Algebra & Number Theory

Volume 12
2018
No. 6

Generalized Fourier coefficients of multiplicative functions

Lilian Matthiesen

msp

ALGEBRA AND NUMBER THEORY 12:6 (2018)
dx.doi.org/10.2140/ant.2018.12.1311

Generalized Fourier coefficients of multiplicative functions

Lilian Matthiesen

We introduce and analyze a general class of not necessarily bounded multiplicative functions, examples of which include the function $n \mapsto \delta^{\omega(n)}$, where $\delta \in \mathbb{R} \setminus \{0\}$ and where ω counts the number of distinct prime factors of n, as well as the function $n \mapsto |\lambda_f(n)|$, where $\lambda_f(n)$ denotes the Fourier coefficients of a primitive holomorphic cusp form.

For this class of functions we show that after applying a W-trick, their elements become orthogonal to polynomial nilsequences. The resulting functions therefore have small uniformity norms of all orders by the Green–Tao–Ziegler inverse theorem, a consequence that will be used in a separate paper in order to asymptotically evaluate linear correlations of multiplicative functions from our class. Our result generalizes work of Green and Tao on the Möbius function.

1.	Introduction	1311
2.	Brief outline of some ideas	1317
3.	A suitable Dirichlet decomposition for $f \in \mathscr{M}_h$	1319
4.	Multiplicative functions in progressions: The class $\mathscr{F}_H$	1322
5.	The W-trick	1349
6.	The noncorrelation result	1355
7.	Linear subsequences of equidistributed nilsequences	1365
8.	Equidistribution of product nilsequences	1370
9.	Proof of Proposition 6.4	1376
	Appendix: Explicit bounds on the correlation of Λ with nilsequences	1392
	Acknowledgements	1398
	References	1398

1. Introduction

Let $f : \mathbb{N} \to \mathbb{C}$ be a multiplicative arithmetic function. Daboussi showed (see [Daboussi and Delange 1974]) that if $|f|$ is bounded by 1, then

$$\frac{1}{x} \sum_{n \leqslant x} f(n) e^{2\pi i \alpha n} = o(x) \tag{1-1}$$

This work was partly supported through a postdoctoral fellowship of the Fondation Sciences Mathématiques de Paris, by the National Science Foundation (Grant No. DMS-1440140) while the author was in residence at MSRI in Spring 2017, and by the Swedish Research Council (Grant No. 2016-05198).
MSC2010: primary 11B30; secondary 11L07, 11N60, 37A45.
Keywords: multiplicative functions, nilsequences, Gowers uniformity norms.

for every irrational α. A detailed proof of the following slightly strengthened version may be found in [Daboussi and Delange 1982]: Suppose that f satisfies

$$\sum_{n\leqslant x}|f(n)|^2 = O(x); \tag{1-2}$$

then (1-1) holds for every irrational α. Montgomery and Vaughan [1977] give explicit error terms for the decay in (1-1) for multiplicative functions that satisfy, in addition to (1-2), a uniform bound at all primes, in the sense that $|f(p)|\leqslant H$ holds for some constant $H\geqslant 1$ and all primes p.

In this paper we will study the closely related question of bounding correlations of multiplicative functions with polynomial nilsequences in place of the exponential function $n\mapsto e^{2\pi i\alpha n}$. A chief concern in this work is to include unbounded multiplicative functions in the analysis. To this end we shall significantly weaken the moment condition (1-2) by decomposing f into a suitable Dirichlet convolution $f=f_1*\cdots*f_t$ and analyzing the correlations of the individual factors with exponentials, or rather nilsequences. The benefit of such a decomposition is that we merely require control on the second moments of the individual factors of the Dirichlet convolution and not of f itself. This essentially allows us to replace (1-2) by the condition that there exists $\theta_f\in(0,1]$ such that

$$\sqrt{\frac{1}{x}\sum_{n\leqslant x}|f_i(n)|^2} \ll (\log x)^{1-\theta_f}\frac{1}{x}\sum_{n\leqslant x}|f_i(n)| \tag{1-3}$$

for all $i\in\{1,\dots,t\}$. To illustrate the difference between these two moment conditions, let us consider a simple example of a function that satisfies (1-3), but neither (1-2) nor

$$\sum_{n\leqslant x}|f(n)|^2 \ll \sum_{n\leqslant x}|f(n)|. \tag{1-4}$$

Example 1.1. For any $t\in\mathbb{N}$, let $d_t(n)=\mathbf{1}*\cdots*\mathbf{1}(n)$ denote the general divisor function, which arises as a t-fold convolution of $\mathbf{1}$. Choosing $f_i=\mathbf{1}$ for each $1\leqslant i\leqslant t$, it is clear that (1-3) holds with $\theta_f=1$. If $t>1$, then neither (1-2) nor (1-4) hold, since

$$\frac{1}{x}\sum_{n\leqslant x}d_t(n)\asymp_t(\log x)^{t-1}, \qquad \text{but} \qquad \frac{1}{x}\sum_{n\leqslant x}d_t^2(n)\asymp_t(\log x)^{t^2-1}.$$

Thus, the second moment is not controlled by the first.

In order to describe the three classes of multiplicative functions that we will be working with here, let us introduce some notation. Throughout this paper, we write

$$S_f(x)=\frac{1}{x}\sum_{n\leqslant x}f(n) \quad \text{and} \quad S_f(x;q,r)=\frac{q}{x}\sum_{\substack{n\leqslant x\\ x\equiv r \pmod q}}f(n)$$

for $x\geqslant 1$ and integers $q,r\in\mathbb{N}$. We furthermore require the following functions w and W:

Definition 1.2. Let $w : \mathbb{N} \to \mathbb{R}$ be an increasing function such that

$$\frac{\log\log x}{\log\log\log x} < w(x) \leqslant \log\log x$$

for all sufficiently large x, and set

$$W(x) = \prod_{p \leqslant w(x)} p.$$

The basic class of function we will be interested in is the following:

Definition 1.3. Given a positive integer $H \geqslant 1$, we let $\mathscr{M}_H$ denote the class of multiplicative arithmetic functions $f : \mathbb{N} \to \mathbb{C}$ such that:

(1) $|f(p^k)| \leqslant H^k$ for all prime powers p^k.

