MÖBIUS DISJOINTNESS FOR NILSEQUENCES ALONG SHORT INTERVALS

XIAOGUANG HE AND ZHIREN WANG

Abstract. For a nilmanifold G/Γ, a 1-Lipschitz continuous function F and the Möbius sequence $\mu(n)$, we prove a bound on the decay of the averaged short interval correlation

$$\frac{1}{HN}\sum_{n\leq N}\Big|\sum_{h\leq H}\mu(n+h)F(g^{n+h}x)\Big|$$

as $H, N\to\infty$. The bound is uniform in $g\in G$, $x\in G/\Gamma$ and F.

Contents

1. Introduction

The Möbius function $\mu:\mathbb{N}\to\{-1,0,1\}$ is defined as follows: $\mu(1)=1$, $\mu(n)=(-1)^k$ when n is the product of k distinct primes and $\mu(n)=0$ otherwise. This is an important function since that $\sum\limits_{n\leq N}\mu(n)=o(N)$ is equivalent to prime number theorem, and that $\sum\limits_{n\leq N}\mu(n)=O_\varepsilon(N^{\frac{1}{2}+\varepsilon})$ for all $\varepsilon>0$ is equivalent to the Riemann Hypothesis.

The Möbius Randomness Law, proposed in [IK04], suggests that reasonable sequences $\xi(n)$ which have significant cancellations with $\mu(n)$, that is

$$\sum_{n\leq N}\mu(n)\xi(n)=o(\sum_{n\leq N}|\xi(n)|).$$

The Möbius Disjointness Conjecture, of Sarnak [Sar09], expects to use observables from zero entropy topological dynamical systems as the sequence ξ.

Conjecture 1.1. *(Möbius Disjointness Conjecture, [Sar09]) Let (X,T) be a topological dynamical system with zero topological entropy. Then*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} f(T^n x)\mu(n) = 0, \forall f\in C(X), \forall x\in X. \tag{1.1}$$

Here, a topological dynamical system is a pair (X,T) consisting of a compact metric space X, and a continuous self-map $T: X\to X$.

There have been many partial results on the Conjecture 1.1. For brevity we will simply refer to the recent comprehensive survey [FKPL18] for the progress in this area, and discuss only the historical developements that are more relevant to this paper.

The special case of Conjecture 1.1 for circle rotations, has been known since 1937 due to Davenport's work [Dav37]. Indeed, Davenport proved in [Dav37] that for all $A>0$,

$$\sup_{\alpha\in\mathbb{R}}\Big|\frac{1}{N}\sum_{n\leq N} e(\alpha n)\Gamma)\Big| \ll_A \log^{-A} N. \tag{1.2}$$

Here $e(u)=e^{2\pi i u}$.

An important extension to this class is the nilsystems, namely tranlations $x\to g.x$ on a compact nilmanifold G/Γ. Such systems are particularly important because of their close relationship to multiple ergodic averages. Functions of the form $n\to f(g^n.x)$ cover all the polynomial and bracket polynomial phases. It was known, as a special case of Ratner's Theorem [Rat91] and its discrete version by Shah [Sha], that every trajectory of such a translation always equidistributes to the union of finitely many translated copies of a closed sub-nilmanifold. This property was extended by Leibman [Lei05] to polynomial orbits in nilmanifolds (see Definition 2.11 for the definition).

Möbius disjointness along orbits of nilsystems, or more generally polynomial orbits, was established by Green and Tao [GT12b] in the following form:

$$\sup_{g,F}\Big|\frac{1}{N}\sum_{n\leq N}\mu(n)F(g(n)\Gamma)\Big| \ll_{m,A} R^{-O_{m,A}(1)}\log^{-A} N, \tag{1.3}$$

where the supremum is taken over all polynomial functions $g:\mathbb{Z}\to G$ with respect to a given nilpotent filtration $G_\bullet$ and all functions $F: G/\Gamma\to\mathbb{C}$ that are 1-Lipschitz. Here $m=\dim G$, and the parameter R records the rationality of the pair $(G_\bullet,\Gamma)$ (see Section 2 for related definitions).

Green-Tao's proof was based on their accompaying paper [GT12a], which effectivized Leibman's Theorem by describing quantitatively how fast a trajectory equidistributes to a subnilmanifold of G/Γ. This was then applied to joinings of two orbits of the forms $\{g(pn)\Gamma\}$ and $\{g(qn)\Gamma\}$. Combined with Vaughan's Identity [Vau97], which is a modern form of the Vinogradov bilinear method, such estimates lead to the orthogonality to the Möbius function.

Another strengthening to Davenport's estimate (1.2) was achieved in the recent breakthrough papers of Matomäki-Radziwiłł [MR16] and Matomäki-Radziwiłł-Tao [MRT15] on averages of non-pretentious multiplicative functions along short intervals. As a consequence, they proved in [MRT15] that for all real-valued 1-bounded multiplicative functions, which in particular include the Möbius and Liouville functions,

$$\sup_{\alpha\in\mathbb{R}} \sum_{n\le N} \left| \sum_{h\le H} \mu(n+h)e(\alpha(n+h)) \right| \, dx \ll \left(\frac{\log\log H}{\log H} + \frac{1}{\log^{1/700} N} \right) HN. \quad (1.4)$$

Such estimates were used to prove an averaged form of the Chowla Conjecture in [MRT15], as well as the logarithmically averaged Chowla and Elliott Conjectures for correlations with either 2 or an odd number of components by Tao [Tao16] and Tao-Teräväinen [TT16]. The theorems in [MR16] and [MRT15] have also yielded many applications to Conjecture 1.1, especially to dynamical systems with strong quasi-periodic behavior (see the survey [FKPL18]). They were also used in Frantzikinakis-Host's proof [FH18] of logarithmically averaged Sarnak Conjecture for ergodic weights. For most of these applications, it is essential to have a uniform decay rate in (1.4) that is independent of the choice of α.

It is natural to seek a further strengthening to (1.2) that combines the theorems of Green-Tao (1.3) and Matomäki-Radziwiłł-Tao (1.4), namely a quantitative bound to Möbius disjointness along short intervals for nilsequences. This is the purpose of the current paper. This question is especially interesting because, as remarked in [Tao16, p34], short interval correlations between multiplicative functions and higher step nilsequences would be useful in the study of logarithmicall averaged Chowla and Elliott conjectures of higher order correlations.

Previously in this direction, Flaminio, Frączek, Kułaga-Przymus, and Lemańczyk [FFKPL19] proved that: if φ is an ergodic unipotent affine automorphism of a compact nilmanifold G/Γ and $x \in G/\Gamma$, $F \in C^0(G/\Gamma)$, then:

$$\frac{1}{N} \sum_{N\le n<2N} \left| \frac{1}{H} \sum_{h\le H} \mu(n+h)F(\varphi^{n+h}(x)) \right| \to 0 \quad (1.5)$$

as $H \to \infty$ and $N/H \to \infty$. Similar results were also shown for polynomial phases by El Abdalaoui-Lema'nczyk-de la Rue in [EALdlR17]. The proofs purely relies on a minor arc argument and uses the bilinear method in the form of the Kátai-Bourgain-Sarnak-Ziegler criterion [Kát86, BSZ13]. The decay estimates in [FFKPL19] and [EALdlR17] are not effective as the dynamics becomes highly quasi-periodic.

The result in this paper produces a uniformly effective bound without requiring ergodicity.

It should also be noted that without the extra average in N, non-trivial bounds on $\left| \frac{1}{H} \sum_{h\le H} \mu(n+h)f(n+h) \right|$ were obtained in the works of Zhan [Zha91], Huang [Hua15, Hua16] and Matomäki-Shao [MS19] when f is a polynomial phase and $H \gg n^\theta$ for some given $\theta \in (0,1)$. ($\theta = \frac{2}{3}$ in [MS19]).

Our main theorem is:

Theorem 1.2. *Suppose G is a connected, simply connected m-dimensional nilpotent Lie group and $\Gamma \subset G$ is a lattice. Then there exists $H_0 = H_0(G,\Gamma) > 0$ and $\epsilon_0 = \epsilon_0(m) > 0$, such that:*

For all $H, N \in \mathbb{N}$ satisfying $H > H_0$ and $(\log N)^{\frac{1}{2}} > \log H$, and $\epsilon \in (\frac{\log\log H}{\log H}, \epsilon_0)$, there exists a set $\mathcal{S} \in [N]$, whose construction depends only on H, N and ϵ, such that

$$N - \#\mathcal{S} \ll_m \epsilon N, \quad (1.6)$$

and

$$\sup_{\substack{\|F\|_{G/\Gamma}\le 1\\ g\in G, x\in G/\Gamma}} \frac{1}{HN}\sum_{n\le N}\Big|\sum_{h\le H} 1_{\mathcal{S}}\mu(n+h)F(g^{n+h}x)\Big| \ll_m H^{-\epsilon}+\delta(H^{\epsilon},N). \tag{1.7}$$

Here, the implied constants depend only on m. $\|F\|_{G/\Gamma}$ stands for the Lipschitz norm of a function F on G/Γ. The construction of the error function $\delta(\cdot,\cdot)>0$ is independent of all the parameters here, and it satisfies $\lim_{N\to\infty}\delta(a,N)=0$ for all $a>0$.

In particular,

$$\sup_{\substack{\|F\|_{G/\Gamma}\le 1\\ g\in G, x\in G/\Gamma}} \frac{1}{HN}\sum_{n\le N}|\Big|\sum_{h\le H} \mu(n+h)F(g^{n+h}x)\Big| \ll_m \epsilon+H^{-\epsilon}+\delta(H^{\epsilon},N). \tag{1.8}$$

The Lipschitz norm of F needs to be defined using a particular Mal'cev basis of the Lie algebra of G that is compatible with Γ. For details, see (2.2).

By taking $\epsilon=\frac{\log\log H}{\log H}$, the following corollary immediately follows:

Corollary 1.3. *Suppose G is a connected, simply connected m-dimensional nilpotent Lie group and $\Gamma\subset G$ is a lattice. Then there exists $H_0=H_0(G,\Gamma)>0$, such that:*

For all $H,N\in\mathbb{N}$ with $H>H_0$ and $(\log N)^{\frac{1}{2}}>\log H$,

$$\sup_{\substack{\|F\|_{G/\Gamma}\le 1\\ g\in G, x\in G/\Gamma}} \frac{1}{HN}\sum_{n\le N}\Big|\sum_{h\le H} \mu(n+h)F(g^{n+h}x)\Big| \ll_m \frac{\log\log H}{\log H}+\delta(\log H,N), \tag{1.9}$$

where the implied constant and the construction of the error function $\delta(\cdot,\cdot)>0$ depends only on m, and δ satisfies $\lim_{N\to\infty}\delta(a,N)=0$ for all $a>0$.

In particular, in the settings of Corollary 1.3,

$$\lim_{H\to\infty}\frac{1}{H}\limsup_{N\to\infty}\frac{1}{N}\sum_{n\le N}\Big|\sum_{h\in H}\mu(n+h)F(g^{n+h}x)\Big|=0, \tag{1.10}$$

uniformly for all $g\in G$, $x\in X$ and functions $F: G/\Gamma\to\mathbb{C}$ from a given uniformly Lipschitz family.

Remark 1.4. Theorem 1.2 and Corollary 1.3 also hold if μ is replaced by the Liouville function λ, or any multiplicative function β that is non-pretentious in the sense $M(\beta\chi,X)\to\infty$ as $X\to\infty$ for all Dirichilet characters χ. For the definition of the quantitiy $M(\cdot,X)$, see Definition 5.1. A more precise version of Theorem 1.2, specifying how $\delta(H^{\epsilon},N)$ depends on the functions $M(\beta\chi,\cdot)$, will be given in Theorem 8.1.

Remark 1.5. Theorem 8.1, and thus Theorem 1.2 and Corollary 1.3, is actually valid for all polynomial sequnces $\{g(n,h)\Gamma\}$ in G/Γ in lieu of $\{g^{n+h}x\}$. This in particular covers orbits of unipotent affine automorphisms as in [FFKPL19].

We now outline the organization of the paper. The strategy in our proof mixes those from [GT12b] and [MRT15]. The main issue is that, while it is known by [GT12a] that when H is sufficiently large, each individual short range orbit $\{g^{n+h}x\}_{1\le h\le H}$ in G/Γ should equidistribute well in a subnilmanifold Y_n, in order to apply the bilinear method, it is necessary to know that the equidstribution behaviors display a similar pattern in Y_n and $Y_{n'}$ when $pn \approx p'n'$ for a pair of bouned prime numbers p', q'. It is for this reason that we choose to view $g(n+h)$, where g is a polynomial in one variable, as a polynomial $g(n,h)$ in two variables n and h. After introducing the background notions in Section 2, in Section 3 we derive a variation of Green-Tao's quantitative version of Leiman's Theorem that better adapts to our situation. Namely, we show that when N and H are both sufficiently large, $\{g(n,h)\Gamma\}_{1\leq h\leq H}$ is equidistributed in some Y_n for a typical $n \leq N$, and the equidistrbution patterns in all such Y_n's are correlated to each other. Section 4 sets up the bilinear methods scheme and separates the estimate into minor and major arcs along each short interval. In the major arc part (Section 5), the Matomäki-Radziwiłł-Tao estimate can be applied as the correspondence $n \to Y_n$ is periodic. In the minor arc part (Section 6), we use Lemma 6.2 to replace the bilinear sum in [MRT15], which becomes a sum of 4-fold products after applying Cauchy-Schwarz and would get too complicated for nilsequences, with one that consists of 2-fold products recording the correlations between short orbits of the form $\{g(n,p(h+r))\}$ and $\{g(n',p'(h+r'))\}$ where $pn \approx p'n'$. The bound of such correlations, for all but a small portion of choices of (n,n',p,p'), will be given by Proposition 6.9 and proved in Section 7 using the aforementionned correlation among equidistribution patterns. Finally, Section 8 merges the minor and major arcs and fixes appropriate parameters to conclude the proof.

Notation 1.6. *In this paper:*

- *$X = O_Y(Z)$ or $X \ll_Y Z$ means that $\frac{X}{Z}$ is bounded by a constant that depends only on Y.*
- *Working under Hypothesis 2.13, we shall assume by default that the implicit constant Y depend on the degree d of the filtration and the dimension m of the nilmanifold, without including m, d in the subscript. For example, $O_A(1)$ will actually stand for $O_{A,m,d}(1)$.*
- *$[N]$ stands for the interval of integers $\{1,\cdots,N\}$.*
- *In the remainder of this paper, many implicit constants $O(1) = O_{m,d}(1)$ will appear. For simplicity, we will use a common constant $C_0 = O_{m,d}(1) \geq 1$ that is large enough for all these purposes. Similarly, from now on the notation $\ll$ will always stand for $\ll_{m,\epsilon}$.*
- *For $\alpha \in \mathbb{R}$, $\|\alpha\|_{\mathbb{R}/\mathbb{Z}}$ denotes $\max_{k\in\mathbb{Z}} |\alpha - k|$.*

Acknowledgments. A large part of this research was done while X.H. was visiting Pennsylvania State University during the 2017-2018 academic year. X.H. thanks the financial support (No. 201706220146) from China Scholarship Council and the hospitiality of Pennsylvania State University that made the visit possible. Z.W. was supported by NSF grants DMS-1501095 and DMS-1753042.

We thank Wen Huang for helpful comments.

2. Background on sequences in nilmanifolds

In this section, we quickly collect all the facts and notions that we will need from Green-Tao's papers [GT12a, §1, §2 & §A] and [GT12b, §3].

A connected, simply connected Lie group G is **nilpotent** if it has a nilpotent **filtration** $G_\bullet$, i.e. a descending sequnce of groups $G = G_1 \supseteq G_2 \supseteq \cdots \supseteq G_d \supseteq G_{d+1} = \{e\}$ such that

$$[G, G_{i-1}] \subseteq G_i, \forall i \geq 2. \tag{2.1}$$

This actually implies $[G_i, G_j] \subseteq G_{i+j}$ for all $i, j \geq 1$. The number d is the **degree** of the filtration $G_\bullet$. The **step** s of G is the degree of the lower central filtration defined by $G_{i+1} = [G, G_i]$.

For all $i \geq d+1$, we will adopt the convention that $G_i = \{e\}$.

Denote by $\mathfrak{g}_i$ the Lie algebra G_i, then $\mathfrak{g}_\bullet = \{\mathfrak{g}_i\}$ is a **filtration of Lie algebras**, i.e. $[\mathfrak{g}, \mathfrak{g}_i] \subseteq \mathfrak{g}_{i+1}$, if and only if G_i is a filtration.

A connected, simply connected nilpotent Lie group G has a lattice Γ if and only if it has an algebraic structure defined over $\mathbb{Q}$. In this case, for a connected Lie subgroup H of G, H is an algebraic subgroup defined over $\mathbb{Q}$ if and only if $H \cap \Gamma$ is a lattice of H. A lattice Γ must be cocompact, and the compact quotient G/Γ is called a **nilmanifold**.

A basis $\mathcal{V} = \{V_1, \cdots, V_m\}$ of $\mathfrak{g}$ is R**-rational** if the structure constants c_{ijk} in the Lie bracket relations $[V_i, V_j] = \sum_k c_{ijk} V_k$ are rational numbers whose heights are bounded by R. Recall that the height of a rational number $\frac{a}{b}$ is $\max(|a|, |b|)$ when a, b are coprime. For nilmanifolds G/Γ, G always has a rational basis. A special kind of rational basis, **Mal'cev basis**, was defined in [Mal49]. A rational basis $\mathcal{V} = \{V_1, \cdots, V_m\}$ is a Mal'cev basis adapted to $(G_\bullet, \Gamma)$ if it satisfies the following properties in [GT12a, Def. 2.1]:

(i) $\{V_j, V_{j+1}, \cdots, V_m\}$ spans an ideal of $\mathfrak{g}$ for all $0 \leq j \leq m$;
(ii) For each $1 \leq i \leq d$ and $m_i = \dim G_i$, the Lie algebra $\mathfrak{g}_i$ of G_i is the linear span of $\{V_{m-m_i+1}, V_{m-m_i+2}, \cdots, V_m\}$;
(iii) There is a diffeomorphism $\psi_\mathcal{V} : G \to \mathbb{R}^m$ determined by
$$\psi_\mathcal{V}\Big(\exp(\omega_1 V_1) \cdots \exp(\omega_m V_m)\Big) = (\omega_1, \cdots, \omega_m);$$
(iv) In the coordinate system $\psi_\mathcal{V}$, $\Gamma = \psi_\mathcal{V}^{-1}(\mathbb{Z}^m)$.

When G has a lattice Γ, there is always a Mal'cev basis adapted to the lower central filtration. In the coordinate system given by $\psi_\mathcal{V}$, the set $\psi_\mathcal{V}^{-1}([0,1)^m)$ will be a fundamental domain of the projection $G \to G/\Gamma$.

In the sequel, we will always assume that G/Γ has a Mal'cev basis $\mathcal{V}$ adapted to $(G_\bullet, \Gamma)$ for some filtration $G_\bullet$, and fix the tuplet $(G, G_\bullet, \Gamma, \mathcal{V})$. In this case, every G_i is a rational subgroup of G, and $\Gamma_i = G_i \cap \Gamma$ is a lattice of G_i.

The nilmanifold G/Γ has a tower structure of principal torus bundles
$$G/\Gamma = G/G_{d+1}\Gamma \to G/G_d\Gamma \to \cdots \to G/G_2\Gamma \to G/G_1\Gamma = \{\mathrm{pt}\},$$
where $G/G_{i+1}\Gamma$ is a principal $G_i/G_{i+1}\Gamma$-bundle over $G/G_i\Gamma$. Remark that here $G_i/G_{i+1}\Gamma \cong \mathbb{T}^{m_i - m_{i+1}}$ is the quotient of the abelian Lie group $G_i/G_{i+1} \cong \mathbb{R}^{m_i - m_{i+1}}$ by the lattice generated by the projections of $V_{m-m_i+1}, \cdots, V_{m-m_{i+1}}$.

