F, G \in L^2(\mathbb{G}_0)$

$$\begin{aligned} \sum_{h, g \in \mathbb{G}_0} F(h) N_m(h^{-1} \cdot g) G(g) &= \int_{\mathcal{C} \times \mathcal{C}} \sum_{h, g \in \mathbb{G}_0} F^\#(h \cdot \mu) N_m(h^{-1} \cdot g) G^\#(g \cdot \nu) d\mu d\nu \\ &= \int_{\mathbb{G}_0^\# \times \mathbb{G}_0^\#} F^\#(y) N_m(y^{-1} \cdot x) G^\#(x) dx dy \\ &\quad + \int_{\mathcal{C} \times \mathcal{C}} \sum_{h, g \in \mathbb{G}_0} F^\#(h \cdot \mu) [N_m(h^{-1} \cdot g) - N_m(\mu^{-1} \cdot h^{-1} \cdot g \cdot \nu)] G^\#(g \cdot \nu) d\mu d\nu. \end{aligned} \quad (4.29)$$

Using (4.13), we have

$$\left\| \sup_{\mu, \nu \in \mathcal{C}} |N_m(x) - N_m(\mu^{-1} \cdot x \cdot \nu)| \right\|_{L_x^1(\mathbb{G}_0)} \lesssim 2^{-J_m/2}.$$

Thus

$$\left| \int_{\mathcal{C} \times \mathcal{C}} \sum_{h, g \in \mathbb{G}_0} F^\#(h \cdot \mu) [N_m(h^{-1} \cdot g) - N_m(\mu^{-1} \cdot h^{-1} \cdot g \cdot \nu)] G^\#(g \cdot \nu) d\mu d\nu \right| \lesssim \|F\|_{L^2} \|G\|_{L^2}.$$

Using (4.29), for (4.27) it suffices to prove that

$$\left| \int_{\mathbb{G}_0^\# \times \mathbb{G}_0^\#} F(y) N_m(y^{-1} \cdot x) G(x) dx dy \right| \lesssim \|F\|_{L^2(\mathbb{G}_0^\#)} \|G\|_{L^2(\mathbb{G}_0^\#)} \quad (4.30)$$

for any $F, G \in L^2(\mathbb{G}_0^\#)$.

We examine the formula (4.11) and define

$$N_m''(x) = \prod_{(l_1, l_2) \in Y_d} 2^{-J_m(l_1 + l_2)} \int_{\mathbb{R}^{|Y_d|}} \sum_{j_1, k_1, \dots, j_r, k_r \in [J_m(1-\kappa), J_m] \cap \mathbb{Z}} P_{j_1, k_1, \dots, j_r, k_r}(\beta) e^{2\pi i(2^{-J_m} \circ x) \cdot \beta} d\beta.$$

Using (3.18)

$$\|N_m - N_m''\|_{L^1(\mathbb{G}_0^\#)} \lesssim 1.$$

Therefore, for (4.30) it suffices to prove that for any $F \in C_0^\infty(\mathbb{G}_0^\#)$

$$\left\| \int_{\mathbb{G}_0^\#} F(y^{-1} \cdot x) N_m''(y) dy \right\|_{L_x^2(\mathbb{G}_0^\#)} \lesssim \|F\|_{L^2(\mathbb{G}_0^\#)}. \quad (4.31)$$

Recalling the definition (3.14) we notice that, for any $F \in C_0^\infty(\mathbb{G}_0^\#)$,

$$\int_{\mathbb{G}_0^\#} F(y^{-1} \cdot x) N_m''(y) dy = \sum_{j_1, k_1, \dots, j_r, k_r \in [J_m(1-\kappa), J_m] \cap \mathbb{Z}} [(H_{j_1}^\#)^* H_{k_1}^\# \dots (H_{j_r}^\#)^* H_{k_r}^\#](F)(x)$$

where, by definition,

$$H_j^\# f(x) = \int_{\mathbb{R}} K_j(t) f(A_0(t)^{-1} \cdot x) dt. \quad (4.32)$$

Therefore, for (4.31) it suffices to prove that

$$\left\| \sum_{j \in [J_m(1-\kappa), J_m] \cap \mathbb{Z}} H_j^\# \right\|_{L^2(\mathbb{G}_0^\#) \rightarrow L^2(\mathbb{G}_0^\#)} \lesssim 1. \quad (4.33)$$

The bound (4.33) is essentially known, as a consequence of Theorem 3.4 in [21]. We can also reprove it easily, using the bounds we have proved so far. As in the proof of Lemma 4.2, using the Cotlar–Stein lemma it suffices to prove that

$$\|H_k^\#((H_j^\#)^* H_j^\#)^r\|_{L^2(\mathbb{G}_0^\#) \rightarrow L^2(\mathbb{G}_0^\#)} + \|(H_k^\#)^*(H_j^\#(H_j^\#)^*)^r\|_{L^2(\mathbb{G}_0^\#) \rightarrow L^2(\mathbb{G}_0^\#)} \lesssim 2^{-\delta(j-k)} \quad (4.34)$$

for some $\delta = \delta(d) > 0$ and any $k \leq j \in [J_m(1-\kappa), J_m] \cap \mathbb{Z}$. The operator $H_k^\#((H_j^\#)^* H_j^\#)^r$ is a convolution operator on the group $\mathbb{G}_0^\#$ defined by the kernel

$$x \rightarrow \int_{\mathbb{R}} K_k(t) M_j''(A_0(t)^{-1} \cdot x) dt,$$

where

$$M_j''(x) = \prod_{(l_1, l_2) \in Y_d} 2^{-j(l_1+l_2)} \int_{\mathbb{R}^{|Y_d|}} P_{j,j,\dots,j,j}(\beta) e^{2\pi i(2^{-j} \circ x) \cdot \beta} d\beta.$$

Using (3.18) and integration by parts, the kernels M_j'' satisfy the same bounds as the kernels M_j defined in (4.3), namely

$$|M_j''(x)| + \sum_{(l_1, l_2) \in Y_d} 2^{j(l_1+l_2)} |\partial_{x_{l_1, l_2}} M_j''(x)| \lesssim (1 + |2^{-j} \circ x|)^{-(4d)^4} \prod_{(l_1, l_2) \in Y_d} 2^{-j(l_1+l_2)}.$$

Using the cancellation assumption $\int_{\mathbb{R}} K_k(t) dt = 0$ in (3.2), it follows that the $L^1(\mathbb{G}_0^\#)$ of the kernel of the operator $H_k^\#((H_j^\#)^* H_j^\#)^r$ is $\lesssim 2^{k-j}$, which suffices to prove the desired bound on the first term in the left-hand side of (4.34). The bound on the second term is similar. This completes the proof of the lemma. $\square$

Finally we verify the main inequalities in (6.21). Proposition 4.1 follows from Lemma 6.2, Lemma 4.2, and Lemma 4.5 below. This completes the proof of Theorem 2.3.

Lemma 4.5. *Assume $J_1, \dots, J_K \in [C(d), \infty)$ satisfy the separation condition*

$$J_{m+1} \geq 2J_m, \quad m = 1, \dots, K-1. \quad (4.35)$$

For $m = 1, \dots, K$ let, as before,

$$S_m = \sum_{j \in [J_m(1-\kappa), J_m] \cap \mathbb{Z}} H_j.$$

Then, for some $\delta = \delta(d) > 0$ and any $m = 1, \dots, K-1$

$$\begin{aligned} & \|S_m[(S_{m+1}^* S_{m+1})^r + \dots + (S_K^* S_K)^r]\|_{L^2 \rightarrow L^2} \\ & + \|S_m^*[(S_{m+1} S_{m+1}^*)^r + \dots + (S_K S_K^*)^r]\|_{L^2 \rightarrow L^2} \lesssim 2^{-\delta m}. \end{aligned} \quad (4.36)$$

Proof of Lemma 4.5. As before, we focus on the bound on the first term in (4.36). We already know from Lemma 4.4 that

$$\|(S_{m+1}^* S_{m+1})^r + \dots + (S_K^* S_K)^r\|_{L^2 \rightarrow L^2} \lesssim 1, \quad m = 1, \dots, K-1,$$

so it remains to prove that composition with the operator S_m contributes an additional factor of $2^{-\delta m}$.

We fix m and apply Proposition 3.2 to the operators $(S_n^* S_n)^r$, $n = m+1, \dots, K$. The contribution of the error terms is clearly acceptable. For $n = m+1, \dots, K$ and $a/q \in \mathcal{S}_{2^{3d^2}\epsilon J_n}$ let

$$U_n^{a/q} F(g) = \sum_{h \in \mathbb{G}_0} F(h^{-1} \cdot g) 2^{2\pi i h \cdot a/q} N_n(h), \quad (4.37)$$

where N_n are the kernels defined in (4.11). After rearranging the sum, for (4.36) it suffices to prove that

$$\sum_{a/q \in \mathcal{S}_\infty} S(a/q) \left\| S_m \sum_{n \in [m+1, K] \cap \mathbb{Z}, 2^{3d^2}\epsilon J_n \geq q} U_n^{a/q} \right\|_{L^2 \rightarrow L^2} \lesssim 2^{-\delta m}.$$

We already know, see (4.12), that

$$\left\| S_m \sum_{n \in [m+1, K] \cap \mathbb{Z}, 2^{3d^2}\epsilon J_n \geq q} U_n^{a/q} \right\|_{L^2 \rightarrow L^2} \lesssim q^{(4d)^4}.$$

In view of the rapid decay of the coefficients $S(a/q)$, see Lemma 3.1, it only remains to estimate the contribution of fractions a/q with denominators q small relative to 2^{J_m} ; more precisely, it remains to prove that for any $m \in [1, K-1] \cap \mathbb{Z}$ and any $a/q \in \mathcal{S}_{2^{\epsilon J_m}}$

$$\left\| S_m \sum_{n \in [m+1, K] \cap \mathbb{Z}} U_n^{a/q} \right\|_{L^2 \rightarrow L^2} \lesssim 2^{-\delta m} q^{(4d)^4}. \quad (4.38)$$

