Discorrelation of multiplicative functions with nilsequences and its application on coefficients of automorphic $L$-functions

Xiaoguang He$^1$ | Mengdi Wang$^2$

1College of Mathematical Sciences, Sichuan University, Chengdu, Sichuan, China
2Department of Mathematics, KTH, Stockholm, Sweden

Correspondence
Mengdi Wang, Department of Mathematics, KTH, Stockholm 100 40, Sweden.
Email: mengdiw@kth.se

Funding information
National Postdoctoral Innovative Talents Support Program, Grant/Award Number: BX20190227; NSFC, Grant/Award Number: 12101427; Fundamental Research Funds for the Central Universities, Grant/Award Number: 2021SCU12109; Göran Gustafsson Foundation

Abstract
We introduce a class of multiplicative functions in which each function satisfies some statistic conditions, and then prove that the above functions are not correlated with finite degree polynomial nilsequences. Besides, we give two applications of this result. One is that the twisting of coefficients of automorphic $L$-function on $GL_m (m \geq 2)$ and polynomial nilsequences has logarithmic decay; the other is that the mean value of the Möbius function, coefficients of automorphic $L$-function, and polynomial nilsequences also has logarithmic decay.

MSC 2020
11L07, 11F30, 37A44 (primary)

1 | INTRODUCTION

Let $a(n), b(n) : \mathbb{N} \to \mathbb{C}$ be two complex valued sequences. Following the literature [9], we say they are “strongly orthogonal” if

$$\sum_{n \leq N} a(n)b(n) = O_A \left( (\log N)^{-A} \sum_{n \leq N} |a(n)b(n)| \right)$$

holds for any $A > 0$ and uniformly for $N \geq 2$. The study of the correlation between two sequences is an important subject in number theory. Here, we list some well-known pieces of literature, taking one of the sequences being the Möbius function as an example.
For \((a(n), b(n)) = (\mu(n), 1)\), the strong orthogonality \(\sum_{n \leq N} \mu(n) \ll Ne^{-c\sqrt{\log N}}\) is essentially equivalent to the prime number theorem.

For \((a(n), b(n)) = (\mu(n), e(n\alpha))\), this is a classical result due to Davenport [3], he proved a strongly orthogonal result by modifying Vinogradov’s method on bilinear forms.

For \((a(n), b(n)) = (\mu(n), e(n^k\alpha))\), Hua [11] generalizes Davenport’s strong orthogonality result from linear phase functions to polynomial phase functions.

For \((a(n), b(n)) = (\mu(n), \Phi(g(n)\Gamma))\), where \(\Phi(g(n)\Gamma)\) is a nilsequence, Green and Tao [9] prove that the Möbius function is strongly orthogonal to polynomial nilsequences.

In this paper, we consider the correlation of nilsequences with a class of multiplicative functions. And in the rest of this section, we would pay attention to the historical developments that are more relevant to the theme of this paper.

Suppose that \(f : \mathbb{N} \rightarrow \mathbb{C}\) is a 1-bounded multiplicative complex valued function, Daboussi [2] proves that for any irrational frequency \(\alpha \in \mathbb{R}/\mathbb{Z}(=\mathbb{T})\),

\[
S(f,\alpha) := \frac{1}{N} \sum_{n \leq N} f(n)e(n\alpha) = o(1).
\]

Montgomery and Vaughan [23] consider a much more general class of multiplicative functions. They relax the 1-bounded condition to the below two conditions

\[
|f(p)| \leq A, \text{ for all primes } p, \tag{1.1}
\]

and

\[
\sum_{n \leq N} |f(n)|^2 \leq A^2N, \text{ for large natural numbers } N, \tag{1.2}
\]

where \(A \geq 1\) is some constant. More formally, they prove that if \(\|q\alpha\| \leq \frac{1}{q}\) for some integer \(1 \leq q \leq N(\log N)^{-3}\), then for every function \(f\) satisfies conditions (1.1) and (1.2), we have

\[
S(f,\alpha) \ll \frac{1}{\sqrt{\phi(q)}} + \frac{1}{\log N}.
\]

Taking note that when \(q \geq (\log N)^{2+\varepsilon}\), the above upper bound would be determined by the second term. Thus, when the function \(f\) retains reasonable decay in arithmetic progressions with modulus less than \((\log N)^{2+\varepsilon}\), we can expect that for any frequency \(\alpha \in \mathbb{T}\),

\[
\frac{1}{N} \sum_{n \leq N} (f(n) - E_{[N]}f)e(n\alpha) \ll (\log N)^{-1}, \tag{1.3}
\]

where we have written \(E_{[N]}f = E_{n \in [N]}f(n) = \frac{1}{N} \sum_{n \in \mathbb{N}} f(n)\) as the average of \(f\) on the discrete interval \([N] = \{1, 2, ..., N\}\). Recently, Jiang, Lü, and Wang [19] weaken the above condition (1.1) that \(f\) takes bounded values in primes to the following two conditions

\[
\sum_{p \leq N} |f(p)|^2 \log p \ll N; \tag{1.4}
\]
and
\[ \sum_{\substack{p \leq N \ \text{is prime}}} |f(p)f(p + h)| \ll \frac{h}{\phi(h)} \frac{N}{(\log N)^2} \] (1.5)

holds for all positive integers \( h \). They can show that when \( f \) is under conditions (1.2), (1.4) and (1.5), the error term of \( S(f, \alpha) \) is similar to the result of Montgomery and Vaughan. Using this result, they also prove that the mean value of coefficients of automorphic \( L \)-functions and linear phase functions has logarithmic decay (see next subsection).

The studying of the correlation between multiplicative functions and polynomial nilsequence in place of the exponential function \( n \to e(n \alpha) \) begins with Green and Tao [8]. In that paper, they obtained the asymptotic formula of linear equations in primes using the information that \( W \)-tricked von Mangoldt function does not correlate with polynomial nilsequences. And very recently, Tao and Teräväinen [28] give the above asymptotic formula a quantitative error term.

To make this paper clear and readable, we are planning to record the definition of nilsequences and related notations here, and refer to [10] for more detailed introductions.

**Definition 1.1.** Let \( G \) be a connected, simply connected nilpotent Lie group, and let \( \Gamma \leq G \) be a lattice. By a filtration \( G_\ast = (G_i)_{i=0}^\infty \) on \( G \), we mean a descending sequence of groups \( G = G_1 \supseteq G_2 \supseteq \cdots \supseteq G_d \supseteq G_{d+1} = \{\text{id}_G\} \) such that
\[ [G, G_{i-1}] \subseteq G_i, \forall i \geq 2. \] (1.6)

This actually implies \([G_i, G_j] \subseteq G_{i+j}\) for all \( i, j \geq 1 \), and \( \Gamma_i := \Gamma \cap G_i \) is a lattice in \( G_i \) for \( i \geq 0 \). The number \( d \) is the degree of the filtration \( G_\ast \). The step \( s \) of \( G \) is the degree of the lower central filtration defined by \( G_{i+1} = [G, G_i] \). A lattice \( \Gamma \) must be cocompact, and the compact quotient \( G/\Gamma \) is called a nilmanifold.

We say that \( g \) is a polynomial sequence with coefficients in \( G_\ast \), and write \( g \in \text{poly}(\mathbb{Z}, G_\ast) \), if \( g : \mathbb{Z} \to G \) satisfies the derivative condition
\[ \partial_{h_1} \cdots \partial_{h_i} g(n) \in G_i \]
for all \( i \geq 0 \), \( n \in \mathbb{Z} \) and all \( h_1, \ldots, h_i \in \mathbb{Z} \), where \( \partial_h g(n) := g(n + h)g(n)^{-1} \) is the discrete derivative with shift \( h \).

The Mal’cev basis \( \mathcal{X} \) (see next section) induces a right invariant metric \( d_G \) on \( G \), which is the largest metric such that \( d(x, y) \leq |\psi_{\mathcal{X}}(xy^{-1})| \) always holds, where \( | \cdot | \) denotes the \( l^\infty \)-norm on \( \mathbb{R}^m \), and \( \psi_{\mathcal{X}} : G \to \mathbb{R} \) is the Mal’cev coordinate map which is defined in [10, (2.1)]. Actually, this in turn induces a metric \( d_{G/\Gamma} \) on \( G/\Gamma \). For a function \( F : G/\Gamma \to \mathbb{C} \), we define its Lipschitz norm as
\[ \|F\|_{\text{Lip}} = \|F\|_\infty + \sup_{x,y \in G/\Gamma, x \neq y} \frac{|F(x) - F(y)|}{d_{G/\Gamma}(x, y)} \] (1.7)
with respect to \( d_{G/\Gamma} \).

Finally, if \( F : G/\Gamma \to \mathbb{C} \) is a Lipschitz function (that is, \( \|F\|_{\text{Lip}} < \infty \)), we call a sequence of the form \( n \mapsto F(g(n)\Gamma) \) a nilsequence.

An interesting question is to generalize the above-described classical results on linear phase functions from Daboussi, Montgomery, Vaughan, and so on, to nilsequences. Frantzikinakis and
Host [6, Theorem 2.2] generalized Daboussi’s consideration to nilsequence cases, and they can prove a qualitative result as follows.

**Theorem 1.2** (Daboussi for nilsequences). Let $G/\Gamma$ be a nilmanifold with finite dimension and $G$, be a finite degree filtration of $G$. Suppose that $g \in \text{poly}(\mathbb{Z},G,)$ is a polynomial sequence and $(g(n)\Gamma)_{n \in \mathbb{N}}$ is totally equidistributed\(^*\) in $G/\Gamma$, then for any 1-bounded multiplicative function $f : \mathbb{Z} \to \mathbb{C}$ and any continuous function $F : G/\Gamma \to \mathbb{C}$ with $\int_{G/\Gamma} F = 0$, we have

$$\lim_{N \to \infty} \mathbb{E}_{n \in [N]} f(n)P(g(n)\Gamma) = 0.$$  

Matthiesen [21] considers a general class of functions which does not need the functions to take bounded values at every integer point but bound at primes and higher power of primes. Let $H \geq 1$ be a fixed number, $\mathcal{M}_H$ denote the class of multiplicative functions in which each function $f : \mathbb{N} \to \mathbb{C}$ satisfies the following conditions.

1. (Bound at prime powers). $|f(p^k)| \leq H^k$ for all prime powers $p^k$.
2. (Accumulation at primes is positive). There is a number $0 < \alpha_f \leq 1$ such that the below inequality holds for all large $N$

$$\frac{1}{N} \sum_{p \leq N} |f(p)| \log p \geq \alpha_f.$$  

Matthiesen [21] then gives a quantitative discorrelation of polynomial nilsequences with $W$-tricked functions from the class $\mathcal{M}_H$, see [21, Theorem 6.1] for precise statement.

We plan to study another class of multiplicative functions, for which we relax the point-to-point information at prime powers but instead we need stronger statistics information.

Let $\mathcal{M}'$ be the class of multiplicative functions $f : \mathbb{N} \to \mathbb{C}$ with the following properties.

1. (Equidistributed in arithmetic progressions with small moduli). There are relatively prime integers $1 \leq b \leq W \ll (\log N)^C$ such that the $W$-tricked function of $f$, named $f(W \cdot + b)$, is equidistributed in long arithmetic progressions. More formally,

$$\left| \frac{\phi(W)}{W} \mathbb{E}_{n \in P} \left( f(Wn + b) - \frac{\phi(W)}{WN} \sum_{n \in [N]} f(Wn + b) \right) \right| \ll \frac{1}{\log N}, \quad (1.8)$$

where $P \subseteq [N]$ is an arithmetic progression of length $|P| \gg N/(\log N)^C$.

2. (The $L^2$-norm at primes is bounded).

$$\sum_{p \leq N} |f(p)|^2 \log p \ll N. \quad (1.9)$$

\(^*\)An infinite sequence $(g(n)\Gamma)_{n \in \mathbb{N}}$ in $G/\Gamma$ is said to be totally equidistributed if for any integers $0 \neq a \in \mathbb{Z}$ and $r \in \mathbb{Z}$ and any continuous function $F : G/\Gamma \to \mathbb{C}$, we have

$$\lim_{n \to \infty} \mathbb{E}_{n \in [N]} F(g(an + r)\Gamma) = \int_{G/\Gamma} F.$$  

Here, the symbol $\int_{G/\Gamma}$ will stand for integration with respect to the unique left invariant probability measure on $G/\Gamma$. 

(3) (The $L^2$-norm is bounded).

$$\mathbb{E}_{n \in [N]} |f(n)|^2 \ll 1; \quad (1.10)$$

and

$$\frac{\phi(W)}{WN} \sum_{n \in N} |f(Wn + b')|^2 \ll 1 \quad (1.11)$$

for all integers $1 \leq b' \leq W$ coprime with $W$.

Note that [9, Proposition A.2] shows that the Möbius function is equidistributed in progressions with small common differences. It can also be seen from [28, Proposition 2.2] that $\Lambda - \Lambda_{\text{Siegel}}$ is equidistributed in progressions with small moduli, where $\Lambda$ is the von Mangoldt function and $\Lambda_{\text{Siegel}}$ is defined in [28, Definition 2.1]. When taking $W = 1$ in (1.8), the von Mangoldt function itself is not equidistributed in progressions with small moduli, but this difficulty can be overcome by a simple affine change of variables known as $W$-trick, which is introduced in [7]. However, for general multiplicative functions, finding such a number $W$ such that $f(W \cdot +b)$ eliminates bias to any residue class with small modulus is much more complicated, we refer the interested readers to [21, section 5] for the inspiring readings.

Remark. When $W \neq 1$ and $W \ll (\log N)^C$, for the function $f(W \cdot +b)$, the logarithmic decay in condition (1.8) is, somehow, also a strict condition, we can weaken this condition to expect a poor error term of Theorem 1.3. However, we are extremely interested in how to obtain a result similar to Montgomery and Vaughan (looks like (1.3)) in the nilsequence settings. And this is the reason that we do such a logarithmic decay hypothesis.

Remark. Assumption (1.11) may be strange at first glance, we try to defend a bit for it. The application of the approach of Montgomery and Vaughan would reduce the question to considering a sum over $pn$ in an arithmetic progression, where $p$ is a prime. Visually speaking, to handle $\sum_{p \leq x, \, pn \equiv b \pmod{W}} f(p)f(n)e(pn\alpha)$, where $f(n)$ and $f(p)$ may be seen as additive coefficients for the moment. When we aim to remove these coefficients and so to separate variables $p$ and $n$, taking note that the summation is on a reduced residue class, at least one of the variables $p$ and $n$ needs to run over all of the reduced residue classes modulus $W$, and this leads to condition (1.11) after applying Cauchy–Schwarz inequality. Moreover, it is easy to verify that when $W = 1$ the condition (1.11) is indeed (1.10).

**Theorem 1.3.** Let $G/\Gamma$ be a nilmanifold of dimension $m_G \geq 1$, $\mathcal{X}$ be a $M_0$-rational Mal’cev basis adapted to $G/\Gamma$ for some $2 \leq M_0 \leq \log N$, and $G$, be a filtration of $G$ of degree $d \geq 1$. Suppose that $g \in \text{poly}(\mathbb{Z},G_{\mathcal{X}})$ is a polynomial sequence and $F : G/\Gamma \to \mathbb{C}$ is a 1-bounded Lipschitz function, then for every function $f \in \mathcal{M}'$, we have

$$\frac{\phi(W)}{WN} \sum_{n \in [N]} (f(Wn + b) - \mathbb{E}_f (N; W, b))F(g(n)\Gamma) \ll_{m_G,d} (1 + \|F\|_{\text{Lip}}) \frac{1}{\log N},$$

where we have written $\mathbb{E}_f (N; W, b) = \frac{\phi(W)}{WN} \sum_{n \in [N]} f(Wn + b)$. 
Coefficient of automorphic $L$-functions twisted by nilsequences

Inspired from the argument of [19], we aim to apply Theorem 1.3 to the coefficients of automorphic $L$-functions. We begin with a brief introduction to the coefficients of automorphic $L$-functions.