A vector $v \in \mathfrak{g}$ is **an** R**-rational combination** of elements in $\mathcal{V}$ if $v = \sum v_j V_j$ where the v_j's are rational numbers of height bounded by R. A subgroup $H \subseteq G$ is R**-rational** with respect to $\mathcal{V}$ if its Lie algebra has a basis consisting of such R-rational combinations.

The Mal'cev basis $\mathcal{V}$ induces a right invariant metric d_G on G, which is the largest metric such that $\mathrm{d}(x, y) \leq |\psi_\mathcal{V}(xy^{-1})|$ always holds. Actually, this in turn induces a metric $\mathrm{d}_{G/\Gamma}$ on G/Γ. For functions $F : G/\Gamma \to \mathbb{C}$, $\|F\|$ will denote the Lipschitz norm
$$\|F\| = \|F\|_{C^0} + \sup_{n \neq y} \frac{|F(x) - F(y)|}{\mathrm{d}_{G/\Gamma}(x, y)} \tag{2.2}$$

with respect to $\mathrm{d}_{G/\Gamma}$. We will also write $\|F\|_{G/\Gamma}$ instead, when it becomes necessary to emphasize that the distance is determined by the Mal'cev basis of G/Γ.

The nilpotent Lie group G is unimodular, and G/Γ has a unique left-invariant probability measure. The notation $\int_{G/\Gamma}$ will refer to the average with respect to this measure.

Since $G/[G,G]$ is abelian and the commutator subgroup $[G,G]$ is a rational subgroup, $(G/\Gamma)/([G,G]/([G,G]\cap\Gamma)) = G/[G,G]\Gamma$ is a quotient torus of the connected abelian Lie group $G/[G,G] \cong \mathbb{R}$, called the **horizontal torus with respect to** $G_\bullet$ of G/Γ.

Definition 2.1. [GT12a, Definition 2.6] *A* **horizontal character** *is a continuous additive homomorphism $\eta : G/[G,G]\Gamma \to \mathbb{R}/\mathbb{Z}$. We remark that η can also be viewed as a continuous group homomorphism $\eta : G \to \mathbb{R}/\mathbb{Z}$ that vanishes on the subgroup $[G,G]\Gamma$.*

Using the coordinate representation $\psi_{\mathcal{V}}$, there exists an integer vector $a \in \mathbb{Z}^m$, supported on the first $m - m_2$ coordinates, such that

$$\eta(g) = a \cdot \psi_{\mathcal{V}}(g)(\mathrm{mod}\ \mathbb{Z}). \tag{2.3}$$

The **modulus** $|\eta|$ of η is defined to be $|a|$. Note η is trivial if and only if $|\eta| = 0$. By abusing notation, we shall also denote by η the linear functional $\eta(v) = a \cdot v$ on $\mathbb{R}^m \cong \mathfrak{g}$.

Definition 2.2. *For a polynomial function $f : [N] \to \mathbb{R}/\mathbb{Z}$ of degree at most d, f can be written as $f(n) = \sum_{i=0}^{d} \alpha_i \binom{n}{i}$. The* $C^\infty([N])$**-norm** *of f is given by*

$$\|f\|_{C^\infty([N])} = \max_{i=0}^{d} N^i \|\alpha_i\|_{\mathbb{R}/\mathbb{Z}}.$$

Lemma 2.3. [GT12b, Lemma 3.2] *If $f(n) = \sum_{i=0}^{d} \beta_i n^i$, then there is an integer $D = O_d(1)$ such that $\|D\beta_i\|_{\mathbb{R}/\mathbb{Z}} \ll_d N^{-i}\|f\|_{C^\infty[N]}$ for all $i = 0, \cdots, d$.*

Lemma 2.4. [GT12a, Lemma 4.5] *Suppose $f(n) = \sum_{i=0}^{d} \beta_i n^i$, $\delta \in (0, \frac{1}{2})$, $\epsilon \in (0, \frac{\delta}{2})$. If $f(n)(\mathrm{mod}\ \mathbb{Z})$ belongs to an interval $I \subseteq \mathbb{R}/\mathbb{Z}$ of length ϵ for at least δN integers $n \in [N]$. Then for some positive integer $D \ll_d \delta^{-O_d(1)}$, $\|Df(\mathrm{mod}\ \mathbb{Z})\|_{C^\infty[N]} \ll_d \epsilon\delta^{-O_d(1)}$.*

For an integer vector $\mathbf{N} \in \mathbb{N}^r$, write $[\mathbf{N}] = [N_1] \times \cdots \times [N_r] \subset \mathbb{Z}^r$.

Definition 2.5. [GT12a, Definition 9.1] *For a multiparatmeter finite sequence $\{g(\mathbf{n})\}_{\mathbf{n}\in[\mathbf{N}]}$ in G and an integer vector $\mathbf{N} \in \mathbb{N}^r$, g is said to be* $(W, \mathbf{N})$**-smooth***, if for all $\mathbf{n} \in [\mathbf{N}]$,*

(1) $\mathrm{d}_G(g(\mathbf{n}), \mathrm{id}_G) \le W$,
(2) $\mathrm{d}_G(g(\mathbf{n}), g(\mathbf{n}+\mathbf{e}_i)) \le \frac{W}{N_i}$ *for all i, where $\mathbf{e}_i$ is the unit vector along the i-th coordinate direction.*

If g_1, g_2 are both $(W, \mathbf{N})$-smooth, and $W \ge R$, where the metric is induced by an R-rational Mal'cev basis, then $g_1 g_2$ is $(W^{O(1)}, \mathbf{N})$ smooth.

Definition 2.6. *An element $g \in G$ is R*-**rational***, if there exists $1 \le r \le R$ such that $g^r \in \Gamma$. An element $z \in G/\Gamma$ is R*-**rational***, if $z = g\Gamma$ for some R-rational group element g.*

Lemma 2.7. [GT12a, Lemma A.11] *Suppose the Mal'cev basis $\mathcal{V}$ adapted to $(G_\bullet, \Gamma)$ is R-rational. With respect to $\mathcal{V}$, if g is R-rational then $\psi_{\mathcal{V}}(g) \in \frac{1}{q}\mathbb{Z}^m$ for some $q \ll R^{O(1)}$. Conversely, if $\psi_{\mathcal{V}}(g) \in \frac{1}{R}\mathbb{Z}^m$ then g is $R^{O(1)}$-rational. Moreover, the product of two R-rational elements is $R^{O(1)}$-rational.*

Definition 2.8. *For a finite arithmetic progression $\mathcal{A} = \{qn + r\}_{n\in[N]}$ in $\mathbb{Z}$, a finite sequence $\{x(n)\}_{n\in\mathcal{A}}$ in G/γ is said to be δ-**equidistributed** in G/Γ if for all Lipschitz function F on G/Γ,*

$$\left| \mathop{\mathbb{E}}_{n\in\mathcal{A}} F(x(n)) - \int_{G/\Gamma} F \right| \le \delta\|F\|;$$

and it is **totally δ-equidistributed** *in G/Γ if the subsequence $\{x(n)\}_{n\in\mathcal{A}'}$ is δ-equidistributed in G/Γ for all arithmetic progressions $\mathcal{A}' \subseteq \mathcal{A}$ of length at least δN.*

Lemma 2.9. *Suppose a Mal'cev basis $\mathcal{V}$ adapted to $(G_\bullet, \Gamma)$ is R-rational where $R \ge 10$. Let η be a non-trivial horizontal character of G/Γ, whose modulus $|\eta|$ is bounded by R with respect to $\mathcal{V}$. If for a polynomial sequence $g \in \mathrm{Poly}(\mathbb{Z}, G_\bullet)$ and $N \gg R$, $\|\eta \circ g\|_{C^\infty([N])} \le R$, then $\{g(n)\Gamma\}_{n\in[N]}$ is not totally $(O(R))^{-1}$-equidistributed.*

Proof. Because $\|\eta\circ g\|_{C^\infty([N])} \le R$, $\|\eta\circ g(n) - \eta\circ g(0)\|_{\mathbb{R}/\mathbb{Z}} \ll RnN^{-1}$. This implies that for the the mapping $\tilde\eta(x) = \exp(2\pi i\eta(x))$ from G/Γ to the unit circle in $\mathbb{C}$, the values of $\tilde\eta(g(n))$ are within distance $\ll R\delta$ to each other for $0 < n \le \delta N$. Using the convention in Notation 1.6, one can assume that the implicit constant here is C_0. In particular,

$$\left| \mathop{\mathbb{E}}_{0<n\le\delta N} \tilde\eta(g(n)\Gamma) \right| > 1 - C_0 R\delta \ge \frac{1}{2}, \tag{2.4}$$

if $\delta < \frac{1}{2}{C_0}^{-1}R^{-1}$. Because η is a non-zero character, $\tilde\eta$ has zero average on G/Γ. In addition, $\|\tilde\eta\|_{G/\Gamma} \le 2\pi|\eta| \le 2\pi R$. It follows that the sequence $\{g(h)\Gamma\}_{h\in[H]}$ is not totally $\min(\frac{1}{2}{C_0}^{-1}R^{-1}, \frac{1}{4\pi}R^{-1})$-equidistributed in G/Γ. ☐

Lemma 2.10. *If $\delta \in (0,1)$ and there exists an interval $\mathcal{A} \subseteq [N]$ of length at least δN such that $\{g(n)\}_{n\in\mathcal{A}}$ is not δ-equidistributed in G/Γ, then for some $N' \in [\frac{\delta^2}{2}N, N]$, $(g(n))_{n\in[N']}$ is not $\frac{\delta^2}{2}$-equidistributed in G/Γ.*

Proof. One may write $\mathcal{A} = \{N_1 < n \le N_2\} = [N_2] \setminus [N_1]$. Write $\theta_i = \frac{N_i}{N}$ and $\theta = \theta_2 - \theta_1$, then $\theta \ge \delta$.

There exists a Lipschitz function F on G'/Γ' with $\int_{G/\Gamma} F = 0$ such that

$$\left| \frac{\theta_2}{\theta} \mathop{\mathbb{E}}_{n\in[N_2]} F(g(n)\Gamma) - \frac{\theta_1}{\theta} \mathop{\mathbb{E}}_{n\in[N_1]} F(g(n)\Gamma) \right| = \left| \mathop{\mathbb{E}}_{n\in\mathcal{A}} F(g(n)\Gamma) \right| > \delta\|F\|.$$

If $\theta_1 \ge \frac{\delta^2}{2}$ and $\left|\mathbb{E}_{n\in[N_1]} F(g(n)\Gamma)\right| > \frac{\delta^2}{2}\|F\|$, then $N_1 \ge \frac{\delta^2}{2}N$ and $(g(n))_{n\in[N_1]}$ is not $\frac{\delta^2}{2}$-equidistributed.

Otherwise, either $\theta_1 < \frac{\delta^2}{2}$ or $\left|\mathbb{E}_{n\in[N_1]} F(g(n)\Gamma)\right| < \frac{\delta^2}{2}\|F\|$. In both cases,

$$\left| \frac{\theta_1}{\theta} \mathop{\mathbb{E}}_{n\in[N_1]} F(g(n)\Gamma) \right| < \frac{\delta^2}{2}\|F\|,$$

and thus

$$\left|\mathop{\mathbb{E}}_{n\in[N_2]} F(g(n)\Gamma)\right| \geq \left|\frac{\theta_2}{\theta}\mathop{\mathbb{E}}_{n\in[N_2]} F(g(n)\Gamma)\right| > \delta\|F\| - \frac{\delta^2}{2}\|F| \geq \frac{\delta}{2}\|F\|.$$

So $(g(n))_{n\in[N_2]}$ is not $\frac{\delta}{2}$-equidistributed. Moreover, $N_2 \geq \theta N \geq \delta N$. □

For a map $g : \mathbb{Z}^r \to G$, the derivative along $\mathbf{h} \in \mathbb{Z}^r$ is

$$\partial_{\mathbf{h}} g(\mathbf{n}) = g(\mathbf{n}+\mathbf{h})g(\mathbf{n})^{-1}. \tag{2.5}$$

Definition 2.11. *A map $g : \mathbb{Z}^r \to G$ is a* **polynomial map with respect to $G_\bullet$** *if for all i and $l_1, \cdots, l_i, n \in \mathbb{Z}$, the i-th derivative $\partial_{l_1}\cdots\partial_{l_i} g(n)$ takes values in G_i. The set of polynomial sequences with respect to $G_\bullet$ is noted by $\mathrm{Poly}(\mathbb{Z}^r, G_\bullet)$.*

The family of $\mathrm{Poly}(\mathbb{Z}^r, G_\bullet)$ is known to be a group (Lazard [Laz54], Leibman [Lei98, Lei02] and Green-Tao [GT12a]). A description of $\mathrm{Poly}(\mathbb{Z}^r, G_\bullet)$ was given in Leibman and Green-Tao's works:

Lemma 2.12. *([Lei10, §4],[GT12a, §6]) Suppose $\mathcal{V}$ is a Mal'cev basis adapted to $(G_\bullet, \Gamma)$, then $g \in \mathrm{Poly}(\mathbb{Z}^r, G)$ if and only if $\psi_{\mathcal{V}}(g(\mathbf{n}))$ has the form*

$$\psi_{\mathcal{V}}(g(\mathbf{n})) = \sum_{\mathbf{j}\in\mathbb{Z}^r_{\geq 0}} \omega_{\mathbf{j}} \binom{n_1}{j_1}\cdots\binom{n_r}{j_r},$$

where $\omega_{\mathbf{j}} \in \mathbb{R}^m$ and $(\omega_{\mathbf{j}})_i = 0$ for all $i \leq m - m_{|\mathbf{j}|}$ with $|\mathbf{j}| = j_1 + \cdots + j_r$.

In particular, if $|\mathbf{j}| > d$, then $m_{|\mathbf{j}|} = 0$ and thus $\omega_{\mathbf{j}} = 0$.

In the rest of this paper we will work under the following work hypothesis

Hypothesis 2.13. *G/Γ is an m-dimensional compact nilmanifold with a degree d rational filtration $G_\bullet$, and $\mathcal{V}$ is an R_0-rational Mal'cev basis adapted to $(G_\bullet, \Gamma)$, where $R_0 > 10$. Moreover, $g \in \mathrm{Poly}(\mathbb{Z}^2, G_\bullet)$ is a polynomial map determined by coefficients $\{\omega_{j,k}\}_{j,k\in\mathbb{Z}_{\geq 0}}$ as in Lemma 2.12. Let $R \geq R_0$ be a parameter to be determined later. In particular, $\mathcal{V}$ is also an R-rational Mal'cev basis adapted to $(G_\bullet, \Gamma)$.*

The formula in Lemma 2.12 writes in this case as:

$$\psi_{\mathcal{V}}(g(n,h)) = \sum_{\substack{j,k\geq 0\\ j+k\leq d}} \omega_{jk}\binom{n}{j}\binom{h}{k}, \tag{2.6}$$

where $(\omega_{jk})_i = 0$ for all $i \leq m - m_{j+k}$.

3. Quantitaive factorization theorem for 2-parameter polynomials

We now state Green-Tao's effectivization of a theorem of Leibman [Lei05], and deduce a variation of it that is refined to our situation.

Proposition 3.1. [GT12a, Theorem 2.9] *Suppose G/Γ is an m-dimensional compact nilmanifold with a degree d rational filtration $G_\bullet$, and $\mathcal{V}$ is an R-rational Mal'cev basis adapted to $(G_\bullet, \Gamma)$ where $R \geq 10$. For $f \in \mathrm{Poly}(\mathbb{Z}, G_\bullet)$, and $N \in \mathbb{N}$ such that $N \gg R^{O(1)}$, at least one of the following holds:*

1. *either $\{f(n)\Gamma\}_{n\in[N]}$ is R^{-1}-equidistributed in G/Γ;*
2. *or there exists a horizontal character η of G/Γ of modulus $|\eta| \leq R^{O(1)}$ such that $\|\eta \circ f\|_{C^\infty([N])} \leq R^{O(1)}$.*

Corollary 3.2. *In Proposition 3.1, one may replace in part (1) the property "R^{-1}-equidistributed" by "totally R^{-1}-equidistributed".*

Proof. Suppose $\{f(n)\Gamma\}_{n\in[N]}$ is not totally R^{-1}-equidistributed. There exist integers $0 \le a < b \le R$, and an interval $\mathcal{A} \subseteq [\frac{N}{b}]$ of length at least $R^{-1}N$, such that the sequence $\{\tilde{f}(n)\Gamma\}_{n\in\mathcal{A}}$ is not R^{-1}-equidistributed, where $\tilde{f}(n) = f(bn+a)$. By Lemma 2.10, there exists $N' < N$ with $N' \ge \frac{1}{2}R^{-2} \cdot \#\mathcal{A} \ge R^{-O(1)}N$ such that $\{\tilde{f}(n)\Gamma\}_{n\in[N']}$ is not R^{-1}-equidistributed. By Proposition 3.1, there exists a horizontal character η such that $0 < |\eta| < R^{O(1)}$ and $\|\eta \circ \tilde{f}\|_{C^\infty([N'])} \le R^{O(1)}$. As $N' \ge R^{-O(1)}N$, this implies that $\|\eta \circ \tilde{f}\|_{C^\infty([N])} \le R^{O(1)}$, which in turn implies by [GT12a, 7.10] that there is a positive integer $D \le R^{O(1)}$ such that $\|D\eta \circ f\|_{C^\infty([N])} \ll R^{O(1)}$. The corollary then follows after replacing η with $D\eta$. □

Corollary 3.3. *Suppose G is an m-dimensional simply connected Lie group with a degree d rational filtration $G_\bullet$, and Γ_j is a lattice in G for $j = 1, 2$ and $\mathcal{V}_j$ is an R-rational Mal'cev basis adapted to $(G_\bullet, \Gamma_j)$. Assume in addition that elements in $\mathcal{V}_2$ are R-rational combinations of elements in $\mathcal{V}_1$.*

For $f \in \mathrm{Poly}(\mathbb{Z}, G_\bullet)$, and $N \in \mathbb{N}$ such that $N \gg R^{O(1)}$, if $\{f(n)\Gamma_1\}_{n\in[N]}$ is not totally R^{-1}-equidistributed in G/Γ_1, then $\{f(n)\Gamma_2\}_{n\in[N]}$ is not totally $R^{-O(1)}$-equidistributed in G/Γ_2.

Proof. By Corollary 3.2, there is a non-trivial horizontal character η of G/Γ_1, i.e. a character $G \to \mathbb{R}/\mathbb{Z}$ that annihilates Γ_1, of size $|\eta|_{\mathcal{V}_1} \le R^{O(1)}$ that satisfies $\|\eta \circ f\|_{C^\infty([N])} \le R^{O(1)}$. Here the modulus $|\eta|_{\mathcal{V}_1} \le R^{O(1)}$ is measured in terms of the basis $\mathcal{V}_1$. Because all elements of $\mathcal{V}_2$ are R-rational combinations of those in $\mathcal{V}_1$, by Lemma 2.7, there is a positive integer $D \le R^{O(1)}$ such that for all $\gamma \in \Gamma_2$, $\gamma^D \in \Gamma_1$ and thus $D\eta(\gamma) = \eta(\gamma^D) = 0$. Then $D\eta$ is a horizontal character of both G/Γ_1 and G/Γ_2 with $|D\eta|_{\mathcal{V}_1} \le R^{O(1)}$. Again, because all elements of $\mathcal{V}_2$ are R-rational combinations of those in $\mathcal{V}_1$, $|D\eta|_{\mathcal{V}_2} \le R^{O(1)}$. After replacing η with $D\eta$, one may assert that:

There exists a non-trivial horizontal character η of G/Γ_2 such that $|\eta|_{\mathcal{V}_2} \le R^{O(1)}$ and $\|\eta \circ f\|_{C^\infty([N])} \le R^{O(1)}$. By Lemma 2.9, $\{f(n)\Gamma_2\}_{n\in[N]}$ fails to be totally $R^{-O(1)}$-equidistributed. □

The following is the refined statement that we will need later, which deals with generic restrictions of a 2-parameter polynomial to one variable.