The kernel of the operator $S_m U_n^{a/q}$, $n \geq m+1$, is

$$g \rightarrow \sum_{t \in \mathbb{Z}} \sum_{j \in [J_m(1-\kappa), J_m]} K_j(t) e^{2\pi i (A_0(t)^{-1} \cdot g) \cdot a/q} N_n(A_0(t)^{-1} \cdot g)$$

which we write as

$$g \rightarrow Z_m(g) N_n(g) + \sum_{t \in \mathbb{Z}} \sum_{j \in [J_m(1-\kappa), J_m]} K_j(t) e^{2\pi i (A_0(t)^{-1} \cdot g) \cdot a/q} [N_n(A_0(t)^{-1} \cdot g) - N_n(g)]$$

where

$$Z_m(g) = \sum_{t \in \mathbb{Z}} \sum_{j \in [J_m(1-\kappa), J_m]} K_j(t) e^{2\pi i (A_0(t)^{-1} \cdot g) \cdot a/q}. \quad (4.39)$$

It follows from (4.13) and the separation condition (4.35) that

$$\left\| \sum_{t \in \mathbb{Z}} \sum_{j \in [J_m(1-\kappa), J_m]} K_j(t) e^{2\pi i(A_0(t)^{-1} \cdot g) \cdot a/q} [N_n(A_0(t)^{-1} \cdot g) - N_n(g)] \right\|_{L_g^1(\mathbb{G}_0)} \lesssim 2^{-J_n/4}$$

Therefore, for (4.38) it remains to prove that for any $m \in [1, K-1] \cap \mathbb{Z}$ and any $a/q \in \mathcal{S}_{2^\epsilon J_m}$

$$\left\| \sum_{h \in G_0} F(h^{-1} \cdot g) \sum_{n \in [m+1, K] \cap \mathbb{Z}} N_n(h) Z_m(h) \right\|_{L^2(\mathbb{G}_0)} \lesssim 2^{-\delta m} q^{(4d)^4} \|F\|_{L^2(\mathbb{G}_0)}. \quad (4.40)$$

We examine now the functions $Z_m : \mathbb{G}_0 \rightarrow \mathbb{C}$ defined in (4.39). Clearly,

$$Z_m(g_1 \cdot h \cdot g_2) = Z_m(g_1 \cdot g_2) \quad \text{for any } g_1, g_2 \in \mathbb{G}_0 \text{ and } h \in \mathbb{H}_q, \quad (4.41)$$

where the subgroup $\mathbb{H}_q$ is defined in (4.16). Moreover, for any $g \in \mathbb{G}_0$,

$$\begin{aligned} |Z_m(g)| &\leq \sum_{y \in Z_q} \left| \sum_{x \in \mathbb{Z}} \sum_{j \in [J_m(1-\kappa), J_m]} K_j(qx + y) e^{2\pi i(A_0(qx+y)^{-1} \cdot g) \cdot a/q} \right| \\ &\leq \sum_{y \in Z_q} \sum_{j \in [J_m(1-\kappa), J_m]} \left| \sum_{x \in \mathbb{Z}} K_j(qx + y) \right|. \end{aligned}$$

It follows from (3.2) and the assumption $q \leq 2^{\epsilon J_m}$ that

$$\sup_{g \in \mathbb{G}_0} |Z_m(g)| \lesssim 2^{-J_m/2}. \quad (4.42)$$

We turn now to the proof of (4.40), which is similar to the proof of (4.12). The functions Z_m replace the oscillatory factors $h \rightarrow e^{2\pi i h \cdot a/q}$; these functions satisfy the identities (4.41) and the estimates (4.42), which provide the additional exponential decay in m . We define the kernels N'_n as in (4.19) and the operators

$$V_n F(g) = \sum_{h \in \mathbb{G}_0} F(h^{-1} \cdot g) N'_n(h) Z_m(h).$$

In view of the Cotlar–Stein lemma, it suffices to prove that for any $n' \geq n \geq m+1$

$$\|V_n V_{n'}^*\|_{L^2 \rightarrow L^2} + \|V_n^* V_{n'}\|_{L^2 \rightarrow L^2} \lesssim 2^{-(n'-n)/100} 2^{-J_m/100}. \quad (4.43)$$

Using (4.18) and (4.41), for any $h \in \mathbb{G}_0$ and $k \in [m+1, K] \cap \mathbb{Z}$

$$\sum_{x \in \mathbb{G}_0} N'_k(x) Z_m(x) \overline{Z_m}(h \cdot x) = \sum_{x \in \mathbb{G}_0} N'_k(x) Z_m(x) \overline{Z_m}(x \cdot h) = 0.$$

Therefore, assuming first that $n' \geq n+1$ in (4.43), we write

$$\begin{aligned} (V_n V_{n'}^*) F(g) &= \sum_{h \in \mathbb{G}_0} F(h \cdot g) \left[\sum_{x \in \mathbb{G}_0} N'_n(x) Z_m(x) \overline{Z_m}(h \cdot x) [\overline{N'_{n'}}(h \cdot x) - \overline{N'_{n'}}(h)] \right], \\ (V_n^* V_{n'}) F(g) &= \sum_{h \in \mathbb{G}_0} F(h^{-1} \cdot g) \left[\sum_{x \in \mathbb{G}_0} \overline{N'_n}(x) \overline{Z_m}(x) Z_m(x \cdot h) [N'_{n'}(x \cdot h) - N'_{n'}(h)] \right]. \end{aligned}$$

Therefore, using (4.42),

$$\begin{aligned} & \|V_n V_{n'}^*\|_{L^2 \rightarrow L^2} + \|V_n^* V_{n'}\|_{L^2 \rightarrow L^2} \\ & \lesssim 2^{-J_m} \left\| \sum_{x \in \mathbb{G}_0} |N'_n(x)| [|N'_{n'}(h \cdot x) - N'_{n'}(h)| + |N'_{n'}(x \cdot h) - N'_{n'}(h)|] \right\|_{L_h^1(\mathbb{G}_0)}, \end{aligned}$$

and the desired bound (4.43) follows from (4.22) and (4.23) in this case.

Finally, to prove (4.43) when $n = n'$, it suffices to prove that

$$\left\| \sum_{h \in \mathbb{G}_0} F(h) N'_n(g \cdot h) Z_m(g \cdot h) \right\|_{L_g^2} \lesssim 2^{-J_m/20} \|F\|_{L^2},$$

for any $F \in L^2(\mathbb{G}_0)$ and $n \geq m+1$. Using the decomposition (4.16)-(4.17), it suffices to prove that

$$\sum_{b \in R_q} \left\| \sum_{x \in \mathbb{H}_q} F(x \cdot b) N'_n(g \cdot x \cdot b) Z_m(g \cdot x \cdot b) \right\|_{L_g^2} \lesssim 2^{-J_m/20} \|F\|_{L^2}.$$

Using (4.41)-(4.42), it suffices to prove that for any functions $F_b \in L^2(\mathbb{H}_q)$, $b \in R_q$,

$$\sum_{b \in R_q} \left\| \sum_{x \in \mathbb{H}_q} F_b(x) N'_n(g \cdot x \cdot b) \right\|_{L_g^2} \lesssim 2^{J_m/4} \left[\sum_{b \in R_q} \|F_b\|_{L^2(\mathbb{H}_q)}^2 \right]^{1/2}.$$

This bound was already proved in Lemma 4.4, see (4.26). $\square$

5. ESTIMATES ON OSCILLATORY SUMS AND OSCILLATORY INTEGRALS

With the notation in section 2, for $r \geq 1$ let $D, \tilde{D} : \mathbb{R}^r \times \mathbb{R}^r \rightarrow \mathbb{G}_0^\#$,

$$\begin{aligned} D((n_1, \dots, n_r), (m_1, \dots, m_r)) &= A_0(n_1)^{-1} \cdot A_0(m_1) \cdot \dots \cdot A_0(n_r)^{-1} \cdot A_0(m_r), \\ \tilde{D}((n_1, \dots, n_r), (m_1, \dots, m_r)) &= A_0(n_1) \cdot A_0(m_1)^{-1} \cdot \dots \cdot A_0(n_r) \cdot A_0(m_r)^{-1}, \end{aligned} \quad (5.1)$$

By definition, we have

$$[A_0(n)]_{l_1 l_2} = \begin{cases} n^{l_1} & \text{if } l_2 = 0, \\ 0 & \text{if } l_2 \geq 1, \end{cases} \quad [A_0(n)^{-1}]_{l_1 l_2} = \begin{cases} -n^{l_1} & \text{if } l_2 = 0, \\ n^{l_1 + l_2} & \text{if } l_2 \geq 1. \end{cases}$$

Thus, for $x = (x_1, \dots, x_r) \in \mathbb{R}^r$ and $y = (y_1, \dots, y_r) \in \mathbb{R}^r$

$$[D(x, y)]_{l_1 l_2} = \begin{cases} \sum_{j=1}^r (y_j^{l_1} - x_j^{l_1}) & \text{if } l_2 = 0, \\ \sum_{1 \leq j_1 < j_2 \leq r} (y_{j_1}^{l_1} - x_{j_1}^{l_1})(y_{j_2}^{l_2} - x_{j_2}^{l_2}) + \sum_{j=1}^r (x_j^{l_1 + l_2} - x_j^{l_1} y_j^{l_2}) & \text{if } l_2 \geq 1, \end{cases} \quad (5.2)$$

and

$$[\tilde{D}(x, y)]_{l_1 l_2} = \begin{cases} \sum_{j=1}^r (x_j^{l_1} - y_j^{l_1}) & \text{if } l_2 = 0, \\ \sum_{1 \leq j_1 < j_2 \leq r} (x_{j_1}^{l_1} - y_{j_1}^{l_1})(x_{j_2}^{l_2} - y_{j_2}^{l_2}) + \sum_{j=1}^r (y_j^{l_1 + l_2} - x_j^{l_1} y_j^{l_2}) & \text{if } l_2 \geq 1. \end{cases} \quad (5.3)$$

The multi-variable polynomials D and $\tilde{D}$ appear when we consider high powers of our singular integral operators, see for example the formula (3.10). In this section we prove two estimates on certain oscillatory sums and integrals involving these polynomials.