(2) There is a positive constant α_f such that

$$\frac{1}{x} \sum_{p \leqslant x} |f(p)| \log p \geqslant \alpha_f$$

for all sufficiently large x.

For the purpose of our main result, Theorem 6.1, it will be necessary to restrict attention to those functions f that admit a so-called W-trick (see Section 5). For this reason, we introduce the subset of elements of $\mathscr{M}_H$ that have stable mean values in certain arithmetic progressions:

Definition 1.4. Let $\mathscr{F}_H \subset \mathscr{M}_H$ be the subset of multiplicative functions f with the following property. Let $x > 1$ be a parameter. Given any constant $C > 0$, there exists a function φ_C with $\varphi_C(x) \to 0$ as $x \to \infty$ such that, whenever $1 \leqslant Q < (\log x)^C$ is a multiple of $W(x)$ and when $A \pmod{Q}$ is a reduced residue, then

$$S_f(x'; Q, A) = S_f(x; Q, A) + O\left(\varphi_C(x) \frac{Q}{\phi(Q)} \frac{1}{\log x} \prod_{\substack{p \leqslant x \\ p \nmid Q}} \left(1 + \frac{|f(p)|}{p}\right)\right) \tag{1-5}$$

for all $x' \in (x(\log x)^{-C}, x)$.

We will discuss this class of functions in detail in Section 4, where we prove several sufficient conditions for $f \in \mathscr{M}_H$ to belong to $\mathscr{F}_H$, or to a related class that will be introduced below. These sufficient conditions, recorded in Propositions 4.4 and 4.10 and Lemmas 4.16 and 4.17, prove to be much easier to verify in practice than the one given in the above definition, not at least because they take a form that allows for applications of the Selberg–Delange method as presented in [Tenenbaum 1995]. As an application of Lemmas 4.16 and 4.17 (see the remarks following their statements), we obtain the following simple criterion applicable to real-valued elements of $\mathscr{M}_H$:

Proposition 1.5. *Suppose that $f \in \mathcal{M}_H$ is real-valued and that it is bounded away from zero at primes, in the sense that there exists $\delta > 0$ and a sign $\epsilon \in \{+, -\}$ such that*

$$\#\{p \leqslant x : \epsilon f(p) \geqslant \delta\} \geqslant \frac{(1+o(1))x}{\log x}, \quad (as\ x \to \infty).$$

Then $f \in \mathcal{F}_H$ if f is nonnegative or if, for every given $C > 0$, there exists a function $\psi_C : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$ with $\psi_C(x) \to 0$ as $x \to \infty$ such that

$$S_{f\chi_0}(x) = O\left(\frac{\psi_C(x)}{\log x} \exp\left(\sum_{p \leqslant x, p \nmid Q} \frac{|f(p)|}{p}\right)\right), \quad (x > 1)$$

for all trivial characters $\chi_0 \pmod Q$ with $Q \in (1, (\log x)^C)$ and $W(x) \mid Q$.

Observe, in particular, that this criterion may be applied to functions that take negative values at all primes, such as the Möbius function. In the latter case, the prime number theorem-type estimate $S_\mu(x) \ll_B (\log x)^{-B}$, which holds for all $x \geqslant 2$ and $B > 0$, implies that all conditions are satisfied; see Example 4.18(i) for details. As an easy consequence of the above proposition, it further follows that any function of the form $f(n) = \delta^{\omega(n)}$ for fixed $\delta > 0$ belongs to $\mathcal{F}_H$. In Section 4D we will show that the function $n \mapsto |\lambda_f(n)|$ belongs to $\mathcal{F}_H$, where $\lambda_f(n)$ denotes the normalized Fourier coefficients of a primitive holomorphic cusp form. This is an example which cannot be deduced from the above proposition.

In Section 6, we will see that in the context of our main result condition (1-5) only needs to hold for slowly varying twists of f. This allows us to slightly weaken the above definition and introduce the following intermediate class of functions $\mathcal{F}_H \subset \mathcal{F}_{H,n^{it}} \subset \mathcal{M}_H$, which will also be discussed in Section 4.

Definition 1.6. Let $\mathcal{F}_{H,n^{it}} \subset \mathcal{M}_H$ denote the subset of functions f with the following property. For every constant $C > 0$ and every sufficiently large $x > 1$, there exists $t_x \in \mathbb{R}$ with $|t_x| \leqslant 2\log x$ such that the function $f_x : n \mapsto f(n)n^{-it_x}$ satisfies (1-5) for all $x' \in (x(\log x)^{-C}, x)$, all $1 \leqslant Q < (\log x)^C$, $W(x) \mid Q$, and all reduced residues $A \pmod Q$. Observe that $\mathcal{F}_H \subset \mathcal{F}_{H,n^{it}}$ since we may take $t_x = 0$ for all x.

Twists of the form $f(n)n^{-it}$ play an important role in the study of multiplicative functions as their behavior is closely linked to that of the mean value of f through Halász's theorem [1968]; see also [Tenenbaum 1995, §III.4.3]. While Halász's theorem concerns bounded functions that are closely related to the constant function **1**, an analogue to this result, applicable to our basic class $\mathcal{M}_H$, has recently been proved independently by Elliott [2017, Theorems 2 and 4] and Tenenbaum [2017, Théorème 1.2]. The next lemma, which we chiefly include for comparison of the error terms in (1-5) and in later results, is a straightforward consequence of their result. The first part is due to Elliott and Kish [2016, Lemma 21].

Lemma 1.7 (Elliott, Kish, Tenenbaum). *Suppose $f \in \mathcal{M}_H$ and that*

$$\sum_{p \leqslant H} \sum_{k \geqslant 2} |f(p^k)| p^{-k} < \infty.$$

Then

$$S_{|f|}(x) \gg \frac{1}{\log x} \exp\biggl(\sum_{p\leqslant x} \frac{|f(p)|}{p}\biggr).$$

Furthermore, we have $|S_f(x)| = o(S_{|f|}(x))$ *unless there exists* $t \in \mathbb{R}$ *such that*

$$\sum_{p \text{ prime}} \frac{|f(p)| - \Re(f(p)p^{it})}{p} < \infty,$$

in which case $|S_f(x)| \asymp S_{|f|}(x)$.