Let $m \geq 2$ be an integer and $\pi$ be an automorphic irreducible cuspidal representation of $GL_m$ over $\mathbb{Q}$ with unitary central character. Denote by $\lambda_{\pi}(n)$ the Dirichlet coefficients of automorphic $L$-function $L(s, \pi)$ attached to $\pi$. Then $\lambda_{\pi}(n)$ is a multiplicative function. We now write $S_1(\alpha, N)$ as the sum of $\lambda_{\pi}$ with the linear phase function, that is,

$$S_1(\alpha, N) = \sum_{n \leq N} \lambda_{\pi}(n)e(n\alpha).$$

The Ramanujan conjecture asserts that

$$|\lambda_{\pi}(p)| \leq m$$

for all primes $p$, and this means that $\lambda_{\pi}(n)$ satisfy the condition (1.1) under Ramanujan conjecture. Besides, in the paper of Jiang, Lü, and Wang [19], they proved that $\lambda_{\pi}(n)$ also satisfy condition (1.2), thus, under Ramanujan conjecture, the result given by Montgomery and Vaughan can be applied to $\lambda_{\pi}(n)$. While, we also concern about the upper bound of $S_1(\alpha, N)$ without any sharp hypothesis.

When $m = 2$, $\lambda_{\pi}(n)$ are the normalized coefficients of a modular form or Maass form, and in this case, [12, Theorem 8.1] shows that

$$S_1(\alpha, N) \ll N^{1/2} \log N.$$

We now view $S_1(\alpha, N)$ as the Fourier coefficient of $\lambda_{\pi}$ at frequency $\alpha$ temporarily, then using of Parseval’s identity and the Rankin–Selberg theory yields that

$$\int_0^1 |S_1(\alpha, N)|^2 d\alpha = \sum_{n \leq N} |\lambda_{\pi}(n)|^2 \sim N,$$

which means that, on average, $S_1(\alpha, N)$ demonstrates square-root cancelation. As a consequence, one cannot expect above exponent $1/2$ could be reduced for all frequencies $\alpha$.

When $m = 3$, Miller [22] proved that

$$S_1(\alpha, N) \ll N^{3/4 + \varepsilon},$$

by using of the Voronoi summation for $GL_3$. For general $m \geq 4$, Jiang, Lü, and Wang [19, Theorem 1.2] proved that

$$S_1(\alpha, N) \ll \frac{N}{\log N}.$$
the coefficients $\lambda_\pi(n)$, thus, reduce the question to verifying conditions (1.2), (1.4), and (1.5) for coefficients $\lambda_\pi(n)$. And, the main business of their paper is to handle the much harder condition (1.5). It is worth pointing out that the logarithmic decay for $S_1(\alpha, N)$ in their result is derived from the approach of Montgomery and Vaughan. And we believe it is a difficult work to gain any power saving result for $S_1(\alpha, N)$, and also needs more clever ideas.

There are also several pieces of literature concerning the correlation of coefficients $\lambda_\pi(n)$ and polynomial phase functions even nilsequences, such as the sum

$$S_k(\alpha, N) = \sum_{n \leq N} \lambda_\pi(n) e(n^k \alpha).$$

Cafferata, Perelli, and Zaccagnini [1] extend the criterion of Bourgain–Sarnak–Ziegler by relaxing some conditions in the criterion to make it efficient for more multiplicative functions, then they applied this (generalized) criterion to holomorphic cusp form $\pi$ on $GL_2$ and prove that

$$S_k(\alpha, N) \ll N \frac{\log \log N}{\log N}.$$

Later, Jiang and Lü [15] extend the above result to $GL_m$ automorphic cuspidal representation $\pi$ under Hypothesis H (Hypothesis H was stated by Rudnick and Sarnak [25]). Afterward, the Hypothesis H is removed by Jiang and Lü in a later paper [16, Lemma 5.3].

As for general nilsequence settings, the only literature we can find is the paper of Matthiesen [21]. Thanks to her Theorem 6.1, one can show that, for each holomorphic cusp form $\pi$ on $GL_2$, the mean value of $\lambda_\pi(n)$ and polynomial nilsequences has some quantitative decay.

Our second result of this paper is to extend the previous consideration to the correlation of general $GL_m (m \geq 2)$ automorphic cuspidal representation $\pi$ over $\mathbb{Q}$ with polynomial nilsequence.

**Theorem 1.4.** Let $G/\Gamma$ be a nilmanifold of dimension $m_G \geq 1$, $\mathcal{X}$ be a $M_0$-rational Mal’cev basis adapted to $G/\Gamma$ for some $2 \leq M_0 \leq \log N$, and $G_n$ be a filtration of $G$ of degree $d \geq 1$. Suppose that $g \in \text{poly}(\mathbb{Z}, G, \cdot)$ is a polynomial sequence and $F : G/\Gamma \to \mathbb{C}$ is a $1$-bounded Lipschitz function, then we have

$$\frac{1}{N} \sum_{n \in [N]} \lambda_\pi(n) F(g(n)\Gamma) \ll_{m_G,d,\pi} (1 + \|F\|_{L^1}) \frac{1}{\log N}.$$

**Remark.** To prove Theorem 1.4, we need to verify $\lambda_\pi(n)$ satisfy the major-arc condition (1.8) with $W = 1$ and the $L^2$-norm bounded conditions (1.9) and (1.10). Thus, compared with the proof of [19], we lessen the work to verify $\lambda_\pi(n)$ satisfying the sieve condition (1.5).

We are willing to take some examples for the readers who are less familiar with nilsequences. Theorem 1.4 shows that the twisting of $\lambda_\pi(n)$ with a polynomial phase function of degree $d \geq 1$ has logarithmic decay, that is,

$$\sum_{n \leq N} \lambda_\pi(n) e(\alpha_d n^d + \cdots + \alpha_1 n + \alpha_0) \ll \frac{N}{\log N},$$
where $\alpha_d, \ldots, \alpha_1, \alpha_0 \in \mathbb{T}$ are arbitrary frequencies; Theorem 1.4 also shows that the following expression has logarithmic decay,

$$
\sum_{n \leq N} \lambda_\pi(n)e(\beta n \lfloor n \alpha \rfloor)\psi(\lfloor \alpha n \rfloor)\psi(\lfloor \beta n \rfloor) \ll \frac{N}{\log N},
$$

where $\lfloor \cdot \rfloor$ is the integer part operator, $\{\cdot\}$ is the fractional part operator, $\alpha, \beta \in \mathbb{T}$ are frequencies, and $\psi : [0, 1] \to \mathbb{C}$ is a Lipschitz function that vanishes near 0 and 1.

It is also an interesting question to consider whether the sequences $\{\mu(n)\}_n$ and $\{\lambda_\pi(n)F(g(n)\Gamma)\}_n$ are orthogonal or not. This is motivated by the Möbius pseudorandomness principle, which predicts that sums of the form $\sum_{n \leq \xi} \mu(n)\xi(n)$ should exhibit some cancelation if $\xi(n)$ is not obviously related to the prime factorization of $n$. See also the Möbius disjointness conjecture of Sarnak [26], which expects to use observables from zero entropy topological dynamical systems as the sequence $\xi(n)$. Our next conclusion shows that the Möbius pseudorandomness principle is true for the sequence $\{\lambda_\pi(n)F(g(n)\Gamma)\}$.

In this direction, the current recording results are from Fouvry and Ganguly [5] when $m = 2$ and Jiang and Lü [14] when $m = 3$, respectively. Combining their results, one has when $m = 2, 3$,

$$
\sum_{n \leq N} \mu(n)\lambda_\pi(n)e(n\alpha) \ll N \exp(-c\sqrt{\log N}),
$$

where $c > 0$ is an absolute constant. In the same paper of Cafferata, Perelli, and Zaccagnini [1], they also have

$$
\sum_{n \leq N} \mu(n)\lambda_\pi(n)e(n^k\alpha) \ll N \frac{\log \log N}{\sqrt{\log N}}.
$$

Jiang and Lü [15] extend the above result to $GL_m$ automorphic cuspidal representation $\pi$ under Hypothesis H and Hypothesis S. Afterward, these two hypotheses are removed in two later papers, respectively—Jiang and Lü [16, Lemma 5.3] removes Hypothesis H and Jiang, Lü, Thorner, and Wang [17, Corollary 4.7] removes Hypothesis S.

Recently, Jiang, Lü, and Wang [18]† prove that

$$
\sum_{n \leq N} \mu(n)\lambda_\pi(n)e(n\alpha) \ll \frac{N}{\log N}
$$

for $\pi$ is self-dual and $\pi \not\cong \pi \otimes \chi$ for any quadratic primitive character $\chi$. In this direction, we can show the following.

**Theorem 1.5.** Assume that $\pi$ is self-dual and $\pi \not\cong \pi \otimes \chi$ for any quadratic primitive character $\chi$, with the notations in Theorem 1.4, we have

$$
\sum_{n \leq N} \mu(n)\lambda_\pi(n)F(g(n)\Gamma) \ll m_{G,d,\pi} \frac{N}{\log N}.
$$

† We would like to thank Yujiao Jiang for posting their paper [18] and explaining their papers.
Remark. The assumption $\pi \not\equiv \pi \otimes \chi$ is in \cite[Theorem 1.2]{17}. Roughly speaking, Jiang, Lü, Thorner, and Wang can obtain an analogy of the Siegel theorem in higher rank groups based on this assumption.

Remark. Following the argument in \cite{5} and \cite{14}, we do not need the additional assumption on $\pi$ for $m = 2$ and $m = 3$, respectively.

2 | PRELIMINARIES AND OUTLINE OF THE PROOF

2.1 | Preliminaries

In this subsection, we quickly collect all the facts and notations that we will need from Green–Tao’s paper \cite{10}.

Denote by $g_\ast$ the Lie algebra $G_\ast$, then $g_\ast = \{g_i\}$ is a filtration of Lie algebras, that is, $[g, g_i] \subseteq g_{i+1}$, if and only if $G_\ast = \{G_i\}$ is a filtration.

Definition 2.1. \cite[Definition 2.1]{10} Let $G/\Gamma$ be an $m$-dimensional nilmanifold and let $G_\ast$ be a filtration. A basis $\mathcal{X} = \{X_1, ..., X_m\}$ for the Lie algebra $g$ over $\mathbb{R}$ is called a Mal’cev basis for $G/\Gamma$ adapted to $G_\ast$ if the following four conditions are satisfied.

(i) $\{X_j, X_{j+1}, ..., X_m\}$ spans an ideal of $g$ for all $0 \leq j \leq m$.

(ii) For each $1 \leq i \leq d$ and $m_i = \dim G_i$, the Lie algebra $g_i$ of $G_i$ is the linear span of $\{X_{m-m_i+1}, X_{m-m_i+2}, ..., X_m\}$.

(iii) There is a diffeomorphism $\psi_{\mathcal{X}} : G \rightarrow \mathbb{R}^m$ determined by

$$
\psi_{\mathcal{X}}(\exp(\omega_1 X_1) \cdots \exp(\omega_m X_m)) = (\omega_1, ..., \omega_m).
$$

(iv) In the coordinate system $\psi_{\mathcal{X}}, \Gamma = \psi_{\mathcal{X}}^{-1}(\mathbb{Z}^m)$.

We say that a Mal’cev basis $\mathcal{X}$ for $G/\Gamma$ is $Q$-rational if all of the structure constants $c_{ijk}$ in the relations

$$
[X_i, X_j] = \sum_k c_{ijk} X_k
$$

are rational with height at most $Q$. The height of a real number $x$ is defined as $\max(|a|, |b|)$ if $x = a/b$ is rational in reduced form.

Definition 2.2. \cite[Definition 1.2]{10} Let $G/\Gamma$ be a nilmanifold.

(1) Given a length $N > 0$ and an error tolerance $\delta > 0$, a finite sequence $(g(n)\Gamma)_{n \in [N]}$ is said to be $\delta$-equidistributed if we have

$$
\left| \mathbb{E}_{n \in [N]} F(g(n)\Gamma) - \int_{G/\Gamma} F \right| \leq \delta \|F\|_{\text{Lip}}
$$

for all Lipschitz functions $F : G/\Gamma \rightarrow \mathbb{C}$.
A finite sequence \((g(n)\Gamma)_{n \in [N]}\) is said to be \(\text{totally } \delta\text{-equidistributed}\) if we have

\[
\left| \mathbb{E}_{n \in P} F(g(n)\Gamma) - \int_{G/\Gamma} F \right| \leq \delta \|F\|_{\text{Lip}}
\]

for all Lipschitz functions \(F : G/\Gamma \to \mathbb{C}\) and all arithmetic progressions \(P \subset [N]\) of length at least \(\delta N\).

The symbol \(\int_{G/\Gamma}\) will stand for integration with respect to the unique left invariant probability measure on \(G/\Gamma\).

**Definition 2.3.** \([10, \text{Definition 1.17}]\) Let \(G/\Gamma\) be a nilmanifold and let \(Q > 0\) be a parameter. We say that \(\gamma \in G\) is \(Q\text{-rational}\) if \(\gamma^r \in \Gamma\) for some integer \(r, 0 < r \leq Q\). A \(Q\text{-rational point}\) is any point in \(G/\Gamma\) of the form \(\gamma \Gamma\) for some \(Q\)-rational group element \(\gamma\). A sequence \((\gamma(n))_{n \in \mathbb{Z}}\) is \(Q\text{-rational}\) if every element \(\gamma(n)\Gamma\) in the sequence is a \(Q\)-rational point.

**Definition 2.4.** \([10, \text{Definition 1.18}]\) Let \(G/\Gamma\) be a nilmanifold with a Mal’cev basis \(\mathcal{X}\). Let \((\varepsilon(n))_{n \in \mathbb{Z}}\) be a sequence in \(G\), and let \(M, N \geq 1\). We say that \((\varepsilon(n))_{n \in \mathbb{Z}}\) is \((M, N)\text{-smooth}\) if we have \(d(\varepsilon(n), id_G) \leq M\) and \(d(\varepsilon(n), \varepsilon(n - 1)) \leq M/N\) for all \(n \in [N]\).

With the above notations, we give the following factorization theorem. See \([10, \text{Theorem 1.19}]\)

**Proposition 2.5** (Green–Taot factorization theorem). Let \(m, d \geq 0\), and let \(M_0, N \geq 1\) and \(A > 0\) be real numbers. Suppose that \(G/\Gamma\) is an \(m\)-dimensional nilmanifold together with a filtration \(G_\ast\) of degree \(d\). Suppose that \(\mathcal{X}\) is an \(M_0\text{-rational}\) Mal’cev basis \(\mathcal{X}\) adapted to \(G_\ast\) and that \(g \in \text{poly}(\mathbb{Z}, G_\ast)\). Then there is an integer \(M\) with \(M_0 \leq M \ll M_0^C/\max(1, m, d(1))\), a rational subgroup \(G' \subseteq G\), a Mal’cev basis \(\mathcal{X}'\) for \(G'/\Gamma'\) in which each element is an \(M\)-rational combination of the elements of \(\mathcal{X}\), and a decomposition \(g = \varepsilon g' \gamma\) into polynomial sequences \(\varepsilon, g', \gamma \in \text{poly}(\mathbb{Z}, G_\ast)\) with the following properties.

1. \(\varepsilon : \mathbb{Z} \to G\) is \((M, N)\text{-smooth}\).
2. \(g' : \mathbb{Z} \to G'\) takes values in \(G'\), and the finite sequence \((g'(n)\Gamma')_{n \in [N]}\) is \(1/M^A\text{-equidistributed}\) in \(G'/\Gamma'\), using the metric \(d_{\mathcal{X}'}\) on \(G'/\Gamma'\).
3. \(\gamma : \mathbb{Z} \to G\) is \(M\text{-rational}\), and \((\gamma(n)\Gamma)_{n \in \mathbb{Z}}\) is periodic with period at most \(M\).