For integers $P \geq 1$ assume $\phi_P^{(j)}, \psi_P^{(j)} : \mathbb{R} \rightarrow \mathbb{R}$, $j = 1, \dots, r$, are C^1 functions with the properties

$$\sup_{j=1, \dots, r} [|\phi_P^{(j)}| + |\psi_P^{(j)}|] \leq \mathbf{1}_{[-P, P]}, \quad \sup_{j=1, \dots, r} \int_{\mathbb{R}} |[\phi_P^{(j)}]'(x)| + |[\psi_P^{(j)}]'(x)| dx \leq 1. \quad (5.4)$$

For $\theta = (\theta_{l_1 l_2})_{(l_1, l_2) \in Y_d} \in \mathbb{R}^{|Y_d|}$, $r \geq 1$, and $P \geq 1$ let

$$S_{P,r}(\theta) = \sum_{n, m \in \mathbb{Z}^r} e^{-2\pi i D((n_1, \dots, n_r), (m_1, \dots, m_r)) \cdot \theta} \phi_P^{(1)}(n_1) \dots \phi_P^{(r)}(n_r) \psi_P^{(1)}(m_1) \dots \psi_P^{(r)}(m_r)$$

and

$$\tilde{S}_{P,r}(\theta) = \sum_{n, m \in \mathbb{Z}^r} e^{-2\pi i \tilde{D}((n_1, \dots, n_r), (m_1, \dots, m_r)) \cdot \theta} \phi_P^{(1)}(n_1) \dots \phi_P^{(r)}(n_r) \psi_P^{(1)}(m_1) \dots \psi_P^{(r)}(m_r).$$

Proposition 5.1. *There is a constant $\overline{C} = \overline{C}(d)$ sufficiently large such that for all $r \geq 1$ and all $\epsilon \in (0, 1/2]$*

$$|S_{P,r}(\theta)| + |\tilde{S}_{P,r}(\theta)| \lesssim_r P^{2r} P^{\overline{C} - r\epsilon/\overline{C}}, \quad P = 1, 2, \dots, \quad (5.5)$$

provided that there is a pair $(l_1, l_2) \in Y_d$ and an irreducible fraction $a/q \in \mathbb{Q}$, $q \in \mathbb{Z}_+^$, such that*

$$|\theta_{l_1 l_2} - a/q| \leq 1/q^2 \text{ and } q \in [P^\epsilon, P^{l_1 + l_2 - \epsilon}].$$

To prove Proposition 5.1 we use a variant of the Weyl method, as in [10] and [3]. We provide all the details, for the sake of self-containedness, with the exception of the following key lemma, see Lemma 3.3 in [10]:

Lemma 5.2. *Assume that $L_1, \dots, L_n : \mathbb{R}^n \rightarrow \mathbb{R}$ are n linear forms, $L_j(u) = \sum_{k=1}^n \lambda_{jk} u_k$, satisfying the symmetry condition*

$$\lambda_{jk} = \lambda_{kj}, \quad j, k = 1, \dots, n. \quad (5.6)$$

Assume that $A > 1$, $Z \in (0, 1]$, and let $U(Z)$ denote the number of points $u \in \mathbb{Z}^n$ satisfying

$$|u| \leq ZA, \quad \sup_{j \in \{1, \dots, n\}} \|L_j(u)\| \leq ZA^{-1},$$

where $\|y\|$ denotes the distance from y to $\mathbb{Z}$ for any $y \in \mathbb{R}$. Then, for any $0 < Z_1 \leq Z_2 \leq 1$,

$$U(Z_2) \lesssim_n (Z_2/Z_1)^n U(Z_1).$$

Proof of Proposition 5.1. We will only prove the estimate for $|S_{P,r}(\theta)|$; the estimate for $|\tilde{S}_{P,r}(\theta)|$ follows by a very similar argument. It follows from (5.4) that $|S_{P,r}(\theta)| \lesssim_r P^{2r}$. Therefore, in proving (5.5) we may assume that $P \geq C_r$ and $r \geq \overline{C}^2/\epsilon$. We divide the proof in several steps.

Step 1. For $n = (n_1, \dots, n_r)$ fixed, let

$$D^0(m) = D^0(m_1, \dots, m_r) = D((n_1, \dots, n_r), (m_1, \dots, m_r)) \in \mathbb{Z}^{|Y_d|},$$

$$\Psi_P^0(m) = \psi_P^{(1)}(m_1) \dots \psi_P^{(r)}(m_r).$$

It suffices to prove that for any $n = (n_1, \dots, n_r) \in \mathbb{Z}^r$ fixed, with $|n_j| \leq P$,

$$|S_{P,r}^n(\theta)| \lesssim_r P^r P^{\bar{C} - r\epsilon/\bar{C}} \quad (5.7)$$

where

$$S_{P,r}^n(\theta) = \sum_{w \in \mathbb{Z}^r} e^{-2\pi i D^0(w) \cdot \theta} \Psi_P^0(w). \quad (5.8)$$

In addition, in view of (5.2),

$$D(w)_{l_1 l_2} \text{ is a polynomial of degree } l_1 + l_2 \text{ in } w \text{ for any } (l_1, l_2) \in Y_d. \quad (5.9)$$

We fix a sequence $0 < \delta_{2d-1} < \dots < \delta_1 < \epsilon$,

$$\delta_l = \epsilon/C_0^l, \quad C_0 = C_0(d) \gg 1. \quad (5.10)$$

Using Dirichlet's lemma, for any $(l_1, l_2) \in Y_d$ one can fix approximations

$$\theta_{l_1 l_2} = \frac{a_{l_1 l_2}}{q_{l_1 l_2}} + \beta_{l_1 l_2}, \quad a_{l_1 l_2}, q_{l_1 l_2} \in \mathbb{Z}, \quad (5.11)$$

$$(a_{l_1 l_2}, q_{l_1 l_2}) = 1, \quad 1 \leq q_{l_1 l_2} \leq P^{l_1 + l_2 - \delta_{l_1 + l_2}}, \quad |\beta_{l_1 l_2}| \leq (q_{l_1 l_2} P^{l_1 + l_2 - \delta_{l_1 + l_2}})^{-1}.$$

In view of the hypothesis, there is $d_0 \in \{1, \dots, 2d-1\}$ such that

$$q_{l_1 l_2} \leq P^{\delta_{l_1 + l_2}} \text{ if } l_1 + l_2 \geq d_0 + 1 \text{ and } q_{l_1 l_2} \geq P^{\delta_{d_0}} \text{ for some } l_1, l_2 \text{ with } l_1 + l_2 = d_0. \quad (5.12)$$

Let

$$D^l(w; v^{(1)}, \dots, v^{(l)}) = D^{l-1}(w + v^{(l)}; v^{(1)}, \dots, v^{(l-1)}) - D^{l-1}(w; v^{(1)}, \dots, v^{(l-1)}),$$

$$\Psi_P^l(w; v^{(1)}, \dots, v^{(l)}) = \Psi_P^{l-1}(w + v^{(l)}; v^{(1)}, \dots, v^{(l-1)}) \Psi_P^{l-1}(w; v^{(1)}, \dots, v^{(l-1)}),$$

for $l = 1, 2, \dots$. Using the formula (5.8),

$$|S_{P,r}^n(\theta)|^2 \leq \sum_{v^{(1)} \in \mathbb{Z}^r} \left| \sum_{w \in \mathbb{Z}^r} e^{-2\pi i (D^0(w + v^{(1)}) - D^0(w)) \cdot \theta} \Psi_P^0(w + v^{(1)}) \Psi_P^0(w) \right|$$

$$\leq \sum_{v^{(1)} \in \mathbb{Z}^r} \left| \sum_{w \in \mathbb{Z}^r} e^{-2\pi i D^1(w; v^{(1)}) \cdot \theta} \Psi_P^1(w; v^{(1)}) \right|.$$

We repeat this estimate $d_0 - 1$ times⁴. Using the Cauchy inequality, it follows that

$$|S_{P,r}^n(\theta)|^{2^{d_0-1}} P^{-r(2^{d_0-1} - d_0)} \lesssim_r \sum_{|v^{(1)}| + \dots + |v^{(d_0-1)}| \lesssim_r P} \left| \sum_{w \in \mathbb{Z}^r} e^{-2\pi i D^{d_0-1}(w; v^{(1)}, \dots, v^{(d_0-1)}) \cdot \theta} \Psi_P^{d_0-1}(w; v^{(1)}, \dots, v^{(d_0-1)}) \right|. \quad (5.13)$$

⁴If $d_0 = 1$ then the formula (5.8) gives already the estimate (5.13).

It follows from (5.9) that $[D^{d_0-1}(w; v^{(1)}, \dots, v^{(d_0-1)})]_{l_1 l_2}$ is a polynomial of degree at most $l_1 + l_2 - d_0 + 1$ in w , for any $v^{(1)}, \dots, v^{(d_0-1)} \in \mathbb{Z}^r$ fixed. Let

$$Q = \prod_{l_1 + l_2 \geq d_0 + 1} q_{l_1 l_2},$$

see (5.11). In view of the assumption (5.12),

$$1 \leq Q \leq P^{2d^2 \delta_{d_0+1}},$$

and we estimate, for any $v^{(1)}, \dots, v^{(d_0-1)} \in \mathbb{Z}^r$ fixed,

$$\begin{aligned} & \left| \sum_{w \in \mathbb{Z}^r} e^{-2\pi i D^{d_0-1}(w; v^{(1)}, \dots, v^{(d_0-1)}) \cdot \theta} \Psi_P^{d_0-1}(w; v^{(1)}, \dots, v^{(d_0-1)}) \right| \\ &= \left| \sum_{w \in \mathbb{Z}^r} e^{-2\pi i \sum_{l_1 + l_2 = d_0} D^{d_0-1}(w; v^{(1)}, \dots, v^{(d_0-1)})_{l_1 l_2} \cdot \theta_{l_1 l_2}} A(w) \right| \\ &\leq \sum_{y \in Z_Q^r} \left| \sum_{x \in \mathbb{Z}^r} e^{-2\pi i \sum_{l_1 + l_2 = d_0} D^{d_0-1}(x; v^{(1)}, \dots, v^{(d_0-1)})_{l_1 l_2} \cdot Q \theta_{l_1 l_2}} A(Qx + y) \right|, \end{aligned} \quad (5.14)$$

where

$$A(w) = e^{-2\pi i \sum_{l_1 + l_2 \geq d_0 + 1} D^{d_0-1}(w; v^{(1)}, \dots, v^{(d_0-1)})_{l_1 l_2} \cdot \theta_{l_1 l_2}} \Psi_P^{d_0-1}(w; v^{(1)}, \dots, v^{(d_0-1)}).$$