Proposition 3.4. *Under Hypothesis 2.13, for $\widetilde{R} \ge R$ and $N, H \in \mathbb{N}$ such that $N, H \gg \widetilde{R}^{O(1)}$, at least one of the following holds:*

1. *either $\{g(n,h)\Gamma\}_{h\in[H]}$ is totally $\widetilde{R}^{-1}$-equidistributed in G/Γ for all but $\widetilde{R}^{-1}N$ values of $n \in [N]$;*
2. *or there exists a horizontal character η of G/Γ of modulus $|\eta| \le \widetilde{R}^{O(1)}$ such that $\|\eta(\omega_{j,k})\|_{\mathbb{R}/\mathbb{Z}} \le \widetilde{R}^{O(1)}N^{-j}H^{-k}$ for all $j, k \ge 0$.*

Proof. Assuming (1) fails, we try to establish (2). For more than $\widetilde{R}^{-1}N$ values of $n \in [N]$, $\{g(n,h)\Gamma\}_{h\in[H]}$ is not totally $\widetilde{R}^{-1}$-equidistributed. For every such n, by Corollary 3.2 there is a horizontal character η with $|\eta| \le \widetilde{R}^{O(1)}$ such that

$$\|\eta \circ g(n,\cdot)\|_{C^\infty([H])} \ll \widetilde{R}^{O(1)}. \tag{3.1}$$

Applying pigeonhole principle to the at least $\widetilde{R}^{-1}N$ values of $n \in [N]$, there is a common η with $0 < |\eta| < \widetilde{R}^{O(1)}$, such that (3.2) holds for at least $\widetilde{R}^{-O(1)}N$ choices of $n \in [N]$. By (2.6), this implies:

$$\left\| \sum_{\substack{j,k\geq 0\\ j+k\leq d}} \binom{n}{j}\binom{\cdot}{k}\eta(\omega_{jk}) \right\|_{C^\infty([H])} \ll \widetilde{R}^{O(1)},$$

which by Definition 2.2 means that

$$\left\| \sum_{j=0}^{d-k} \binom{n}{j}\eta(\omega_{jk}) \right\|_{\mathbb{R}/\mathbb{Z}} \ll \widetilde{R}^{O(1)}H^{-k},\ \forall k = 0, \cdots, d. \tag{3.2}$$

As this inequality holds for $\widetilde{R}^{-O(1)}N$ choices of $n \in [N]$, by Lemma 2.4 there is a positive integer $D > 0$ such that

$$\left\| D\sum_{j=0}^{d-k} \binom{\cdot}{j}\eta(\omega_{jk}) \right\|_{C^\infty([N])} \ll \widetilde{R}^{O(1)}H^{-k}\cdot \widetilde{R}^{O(1)} = \widetilde{R}^{O(1)}H^{-k},\ \forall k = 0, \cdots, d.$$

In other words,

$$\|D\eta(\omega_{jk})\|_{\mathbb{R}/\mathbb{Z}} \ll \widetilde{R}^{O(1)}H^{-k}N^{-j}, \forall k, j \geq 0 \text{ such that } k + j \leq d. \tag{3.3}$$

This is exactly the desired conclusion after replacing η with $D\eta$. □

Lemma 3.5. *If Case 3.4.(2) holds in Proposition 3.4, then there is a decomposition $g = \epsilon g'\gamma$ with $\epsilon, g', \gamma \in \mathrm{Poly}(\mathbb{Z}^2, G)$ such that:*

(1) ϵ is $(\widetilde{R}^{O(1)}, (N, H))$-smooth;
(2) $\eta \circ g' = 0$ while regarding $\eta : G/\Gamma \to \mathbb{R}/\mathbb{Z}$ as a morphism from G to $\mathbb{R}$;
(3) $\gamma(n, h)$ is $\widetilde{R}^{O(1)}$-rational for all $n, h \in \mathbb{Z}$.

Proof. The proof is the same as that of [GT12a, Lemma 9.2] except that we are not reducing to the case $g(0) = \mathrm{id}$. For completeness, we give a sketch.

For all integer pairs $j, k \geq 0$ with $j+k \leq d$, choose $u_{jk} \in \mathbb{R}^m$ such that $\eta(u_{jk}) \in \mathbb{Z}$ and $|\omega_{jk} - u_{jk}| \ll \widetilde{R}^{O(1)}N^{-j}H^{-k}$, and $v_{jk} \in \mathbb{Q}^m$ such that $\eta(u_{jk}) = \eta(v_{jk})$, where η is regarded as an $\mathbb{R}$-valued linear functional from $\mathbb{R}^m \cong \mathfrak{g}$. This can be done while requiring that $(u_{jk})_i = (v_{jk})_i = 0$ for all $i \leq m - m_{j+k}$. Furthermore, one can require that $v_{j,k}$ is from $(\frac{1}{D}\mathbb{Z})^m$ for some integer $1 \leq D \leq \widetilde{R}^{O(1)}$.

Then define ϵ, g' and γ by

$$\psi_{\mathcal{V}}(\epsilon(n,h)) = \sum_{\substack{j,k\geq 0\\ j+k\leq d}} (\omega_{jk} - u_{jk})\binom{n}{j}\binom{h}{k},\ \psi_{\mathcal{V}}(\gamma(n,h)) = \sum_{\substack{j,k\geq 0\\ j+k\leq d}} v_{jk}\binom{n}{j}\binom{h}{k},$$

and $g'(n, h) = \epsilon(n, h)^{-1}g(n, h)\gamma(n, h)^{-1}$. Then by Lemma 2.12, ϵ, γ belong to $\mathrm{Poly}(\mathbb{Z}^2, G_\bullet)$ and hence so does g' as $\mathrm{Poly}(\mathbb{Z}^2, G_\bullet)$ is a group.

By the bound on $|\omega_{jk} - v_{jk}|$, we know that for all $(n, h) \in [N] \times [H]$,

$$|\psi_{\mathcal{V}}(\epsilon(n+1,h)) - \psi_{\mathcal{V}}(\epsilon(n,h))| \ll \sum_{\substack{j\geq 1,k\geq 0\\ j+k\leq d}} \widetilde{R}^{O(1)}N^{-j}H^{-k}\cdot n^{j-1}h^k \ll \widetilde{R}^{O(1)}N^{-1}$$

and similarly $|\psi_{\mathcal{V}}(\epsilon(n, h+1)) - \psi_{\mathcal{V}}(\epsilon(n, h))| \ll \widetilde{R}^{O(1)}H^{-1}$. Moreover, $|\psi_{\mathcal{V}}(\epsilon(0, 0))| = |\omega_{00} - v_{00}| \ll \widetilde{R}^{O(1)}$. These inqualities guarantee property (1) for ϵ by [GT12a, Lemma A.5].

Property (2) holds as

$$\begin{aligned}&\eta(g'(n,h))\\=&\eta(g(n,h))-\eta(\epsilon(n,h))-\eta(\gamma(n,h))\\=&\sum_{\substack{j,k\geq 0\\ j+k\leq d}}\eta(\omega_{jk})\binom{n}{j}\binom{h}{k}-\sum_{\substack{j,k\geq 0\\ j+k\leq d}}\eta(\omega_{jk}-u_{jk})\binom{n}{j}\binom{h}{k}-\sum_{\substack{j,k\geq 0\\ j+k\leq d}}\eta(v_{jk})\binom{n}{j}\binom{h}{k}\\=&0.\end{aligned}$$

Finally, it follows from Lemma 2.7 that γ is $\widetilde{R}^{(O(1)}$-rational. This also implies by [GT12a, Lemma A.12] (or rather the natural multiparameter extension of it) that for some positive integer $q \ll (\widetilde{R}^{O(1)})^{O(1)} \ll \widetilde{R}^{O(1)}$, $\gamma(n,h)\Gamma$ is $q\mathbb{Z}^2$-periodic. Thus we have property (3). □

Using this, Green-Tao's factorization theorem [GT12a, Theorems 1.19 & 10.2] can be easily refined to the following:

Theorem 3.6. *Under Hypothesis 2.13, for $B \geq 1$, $N, H \in \mathbb{N}$ such that $N, H \gg R^{O(1)}$, there exists an integer $W \in [R, R^{O(B^m)}]$, a W-rational subgroup $G' \subseteq G$, a W-rational Mal'cev basis $\mathcal{V}'$ adapted to $(G'_\bullet, G' \cap \Gamma)$ consisting of W-rational combinations of vector in $\mathcal{V}$, and a decomposition $g = \epsilon g' \gamma$ with $\epsilon, g', \gamma \in \mathrm{Poly}(\mathbb{Z}^2, G_\bullet)$ such that:*

1. *ϵ is $(W, (N, H))$-smooth.*
2. *g' takes value in G'. And, with respect to the metric induced by $\mathcal{V}'$ on G'/Γ', $\{g'(n,h)\}_{h\in[H]}$ is totally W^{-B}-equidistributed for all but at most $W^{-B}N$ values of $n \in [N]$;*
3. *$\gamma(n,h)$ is W-rational for all $n, h \in \mathbb{Z}$. Moreover for some $1 \leq q \leq W$, $\{\gamma(n,h)\Gamma\}_{(n,h)\in\mathbb{Z}^2}$ is $q\mathbb{Z}^2$-periodic.*

Proof. We start with the squence $g(n)$ apply Proposition 3.4 with $\widetilde{R} = R^B$. If Case 3.4.(1) holds, then the theorem is true for $G' = G$, $W = R$, $\epsilon(n,h) = \gamma(n,h) = \mathrm{id}$ and $g' = g$.

If Case 3.4.(2) holds for a non-trivial horizontal character η_1 of G/Γ of norm $\ll \widetilde{R}^{O(1)}$ and Lemma 3.5 applies, yielding a decomposition $g = \epsilon_1 g_1' \gamma_1$. In this case, let $G_1' = \ker_G \eta_1$ and $\Gamma_1' = G_1' \cap \Gamma$. Then $(G_1')_\bullet = \{(G_1')_i\}_{i\geq 0} = \{G_1' \cap G_i\}_{i\geq 0}$ is a filtration of G_1'. Notice that each $(G_1')_i$ is a $\widetilde{R}^{O(1)}$-rational subgroup. For $R_1 = \widetilde{R}^{O(1)} = R^{O(B)}$, by [GT12a, Lemma A.10] G_1 has an R_1-rational Mal'cev basis $\mathcal{V}_1$ adapted to $((G_1)_\bullet, \Gamma_1')$ consisting of R_1-rational combinations of vector in $\mathcal{V}$.

We then again to apply Proposition 3.4 with $\widetilde{R} = R_1^B$, and apply Lemma 3.5 if necessary, to the sequence $\{g_1'(n)\Gamma_1'\}$ in G_1/Γ_1'. The argument is iterated if Case 3.4.(2) holds in every step. So in the k-th step, we will apply Proposition 3.4 with $\widetilde{R} = R_{k-1}^B$, and obtain, with $R_k = \left(R_{k-1}^B\right)^{O(1)} = (R_{k-1})^{O(B)}$:

- a non-trivial horizontal charcter η_k of G_{k-1}'/Γ_{k-1}' of norm $\ll R_k$;
- an R_k-rational Mal'cev basis $\mathcal{V}_k$ adapted to $((G_k')_\bullet, \Gamma_k')$ consisting of R_k-rational combinations of vector in $\mathcal{V}_{k-1}$, where $G_k' = \ker_{G_{k-1}} \eta_k$ and $(G_k')_i = G_k' \cap G_i$;
- a decomposition $g_{k-1}' = \epsilon_k g_k' \gamma_k$ in the group $\mathrm{Poly}(\mathbb{Z}^2, (G_{k-1})_\bullet)$,

such that:

- ϵ is $(R_k, (N, H))$-smooth with respect to the metric induced by $\mathcal{V}_{k-1}$ on G'_{k-1};
- g'_k takes value in G'_k, and thus $g'_k \in \mathrm{Poly}(\mathbb{Z}^2, (G'_k)_\bullet)$;
- γ'_k is R_k-rational with respect to the Mal'cev basis $\mathcal{V}_{k-1}$.

As $\dim G'_k$ strictly decreases, the process must stop at some $k \le m$. This means Case 3.4.(1) holds, i.e. $\{g'_k(n,h)\Gamma_k\}_{h\in[H]}$ is totally R_k^{-B}-equidistributed in G'_k/Γ'_k for all but $R_k^{-B}N$ values of $n \in [N]$.

Write $g = \epsilon g' \gamma$ where $\epsilon = \epsilon_1 \cdots \epsilon_k$, $g' = g'_k$ and $\gamma = \gamma_k \cdots \gamma_1$, $G' = G'_k$, $\mathcal{V}' = \mathcal{V}_k$ and $W = R_k$. Notice that since for each j, $\epsilon_j \in \mathrm{Poly}(\mathbb{Z}^2, (G'_j)_\bullet) \subseteq \mathrm{Poly}(\mathbb{Z}^2, G_\bullet)$ and $\mathrm{Poly}(\mathbb{Z}^2, G_\bullet)$ is a group, $\epsilon \subseteq \mathrm{Poly}(\mathbb{Z}^2, G_\bullet)$. Similarly γ is in $\mathrm{Poly}(\mathbb{Z}^2, G_\bullet)$ and so is g'.

It was shown above that the property (2) in the theorem holds for g'. The properties (1) and the W-rationality in (3) follow in the same way as in the proof of [GT12a, Theorem 10.2], after replacing W with $W^{O(1)}$ if necessary. Furthermore, by a multiparameter version of [GT12a, Lemma A.12], the 2-parameter sequence $\{\gamma'(n,h)\Gamma\}_{(n,h)\in\mathbb{Z}^2}$ is $q\mathbb{Z}^2$-periodic for some $q \ll W^{O(1)}$. Once again by replacing W with $W^{O(1)}$, we obtain the property (3) for γ.

Finally, remark that as $k \le m$, $R_k \ll R^{O(B^m)}$ and $W \ll R_k^{O(1)} \ll R^{O(B^m)}$. $\square$

4. Separation of major and minor arcs

From now on, we work under Hypothesis 2.13.

Notation 4.1. *Suppose $C_0 = O(1)$ is sufficiently large, and $B_1 \ge 10C_0$. Let N, H, and g be as in Theorem 3.6, applied with $B = B_1$. Also let ϵ, g', γ, W, q, G' and $\mathcal{V}'$ be as in the conclusion of the theorem. Without loss of generality, we may assume $R \ge 10$. In addition, after replacing the period q with a multiple of it if necessary, we may assume $q \in (\frac{W}{2}, W]$.*

Because $W \in [R, R^{O({B_1}^m)}]$, we will fix a constant $C_1 = O_{m,d}(1) \ge 1$ and assume

$$W \in [R, R^{C_1 {B_1}^m}]. \tag{4.1}$$

Let $F : G/\Gamma \to \mathbb{C}$ be a function with $\|F\| \le 1$. For every $n > 0$, choose θ_n from the unit circle such that

$$\Big|\sum_{h\le H} \beta(n+h)F(g(n,h)\Gamma)\Big| = \theta_n \sum_{h\le H} \beta(n+h)F(g(n,h)\Gamma). \tag{4.2}$$

Split $(0, H]$ into W^2 subintervals $I_1, \cdots, I_k$ of equal lengths $W^{-2}H$. Then for each n, the arithmetic progression $[H]$ is decomposed as the disjoint union

$$[H] = \bigsqcup_{\mathbf{j}\in\mathcal{J}} \mathcal{I}_{n,\mathbf{j}}$$

of arithmetic progressions

$$\mathcal{I}_{n,\mathbf{j}} = \{h \in I_k \cap \mathbb{N} : n + h \equiv j (\mathrm{mod}\ q)\},$$

where

$$\mathcal{J} = \{(k,j) : 1 \le k \le W^2, 0 \le j \le q-1\}. \tag{4.3}$$

Remark that

$$\#\mathcal{J} = W^2 q \in (\frac{1}{2}W^3, W^3]. \tag{4.4}$$

Thus the length of the arithmetic progression $\mathcal{I}_{n,\mathbf{j}}$ satisfies

$$\#\mathcal{I}_{n,\mathbf{j}} \in [W^{-3}H, 2W^{-3}H) \tag{4.5}$$

Because ϵ is $(W,(N,H))$-smooth, $\mathrm{d}_G(\epsilon(n,h),\mathrm{id}_G) \leq W$ for all $(n,h) \in [N]\times[H]$. Moreover, for any given $1 \leq k \leq W^2$, $\mathrm{d}_G(\epsilon(n,h),\epsilon(n,h')) \leq \frac{W}{H}\cdot W^{-2}H \leq W^{-1}$ for all $h,h' \in I_{n,k}$.

For a given pair $(n,\mathbf{j}) = (n,k,j)$, Choose $\epsilon_{n,\mathbf{j}} = \epsilon(n,h)$ for the smallest $h \in \mathcal{I}_{n,\mathbf{j}}$. As $\mathcal{I}_{n,\mathbf{j}} \subseteq I_{n,k}$, we know

$$d_G(\epsilon_{n,\mathbf{j}},\epsilon(n,h)) \leq W^{-1}, \forall h \in \mathcal{I}_{n,\mathbf{j}}. \tag{4.6}$$

Then

$$\mathrm{d}_G(\epsilon_{n,\mathbf{j}},\mathrm{id}_G) \ll W. \tag{4.7}$$

Choose a rational element $\gamma_{n,\mathbf{j}}$ from such that $\gamma_{n,\mathbf{j}}\Gamma = \gamma(n,h)\Gamma$ for any $h \in \mathcal{I}_{n,\mathbf{j}}$. The value of $\gamma_{n,\mathbf{j}}$ can in fact be chosen to be independent of the choice of $h \in \mathcal{I}_{n,\mathbf{j}}$ and q-periodic in n, because $\mathcal{I}_{n,\mathbf{j}} \subset q\mathbb{Z}+j$ and $\gamma(n,h)$ is q-periodic in both n and h. As $\gamma(n,h)$ is W-rational, and $\gamma_{n,\mathbf{j}} = \gamma(n,h)\xi$ for some $\xi \in \Gamma$, $\gamma_{n,\mathbf{j}}$ is $W^{O(1)}$-rational by Lemma 2.7. Moreover, we may choose $\gamma_{n,\mathbf{j}}$ from the fundamental domain $\psi_{\mathcal{V}}^{-1}([0,1)^m)$. In particular, by [GT12a, Lemma A.4],

$$\mathrm{d}_G(\gamma_{n,\mathbf{j}},\mathrm{id}_G) \ll R^{O(1)}. \tag{4.8}$$

Define $G_{n,\mathbf{j}}$ by $G_{n,\mathbf{j}} = \gamma_{n,\mathbf{j}}^{-1}G'\gamma_{n,\mathbf{j}}$ and $\Gamma_{n,\mathbf{j}} = G_{n,\mathbf{j}} \cap \Gamma$.

Lemma 4.2. *The following properties are true:*

(1) $G_{n,\mathbf{j}}$ is a $W^{O(1)}$-rational subgroup and $\Gamma_{n,\mathbf{j}}$ is a lattice of it;
(2) The assignments $G_{n,\mathbf{j}}$ and $\Gamma_{n,\mathbf{j}}$ are q-periodic in n;
(3) $G_{n,\mathbf{j}}$ has a $W^{O(1)}$ -rational Mal'cev basis $\mathcal{V}_{n,\mathbf{j}}$ adapted to $((G_{n,\mathbf{j}})_\bullet,\Gamma_{n,\mathbf{j}})$ that consists of $W^{O(1)}$-rational combinations of elements from $\mathcal{V}$. Here $(G_{n,\mathbf{j}})_\bullet$ consists of the subgroups $(G_{n,\mathbf{j}})_i = G_{n,\mathbf{j}} \cap G_i$.