The next result is the key theorem to verify whether a sequence is equidistributed or not, but before it, we introduce the so-called smoothness norm.

**Definition 2.6.** \([10, \text{Definition 2.7}]\) Suppose that \(g : \mathbb{Z} \to \mathbb{T}\) is a polynomial sequence of degree \(d\). Then \(g\) may be written uniquely as

\[
g(n) = \alpha_0 + \alpha_1 \binom{n}{1} + \cdots + \alpha_d \binom{n}{d}.
\]

For any \(N > 0\), we define the smoothness norm as

\[
\|g\|_{C^\infty[N]} := \sup_{1 \leq j \leq d} N^j \|\alpha_j\|.
\]
Proposition 2.7 (Quantitative Leibman theorem). Let $m_G, d \geq 0$, $0 < \delta < 1/2$, and $N \geq 1$. Let $G/\Gamma$ be an $m_G$-dimensional nilmanifold together with a filtration of degree $d$ and a $1/\delta$-rational Mal’cev basis $\mathcal{X}$ adapted to this filtration. Suppose that $g \in \text{poly}(\mathbb{Z}, G)$, If $(g(n)\Gamma)_{n \in \mathbb{N}}$ is not $\delta$-equidistributed in $G/\Gamma$, then there is a horizontal character $\eta$ with $0 < |\eta| \ll \delta^{-O(m_G, d)}$ such that

$$\|\eta \circ g\|_{C^\infty[\mathbb{N}]} \ll \delta^{-O(m_G, d)}.$$ 

Proof. See [10, Theorem 2.9].

2.2 Motivation and outline of the argument

For convenience, in the rest of this paper we would allow all of the asymptotic notions, such as $O(\cdot), o(\cdot), \gg, \ll$ and so on, to depend on the degree $d$ and the dimension $m_G$.

We actually dawn on the ideas from the three papers [9], [21], and [23] to prove Theorem 1.3. As we said in introduction, Montgomery and Vaughan [23] deal with the correlation of a class of multiplicative functions with linear phase functions. For doing this, they first reduce the exponential sum $\mathbb{E}_{n \in \mathbb{N}} f(n) e(n \alpha)$ with $\alpha \neq 0$ to the one in which variables are in terms of primes $p$, that is $\mathbb{E}_{p \sim U} \mathbb{E}_{p \sim N} \log p f(p) f(n) e(p n \alpha)$. The using of Cauchy–Schwarz inequality shows that it is bounded by $\|f\|_{L^2[\mathbb{N}/U]}$ times a half power of

$$\mathbb{E}_{p, p' \sim U} \log p \log p' f(p) f(p') \mathbb{E}_{n \in \mathbb{N}}\left(\frac{n}{p} - \frac{n}{p'}\right) e((p - p')n \alpha).$$

Because the number of pairs $(p, p') \in (U, 2U)^2$ such that $p = p'$ is only $U$, one can show that in this case the above expression is pretty small so is negligible; while, Montgomery and Vaughan also show that the sequence $\{n(p - p')\alpha \text{ (mod 1)}\}_{p}$ is equidistributed for almost all pairs $(p, p')$ whenever $\alpha \neq 0$, and this leads them to a desired bound. It seems that their approach can be generalized to deal with the correlation of $f$ with polynomial phase functions if $f$ satisfies conditions (1.1) and (1.2). In what follows, we provide a heuristic argument for the generalization. Suppose that $k \geq 2$ is a natural number and $0 < \delta < 1$ is a parameter, we aim to show that the following inequality holds only if $\alpha$ is $\delta^{-O(1)}$-rational ($\|q\alpha\| \ll \delta^{-O(1)}/N^k$ for some $1 \leq q \ll \delta^{-O(1)}$)

$$\delta \ll \left|\mathbb{E}_{p \sim U} \mathbb{E}_{p \sim N} \log p f(p) f(n) e(p^k n^k \alpha)\right|.$$

First, it can be seen from Hölder’s inequality with some large integer $s = s(k)$ that

$$\delta \ll \left(\mathbb{E}_{n \sim N/U} |f(n)|^{2k} \right)^{\frac{2k-1}{2s}} \left(\mathbb{E}_{n \sim \mathbb{N}} \log p f(p) e(p^k n^k \alpha)\right)^{\frac{1}{2s}}.$$

The assumptions $f(p)$ is bounded for every prime $p$ and $\mathbb{E}_{n} |f(n)|^2 \ll 1$ (or a weaker assumption $\mathbb{E}_{n} |f(n)|^{\frac{2k}{2s-1}} \ll 1$) then lead us to

$$\delta^{2s} \ll \mathbb{E}_{p_1, \ldots, p_{2s} \sim U} \log p_1 \log p_2 \cdots \log p_{2s} \left|\mathbb{E}_{n \sim N/U} e\left((p_1^k + \cdots + p_s^k - \cdots - p_{2s}^k) n^k \alpha\right)\right|.$$
It follows from [27, Lemma 1.1.16] that there is \( q \ll \delta^{-O(1)} \) and a set \( P \subset (U, 2U)^{2s} \) with \( |P| \gg \delta^{O(1)} U^{2s} \) such that for each \( \vec{p} = (p_1, ..., p_{2s}) \in P \)

\[
\|q(p_1^k + \cdots + p_s^k - \cdots - p_{2s}^k)\alpha\| \ll \delta^{-O(1)} \left( \frac{U}{N} \right)^k.
\]

If we can show that \( \{p_1^k + \cdots + p_s^k - \cdots - p_{2s}^k\} \in P \) forms a dense (up to a factor \( \delta^{O(1)} \) ) subset of the interval \( [1, O(U^k)] \), the desired conclusion (\( \alpha \) is \( \delta^{-O(1)} \)-rational) would follow from [9, Lemma 3.4]. Practically, the fact that \( \{p_1^k + \cdots + p_s^k - \cdots - p_{2s}^k\} \in P \) is dense is a consequence of [9, Lemma 3.3] if \( s \) is large compared with the degree \( k \). In this paper, however, we are more interested in the correlation of multiplicative functions under weaker conditions (weaker than (1.1) and (1.2)).

In our understanding, the main result of Jiang, Lü, and Wang [19] is that they verify the function \( \lambda_\pi \) satisfies conditions (1.2–1.5), and then they apply the thought of Montgomery and Vaughan to show that \( \lambda_\pi \) does not correlate with linear phase functions. Fortunately, from their proof, we find that the function \( \lambda_\pi \) has good major-arc behavior, and this motivates us to study the (dis-)correlation of nilsequences with a class of functions that have good major-arc behavior (satisfy condition (1.8)).

The first step is to apply Green–Tao factorization theorem, Theorem 2.5, to decompose the polynomial \( g \) as a product \( \varepsilon g_\gamma \), where \( \varepsilon \) is a smooth factor, \( \gamma \) is rational, and \( g_\gamma \) is highly equidistributed in a closed subgroup \( G' \subseteq G \). We then eliminate the rather harmless factors \( \varepsilon \) and \( \gamma \) just as Green and Tao did in their [9, section 2]. This leaves us to deal with the highly equidistributed sequences with respect to arithmetic progressions with small moduli, say,

\[
\mathbb{E}_P \mathbb{E}_{n \in P} f(n) F_P(g_p(n) \Gamma_P),
\]

where \( P \) denotes those arithmetic progressions, and \( g_p \) is highly equidistributed for each \( P \). The next step is to utilize the approach of Montgomery and Vaughan, to split the function \( f(n) \) as a product \( f(p)f(n) \), where \( p \) is a prime number, so that we are in the position that

\[
\mathbb{E}_P \mathbb{E}_{p \in P} \log p f(p)f(n) F_P(g_p(pn) \Gamma_P).
\]

Here, we simplify the situation we would actually face, such as \( pn \) is in a reduced residue class \( pn \equiv b \mod W \). Unlike in the work of Matthiassen [21], our functions \( f \) do not have enough information in each progression \( P \) which fails us to deal with the summation over \( pn \) in progressions \( P \) piece by piece. We compromise with taking the above outside sum (counting progression) all the time, and view \( F_P \) as a piecewise function which supports on \( G/\Gamma \) and has very nice local properties in the sense that \( F_P(g_p(-\cdot) \Gamma_P) \) is equidistributed in each progression \( P \).

The application of Cauchy–Schwarz inequality allows us to transfer the matter to understanding the equidistribution of product of nilsequences, and the correlation of von Mangoldt function with product of nilsequences,

\[
\mathbb{E}_n F_P(g_p(pn) \Gamma_P) \overline{F_P(g_p(p'n) \Gamma_P)}
\]

and

\[
\mathbb{E}_p \Lambda(p) F_P(g_p(pn) \Gamma_P) \overline{F_P(g_p(p'n) \Gamma_P)}.
\]
We then take care of the two expressions by showing that almost all product sequences \((g_p(p \cdot), g_p(p' \cdot))\) and \((g_p(n \cdot), g_p(n' \cdot))\) are equidistributed whenever \(g_p\) is highly equidistributed, making use of the quantitative Leibman theorem (Proposition 2.7); and also by using bilinear form method adapted to polynomial nilsequences.

3 | REDUCING TO EQUIDISTRIBUTED CASES

As described in Section 2, we are going to factorize the polynomial sequence \(g\) into a smooth factor \(\epsilon\), a rational factor \(\gamma\) and a totally equidistributed factor \(g'\), then eliminate the rather harmless factors \(\epsilon\) and \(\gamma\) in the light of condition (1.8). This will allow us to focus on the highly equidistributed polynomial sequences. The overall strategy of the reduction follows from [9, section 2].

Without loss of generality, we may normalize the Lipschitz function \(F\) so that \(\|F\|_{\text{Lip}} = 1\), assume that \(A > 1\) is a large number to be specified, and the parameter \(M_0\) in Theorem 1.3 is equal to the maximal value, that is \(M_0 \geq \log N\).

We begin our calculation with applying Proposition 2.5 to the polynomial sequence \(g\) to find an integer \(M\) with \(M_0 \leq M \leq M_0^{O(1)}\), a rational subgroup \(G' \subseteq G\); a Mal’cev basis \(\mathcal{X}'\) adapted to \((G')\); for which elements are \(M\)-rational combinations of elements of \(\mathcal{X}\); and a decomposition \(g = \epsilon g' \gamma\), where \((\epsilon(n))_{n \in \mathbb{Z}}\) is \((M, N)\)-smooth, \((g'(n) \Gamma'))_{n \in \mathbb{Z}}\) is totally \(M^A\)-equidistributed in \(G'/\Gamma'\), and \((\gamma(n))_{n \in \mathbb{Z}}\) is periodic with period \(q \leq M\). Immediately, on writing \(g\) as the product \(\epsilon g' \gamma\), one has

\[
\frac{\phi(W)}{WN} \sum_{n \in [N]} \left( f(Wn + b) - \mathbb{E}_f(N; W, b) \right) F(g(n) \Gamma') = \frac{\phi(W)}{WN} \sum_{n \in [N]} \left( f(Wn + b) - \mathbb{E}_f(N; W, b) \right) F(\epsilon(n) g'(n) \gamma(n) \Gamma').
\]

Since the sequence \((\gamma(n))_{n \in \mathbb{Z}}\) is periodic with period \(q \leq M\), we can partition the discrete interval \([N]\) into arithmetic progressions with common difference \(q\) so that the fractional part of \(\gamma\) with respect to \(\Gamma\) takes constant values in each progression. In practice, let \(P_i\) be the largest progression of the form \(i + q \cdot \{0, 1, 2, \ldots\}\) inside of \([N]\) and of common difference \(q\) and starting point \(i = \{0, 1, \ldots, q - 1\}\), respectively. Then for every integer \(0 \leq i < q\), there is a number \(\gamma_i\) such that \(\gamma_i \Gamma = \gamma(P_i) \Gamma\); besides, from [9, section 2], all the coordinates \(\phi_{\mathcal{X}}(\gamma_i)\) are rationals with height at most \(O(M^{O(1)})\). Therefore, we see quickly that

\[
\frac{\phi(W)}{WN} \sum_{n \in [N]} \left( f(Wn + b) - \mathbb{E}_f(N; W, b) \right) F(\epsilon(n) g'(n) \gamma(n) \Gamma') = \frac{\phi(W)}{WN} \sum_{0 \leq i < q} \sum_{n \in P_i} \left( f(Wn + b) - \mathbb{E}_f(N; W, b) \right) F(\epsilon(n) g'(n) \gamma_i \Gamma').
\]

In the next step, we would like to split \(P_i\) further into sub-progressions such that \(\epsilon\) is approximately constant in each of these sub-progressions. For this purpose, we fix a number \(n_0 \in \mathbb{N}\) for the moment and consider the difference

\[
\left| F(\epsilon(n_0) g'(n) \gamma_i \Gamma) - F(\epsilon(n) g'(n) \gamma_i \Gamma) \right|.
\]
Taking note that the function $F : G/\Gamma \to \mathbb{C}$ has Lipschitz norm 1, one can deduce from the definition of Lipschitz norm that

$$\left| F(\varepsilon(n_0)g'(n)\gamma_i\Gamma) - F(\varepsilon(n)g'(n)\gamma_i\Gamma) \right| \leq d_{\chi}(\varepsilon(n_0)g'(n)\gamma_i, \varepsilon(n)g'(n)\gamma_i).$$

From the perspective of the operator $d_{\chi}$ is right-invariant, one has

$$d_{\chi}(\varepsilon(n_0)g'(n)\gamma_i, \varepsilon(n)g'(n)\gamma_i) = d_{\chi}(\varepsilon(n_0), \varepsilon(n)).$$

The assumption $\varepsilon$ is $(M,N)$-smooth then yields that

$$d_{\chi}(\varepsilon(n_0), \varepsilon(n)) \leq |n - n_0| \sup_{m \in [n_0, n]} d_{\chi}(\varepsilon(m - 1), \varepsilon(m)) \leq \frac{M|n - n_0|}{N}. $$

Thus, we can conclude that

$$\left| F(\varepsilon(n_0)g'(n)\gamma_i\Gamma) - F(\varepsilon(n)g'(n)\gamma_i\Gamma) \right| \ll \frac{1}{\log N}, \quad (3.1)$$

whenever $|n - n_0| \leq \frac{N}{M \log N}$. Taking the advantage of above analysis, one can decompose $P_i$ into sub-progressions $P_{i,j}$ in which each $P_{i,j}$ has diameter length at most $O\left(\frac{N}{M \log N}\right)$, and as a consequence, there are at most $O(M \log N)$ disjoint progressions. By fixing an element $\varepsilon_{i,j} \in \{\varepsilon(P_{i,j})\}$ for each progression $P_{i,j}$, plainly, we have

$$\frac{\phi(W)}{WN} \sum_{0 \leq i < q} \sum_{n \in P_i} (f(Wn + b) - \mathbb{E}_f(N;W,b)) F(\varepsilon(n)g'(n)\gamma_i\Gamma)$$

$$= \frac{\phi(W)}{WN} \sum_{i,j} \sum_{n \in P_{i,j}} (f(Wn + b) - \mathbb{E}_f(N;W,b)) F(\varepsilon_{i,j}g'(n)\gamma_i\Gamma)$$

$$+ \frac{\phi(W)}{WN} \sum_{i,j} \sum_{n \in P_{i,j}} (f(Wn + b) - \mathbb{E}_f(N;W,b)) \{F(\varepsilon(n)g'(n)\gamma_i\Gamma) - F(\varepsilon_{i,j}g'(n)\gamma_i\Gamma)\}. $$

Since it follows from Cauchy–Schwarz inequality and condition (1.11) that

$$\frac{\phi(W)}{WN} \sum_{n \in N} |f(Wn + b')| \ll N^{-1} \left( \frac{\phi(W)}{W} \right)^{1/2} \left( \frac{\phi(W)}{WN} \sum_{n \in [N]} |f(Wn + b)|^2 \right)^{1/2} \ll 1 \quad (3.2)$$

holds for every integer $1 \leq b' \leq W$ coprime with $W$. We are able to bound the above second term as follows due to inequality (3.1) and (3.2)

$$\frac{\phi(W)}{WN} \sum_{i,j} \sup_{n \in P_{i,j}} \left| F(\varepsilon_{i,j}g'(n)\gamma_i\Gamma) - F(\varepsilon(n)g'(n)\gamma_i\Gamma) \right| \sum_{n \in P_{i,j}} \left( |f(Wn + b)| + |\mathbb{E}_f(N;W,b)| \right)$$

$$\ll \frac{1}{\log N} \frac{\phi(W)}{WN} \sum_{n \in [N]} |f(Wn + b)| \ll \frac{1}{\log N}, $$
thus above term can be absorbed by the error term of Theorem 1.3. Besides, it also follows from the assumption $\varepsilon$ is $(M, N)$-smooth that $d_\chi(\varepsilon_{i,j}, \text{id}_G) \leq M$ and thus, by [10, Lemma A.4], $\psi_\chi(\varepsilon_{i,j}) \ll M^{O(1)}$.