We examine now the function $A'(x) = A(Qx + y)$, $y \in Z_Q^r$ fixed. Using (5.11),

$$\begin{aligned} A'(x) &= A''(y, v^{(1)}, \dots, v^{(d_0-1)}) \\ &\quad \times \Psi_P^{d_0-1}(Qx + y; v^{(1)}, \dots, v^{(d_0-1)}) e^{-2\pi i \sum_{l_1 + l_2 \geq d_0 + 1} D^{d_0-1}(Qx + y; v^{(1)}, \dots, v^{(d_0-1)})_{l_1 l_2} \cdot \beta_{l_1 l_2}}, \end{aligned}$$

where $x \in \mathbb{Z}^r$ and $|A''(y, v^{(1)}, \dots, v^{(d_0-1)})| = 1$. By definition, see also (5.2), it is easy to see that $D^{d_0-1}(w; v^{(1)}, \dots, v^{(d_0-1)})_{l_1 l_2}$ is a polynomial of degree at most $l_1 + l_2$ in $w, n, v^{(1)}, \dots, v^{(d_0-1)}$ with coefficients $\lesssim_r 1$. Since $|\beta_{l_1 l_2}| \leq P^{-l_1 - l_2 + \delta_{d_0+1}}$ and $1 \leq Q \leq P^{2d^2 \delta_{d_0+1}}$,

$$\sup_{|x| \lesssim_r P} \left| \partial_{x_1}^{\sigma_1} \dots \partial_{x_r}^{\sigma_r} e^{-2\pi i \sum_{l_1 + l_2 \geq d_0 + 1} D^{d_0-1}(Qx + y; v^{(1)}, \dots, v^{(d_0-1)})_{l_1 l_2} \cdot \beta_{l_1 l_2}} \right| \lesssim_r P^{(-1 + 4d^2 \delta_{d_0+1})(\sigma_1 + \dots + \sigma_{2r})}$$

for all $y \in Z_Q^r$, all $n, v^{(1)}, \dots, v^{(d_0-1)} \in \mathbb{Z}^r$ with $|n| + |v^{(1)}| + \dots + |v^{(d_0-1)}| \lesssim_r P$, and $\sigma_1, \dots, \sigma_r \in \{0, 1\}$. Therefore, by summation by parts, it follows from (5.14) that

$$\begin{aligned} & \left| \sum_{w \in \mathbb{Z}^r} e^{-2\pi i D^{d_0-1}(w; v^{(1)}, \dots, v^{(d_0-1)}) \cdot \theta} \Psi_P^{d_0-1}(w; v^{(1)}, \dots, v^{(d_0-1)}) \right| \\ &\lesssim_r P^{20rd^2 \delta_{d_0+1}} \sup_{a_j, b_j \in [-2P, 2P]} \left| \sum_{x_j \in [a_j, b_j] \cap \mathbb{Z}} e^{-2\pi i \sum_{l_1 + l_2 = d_0} D^{d_0-1}(x; v^{(1)}, \dots, v^{(d_0-1)})_{l_1 l_2} \cdot Q \theta_{l_1 l_2}} \right| \\ &\lesssim_r P^{20rd^2 \delta_{d_0+1}} \prod_{j=1}^r \min(P, \|B_j(v^{(1)}, \dots, v^{(d_0-1)})\|^{-1}), \end{aligned}$$

where

$$B_j(v^{(1)}, \dots, v^{(d_0-1)}) = \frac{d}{dx_j} \left[\sum_{l_1+l_2=d_0} D^{d_0-1}(x; v^{(1)}, \dots, v^{(d_0-1)})_{l_1 l_2} \cdot Q\theta_{l_1 l_2} \right]. \quad (5.15)$$

In view of (5.13), it remains to prove that

$$\sum_{|v^{(1)}|+\dots+|v^{(d_0-1)}|\leq P} \prod_{j=1}^r \min(P, \|B_j(v^{(1)}, \dots, v^{(d_0-1)})\|^{-1}) \lesssim_r P^{rd_0} P^{\overline{C}} P^{-40rd^2\delta_{d_0+1}}, \quad (5.16)$$

assuming that $P^{\delta_{d_0}} \leq q_{l_1 l_2} \leq P^{l_1+l_2-\delta_{d_0}}$ for some $(l_1, l_2) \in Y_d$ with $l_1 + l_2 = d_0$, see (5.12).

For later use, we provide below a description of the functions B_j , $j = 1, \dots, r$. Assuming that $l_1 + l_2 = d_0$ and

$$D(w)_{l_1 l_2} = \sum_{j_1, \dots, j_{d_0}=1}^r \lambda_{j_1 \dots j_{d_0}}^{l_1 l_2} w_{j_1} \cdot \dots \cdot w_{j_{d_0}} \quad (5.17)$$

for some real-valued coefficients $\lambda_{j_1 \dots j_{d_0}}^{l_1 l_2}$ satisfying the symmetry condition

$$\lambda_{j_1 \dots j_{d_0}}^{l_1 l_2} = \lambda_{j_{\sigma(1)} \dots j_{\sigma(d_0)}}^{l_1 l_2} \quad \text{for any permutation } \sigma \text{ of the set } \{1, \dots, d_0\}, \quad (5.18)$$

it follows from the definition that

$$B_j(v^{(1)}, \dots, v^{(d_0-1)}) = d_0! \sum_{l_1+l_2=d_0} Q\theta_{l_1 l_2} \sum_{j_1, \dots, j_{d_0-1}=1}^r \lambda_{j_1 \dots j_{d_0-1}}^{l_1 l_2} v_{j_1}^{(1)} \cdot \dots \cdot v_{j_{d_0-1}}^{(d_0-1)}. \quad (5.19)$$

The claim (5.16) is easy to verify if $(l_1, l_2) = (1, 0)$, using directly the definition (5.1). Therefore, we will assume from now on that $2 \leq d_0 \leq 2d - 1$.

Step 2. We show now that it suffices to prove that

$$\begin{aligned} & \left| \{v^{(1)}, \dots, v^{(d_0-1)} \in B_{\mathbb{Z}^r}(P) : \sup_{j=1, \dots, r} \|B_j(v^{(1)}, \dots, v^{(d_0-1)})\| \leq P^{-1}\} \right| \\ & \lesssim_r P^{r(d_0-1)} P^{\overline{C}} P^{-80rd^2\delta_{d_0+1}}, \end{aligned} \quad (5.20)$$

where, by definition, $B_{\mathbb{Z}^m}(R) = \{v \in \mathbb{Z}^m : |v| \leq R\}$. Indeed, assuming (5.20), it follows that

$$\sum_{v^{(2)}, \dots, v^{(d_0-1)} \in B_{\mathbb{Z}^r}(P)} N_1(v^{(2)}, \dots, v^{(d_0-1)}) \lesssim_r P^{r(d_0-1)} P^{\overline{C}} P^{-80rd^2\delta_{d_0+1}},$$

where, for any $v^{(2)}, \dots, v^{(d_0-1)} \in B_{\mathbb{Z}^r}(P)$,

$$N_1(v^{(2)}, \dots, v^{(d_0-1)}) = \left| \{v^{(1)} \in B_{\mathbb{Z}^r}(P) : \|B_j(v^{(1)}, \dots, v^{(d_0-1)})\| \leq P^{-1}, j = 1, \dots, r\} \right|.$$

On the other hand, arguing as in [10, Lemma 3.2], for any $v^{(2)}, \dots, v^{(d_0-1)} \in B_{\mathbb{Z}^r}(P)$

$$\sum_{v^{(1)} \in B_{\mathbb{Z}^r}(P)} \prod_{j=1}^r \min(P, \|B_j(v^{(1)}, \dots, v^{(d_0-1)})\|^{-1}) \lesssim_r N_1(v^{(2)}, \dots, v^{(d_0-1)}) (P \log P)^r.$$

The desired bound (5.16) follows from these two estimates.

Step 3. Let

$$\rho = (\delta_{d_0+1}\delta_{d_0})^{1/2}, \quad \delta_{d_0+1} \ll \rho \ll \delta_{d_0}.$$

We show now that it suffices to prove that

$$\begin{aligned} |\{v^{(1)}, \dots, v^{(d_0-1)} \in B_{\mathbb{Z}^r}(P^\rho) : \sup_{j=1, \dots, r} \|B_j(v^{(1)}, \dots, v^{(d_0-1)})\| \leq P^{-d_0+(d_0-1)\rho}\}| \\ \lesssim_r P^{r(d_0-1)\rho} P^{C_1-r\rho/C_1}. \end{aligned} \quad (5.21)$$

for some constant $C_1 = C_1(d)$ sufficiently large. To prove that (5.21) implies (5.20), we prove that for $l = 0, \dots, d_0 - 1$ the number $N_{\rho, l}$ of solutions

$$\begin{aligned} v^{(1)}, \dots, v^{(l)} \in B_{\mathbb{Z}^r}(P), \quad v^{(l+1)}, \dots, v^{(d_0-1)} \in B_{\mathbb{Z}^r}(P^\rho), \\ \sup_{j=1, \dots, r} \|B_j(v^{(1)}, \dots, v^{(d_0-1)})\| \leq P^{-(d_0-l)+(d_0-1-l)\rho}, \end{aligned} \quad (5.22)$$

satisfies

$$N_{\rho, l} \lesssim_r P^{rl(1-\rho)} P^{r(d_0-1)\rho} P^{C_1-r\rho/C_1}. \quad (5.23)$$

In the case $l = 0$ this is equivalent to the assumption (5.21). The claim (5.23) follows by induction over l , using Lemma 5.2 at each step. The symmetry condition (5.6) is satisfied, in view of (5.17)-(5.19). The case $l = d_0 - 1$ gives the desired conclusion (5.20).