Proof. Because $\gamma_{n,\mathbf{j}}$ is $W^{O(1)}$-rational and G' is a W-rational subgroup, by [GT12a, Lemma A.13], $G_{n,\mathbf{j}}$ is a $W^{O(1)}$-rational subgroup. As $\gamma_{n,\mathbf{j}}$ is q-periodic in n, so are the correspondences from $(n,\mathbf{j})$ to $G_{n,\mathbf{j}}$ and $\Gamma_{n,\mathbf{j}}$. The last property is given by [GT12a, Proposition A.10]. □

Define $g_{n,\mathbf{j}}(h) = \gamma_{n,\mathbf{j}}^{-1}g'(n,h)\gamma_{n,\mathbf{j}} \in G_{n,\mathbf{j}}$. Then $g_{n,\mathbf{j}} \in \mathrm{Poly}(\mathbb{Z},(G_{n,\mathbf{j}})_\bullet)$ and

$$\begin{aligned} g(n,h)\Gamma =&\epsilon(n,h)g'(n,h)\gamma(n,h)\Gamma = \epsilon(n,h)g'(n,h)\gamma_{n,\mathbf{j}}\Gamma \\ =&\epsilon(n,h)\gamma_{n,\mathbf{j}}g_{n,\mathbf{j}}(h)\Gamma,\ \forall h \in \mathcal{I}_{n,\mathbf{j}}. \end{aligned} \tag{4.9}$$

We then define a new function $F_{n,\mathbf{j}} : G_{n,\mathbf{j}}/\Gamma_{n,\mathbf{j}} \to \mathbb{C}$ by

$$F_{n,\mathbf{j}}(g\Gamma_{n,\mathbf{j}}) = \theta_n F(\epsilon_{n,\mathbf{j}}\gamma_{n,\mathbf{j}}g\Gamma). \tag{4.10}$$

Note that $F_{n,\mathbf{j}}$ is well-defined because if $g = \hat{g}\eta$ with $\eta \in \Gamma_{n,\mathbf{j}} \subset \Gamma$, then $g\Gamma = \hat{g}\Gamma$.

By (4.7), (4.8) and [GT12a, Lemma A.5] and

$$\|F_{n,\mathbf{j}}\|_{G_{n,\mathbf{j}}/\Gamma_{n,\mathbf{j}}} \leq (WR^{O(1)})^{O(1)}\|F\|_{G/\Gamma} \leq W^{O(1)}. \tag{4.11}$$

Lemma 4.3. *Suppose $C_0 = O(1)$ is sufficiently large and $B_1 \geq 10C_0$. There exists a subset $\mathcal{N} \subseteq [N]$ such that*

$$\#\mathcal{N} \geq (1 - W^{-B_1})N \tag{4.12}$$

and for all $(n,\mathbf{j}) \in \mathcal{N}\times\mathcal{J}$, the sequence $\{g_{n,\mathbf{j}}(h)\Gamma_{n,\mathbf{j}}\}_{h\in[H]}$ is totally $W^{-C_0{}^{-1}B_1}$-equidistributed in $G_{n,\mathbf{j}}/\Gamma_{n,\mathbf{j}}$.

Proof. By property (2) in Theorem 3.6, it suffices to show that if $\{g_{n,\mathbf{j}}(h)\Gamma_{n,\mathbf{j}}\}_{h\in[H]}$ is not totally $W^{-C_0{}^{-1}B_1}$-equidistributed, then $\{g'(n,h)\Gamma'\}_{h\in[H]}$ is not totally W^{-B_1}-equidistributed in G'/Γ'.

Consider the lattice $\Gamma'_{n,\mathbf{j}} = \gamma_{n,\mathbf{j}}\Gamma_{n,\mathbf{j}}\gamma_{n,\mathbf{j}}$ in G'. Then $G'/\Gamma'_{n,\mathbf{j}}$ is isomorphic to $G_{n,\mathbf{j}}/\Gamma_{n,\mathbf{j}}$ via the conjugacy $\mathrm{Ad}_{\gamma_{n,\mathbf{j}}}$ by $\gamma_{n,\mathbf{j}}$. Let $\mathcal{V}'_{n,\mathbf{j}}$ be the image of $\mathcal{V}_{n,\mathbf{j}}$ under $\mathrm{Ad}_{\gamma_{n,\mathbf{j}}}$, then it is a Mal'cev basis adapted to $(G'_\bullet, \Gamma'_{n,\mathbf{j}})$. Because of the bound (4.8) and [GT12a, Lemma A.5], $\mathrm{Ad}_{\gamma_{n,\mathbf{j}}}$ is $R^{O(1)}$-Lipschitz continuous. As $W \geq R$ and $g'(n,h) = \mathrm{Ad}_{\gamma_{n,\mathbf{j}}} g_{n,\mathbf{j}}(h)$, the sequence $\{g'(n,h)\Gamma'_{n,\mathbf{j}}\}_{h\in[H]}$ fails to be totally $W^{-C_0{}^{-1}B_1-O(1)}$-equidistributed in $G_{n,\mathbf{j}}/\Gamma'_{n,\mathbf{j}}$, with respect to the metric induced by $\mathcal{V}'_{n,\mathbf{j}}$.

Moreover, because $\gamma_{n,\mathbf{j}}$ is W-rational and satisfies the bound (4.8), it is a rational element of height bounded by $W^{O(1)}$. Since $\mathcal{V}_{n,\mathbf{j}}$ consists of $W^{O(1)}$-rational combinations of elements of $\mathcal{V}$, by [GT12a, Lemma A.11], so does $\mathcal{V}'_{n,\mathbf{j}}$. We also know that $\mathcal{V}'$ consists of W-rational combinations of elements from $\mathcal{V}$. Because they are both Mal'cev basis of G', it follows that $\mathcal{V}'$ consists of $W^{O(1)}$-rational combinations of elements from $\mathcal{V}'_{n,\mathbf{j}}$. Hence by Corollary 3.3, the sequence $\{g'(n,h)\Gamma'\}_{h\in[H]}$ fails to be totally $W^{-O(C_0{}^{-1}B_1+O(1))}$-equidistributed in $G_{n,\mathbf{j}}/\Gamma'$, with respect to the metric induced by $\mathcal{V}'$. As it will be assumed that $B_1 \geq 10C_0$, the lemma follows after updating the value of the constant $C_0 = O(1)$. □

By (4.9), (4.11) and (4.6), for all $h \in \mathcal{I}_{n,\mathbf{j}}$,

$$\mathrm{d}_{G/\Gamma}(\epsilon_{n,\mathbf{j}}\gamma_{n,\mathbf{j}}g_{n,\mathbf{j}}(h)\Gamma, g(n,h)\Gamma) \leq W^{-1}, \tag{4.13}$$

and

$$|F_{n,\mathbf{j}}(g_{n,\mathbf{j}}(h)\Gamma_{n,\mathbf{j}}) - \theta_n F(g(n,h)\Gamma)| \leq W^{-1}\|F\|. \tag{4.14}$$

Lemma 4.4. *For all Lipschitz function F on G/Γ, the sum*

$$\sum_{n\leq N}\Big|\sum_{h\leq H}\beta(n+h)F(g(n,h)\Gamma)\Big| \tag{4.15}$$

is approximated by

$$\sum_{n\leq N}\sum_{\mathbf{j}\in\mathcal{J}}\sum_{h\in\mathcal{I}_{n,\mathbf{j}}}\beta(n+h)F_{n,\mathbf{j}}(g_{n,\mathbf{j}}(h)\Gamma_{n,\mathbf{j}}), \tag{4.16}$$

up to an error bounded by $W^{-1}HN$.

Proof. As $[H] = \bigsqcup_{\mathbf{j}\in\mathcal{J}}\mathcal{I}_{n,\mathbf{j}}$, the claim follows from (4.2) and (4.14). □

For each triple $(n,\mathbf{j})$, decompose $F_{n,\mathbf{j}}$ as $\tilde{F}_{n,\mathbf{j}} + E_{n,\mathbf{j}}$ where $E_{n,\mathbf{j}} = \int_{G_{n,\mathbf{j}}/\Gamma_{n,\mathbf{j}}} F_{n,\mathbf{j}}$ is a constant and $\widetilde{F}_{n,\mathbf{j}}$ has zero average on $G_{n,\mathbf{j}}/\Gamma_{n,\mathbf{j}}$. Then (4.16) splits into the sum of a major arc part

$$\sum_{n\leq N}\sum_{\mathbf{j}\in\mathcal{J}}\sum_{h\in\mathcal{I}_{n,\mathbf{j}}}E_{n,\mathbf{j}}\beta(n+h). \tag{4.17}$$

and a minor art part

$$\sum_{n\leq N}\sum_{\mathbf{j}\in\mathcal{J}}\sum_{h\in\mathcal{I}_{n,\mathbf{j}}}\beta(n+h)\widetilde{F}_{n,\mathbf{j}}(g_{n,\mathbf{j}}(h)\Gamma_{n,\mathbf{j}}), \tag{4.18}$$

Note that,

$$|E_{n,\mathbf{j}}| \leq 1, \tag{4.19}$$

$$\|\widetilde{F}_{n,\mathbf{j}}\|_{G_{n,\mathbf{j}}/\Gamma_{n,\mathbf{j}}} \le 2\|F_{n,\mathbf{j}}\|_{G_{n,\mathbf{j}}/\Gamma_{n,\mathbf{j}}} \ll W^{O(1)}. \tag{4.20}$$

$$\|\widetilde{F}_{n,\mathbf{j}}\|_{C^0(G_{n,\mathbf{j}}/\Gamma_{n,\mathbf{j}})} \le 2. \tag{4.21}$$

5. Major arc estimate

The major arc estimate will concern only multiplicative functions β that are non-pretentious as defined by Granville and Soundararajan [GS07]. Given two 1-bounded multiplicative functions β, β' and a parameter $X \ge 1$, a distance $\mathbb{D}(\beta, \beta'; X) \in [0, +\infty)$ is defined by the formula

$$\mathbb{D}(\beta, \beta'; X) := \left(\sum_{p \le X} \frac{1 - \mathrm{Re}(\beta(p)\overline{\beta'(p)})}{p} \right)^{1/2}.$$

It is known that this gives a (pseudo-)metric on 1-bounded multiplicative functions; see [GS07, Lemma 3.1]. Moreover, let

$$M(\beta; X) := \inf_{|t| \le X} \mathbb{D}(\beta, n \mapsto n^{it}; X)^2 \tag{5.1}$$

and

$$\begin{aligned} M(\beta; X, Y) :&= \inf_{q \le Y; \chi(q)} M(\beta\overline{\chi}; X) \\ &= \inf_{|t| \le X; q \le Y; \chi(q)} \mathbb{D}(\beta, n \mapsto \chi(n) n^{it}; X)^2, \end{aligned} \tag{5.2}$$

where χ ranges over all Dirichlet characters of modulus $q \le Y$.

In addition, we also define

$$\widetilde{M}(\beta, X, Y) = \inf_{X' \ge X} M(\beta, X', Y). \tag{5.3}$$

Remark that $\widetilde{M}$ is increasing in X and decreasing in Y.

Instead of (4.17), we will first estimate

$$\sum_{n \le N} \sum_{\mathbf{j} \in \mathcal{J}} \sum_{h \in \mathcal{I}_{n,\mathbf{j}}} E_{n,\mathbf{j}} 1_{\mathcal{S}} \beta(n+h). \tag{5.4}$$

In this part, we will prove

Proposition 5.1. *Assuming Hypothesis 2.13, Notation 4.1 and the following inequalities:*

$$\frac{\log\log H}{\log H} < \epsilon < \frac{1}{500};\ 10 \le R_0 \le R \le H^{\frac{\epsilon}{C_1 B_1 m}}; \log H < (\log N)^{\frac{1}{2}}. \tag{5.5}$$

Then for all 1-bounded multiplicative function $\beta : \mathbb{N} \to \mathbb{C}$ and function $F : G/\Gamma \to \mathbb{C}$ with $\|F\| \le 1$, there exists a subset $\mathcal{S} \subseteq [0, N] \cap \mathbb{N}$ with $N - \#\mathcal{S} \ll \epsilon N$, such that

$$\begin{aligned} &\Big| \sum_{n \le N} \sum_{\mathbf{j} \in \mathcal{J}} \sum_{h \in \mathcal{I}_{n,\mathbf{j}}} E_{n,\mathbf{j}} 1_{\mathcal{S}} \beta(n+h) \Big| \\ &\ll \Big(W^{-\frac{1}{4}} + W^2 e^{-\frac{1}{2}\widetilde{M}(\beta, \frac{N}{W^5}, W)} \widetilde{M}(\beta, \frac{N}{W^5}, W)^{\frac{1}{2}} + W^2 (\log \frac{N}{W^5})^{-\frac{1}{100}} \Big) HN. \end{aligned} \tag{5.6}$$

Moreover, the choice of $\mathcal{S}$ only depends on H, N, and ϵ.

This will result from the following more precise statement.

Proposition 5.2. *Assume the settings of Theorem 3.6, and inequalities*

$$10 \le P_1 < Q_1 \le \exp\left((\log N)^{\frac{1}{2}}\right),\ (\log Q_1)^{480} < P_1; \tag{5.7}$$

$$W^{96} \le P_1 < Q_1 \le W^{-4}H. \tag{5.8}$$

Then there exists a subset $\mathcal{S} \subseteq [0, N] \cap \mathbb{N}$ *with*

$$N - \#\mathcal{S} \ll \frac{\log P_1}{\log Q_1} N, \tag{5.9}$$

such that for all 1*-bounded multiplicative function* $\beta : \mathbb{N} \to \mathbb{C}$ *and function* $F : G/\Gamma \to \mathbb{C}$ *with* $\|F\| \le 1$*,*

$$\begin{aligned}&\Big|\sum_{n\le N}\sum_{\mathbf{j}\in\mathcal{J}}\sum_{h\in\mathcal{I}_{n,\mathbf{j}}} E_{n,\mathbf{j}} 1_{\mathcal{S}} \beta(n+h)\Big| \\ \ll&\Big(W^{-\frac{1}{4}} + W^2 e^{-\frac{1}{2}\widetilde{M}(\beta,\frac{N}{W^5},W)}\widetilde{M}(\beta,\frac{N}{W^5},W)^{\frac{1}{2}} + W^2(\log\frac{N}{W^5})^{-\frac{1}{100}} \\ &+\frac{(\log H)^{\frac{1}{6}}}{P_1^{\frac{1}{48}}}\Big)HN.\end{aligned} \tag{5.10}$$

Moreover, the choice of $\mathcal{S}$ *only depends on* H*,* N*,* P_1 *and* Q_1*.*

Proof of Proposition 5.1 assuming Proposition 5.2. Let $Q_1 = H^{\frac{96}{100}}$ and $P_1 = Q_1^{500\epsilon}$. The inequalities in (5.5), together with the fact that $W \in [R, R^{C_1 B_1{}^m}]$, imply $W < H^\epsilon < H^{\frac{1}{500}}$, $Q_1 < W^{-4}H$, and $P_1 = H^{480\epsilon}$, which in turn guarantee (5.7) and (5.8).

We also have

$$\frac{(\log H)^{\frac{1}{6}}}{P_1^{\frac{1}{48}}} \le \frac{(\log H)^{\frac{1}{6}}}{H^{2\epsilon}} < H^{(-2+\frac{1}{6})\epsilon} < H^{-\epsilon} < W^{-1},$$

and

$$\frac{\log P_1}{\log Q_1} = 500\epsilon \ll \epsilon.$$

So Proposition 5.1 follows from (5.10). Notice that $\mathcal{S}$ only depends on N, H, P_1 and Q_1, where as P_1 and Q_1 are determined by H and ϵ. □

The following constants are defined in [MRT15, §2]:

Definition 5.3. *Given* P_1*,* Q_1 *as in (5.7), let* P_r*,* Q_r *be inductively defined by*

$$P_r = \exp(r^{4r}(\log Q_1)^{r-1}\log P_1),\ Q_r = \exp(r^{4r+2}(\log Q_1)^r).$$

Let r_+ *be the largest index such that* $Q_{r_+} \le \exp\left(\frac{(\log N)^{\frac{1}{2}}}{2}\right)$*. Also define*

$$\mathcal{S}_{P_1,Q_1,N} = \{n \le N : n \text{ has at least one prime factor in } [P_r, Q_r], \forall 1 \le r \le r_+\}.$$

Lemma 5.4. [MRT15, Lemma 2.2] $\#(\{1 \le n \le N\} \setminus \mathcal{S}_{P_1,Q_1,N}) \ll \frac{\log P_1}{\log Q_1} N$.

In addition to the conditions in Definition 5.3, we shall also assume $H \ll N$ and (5.8). We will also write simply

$$\mathcal{S} = \mathcal{S}_{P_1,Q_1,N} \tag{5.11}$$

when it does not cause ambiguity. Clearly, the construction of $\mathcal{S}$ depends only on N, P_1 and Q_1.

Following [MRT15, p2177-2178], denote by $\hat{\beta}$ the completely multiplicative function determined by $\hat{\beta}(p) = \beta(p)$ for all prime numbers p. Then the Dirichlet inverse of $\hat{\beta}$ is $\mu\hat{\beta}$, and thus $\beta = \hat{\beta} * \eta$, where $\eta = \beta * \mu\hat{\beta}$ is the Dirichlet convolution between β and $\mu\hat{\beta}$. Then the function η is multiplicative, bounded by 2 in absolute value, and satisfies

$$\sum_{n=1}^{\infty} |\eta(n)| n^{-(\frac{1}{2}+\sigma)} = O_\epsilon(1) \tag{5.12}$$

for all $\sigma > 0$. Note that $\mathbb{D}(\beta, \beta'; N) = \mathbb{D}(\hat{\beta}, \beta'; N)$ for all β'.

For $1 \le k \le W^2$ let

$$f_{n,k}(h) = \sum_{j=0}^{q-1} E_{n,(k,j)} 1_{\mathcal{I}_{n,(k,j)}}(h)$$

on $I_{n,k}$. Then $f_{n,k}$ is bounded by 1 in absolute value and q-periodic on $I_k \cap \mathbb{N}$. Furthermore,

$$\begin{aligned}(5.4) &= \sum_{n\le N} \sum_{k \le W^2} \sum_{h \in I_k \cap \mathbb{N}} 1_{\mathcal{S}} \beta(n+h) f_{n,k}(h) \\ &= \sum_{n\le N} \sum_{k \le W^2} \sum_{a \in \mathbb{N}} \eta(a) \sum_{\substack{b \in \mathbb{N} \\ ab \in n+I_k}} 1_{\mathcal{S}}(ab) \hat{\beta}(b) f_{n,k}(ab-n)\end{aligned} \tag{5.13}$$

By (5.12), the contribution of terms with $a > W$ is bounded:

Lemma 5.5. $\left| \sum_{n\le N} \sum_{k \le W^2} \sum_{a > W} \eta(a) \sum_{\substack{b \in \mathbb{N} \\ ab \in n+I_k}} 1_{\mathcal{S}}(ab) \hat{\beta}(b) f_{n,k}(ab-n) \right| \ll W^{-\frac{1}{4}} HN.$

Proof. For every $x \in [0, N]$ and $k \le W^2$,

$$\begin{aligned}&\left| \sum_{a > W} \eta(a) \sum_{\substack{b \in \mathbb{N} \\ ab \in n+I_k}} 1_{\mathcal{S}}(ab) \hat{\beta}(b) f_{n,k}(ab-n) \right| \\ \le &\sum_{a > W} |\eta(a)| \cdot 2a^{-1} W^{-2} H \le \sum_{a > W} |\eta(a)| a^{-\frac{3}{4}} \cdot 2W^{-\frac{1}{4}} \cdot W^{-2} H \\ \ll &W^{-\frac{1}{4}} \cdot W^{-2} H.\end{aligned} \tag{5.14}$$

The lemma follows by summing over $1 \le k \le W^2$ and $n \le N$. $\square$

Next, we aim to bound

$$\begin{aligned}&\sum_{n\le N} \sum_{k \le W^2} \sum_{a \le W} \eta(a) \sum_{\substack{b \in \mathbb{N} \\ ab \in n+I_k}} 1_{\mathcal{S}}(ab) \hat{\beta}(b) f_{n,k}(ab-n) \\ = &\sum_{n\le N} \sum_{k \le W^2} \sum_{a \le W} \eta(a) \sum_{\substack{b \in \mathbb{N} \\ ab \in n+I_k}} 1_{\mathcal{S}} \hat{\beta}(b) f_{n,k}(ab-n).\end{aligned} \tag{5.15}$$

Here the equality is because of the fact that, as $a \le W < P_1 < Q_1$, $b \in \mathcal{S}$ if and only if $ab \in \mathcal{S}$.