Returning to the basic class $\mathscr{M}_H$, let us record the lemma that shows that every element of $\mathscr{M}_H$ does indeed admit a Dirichlet decomposition with the properties described at the beginning of this introduction. To be precise, the lemma below corresponds to $\theta_f = \frac{1}{2}$ in (1-3). We will prove this lemma in Section 3. In accordance with the earlier discussion, this lemma will only be needed in the case where f is unbounded, i.e., when $H > 1$.

Lemma 1.8. *(Dirichlet decomposition) Let* $f \in \mathscr{M}_H$ *and let* h *be the multiplicative function defined as*

$$h(p^k) = \begin{cases} f(p)/H & \text{if } k = 1, \\ 0 & \text{if } k > 1. \end{cases} \tag{1-6}$$

Let h^{*H} *denote the* H*-fold convolution of* h *with itself. Then*

$$f = h^{*H} * h',$$

where h' *is a multiplicative function that satisfies* $h'(p) = 0$ *at primes and* $|h'(p^k)| \leqslant (2H)^k$ *at prime powers.*

Let $f = f_1 * \cdots * f_H$ *with* $f_i = h$ *for all but one of the factors and* $f_i = h * h'$ *for the remaining one. If* $x > 1$ *and if* $Q \leqslant x^{1/2}$ *is an integer multiple of* $W(x)$*, then the following bound holds for all* $A \in (\mathbb{Z}/Q\mathbb{Z})^*$:

$$\sum_{\substack{D\leqslant x^{1-1/H} \\ \gcd(D,Q)=1}} \sum_{d_1\cdots d_{H-1}=D} \frac{|f_1(d_1)\cdots f_{H-1}(d_{H-1})|}{D} \sqrt{\frac{DQ}{x} \sum_{\substack{n\leqslant x/D \\ nD\equiv A \ (\mathrm{mod}\ Q)}} |f_H(n)|^2} \ll (\log x)^{1/2} \frac{Q}{\phi(Q)} \frac{1}{\log x} \prod_{\substack{p\leqslant x \\ p\nmid Q}} \biggl(1 + \frac{|f(p)|}{p}\biggr). \tag{1-7}$$

Aim and motivation. As mentioned before, the purpose of this paper is to study correlations of multiplicative functions, more specifically of functions from $\mathscr{M}_H$, with polynomial nilsequences. In general, such correlations can only be shown to be small if either the nilsequence is highly equidistributed or else if the multiplicative function is equidistributed in progressions with short common difference. We will consider both cases, the former in Proposition 6.4 and the latter in Theorem 6.1. In accordance with this restriction, the latter result only applies to the subsets $\mathscr{F}_H$ and $\mathscr{F}_{H,n^{it}}$ whose elements admit a W-trick as we will establish in Section 5. Restricting attention to the class $\mathscr{F}_H$ for now, then "W-trick" roughly means the following here. For every $f \in \mathscr{F}_H$ there is a product $\widetilde{W} = \widetilde{W}(x)$ of small prime powers such that f has a constant average value in all suitable subprogressions of $\{n \equiv A \pmod{\widetilde{W}}\}$ for every fixed residue $A \in (\mathbb{Z}/\widetilde{W}\mathbb{Z})^*$. Instead of bounding Fourier coefficients of f as in (1-1), we aim to show that every $f \in \mathscr{F}_H$ satisfies[1]

$$\frac{\widetilde{W}}{x} \sum_{n \leqslant x/\widetilde{W}} (f(\widetilde{W}n + A) - S_f(x; \widetilde{W}, A)) F(g(n)\Gamma) = o_{G/\Gamma}\left(\frac{1}{\log x} \frac{\widetilde{W}}{\phi(\widetilde{W})} \prod_{p \leqslant x,\ p \nmid \widetilde{W}} \left(1 + \frac{|f(p)|}{p} \right) \right) \tag{1-8}$$

for all 1-bounded polynomial nilsequences $F(g(n)\Gamma)$ of bounded degree and bounded Lipschitz constant that are defined with respect to a nilmanifold G/Γ of bounded step and bounded dimension. The precise statement will be given in Section 6. This result can be viewed as a generalization of work of Green and Tao [2012a] who were the first to study correlations of the form (1-8) and who prove (1-8) for the Möbius function. In fact, we borrow their approach to reduce Theorem 6.1 to Proposition 6.4 in Section 6 and we work with their techniques in Sections 7 and 8.

Note carefully that the bound proposed in (1-8) is nontrivial even in the case where the function f satisfies $S_f(x) = o(1)$, i.e., even for a function like $f(n) = \delta^{\omega(n)}$ with $\delta \in (0, 1)$, which satisfies

$$S_f(x) \sim (\log x)^{\delta - 1} \asymp \frac{1}{\log x} \prod_{p \leqslant x} \left(1 + \frac{|f(p)|}{p} \right) = o(1).$$

To see this, we observe that Lemma 1.7 and Shiu's lemma [1980, Theorem 1] imply that the error term in (1-8) is, at least for a positive proportion of the reduced residues $A \pmod{\widetilde{W}}$, of the form o("the trivial upper bound"), which is the bound obtained by inserting absolute values everywhere.

The interest in estimates of the form (1-8) lies in the fact that the Green–Tao–Ziegler inverse theorem [Green et al. 2012] allows one to deduce that $f(\widetilde{W}n + A) - S_f(x; \widetilde{W}, A)$ has small U^k-norms of all orders, where "small" may depend on k. Employing the nilpotent Hardy–Littlewood method of Green and Tao [2010], this in turn allows one to deduce asymptotic formulae for expressions of the form

$$\sum_{\boldsymbol{x} \in K \cap \mathbb{Z}^s} f(\varphi_1(\boldsymbol{x}) + a_1) \cdots f(\varphi_r(\boldsymbol{x}) + a_r), \tag{1-9}$$

for $a_1, \dots, a_r \in \mathbb{Z}$, pairwise nonproportional linear forms $\varphi_1, \dots, \varphi_r : \mathbb{Z}^s \to \mathbb{Z}$ and convex $K \subset \mathbb{R}^s$, *provided* that f has a sufficiently pseudorandom majorant function. We construct such pseudorandom majorants in the companion paper [Matthiesen 2016], which also addresses the question of evaluating (1-9) for functions $f \in \mathscr{F}_{H,n^{it}}$ with the property that $|f(n)| \ll_\varepsilon n^\varepsilon$ for all $\varepsilon > 0$.