Step 4. For $j = 1, \dots, r$ and $(l_1, l_2) \in Y_d$ with $l_1 + l_2 = d_0$ let

$$A_j^{l_1 l_2}(v^{(1)}, \dots, v^{(d_0-1)}) = d_0! \sum_{j_1, \dots, j_{d_0-1}=1}^r \lambda_{j_1 \dots j_{d_0-1} j}^{l_1 l_2} v_{j_1}^{(1)} \dots v_{j_{d_0-1}}^{(d_0-1)}, \quad (5.24)$$

see (5.17)-(5.19). For any $v^{(1)}, \dots, v^{(d_0-1)}$ fixed we think of $A_j^{l_1 l_2}(v^{(1)}, \dots, v^{(d_0-1)})$ as a $r \times d_1$ matrix, where

$$d_1 = |Y_{d, d_0}|, \quad Y_{d, d_0} = \{(l_1, l_2) \in Y_d : l_1 + l_2 = d_0\}.$$

We show now that

$$\begin{aligned} \{v^{(1)}, \dots, v^{(d_0-1)} \in B_{\mathbb{Z}^r}(P^\rho) : \sup_{j=1, \dots, r} \|B_j(v^{(1)}, \dots, v^{(d_0-1)})\| \leq P^{-d_0+(d_0-1)\rho}\} \\ \subseteq \{v^{(1)}, \dots, v^{(d_0-1)} \in B_{\mathbb{Z}^r}(P^\rho) : \text{rank}[A_j^{l_1 l_2}(v^{(1)}, \dots, v^{(d_0-1)})] \leq d_1 - 1\}, \end{aligned} \quad (5.25)$$

provided that the constant C_0 fixed in (5.10) is sufficiently large (depending only on d). To see this, as in the proof of Lemma 2.5 in [3], assume, for contradiction, that $\sup_{j=1, \dots, r} \|B_j(v^{(1)}, \dots, v^{(d_0-1)})\| \leq P^{-d_0+(d_0-1)\rho}$ for some $v^{(1)}, \dots, v^{(d_0-1)} \in B_{\mathbb{Z}^r}(P^\rho)$ for which $\text{rank}[A_j^{l_1 l_2}(v^{(1)}, \dots, v^{(d_0-1)})] = d_1$. Notice that

$$B_j(v^{(1)}, \dots, v^{(d_0-1)}) = \sum_{l_1+l_2=d_0} Q\theta_{l_1 l_2} A_j^{l_1 l_2}(v^{(1)}, \dots, v^{(d_0-1)}).$$

We could then solve the linear system in the variables $Q\theta_{l_1 l_2}$ to deduce that

$$Q\theta_{l_1 l_2} = \frac{m_{l_1 l_2}}{n_{l_1 l_2}} + \delta_{l_1 l_2}, \quad m_{l_1 l_2}, n_{l_1 l_2} \in \mathbb{Z}, \quad 1 \leq n_{l_1 l_2} \lesssim_r P^{d_1(d_0-1)\rho}, \quad |\delta_{l_1 l_2}| \lesssim_r P^{-d_0+d_1(d_0-1)\rho}$$

for any $(l_1, l_2) \in Y_{d, d_0}$. Recalling the bound $1 \leq Q \leq P^{2d^2\delta_{d_0+1}}$ and the definition $\rho = (\delta_{d_0+1}\delta_{d_0})^{1/2}$, this is clearly in contradiction with (5.11)-(5.12) if P is sufficiently large relative to r and $C_0 = \delta_{d_0}/\delta_{d_0+1}$ is sufficiently large relative to d .

Therefore, for (5.21) it suffices to prove that

$$\begin{aligned} |\{v^{(1)}, \dots, v^{(d_0-1)} \in B_{\mathbb{Z}^r}(P^\rho) : \text{rank}[A_j^{l_1 l_2}(v^{(1)}, \dots, v^{(d_0-1)})] \leq d_1 - 1\}| \\ \lesssim_r P^{r(d_0-1)\rho} P^{C_1 - r\rho/C_1}. \end{aligned} \quad (5.26)$$

Recall that (see (5.2))

$$D^0(m)_{l_1 l_2} = \begin{cases} \sum_{1 \leq j \leq r} m_j^{l_1} + R_{l_1 l_2}^0(m) & \text{if } (l_1, l_2) = (d_0, 0), \\ \sum_{1 \leq j_1 < j_2 \leq r} m_{j_1}^{l_1} m_{j_2}^{l_2} + R_{l_1 l_2}^0(m) & \text{if } (l_1, l_2) \in Y_{d, d_0}, \quad l_2 \geq 1, \end{cases} \quad (5.27)$$

where $R_{l_1 l_2}^0$ are polynomials in m of degree at most $d_0 - 1$. These polynomials give no contribution to the values of $A_j^{l_1 l_2}(v^{(1)}, \dots, v^{(d_0-1)})$. Using the definitions, it follows that for fixed $1 \leq j \leq r$

$$A_j^{l_1 l_2}(v^{(1)}, \dots, v^{(d_0-1)}) = d_0 \sum_{\sigma} v_j^{(\sigma_1)} \dots v_j^{(\sigma_{d_0-1})} \quad (5.28)$$

if $(l_1, l_2) = (d_0, 0)$, and

$$\begin{aligned} A_j^{l_1 l_2}(v^{(1)}, \dots, v^{(d_0-1)}) &= l_1 \sum_{\sigma} \sum_{j < k \leq r} v_j^{(\sigma_1)} \dots v_j^{(\sigma_{l_1-1})} v_k^{(\sigma_{l_1})} \dots v_k^{(\sigma_{d_0-1})} \\ &\quad + l_2 \sum_{\sigma} \sum_{1 \leq k < j} v_k^{(\sigma_1)} \dots v_k^{(\sigma_{l_1})} v_j^{(\sigma_{l_1+1})} \dots v_j^{(\sigma_{d_0-1})}, \end{aligned} \quad (5.29)$$

if $(l_1, l_2) \in Y_{d, d_0}$, $l_2 \geq 1$. Here $\sigma = (\sigma_1, \dots, \sigma_{d_0-1})$ runs through all the permutations of the set $\{1, 2, \dots, d_0 - 1\}$.

Step 5. We examine now the set in the left-hand side of (5.26). Since the matrix coefficients $A_j^{l_1 l_2}(v^{(1)}, \dots, v^{(d_0-1)})$ are integers and of size $\lesssim P^{(d_0-1)\rho}$, it is easy to see from Cramer's rule that if $\text{rank}[A_j^{l_1 l_2}(v^{(1)}, \dots, v^{(d_0-1)})] \leq d_1 - 1$ for some $v^{(1)}, \dots, v^{(d_0-1)} \in B_{\mathbb{Z}^r}(P^\rho)$ then there exists a set of integers $b_{l_1 l_2}$ not all zero of size $|b_{l_1 l_2}| \lesssim P^{C_2\rho}$ (with a constant C_2 depending only on d), such that

$$\sum_{(l_1, l_2) \in Y_{d, d_0}} b_{l_1 l_2} A_j^{l_1 l_2}(v^{(1)}, \dots, v^{(d_0-1)}) = 0 \quad \text{for all } 1 \leq j \leq r. \quad (5.30)$$

For a given permutation $\sigma = (\sigma_1, \dots, \sigma_{d_0-1})$ and a given pair $(l_1, l_2) \in Y_{d, d_0}$ such that $l_2 \geq 1$, define

$$T_{l_1 l_2}^\sigma = \sum_{k=1}^r v_k^{(\sigma_{l_1})} \dots v_k^{(\sigma_{d_0-1})}.$$

We define, compare with 5.29,

$$\begin{aligned} \tilde{A}_j^{l_1 l_2}(v^{(1)}, \dots, v^{(d_0-1)}) &= l_2 \sum_{\sigma} \sum_{1 \leq k < j} v_k^{(\sigma_1)} \dots v_k^{(\sigma_{l_1})} v_j^{(\sigma_{l_1+1})} \dots v_j^{(\sigma_{d_0-1})} \\ &\quad - l_1 \sum_{\sigma} \sum_{1 \leq k \leq j} v_j^{(\sigma_1)} \dots v_j^{(\sigma_{l_1-1})} v_k^{(\sigma_{l_1})} \dots v_k^{(\sigma_{d_0-1})} \\ &\quad + l_1 \sum_{\sigma} T_{l_1 l_2}^\sigma v_j^{(\sigma_1)} \dots v_j^{(\sigma_{l_1-1})}. \end{aligned} \quad (5.31)$$

The advantage of formula 5.31 is that for any fixed values of the parameters $T_{l_1 l_2}^\sigma$, the quantities $\tilde{A}_j^{l_1 l_2}(v^{(1)}, \dots, v^{(d_0-1)})$ depend only on the variables $v_k^{(m)}$ for $1 \leq m \leq d_0 - 1$ and $1 \leq k \leq j$. We define also, compare with (5.28),

$$A_j^{d_0 0}(v^{(1)}, \dots, v^{(d_0-1)}) = A_j^{d_0 0}(v^{(1)}, \dots, v^{(d_0-1)}) = d_0 \sum_{\sigma} v_j^{(\sigma_1)} \dots v_j^{(\sigma_{d_0-1})}.$$

Using these definitions and (5.30), we conclude that if $(v^{(1)}, \dots, v^{(d_0-1)})$ is an element of the set in the left-hand side of (5.26) then there are integers $b_{l_1 l_2}$ (not all zero) and $T_{l_1 l_2}^\sigma$ in $[-P, P]$ such that

$$\sum_{(l_1, l_2) \in Y_{d, d_0}} b_{l_1 l_2} \tilde{A}_j^{l_1 l_2}(v^{(1)}, \dots, v^{(d_0-1)}) = 0 \quad \text{for all } 1 \leq j \leq r.$$

Therefore, for (5.26) it suffices to prove that for any integers $b_{l_1 l_2}$ (not all zero) and $T_{l_1 l_2}^\sigma$ in $[-P, P]$

$$\begin{aligned} &|\{v^{(1)}, \dots, v^{(d_0-1)} \in B_{\mathbb{Z}^r}(P^\rho) : \\ &\quad \sum_{(l_1, l_2) \in Y_{d, d_0}} b_{l_1 l_2} \tilde{A}_j^{l_1 l_2}(v^{(1)}, \dots, v^{(d_0-1)}) = 0 \text{ for all } 1 \leq j \leq r\}| \lesssim_r P^{r(d_0-1)\rho} P^{-r\rho/C_1}. \end{aligned} \quad (5.32)$$

Step 6. Finally, we prove (5.32) using the simple Lemma 5.3 below. Let $1 \leq j \leq r$ be a given even integer. For any given choice of the parameters $b_{l_1 l_2}$ (not all zero), $T_{l_1 l_2}^\sigma$ and for any given values of the variables $v_k^{(h)}$, $1 \leq h \leq d_0 - 1, 1 \leq k \leq j - 2$ we claim that

$$\sum_{(l_1, l_2) \in Y_{d, d_0}} b_{l_1 l_2} \tilde{A}_j^{l_1 l_2}(v^{(1)}, \dots, v^{(d_0-1)}) \quad (5.33)$$

is not identically zero as a polynomial in the variables $v_{j-1}^{(1)}, v_j^{(1)}, \dots, v_{j-1}^{(d_0-1)}, v_j^{(d_0-1)}$. Indeed, if $b_{l_1 l_2} \neq 0$ for a pair $(l_1, l_2) \in Y_{d, d_0} \setminus \{(d_0, 0)\}$, then, for any permutation σ , the expression 5.33 contains the term

$$b_{l_1 l_2} l_2 v_{j-1}^{(\sigma_1)} \dots v_{j-1}^{(\sigma_{l_1})} v_j^{(\sigma_{l_1+1})} \dots v_j^{(\sigma_{d_0-1})}.$$

If, on the other hand, $b_{d_0 0} \neq 0$ for $l_1 = d_0 - 1, l_2 = 0$ but $b_{l_1 l_2} = 0$ for all pairs $(l_1, l_2) \in Y_{d, d_0} \setminus \{(d_0, 0)\}$, then the expression 5.33 takes the form

$$b_{d_0 0} d_0 \sum_{\sigma} v_j^{(\sigma_1)} \dots v_j^{(\sigma_{d_0-1})}$$

which is not identically zero.