It leaves us to concentrate on settling the above first term. When the totally equidistributed part $(g'(n)\Gamma')_{n \in [N]}$ is trivial, that is, $g'(n) = \text{id}_G \mod \Gamma$, it becomes

$$\frac{\phi(W)}{WN} \sum_{0 \leq i < q} \sum_{n \in P_i} c_{i,j} \left( f(Wn + b) - \mathbb{E}_f(N; W, b) \right)$$

for some $c_{i,j} \in \mathbb{C}$ with $|c_{i,j}| \leq 1$, in view of $F : G/\Gamma \to \mathbb{C}$ is a 1-bounded function. It then can be easily verified from condition (1.8) that this term is bounded by $O\left(\frac{1}{\log N}\right)$, and the theorem follows.

Hence, in the rest part of this paper we would always assume that $(g'(n)\Gamma')_{n \in [N]}$ is not trivial.

For a fixed pair $(i, j)$, write $(H_i)_\cdot = \gamma_i^{-1}G'\gamma_i$, and let $g_i \in \text{poly}(Z, (H_i)_\cdot)$ be the polynomial sequence defined via $g_i(n) = \gamma_i^{-1}g'(n)\gamma_i$. Besides, suppose that $\Lambda_i = \Gamma \cap H_i$, and define the 1-bounded Lipschitz function $F_{i,j} : H_i/\Lambda_i \to \mathbb{C}$ in the manner of $F_{i,j}(x\Lambda_i) = F(\varepsilon_{i,j}\gamma_i x\Gamma)$. Thus, for each pair $(i, j)$,

$$\sum_{n \in P_{i,j}} \left( f(Wn + b) - \mathbb{E}_f(N; W, b) \right) F(\varepsilon_{i,j}g'(n)\gamma_i \Gamma)$$

$$= \sum_{n \in P_{i,j}} \left( f(Wn + b) - \mathbb{E}_f(N; W, b) \right) F_{i,j}(g_i(n)\Lambda_i).$$

Now, assume that $\mu_{i,j} = \int_{H_i/\Lambda_i} F_{i,j}$ as the mean value of $F_{i,j}$ on the nilmanifold $H_i/\Lambda_i$ and then rewrite $F_{i,j}$ as $(F_{i,j} - \mu_{i,j}) + \mu_{i,j}$. Clearly, $F_{i,j} - \mu_{i,j}$ is a bounded Lipschitz function and $\int_{H_i/\Lambda_i} (F_{i,j} - \mu_{i,j}) = 0$. Moreover, taking note that $P_{i,j}$ are progressions of size $|P_{i,j}| \gg \frac{N}{qM \log N} \gg \frac{N}{qM^2}$, it can be deduced from condition (1.8), as well as $|\mu_{i,j}| \ll 1$ for all pairs $(i, j)$, that

$$\left| \frac{\phi(W)}{WN} \sum_{i,j} \sum_{n \in P_{i,j}} \mu_{i,j} \left( f(Wn + b) - \mathbb{E}_f(N; W, b) \right) \right| \ll \frac{\phi(W)}{WN} \sum_{i,j} \sum_{n \in P_{i,j}} \left| f(Wn + b) - \mathbb{E}_f(N; W, b) \right| \ll \frac{1}{\log N}.$$

Thus, without loss of generality, we may assume that $F_{i,j} : H_i/\Lambda_i \to \mathbb{C}$ is a bounded Lipschitz function with $\int_{H_i/\Lambda_i} F_{i,j} = 0$. Just sum up what we have agreed so far, we are in the position that

$$\frac{\phi(W)}{WN} \sum_{n \in [N]} \left( f(Wn + b) - \mathbb{E}_f(N; W, b) \right) F(g(n)\Gamma)$$

$$= \frac{\phi(W)}{WN} \sum_{i,j} \sum_{n \in P_{i,j}} \left( f(Wn + b) - \mathbb{E}_f(N; W, b) \right) F_{i,j}(g_i(n)\Lambda_i) + O\left(\frac{1}{\log N}\right). \quad (3.3)$$
With the help of [9, Claim in section 2], together with \( \psi, \psi(x, \varepsilon_{i, j}, \gamma_i) \ll M^{O(1)} \), one may find a Mal’cev basis \( \mathcal{Y}_i \) for \( H_i / \Lambda_i \) adapted to \( (H_i)_* \), for which each \( \mathcal{Y}_i \) is an \( M^{O(1)} \)-rational combination of elements of \( \mathcal{X}_i \); besides, we also have \( \|F_{i, j}\|_{lip} \ll M^{O(1)} \) and \( g_i \in \poly(\mathcal{Z}, (H_i)_*) \) with \( (g_i(n))_{n \in [N]} \) is totally \( M^{-cA+O(1)} \)-equidistributed for some constant \( c > 0 \). Besides, it can be seen from (3.2) that

\[
\mathbb{E}_f(N; W, b) = \frac{\phi(W)}{WN} \sum_{n \in [N]} f(Wn + b) \ll 1.
\]

Therefore, using the fact that \( (g_i(n))_{n \in [N]} \) is totally \( M^{-cA+O(1)} \)-equidistributed, and noting that \( \|F_{i, j}\|_{lip} \ll M^{-O(1)} \) and \( \int_{H_i / \Lambda_i} F_{i, j} = 0 \), we can get

\[
\frac{\phi(W)}{WN} \sum_{i,j} |\mathbb{E}_f(N; W, b)| \cdot \left| \sum_{n \in P_{i,j}} F_{i,j}(g_i(n)\Lambda_i) \right| \ll \frac{\phi(W)}{WN} \sum_{i,j} \left| \sum_{n \in P_{i,j}} F_{i,j}(g_i(n)\Lambda_i) \right| 
\ll M^{-cA+O(1)} \frac{\phi(W)}{WN} \sum_{i,j} |P_{i,j}| \ll M^{-cA+O(1)}.
\]

In view of (3.3) and above inequality, we see that to prove Theorem 1.3 it suffices to show that

\[
\frac{\phi(W)}{WN} \sum_{i,j} \sum_{n \in P_{i,j}} f(Wn + b)F_{i,j}(g_i(n)\Lambda_i) \ll \frac{1}{\log N}.
\]

And this is the main business of the next two sections.

4 | APPLYING THE APPROACH OF MONTGOMERY AND VAUGHAN

For clarity and completeness, we restate our task here again. Let \( A > 1 \) be a large number and \( \log N \ll M \ll (\log N)^{O(A)} \). Let \( G' \subseteq G \) be a subgroup of \( G \), \( g' \in \poly(\mathcal{Z}, G_*) \) and \( (g'(n))_{n \in [N]} \) be a totally \( M^{-cA} \)-equidistributed sequence for some constant \( c > 0 \). Suppose that \( P_{i,j} \) are pairwise disjoint arithmetic progressions with common difference \( q \leq M \) and \( \ell^{P_{i,j}} \geq N \). Suppose that \( (\varepsilon_{i, j})_{i, j} \) and \( (\gamma_i)_i \) are sequences of \( G \), \( \Lambda_i \) is \( (H_i) \)-rational Mal’cev basis adapted to the filtration \( (H_i)_* \). Suppose that \( F_{i,j}(x\Lambda_i) = F(\varepsilon_{i,j} x \gamma_i) \) with \( \|F_{i,j}\|_{lip} \ll M^{-O(1)} \) and \( \int_{H_i / \Lambda_i} F_{i,j} = 0 \), and \( (g_i(n)\Lambda_i)_{n \in [N]} \) is totally \( M^{-A} \)-equidistributed with \( g_i = \gamma_i^{-1} g' \gamma_i \). We will focus on handling the expression

\[
\frac{\phi(W)}{WN} \sum_{i,j} \sum_{n \in P_{i,j}} f(Wn + b)F_{i,j}(g_i(n)\Lambda_i),
\]

where \( f \) is an arbitrary function from \( M' \).

On employing the approach of Montgomery and Vaughan, we would like to add a logarithmic factor into above expression, and the aim of this section is to transfer the question to understand the equidistribution matter along with variables in terms of the prime numbers.
Lemma 4.1. Suppose that $H_i/\Lambda_i$ are nilmanifolds, and that for each $i$ there is a $M^{O(1)}$-rational Mal’cev basis $\mathcal{Y}_i$ adapted to filtration $(H_i)_*$, where $\log N \leq M \leq (\log N)^C$. Suppose that $P_{i,j}$ are disjoint arithmetic progressions with common difference $q \leq M$ and length $\Omega(\frac{N}{qM^2})$, and the union of $P_{i,j}$ is the discrete interval $[N]$. Assume that $g_i \in \text{poly}(\mathbb{Z},(H_i)_*)$ are polynomial sequences and $F_{i,j} : H_i/\Lambda_i \to \mathbb{C}$ are bounded functions, then for any multiplicative function $f \in \mathcal{M}'$, we have

$$\phi(W) \frac{1}{WN} \sum_{i,j} \sum_{n \in P_{i,j}} f(Wn+b)F_{i,j}(g_i(n)\Lambda_i) \ll \frac{1}{\log N} \left| \phi(W) \frac{1}{WN} \sum_{i,j} \sum_{p_n \in W \cdot P_{i,j} + b} \log p f(p)f(n)F_{i,j} \left( g_i \left( \frac{pn-b}{W} \right) \Lambda_i \right) \right| + \frac{1}{\log N}.$$ 

Proof. We start the proof by adding a suitable log-factor into the target expression as follows

$$\phi(W) \frac{1}{WN} \sum_{i,j} \sum_{n \in P_{i,j}} f(Wn+b)F_{i,j}(g_i(n)\Lambda_i) \log \frac{WN+b}{Wn+b}.$$ 

Directly, on the one hand, it equals to

$$\phi(W) \frac{1}{WN} \sum_{i,j} \sum_{n \in P_{i,j}} \left\{ \log(WN+b) - \log(Wn+b) \right\} f(Wn+b)F_{i,j}(g_i(n)\Lambda_i).$$ 

On the other hand, from the assumption that $F_{i,j}$ are bounded functions and all progressions $\{P_{i,j}\}_{i,j}$ form a partition of $[N]$, making use of Cauchy–Schwarz inequality and condition (1.11) gives that

$$\phi(W) \frac{1}{WN} \sum_{i,j} \sum_{n \in P_{i,j}} f(Wn+b)F_{i,j}(g_i(n)\Lambda_i) \log \frac{WN+b}{Wn+b} \ll \left( \phi(W) \frac{1}{WN} \sum_{n \in [N]} |f(Wn+b)|^2 \right)^{1/2} \left( \phi(W) \frac{1}{WN} \sum_{n \in [N]} (\log(WN+b) - \log(Wn+b))^2 \right)^{1/2}$$

$$\ll \left( \phi(W) \frac{1}{WN} \sum_{n \in [N]} \left( \log^2(N + \frac{b}{W}) - 2 \log(N + \frac{b}{W}) \log(n + \frac{b}{W}) + \log^2(n + \frac{b}{W}) \right) \right)^{1/2} \ll 1.$$ 

Combining the above two estimates, it turns out that

$$\phi(W) \frac{1}{WN} \sum_{i,j} \sum_{n \in P_{i,j}} f(Wn+b)F_{i,j}(g_i(n)\Lambda_i) \ll \frac{1}{\log N} \left| \phi(W) \frac{1}{WN} \sum_{i,j} \sum_{n \in P_{i,j}} \log(Wn+b) f(Wn+b)F_{i,j}(g_i(n)\Lambda_i) \right| + \frac{1}{\log N}.$$
By denoting \( x = Wn + b \) with \( n \in P_{i,j} \), one has \( x \in W \cdot P_{i,j} + b \), and then rename \( x \) as \( n \) to get that

\[
\sum_{n \in P_{i,j}} \log(Wn + b)f(Wn + b)F_{i,j}(g_{i}(n)\Lambda_{i}) = \sum_{n \in W \cdot P_{i,j} + b} \log n f(n)F_{i,j}\left(g_{i}\left(\frac{n - b}{W}\right)\Lambda_{i}\right).
\]

The formula \( \log n = \sum_{m|n} \Lambda(m) \), which can be viewed as a generation of the fundamental theorem of arithmetic, then yields that the above identity is also equal to

\[
\sum_{mn \in W \cdot P_{i,j} + b} \Lambda(m)f(mn)F_{i,j}\left(g_{i}\left(\frac{mn - b}{W}\right)\Lambda_{i}\right).
\]

In light of \( \Lambda(m) \) is zero unless \( m \) is a prime power, we may rewrite the above expression as

\[
\sum_{p^n \in W \cdot P_{i,j} + b} \log p f(p)f(n)F_{i,j}\left(g_{i}\left(\frac{pn - b}{W}\right)\Lambda_{i}\right)
\]

\[+ \sum_{p^n \in W \cdot P_{i,j} + b} \log p \{f(pn) - f(p)f(n)\}F_{i,j}\left(g_{i}\left(\frac{pn - b}{W}\right)\Lambda_{i}\right)
\]

\[+ \sum_{p,k \geq 2} \sum_{p^k n \in W \cdot P_{i,j} + b} \log p f(p^k n)F_{i,j}\left(g_{i}\left(\frac{p^k n - b}{W}\right)\Lambda_{i}\right).
\]

We then quickly see that this lemma follows if we can prove that the contribution from the collection of all progressions \( P_{i,j} \) in terms of the above latter two terms times \( \frac{\phi(W)}{WN} \) is bounded by 1.