Strategy and related work. Our overall strategy is to decompose the given multiplicative function via Dirichlet decomposition in such a way that we can employ the Montgomery–Vaughan approach to the individual factors. This approach reduces matters to bounding correlations of sequences defined in terms of primes. One type of correlation that appears will be handled with the help of Green and Tao's bound [2010, Proposition 10.2] on the correlation of the "W-tricked von Mangoldt function" with nilsequences.

[1] This statement needs to be slightly adapted if $f \in \mathscr{F}_{H,n^{it}}$.

Carrying out the Montgomery–Vaughan approach in the nilsequences setting makes it necessary to understand the equidistribution properties of certain families of product nilsequences which result from an application of the Cauchy–Schwarz inequality. These product sequences are studied in Section 8 refining techniques introduced in [Green and Tao 2012a]. More precisely, we show that most of these products are equidistributed provided the original sequence that these products are derived from was equidistributed. The latter can be achieved by the Green–Tao factorization theorem for nilsequences from [Green and Tao 2012b].

The question studied in this paper is in spirit related to that of Bourgain–Sarnak–Ziegler [Bourgain et al. 2013], who use an orthogonality criterion that can be proved employing ideas that go back to Daboussi and Delange [1974] (see also [Harper 2011] and [Tao 2011]). Invoking the orthogonality criterion in the form it is presented in [Kátai 1986], recent and very substantial work of Frantzikinakis and Host [2017] shows that every bounded multiplicative function can be decomposed into the sum of a Gowers-uniform function, a structured part and an error term. This error term is small in the sense that the integral of the error term over the space of all 1-bounded multiplicative functions is small. While their result provides no information on the quality of the error term of individual functions, it allows one to study simultaneously all bounded multiplicative functions.

The point of view taken in the present work is a different one: we have applications to explicit multiplicative functions in mind. For many multiplicative functions f that appear naturally in number theoretic contexts, the mean value $\frac{1}{x}\sum_{n\leqslant x} f(x)$ is described by a reasonably nice function in x, and one can hope to be able to verify the conditions from Definitions 1.3 and 1.4 (or 1.6) for such functions. In order to deduce asymptotic formulae for expressions as in (1-9), it is important that the bound on the correlation (1-8) improves at least on the trivial bound given by the average value of $|f|$. Thus, we need to be able to understand these bounds for individual functions f. We establish a noncorrelation result (Theorem 6.1) with an explicit bound that preserves information on $|f|$ just as in (1-8). An important feature of this work is that it applies to a large class of unbounded functions.

Notation. The following, perhaps unusual, piece of notation will be used throughout the paper: Suppose $\delta \in (0,1)$, we write $x = \delta^{-O(1)}$ instead of $x = (1/\delta)^{O(1)}$ to indicate that there is a constant $0 \leqslant C \ll 1$ such that $x = (1/\delta)^C$.

Convention. If the statement of a result contains Vinogradov or O-notation in the assumptions, the implied constants in the conclusion may depend on all implied constants from the assumptions.

2. Brief outline of some ideas

In this section we give a very rough outline of the ideas behind the application of the Montgomery–Vaughan approach in the nilsequences setting, making a number of simplifications for the benefit of the exposition. The main idea of Montgomery and Vaughan [1977] is to introduce a log factor into the Fourier coefficient that we wish to analyze. Let $f : \mathbb{N} \to \mathbb{R}$ be a multiplicative function that satisfies $|f(p)| \leqslant H$

for some constant $H \geqslant 1$ and all primes p and suppose (1-4) holds. Then we have

$$\sum_{n\leqslant N} f(n)e(n\alpha)\log\frac{N}{n} \leqslant \biggl(\sum_{n\leqslant N}\Bigl(\log\frac{N}{n}\Bigr)^2\biggr)^{1/2}\biggl(\sum_{n\leqslant N}|f(n)|^2\biggr)^{1/2} \ll N^{1/2}\biggl(\sum_{n\leqslant N}|f(n)|^2\biggr)^{1/2},$$

and thus

$$\log N\biggl(\frac{1}{N}\sum_{n\leqslant N} f(n)e(n\alpha)\biggr) \ll \biggl(\frac{1}{N}\sum_{n\leqslant N}|f(n)|^2\biggr)^{1/2} + \biggl|\frac{1}{N}\sum_{n\leqslant N} f(n)e(n\alpha)\log n\biggr|.$$

The first term in the bound is handled by the assumptions on f, that is, by assuming that (1-4) holds. To bound the second term, one invokes the identity $\log n = \sum_{d|n}\Lambda(d)$, which reduces the task to bounding the expression

$$\sum_{nm\leqslant N} f(nm)\Lambda(m)e(nm\alpha).$$

This in turn may be reduced to the task of bounding

$$\sum_{np\leqslant N} f(n)f(p)\Lambda(p)e(pn\alpha),$$

where p runs over primes. Applying the Cauchy–Schwarz inequality and smoothing, it furthermore suffices to estimate expressions of the form

$$\sum_{p,p'} f(p)f(p')\log(p)\log(p')\sum_{n} w(n)e((p-p')n\alpha),$$

where p and p' run over primes and where w is a smooth weight function. One employs a standard sieve estimate to bound $\#\{(p,p') : p-p'=h\}$ for fixed h. Standard exponential sum estimates and a delicate decomposition of the summation ranges for n, p, p' yield an explicit bound on $(1/N)\sum_{n\leqslant N} f(n)e(n\alpha)$.