Therefore we may apply estimate 5.34 repeatedly for $j = 2, 4, \dots$. It follows that the number of solutions $(v^{(1)}, \dots, v^{(d_0-1)}) \in B_{\mathbb{Z}(d_0-1)r}(P^\rho)$ of the system of equations in (5.32) is $\lesssim P^{r(d_0-1)\rho-r\rho/2}$, as desired. $\square$

Lemma 5.3. *Assume that $P = P(x_1, \dots, x_s)$ is a polynomial of degree d in s variables which is not identically 0, and $A \subseteq \mathbb{R}$. Then*

$$|\{(x_1, \dots, x_s) \in A^s : P(x_1, \dots, x_s) = 0\}| \leq d|A|^{s-1}. \quad (5.34)$$

Proof of Lemma 5.3. The statement is immediate when $s = 1$ or $d = 1$. We proceed by induction. Without loss of generality assume that

$$P(x_1, \dots, x_s) = Q(x_2, \dots, x_s)x_1^{d_1} + R(x_1, \dots, x_s)$$

where $Q(x_2, \dots, x_s)$ is a polynomial of degree at most $d - d_1$ not identically zero. If $Q(x_2, \dots, x_s) \neq 0$ then there are at most d_1 values of x_1 for which $P(x_1, x_2, \dots, x_s) = 0$. Thus, by induction, the left-hand side of 5.34 is estimated by

$$d_1|A|^{s-1} + (d - d_1)|A|^{s-2}|A| = d|A|^{s-1},$$

as desired. $\square$

We conclude this section with an estimate on an oscillatory integral. We think of $D, \tilde{D}$ as functions defined on $\mathbb{R}^r \times \mathbb{R}^r$ taking values in $\mathbb{R}^{|Y_d|}$, given by (5.2) and (5.3).

Lemma 5.4. *Assume $\Phi : \mathbb{R}^r \times \mathbb{R}^r \rightarrow \mathbb{R}$ satisfies*

$$|\partial_{x_1}^{\sigma_1} \dots \partial_{x_r}^{\sigma_r} \partial_{y_1}^{\vartheta_1} \dots \partial_{y_r}^{\vartheta_r} \Phi(x, y)| \leq \mathbf{1}_{B_{\mathbb{R}^r}(1)}(x) \mathbf{1}_{B_{\mathbb{R}^r}(1)}(y) \quad (5.35)$$

for any $\sigma_1, \dots, \sigma_r, \vartheta_1, \dots, \vartheta_r \in \{0, 1\}$, where $B_{\mathbb{R}^m}(C) = \{x \in \mathbb{R}^m : |x| \leq C\}$. Then there is a constant $\overline{C} = \overline{C}(d)$ sufficiently large such that for any $\beta \in \mathbb{R}^{|Y_d|}$,

$$\left| \int_{\mathbb{R}^r \times \mathbb{R}^r} \Phi(x, y) e^{-2\pi i D(x, y) \cdot \beta} dx dy \right| + \left| \int_{\mathbb{R}^r \times \mathbb{R}^r} \Phi(x, y) e^{-2\pi i \tilde{D}(x, y) \cdot \beta} dx dy \right| \lesssim_r (1 + |\beta|)^{\overline{C} - r/\overline{C}}. \quad (5.36)$$

Proof of Lemma 5.4. We will only prove the estimate on the first term in the left-hand side of (5.36), using Proposition 5.1. Let C_0, C_1 are suitably large fixed constants (depending on the constant $\bar{C}$ in Proposition 5.1), assume $|\beta| \geq C_0$, and choose $\epsilon = C_1/r$ in Proposition 5.1. Assume that

$$(n_1, n_2) \in Y_d, \quad |\beta_{n_1 n_2}| = \sup_{(l_1, l_2) \in Y_d} |\beta_{l_1 l_2}|.$$

Let P be a positive number, so that $P \approx |\beta|^{1/\epsilon}$ and $q := P^{n_1+n_2} |\beta_{n_1 n_2}|^{-1}$ is an integer.

By rescaling one may write

$$I_D(\beta) := \int_{\mathbb{R}^r \times \mathbb{R}^r} \Phi(x, y) e^{-2\pi i D(x, y) \cdot \beta} dx dy = P^{-2r} \int_{\mathbb{R}^r \times \mathbb{R}^r} \Phi\left(\frac{x}{P}, \frac{y}{P}\right) e^{-2\pi i D(x, y) \cdot \theta} dx dy$$

where

$$\theta_{l_1 l_2} = P^{-(l_1+l_2)} \beta_{l_1 l_2}, \quad (l_1, l_2) \in Y_d.$$

Note that $\theta_{n_1 n_2} = \pm 1/q$ with $q \approx P^{n_1+n_2-\epsilon}$. Therefore, by Proposition 5.1, one has the estimate

$$P^{-2r} S_{P,r}(\theta) := P^{-2r} \sum_{(n,m) \in \mathbb{Z}^r \times \mathbb{Z}^r} \Phi(n/P, m/P) e^{-2\pi i D(n,m) \cdot \theta} \lesssim_r P^{\bar{C}-r\epsilon/\bar{C}} \lesssim_r P^{-1}.$$

On the other hand writing $x = n + s$, $y = m + t$ with $m, n \in \mathbb{Z}^r$ and $s, t \in [0, 1)^r$ it is easy to see that

$$|I_D(\beta) - P^{-2r} S_{P,r}(\theta)| \lesssim \sum_{(l_1, l_2) \in Y_d} |\theta_{l_1 l_2}| P^{l_1+l_2-1} + P^{-1} \lesssim P^{-1/2}.$$

This gives the estimate $|I_D(\beta)| \lesssim_r |\beta|^{-r/2C_1}$ for $|\beta| \geq C_0$ and the lemma follows. $\square$

6. AN ALMOST ORTHOGONALITY LEMMA

We assume that H is a Hilbert space, $S_m \in \mathcal{L}(H)$, $m = 1, \dots, K$, are self-adjoint operators, and

$$\|S_m\| \leq 1, \quad m = 1, \dots, K. \quad (6.1)$$

Let

$$I = \{0, 1\}, \quad S_{m,0} = S_m, \quad S_{m,1} = 0.$$

For any dyadic integer p we define

$$B_p = \sup_{i_1, \dots, i_K \in I} \|S_{1,i_1}^p + S_{2,i_2}^p + \dots + S_{K,i_K}^p\|. \quad (6.2)$$

and, for any $m = 1, \dots, K-1$ and dyadic integer p

$$\gamma_{m,p} = \sup_{i_m, \dots, i_K \in I} \|S_{m,i_m} (S_{m+1,i_{m+1}}^p + \dots + S_{K,i_K}^p)\|. \quad (6.3)$$

We start with a lemma:

Lemma 6.1. *Assume that $S_{m,i}$, B_p , γ_{mp} are as above and that there are constants $\delta_0 > 0$, $A \geq 1$ and a dyadic integer p_0 such that*

$$\gamma_{m,p_0} \leq A2^{-\delta_0 m}(B_{p_0} + 1) \quad \text{for } m = 1, \dots, K-1. \quad (6.4)$$

Then

$$\begin{aligned} B_1 &\leq C(\delta_0, A, p_0), \\ \gamma_{m,1} &\leq C(\delta_0, A, p_0)2^{-\delta'_0 m}, \quad m = 1, \dots, K-1, \end{aligned} \quad (6.5)$$

for some constants $C = C(\delta_0, A, p_0) \in [1, \infty)$ and $\delta'_0 = \delta'_0(\delta_0, A, p_0) > 0$.

Proof. We prove the lemma in two steps.

Step 1. We show first that

$$B_{p_0} \leq C(\delta_0, A). \quad (6.6)$$

Assume $p \geq p_0$ is a dyadic integer and fix $i_1, \dots, i_K \in I$ such that the supremum in (6.2) is attained. Then, using self-adjointness and (6.1), we write

$$\begin{aligned} B_p^2 &= \|(S_{1,i_1}^p + S_{2,i_2}^p + \dots + S_{K,i_K}^p)^2\| \\ &\leq \|S_{1,i_1}^{2p} + \dots + S_{K,i_K}^{2p}\| + 2 \sum_{m=1}^{K-1} \|S_{m,i_m}^p (S_{m+1,i_{m+1}}^p + \dots + S_{K,i_K}^p)\| \\ &\leq B_{2p} + 2 \sum_{m=1}^{K-1} \gamma_{m,p}. \end{aligned} \quad (6.7)$$

We estimate also $\gamma_{m,2p}$. For any $j_m, \dots, j_K \in I$

$$\begin{aligned} \|S_{m,j_m}(S_{m+1,j_{m+1}}^{2p} + \dots + S_{K,j_K}^{2p})\| &\leq \|S_{m,j_m}(S_{m+1,j_{m+1}}^p + \dots + S_{K,j_K}^p)^2\| \\ &\quad + 2 \sum_{m'=m+1}^{K-1} \|S_{m',j_{m'}}^p (S_{m'+1,j_{m'+1}}^p + \dots + S_{K,j_K}^p)\| \\ &\leq B_p \gamma_{m,p} + 2 \sum_{m'=m+1}^{K-1} \gamma_{m',p}, \end{aligned}$$

using (6.1) and the identity

$$\begin{aligned} S_{m+1,j_{m+1}}^{2p} + \dots + S_{K,j_K}^{2p} &= (S_{m+1,j_{m+1}}^p + \dots + S_{K,j_K}^p)^2 \\ &\quad - \sum_{m'=m+1}^{K-1} S_{m',j_{m'}}^p (S_{m'+1,j_{m'+1}}^p + \dots + S_{K,j_K}^p) - (S_{m'+1,j_{m'+1}}^p + \dots + S_{K,j_K}^p) S_{m',j_{m'}}^p. \end{aligned}$$