Since \( f \) is multiplicative, \( f(pn) - f(p)f(n) \) vanishes unless \( p \) is a divisor of \( n \). On recalling the facts that \( F_{i,j} \) is a bounded function and the collection of \( P_{i,j} \) forms a partition of \([N]\), one has

\[
\left|\frac{\phi(W)}{WN} \sum_{i,j} \sum_{p^n \in W \cdot P_{i,j} + b} \log p \{f(pn) - f(p)f(n)\}F_{i,j}\left(g_{i}\left(\frac{pn - b}{W}\right)\Lambda_{i}\right)\right|
\]

\[\ll \frac{\phi(W)}{WN} \sum_{p,k \geq 2} \sum_{p^k n \in [WN+b]} \log p \{|f(p^k)||f(n)| + |f(p)||f(p^{k-1})||f(n)|\}.
\]

We deal with above first term at first. As \( (b, W) = 1 \) and \( 1 \leq b \leq W \ll (\log N)^C \) for some constant \( C > 0 \), so that \( b \) is quite smaller than \( N \), and thus, the first summation term is bounded by

\[
\frac{\phi(W)}{WN} \sum_{p,k \geq 2} \sum_{(p,W)=1} \log p |f(p^k)| \sum_{n \leq W^k/p^k} |f(n)| \ll \sum_{p,k \geq 2} \frac{\log p |f(p^k)|}{p^k},
\]

where we have applied the following inequality (3.2) to the inner summation.
On applying Cauchy–Schwarz inequality another time, one can deduce from condition (1.9) that
\[
\sum_{p,k \geq 2} \log p \frac{|f(p^k)|}{p^k} \leq \left( \sum_{p,k \geq 2} \frac{(\log p)^2}{p^{3k/4}} \right)^{1/2} \left( \sum_{p,k \geq 2} \frac{|f(p^k)|^2}{p^{5k/4}} \right)^{1/2} \ll 1.
\]

Besides, there is no difficulty to find that in the same manner we can also have
\[
\left| \frac{\phi(W)}{WN} \sum_{i,j} \sum_{p,k \geq 2} \sum_{p^k n \in W \cdot P_{i,j} + b} \log p f(p^k n) F_{i,j} \left( g_i \left( \frac{p^k n - b}{W} \right) \Lambda_i \right) \right| \ll 1.
\]

It, therefore, remains to calculate the following
\[
\frac{\phi(W)}{WN} \sum_{p,k \geq 2} \log p \frac{|f(p)||f(p^{k-1})|}{p^{1/3}} \sum_{n \leq WN/p^k} |f(n)| \ll \sum_{p,k \geq 1} \frac{|f(p)||f(p^k)|}{p^{1/3}p^{k/3}} \log p p^{-2/3} p^{-2k/3}.
\]

Using the inequality \(ab \leq |a|^2 + |b|^2\), together with condition (1.9), it is bounded by
\[
\sum_{p \geq 1} \frac{|f(p)|^2}{p^{3/3}} \log p \sum_{k \geq 1} p^{-2k/3} + \sum_{p,k \geq 1} \frac{|f(p^k)|^2}{p^{4k/3}} \log p p^{-2/3} \ll 1.
\]

We then finish the proof of this lemma.

The next step is to decompose the summation range of prime \(p\) into two domains, according to whether it is close to \(N\) or not. Let
\[
U = N^{2/3}
\]
be a cutoff parameter, it is clear that
\[
\frac{\phi(W)}{WN} \sum_{i,j} \sum_{p \in W \cdot P_{i,j} + b} \log p f(p)f(n) F_{i,j} \left( g_i \left( \frac{pn - b}{W} \right) \Lambda_i \right) = \frac{\phi(W)}{WN} \sum_{i,j} \left\{ \sum_{p \in U} + \sum_{U < p \leq N} \right\} \sum_{p \in W \cdot P_{i,j} + b} \log p f(p)f(n) F_{i,j} \left( g_i \left( \frac{pn - b}{W} \right) \Lambda_i \right). \tag{4.2}
\]

And we are going to handle the above two summations, respectively, in the next section. The proof idea can be explained as follows. Suppose that \(g\) is a highly equidistributed polynomial, when \(p\) is small, then \(g_p = g(p \cdot)\) would also look like equidistributed and thus the summation \(\sum_{n} F(g_p(n)\Gamma)\) should have some cancelation; while when \(p\) is large, since \(pn\) is no more than \(WN\) with \(W \ll (\log N)^C\), \(n\) should be small, and in this case the sum \(\sum_{p} F(g_n(p)\Gamma)\) also has some cancelation.
5 | EQUIDISTRIBUTION OF PRODUCT NILSEQUENCES

In this section, we will handle the two representations in (4.2) in turns.

5.1 | When $p$ is small

Lemma 5.1.

\[
\frac{\phi(W)}{WN} \sum_{i,j} \sum_{p \leq U} \sum_{p n \in W - P_{i,j} + b} \log p f(p)f(n)F_{i,j} \left( g_i \left( \frac{pn - b}{W} \right) \Lambda_i \right) \ll 1,
\]

where all of the parameters are in Section 4.

Proof. To begin with, we split $p \leq U$ into dyadic ranges $p \sim N_k$ with $N_k = 2^{-k} U$ and $0 \leq k \leq \frac{2 \log N}{3 \log 2}$. Let $P_{i,j} = a_{i,j} + q \cdot [X]$, where $a_{i,j} + q$ is the starting point of the progression $P_{i,j}$ and, by assumption, we have $X \geq \frac{N}{q M^2}$. After decomposing the sum in the congruence condition $p n \equiv W a_{i,j} + b (\text{mod } W q)$ further, we are in the position that

\[
\frac{\phi(W)}{WN} \sum_{i,j,k} \sum_{a_{1,2} (\text{mod } W q)} \sum_{p \sim N_k} \sum_{n \equiv a_1 (\text{mod } W q)} \log p f(p)f(n)F_{i,j} \left( g_i \left( \frac{pn - b}{W} \right) \Lambda_i \right),
\]

where $n$ ranges over $\frac{W a_{i,j} + b}{p} \leq n \leq \frac{W q X + W a_{i,j} + b}{p}$ and obeys the congruence condition $n \equiv a_2 (\text{mod } W q)$. Switching the summation order of variables $p$ and $n$, and using Cauchy–Schwarz inequality, one has

\[
\sum_{p \equiv a_1 (\text{mod } W q)} \sum_{n \equiv a_2 (\text{mod } W q)} \log p f(p)f(n)F_{i,j} \left( g_i \left( \frac{pn - b}{W} \right) \Lambda_i \right) \\
\leq \left( \sum_n |f(n)|^2 \right)^{1/2} \left( \sum_n \sum_{p \sim N_k} \log p f(p)f(n)F_{i,j} \left( g_i \left( \frac{pn - b}{W} \right) \Lambda_i \right)^2 \right)^{1/2},
\]

where $n$ ranges over the interval $\frac{W a_{i,j} + b}{N_k} \leq n \leq \frac{W q X + W a_{i,j} + b}{N_k}$. The first factor is easily to calculate. Indeed, condition (1.10) gives that

\[
\left( \sum_n \frac{|f(n)|^2}{W q X + W a_{i,j} + b} - \sum_n \frac{|f(n)|^2}{W a_{i,j} + b} \right)^{1/2} \ll \left( \frac{W q X}{N_k} \right)^{1/2}.
\]
For the second factor, expanding the square to double the variables \( p \) and then swapping the order of summation to get that

\[
\left( \sum_{p, p' \sim N_k} \log p \log p' f(p) f(p') \sum_n F_{i,j} \left( g_i \left( \frac{pn - b}{W} \right) \Lambda_i \right) F_{i,j} \left( g_i \left( \frac{p'n - b}{W} \right) \Lambda_i \right) \right)^{1/2},
\]

where \( n \) ranges over \( Wa_{i,j} + b \leq pn, p'n \leq WqX + Wa_{i,j} + b \) and is also in the residue class \( n \equiv a_2 \pmod{Wq} \).

Let \( n = a_2 + Wqm \), taking note that \( p \equiv a_1 \pmod{Wq} \), \( a_1a_2 \equiv b \pmod{Wq} \), as well as \( 1 \leq a_1 \leq Wq \), we can conclude that there must be integers \( |a_p|, |a_{p'}| \leq qN_k \) such that \( \frac{pn - b}{W} = qpm + a_p \) and \( \frac{p'n - b}{W} = qp'm + a_{p'} \), respectively. Thus, above inner summation over \( n \) is indeed

\[
\sum_{m \in \max \{p, p'\}} F_{i,j} \left( g_i(qpm + a_p)\Lambda_i \right) F_{i,j} \left( g_i(qp'm + a_{p'})\Lambda_i \right),
\]

where \( I \subseteq \lfloor N/q \rfloor \) is a discrete subinterval of length at least \( X \).

Thanks to the sequences \( g_i \) are totally \( M^{-A} \)-equidistributed, there is a chance to show that for almost all pairs of primes \((p, p') \in (N_k, 2N_k]^2\) the multiplicative sequence \((g_i(qp \cdot + a_p), g_i(qp' \cdot + a_{p'})) \in \text{poly}(\mathbb{Z}, (H_i, \times (H_i),))\) is equidistributed whenever this sequence is not too short. The following result is similar to [21, Proposition 8.1]. And we hope the proof could serve as an easy-reading exposition of [21, Proposition 8.1].

**Lemma 5.2 (Equidistribution of product nilsequences).** Suppose that \( 0 < \delta < 1/2 \) is a number. There are two small numbers \( 0 < c < c' < 1 \) such that the following statement holds uniformly for integers \( K \) with \( 1 \leq K \leq \delta'^c N \).

Suppose that \( H_i \) is a filtration of finite degree and of a finite-dimensional nilmanifold \( H/\Lambda \). Suppose that \( g \in \text{poly}(\mathbb{Z}, H_i) \) is a polynomial and the sequence \((g(n)\Lambda)_{n \in \lfloor N \rfloor} \) is totally \( \delta \)-equidistributed. Suppose that \( q \ll \delta^{-O(1)} \), \( a_p \) and \( a_{p'} \) are integers satisfying \( |a_p|, |a_{p'}| \leq qK \), and \( I \subseteq \lfloor N/q \rfloor \) is an interval of length \( |I| \geq \delta^{O(1)} N \). Assume further that \( F : H/\Lambda \to \mathbb{C} \) is a Lipschitz function with \( \int_{H/\Lambda} F = 0 \).

Write \( E_K \) as the set of pairs \((p, p') \in (K, 2K]^2\) for which

\[
\# \left\{ m : m \in \frac{I}{\max \{p, p'\}} \right\} \gg \delta^{O(c)} N/K,
\]

and

\[
\left| \sum_{m \in \frac{I}{\max \{p, p'\}}} F(g(qpm + a_p)\Lambda)F(g(qp'm + a_{p'})\Lambda) \right| > (1 + \|F\|_{\text{Lip}})\delta^{O(c)} N/K.
\]
Then we have

\[ \# \mathcal{E}_K \ll \delta^{O(c)} K^2. \]

**Proof.** In the following, we only consider those pairs \((p, p') \in (K, 2K]^2\) such that

\[ \# \left\{ m : m \in \frac{I}{\max \{p, p'\}} \right\} \gg \delta^{O(c)} N/K. \]

Now assume for contradiction that \( \# \mathcal{E}_K \gg \delta^{O(c)} K^2 \), which means that there are \( \gg \delta^{O(c)} K^2 \) pairs of primes \((p, p') \in (K, 2K]^2\) such that

\[ \left| \sum_{m \in I \atop \max \{p, p'\}} F(g(q pm + a_p) \Lambda)F(g(q p' m + a_{p'}) \Lambda) \right| > (1 + \|F\|_{\text{Lip}}) \delta^{O(c)} N/K. \]

For such a pair \((p, p') \in (K, 2K]^2\), we define \( g_{p, p'}(n) = (g(q pn + a_p), g(q p'n + a_{p'})) \) as a new polynomial. Then we have \( g_{p, p'} \in \text{poly}(\mathbb{Z}, H \times H) \). Besides, set a Lipchitz function \( \tilde{F} : H \times \mathbb{R} \to \mathbb{C} \) via \( \tilde{F}(\gamma, \gamma') = F(\gamma) \overline{F(\gamma)} \), then \( \int_{H/\Lambda \times H/\Lambda} \tilde{F} = 0 \) and \( \|\tilde{F}\|_{\text{Lip}} \ll \|F\|_{\text{Lip}} \). Moreover, the above inequality can be rewritten as — there are \( \gg \delta^{O(c)} K^2 \) pairs of \((p, p') \in (K, 2K]^2\) such that

\[ \left| \sum_{m \in I \atop \max \{p, p'\}} F(g_{p, p'}(m) \Lambda \times \Lambda) \right| \gg (1 + \|F\|_{\text{Lip}}) \delta^{O(c)} N/K. \]

Clearly, it can be deduced from Definition 2.2 and the assumption \( \# \{m : m \in \frac{I}{\max \{p, p'\}} \} \gg \delta^{O(c)} N/K \) that there are \( \gg \delta^{O(c)} K^2 \) of \((p, p') \in (K, 2K]^2\) such that the corresponding polynomial sequences \((g_{p, p'}(n))_{n \in I \atop \max \{p, p'\}}\) fails to be \( \delta^{O(c)} \)-equidistributed. It then follows from Theorem 2.7 that for each pair of primes \((p, p')\), there is a nontrivial horizontal character \( 0 < |\psi_{p, p'}| \ll \delta^{-O(c)} \) such that

\[ \left\| \psi_{p, p'} \circ g_{p, p'} \right\|_{C^\infty \left[ \frac{I}{\max \{p, p'\}} \right]} \ll \delta^{O(c)}. \]

On the other hand, for any nontrivial horizontal character \( \psi : H \times H \to \mathbb{T} \) of modulus \( \ll \delta^{-O(c)} \), define a set

\[ S_\psi = \{(p, p') \in (K, 2K]^2 : \psi_{p, p'} = \psi \}. \]

Since the total number of characters \( \psi_{p, p'} \) is at least \( \Omega(\delta^{O(c)} K^2) \), it can be deduced from pigeonhole principle that there is a nontrivial horizontal character \( 0 < |\psi| \ll \delta^{-O(c)} \) such that \( |S_\psi| \gg \delta^{O(c)} K^2 \). Fixing a character \( \psi \) with \( |S_\psi| \gg \delta^{O(c)} K^2 \), we thus have

\[ \left\| \psi \circ g_{p, p'} \right\|_{C^\infty \left[ \frac{I}{\max \{p, p'\}} \right]} \ll \delta^{O(c)} \]
holds for \( \gg \delta^{O(c)}K^2 \) pairs of \((p, p') \in (K, 2K)^2\). Taking \( \psi = \psi_1 \oplus \psi_2 \) where \( \psi_1, \psi_2 : H \to \mathbb{T} \) and \( \psi_1 \) is nontrivial, and let
\[
\begin{align*}
(\psi_1 \circ g)(n) &= \alpha_d n^d + \cdots + \alpha_1 n + \alpha_0; \\
(\psi_2 \circ g)(n) &= \alpha'_d n^d + \cdots + \alpha'_1 n + \alpha'_0.
\end{align*}
\]

Hence,
\[
(\psi \circ g_{p, p'})(n) = \alpha_d (qpn + a_p)^d + \cdots + \alpha_0 + \alpha'_d (qp'n + a_{p'})^d + \cdots + \alpha'_0
\]
\[
= \sum_{1 \leq j \leq d} n^j \sum_{i \leq j \leq d} \left( \begin{array}{c} i \\ j \end{array} \right) q^j (p^i a_{p}^{-j} \alpha_i + p'^i a_{p'}^{-j} \alpha'_i) + \tilde{\alpha}_0.
\]

From the definition of smoothness norm in Definition 2.6, one may find that there are \( \gg \delta^{O(c)}K^2 \) of pairs \((p, p') \in (K, 2K)^2\) such that
\[
\|q^j (p^i a_{p}^{-j} \alpha_i + p'^i a_{p'}^{-j} \alpha'_i)\| \ll \delta^{-O(c)} \left( \frac{K}{N} \right)^j
\]
for all \( 1 \leq j \leq d \), as \( \#\{m : m \in I \max \{p, p'\}\} \gg \delta^{-O(c)} N/K \). We claim that there is a non-zero integer
\[0 < q' \ll \delta^{-O(c)}\]
such that
\[
\|q' \alpha_j\| \ll \delta^{-O(c)} N^{-j} \quad \text{for all } 1 \leq j \leq d.
\]

When \( j = d \), from (5.6), there are \( \gg \delta^{O(c)}K^2 \) of pairs \((p, p') \in (K, 2K)^2\) such that \( \|q^d p^d \alpha_d + q^d p'^d \alpha'_d\| \ll \delta^{-O(c)} \left( \frac{K}{N} \right)^d \). By pigeonhole principle there is some \( p' \sim K \) such that at least \( \gg \delta^{O(c)}K \) of \( p \sim K \) satisfying the sequence \( \{q^d p^d \alpha (\mod \mathbb{Z})\} \) stays in an interval of length \( O(\delta^{-O(c)} \left( \frac{K}{N} \right)^d) \).