We seek to employ the above approach to correlations of the form

$$\frac{1}{N}\sum_{n\leqslant N}\biggl(f(n)-\frac{1}{N}\sum_{m\leqslant N} f(m)\biggr)F(g(n)\Gamma)$$

for multiplicative f. One problem we face is that the above approach makes substantial use of the strong equidistribution properties of the exponential functions $e((p-p')n\alpha)$ for distinct primes p, p'. A general polynomial sequence $(g(n)\Gamma)_{n\leqslant N}$ on a nilmanifold G/Γ may, on the other hand, not even be equidistributed. This problem is resolved by an application of the factorization theorem for polynomial sequences from [Green and Tao 2012b], which allows us to assume that $(g(n)\Gamma)_{n\leqslant N}$ is equidistributed in G/Γ if f is equidistributed in progressions to small moduli. The latter will be arranged for by employing a W-trick. As above, we then consider the following expression, which we split into sums over large and small primes, respectively, with respect to a suitable cutoff parameter X:

$$\frac{1}{N}\sum_{mp\leqslant N} f(m)f(p)\Lambda(p)F(g(mp)\Gamma) = \frac{1}{N}\sum_{m\leqslant X}\sum_{p\leqslant N/m} f(m)f(p)\Lambda(p)F(g(mp)\Gamma) + \frac{1}{N}\sum_{m>X}\sum_{p\leqslant N/m} f(m)f(p)\Lambda(p)F(g(mp)\Gamma).$$

Applying Cauchy–Schwarz to both terms shows that it suffices to understand correlations of the form

$$\sum_{m,m'} f(m)f(m')\sum_{p}\Lambda(p)F(g(mp)\Gamma)\overline{F(g(m'p)\Gamma)}$$

and

$$\sum_{p,p'} f(p)f(p')\Lambda(p)\Lambda(p')\sum_{m}F(g(pm)\Gamma)\overline{F(g(p'm)\Gamma)}.$$

Choosing X suitably, only the first of these correlations matters. We shall bound this correlation by employing Green and Tao's result that the W-tricked von Mangoldt function is orthogonal to nilsequences. The necessary equidistribution properties of the sequences $n \mapsto F(g(mn)\Gamma)\overline{F(g(m'n)\Gamma)}$ will be established in Sections 7 and 8. The problem of extending the above method to functions from $\mathscr{M}_H$ will be addressed at the beginning of Section 9. For this purpose the moment condition (1-4) will be replaced by Lemma 1.8.

3. A suitable Dirichlet decomposition for $f \in \mathscr{M}_h$

In this section we prove Lemma 1.8, which shows that every function $f \in \mathscr{M}_H$ has a decomposition $f = f_1 * \cdots * f_H$ into multiplicative functions f_i such that the L^2-norms of the f_i are controlled on average by the mean value of f. This lemma will replace the much more restrictive condition (1-4) in our application of the Montgomery–Vaughan approach outlined in the previous section. Before we prove Lemma 1.8, let us record a straightforward consequence of [Shiu 1980, Theorem 1] that will be used.

Lemma 3.1 (Shiu). *Let H be a positive integer and suppose $f : \mathbb{N} \to \mathbb{R}$ is a nonnegative multiplicative function satisfying $f(p^k) \leqslant H^k$ at all prime powers p^k. Let $W = W(x)$ be as before, let $q > 0$ be an integer and let $A' \in (\mathbb{Z}/Wq\mathbb{Z})^*$. Then*

$$\sum_{\substack{x-y<n\leqslant x\\ n\equiv A' \pmod{Wq}}} f(n) \ll \frac{y}{\phi(Wq)}\frac{1}{\log x}\exp\biggl(\sum_{\substack{w(x)<p\leqslant x\\ p\nmid q}}\frac{f(p)}{p}\biggr), \tag{3-1}$$

uniformly in A', q and y, provided that $q \leqslant y^{1/2}$ and $x^{1/2} \leqslant y \leqslant x$.

Proof. This lemma differs from [Shiu 1980, Theorem 1] in that it does not concern short intervals but at the same time it does not require f to satisfy $f(n) \ll_\varepsilon n^\varepsilon$. Shiu's result works with a summation range of the form $x - y < n \leqslant x$, where $x^\beta < y \leqslant x$, $\beta \in \bigl(0, \frac{1}{2}\bigr)$. Thus, in our case the parameter β can be regarded as fixed. As observed in [Nair and Tenenbaum 1998], the proof of [Shiu 1980, Theorem 1] only requires the condition $f(n) \ll_\varepsilon n^\varepsilon$ to hold for one fixed value of ε once β is fixed.

Note that any integer $n \equiv A' \pmod{Wq}$ is free from prime divisors $p < w(x)$. Thus, $f(n) \leqslant H^{\Omega(n)} \leqslant n^{\log H/\log w(x)}$. Given any $\varepsilon > 0$, we deduce that $n \equiv A' \pmod{Wq}$ implies $f(n) \leqslant n^{\varepsilon}$ provided x is sufficiently large. □

Proof of Lemma 1.8. Let h and h' be as in the statement of the lemma. We begin by showing that

$$|h'(p^k)| \leqslant (2H)^k, \tag{3-2}$$

using induction. Since $h'(p) = 0$, the inequality holds for $k = 1$. To analyze the general case, note that, since $h(p^k) = 0$ whenever $k \geqslant 2$, we have

$$h^{*H}(p^k) = \binom{H}{k} h^k(p) = \binom{H}{k} \frac{f^k(p)}{H^k}$$

for $1 \leqslant k \leqslant H$, and $h^{*H}(p^k) = 0$ if $k > H$. Thus, $f = h' * h^{*H}$ implies that

$$h'(p^k) = f(p^k) - \sum_{j=1}^{\min(k,H)} h'(p^{k-j}) h^{*H}(p^j).$$

Suppose now that $k \geqslant 2$ and that the inequality holds for all $j < k$. Then, invoking also (1) of Definition 1.3, we have

$$|h'(p^k)| < H^k + \sum_{j=1}^{\min(k,H)} (2H)^{k-j} \binom{H}{j} < (2H)^k \biggl(2^{-k} + \sum_{j=1}^{\min(k,H)} \frac{1}{j!2^j} \biggr) < (2H)^k,$$

as claimed.