Given $a \le W$, $k \le W^2 < P_1$ and $n \le N$, decompose $\{b \in \mathbb{N} : ab \in n + I_k\}$ according to $u = \gcd(b, q)$:

$$\begin{aligned}
&\sum_{\substack{b\in\mathbb{N}\\ ab\in n+I_k}} 1_{\mathcal{S}}\hat{\beta}(b) f_{n,k}(ab-n)\\
=&\sum_{u|q}\sum_{\substack{ab\in n+I_k\\ (b,q)=u}} 1_{\mathcal{S}}\hat{\beta}(b) f_{n,k}(ab-n)\\
=&\sum_{u|q}\hat{\beta}(u)\sum_{\substack{auv\in n+I_k\\ (v,\frac{q}{u})=1}} 1_{\mathcal{S}}\hat{\beta}(v) f_{n,k}(auv-n),
\end{aligned}\tag{5.16}$$

where the last equality uses the fact that $1_{\mathcal{S}}(uv)\hat{\beta}(uv) = 1_{\mathcal{S}}(v)\hat{\beta}(u)\hat{\beta}(v)$, which is because $\hat{\beta}$ is completely multiplicative and $u \le q \le W < P_1$.

The Dirichlet characters of conductor $\frac{q}{u}$ form an orthonormal basis of the l^2-space on the finite abelian group $\left(\mathbb{Z}/(\frac{q}{u})\mathbb{Z}\right)^{\times}$.

Since $f_{n,k,a,u} : v \to f_{n,k}(auv-n)1_{(v,\frac{q}{u})=1}$ is $\frac{q}{u}$-periodic, it can be decomposed as a linear combination $\sum_{\chi \bmod^* \frac{q}{u}} w_{n,k,a,u,\chi}\chi$ of such characters. Then,

$$\sum_{\chi \bmod^* \frac{q}{u}} |w_{n,k,a,u,\chi}|^2 \le \|f_{n,k,a,u}\|_{l^\infty} \le 1. \tag{5.17}$$

It follows from this and (5.16) that, by Cauchy-Schwarz inequality,

$$\begin{aligned}
&\Big|\sum_{\substack{b\in\mathbb{N}\\ ab\in n+I_k}} 1_{\mathcal{S}}\hat{\beta}(b) f_{n,k}(ab-n)\Big|^2\\
=&\Big|\sum_{u|q}\hat{\beta}(u)\sum_{\chi \bmod^* \frac{q}{u}} w_{n,k,a,u,\chi}\sum_{\substack{v\in\mathbb{N}\\ auv\in n+I_k}} 1_{\mathcal{S}}\hat{\beta}(v)\chi(v)\Big|^2\\
\le&\Big(\sum_{u|q}|\hat{\beta}(u)|^2\Big)\cdot\Bigg(\sum_{u|q}\Big|\sum_{\chi \bmod^* \frac{q}{u}} w_{n,k,a,u,\chi}\sum_{v\in(\frac{n}{au}+\frac{1}{au}I_k)\cap\mathbb{N}} 1_{\mathcal{S}}\hat{\beta}(v)\chi(v)\Big|^2\Bigg)\\
\le& q\Bigg(\sum_{u|q}\Big(\sum_{\chi \bmod^* \frac{q}{u}}|w_{n,k,a,u,\chi}|^2\Big)\Big(\sum_{\chi \bmod^* \frac{q}{u}}\Big|\sum_{v\in(\frac{n}{au}+\frac{1}{au}I_k)\cap\mathbb{N}} 1_{\mathcal{S}}\hat{\beta}(v)\chi(v)\Big|^2\Big)\Bigg)\\
\le& q\Bigg(\sum_{\substack{u|q\\ \chi \bmod^* \frac{q}{u}}}\Big|\sum_{v\in(\frac{n}{au}+\frac{1}{au}I_k)\cap\mathbb{N}} 1_{\mathcal{S}}\hat{\beta}(v)\chi(v)\Big|^2\Bigg).
\end{aligned}\tag{5.18}$$

Therefore, again by Cauchy-Schwarz inequality,

$$\begin{aligned}
&\Big|\sum_{n\le N}\sum_{\substack{b\in\mathbb{N}\\ ab\in n+I_k}} 1_{\mathcal{S}}\hat{\beta}(b)f_{n,k}(ab-n)\Big|^2\\
\le& N\sum_{n\le N}\Big|\sum_{\substack{b\in\mathbb{N}\\ ab\in n+I_k}} 1_{\mathcal{S}}\hat{\beta}(b)f_{n,k}(ab-n)\Big|^2\\
\le& N\sum_{n\le N} q\sum_{\substack{u|q\\ \chi \bmod^* \frac{q}{u}}}\Big|\sum_{v\in(\frac{n}{au}+\frac{1}{au}I_k)\cap\mathbb{N}} 1_{\mathcal{S}}\hat{\beta}(v)\chi(v)\Big|^2\\
\le& WN\sum_{n\le N}\sum_{\substack{u\le W\\ \mathrm{cond}\chi\le\frac{W}{u}}}\Big|\sum_{v\in(\frac{n}{au}+\frac{1}{au}I_k)\cap\mathbb{N}} 1_{\mathcal{S}}\hat{\beta}(v)\chi(v)\Big|^2\\
\le& WN\sum_{\substack{u\le W\\ \mathrm{cond}\chi\le\frac{W}{u}}} au\sum_{n\le\frac{N}{au}}\Big|\sum_{v\in(n+\frac{1}{au}I_k)\cap\mathbb{N}} 1_{\mathcal{S}_{P_1,Q_1,\frac{N+H}{au}}}\hat{\beta}(v)\chi(v)\Big|^2.
\end{aligned}\tag{5.19}$$

The sum within (5.19) is controlled by the work of Matomäki-Radziwiłł-Tao on averages of multiplicative functions on short intervals.

Theorem 5.6. *(Matomäki-Radziwiłł-Tao)* [MRT15, Thm A.2] *Suppose that* $10 < P_1 < Q_1 < H$ *and* $(\log Q_1)^{480} < P_1$*, then for all sufficiently large* N*, 1-bounded multiplicative function* β *and Dirichlet character of modulus bounded by* Y*,*

$$\begin{aligned}
&\sum_{N<n\le 2N}\Big|\sum_{n\le v\le n+H_0} 1_{\mathcal{S}_{P_1,Q_1,2N+H_0}}\beta(v)\chi(v)\Big|^2\\
\ll&\Big(e^{-M(\beta,N,Y)}M(\beta,N,Y)+\frac{(\log H_0)^{\frac{1}{3}}}{P_1^{\frac{1}{12}}}+(\log N)^{-\frac{1}{50}}\Big)H_0^2N,
\end{aligned}$$

where $M(\beta,N,Y)$ *is defined by* (5.2).

Corollary 5.7. *Assuming the conditions* (5.7) *and* (5.8)*, for all positive integers* $k\le W^2, T\le W^2$*, 1-bounded multiplicative function* β*, and primitive characters* χ *of conductor bounded by* W*,*

$$\begin{aligned}
&T\sum_{n\le\frac{N}{T}}\Big|\sum_{v\in(n+\frac{1}{T}I_k)\cap\mathbb{N}} 1_{\mathcal{S}_{P_1,Q_1,\frac{N+H}{T}}}\hat{\beta}(v)\chi(v)\Big|^2\\
\ll&\Big(W^{-7}+e^{-\widetilde{M}(\beta,\frac{N}{W^5},W)}\widetilde{M}(\beta,\frac{N}{W^5},W)+\frac{(\log H)^{\frac{1}{3}}}{P_1^{\frac{1}{12}}}+\big(\log\frac{N}{W^5}\big)^{-\frac{1}{50}}\Big)\frac{H^2N}{T^2}.
\end{aligned}$$

Proof. Decompose $[0,\frac{N}{T}]$ into dyadic intervals $[\frac{N}{2^iT},\frac{N}{2^{i-1}T}]$ for $i=1,\cdots,\lceil 3\log_2 W\rceil$, and $[0,\frac{N}{2^{\lceil 3\log_2 W\rceil}T}]$.

The contribution of the last interval can be bound trivially by

$$T\cdot\frac{N}{W^3T}\cdot\Big(\frac{H}{W^2T}\Big)^2\ll W^{-7}\frac{H^2N}{T^2}.$$

By Theorem 5.6, with $H_0 = \frac{H}{W^2T} \leq W^{-2}H$, the contribution from the dyadic intervals is

$$\ll \sum_{i \leq \lceil 3\log_2 W \rceil} \Big(e^{-M(\hat{\beta}, \frac{N}{2^iT}, W)} M(\hat{\beta}, \frac{N}{2^iT}, W) + \frac{(\log H)^{\frac{1}{3}}}{P_1^{\frac{1}{12}}} + (\log \frac{N}{2^iT})^{-\frac{1}{50}} \Big) \frac{H^2N}{2^{2i}T^2}$$
$$\ll \Big(e^{-\widetilde{M}(\hat{\beta}, \frac{N}{W^5}, W)} \widetilde{M}(\hat{\beta}, \frac{N}{W^5}, W) + \frac{(\log H)^{\frac{1}{3}}}{P_1^{\frac{1}{12}}} + (\log \frac{N}{W^5})^{-\frac{1}{50}} \Big) \frac{H^2N}{T^2}.$$

The corollary follows because $\widetilde{M}(\beta, \cdot, \cdot)$ and $\widetilde{M}(\hat{\beta}, \cdot, \cdot)$ have the same value. $\square$

Therefore, with $\square$ denoting the bracketed coefficient in Corollary 5.7,

$$(5.19) \ll WN \sum_{u \leq W} \frac{W}{u} \cdot \square \frac{H^2N}{(au)^2} \ll \square \frac{W^2H^2N^2}{a^2}. \tag{5.20}$$

In other words,

$$\Big| \sum_{n \leq N} \sum_{\substack{b \in \mathbb{N} \\ ab \in n + I_k}} 1_{\mathcal{S}} \hat{\beta}(b) f_{n,k}(ab - n) \Big| \ll a^{-1} \square^{\frac{1}{2}} WHN \tag{5.21}$$

for all $a \leq W$, $k \leq W^2$.

Lemma 5.8. *Assuming the conditions* (5.7) *and* (5.8)*, we have*

$$\Big| \sum_{n \leq N} \sum_{k \leq W^2} \sum_{a \leq W} \eta(a) \sum_{\substack{b \in \mathbb{N} \\ ab \in n + I_k}} 1_{\mathcal{S}} \hat{\beta}(b) f_{n,k}(ab - n) \Big| \ll \square^{\frac{1}{2}} W^3 HN.$$

Proof. Summing (5.21) over k and a, one can see that the left hand side is bounded by

$$\sum_{a \leq W} \eta(a) a^{-1} \square^{\frac{1}{2}} W^3 HN.$$

which is in turn by (5.12) bounded by the right hand side up to a multiplicative constant. $\square$

Proof of Proposition 5.2. By merging Lemmas 5.5, Lemma 5.8 into (5.13), we see that

$$\begin{aligned} &|(5.4)| \\ \ll& W^{-\frac{1}{4}} HN + W^2 \Big(W^{-5} + e^{-\widetilde{M}(\beta, \frac{N}{W^5}, W)} \widetilde{M}(\beta, \frac{N}{W^5}, W) + \frac{(\log H)^{\frac{1}{3}}}{P_1^{\frac{1}{12}}} \\ &+ (\log \frac{N}{W^5})^{-\frac{1}{50}} \Big)^{\frac{1}{2}} HN \\ \ll& \Big(W^{-\frac{1}{4}} + W^2 e^{-\frac{1}{2}\widetilde{M}(\beta, \frac{N}{W^5}, W)} \widetilde{M}(\beta, \frac{N}{W^5}, W)^{\frac{1}{2}} + W^2 (\log \frac{N}{W^5})^{-\frac{1}{100}} \\ &+ W^2 \frac{(\log H)^{\frac{1}{6}}}{P_1^{\frac{1}{24}}} \Big) HN, \end{aligned}$$

which is in turn bounded by the right hand side up to a constant multiple.

The proposition follows, thanks to Lemma 5.4 and the fact that $W^2 \leq P_1^{\frac{1}{48}}$. $\square$

6. Minor arc estimate

In Sections 6 and 7, we will provide a bound to (4.18) under appropriate hypothesis.

Proposition 6.1. *Assuming Hypothesis 2.13 and Notation 4.1, the constant C_0 being sufficiently large, and the following inequalities:*

$$0 < \epsilon < \frac{1}{100}; B_1 \geq C_0; 10 \leq R_0 \leq R \leq H^{\frac{\epsilon}{C_1 B_1^{m+1}}}. \tag{6.1}$$

Then for all 1-bounded multiplicative function $\beta : \mathbb{N} \to \mathbb{C}$ and function $F : G/\Gamma \to \mathbb{C}$ with $\|F\| \leq 1$, there exists a subset $\mathcal{S} \subseteq [0, N] \cap \mathbb{N}$ with $N - \#\mathcal{S} \ll \epsilon N$, such that

$$\begin{aligned} &\Big| \sum_{n \leq N} \sum_{\mathbf{j} \in \mathcal{J}} \sum_{h \in \mathcal{I}_{n,\mathbf{j}}} 1_{\mathcal{S}} \beta(n+h) \widetilde{F}_{n,\mathbf{j}}(g_{n,\mathbf{j}}(h)\Gamma_{n,\mathbf{j}}) \Big| \\ \ll &(W^{-C_0{}^{-1} B_1} \log H + H^{-\epsilon}) HN. \end{aligned} \tag{6.2}$$

Moreover, the choice of $\mathcal{S}$ only depends on H, N, and ϵ.

Following [MRT15, §3], let $\mathcal{P}$ be the set of primes in $[P_1, Q_1]$ for some fixed values $W < P_1 < Q_1 < H$. A priori, P_1, Q_1 do not have to assume the same values as in §5.

Lemma 6.2. *Under the assumptions of Proposition 6.1, there exists a subset $\mathcal{S} \subseteq [0, N] \cap \mathbb{N}$ with $N - \#\mathcal{S} \ll \frac{\log P_1}{\log Q_1} N$, such that for all $n \leq N$,*

$$\sum_{\substack{h \leq H \\ n+h \in \mathcal{S}}} \left| \beta(n+h) - \sum_{p \in \mathcal{P}} \sum_{l \in \mathbb{N}} \frac{1_{pl=n+h} \beta(p)\beta(l)}{1 + \#\{q \in \mathcal{P} : q | l\}} \right| \ll \frac{H}{P_1}.$$

The construction of $\mathcal{S}$ only depends on N and P_1, Q_1.

Proof. Define

$$\mathcal{S} = \{n \leq N : \exists p \in \mathcal{P}, p | n\}$$

and

$$\mathcal{F} = \{n \in \mathbb{N} \leq N : p^2 \nmid n, \forall p \in \mathcal{P}\}.$$

Note that these definitions depends only on N, P_1 and Q_1.

By Lemma 5.4, $N - \#\mathcal{S} \ll \frac{\log P_1}{\log Q_1} N$.

Decompose the sum on the left hand side as $\sum_{\substack{h \leq H \\ n+h \in S \setminus \mathcal{F}}} + \sum_{\substack{h \leq H \\ n+h \in \mathcal{S} \cap \mathcal{F}}}$. We will bound the two components separately.

Remark first that, when $n + h \in \mathcal{S}$,

$$\begin{aligned} \sum_{p \in \mathcal{P}} \sum_{l \in \mathbb{N}} \frac{1_{pl=n+h}}{1 + \#\{q \in \mathcal{P} : q | l\}} &= \sum_{p \in \mathcal{P}} \sum_{l \in \mathbb{N}} \frac{1_{pl=n+h}}{1_{p^2 | n+h} + \#\{q \in \mathcal{P} : q | n\}} \\ &\leq \sum_{p \in \mathcal{P}} \frac{1_{p | n+h}}{\#\{q \in \mathcal{P} : q | n\}} = 1. \end{aligned} \tag{6.3}$$

In particular, the equality holds when $n \in \mathcal{S} \cap \mathcal{F}$.

If $n+h \in \mathcal{S} \cap \mathcal{F}$, then for all $p \in \mathcal{P}$ and $l \in \mathbb{N}$ such that $pl = n+h$, $p \nmid l$ and thus $\beta(n+h) = \beta(p)\beta(l)$. Hence

$$\left|\beta(n+h) - \sum_{p\in\mathcal{P}} \sum_{l\in\mathbb{N}} \frac{1_{pl=n+h}\beta(p)\beta(l)}{1+\#\{q\in\mathcal{P} : q|l\}}\right| = \left|\beta(n+h) - \sum_{p\in\mathcal{P}} \sum_{l\in\mathbb{N}} \frac{1_{pl=n+h}\beta(n+h)}{1+\#\{q\in\mathcal{P} : q|l\}}\right| = 0.$$

So

$$\sum_{\substack{h\le H\\ n+h\in\mathcal{S}\cap\mathcal{F}}} = 0 \tag{6.4}$$

On the other hand, if $n+h \in \mathcal{S}\setminus\mathcal{F}$, then

$$\left|\beta(n+h) - \sum_{p\in\mathcal{P}} \sum_{l\in\mathbb{N}} \frac{1_{pl=n+h}\beta(p)\beta(l)}{1+\#\{q\in\mathcal{P} : q|l\}}\right| \le 1 + \sum_{p\in\mathcal{P}} \sum_{l\in\mathbb{N}} \frac{1_{pl=n+h}}{1+\#\{q\in\mathcal{P} : q|l\}} \le 2.$$

So

$$\sum_{\substack{h\le H\\ n+h\in\mathcal{S}\setminus\mathcal{F}}} \le 2 \sum_{\substack{h\le H\\ n+h\in\mathcal{S}\setminus\mathcal{F}}} 1 \le 2\sum_{h\le H}\sum_{p\in\mathcal{P}} 1_{p^2|n+h} \le 2\sum_{p\ge P_1} \frac{H}{p^2} \ll \frac{H}{P_1}. \tag{6.5}$$

It now suffices to add together (6.4) and (6.5). ∎

Corollary 6.3. *The integral*

$$\sum_{n\le N}\sum_{\mathbf{j}\in\mathcal{J}}\sum_{h\in\mathcal{I}_{n,\mathbf{j}}} 1_{\mathcal{S}}\beta(n+h)\widetilde{F}_{n,\mathbf{j}}(g_{n,\mathbf{j}}(h)\Gamma_{n,\mathbf{j}}), \tag{6.6}$$

is approximated by

$$\sum_{n\in\mathcal{N}}\sum_{\mathbf{j}\in\mathcal{J}}\sum_{p\in\mathcal{P}}\sum_{l\in\mathbb{N}} \frac{1_{pl\in n+\mathcal{I}_{n,\mathbf{j}}}\beta(p)\beta(l)}{1+\#\{q\in\mathcal{P} : q|l\}}\widetilde{F}_{n,\mathbf{j}}(g_{n,\mathbf{j}}(pl)\Gamma_{n,\mathbf{j}}) \tag{6.7}$$

within an error of $O(P_1^{-1} + W^{-B_1})\cdot HN$.

Here the set $\mathcal{N} \subseteq [N]$ is chosen as in (4.12).