Here we assume that \( \delta^{-O(c)} \left( \frac{K}{N} \right)^d \ll \delta^{-O(1)} \), this can be done by taking \( 0 < c' < 1 \) larger than \( c \), recalling that \( K \leq \delta^{O(c')}N \) (we will find that \( c \) is sufficiently small). The application of [21, Lemma 8.4] yields that there are \( \gg \delta^{O(c)}K^d \) of integers \( n \leq 2^d K^d \) such that \( \{q^d n \alpha (\mod \mathbb{Z})\} \) stays in an interval of length \( O(\delta^{-O(c)} \left( \frac{K}{N} \right)^d) \). Then making use of [27, Lemma 1.1.14] together with the fact \( q \ll \delta^{-O(c)} \) ensures that there is \( 0 < q_d \ll \delta^{-O(c)} \) such that \( \|q_d \alpha_d\| \ll \delta^{-O(c)} N^{-d} \).

We now assume that there is a group of integers \( 0 < q_{d+1}, \ldots, q_d \ll \delta^{-O(c)} \) such that \( \|q_i \alpha_i\| \ll \delta^{-O(c)} N^{-i} \) holds for all \( j < i \leq d \). And consider the case \( j \). Also pigeonhole gives some \( p' \sim K \) such that for at least \( \delta^{O(c)} \) proportion of \( p \sim K \) such that \( \{(\sum_{j \leq i \leq d} q^j p^j a_{p}^{-j} \alpha_i (\mod \mathbb{Z}))\} \) lies in an interval of length at most \( O(\delta^{-O(c)} \left( \frac{K}{N} \right)^j) \). While, from the assumption one has when \( i > j \),
\[
\|q^j p^j a_{p}^{-j} \alpha_i\| \leq p^j a_{p}^{-j} \|q^j \alpha_i\| \ll \delta^{-O(c)} \left( \frac{K}{N} \right)^j,
\]
as \( |a_p| \leq qK, p \sim K \) and \( \|q_i \alpha_i\| \ll \delta^{-O(c)} N^{-i} \) for some \( q_i \ll \delta^{-O(c)} \). This means that the sequence \( \{q^j p^j a_{p}^{-j} \alpha_i (\mod \mathbb{Z})\} \) moves very slowly apart from the interval \( \{(\sum_{j \leq i \leq d} q^j p^j a_{p}^{-j} \alpha_i (\mod \mathbb{Z}))\} \), and thus, we can conclude that there are at least \( \delta^{O(c)} \) proportion of \( p \sim K \) such that \( \{q^j p^j \alpha_j (\mod \mathbb{Z})\} \) stays in an interval of length at most \( O(\delta^{-O(c)} \left( \frac{K}{N} \right)^j) \). Another application of [21, Lemma 8.4] and [27, Lemma 1.1.14] gives an integer \( 0 < q_j \ll \delta^{-O(c)} \) such that \( \|q_j \alpha_j\| \ll \delta^{-O(c)} N^{-j} \). Therefore, the claim follows from taking \( q' \) as the least common multiple of \( q_1, \ldots, q_d \).
Now write $\psi' = q' \cdot \psi$, then we see that $\|\psi' \circ g(n)\| \ll \delta^{-O(c)} \frac{n}{N}$ for every positive integer $n$. Let $N' = \delta^{O(cC)} N$ for some large $C \geq 1$ such that when $n \in [N']$, we can get

$$\|\psi' \circ g(n)\| \leq \frac{1}{10}.$$ 

We then set $F' : H/\Lambda \to \mathbb{C}$ to be the function $F' = \eta \psi'$, where $\eta : \mathbb{T} \to \mathbb{C}$ is a function with mean zero, bounded Lipschitz norm and equals 1 on the interval $[-\frac{1}{10}, \frac{1}{10}]$. Then we have $\int_{H/\Lambda} F' = 0$, $\|F'\|_{\text{Lip}} \ll \delta^{-O(c)}$, and

$$\left| \mathbb{E}_{n \in [N]} F'((g(n)\Lambda) \right| \geq \delta \|F'\|_{\text{Lip}}$$

provided that $c$ is sufficiently small. Thereby, we conclude the proof by noting that above inequality contradicts with the assumption that $(g(n)\Lambda)_{n \in [N]}$ is $\delta$-equidistributed. \qed

We now back our attention to the representation (5.2) and utilizing the full strength of Lemma 5.2. Recalling that the sequence $(g_i(n))_{n \in [N]}$ is totally $M^{-A}$-equidistributed for each $i$, where

$$\log N \leq M \ll (\log N)^{O_A(1)}, \quad (5.8)$$

and $q \leq M$ and $|I| \geq \frac{N}{qM^2}$, one can invoke Lemma 5.2 with $\delta = M^{-A}$. To begin with, we split the sum range over $p, p' \sim N_k$ into two cases according to whether $\#\{m : m \in \frac{I}{\max\{p, p'\}}\} \gg M^{-O(cA)} N$ or not. If $\{m : m \in \frac{I}{\max\{p, p'\}}\}$ does not contain many elements (less than $CM^{-O(cA)} N$), one can show that the total contribution in (5.1) is negligible. In practice, The fact that $F_{i, j}$ is a bounded function leads to

$$\sum_{m \in \frac{I}{\max\{p, p'\}}} F_{i, j}(g_i(qpm + a_{p})\Lambda_i) F_{i, j}(g_i(qp'm + a_{p'})\Lambda_i) \ll \# \left\{ m : m \in \frac{I}{\max\{p, p'\}} \right\}.$$ 

From the perspective of the assumption, it is bounded by $O(M^{-O(cA)} N)$. Substituting it into (5.4) to obtain that (5.4) is no more than

$$\left(M^{-O(cA)} N \sum_{p, p' \sim N_k} \log p \log p' |f(p)| |f(p')| \right)^{1/2} = M^{-O(cA)} N^{1/2} \sum_{p \sim N_k} \log p |f(p)| \ll M^{-O(cA)} (NN_k)^{1/2},$$

where the last inequality follows from condition (1.9) and Cauchy–Schwarz inequality. Substituting this inequality and (5.3) into (5.2) yields that (5.2) is bounded by $O(M^{-O(cA)} - 1 W^{1/2} N)$, due to $X \leq \frac{N}{qM^2}$. Thus, it can be seen that the contribution to (5.1) is bounded by

$$q W^{1/2} M^{-O(cA)} \sum_{k \leq \log N} \# \left\{ (a_1, a_2) \in [0, W - 1)^2 : a_1 a_2 \equiv b \pmod{W} \right\} \ll 1.$$
just by using the facts that \( W \leq (\log N)^C \), \( q \leq M \) and (5.8) and by enlarging \( A > 1 \) if necessary.

Hence, we just need to consider those pairs \((p, p')\) for which \( \frac{1}{\max\{p, p'\}} \gg M^{-O(cA) N} \). Decomposing the summation over prime pairs \((p, p')\) in (5.4) according to whether \((p, p')\) in the exceptional set \( E_k \), it allows us to conclude that (5.4) is indeed a half power of the following expression

\[
\left\{ \sum_{(p, p') \in E_{N_k}} + \sum_{(p, p') \in (N_k, 2N_k] \setminus E_{N_k}} \right\} \log p \log p' |f(p)||f(p')|
\]

\[
\times \left| \sum_{m \in \max\{p, p'\}} F_{i, j} \left( g_i(qpm + a_p)\Lambda_i \right) F_{i, j} \left( g_i(qp'm + a_{p'})\Lambda_i \right) \right|.
\]

The application of Lemma 5.2 shows that \( \#E_{N_k} \ll M^{-O(cA) N_k^2} \), and thus the above first summation term is bounded by

\[
\sum_{(p, p') \in (N_k, 2N_k]} \log p \log p' |f(p)||f(p')|1_{(p, p') \in E_{N_k}} \cdot \frac{|I|}{\max\{p, p'\}}
\]

due to \( F_{i, j} \) is a bounded function. By making use of Cauchy–Schwarz inequality and condition (1.9), as well as noting that \( |I| = \Theta(\frac{N}{qM^2}) \), one may find it is bounded by

\[
\frac{N}{qM^2N_k} \left( \sum_{p \sim N_k} (\log p f(p))^2 \right) \left( \#E_{N_k} \right)^{1/2}
\ll q^{-1} M^{-O(cA)} N \sum_{p \sim N_k} \log p |f(p)|^2 \ll M^{-O(cA)} q^{-1} NN_k.
\]

Substituting the above inequality and (5.3) into representation (5.1) and taking \( A \geq 1 \) large enough, the total contribution to (5.1) can be bounded by \( O(1) \).

It, therefore, remains to calculate the case \( \frac{1}{\max\{p, p'\}} \gg M^{-O(cA) N} \) and \((p, p') \notin E_{N_k} \). Another application of Lemma 5.2 gives that

\[
\sum_{(p, p') \notin E_{N_k}} \log p \log p' |f(p)||f(p')| \sum_{m \in \max\{p, p'\}} F_{i, j} \left( g_i(qpm + a_p)\Lambda_i \right) F_{i, j} \left( g_i(qp'm + a_{p'})\Lambda_i \right)
\ll M^{-O(cA)} \frac{N}{N_k} \left( \sum_{p \sim N_k} \log p f(p) \right)^2 \ll M^{-O(cA)} N \sum_{p \sim N_k} \log p \sum_{p \sim N_k} \log p |f(p)|^2
\ll M^{-O(cA)} NN_k.
\]

And in this case it can also be shown that the contribution to (5.1) is bounded by \( O(1) \).
To sum up, in order to prove Lemma 5.1 we split the prime pairs \((p, p') \in (N_k, 2N_k]\) into three cases: the interval \(\frac{I}{\max\{p, p'\}}\) is short; there are many elements in the interval \(\frac{I}{\max\{p, p'\}}\) but \((p, p')\) belongs to the exceptional set \(\mathcal{E}_{N_k}\); and there are many elements in the interval \(\frac{I}{\max\{p, p'\}}\) and \((p, p')\) is out of \(\mathcal{E}_{N_k}\) so that (5.5) has \(M^{-O(cA)}\) decay. We then complete the deduction by showing that in each case the contribution to (5.1) is bounded by \(O(1)\).

\[\square\]

5.2 When \(p\) is large

The task of this subsection is to prove the following lemma. And, clearly, Theorem 1.3 follows from this lemma and Lemma 5.1.

**Lemma 5.3.** With all parameters the same as (4.2), we have

\[
\phi(W) \frac{1}{WN} \sum_{U < p \leq N} \sum_{i,j} \sum_{p \in \mathcal{P}_{i,j} + b} \log p \ f(p) f(n) \Phi_{i,j} \left( g_i \left( \frac{pn - b}{W} \right) \Lambda_i \right) \ll 1.
\]

**Proof.** The proof, to some extent, is similar to Lemma 5.1. We also start with splitting the interval \(U < p \leq N\) into dyadic ranges, say \(p \sim U_k\) with \(U_k = 2^k U\) and \(0 \leq k \leq \frac{\log N}{3 \log 2}\). For a fixed number \(0 \leq k \leq \frac{\log N}{3 \log 2}\) and a fixed pair \((i, j)\), consider the following representation

\[
\sum_{\substack{p \sim U_k \\
p \equiv a_1 (\text{mod } Wq) \geq x \}} \log p \ f(p) \sum_{\substack{w a_{i,j} + b \\
 \leq n \leq \frac{w q x + w a_{i,j} + b}{p} \leq Wq \}} f(n) \Phi_{i,j} \left( g_i \left( \frac{pn - b}{W} \right) \Lambda_i \right).
\]

Here, as before, letting \(P_{i,j} = a_{i,j} + q \cdot [X]\), and both \(a_1\) and \(a_2\) run over the residue classes modulus \(Wq\) and satisfy the relationship \(a_1 a_2 \equiv W a_{i,j} + b (\text{mod } Wq)\). The application of Cauchy–Schwarz inequality shows that it is bounded by

\[
\left( \sum_{\substack{p \sim U_k \\
p \equiv a_1 (\text{mod } Wq) \geq x \}} \log p \ f(p)^2 \right)^{1/2} \left( \sum_{\substack{w a_{i,j} + b \\
 \leq n, n' \leq \frac{w q x + w a_{i,j} + b}{p} \leq Wq \}} f(n) f(n') \right)^{1/2}
\]

\[\times \sum_{\substack{p \sim U_k \\
p \equiv a_1 (\text{mod } Wq) \geq x \}} \log p \Phi_{i,j} \left( g_i \left( \frac{pn - b}{W} \right) \Lambda_i \right) \Phi_{i,j} \left( g_i \left( \frac{pn' - b}{W} \right) \Lambda_i \right) \right)^{1/2}.
\]

(5.9)
Condition (1.9) implies that the first factor is bounded by $O(U_k^{1/2})$. After changing of variables, we see that the sum over $p$ in above second factor is indeed bounded by

$$\sum_{m \sim \frac{U_k}{Wq}} \Lambda(Wqm + a_1)F_{i,j}(g_i(qmn + c_n)\Lambda_i)F_{i,j}(g_i(qmn' + c_n')\Lambda_i),$$

since the contribution of the higher powers is $O(U_k^{1/2 + \varepsilon})$ which is admissible in view of the estimate of the above expression, see (5.12). When $(a_1, Wq) \neq 1$, since $F_{i,j}$ is a bounded function, it is easy to see that in this case above expression is bounded by $\log U_k \cdot \#\{p \sim U_k : p \equiv a_1 \pmod{Wq}\} \ll \log U_k$, thus the contribution is negligible.

In the following, we can assume that $(a_1, Wq) = 1$. A possible way to deal with this expression is to follow the idea of [21, Lemma 9.5], to transfer $|E_m \Lambda(Wqm + a_1)F(g'(m)\Gamma)|$ to $|E_m F(g'(m)\Gamma)|$ with some error term. In [21, Lemma 9.5], this error term is double logarithmic decay which is far too large than we expect. Indeed, one can check that to make Theorem 1.3 has logarithmic decay, this error term should be strongly logarithmic decay. This seems achievable, because [28, theorem 2.7] shows that the average of $\Lambda - \Lambda_{\text{Siegel}}$ and polynomial nilsequences has pseudopolynomial decay. It, thus, needs to calculate the average $E_{n \in [N]}(\Lambda_{\text{Siegel}} - 1)(Wn + b)F(g'(n)\Gamma)$. [28, Proposition 7.1] tells us that it is bounded by $O(M^{O(1)}q^{O(1)}_{\text{Siegel}})$. Because in our case $M \ll (\log N)^{O(1)}$, if $q_{\text{Siegel}} \ll_A (\log N)^{-A}$, we have the chance to show that $E_{n \in [N]}(\Lambda_{\text{Siegel}} - 1)(Wn + b)F(g'(n)\Gamma)$ is logarithmic decay. However, the assumption $q_{\text{Siegel}} \ll_A (\log N)^{-A}$ will never happen in the case of $W \ll (\log N)^{O(1)}$, on noting from the analysis surrounds formula (5.6) in [28, P. 19].

Our strategy is as follows. When $p$ is a large number, that is $p > U = N^{2/3}$, due to $pn, pn' \leq WN$, both $n$ and $n'$ should be small numbers. Inspired by the proof idea of Lemma 5.1, when polynomial $g$ is highly equidistributed, we guess for almost all pairs of $(n, n')$ the product polynomial $(g(n \cdot), g(n' \cdot))$ should be equidistributed. It, therefore, can be reduced to consider the correlation of an equidistributed nilsequence with the shifted von Mangoldt function. And this can be calculated by the so-called bi-linear form method adapted to polynomial nilsequences, see [9, section 3] as an example.