To prove (1-7), suppose that $f_H = h$ or $h * h'$. By Shiu's bound, we have

$$\frac{DQ}{x} \sum_{\substack{n \leqslant x/D \\ n \equiv A \pmod{Q}}} f_H^2(n) \ll \frac{1}{\log(x/D)} \frac{Q}{\phi(Q)} \prod_{\substack{p \leqslant x \\ p \nmid Q}} \biggl(1 + \frac{|f(p)|}{Hp} \biggr),$$

where we used the trivial inequality $f_H(p)^2 \leqslant |f_H(p)| = |f(p)|/H$ and extended the product over primes up to x. Multiplying the right-hand side with

$$\frac{Q}{\phi(Q)} \prod_{\substack{p \leqslant x \\ p \nmid Q}} \biggl(1 + \frac{|f(p)|}{Hp} \biggr) \gg 1,$$

and observing that $\log(x/D) \asymp_H \log x$, we obtain

$$\sqrt{\frac{DQ}{x} \sum_{\substack{n \leqslant x/D \\ n \equiv A \pmod{Q}}} f_i^2(n)} \ll_H \frac{1}{(\log x)^{1/2}} \frac{Q}{\phi(Q)} \prod_{\substack{p \leqslant x \\ p \nmid Q}} \biggl(1 + \frac{|f(p)|}{Hp} \biggr).$$

Thus, the left-hand side of (1-7) is bounded by

$$\begin{aligned}
&\ll_H \frac{1}{(\log x)^{1/2}}\frac{Q}{\phi(Q)}\prod_{\substack{p\leqslant x\\ p\nmid Q}}\Big(1+\frac{|f(p)|}{Hp}\Big)\sum_{\substack{D\leqslant x^{1-1/H}\\ \gcd(D,Q)=1}}\ \sum_{d_1\cdots d_{H-1}=D}\frac{|f_1(d_1)\cdots f_{H-1}(d_{H-1})|}{D}\\
&\ll_H \frac{1}{(\log x)^{1/2}}\frac{Q}{\phi(Q)}\prod_{\substack{p\leqslant x\\ p\nmid Q}}\Big(1+\frac{|f(p)|}{Hp}\Big)\Big(1+\frac{(H-1)|f(p)|}{Hp}\Big)\Big(1+\sum_{k\geqslant 2}\frac{|f_1*\cdots*f_{H-1}(p^k)|}{p^k}\Big)\\
&\ll_H \frac{1}{(\log x)^{1/2}}\frac{Q}{\phi(Q)}\prod_{\substack{p\leqslant x\\ p\nmid Q}}\Big(1+\frac{|f(p)|}{p}\Big)\Big(1+\frac{(H-1)}{p^2}\Big)\Big(1+\sum_{k\geqslant 2}\frac{|f_1*\cdots*f_{H-1}(p^k)|}{p^k}\Big).
\end{aligned}$$

The above is now seen to have the claimed bound

$$\ll_H \frac{1}{(\log x)^{1/2}}\frac{Q}{\phi(Q)}\prod_{\substack{p\leqslant x\\ p\nmid Q}}\Big(1+\frac{|f(p)|}{p}\Big)$$

for all sufficiently large x, providing

$$\sum_{w(x)<p\leqslant x}\ \sum_{k\geqslant 2}\frac{|f_1*\cdots*f_{H-1}(p^k)|}{p^k}\ll_H 1.$$

To show the latter, note that $f_1*\cdots*f_{H-1}$ equals either $h^{*(H-1)}$ or $h^{*(H-1)}*h'$. Similarly as in the first part of this proof, we have

$$|h^{*(H-1)}(p^k)|\leqslant\binom{H-1}{k}\leqslant\frac{H^k}{k!}$$

for $k<H$ and $h^{*(H-1)}(p^k)=0$ for $k\geqslant H$, and, consequently,

$$\begin{aligned}
|(h^{*(H-1)}*h')(p^k)| &= \sum_{j=0}^{\min(k,H-1)}|h'(p^{k-j})h^{*(H-1)}(p^j)|\\
&\leqslant \sum_{j=0}^{\min(k,H-1)}(2H)^{k-j}H^j\leqslant 2(2H)^k.
\end{aligned}$$

Thus, if x is large enough that $w(x)>4H$, then

$$\begin{aligned}
\sum_{w(x)<p\leqslant x}\ \sum_{k\geqslant 2}\frac{|f_1*\cdots*f_{H-1}(p^k)|}{p^k} &\leqslant 2\sum_{w(x)<p\leqslant x}\ \sum_{k\geqslant 2}\Big(\frac{2H}{p}\Big)^k\leqslant 8H^2\sum_{w(x)<p\leqslant x}\frac{1}{p^2}\Big(1-\frac{2H}{p}\Big)^{-1}\\
&\leqslant 16H^2\sum_{w(x)<p\leqslant x}\frac{1}{p^2}\ll_H 1,
\end{aligned}$$

which completes the proof. □

4. Multiplicative functions in progressions: The class of functions $\mathscr{F}_H$

Both of the conditions that define $\mathscr{M}_H$ are natural and simple conditions on the behavior of $|f|$ at prime powers. Our aim in this section is to discuss the more complicated stability condition (1-5) on mean values in progressions, that defines the class $\mathscr{F}_H$. We will prove two sufficient conditions, recorded in Propositions 4.4 and 4.10, for the bound (1-5) to hold and apply these to provide several examples of natural functions that belong to $\mathscr{F}_H$. In particular, we deduce a simple criterion, see Lemma 4.16, for a nonnegative function to belong to $\mathscr{F}_H$.

4A. ***A sufficient condition for $f \in \mathscr{F}_H$.*** The main tool in our analysis of (1-5) will be the following consequence of the "pretentious large sieve", which lets one bound the tail of the character sum expansion of $S_h(x; q, a)$ for any bounded multiplicative function h and thereby simplifies the task of analyzing the expression $S_h(x; q, a)$.