Proof. The corollary directly follows from the lemma above and the inequality (4.12). ∎

Take $P_1 = 2^{s_-}$ and $Q_1 = 2^{s_+}$ for integers $s_- < s_+$. The expression (6.7) splits into the sum

$$\sum_{s\in(s_-,s_+]}\sum_{n\in\mathcal{N}}\sum_{\mathbf{j}\in\mathcal{J}}\sum_{p\in(2^{s-1},2^s]}\sum_{l\in\mathbb{N}} \frac{1_{pl\in n+\mathcal{I}_{n,\mathbf{j}}}\beta(p)\beta(l)}{1+\#\{q\in\mathcal{P} : q|l\}}\widetilde{F}_{n,\mathbf{j}}(g_{\mathbf{j}}(n)\Gamma_{n,\mathbf{j}}), \tag{6.8}$$

over all integers $s \in [s_-, s_+]$.

Notation 6.4. *Here and below, the letter* p*, as well as* p_1*,* p_2*, will always refer to prime numbers only.*

Observe that, for all given s,

$$\begin{aligned}(6.9)\quad &\Big|\sum_{n\in\mathcal{N}}\sum_{\mathbf{j}\in\mathcal{J}}\sum_{p\in(2^{s-1},2^s]}\sum_{l\in\mathbb{N}}\frac{1_{pl\in n+\mathcal{I}_{n,\mathbf{j}}}\beta(p)\beta(l)}{1+\#\{q\in\mathcal{P}:q|l\}}\widetilde{F}_{n,\mathbf{j}}(g_{n,\mathbf{j}}(pl)\Gamma_{n,\mathbf{j}})\Big|\\ \le&\sum_{l\in\mathbb{N}}\frac{|\beta(l)|}{1+\#\{q\in\mathcal{P}:q|l\}}\Big|\sum_{n\in\mathcal{N}}\sum_{\mathbf{j}\in\mathcal{J}}\sum_{\substack{p\in(2^{s-1},2^s]\\ pl\in n+\mathcal{I}_{n,\mathbf{j}}}}\beta(p)\widetilde{F}_{n,\mathbf{j}}(g_{n,\mathbf{j}}(pl)\Gamma_{n,\mathbf{j}})\Big|\\ \le&\sum_{l\le\frac{N+H}{2^{s-1}}}\Big|\sum_{n\in\mathcal{N}}\sum_{\mathbf{j}\in\mathcal{J}}\sum_{\substack{p\in(2^{s-1},2^s]\\ pl\in n+\mathcal{I}_{n,\mathbf{j}}}}\beta(p)\widetilde{F}_{n,\mathbf{j}}(g_{n,\mathbf{j}}(pl)\Gamma_{n,\mathbf{j}})\Big|\\ \ll&2^{-\frac{s}{2}}N^{\frac12}\Big(\sum_{l\le\frac{N+H}{2^{s-1}}}\Big|\sum_{n\in\mathcal{N}}\sum_{\mathbf{j}\in\mathcal{J}}\sum_{\substack{p\in(2^{s-1},2^s]\\ pl\in n+\mathcal{I}_{n,\mathbf{j}}}}\beta(p)\widetilde{F}_{n,\mathbf{j}}(g_{n,\mathbf{j}}(pl)\Gamma_{n,\mathbf{j}})\Big|^2\Big)^{\frac12}.\end{aligned}$$

This is because if $\mathbf{j}=(k,j)$ and $pl\in n+\mathcal{I}_{n,\mathbf{j}}$, then $2^{s-1}l\le pl\le N+H$.

For a configuration $\mathbf{n}=(n,\mathbf{j})=(n,k,j)\in\mathcal{N}\times\mathcal{J}$, define an arithmetic progression

$$\mathcal{A}_{\mathbf{n},p}=\{l\in\mathbb{N}:pl\in n+\mathcal{I}_{n,\mathbf{j}}\}=\{l\in\mathbb{N}:pl-n\in I_k,pl\equiv j(\mathrm{mod}\ q)\}\tag{6.10}$$

For two such given configurations

$$\mathbf{n}_1=(n_1,\mathbf{j}_1)=(n_1,k_1,j_1),\mathbf{n}_2=(n_2,\mathbf{j}_2)=(n_2,k_2,j_2)\in\mathcal{N}\times\mathcal{J},$$

write

$$\mathcal{A}_{\mathbf{n}_1,\mathbf{n}_2,p_1,p_2}=\mathcal{A}_{\mathbf{n}_1,p_1}\cap\mathcal{A}_{\mathbf{n}_2,p_2}.\tag{6.11}$$

Then

$$\begin{aligned}(6.12)\quad &\sum_{l\le\frac{N+H}{2^{s-1}}}\Big|\sum_{n\in\mathcal{N}}\sum_{\mathbf{j}\in\mathcal{J}}\sum_{\substack{p\in(2^{s-1},2^s]\\ pl\in n+\mathcal{I}_{n,\mathbf{j}}}}\beta(p)\widetilde{F}_{n,\mathbf{j}}(g_{n,\mathbf{j}}(pl)\Gamma_{n,\mathbf{j}})\Big|^2\\ =&\sum_{\mathbf{n}_1,\mathbf{n}_2\in\mathcal{N}\times\mathcal{J}}\sum_{p_1,p_2\in(2^{s-1},2^s]}\sum_{l\in\mathcal{A}_{\mathbf{n}_1,\mathbf{n}_2,p_1,p_2}}\beta(p_1)\overline{\beta(p_2)}\\ &\widetilde{F}_{\mathbf{n}_1}(g_{\mathbf{n}_1}(p_1l)\Gamma_{\mathbf{n}_1})\overline{\widetilde{F}_{\mathbf{n}_2}(g_{\mathbf{n}_2}(p_2l)\Gamma_{\mathbf{n}_2})}\end{aligned}$$

It will be useful to have an upper bound on the size of $\mathcal{A}_{\mathbf{n}_1,\mathbf{n}_2,p_1,p_2}$.

Lemma 6.5. *If $p_1>W$, then $\#\mathcal{A}_{\mathbf{n}_1,\mathbf{n}_2,p_1,p_2}\ll p_1^{-1}W^{-3}H$.*

Proof. For a prime $p>W$, p is coprime to $q\in(\frac{W}{2},W]$. The arithmetic progression $\mathcal{A}_{\mathbf{n},p}$ from (6.10) is bounded in length by

$$\#\mathcal{A}_{\mathbf{n},p}\le q^{-1}p^{-1}|I_k|\le 2p^{-1}W^{-1}W^{-2}H=2p^{-1}W^{-3}H.\tag{6.13}$$

The lemma follows because $\mathcal{A}_{\mathbf{n}_1,\mathbf{n}_2,p_1,p_2}=\mathcal{A}_{\mathbf{n}_1,p_1}\cap\mathcal{A}_{\mathbf{n}_2,p_2}$. □

We remark that, on the other hand, if $H\ge 4pW^3$, then we also have

$$\#\mathcal{A}_{\mathbf{n},p}\ge q^{-1}(p^{-1}|I_k|-1)-1\ge\frac12 q^{-1}p^{-1}|I_k|\ge\frac12 p^{-1}W^{-3}H.\tag{6.14}$$

We first take the sum when the length of $\mathcal{A}_{\mathbf{n}_1,\mathbf{n}_2,p_1,p_2}$ is bounded by $2^{-s}W^{-(B_2+3)}H$ where $B_2\ge 10$ and will be determined later. This part of (6.12) is easily bounded as below.

Proposition 6.6. *For $B_2 \geq 10$, the expression*

$$\sum_{\mathbf{n}_1,\mathbf{n}_2\in\mathcal{N}\times\mathcal{J}} \sum_{\substack{p_1,p_2\in(2^{s-1},2^s] \\ \#\mathcal{A}_{\mathbf{n}_1,\mathbf{n}_2,p_1,p_2}<2^{-s}W^{-(B_2+3)}H}} \sum_{l\in\mathcal{A}_{\mathbf{n}_1,\mathbf{n}_2,p_1,p_2}} \beta(p_1)\overline{\beta(p_2)} \widetilde{F}_{\mathbf{n}_1}(g_{\mathbf{n}_1}(p_1 l)\Gamma_{\mathbf{n}_1})\overline{\widetilde{F}_{\mathbf{n}_2}(g_{\mathbf{n}_2}(p_2 l)\Gamma_{\mathbf{n}_2}))} \tag{6.15}$$

satisfies $|(6.15)| \ll 2^s W^{-B_2} H^2 N$.

Proof.

$$\begin{aligned}|(6.15)| \leq & \Bigg|\sum_{\mathbf{n}_1,\mathbf{n}_2\in\mathcal{N}\times\mathcal{J}} \sum_{\substack{p_1,p_2\in(2^{s-1},2^s] \\ \#\mathcal{A}_{\mathbf{n}_1,\mathbf{n}_2,p_1,p_2}<2^{-s}W^{-(B_2+3)}H}} 2^{-s}W^{-(B_2+3)}H\Bigg| \\ \leq & 2^{-s}W^{-(B_2+3)}H \sum_{p_1,p_2\in(2^{s-1},2^s]} \sum_{\mathbf{n}_1,\mathbf{n}_2\in\mathcal{N}\times\mathcal{J}} 1_{\mathcal{A}_{\mathbf{n}_1,\mathbf{n}_2,p_1,p_2}\neq\emptyset} \\ \ll & 2^{-s}W^{-(B_2+3)}H\cdot 2^{2s}\cdot W^3 N\cdot H = 2^s W^{-B_2}H^2 N\end{aligned}$$

□

Here the last inequality follows from (4.4) and the lemma below.

Lemma 6.7. *If $2^s \geq W \geq 10$, then for all $\mathbf{n}_1 \in \mathcal{N}\times\mathcal{J}$ and $p_1, p_2 \in (2^{s-1}, 2^s]$,*

$$\#\{\mathbf{n}_2 \in \mathcal{N}\times\mathcal{J} : \ \mathcal{A}_{\mathbf{n}_1,\mathbf{n}_2,p_1,p_2} \neq \emptyset\} \ll H.$$

Proof. Notice that if in $\mathbf{n}_2 = (n_2, k_2, j_2)$, k_2 is given, then $\mathcal{A}_{\mathbf{n}_1,\mathbf{n}_2,p_1,p_2} \neq \emptyset$ implies $(\frac{n_1}{p_1} + \frac{1}{p_1}I_{k_1}) \cap (\frac{n_2}{p_2} + \frac{1}{p_2}I_{k_2}) \neq \emptyset$. This is true only if n_2 belongs to an interval whose length is at most

$$\frac{p_2}{p_1}|I_{k_1}| + |I_{k_2}| \leq 2W^{-2}H + W^{-2}H = 3W^{-2}H.$$

Moreover, the congruence class of elements in $\mathcal{A}_{\mathbf{n}_1,p_1}$ modulo q is determined by $\mathbf{n}_1$ and p_1. This congruence class, together with n_2 and p_2, in turn determines a unique choice of the remainder j_2 modulo q in order for $\mathcal{A}_{\mathbf{n}_1,\mathbf{n}_2,p_1,p_2} = \mathcal{A}_{\mathbf{n}_1,p_1} \cap \mathcal{A}_{\mathbf{n}_2,p_2}$.

Therefore, $\sum_{\mathbf{n}_2\in\mathcal{N}\times\mathcal{J}} 1_{\mathcal{A}_{\mathbf{n}_1,\mathbf{n}_2,p_1,p_2}\neq\emptyset} \ll \sum_{k_2\leq W^2} W^{-2}H = H.$ □

We now focus on intersections with $\#\mathcal{A}_{\mathbf{n}_1,\mathbf{n}_2,p_1,p_2} \geq 2^{-s}W^{-(B_2+3)}H$.

Definition 6.8. *For $s \in [s_-, s_+]$, $\mathbf{n}_1 \in \mathcal{N}\times\mathcal{J}$, prime number $p_1 \in (2^{s-1}, 2^s]$ and a parameter $B_2 \geq 10$, the set $\Omega_{s,\mathbf{n}_1,p_1,B_2}$ is defined to be the set of all configurations $(\mathbf{n}_2, p_2) \in \mathcal{N}\times\mathcal{J}\times(2^{s-1}, 2^s]$ such that:*

(i) p_2 *is prime;*
(ii) $\#\mathcal{A}_{\mathbf{n}_1,\mathbf{n}_2,p_1,p_2} \geq 2^{-s}W^{-(B_2+3)}H$;
(iii)

$$\Bigg|\sum_{l\in\mathcal{A}_{\mathbf{n}_1,\mathbf{n}_2,p_1,p_2}} \widetilde{F}_{\mathbf{n}_1}(g_{\mathbf{n}_1}(p_1 l - n_1)\Gamma_{\mathbf{n}_1})\overline{\widetilde{F}_{\mathbf{n}_2}(g_{\mathbf{n}_2}(p_2 l - n_1)\Gamma_{\mathbf{n}_2})}\Bigg| \geq W^{-B_2}\#\mathcal{A}_{\mathbf{n}_1,\mathbf{n}_2,p_1,p_2}.$$

Proposition 6.9. *One can choose the constant $C_0 = O(1) \geq 10$ to be sufficiently large, such that: if*

$$W \geq 10, B_2 \geq 10, B_1 \geq C_0 B_2, H \geq \max(W^{B_1}, 2^{10s}), \tag{6.16}$$

then for all pairs $(\mathbf{n}_1, p_1)$, where $\mathbf{n}_1 \subset \mathcal{N} \times \mathcal{J}$ and $p_1 \in (2^{s-1}, 2^s]$,

$$\#\Omega_{s,\mathbf{n}_1,p_1,B_2} < 2^s W^{-B_2} H.$$

The proof of the proposition is postponed to the next section.

Proposition 6.10. *In the settings of Proposition 6.9, the expression*

$$\begin{aligned} \sum_{\mathbf{n}_1,\mathbf{n}_2 \in \mathcal{N}\times\mathcal{J}} \sum_{\substack{p_1,p_2 \in (2^{s-1},2^s] \\ \#\mathcal{A}_{\mathbf{n}_1,\mathbf{n}_2,p_1,p_2} \geq 2^{-s} W^{-(B_2+3)} H}} \sum_{l \in \mathcal{A}_{\mathbf{n}_1,\mathbf{n}_2,p_1,p_2}} \beta(p_1)\overline{\beta(p_2)} \\ \widetilde{F}_{\mathbf{n}_1}(g_{\mathbf{n}_1}(p_1 l - n_1)\Gamma_{\mathbf{n}_1})\overline{\widetilde{F}_{\mathbf{n}_2}(g_{\mathbf{n}_2}(p_2 l - n_2)\Gamma_{\mathbf{n}_2})} \end{aligned} \tag{6.17}$$

satisfies $|(6.17)| \ll 2^s W^{-B_2} H^2 N$.

Proof. As $|\beta| \leq 1$ and $\|\widetilde{F}_{\mathbf{n}}\|_{C^0} \leq 2$ for all $\mathbf{n}$, in $|(6.17)|$, using Lemma 6.5 and Proposition 6.9, the contribution from configuration with $(\mathbf{n}_2, p_2) \in \Omega_{s,\mathbf{n}_1,p_1,B_2}$ is bounded by

$$\begin{aligned} &(\#\mathcal{N} \cdot \#\mathcal{J}) \cdot 2^s \cdot (\max_{\mathbf{n}_1,p_1} \#\Omega_{s,\mathbf{n}_1,p_1,B_2})(\max_{\mathbf{n}_1,\mathbf{n}_2,p_1,p_2} \#\mathcal{A}_{\mathbf{n}_1,\mathbf{n}_2,p_1,p_2}) \cdot 4 \\ \ll& NW^3 \cdot 2^s \cdot 2^s W^{-B_2} H \cdot 2p^{-1} W^{-3} H \\ \ll& 2^s W^{-B_2} H^2 N. \end{aligned} \tag{6.18}$$

And the contribution from out of this collection is bounded, thanks to Lemma 6.5, Lemma 6.7 and the construction of $\Omega_{s,\mathbf{n}_1,p_1,B_2}$, by

$$\begin{aligned} &(\#\mathcal{N} \cdot \#\mathcal{J}) \cdot 2^{2s} \cdot \max_{\mathbf{n}_1,p_1,p_2} \sum_{\substack{\mathbf{n}_2 \in \mathcal{N}\times\mathcal{J} \\ \mathcal{A}_{\mathbf{n}_1,\mathbf{n}_2,p_1,p_2} \neq \emptyset}} \\ &\quad \Big| \sum_{l \in \mathcal{A}_{\mathbf{n}_1,\mathbf{n}_2,p_1,p_2}} \widetilde{F}_{\mathbf{n}_1}(g_{\mathbf{n}_1}(p_1 l - n_1)\Gamma_{\mathbf{n}_1})\overline{\widetilde{F}_{\mathbf{n}_2}(g_{\mathbf{n}_2}(p_2 l - n_1)\Gamma_{\mathbf{n}_2})} \Big| \\ \ll& NW^3 \cdot 2^{2s} \cdot H \cdot W^{-B_2} \max_{\mathbf{n}_1,\mathbf{n}_2,p_1,p_2} \#\mathcal{A}_{\mathbf{n}_1,\mathbf{n}_2,p_1,p_2} \\ \ll& NW^3 \cdot 2^{2s} \cdot H \cdot W^{-B_2} 2^{-s} W^{-3} H \\ =& 2^s W^{-B_2} H^2 N. \end{aligned} \tag{6.19}$$

The lemma follows by combining these two bounds. ☐

Now adding up the estimates from Propositions 6.6 and 6.10 leads to the proof of Proposition 6.1.

Proof of Proposition 6.1. By Propositions 6.6 and 6.10, when C_0 is sufficiently large, under assumptions (6.16), we have

$$\begin{aligned} (6.9) \ll& 2^{-\frac{s}{2}} N^{\frac{1}{2}} (6.12)^{\frac{1}{2}} \leq 2^{-\frac{s}{2}} N^{\frac{1}{2}} ((6.15) + (6.17))^{\frac{1}{2}} \\ \ll& 2^{-\frac{s}{2}} N^{\frac{1}{2}} \cdot 2^{\frac{s}{2}} W^{-\frac{B_2}{2}} H N^{\frac{1}{2}} \\ =& W^{-\frac{B_2}{2}} HN. \end{aligned} \tag{6.20}$$

Hence,

$$|(6.7)| = |(6.8)| \leq \sum_{s\in(s_-,s_+]} (6.9) \leq s_+ W^{-\frac{B_2}{2}} HN, \tag{6.21}$$

and by Corollary 6.7,

$$\begin{aligned} |(6.6)| \leq & |(6.7)| + (2^{-s_-} + W^{-B_1})HN \\ \ll & (s_+ W^{-\frac{B_2}{2}} + 2^{-s_-} + W^{-B_1})HN. \end{aligned} \tag{6.22}$$

We now set the parameters s_-, s_+, B_1 and B_2. Let $s_+ = \lfloor \frac{1}{10} \log H \rfloor$. and $s_- = \lfloor 20\epsilon s_+ \rfloor$. This guarantees that $N - \#\mathcal{S} \ll \frac{s_-}{s_+} N \leq \epsilon N$. Moreover, $2^{-s_-} < H^{-\epsilon}$.

Assume in addition that $B_1 \geq 10C_0$ and let $B_2 = {C_0}^{-1} B_1$. The inequalities in (6.1), together with the fact that $W \in [R, R^{C_1 {B_1}^m}]$, imply $W^{B_1} < R^{C_1 {B_1}^{m+1}} < H^\epsilon < H$. This also implies for all $s \in (s_-, s_+)$, $2^s > 2^{s_-} > H^\epsilon > W$. So all conditions in (6.16) are verified.

(6.22) now yields

$$\begin{aligned} |(6.6)| \ll & (W^{-\frac{{C_0}^{-1} B_1}{2}} \log H + H^{-\epsilon} + W^{-B_1})HN \\ \ll & (W^{-\frac{{C_0}^{-1} B_1}{2}} \log H + H^{-\epsilon})HN. \end{aligned} \tag{6.23}$$

Finally, to complete the proof, one only needs to replace the value of the constant C_0 with $10C_0$. □

7. Proof of Proposition 6.9

This part contains the proof of Proposition 6.9 by contradiction. In the rest of Section 7, we will assume that t, s, $\mathbf{n}_1$, p_1 are all fixed. For brevity, we will replace the notations $\mathbf{n}_2$ and p_2 with $\mathbf{n}$ and p.