For technical reason, we would decompose the summation interval of $n, n'$ in the following representation into dyadic ranges, say $n, n' \sim V_l$ with $V_l = [\frac{W_a}{U_k}, \frac{W_a + b}{U_k}]$ and $0 \leq l \ll \log N$.

$$\sum_{n} f(n)f(n') \sum_{m \sim \frac{Wq}{U_k}} \Lambda(Wqm + a_1)F_{i,j}(g_i(qmn + c_n)\Lambda_i)\overline{F_{i,j}(g_i(qmn' + c_n')\Lambda_i)},$$

where both $n$ and $n'$ are supported in the interval $[\frac{W_a}{U_k}, \frac{WqX + W_a + b}{U_k}]$ and subject to $n \equiv n' \equiv a_2 \pmod{Wq}$. Let $g_{i; n, m'} : \mathbb{Z} \to H_i/\Lambda_i \times H_i/\Lambda_i$ be the polynomial defined by

$$g_{i; n, m'}(m) = (g_i(qnm + c_n), g_i(qnm' + c_n')).$$

Claim. There is a small constant $0 < c < 1$ such that the following statement holds uniformly for $V_l$ with $0 \leq l \ll \log N$. Suppose that $(g(m))_{m \sim \frac{Wq}{U_k}}$ is totally $M^{-A}$-equidistributed. When the

\[\text{We would like to thank Joni Teräväinen for explaining their paper to us.}\]
pairs \((n, n') \in (V_1, 2V_1]\) is outside of a subset of size bounded by \(O(M^{-O(c_A)}V_1^2)\), the corresponding product polynomial sequence \((g_{n,n'}(m))_{m \sim \frac{U_k}{Wq}}\) is totally \(M^{-O(c_A)}\)-equidistributed.

Since the treatment for every \(i\) is the same, in the following we fix one such \(i\) and omit the letter \(i\). Similar to the proof of Lemma 5.2, we assume that the sequence \((g_{n,n'}(m))_{m \sim \frac{U_k}{Wq}}\) fails to be \(M^{-O(c_A)}\)-equidistributed. Then from quantitative Leibman theorem (Theorem 2.7), there is a nontrivial horizontal character of modulus \(0 < \psi \ll M^{-O(c_A)}\) such that

\[
\|\psi \circ g_{n,n'}\|_{C^\infty \left[ \frac{U_k}{Wq}, \frac{2U_k}{Wq} \right]} \ll M^{-O(c_A)}.
\]

One can see from the proof of Lemma 5.2 that this contradicts with the assumption \((g(m))_{m \sim \frac{U_k}{Wq}}\) is \(M^{-A}\)-equidistributed when we take \(0 < c < 1\) sufficiently small. Just need to pay attention to one thing, the prime pairs \((p, p')\) in Lemma 5.2 are replaced by integer pairs \((n, n')\) here, but this does not matter because it can be achieved by replacing [21, Lemma 8.4] by [9, Lemma 3.3] in the proof. The claim then follows by an application of [21, Lemma 7.2], which asserts that if \((g(m))_{m \in \mathbb{N}}\) is \(\delta\)-equidistributed, then there is a number \(0 < c < 1\) such that \((g(m))_{m \in \mathbb{N}}\) is totally \(\delta c\)-equidistributed.

Now we write \(\bar{F}_{i,j}(\gamma, \gamma') = F_{i,j}(\gamma)F_{i,j}(\gamma')\) as a Lipschitz function supported on \(H_i/\Lambda_i \times H_i/\Lambda_i\). Then we have the information that \(\int_{H_i/\Lambda_i \times H_i/\Lambda_i} \bar{F}_{i,j} = 0\) and \(\|\bar{F}\|_{\text{Lip}} \ll M^{-O(1)}\). Next, we aim to prove when \((g_{n,n'}(m))_{m \sim \frac{U_k}{Wq}}\) is totally \(M^{-O(c_A)}\)-equidistributed, the following inequality holds

\[
\sum_{m \sim \frac{U_k}{Wq}} \Lambda(Wqm + a_1)\bar{F}_{i,j}(g_{n,n'}(m) \Lambda_i \times \Lambda_i) \ll M^{-O(c_A)} \frac{U_k}{Wq}.
\]

Here, \(a_1\) is less than \(Wq\) and coprime with \(Wq\). This, indeed, follows from the next lemma. And a brief proof of the below lemma will be given after the proof of Lemma 5.3.

**Lemma 5.4** (von Mangoldt function disjoints with equidistributed nilsequences). Let \(0 < \delta < 1/2\) be a parameter. Suppose that \(G/\Gamma\) is a nilmanifold of finite dimension \(m_G\), \(G\) is a filtration of \(G/\Gamma\) and of degree \(d\), and \(\mathcal{Y}\) is a \(\delta^{-O(1)}\)-rational Mal’cev basis adapted to \(G\). Suppose that \(g \in \text{poly}(\mathbb{Z}, G/\Gamma)\) is a polynomial and \((g(n))_{n \in \mathbb{N}}\) is totally \(\delta\)-equidistributed. Then for any coprime pair \(1 \leq b \leq W \ll \delta^{-O(1)}\) and any Lipschitz function \(F: \mathbb{Z} \to \mathbb{C}\) with \(\int_{G/\Gamma} F = 0\) and \(\|F\|_{\text{Lip}} \ll \delta^{-O(1)}\), we have

\[
\mathbb{E}_{n \in \mathbb{N}} \Lambda(Wn + b)F(g(n)\Gamma) \ll \delta^{-O(1)}.
\]

We now splice all the ingredients to finish the proof of Lemma 5.3. For a fixed number \(0 \leq l \ll \log N\), divide the pairs \((n, n') \in (V_1, 2V_1]^2\) into two sets. One is that (5.11) holds for each pair \((n, n')\) in this set, and we call the other one as the exceptional set \(E_l\). In view of the claim, the size of the exceptional set \(E_l\) is at most \(O(M^{-O(c_A)}V_1^2)\), besides, for \((n, n') \in E_l\), we note that

\[
\sum_{m \sim \frac{U_k}{Wq}} \Lambda(Wqm + a_1)\bar{F}_{i,j}(g_{n,n'}(m) \Lambda_i \times \Lambda_i) \ll \sum_{m \sim \frac{U_k}{Wq}} \Lambda(Wqm + a_1) \ll \frac{U_k}{\phi(Wq)},
\]

(5.12)
as $F_{i,j}$ is bounded and $(Wq, a_1) = 1$. One can now see from (5.11), together with above information for exceptional set, that (5.10) is bounded by

$$
\sum_{0 \leq l < \log N} \left\{ M^{-O(cA)} \frac{U_k}{Wq} \sum_{n, n' \sim V_l} |f(n)| |f(n')| + \frac{U_k}{\phi(Wq)} \sum_{(n, n') \in (V_l, 2V_l)} |f(n)| |f(n')| 1_{E_l}(n, n') \right\}.
$$

We now compute above two sum terms in turns. Summing over the dyadic intervals $(V_l, 2V_l]$, and then using Cauchy–Schwarz inequality and condition (1.10), one has the first term is bounded by

$$
M^{-O(cA)} \frac{U_k}{Wq} \left( \sum_{WqX \leq n \leq WqX + Wai, j + b} |f(n)| \right)^2 \ll M^{-O(cA)} \frac{U_k}{Wq} \left( \sum_{WqX \leq n \leq WqX + Wai, j + b} |f(n)| \right)^2 \ll M^{-O(cA)} \frac{WqX^2}{U_k}.
$$

Besides, Cauchy–Schwarz inequality and condition (1.10) and summing over the dyadic intervals $(V_l, 2V_l]$ also shows that the second term is bounded by

$$
\frac{U_k}{\phi(Wq)} \left( \sum_{l < \log N} \sum_{n \in V_l} |f(n)| \right)^2 \ll \frac{U_k}{\phi(Wq)} \left( \sum_{l < \log N} \sum_{(n, n') \in (V_l, 2V_l]} |f(n)| \right)^2 \ll \frac{M^{-O(cA)}(\log N)^{1/2} N X}{U_k}.
$$

Recalling that $X \gg \frac{N}{qM^2}$ and $\log N \ll M \ll (\log N)^{O_A(1)}$, when we take $A > 1$ sufficiently large, the factors $(\log N)^{1/2}$ and $N/X = qM^2$ can be absorbed by $M^{-O(cA)}$, thus, we can assume that both of the above two terms are bounded by $O(M^{-O(cA)} X^2/U_k)$. Substitute it into (5.9), as well as recalling that the first factor in (5.9) is no more than $O(U_k^{1/2})$, we can conclude that (5.9) is bounded by $O(M^{-O(cA)} X)$. Splicing the intervals $(U_k, 2U_k]$ with $0 \leq k \leq \frac{\log N}{3\log 2}$, one therefore has

$$
\frac{\phi(W)}{WN} \sum_{U < p \leq \log N} \sum_{i, j} \sum_{p n \in W - P_{i,j} + b} \log p f(p) f(n) F_{i,j} \left( g_i \left( \frac{pn - b}{W} \right) \right) \Lambda_i \ll \frac{\phi(W)}{WN} \sum_{k < \log N} \sum_{i,j} \sum_{a_1, a_2 \equiv Wai, j + b (mod Wq)} M^{-O(cA)} X \ll 1,
$$
as $X = \Theta\left(\frac{N}{qM^2}\right)$ and letting $A > 1$ sufficiently large. Therefore, we complete the proof of this lemma.

A brief proof of Lemma 5.4. On recalling the well-known Vaughan’s identity [29]

$$\Lambda(n) = \Lambda(n)1_{n \leq N^{1/3}} - \sum_{d \leq N^{2/3}} a_d 1_{d|n} + \sum_{d \leq N^{1/3}} \mu(d) 1_{d|n} \log \frac{n}{d} + \sum_{d, w > N^{1/3}} \Lambda(d) b_w 1_{dw = n},$$

where $a_d = \sum_{bc = d, b, c \leq N^{1/3}} \mu(b) \Lambda(c)$ and $b_w = \sum_{c | w : c > N^{1/3}} \mu(c)$. The first term is a negligible one, we call the second and fourth term as Type I sum and Type II sum, respectively, besides, the third one can be expressed as a convex combination of Type I sum by using the identity

$$\log \frac{n}{d} = \log N - \int_1^{N/t} \frac{dt}{t} = \log d,$$

and then by absorbing all the various logarithmic factors into the divisor-bounded coefficients. Therefore, the expression $\sum_{n \leq N} \Lambda(Wn + b)F(g(n) \Gamma)$ can be expressed as follows,

$$\sum_{n \leq N} \Lambda(n)F\left(g\left(\frac{n - b}{W}\right)\Gamma\right) + \sum_{n \leq N} \sum_{d \leq N^{2/3}} a_d F\left(g\left(\frac{n - b}{W}\right)\Gamma\right) + \sum_{d, w \leq N^{1/3}} a_d b_w F\left(g\left(\frac{dw - b}{W}\right)\Gamma\right),$$

where $P = b + W \cdot [N]$ is the arithmetic progression, $a_d \ll (\log N)^O(1) \tau_{O(1)}(d)$, and $b_w \ll (\log N)^O(1) \tau_{O(1)}(w)$. Plainly, the first term is bounded by $O(N^{1/2})$ (say) so is negligible, and the latter two terms can be dealt with in the same manner of [9, section 3], also of [28, section 7].

6 | PROOF OF THEOREMS 1.4 AND 1.5

We first give some standard facts about $L$-functions which can be found in [25, section 2] and [13, chapter 5].

Let $m \geq 2$ be an integer, $\pi = \bigotimes_p \pi_p$ be a normalized irreducible cuspidal automorphic representations of $GL_m$ over $\mathbb{Q}$, which means that $\pi$ has unitary central character. The $L$-function $L(s, \pi)$ associated to $\pi$ is defined by

$$L(s, \pi) = \sum_{n=1}^{\infty} \frac{\lambda_\pi(n)}{n^s} = \prod_p \prod_{j=1}^{m} \left(1 - \frac{\alpha_{j, \pi}(p)}{p^s}\right)^{-1},$$

for some suitable complex numbers $\alpha_{j, \pi}(p)$. The Dirichlet series converge absolutely for $\Re(s) > 1$.

There is also an archimedean local factor $L(s, \pi_\infty)$. There are $m$ complex Langlands parameters $\mu_\pi(f)$, we define

$$L(s, \pi_\infty) = \pi^{-\frac{m}{2}} \prod_{j=1}^{m} \left(\frac{s + \mu_\pi(f)}{2}\right).$$
The generalized Ramanujan conjecture and Selberg conjecture assert that

\[ |\alpha_{j,\pi}(p)| \leq 1 \text{ and } \Re(\mu_{\pi}(j)) \geq 0. \]

The Ramanujan conjecture was proved by Deligne [4] for holomorphic cusp form on \( GL_2 \). According to Luo, Rudnick, and Sarnak [20] and Müller and Speh [24], we know that

\[ |\alpha_{j,\pi}(p)| \leq p^{\frac{1}{2} - \frac{1}{m^2 + 1}} \text{ and } \Re(\mu_{\pi}(j)) \geq -\left(\frac{1}{2} - \frac{1}{m^2 + 1}\right). \] (6.1)

To define the functional equation for \( L(s, \pi) \), we give the contragredient of \( \pi \). Let \( \tilde{\pi} \) be the contragredient of \( \pi \) which is also an irreducible cuspidal automorphic representations of \( GL_m \) over \( \mathbb{Q} \). For each \( p \leq \infty \), we have

\[ \{\alpha_{j,\pi}(p) : 1 \leq j \leq m\} = \{\alpha_{j,\pi}(p): 1 \leq j \leq m\} \]

and

\[ \{\mu_{\tilde{\pi}}(j) : 1 \leq j \leq m\} = \{\mu_{\pi}(j) : 1 \leq j \leq m\}. \]

Define the completed \( L \)-function

\[ \Lambda(s, \pi) = N_{\pi}^{s/2}L(s, \pi)L(s, \pi_{\infty}), \]

where \( N_{\pi} \) be the conductor of \( \pi \). Then, \( \Lambda(s, \pi) \) extends to an entire function and satisfies the functional equation

\[ \Lambda(s, \pi) = \xi(\pi)\Lambda(1 - s, \tilde{\pi}), \]

where \( \xi(\pi) \) is a complex number of modulus 1.

Let \( \chi \) be a primitive Dirichlet character modulo \( q \). Then the twisted \( L \)-function is defined by

\[ L(s, \pi \times \chi) = \prod_{p} \prod_{j=1}^{m} \left(1 - \frac{\alpha_{j,\pi \times \chi}(p)}{p^s}\right)^{-1}. \]

When \( p \nmid q \), we have

\[ \{\alpha_{j,\pi \times \chi}(p) : 1 \leq j \leq m\} = \{\alpha_{j,\pi}(p)\chi(p) : 1 \leq j \leq m\}. \]

Thus, we get

\[ \sum_{n=1}^{\infty} \frac{\lambda_{\pi}(n)\chi(n)}{n^s} = \prod_{p} \prod_{j=1}^{m} \left(1 - \frac{\alpha_{j,\pi}(p)\chi(p)}{p^s}\right)^{-1} = L(s, \pi \times \chi)\prod_{p\mid q} \prod_{j=1}^{m} \left(1 - \frac{\alpha_{j,\pi \times \chi}(p)}{p^s}\right). \] (6.2)
We also need the following convexity bound for $L(s, \pi \times \chi)$, see [19, Lemma 3.2].

$$L\left(\frac{1}{2} + it, \pi \times \chi\right) \ll (q(1 + |t|))^{\frac{m}{2} + \varepsilon}. \tag{6.3}$$

Next, we prove that $\lambda_{\pi}(n) \in \mathcal{M}'$ with $W = 1$, then Theorem 1.4 directly follows from Theorem 1.3. The conditions (1.10) and (1.9) can be easily deduced from the Rankin–Selberg theory and the prime number theorem of Rankin–Selberg L-functions, see [19, (5.2) and p. 631], that is,

$$\sum_{n \leq N} |\lambda_{\pi}(n)|^2 \ll N, \tag{6.4}$$

and

$$\sum_{p \leq N} |\lambda_{\pi}(p)|^2 \log p \ll N. \tag{6.5}$$

Then, we just need to prove that $\lambda_{\pi}(n)$ satisfy condition (1.8) with $W = 1$. Take $P = \{n \equiv b(\text{mod } q), n \leq N\}$ with $q \ll (\log N)^C$, we consider the following sum

$$\sum_{n \leq N \atop n \equiv b(\text{mod } q)} \lambda_{\pi}(n).$$

We will consider the following two cases according to $(b, q) = 1$ or $(b, q) > 1$.