Thus, for any $m = 1, \dots, K$ and any dyadic integer $p \geq p_0$

$$\gamma_{m,2p} \leq B_p \gamma_{m,p} + 2 \sum_{m'=m+1}^{K-1} \gamma_{m',p}. \quad (6.8)$$

We use now inequalities (6.4), (6.7), and (6.8) to prove (6.6). Let

$$L = L(\delta_0) = \sum_{m=0}^{\infty} 2^{-\delta_0 m}.$$

Let $p_1 \geq p_0$ denote the smallest dyadic integer for which $B_{p_1} \leq (100LA)^{p_1}$. Such p_1 exists because $B_p \leq K$, using (6.1). The bound (6.6) follows if $p_1 = p_0$. Otherwise we have, for any dyadic integer $p \in [p_0, p_1)$ and any $m = 1, \dots, K$

$$\begin{aligned} B_p &> (100LA)^p; \\ B_p^2 &\leq B_{2p} + 2 \sum_{m=1}^{K-1} \gamma_{m,p}; \\ \gamma_{m,2p} &\leq B_p \gamma_{m,p} + 2 \sum_{m'=m}^{K-1} \gamma_{m',p}. \end{aligned} \tag{6.9}$$

It follows from the second equation of (6.9) and (6.4) that

$$B_{p_0}^2 \leq B_{2p_0} + 4ALB_{p_0}.$$

Using the first equation of (6.9) it follows that

$$B_{p_0}^2 \leq 2B_{2p_0}.$$

Using the third equation of (6.9) and (6.4) it follows that

$$\gamma_{m,2p_0} \leq B_{p_0} 2A 2^{-\delta_0 m} B_{p_0} + 4ALB_{p_0} 2^{-\delta_0 m} \leq 2^{-\delta_0 m} B_{2p_0} (8A),$$

using $B_{p_0} \geq 2L$ and $B_{p_0}^2 \leq 2B_{2p_0}$.

More generally, we prove by induction that for any dyadic integer $p \in [p_0, p_1)$ and any $m = 1, \dots, K$

$$B_p^2 \leq 2B_{2p} \quad \text{and} \quad \gamma_{m,2p} \leq 2^{-\delta_0 m} B_{2p} (4A)^{2p}. \tag{6.10}$$

This was already proved above for $p = p_0$. Assume $p \in [2p_0, p_1)$ is a dyadic integer. It follows from the second inequality in (6.9) and the induction hypothesis that

$$B_p^2 \leq B_{2p} + 2L(4A)^p B_p.$$

Since $B_p > (100LA)^p$, this gives the first inequality in (6.10). Using the third inequality in (6.9) and the induction hypothesis,

$$\gamma_{m,2p} \leq B_p 2^{-\delta_0 m} B_p (4A)^p + 2 \cdot 2^{-\delta_0 m} L B_p (4A)^p \leq 2^{-\delta_0 m} B_{2p} (4A)^{2p},$$

using $B_p^2 \leq 2B_{2p}$ and $B_p \geq 2L$. By induction, this completes the proof of (6.10).

Recall now that $B_{p_1} \leq (100LA)^{p_1}$. Thus, using only the first inequality in (6.10),

$$\begin{aligned} B_{p_1/2} &\leq 2^{1/2}(100LA)^{p_1/2} \\ B_{p_1/4} &\leq 2^{1/2}2^{1/4}(100LA)^{p_1/4} \\ &\dots \\ B_{p_1/2^l} &\leq 2^{1/2}2^{1/4} \dots 2^{1/2^l}(100LA)^{p_1/2^l}. \end{aligned}$$

The bound (6.6) follows by letting $2^l = p_1/p_0$.

Step 2. We prove now the bound (6.5). It follows from (6.4) and (6.6) that

$$B_{p_0} \leq A' \quad \text{and} \quad \gamma_{m,p_0} \leq A'2^{-\delta_0 m} \quad \text{for } m = 1, \dots, K, \quad (6.11)$$

for some constant $A' = A'(\delta_0, A)$. We would like to prove that, for some constant $A'' = A''(A', \delta_0)$

$$B_{p_0/2} \leq A'' \quad \text{and} \quad \gamma_{m,p_0/2} \leq A''2^{-\delta_0 m/4} \quad \text{for } m = 1, \dots, K. \quad (6.12)$$

We would then be able to prove (6.5) by repeating this step finitely many times.

We may assume $p_0 \geq 2$ and look at $B_{p_0/2}$. Fix $i_1, \dots, i_K \in I$ which attain the supremum in the definition of $B_{p_0/2}$ and write

$$\begin{aligned} B_{p_0/2}^2 &= \|(S_{1,i_1}^{p_0/2} + \dots + S_{K,i_K}^{p_0/2})^2\| \leq \|S_{1,i_1}^{p_0} + \dots + S_{K,i_K}^{p_0}\| \\ &\quad + 2 \sum_{m=1}^{K-1} \|S_{m,i_m}^{p_0/2} (S_{m+1,i_{m+1}}^{p_0/2} + \dots + S_{K,i_K}^{p_0/2})\| \\ &\leq A' + 2 \sum_{m=1}^{K-1} \|S_{m,i_m} (S_{m+1,i_{m+1}}^{p_0/2} + \dots + S_{K,i_K}^{p_0/2})\|, \end{aligned} \quad (6.13)$$

using (6.1). Let

$$Q = \sup_{m=1, \dots, K-1} \sup_{j_m, \dots, j_K \in I} 2^{\delta_0 m/4} \|S_{m,j_m} (S_{m+1,j_{m+1}}^{p_0/2} + \dots + S_{K,j_K}^{p_0/2})\|. \quad (6.14)$$

Fix $m, j_m, \dots, j_K$ such that the supremum in (6.14) is attained. Then we have

$$\begin{aligned} Q &= 2^{\delta_0 m/4} \|S_{m,j_m} (S_{m+1,j_{m+1}}^{p_0/2} + \dots + S_{K,j_K}^{p_0/2})\| \leq 2^{\delta_0 m/4} \sum_{m'=m+1}^{8m} \|S_{m,j_m} S_{m',j_{m'}}^{p_0/2}\| \\ &\quad + 2^{\delta_0 m/4} \|S_{m,j_m} (S_{8m+1,j_{8m+1}}^{p_0/2} + \dots + S_{K,j_K}^{p_0/2})\|. \end{aligned} \quad (6.15)$$

Now, using the second inequality in (6.11) and the definition of Q in (6.14), (6.1), selfadjointness, and the hypothesis $S_{m,0} = 0$

$$\|S_{m,j_m} S_{m',j_{m'}}^{p_0/2}\|^2 \leq \|S_{m,j_m} S_{m',j_{m'}}^{p_0} S_{m,j_m}\| \leq A'2^{-\delta_0 m},$$

and

$$\begin{aligned}
\|S_{m,j_m}(S_{8m+1,j_{8m+1}}^{p_0/2} + \dots + S_{K,j_K}^{p_0/2})\|^2 &\leq \|S_{m,j_m}(S_{8m+1,j_{8m+1}}^{p_0} + \dots + S_{K,j_K}^{p_0})\| \\
&+ 2 \sum_{m' \geq 8m} \|S_{m',j_{m'}}^{p_0/2}(S_{m'+1,j_{m'+1}}^{p_0/2} + \dots + S_{K,j_K}^{p_0/2})\| \\
&\leq A'2^{-\delta_0 m} + 2 \sum_{m' \geq 8m} Q2^{-\delta_0 m'} \leq 2^{-\delta_0 m}(A' + 2LQ).
\end{aligned}$$

Therefore, it follows from (6.15) and the last two inequalities that

$$Q \leq 2^{\delta_0 m/4} 2^{-\delta_0 m/2} \sqrt{A'}(7m) + 2^{\delta_0 m/4} 2^{-\delta_0 m/2} \sqrt{A' + 2LQ} \leq C_{\delta_0} \sqrt{A' + 2LQ}.$$

It follows that $Q \leq C(\delta_0, A')$. In view of the definition (6.14), this proves the second inequality in (6.12). The first inequality in (6.12) follows from (6.13). This completes the proof of the lemma. $\square$

We will need a version of this lemma for non-selfadjoint operators.

Lemma 6.2. *Assume that H is a Hilbert space, $S_m \in \mathcal{L}(H)$, $m = 1, \dots, K$, and*

$$\|S_m\| \leq 1, \quad m = 1, \dots, K. \quad (6.16)$$

Let

$$I = \{0, 1\}, \quad S_{m,0} = S_m, \quad S_{m,1} = 0.$$

For any dyadic integer p we define

$$\begin{aligned}
D_p &= \sup_{i_1, \dots, i_K \in I} \|(S_{1,i_1} S_{1,i_1}^*)^p + \dots + (S_{K,i_K} S_{K,i_K}^*)^p\|, \\
\tilde{D}_p &= \sup_{i_1, \dots, i_K \in I} \|(S_{1,i_1}^* S_{1,i_1})^p + \dots + (S_{K,i_K}^* S_{K,i_K})^p\|.
\end{aligned} \quad (6.17)$$

For any $m = 1, \dots, K-1$ and dyadic integer p we define

$$\begin{aligned}
\mu_{m,p} &= \sup_{i_m, \dots, i_K \in I} \|(S_{m,i_m} S_{m,i_m}^*)[(S_{m+1,i_{m+1}} S_{m+1,i_{m+1}}^*)^p + \dots + (S_{K,i_K} S_{K,i_K}^*)^p]\|, \\
\tilde{\mu}_{m,p} &= \sup_{i_m, \dots, i_K \in I} \|(S_{m,i_m}^* S_{m,i_m})[(S_{m+1,i_{m+1}}^* S_{m+1,i_{m+1}})^p + \dots + (S_{K,i_K}^* S_{K,i_K})^p]\|.
\end{aligned} \quad (6.18)$$