Because one may choose the constant C_0 as long as it depends only on m and d, instead of (6.16) we will assume instead:

$$2^s > W \geq 10, B_2 \geq 10, B_1 \geq 10{C_0}^2 B_2, H \geq \max(W^{B_1}, 2^{10s}), \tag{7.1}$$

In order to get contradiction, suppose for $\mathbf{n}_1 \in \mathcal{N} \times \mathcal{J}$ and $p_1 \in (2^{s-1}, 2^s]$,

$$\#\Omega_{s,\mathbf{n}_1,p_1,B_2} \geq 2^s W^{-B_2} H. \tag{7.2}$$

Let $(\mathbf{n}, p)$ be an element of $\Omega_{s,\mathbf{n}_1,p_1,B_2}$, then $p_1, p \geq 2^s > W \geq q$. By the proof of Lemma 6.5, as $\mathcal{A}_{\mathbf{n}_1,\mathbf{n},p_1,p}$ is the intersection of two finite arithmetic progressions $\mathcal{A}_{\mathbf{n}_1,p_1}$, $\mathcal{A}_{\mathbf{n},p}$ of step length q, it also has step length q itself whenever it is non-empty.

Since $\mathbf{n}_1$ and p_1 are fixed, the arithmetic progression $\mathcal{A}_{\mathbf{n}_1,p_1}$ can be parametrized as $\{qt + r : t \in [T]\}$ for some $r \in \mathbb{Z}$. Here by (6.13)

$$T = \#\mathcal{A}_{\mathbf{n}_1,p_1} \leq 4 \cdot 2^{-s} W^{-3} H. \tag{7.3}$$

When $(\mathbf{n}, p) \in \Omega_{s,\mathbf{n}_1,p_1,B_2}$, the subsequence $\mathcal{A}_{\mathbf{n}_1,\mathbf{n},p_1,p}$ has the form $\{qt + r : t \in \mathcal{A}'_{\mathbf{n},p}\}$ where $\mathcal{A}'_{\mathbf{n},p}$ is a subinterval of integers in $[T]$ of length $\#\mathcal{A}_{\mathbf{n}_1,\mathbf{n},p_1,p} \geq 2^{-s} W^{-B_2} H$.

The conditions (ii) and (iii) on $\Omega_{s,\mathbf{n}_1,p_1,B_2}$ in Definition 6.8 can be rewritten as

$$\mathcal{A}'_{\mathbf{n},p} \geq 2^{-s} W^{-(B_2+3)} H \tag{7.4}$$

and

$$\left|\sum_{t\in\mathcal{A}'_{\mathbf{n},p}} \widetilde{F}_{\mathbf{n}_1}(g_{\mathbf{n}_1}(p_1(qt+r)-n_1)\Gamma_{\mathbf{n}_1})\overline{\widetilde{F}_{\mathbf{n}}(g_{\mathbf{n}}(p(qt+r)-n)\Gamma_{\mathbf{n}})}\right| \geq W^{-B_2}\#\mathcal{A}'_{\mathbf{n},p} \tag{7.5}$$

For every configuration $(\mathbf{n},p)=(n,\mathbf{j},p)=(n,k,j,p)\in\Omega_{s,\mathbf{n}_1,p_1,B_2}$. Define polynomial sequences $g_{\mathbf{n},p},\tilde{g}_{\mathbf{n},p}:\mathbb{Z}\to G_{\mathbf{n}_1}\times G_{\mathbf{n}}$ by

$$g_{\mathbf{n},p}(l)=\big(g_{\mathbf{n}_1}(p_1l-n_1),g_{\mathbf{n}}(pl-n)\big);\ \tilde{g}_{\mathbf{n},p}(t)=g_{\mathbf{n},p}(qt+r). \tag{7.6}$$

Note that the definition of $\tilde{g}_{\mathbf{n},p}$ depends on the choice of $\mathbf{n}$.

Then $g_{\mathbf{n},p},\tilde{g}_{\mathbf{n},p}\in\mathrm{Poly}(\mathbb{Z},(G_{\mathbf{n}_1})_\bullet\times(G_{\mathbf{n}})_\bullet)$. From (4.21), (7.3), (7.4) and (7.5), we know that the sequence $(\tilde{g}_{\mathbf{n},p}(t)(\Gamma\times\Gamma))_{t\in\mathcal{A}'_{\mathbf{n},p}}$ is not totally $2^{-2}W^{-B_2}$-equidistributed in $(G_{\mathbf{n}_1}/\Gamma_{\mathbf{n}_1})\times(G_{\mathbf{n}}/\Gamma_{\mathbf{n}})$. Then by Lemma 2.10, for a shorter length $T'_{\mathbf{n},p}\geq 2^{-5}W^{-2B_2}T$, the sequence $(\tilde{g}_{\mathbf{n},p}(t)(\Gamma\times\Gamma))_{t\in[T'_{\mathbf{n},p}]}$ fails to be $2^{-5}W^{-2B_2}$-equidistributed in $(G_{\mathbf{n}_1}/\Gamma_{\mathbf{n}_1})\times(G_{\mathbf{n}}/\Gamma_{\mathbf{n}})$.

By Proposition 3.1, there exists a horizontal character $\eta_{\mathbf{n},p}$ of $(G_{\mathbf{n}_1}/\Gamma_{\mathbf{n}_1})\times(G_{\mathbf{n}}/\Gamma_{\mathbf{n}})$ such that

$$0<|\eta_{\mathbf{n},p}|<W^{O(B_2)} \tag{7.7}$$

and $\|\eta_{\mathbf{n},p}\circ\tilde{g}_{\mathbf{n},p}\|_{C^\infty([T'_{\mathbf{n},p}])}\ll W^{O(B_2)}$. As $T'_{\mathbf{n},p}\gg W^{-2B_2}T$, this implies that

$$\|\eta_{\mathbf{n},p}\circ\tilde{g}_{\mathbf{n},p}\|_{C^\infty([T])}\ll W^{O(B_2)}. \tag{7.8}$$

Here the norm $|\eta_{\mathbf{n},p}|$ is measured in terms of the Mal'cev basis $\mathcal{V}_{\mathbf{n}}\cup\mathcal{V}_{\mathbf{n}'}$, where $\mathcal{V}_{\mathbf{n}}=\mathcal{V}_{n,\mathbf{j}}$ and $\mathcal{V}_{\mathbf{n}_1}=\mathcal{V}_{n_1,\mathbf{j}_1}$ are defined in Section 4.

Recall from our construction in Section 4 that the sequences $G_{\mathbf{n}}$, $\Gamma_{\mathbf{n}}$, $\mathcal{V}_{\mathbf{n}}$ are determined by $\gamma_{\mathbf{n}}$, which in turn depends only on the variables n, j in $\mathbf{n}=(n,k,j)$ and is q-periodic in n. So there are γ_*, G_*, Γ_*, $\mathcal{V}_*$ such that for at least $q^{-2}\#\Omega_{s.\mathbf{n}_1,p_1}$ choices of $(\mathbf{n},p)\in\Omega_{s,\mathbf{n}_1,p_1,B_2}$,

$$(\gamma_{\mathbf{n}},G_{\mathbf{n}},\Gamma_{\mathbf{n}},\mathcal{V}_{\mathbf{n}})=(\gamma_*,G_*,\Gamma_*,\mathcal{V}_*). \tag{7.9}$$

Note that the choices of horizontal characters satisfying (7.13) is bounded by $W^{O(B_2)}$. Given (7.2) and that $q\leq W$, by pigeonhole principle, we can find some horizontal character η of $(G_{\mathbf{n}_1}/\Gamma_{\mathbf{n}_1})\times(G_*/\Gamma_*)$ such that for a set Ω_* of at least $2^sW^{-O(B_2)}H$ choices of $(\mathbf{n},p)\in\Omega_{s,\mathbf{n}_1,p_1,B_2}$, (7.9) holds and $\eta_{\mathbf{n},p}=\eta$.

Therefore,

$$\|\eta\circ\tilde{g}_{\mathbf{n},p}\|_{C^\infty[T]}\ll W^{O(B_2)} \tag{7.10}$$

holds for at least $2^sW^{-O(B_2)}H$ choices of $(\mathbf{n},p)\in\Omega_{s,\mathbf{n}_1,p_1,B_2}$. In particular, because of the fact $\#\mathcal{J}\leq W^3$ and Lemma 6.7, there is a set $\mathcal{P}_{s,\mathbf{n}_1,p_1}\subseteq\{p\text{ prime:}p\in(2^{s-1},2^s]\}$ of size

$$\#\mathcal{P}_{s,\mathbf{n}_1,p_1}\gg 2^sW^{-O(B_2)}, \tag{7.11}$$

such that for all $p\in\mathcal{P}_{s,\mathbf{n}_1,p_1}$, there are at least $W^{-O(B_2)}H$ choices of n, such that for some $\mathbf{j}$, the configuration $\mathbf{n}=(n,\mathbf{j})$ satisfies $(\mathbf{n},p)\in\Omega_{s,\mathbf{n}_1,p,B_2}$ and (7.10).

Recall that $g_{\mathbf{n}}(h)=\gamma_{\mathbf{n}}^{-1}g'(n,h)\gamma_{\mathbf{n}}$. So for the polynomial $g_*(n,h)=\gamma_*^{-1}g'(n,h)\gamma_*$ and every $(\mathbf{n},p)\in\Omega_{s,\mathbf{n}_1,p_1,B_2}$, $g_{\mathbf{n}}(h)=g_*(n,h)$ where n is the first coordinate of $\mathbf{n}=(n,k,j)$. In this case,

$$\tilde{g}_{\mathbf{n},p}(t)=\big(g_{\mathbf{n}_1}(p_1(qt+r)-n_1),g_*(n,p(qt+r)-n)\big). \tag{7.12}$$

Write $\eta = \eta_{(1)} \oplus \eta_{(2)}$, where $\eta_{(1)}$ and $\eta_{(2)}$ are respectively horizontal characters of $G_{\mathbf{n}_1}/\Gamma_{\mathbf{n}_1}$ and G_*/Γ_* and at least one of them is non-zero. Then $\eta_{(1)} \circ g_{\mathbf{n}_1} : \mathbb{Z} \to \mathbb{R}$ and $\eta_{(1)} \circ g_* : \mathbb{Z}^2 \to \mathbb{R}$ are polynomials of total degree bounded by d, where d is the step of nilpotency of $G_\bullet$. As p_1, r, q, $\mathbf{n}_1$ are all fixed, one can write

$$\eta_{(1)} \circ g_{\mathbf{n}_1}(t) = \sum_{l=0}^{d} \alpha_l t^l. \tag{7.13}$$

$$\eta_{(1)} \circ g_*(n,h) = \sum_{\substack{l_1,l_2\geq 0\\ l_1+l_2\leq d}} \beta^*_{l_1,l_2} n^{l_1} h^{l_2}. \tag{7.14}$$

We now parametrize $\eta_{(2)} \circ g_*$ in a better way. When $(\mathbf{n}, p) \in \Omega_{s,\mathbf{n}_1,p_1,B_2}$, $\mathcal{A}_{\mathbf{n}_1,n,p_1,p} \neq \emptyset$. So we can fix an $t_0 = t_0(\mathbf{n}, p) \in [T]$ such that $p(qt_0 + r) - n \in \mathcal{I}_\mathbf{n} \subset [H]$. On the other hand, because $t_0 \leq T = \#\mathcal{A}_{\mathbf{n}_1,p_1}$, by (6.13), $0 < pqt_0 \leq 2pq \cdot q^{-1}p_1^{-1}W^{-2}H \leq 4W^{-2}H$. Thus $pr - n \in [-4W^{-2}H, H] \subseteq (-H, H]$. We will write $b = pr - n + H$. Then $b \in [2H]$. For $u \in \mathbb{Z}$, we can write

$$\begin{aligned} &\eta_{(2)} \circ g_*(n, qu + pr - n)\\ =&\eta_{(2)} \circ g_*(pr + H - b, qu + b - H)\\ =&\sum_{\substack{l_1,l_2\geq 0\\ l_1+l_2\leq d}} \beta^*_{l_1,l_2}(pr + H - b)^{l_1}(qu + b - H)^{l_2}\\ =:&\sum_{\substack{l_1,l_2,i\geq 0\\ l_1+l_2+i\leq d}} \beta_{l_1,l_2,i}\, p^{l_1} u^{l_2} b^i \end{aligned} \tag{7.15}$$

In particular, for $u = pt$, we have

$$\begin{aligned} &\eta_{(2)} \circ g_*(n, p(qt + r) - n)\\ =&\eta_{(2)} \circ g_*(pr + H - b, q(pt) + b - H)\\ =&\sum_{\substack{l_1,l_2,i\geq 0\\ l_1+l_2+i\leq d}} \beta_{l_1,l_2,i}\, p^{l_1}(pt)^{l_2} b^i\\ =&\sum_{l=0}^{d}\sum_{l'=l}^{d}\sum_{i=0}^{d-l'} \beta_{l'-l,l,i}\, p^{l'} b^i t^l \end{aligned} \tag{7.16}$$

then

$$\eta \circ \tilde{g}_{\mathbf{n},p}(t) = \sum_{l=0}^{d}(\alpha_l + \sum_{l'=l}^{d}\sum_{i=0}^{d-l'} \beta_{l'-l,l,i}\, p^{l'} b^i) t^l, \tag{7.17}$$

where the coefficients $\beta_{l'-l,l,i}$ are independent of p,b and t (but depend on $\mathbf{n}_1$, p_1 and H).

The earlier discussion asserts that for all $p \in \mathcal{P}_{s,\mathbf{n}_1,p_1}$, there are is a subset $\mathcal{B}_{s,\mathbf{n}_1,p_1,p} \subseteq [2H]$ whose size satisfies

$$\#\mathcal{B}_{s,\mathbf{n}_1,p_1,p} \gg W^{-O(B_2)} H \tag{7.18}$$

such that for all $b \in \mathcal{B}_{s,\mathbf{n}_1,p_1,p}$, $\|(7.17) \pmod{\mathbb{Z}}\|_{C^\infty([T])} \ll W^{O(B_2)}$, where (7.17) is regarded as a polynomial in t.

For such pairs (p, b), by Lemma 2.3 and (7.3), we can find a positive integer $Z_1 \ll O(1)$ such that for all $0 \le l \le d$,

$$\Big\| Z_1(\alpha_l + \sum_{l'=l}^{d} \sum_{i=0}^{d-l'} \beta_{l'-l,l,i} p^{l'} b^i) \Big\|_{\mathbb{R}/\mathbb{Z}} \ll W^{O(B_2)} T^{-l} \ll 2^{ls} W^{O(B_2)} H^{-l}. \tag{7.19}$$

By using pigeonhole principle, one can make Z_1 independent of b after substituting $\mathcal{B}_{s,\mathbf{n}_1,p_1,p}$ with a smaller subset whose size still satisfies the lower bound (7.11).

We now view $Z_1(\alpha_l + \sum_{l'=l}^{d} \sum_{i=0}^{d-l'} \beta_{l'-l,l,i} p^{l'} b^i)$ as a polynomial of b. Applying Lemma 2.4 (with $\epsilon = 2^{ls} W^{O(B_2)} H^{-l}$ and $\delta = W^{-O(B_2)}$), we reduce from (7.19) that there is a positive integer $Z_2 \ll W^{O(B_2)}$ such that

$$\Big\| Z_2 Z_1(\alpha_l + \sum_{l'=l}^{d} \sum_{i=0}^{d-l'} \beta_{l'-l,l,i} p^{l'} b^i) \ (\mathrm{mod}\ \mathbb{Z}) \Big\|_{C^\infty[2H]} \ll 2^{ls} W^{O(B_2)} H^{-l}, \tag{7.20}$$

Again by Lemma 2.3, for all $p \in \mathcal{P}_{s,\mathbf{n}_1,p_1}$, there is a positive integer $Z_3 \ll O(1)$, such that for all $i \ge 1$, $l \ge 0$ such that $i + l \le d$,

$$\Big\| Z_3 Z_2 Z_1 \sum_{l'=l}^{d-i} \beta_{l'-l,l,i} p^{l'} \Big\|_{\mathbb{R}/\mathbb{Z}} \ll 2^{ls} W^{O(B_2)} H^{-i-l}. \tag{7.21}$$

And, when $i = 0$, for all $0 \le l \le d$,

$$\Big\| Z_3 Z_2 Z_1(\alpha_l + \sum_{l'=l}^{d} \beta_{l'-l,l,0} p^{l'}) \Big\|_{\mathbb{R}/\mathbb{Z}} \ll 2^{ls} W^{O(B_2)} H^{-l}. \tag{7.22}$$

Lemma 2.4 applies again, with respect to the variable $p \in [2^s]$, with $\epsilon = 2^{ls} W^{O(B_2)} H^{-l}$, $\delta = W^{-O(B_2)}$, and yields a positive integer $Z_4 \ll W^{O(B_2)}$ that:

For all $i \ge 1$, $0 \le l \le d$ subject to $i + l' \le d$,

$$\Big\| Z_4 Z_3 Z_2 Z_1 \sum_{l'=l}^{d-i} \beta_{l'-l,l,i} p^{l'} \ (\mathrm{mod}\ \mathbb{Z}) \Big\|_{C^\infty([2^s])} \ll 2^{ls} W^{O(B_2)} H^{-i-l}; \tag{7.23}$$

and for $i = 0$ and $0 \le l \le d$,

$$\Big\| Z_4 Z_3 Z_2 Z_1(\alpha_l + \sum_{l'=l}^{d} \beta_{l,l',0} p^{l'}) \ (\mathrm{mod}\ \mathbb{Z}) \Big\|_{C^\infty([2^s])} \ll 2^{ls} W^{O(B_2)} H^{-l}. \tag{7.24}$$

A final round of application of Lemma 2.3 tells us that, for a positive integer $Z_5 \ll O(1)$, the following properties hold:

For all $i \ge 1$, $0 \le l \le l' \le d$ subject to $i + l \le d$,

$$\Big\| Z_5 Z_4 Z_3 Z_2 Z_1 \beta_{l'-l,l,i} \Big\|_{\mathbb{R}/\mathbb{Z}} \ll 2^{(l-l')s} W^{O(B_2)} H^{-i-l}; \tag{7.25}$$

in addition, for $i = 0$ and $0 \le l \le l' \le d$ with $l' \ge 1$, (7.25) also holds.

Write $Z = Z_5 Z_4 Z_3 Z_2 Z_1$, which is an integer that is independent of b and t, and satisfies $Z \ll W^{O(B_2)}$. Thus the character $Z\eta_{(2)}$ satisfies

$$|Z\eta_{(2)}| \ll |Z| \cdot |\eta| \ll W^{O(B_2)}. \tag{7.26}$$

As we state in Notation 1.6, one choose a sufficiently large constant $C_0 = O(1) \ge 10$ which serves as the implicit constants both in the exponent of $W^{O(B_2)}$ of (7.25) and in (7.26). Now (7.25) writes

$$\Big\|Z\beta_{l'-l,l,i}\Big\|_{\mathbb{R}/\mathbb{Z}} \ll 2^{(l-l')s}W^{C_0B_2}H^{-i-l}; \tag{7.27}$$

In other words, the inequality

$$\Big\|Z\beta_{l_1,l_2,i}\Big\|_{\mathbb{R}/\mathbb{Z}} \ll 2^{-l_1s}W^{C_0B_2}H^{-i-l_2} \tag{7.28}$$

holds for all integer triples (l_1, l_2, i) such that $l_1, l_2, i \geq 0$, $l_1 + l_2 + i \leq d$ and l_1, l_2, i are not simultaneously equal to 0.