Proposition 4.1 ([Granville and Soundararajan $\geq$ 2018]; cf. [Balog et al. 2013, Lemma 3.1; Granville 2009, Theorem 2; Granville et al. 2017, Theorem 1.8). *Let $C > 0$ be fixed and let h be a bounded multiplicative function. For any given x, consider the set of primitive characters of conductor at most $(\log x)^C$ and enumerate them as $\chi_1, \chi_2, \ldots$ in such a way that $|S_{h\overline{\chi}_1}(x)| \geqslant |S_{h\overline{\chi}_2}(x)| \geqslant \cdots$. If x is sufficiently large, then the following holds for all $x^{1/2} \leqslant X \leqslant x$ and $q \leqslant (\log x)^C$. Let $\mathscr{C}$ be any set of characters modulo q, $q \leqslant (\log x)^C$, which does not contain characters induced by $\chi_1, \ldots, \chi_k$, where $k \geqslant 2$. Then*

$$\left|\frac{1}{\phi(q)} \sum_{\chi \in \mathscr{C}} \chi(a) \sum_{n \leqslant X} h(n)\overline{\chi}(n)\right| \ll_C \frac{e^{O_C(\sqrt{k})} X}{q} \left(\frac{\log\log x}{\log x}\right)^{1 - \frac{1}{\sqrt{k}}} \log \frac{\log x}{\log\log x} \prod_{p \leqslant q,\, p \nmid q} \left(1 + \frac{|h(p)| - 1}{p}\right).$$

To deduce a sufficient condition for (1-5) we first extend this result to all unbounded elements of $\mathscr{M}_H$.

Corollary 4.2. *Let $f \in \mathscr{M}_H$ and set $h = f$ if $H = 1$. If $H > 1$, let h be the multiplicative function defined in (1-6) so that $f = h^{*H} * h'$ for a multiplicative function h' with support in the square-full numbers. Let $C > 0$ be a constant, let $\varepsilon = \frac{1}{2}\min(1, \alpha_f/H)$, and set $k = \lceil \varepsilon^{-2} \rceil \geqslant 2$ and $k' = \lceil \log_2(4H) \rceil$. For each $j \in \{0, \ldots, k'\}$, let $\mathscr{E}_j = \{\chi_1^{(j)}, \ldots, \chi_k^{(j)}\}$ denote the set consisting of the first k primitive characters of conductor at most $(\log x^{1/2^j})^C$ defined by Proposition 4.1 when applied to h and with x replaced by $x^{1/2^j}$.*

If x is sufficiently large, the following holds for all $x^{1/2} \leqslant y \leqslant x$ and all integer multiplies $0 < Q \leqslant (\log x^{1/(8H)})^C$ of $W(x)$. Let $\mathscr{C}$ be any set of characters modulo Q which does not contain characters induced by any $\chi \in \mathscr{E} := \mathscr{E}_0 \cup \cdots \cup \mathscr{E}_{k'}$, then

$$\left|S_f(y; Q, a) - \frac{Q}{y}\frac{1}{\phi(Q)} \sum_{\substack{\chi \pmod{Q} \\ \chi \notin \mathscr{C}}} \chi(a) \sum_{n \leqslant y} f(n)\overline{\chi}(n)\right|$$
$$= \frac{Q}{\phi(Q)}\left|\sum_{\chi \in \mathscr{C}} \chi(a) \frac{1}{y} \sum_{n \leqslant y} f(n)\overline{\chi}(n)\right| \ll_{C,H,\alpha_f} \frac{1}{(\log x)^{1+\alpha_f/(3H)}} \frac{Q}{\phi(Q)} \exp\left(\sum_{p \leqslant x,\, p \nmid Q} \frac{|f(p)|}{p}\right).$$

The proof of this corollary makes use of the following lemma about the contribution of the sparse function h'.

Lemma 4.3. *Let $H>1$, $f\in\mathscr{M}_H$ and let $f=h^{*H}*h'$ be the decomposition from (1-6). Let $g:\mathbb{N}\to\mathbb{C}$ be a bounded completely multiplicative function that vanishes at all primes $p\leqslant w$ for a fixed $w\geqslant(2H)^{16}$, and let $\delta\in(0,1)$. Then, if $x^{1/2}\leqslant y\leqslant x$, we have*

$$\left|\frac{1}{y}\sum_{n\leqslant y}f(n)b(n)\right|\leqslant\sum_{n_1\leqslant y^\delta}\frac{|h'(n_1)b(n_1)|}{n_1}\left|\frac{n_1}{y}\sum_{n_2\leqslant\frac{y}{n_1}}h^{*H}(n_2)b(n_2)\right|+O(x^{-\frac{\delta}{8}}(\log y)^{O(H)}).$$

Proof. Recall that h' is supported on square-full numbers only and that $|h'(p^k)|\leqslant(2H)^k$ by (3-2). Since b is completely multiplicative, we have

$$\begin{aligned}\sum_{n\leqslant y}f(n)b(n)&=\sum_{n_1n_2\leqslant y}h'(n_1)h^{*H}(n_2)b(n_1)b(n_2)\\&\leqslant\sum_{n_1\leqslant y^\delta}|h'(n_1)b(n_1)|\left|\sum_{n_2\leqslant y/n_1}h^{*H}(n_2)b(n_2)\right|+y\sum_{n_1>y^\delta}\frac{|h'(n_1)b(n_1)|}{n_1}\sum_{n_2\leqslant y/n_1}\frac{|h^{*H}(n_2)|}{n_2}\\&\leqslant\sum_{n_1\leqslant y^\delta}|h'(n_1)b(n_1)|\left|\sum_{n_2\leqslant y/n_1}h^{*H}(n_2)b(n_2)\right|+y(\log y)^{O(H)}\sum_{\substack{n_1\geqslant y^\delta\\p|n_1\Rightarrow p>w}}\frac{|h'(n_1)|}{n_1}.\end{aligned}$$