Case I: $(b, q) = 1$. Applying the orthogonality of Dirichlet characters, we have

$$\sum_{n \leq N \atop n \equiv b(\text{mod } q)} \lambda_{\pi}(n) = \frac{1}{\varphi(q)} \sum_{\chi \text{ mod } q} \overline{\chi}(b) \sum_{n \leq N} \lambda_{\pi}(n) \chi(n). \tag{6.6}$$

It suffices to estimate the sum

$$\sum_{n \leq N} \lambda_{\pi}(n) \chi(n) \tag{6.7}$$

for any $\chi(\text{mod } q)$. This is the same with [19, (5.4)], they proved that

$$\sum_{n \leq N} \lambda_{\pi}(n) \chi(n) \ll q^{\frac{m} {2m+4} + \varepsilon} N^{\frac{m+1}{m+2} + \varepsilon}. \tag{6.8}$$

Taking (6.8) into (6.6), we get

$$\sum_{n \leq N \atop n \equiv b(\text{mod } q)} \lambda_{\pi}(n) \ll q^{\frac{m} {2m+4} + \varepsilon} N^{\frac{m+1}{m+2} + \varepsilon}. \tag{6.9}$$

This is also true for $q = 1$. We confirm that $\lambda_{\pi}(n)$ satisfy condition (1.8) with $W = 1$.

Case II: $(b, q) > 1$. First, we consider that $n$ runs over square-free numbers. Let $d = (b, q)$, $b = b'd$ and $q = q'd$, then $(b', q') = 1$ and $(d, n/d) = 1$. Denote $\chi_{d}$ be the principal character modulo
\[
\sum_{\substack{n \in \mathcal{N} \\ \nu(n) \equiv b \pmod{q} \atop n \text{ square-free}}} \lambda_{\pi}(n) = \lambda_{\pi}(d) \sum_{\substack{l \in \mathcal{N}/d \\ l \equiv b' \pmod{q'} \atop \nu(l) \text{ square-free}}} \lambda_{\pi}(l)\chi_d(l) \\
= \frac{\lambda_{\pi}(d)}{\varphi(q')} \sum_{\chi \bmod q'} \chi(b') \sum_{\substack{l \in \mathcal{N}/d \\ \nu(l) \text{ square-free}}} \lambda_{\pi}(l)(\chi\chi_d)(l).
\]

(6.10)

Since \(\chi\chi_d\) is a character modulo \(q'd = q\), we use (6.8) to the inner sum, we have

\[
\sum_{\substack{n \in \mathcal{N} \\ \nu(n) \equiv b \pmod{q} \atop n \text{ square-free}}} \lambda_{\pi}(n) \ll q \frac{m}{3m+4} + \varepsilon N \frac{m}{m+2} + \varepsilon.
\]

(6.11)

Here, we use \(\lambda_{\pi}(d) \ll d^\frac{1}{2} - \frac{1}{m+1} + \varepsilon\).

Next we remove the restriction that \(n\) is square-free. Any arbitrary positive integer \(n\) can be represented in a unique way as

\[n = kl, \quad k \text{ is square-free}, \quad l \text{ is square-full}, \quad (k, l) = 1.\]

Hence, we have

\[
\sum_{\substack{n \in \mathcal{N} \\ \nu(n) \equiv b \pmod{q} \atop n \text{ square-free}}} \lambda_{\pi}(n) = \sum_{\substack{l \in \mathcal{N} \\ \nu(l) \text{ square-full}, kl \text{ square-free}}} \lambda_{\pi}(l) \lambda_{\pi}(k) = \sum_{\substack{l \in \mathcal{N} \\ \nu(l) \text{ square-full}}} \lambda_{\pi}(l) \sum_{\substack{k \in \mathcal{N}/l \\ \nu(k) \text{ square-free}}} \lambda_{\pi}(k).
\]

We split the sum over \(l\) into two parts: \(1 \leq l \leq N^{4/5}\) and \(N^{4/5} < l \leq N\). If \(1 \leq l \leq N^{4/5}\), then \(N/l \geq N^{1/5}\). We use (6.11) for the inner sum, we have

\[
\sum_{\substack{n \in \mathcal{N} \\ \nu(n) \equiv b \pmod{q} \atop l \in N^{4/5}, n \text{ square-full}}} \lambda_{\pi}(n) \ll q \frac{m}{3m+4} + \varepsilon N \frac{m+1}{m+2} + \varepsilon \sum_{l \in N^{4/5}} \lambda_{\pi}(l) \ll q \frac{m}{3m+4} + \varepsilon N \frac{5m+9}{5m+10} + \varepsilon.
\]

(6.12)

Here we use (6.4) and the Cauchy–Schwarz inequality to get

\[
\sum_{l \in N^{4/5}} \frac{|\lambda_{\pi}(l)|}{l \text{ square-full}} \ll N \frac{4}{5(m+2)} \left( \sum_{l \in N^{4/5}} \frac{1}{l} \right)^{1/2} \left( \sum_{l \in N^{4/5}} \frac{|\lambda_{\pi}(l)|^2}{l} \right)^{1/2} \ll N \frac{4}{5(m+2)} \log N.
\]
If $N^{4/5} < l \leq N$, we use (6.4) and the Cauchy–Schwarz inequality

$$\sum_{n \in \mathbb{N} \atop n \equiv b \pmod q} \lambda_{\pi}(n) \ll \sum_{l \leq N, l \text{ square-full}} |\lambda_{\pi}(l)| \sum_{k \leq N^{1/5}, k \text{ square-free}} |\lambda_{\pi}(k)|$$

$$\ll \left( \sum_{l \leq N, l \text{ square-full}} 1 \right)^{1/2} \left( \sum_{l \leq N} |\lambda_{\pi}(l)|^2 \right)^{1/2} \left( \sum_{k \leq N^{1/5}} 1 \right)^{1/2} \left( \sum_{k \leq N^{1/5}} |\lambda_{\pi}(k)|^2 \right)^{1/2}$$

$$\ll N^{19/20}.$$  

(6.13)

From (6.12) and (6.13), we have

$$\sum_{n \leq N \atop n \equiv b \pmod q} \mu(n) \lambda_{\pi}(n) \ll q^{\frac{m}{2m+4} + \varepsilon} N^{\frac{5m+9}{5m+10} + \varepsilon}.$$  

We confirm that $\lambda_{\pi}(n)$ satisfy condition (1.8). This completes the proof of Theorem 1.4.

To prove Theorem 1.5, we prove that $\mu(n)\lambda_{\pi}(n) \in M'$ with $W = 1$. From (6.4) and (6.5), we know that

$$\sum_{n \leq N} |\mu(n)\lambda_{\pi}(n)|^2 \ll \sum_{n \leq N} |\lambda_{\pi}(n)|^2 \ll N,$$  

(6.14)

and

$$\sum_{p \leq N} |\mu(p)\lambda_{\pi}(p)|^2 \log p \ll N.$$  

(6.15)

This says that $\mu(n)\lambda_{\pi}(n)$ satisfies conditions (1.10) and (1.9).

Next we consider the following sum

$$\sum_{n \leq N \atop n \equiv b \pmod q} \mu(n)\lambda_{\pi}(n),$$

with $q \leq (\log N)^C$. We use the recent result of Jiang, Lü, and Wang [18, section 5], for $\pi$ is self-dual and $\pi \not\cong \pi \otimes \chi$ for any quadratic primitive character $\chi$, they prove that

$$\sum_{n \leq N} \mu(n)\lambda_{\pi}(n)\chi(n) \ll N \exp \left( -c \sqrt{\log N} \right)$$  

(6.16)

for any Dirichlet character $\chi \pmod q$, where $c > 0$ is a constant.

Let $d = (b, q)$, $b = b'd$ and $q = q'd$, then $(b', q') = 1$. Denote $\chi_d$ be the principal character modulo $d$. Then we have
\[
\sum_{\substack{n \in \mathbb{N} \\
 n \equiv b \pmod{q}}} \mu(n) \lambda_\pi(n) = \sum_{\substack{d \mid N \\
 \gcd(d, q') = 1}} \mu(d) \lambda_\pi(d) \sum_{\substack{n \in \mathbb{N} \\
 n \equiv b' \pmod{q'}}} \mu(n) \lambda_\pi(n) \chi_d(n)
\]

\[
= \mu(d) \lambda_\pi(d) \sum_{\substack{n \in \mathbb{N} \\
 n \equiv b' \pmod{q'}}} \mu(n) \lambda_\pi(n) \chi_d(n)
\]

\[
= \frac{\mu(d) \lambda_\pi(d)}{\varphi(q')} \sum_{\chi \pmod{q'}} \chi(b') \sum_{\substack{n \in \mathbb{N} \\
 n \equiv b' \pmod{q'}}} \mu(n) \lambda_\pi(n) \chi_d(n).
\]

Since \( \chi \chi_d \) is a character modulo \( q'd = q \), we use (6.16) to the inner sum, we have

\[
\sum_{\substack{n \in \mathbb{N} \\
 n \equiv b \pmod{q}}} \mu(n) \lambda_\pi(n) \ll N \exp \left( -c \sqrt{\log(N/q)} \right).
\]

Here, we use \( \lambda_\pi(d) \ll d^{1 - \frac{1}{m^2 + 1} + \varepsilon} \). We confirm that \( \mu(n) \lambda_\pi(n) \) satisfy condition (1.8) with \( W = 1 \). This completes the proof of Theorem 1.5.

**ACKNOWLEDGEMENTS**

The authors are grateful to Yujiao Jiang, Lilian Matthiesen, Joni Teräväinen, and Zhiwei Wang for their helpful discussions. The authors also appreciate comments and suggestions from the referee which help us to improve this article. Part of this work was written when Mengdi Wang visits Shandong University, she thanks the warm and generous hospitality from Shandong University. Xiaoguang He is supported by the National Postdoctoral Innovative Talents Support Program (Grant Number: BX20190227), NSFC (Grant Number: 12101427), and the Fundamental Research Funds for the Central Universities, SCU (Grant Number: 2021SCU12109). Mengdi Wang is partially supported by the Göran Gustafsson Foundation.

**JOURNAL INFORMATION**

*Mathematika* is owned by University College London and published by the London Mathematical Society. All surplus income from the publication of *Mathematika* is returned to mathematicians and mathematics research via the Society’s research grants, conference grants, prizes, initiatives for early career researchers and the promotion of mathematics.

**REFERENCES**

1. M. Cafferata, A. Perelli, and A. Zaccagnini, *An extension of the Bourgain–Sarnak–Ziegler theorem with modular applications*, Q. J. Math. **71** (2020), no. 1, 359–377.
2. H. Daboussi, *Fonctions multiplicatives presque périodiques B*, Astérisque **24** (1975), no. 25, 321–324.
3. H. Davenport, *On some infinite series involving arithmetical functions (II)*, Quart. J. Math. **8** (1937), 313–320.
4. P. Deligne, *La conjecture de Weil. I*, Inst. Hautes Études Sci. Publ. Math. **1** (1974), no. 43, 273–307.
5. É. Fouvry and S. Ganguly, *Strong orthogonality between the Möbius function, additive characters and Fourier coefficients of cusp forms*, Compos. Math. **150** (2014), no. 5, 763–797.
6. N. Frantzikinakis and B. Host, *Higher order Fourier analysis of multiplicative functions and applications*, J. Amer. Math. Soc. **30** (2017), no. 1, 67–157.
7. B. Green and T. Tao, *The primes contain arbitrarily long arithmetic progressions*, Ann. of Math. (2) **167** (2008), no. 2, 481–547.
8. B. Green and T. Tao, *Linear equations in primes*, Ann. of Math. (2) **171** (2010), no. 3, 1753–1850.
9. B. Green and T. Tao, The Möbius function is strongly orthogonal to nilsequences, Ann. of Math. (2) 175 (2012), no. 2, 541–566.
10. B. Green and T. Tao, The quantitative behaviour of polynomial orbits on nilmanifolds, Ann. of Math. (2) 175 (2012), no. 2, 465–540.
11. L. K. Hua, Additive theory of prime numbers, Translations of Mathematical Monographs, vol. 13, American Mathematical Society, Providence, RI, 1965.
12. H. Iwaniec, Spectral methods of automorphic forms, Graduate Studies in Mathematics, 2nd ed., vol. 53, American Mathematical Society, Providence, RI, Revista Matemática Iberoamericana, Madrid, 2002.
13. H. Iwaniec and E. Kowalski, Analytic number theory, American Mathematical Society Colloquium Publications, vol. 53, American Mathematical Society, Providence, RI, 2004.
14. Y. J. Jiang and G. S. Lü, On automorphic analogues of the Möbius randomness principle, J. Number Theory 197 (2019), 268–296.
15. Y. J. Jiang and G. S. Lü, The generalized Bourgain–Sarnak–Ziegler criterion and its application to additively twisted sums on GL_m, Sci. China Math. 64 (2021), no. 10, 2207–2230.
16. Y. J. Jiang and G. S. Lü, Cancellation in algebraic twisted sums on GL_m, Forum Mathematicum 33 (2021), no. 4, 1061–1082.
17. Y. J. Jiang, G. S. Lü, J. Thorner, and Z. H. Wang, A Bombieri–Vinogradov theorem for higher rank groups, to appear in Int. Math. Res. Not. (2021). https://doi.org/10.1093/imrn/rnab261
18. Y. J. Jiang, G. S. Lü, and Z. H. Wang, Möbius randomness law for GL(m) automorphic L-functions twisted by additive characters, Preprint.
19. Y. J. Jiang, G. S. Lü, and Z. W. Wang, Exponential sums with multiplicative coefficients without the ramanujan conjecture, Math. Ann. 379 (2021), no. 1, 589–632.
20. W. Z. Luo, Z. Rudnick, and P. Sarnak, On Selberg’s eigenvalue conjecture, Geom. Funct. Anal. 5 (1995), no. 2, 387–401.
21. L. Matthiesen, Generalized fourier coefficients of multiplicative functions, Algebra Number Theory 12 (2018), no. 6, 1311–1400.
22. S. D. Miller, Cancellation in additively twisted sums on GL(n), Amer. J. Math. 128 (2006), no. 3, 699–729.
23. H. L. Montgomery and R. C. Vaughan, Exponential sums with multiplicative coefficients, Invent. Math. 43 (1977), no. 1, 69–82.
24. W. Müller and B. Speh, Absolute convergence of the spectral side of the Arthur trace formula for GL_n, Geom. Funct. Anal. 14 (2004), no. 1, 58–93. (With an appendix by E. M. Lapid.)
25. Z. Rudnick and P. Sarnak, Zeros of principal L-functions and random matrix theory, Duke Math. J. 81 (1996), no. 2, 269–322.
26. P. Sarnak, Three lectures on the möbius function randomness and dynamics. www.math.ias.edu/files/wam/2011/PSMmobius.pdf, 2011.
27. T. Tao, Higher order Fourier analysis, vol. 142, American Mathematical Society, Providence, RI, 2012.
28. T. Tao and J. Teräväinen, Quantitative bounds for Gowers uniformity of the Möbius and von Mangoldt functions, arXiv:2107.02158, 2021.
29. R. C. Vaughan, Sommes trigonométriques sur les nombres premiers, C. R. Acad. Sci. Paris Sér. A-B 285 (1977), no. 16, A981–A983.