Lemma 7.1. *One can choose the constant $C_0 = C_0(m, d) \geq 10$ to be sufficiently large, such that :*

If (7.1) and (7.2) both hold then for every configuration $(\mathbf{n}, p) \in \Omega_{s,\mathbf{n}_1,p_1,B_2}$, the sequence $\{g_{\mathbf{n}}(h)\Gamma_{\mathbf{n}}\}_{h\in[H]}$ is not totally $W^{-C_0{}^{-1}B_1}$-equidistibuted in $G_{\mathbf{n}}/\Gamma_{\mathbf{n}}$.

Proof. Let r and b be as above. Set $\mathcal{U}_{\mathbf{n},p} = \{u \in \mathbb{Z} : qu + b - H \in [H]\}$. Then $\mathcal{U}_{\mathbf{n},p}$ is an interval of integers, whose length satisfies $\frac{H}{q} - 1 < \#\mathcal{U}_{\mathbf{n},p} < \frac{H}{q} + 1$. Moreover, as $0 < b \leq 2H$, every $u \in \mathcal{U}_{\mathbf{n},p}$ satisfies $|u| \leq \frac{2H}{q}$.

Fix any subinterval $\mathcal{U}'_{\mathbf{n},p} \subset \mathcal{U}_{\mathbf{n},p}$ of integers, that is of length $\lceil \frac{2W^{-2C_0B_2-3}H}{q} \rceil$. We note that because of (7.1), $\#\mathcal{U}'_{\mathbf{n},p} \geq 10$. Then for any $u_1, u_2 \in \mathcal{U}'$, by (7.16),

$$\begin{aligned}
&\|Z\eta_{(2)} \circ g_*(n, qu_1 + b - H) - Z\eta_{(2)} \circ g_*(n, qu_2 + b - H)\|_{\mathbb{R}/\mathbb{Z}} \\
=&\Big\|Z \sum_{\substack{l_1,l_2,i\geq 0\\ l_1+l_2+i\leq d}} \beta_{l_1,l_2,i}p^{l_1}b^i(u_1^{l_2} - u_2^{l_2})\Big\|_{\mathbb{R}/\mathbb{Z}} \\
=&\Big\|Z \sum_{\substack{l_1,l_2,i\geq 0\\ l_1+l_2+i\leq d}} \beta_{l_1,l_2,i}p^{l_1}b^i(u_1 - u_2)\sum_{h=0}^{l_2-1} u_1^h u_2^{l_2-1-h}\Big\|_{\mathbb{R}/\mathbb{Z}} \\
\ll& \sum_{\substack{l_1,i\geq 0;l_2\geq 1\\ l_1+l_2+i\leq d}} 2^{-l_1s}W^{C_0B_2}H^{-i-l_2} \cdot (2^s)^{l_1}(2H)^i\Big(\frac{W^{-2C_0B_2-3}H}{q}\Big)\Big(\frac{H}{q}\Big)^{l_2-1} \\
=& \sum_{\substack{l_1,i\geq 0;l_2\geq 1\\ l_1+l_2+i\leq d}} (W^{-C_0B_2-3})q^{-l_2} \\
\ll& W^{-C_0B_2}.
\end{aligned} \tag{7.29}$$

This implies that for the the mapping $\tilde{\eta}(x) = \exp(2\pi i Z\eta_{(2)}(x))$ from G/Γ to the unit circle in $\mathbb{C}$, the values of $\tilde{\eta}(g_{\mathbf{n}}(h))$ are within distance $\ll W^{-C_0B_2}$ to each other for $h \in \{qu + b - H : u \in \mathcal{U}'_{\mathbf{n},p}\}$. Again, using the convention in Notation 1.6, one can assume that the implicit constant here is C_0. In particular,

$$\Big|\mathop{\mathbb{E}}_{h\in\{qu+b-H:u\in\mathcal{U}'_{\mathbf{n},p}\}} \tilde{\eta}(g_{\mathbf{n}}(h)\Gamma_{\mathbf{n}})\Big| > 1 - C_0W^{-C_0B_2} \geq \frac{1}{2}, \tag{7.30}$$

as we assumed C_0, B_2 and W are all bounded by 10 from below. Because $Z\eta$ is a non-zero character, $\tilde{\eta}$ has zero average on $G_{\mathbf{n}}/\Gamma_{\mathbf{n}}$. In addition, $\|\tilde{\eta}\|_{G_{\mathbf{n}}/\Gamma_{\mathbf{n}}} \ll |Z\eta_{(2)}| \leq W^{C_0B_2}$.

Now note that $\{qu + b - H : u \in \mathcal{U}'_{\mathbf{n},p}\} \subseteq [H]$ is an arithmetic progression whose length is greater than $W^{-2C_0B_2-4}H$. It follows that the sequence $\{g_{\mathbf{n}}(h)\Gamma_{\mathbf{n}}\}_{h\in[H]}$ is not totally $\min(W^{-2C_0B_2-4}, \frac{1}{2}W^{-C_0B_2})$-equidistributed in $G_{\mathbf{n}}/\Gamma_{\mathbf{n}}$.

To finish the proof of Lemma 7.1, it suffices to notice that by the assumptions in (7.1), $\min(W^{-2C_0B_2-4}, \frac{1}{2}W^{-C_0B_2}) \geq W^{-C_0{}^{-1}B_1}$. □

Proof of Proposition 6.9. Recall that after redefining C_0 we may assume (7.1) instead of (6.16). By Lemma 7.1, and the construction of $\mathcal{N}$ in Lemma 4.3, if (7.2) holds, then for all $\mathbf{n} \in \Omega_{s,\mathbf{n}_1,p_1,B_2}$, $\mathbf{n} \notin \mathcal{N} \times \mathcal{J}$. This contradicts the definition of $\Omega_{s,\mathbf{n}_1,p_1,B_2}$, which requires $\mathbf{n} \in \mathcal{N} \times \mathcal{J}$. Therefore, (7.2) is false for all $\mathbf{n}_1 \in \mathcal{N} \times \mathcal{J}$; in other words, Proposition 6.9 is true. □

8. Proof of the main theorem

Theorem 1.2 will follow from

Theorem 8.1. *Suppose G is a connected, simply connected nilpotent Lie group and $\Gamma \subset G$ is a lattice. Assume that there exists an R_0-rational Mal'cev basis $\mathcal{V}$ of the Lie algebra G adapted to a nilpotent filtration $G_\bullet$ and the lattice Γ. Then there are constants $C, \epsilon_0 > 0$ that only depend on the dimension m of G, such that for all $g \in \mathrm{Poly}(\mathbb{Z}^2, G_\bullet)$, 1-bounded multiplicative function $\beta : \mathbb{N} \to \mathbb{C}$, and continuous function $F : G/\Gamma \to \mathbb{R}$, $H, N \in \mathbb{N}$, $\epsilon > 0$, if*

$$\max\Big(\frac{\log R_0}{\log H}, \frac{\log\log H}{\log H}\Big) < \epsilon < \epsilon_0;\ \log H < (\log N)^{\frac{1}{2}},. \tag{8.1}$$

then

$$\begin{aligned}&\frac{1}{HN}\sum_{n\le N}\Big|\sum_{h\le H}\beta(n+h)F(g(n,h)\Gamma)\Big| \\ &\ll\Big(H^{-\epsilon} + H^{C\epsilon}e^{-\frac{1}{2}\widetilde{M}(\beta,\frac{N}{H^{C\epsilon}},H^{C\epsilon})}\widetilde{M}(\beta,\frac{N}{H^{C\epsilon}},H^{C\epsilon})^{\frac{1}{2}} \\ &\qquad + H^{C\epsilon}(\log\frac{N}{H^{C\epsilon}})^{-\frac{1}{100}}\Big)HN.\end{aligned} \tag{8.2}$$

Proof. Let $B_1 = C_0$, $C_2 = C_1B_1{}^{m+1} = O(1)$ and $R = H^{C_2{}^{-1}\epsilon'}$. Combining Propositions 5.1 and 6.1, we know that if the following inequalities hold :

$$\frac{\log\log H}{\log H} < \epsilon' < \frac{1}{500}; H^{C_2{}^{-1}\epsilon'} \geq R_0 \geq 10;\ \log H < (\log N)^{\frac{1}{2}}. \tag{8.3}$$

then there exists a subset $\mathcal{S} \subseteq [0, N] \cap \mathbb{N}$, determined by H, N, and ϵ', with $N - \#\mathcal{S} \ll \epsilon' N$, such that

$$\begin{aligned}&\sum_{n\le N}\Big|\sum_{h\le H}\beta(n+h)F(g(n,h)\Gamma)\Big| \\ &\ll\Big(W^{-1}\log H + H^{-\epsilon'} + W^{-\frac{1}{4}} \\ &\qquad + W^2e^{-\frac{1}{2}\widetilde{M}(\beta,\frac{N}{W^5},W)}\widetilde{M}(\beta,\frac{N}{W^5},W)^{\frac{1}{2}} + W^2(\log\frac{N}{W^5})^{-\frac{1}{100}}\Big)HN \\ &\ll\Big(H^{-C_2{}^{-1}\epsilon'}\log H + H^{2\epsilon'}e^{-\frac{1}{2}\widetilde{M}(\beta,\frac{N}{H^{5\epsilon'}},H^{\epsilon'})}\widetilde{M}(\beta,\frac{N}{H^{5\epsilon'}},H^{\epsilon'})^{\frac{1}{2}} \\ &\qquad + H^{2\epsilon'}(\log\frac{N}{H^{5\epsilon'}})^{-\frac{1}{100}}\Big)HN.\end{aligned} \tag{8.4}$$

Here we used the fact that $W \in [R, R^{C_1B_1{}^m}] \subseteq [H^{C_2{}^{-1}\epsilon'}, H^{\epsilon'}]$, and that the function $\widetilde{M}(\beta, \frac{N}{W^5}, W)$ is decreasing in W. The set $\mathcal{S}$ is the union of both the exceptional sets from Propositions 5.1 and 6.1.

We now rewrite $\epsilon = \frac{1}{2}{C_2}^{-1}\epsilon'$ and assume $\epsilon > \frac{\log\log H}{\log H}$. Then $H^\epsilon > \log H$ and

$$H^{-{C_2}^{-1}\epsilon'}\log H = H^{-2\epsilon}\log H < H^{-\epsilon}.$$

Note that (8.3) implies (8.1). So (8.4) becomes

$$\begin{aligned}&\sum_{n\le N}\Big|\sum_{h\le H}\beta(n+h)F(g(n,h)\Gamma)\Big|\\ \ll&\Big(H^{-\epsilon}+H^{4C_2\epsilon}e^{-\frac{1}{2}\widetilde{M}(\beta,\frac{N}{H^{10C_2\epsilon}},H^{2C_2\epsilon})}\widetilde{M}(\beta,\frac{N}{H^{10C_2\epsilon}},H^{2C_2\epsilon})^{\frac{1}{2}}\\ &+H^{4C_2\epsilon}(\log\frac{N}{H^{10C_2\epsilon}})^{-\frac{1}{100}}\Big)HN.\end{aligned}\tag{8.5}$$

The theorem follows by letting $C = 10C_2$ and $\epsilon_0 = \frac{1}{1000C_2}$, which only on m and d. But as $d \le m$, the dependence on d can be suppressed. □

Proof of Theorem 1.2. First choose $R_0 \ge 10$ such that $\mathfrak{g}$ has an R_0-rational Mal'cev basis with respect to the lower central series filtration $G_\bullet$ and lattice Γ. We then fix H_0 such that $\log H_0 \ge R_0$.

Notice that $f(n,h) = g^{n+h}x \in G/\Gamma$ is a polynomial map from $\mathrm{Poly}(\mathbb{Z}^2, G_\bullet)$. Furthermore, in (8.1), $\max\left(\frac{\log R_0}{\log H}, \frac{\log\log H}{\log H}\right) = \frac{\log\log H}{\log H}$ for all $H > H_0$. Hence Theorem 8.1 can be applied. The output is (1.6) and (1.9), with

$$\delta(a,N) = a^C e^{-\frac{1}{2}\widetilde{M}(\beta,\frac{N}{a^C},a^C)}\widetilde{M}(\beta,\frac{N}{a^C},a^C)^{\frac{1}{2}} + a^C(\log\frac{N}{a^C})^{-\frac{1}{100}}.$$

We need to show $\lim\limits_{N\to\infty}\delta(a,N) = 0$ for all $a > 0$, which is equivalent to that

$$\lim_{X\to\infty}\widetilde{M}(\beta,X,Y) = \infty,\ \forall Y > 0.\tag{8.6}$$

When β is the Möbius function μ or the Liouville function λ, it is known that $\lim_{N\to\infty}\frac{1}{X}\sum_{n\le X}\beta(n)\chi(n) = 0$. By Halász's Theorem [Hal68], for any given Dirichlet character χ, $\lim\limits_{X\to\infty}\mathbb{D}(\beta\chi,1,X) = \infty$. Moreover, [MRT15, Lemma C.1], which is based on an argument of Granville and Soundararajan [GS07], guarantees that

$$\inf_{|t|\le X}\mathbb{D}(\beta\chi,n^{it},X) \ge \frac{1}{4}\min\left(\sqrt{\log\log X}, D(\beta\chi,1,X)\right) + O(1).$$

Therefore, for all Dirichlet characters χ, $M(\beta\chi,X)\to\infty$ as $X\to\infty$. This implies (8.6) by construction (5.3) of $\widetilde{M}$.

Finally, it remains to show (1.8). To see this, it suffices to notice that, because because $N > \exp((\log H)^2) = H^{\log H} > H\log H > H\epsilon^{-1}$,

$$\begin{aligned}&\frac{1}{HN}\left|\sum_{n=1}^{N}\Big|\sum_{l=n+1}^{n+H}1_{\mathcal{S}}\mu(l)F(g^lx)\Big| - \sum_{n=1}^{N}\Big|\sum_{l=n+1}^{n+H}\mu(l)F(g^lx)\Big|\right|\\ \le&\frac{1}{HN}\left|\sum_{n=1}^{N}\#((n,n+H]\backslash\mathcal{S})\right| \le \frac{1}{HN}\cdot H\#([N+H]\backslash\mathcal{S})\\ \ll&\frac{1}{N}(\epsilon N + H) \ll \epsilon.\end{aligned}$$

So (1.8) can be deduced from (1.9). □

References

[BSZ13] J. Bourgain, P. Sarnak, and T. Ziegler, *Disjointness of Möbius from horocycle flows*, From Fourier analysis and number theory to Radon transforms and geometry, Dev. Math., vol. 28, Springer, New York, 2013, pp. 67–83.

[Dav37] H. Davenport, *On some infinite series involving arithmetical functions II*, Quat. J. Math. **8** (1937), 313–320.

[EALdlR17] E. H. El Abdalaoui, M. Lemańczyk, and T. de la Rue, *Automorphisms with quasi-discrete spectrum, multiplicative functions and average orthogonality along short intervals*, International Mathematics Research Notices (2017), 4350–4368.

[FKPL18] S. Ferenczi, J. Kułaga-Przymus, and M. Lemańczyk, *Sarnak's conjecture: what's new*, Ergodic theory and dynamical systems in their interactions with arithmetics and combinatorics, 2018, pp. 163–235.

[FFKPL19] L. Flaminio, K. Frączek, J. Kułaga-Przymus, and M. Lemańczyk, *Approximate orthogonality of powers for ergodic affine unipotent diffeomorphisms on nilmanifolds*, Studia Math. **244** (2019), no. 1, 43–97.

[FH18] N. Frantzikinakis and B. Host, *The logarithmic Sarnak conjecture for ergodic weights*, Ann. of Math. (2) **187** (2018), no. 3, 869–931.

[GS07] A. Granville and K. Soundararajan, *Large character sums: pretentious characters and the Pólya-Vinogradov theorem*, J. Amer. Math. Soc. **20** (2007), no. 2, 357–384.

[GT12a] B. Green and T. Tao, *The quantitative behaviour of polynomial orbits on nilmanifolds*, Ann. of Math. (2) **175** (2012), no. 2, 465–540.

[GT12b] B. Green and T. Tao, *The Möbius function is strongly orthogonal to nilsequences*, Ann. of Math. (2) **175** (2012), no. 2, 541–566.

[Hal68] G. Halász, *Über die Mittelwerte multiplikativer zahlentheoretischer Funktionen*, Acta Math. Acad. Sci. Hungar. **19** (1968), 365–403 (German).

[Hua15] B. Huang, *Strong orthogonality between the Möbius function and nonlinear exponential functions in short intervals*, Int. Math. Res. Not. IMRN **23** (2015), 12713–12736.

[Hua16] B. Huang, *Exponential sums over primes in short intervals and an application to the Waring-Goldbach problem*, Mathematika **62** (2016), no. 2, 508–523.

[IK04] H. Iwaniec and E. Kowalski, *Analytic number theory*, American Mathematical Society Colloquium Publications, vol. 53, American Mathematical Society, Providence, RI, 2004.

[Kát86] I. Kátai, *A remark on a theorem of H. Daboussi*, Acta Math. Hungar. **47** (1986), no. 1-2, 223–225.

[Laz54] M. Lazard, *Sur les groupes nilpotents et les anneaux de Lie*, Ann. Sci. Ecole Norm. Sup. (3) **71** (1954), 101–190 (French).

[Lei98] A. Leibman, *Polynomial sequences in groups*, J. Algebra **201** (1998), no. 1, 189–206.

[Lei02] A. Leibman, *Polynomial mappings of groups*, Israel J. Math. **129** (2002), 29–60.

[Lei05] A. Leibman, *Pointwise convergence of ergodic averages for polynomial sequences of translations on a nilmanifold*, Ergodic Theory Dynam. Systems **25** (2005), no. 1, 201–213.

[Lei10] A. Leibman, *Orbit of the diagonal in the power of a nilmanifold*, Trans. Amer. Math. Soc. **362** (2010), no. 3, 1619–1658.

[Mal49] A. I. Mal′cev, *On a class of homogeneous spaces*, Izvestiya Akad. Nauk. SSSR. Ser. Mat. **13** (1949), 9–32 (Russian).

[MR16] K. Matomäki and M. Radziwiłł, *Multiplicative functions in short intervals*, Ann. of Math. (2) **183** (2016), no. 3, 1015–1056.

[MRT15] K. Matomäki, M. Radziwiłł, and T. Tao, *An averaged form of Chowla's conjecture*, Algebra Number Theory **9** (2015), no. 9, 2167–2196.

[MS19] K. Matomäki and X. Shao, *Discorrelation between primes in short intervals and polynomial phases*, preprint (2019).

[Rat91] M. Ratner, *Raghunathan's topological conjecture and distributions of unipotent flows*, Duke Math. J. **63** (1991), no. 1, 235–280.

[Sar09] P. Sarnak, *Three lectures on the Möbius function, randomness and dynamics*, lecture notes, IAS (2009).

[Sha] N. A. Shah, *Invariant measures and orbit closures on homogeneous spaces for actions of subgroups generated by unipotent elements*, Lie groups and ergodic theory

(Mumbai, 1996), Tata Inst. Fund. Res. Stud. Math., vol. 14, Tata Inst. Fund. Res., Bombay.

[Tao16] T. Tao, *The logarithmically averaged Chowla and Elliott conjectures for two-point correlations*, Forum Math. Pi **4** (2016), e8, 36.

[TT16] T. Tao and J. Teräväinen, *Odd order cases of the logarithmically averaged Chowla conjecture*, preprint (2016).

[Vau97] R. C. Vaughan, *The Hardy-Littlewood method*, 2nd ed., Cambridge Tracts in Mathematics, vol. 125, Cambridge University Press, Cambridge, 1997.

[Zha91] T. Zhan, *On the representation of large odd integer as a sum of three almost equal primes*, Acta Math. Sinica (N.S.) **7** (1991), no. 3, 259–272. A Chinese summary appears in Acta Math. Sinica **35** (1992), no. 4, 575.

Shandong University, Jinan, Shandong 250100, China
hexiaoguangsdu@gmail.com

Pennsylvania State University, University Park, PA 16802, USA
zhirenw@psu.edu