Assume that

$$\mu_{m,p_0} \leq A2^{-\delta_0 m}(D_{p_0} + 1) \text{ and } \tilde{\mu}_{m,p_0} \leq A2^{-\delta_0 m}(\tilde{D}_{p_0} + 1), \quad m = 1, \dots, K-1, \quad (6.19)$$

for some dyadic integer p_0 and some numbers $A \geq 1$ and $\delta_0 > 0$. Then

$$\|S_1 + \dots + S_K\| \leq C(\delta_0, A, p_0). \quad (6.20)$$

Remark 6.3. A simplified version of the lemma, which is used in the paper, is the following: assume that H is a Hilbert space, $S_m \in \mathcal{L}(H)$, $m = 1, \dots, K$, and let $S_{m,0} = S_m$, $S_{m,1} = 0$. Assume that, for all $m = 1, \dots, K$,

$$\begin{aligned} \sup_{m \in \{1, \dots, K\}} \|S_m\| &\leq 1, \\ \sup_{i_m, \dots, i_K \in I} \|S_{m,i_m}^* [(S_{m+1,i_{m+1}} S_{m+1,i_{m+1}}^*)^{p_0} + \dots + (S_{K,i_K} S_{K,i_K}^*)^{p_0}]\| &\leq A 2^{-\delta_0 m}, \\ \sup_{i_m, \dots, i_K \in I} \|S_{m,i_m} [(S_{m+1,i_{m+1}}^* S_{m+1,i_{m+1}})^{p_0} + \dots + (S_{K,i_K}^* S_{K,i_K})^{p_0}]\| &\leq A 2^{-\delta_0 m}. \end{aligned} \quad (6.21)$$

Then

$$\|S_1 + \dots + S_K\| \leq C(\delta_0, A, p_0).$$

Proof of Lemma 6.2. We apply Lemma 6.1 to the operators $S_m S_m^*$ and $S_m^* S_m$. It follows that there are constants $\bar{A} \geq 1$ and $\bar{\delta} > 0$ depending only on δ_0, A, P_0 such that

$$D_1 + \tilde{D}_1 \leq \bar{A}, \quad \mu_{m,1} + \tilde{\mu}_{m,1} \leq \bar{A} 2^{-\bar{\delta} m}, \quad m = 1, \dots, K. \quad (6.22)$$

For any $m = 1, \dots, K-1$ let

$$\begin{aligned} \nu_m &= \sup_{i_m, \dots, i_K \in I} \|S_{m,i_m}^* [(S_{m+1,i_{m+1}} S_{m+1,i_{m+1}}^*) + \dots + (S_{K,i_K} S_{K,i_K}^*)]\|, \\ \tilde{\nu}_m &= \sup_{i_m, \dots, i_K \in I} \|S_{m,i_m} [(S_{m+1,i_{m+1}}^* S_{m+1,i_{m+1}}) + \dots + (S_{K,i_K}^* S_{K,i_K})]\|. \end{aligned}$$

Clearly, for any $m = 1, \dots, K-1$

$$\nu_m^2 \leq D_1 \mu_{m,1}, \quad \tilde{\nu}_m^2 \leq \tilde{D}_1 \tilde{\mu}_{m,1}.$$

Therefore, using (6.22),

$$\nu_m + \tilde{\nu}_m \leq 2\bar{A} 2^{-\bar{\delta} m/2}, \quad m = 1, \dots, K. \quad (6.23)$$

Clearly

$$\|S_1 + \dots + S_K\|^2 \leq \|S_1 S_1^* + \dots + S_K S_K^*\| + 2 \sum_{m=1}^{K-1} \|S_m (S_{m+1}^* + \dots + S_K^*)\|.$$

Since $D_1 \leq \bar{A}$, for (6.20) it suffices to prove that

$$\|S_m (S_{m+1}^* + \dots + S_K^*)\| \leq A' 2^{-\bar{\delta} m/8}, \quad m = 1, \dots, K-1. \quad (6.24)$$

Let

$$\begin{aligned} Q &= \sup_{m=1, \dots, K-1} \sup_{i_m, \dots, i_K \in I} 2^{\bar{\delta} m/8} \|S_{m,i_m} (S_{m+1,i_{m+1}}^* + \dots + S_{K,i_K}^*)\|, \\ \tilde{Q} &= \sup_{m=1, \dots, K-1} \sup_{i_m, \dots, i_K \in I} 2^{\bar{\delta} m/8} \|S_{m,i_m}^* (S_{m+1,i_{m+1}} + \dots + S_{K,i_K})\|. \end{aligned}$$

Fix $m, i_m, \dots, i_K$ such that the supremum in the definition of Q is attained. Then

$$Q \leq 2^{\bar{\delta}m/8} \sum_{m'=m+1}^{8m} \|S_{m,i_m} S_{m',i_{m'}}^*\| + 2^{\bar{\delta}m/8} \|S_{m,i_m} (S_{8m+1,i_{8m+1}}^* + \dots + S_{K,i_K}^*)\|. \quad (6.25)$$

For any $m' \in [m+1, 8m] \cap \mathbb{Z}$ we have, using (6.23),

$$\|S_{m,i_m} S_{m',i_{m'}}^*\| \leq \|S_{m,i_m} S_{m',i_{m'}}^* S_{m',i_{m'}}\|^{1/2} \leq \tilde{\nu}_m^{1/2} \leq 2\bar{A}2^{-\bar{\delta}m/4}.$$

Using $\|S_m\| \leq 1$ and the definitions, it follows that

$$\begin{aligned} & \|S_{m,i_m} (S_{8m+1,i_{8m+1}}^* + \dots + S_{K,i_K}^*)\|^2 \\ & \leq \|S_{m,i_m} (S_{8m+1,i_{8m+1}}^* + \dots + S_{K,i_K}^*) (S_{8m+1,i_{8m+1}} + \dots + S_{K,i_K})\| \\ & \leq \tilde{\nu}_m + 2 \sum_{m''=8m+1}^K \|S_{m'',i_{m''}}^* (S_{m''+1,i_{m''+1}} + \dots + S_{K,i_K})\| \\ & \leq \tilde{\nu}_m + 2 \sum_{m''=8m+1}^K \tilde{Q} 2^{-\bar{\delta}m''/8}. \end{aligned}$$

Therefore, using (6.23) and (6.25),

$$Q \leq C(\bar{\delta}, \bar{A})(1 + \tilde{Q}^{1/2}).$$

A similar argument shows that

$$\tilde{Q} \leq C(\bar{\delta}, \bar{A})(1 + Q^{1/2}),$$

and the desired bound (6.24) follows. $\square$

REFERENCES

- [1] G. I. Arkhipov and K. I. Oskolkov, On a special trigonometric series and its applications, *Math. USSR-Sb.* **62** (1989), 145–155.
- [2] V. Bergelson and A. Leibman, A nilpotent Roth theorem, *Invent. Math.* **147** (2002), 429–470.
- [3] B. J. Birch, Forms in many variables, *Proc. R. Soc. Lond. A* **265** (1962), 245–263.
- [4] J. Bourgain, On the maximal ergodic theorem for certain subsets of the integers, *Israel J. Math.* **61** (1988), 39–72.
- [5] J. Bourgain, On the pointwise ergodic theorem on L^p for arithmetic sets, *Israel J. Math.* **61** (1988), 73–84.
- [6] J. Bourgain, Pointwise ergodic theorems for arithmetic sets, with an appendix by the author, H. Furstenberg, Y. Katznelson and D. S. Ornstein, *Inst. Hautes Études Sci. Publ. Math.* **69** (1989), 5–45.
- [7] L. J. Corwin and F. P. Greenleaf, *Representations of nilpotent Lie groups and their applications. Part I. Basic theory and examples*, Cambridge Studies in Advanced Mathematics **18**, Cambridge University Press, Cambridge (1990).
- [8] M. Christ, Hilbert Transforms Along Curves: I. Nilpotent Groups, *Ann. Math.* **122/3** (1985), 575–596.
- [9] M. Christ, A. Nagel, E. M. Stein, and S. Wainger, Singular and maximal Radon transforms: analysis and geometry, *Ann. Math.* **150** (1999), 489–577.

- [10] H. Davenport, Cubic forms in thirty-two variables, *Phil. Trans. R. Soc. Lond. A* **251** (1959), 193–232.
- [11] J. Garcia-Cuerva and J. L. Rubio de Francia, *Weighted norm inequalities and related topics*, North-Holland Mathematics Studies **116**, North-Holland Publishing Co., Amsterdam (1985).
- [12] A. D. Ionescu, A. Magyar, E. M. Stein, and S. Wainger, Discrete Radon transforms and applications to ergodic theory, *Acta Math.* **198** (2007), 231–298.
- [13] A. D. Ionescu and S. Wainger, L^p boundedness of discrete singular Radon transforms, *J. Amer. Math. Soc.* **19** (2006), 357–383.
- [14] A. Leibman, Convergence of multiple ergodic averages along polynomials of several variables, *Israel J. Math.* **146** (2005), 303–315.
- [15] A. I. Malcev, On a class of homogeneous spaces, *Izvestia Acad. Nauk SSSR Ser. Math.* **13** (1949), 9–32.
- [16] A. Magyar, E. M. Stein, and S. Wainger, Maximal operators associated to discrete subgroups of nilpotent Lie groups, *J. Anal. Math.* **101** (2007), 257–312.
- [17] D. Oberlin, Two discrete fractional integrals, *Math. Res. Lett.* **8** (2001), 1–6.
- [18] L. Pierce, A note on twisted discrete singular Radon transforms, *Math. Res. Lett.* **17** (2010), 701–720.
- [19] L. Pierce, Discrete fractional Radon transforms and quadratic forms, *Duke Math. J.* (to appear).
- [20] F. Ricci and E. M. Stein, Harmonic analysis on nilpotent groups and singular integrals I. Oscillatory integrals, *J. Funct. Anal.* **73** (1987), 179–194.
- [21] F. Ricci and E. M. Stein, Harmonic analysis on nilpotent groups and singular integrals II. Singular kernels supported on manifolds, *J. Funct. Anal.* **78** (1988), 56–84.
- [22] J. L. Rubio de Francia, A Littlewood–Paley inequality for arbitrary intervals, *Revista Matemática Iberoamericana* **1** (1985), 1–14.
- [23] E. M. Stein, *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*, Princeton University Press, Princeton (1993).
- [24] E. M. Stein and S. Wainger, Discrete analogues of singular Radon transforms, *Bull. Amer. Math. Soc.* **23** (1990), 537–544.
- [25] E. M. Stein and S. Wainger, Discrete analogues in harmonic analysis I: ℓ^2 estimates for singular Radon transforms, *Amer. J. Math.* **121** (1999), 1291–1336.
- [26] E. M. Stein and S. Wainger, Two discrete fractional integral operators revisited, *J. Analyse Math.* **87** (2002), 451–479.

PRINCETON UNIVERSITY

E-mail address: aionescu@math.princeton.edu

UNIVERSITY OF BRITISH COLUMBIA

E-mail address: magyar@math.ubc.ca

UNIVERSITY OF WISCONSIN–MADISON

E-mail address: wainger@math.wisc.edu