By decomposing every square-full number n_1 as m^2d with $d\mid m$, we obtain the following bound for the sum in the final term:

$$\begin{aligned}\sum_{\substack{n_1\geqslant y^\delta\\p|n_1\Rightarrow p>w}}\frac{|h'(n_1)|}{n_1}&\leqslant\sum_{\substack{m\geqslant y^{\delta/3}\\p|m\Rightarrow p>w}}\frac{(2H)^{\Omega(m^2)}}{m^2}\sum_{d|m}\frac{(2H)^{\Omega(d)}}{d}\leqslant\sum_{\substack{m\geqslant y^{\delta/3}\\p|m\Rightarrow p>w}}\frac{(2H)^{\frac{2\log m}{\log w}}}{m^2}d(m)\\&\ll\sum_{\substack{m\geqslant y^{\delta/3}\\p|m\Rightarrow p>w}}m^{-2+\frac{2\log(2H)}{\log w}+\frac{1}{8}}\ll\sum_{\substack{m\leqslant y^{\delta/3}\\p|m\Rightarrow p>w}}m^{-2+\frac{1}{4}}\ll y^{-\frac{\delta}{4}},\end{aligned}\tag{4-1}$$

where we used the bound $d(n)\ll n^{1/8}$. Combining the two bounds above completes the proof. □

Proof of Corollary 4.2. To start with, we consider the bounded multiplicative function h. Note that Proposition 4.1 applies to values of X with $x^{1/2}\leqslant X\leqslant x$. Our application, will, however, require a range of the form $x^{1/(4H)}\leqslant X\leqslant x$. For this reason, we will apply Proposition 4.1 once with x replaced by $x^{1/2^j}$ for each $j\in\{0,\dots,\lceil\log_2(4H)\rceil\}$. If $\mathscr{C}$ is as in the statement of the corollary, then Proposition 4.1 shows that for all $Q\leqslant(\log x^{1/(8H)})^C$ and for all $x^{1/(4H)}<X\leqslant x$, we have

$$\begin{aligned}\frac{1}{X}\frac{Q}{\phi(Q)}\left|\sum_{\chi\in\mathscr{C}}\chi(A)\sum_{n\leqslant X}h(n)\overline{\chi}(n)\right|&\ll_{C,H,\alpha_f}\left(\frac{\log\log x}{\log x}\right)^{1-1/\sqrt{k}}\log\left(\frac{\log x}{\log\log x}\right)\\&\ll_{C,H,\alpha_f}(\log x)^{-1+\alpha_f/(2H)}(\log\log x)^2,\end{aligned}$$

since $1/\sqrt{k} = 1/\sqrt{[\varepsilon^{-2}]} \leqslant \varepsilon = \frac{1}{2}\min(1, \alpha_f/H) \leqslant \alpha_f/(2H)$.

By property (2) of Definition 1.3, we have

$$\frac{Q}{\phi(Q)} \exp\Bigl(\sum_{\substack{p\leqslant x\\ p\nmid Q}} \frac{|h(p)|}{p}\Bigr) \geqslant \exp\Bigl(\sum_{\substack{p\leqslant x\\ p\nmid Q}} \frac{|f(p)|}{Hp}\Bigr) \geqslant \Bigl(\frac{\log x}{C\log\log x}\Bigr)^{\alpha_f/H},$$

and, thus,

$$\begin{aligned}\frac{1}{X}\frac{Q}{\phi(Q)}\Bigl|\sum_{\chi\in\mathscr{C}}\chi(A)\sum_{n\leqslant X} h(n)\overline{\chi}(n)\Bigr| &\ll_{C,H,\alpha_f} \frac{(\log\log x)^{2+\alpha_f/H}}{(\log x)^{1+\alpha_f/(2H)}}\frac{Q}{\phi(Q)}\exp\Bigl(\sum_{\substack{p\leqslant x\\ p\nmid Q}} \frac{|h(p)|}{p}\Bigr)\\ &\ll_{C,H,\alpha_f} \frac{1}{(\log x)^{1+\alpha_f/(3H)}}\frac{Q}{\phi(Q)}\exp\Bigl(\sum_{\substack{p\leqslant x\\ p\nmid Q}} \frac{|h(p)|}{p}\Bigr).\end{aligned} \tag{4-2}$$

To handle the case where $H > 1$, consider the decomposition $f = h^{*H} * h'$ with h as in (1-6). If $x^{1/2} \leqslant y \leqslant x$, then Lemma 4.3 implies that for any $\delta \in (0, 1)$,

$$\frac{1}{y}\sum_{\chi\in\mathscr{C}}\chi(A)\sum_{n\leqslant y} f(n)\overline{\chi}(n) \leqslant \frac{1}{y}\sum_{\chi\in\mathscr{C}}\sum_{d_0\leqslant y^\delta}|h'(d_0)\overline{\chi}(d_0)|\Bigl|\sum_{d\leqslant y/d_0} h^{*H}(d)\overline{\chi}(d)\Bigr| + O(x^{-\delta/8}(\log x)^{O(H)}). \tag{4-3}$$

The error term in this bound is acceptable. A generalization of the hyperbola method applied to the sum over d (see Section 9A for a deduction) shows that the main term satisfies

$$\begin{aligned}&\frac{1}{y}\sum_{\chi\in\mathscr{C}}\sum_{d_0\leqslant y^\delta}|h'(d_0)\overline{\chi}(d_0)|\Bigl|\sum_{d\leqslant y/d_0} h^{*H}(d)\overline{\chi}(d)\Bigr|\\ &\quad\leqslant \sum_{\substack{d_0\leqslant y^\delta\\ p\mid d_0\Rightarrow p>w(x)}} \frac{|h'(d_0)|}{d_0}\sum_{D\leqslant (y/d_0)^{1-1/H}}\sum_{d_1\cdots d_{H-1}=D}\frac{|h(d_1)\cdots h(d_{H-1})||\overline{\chi}(D)|}{D}\\ &\qquad\times\sum_{i=1}^{H}\frac{Dd_0}{y}\Bigl|\sum_{\chi\in\mathscr{C}}\chi(A)\sum_{\substack{n:\\ (y/d_0)^{1-1/H}\max(d_1,\dots,d_{i-1})\\ \leqslant Dn\leqslant y/d_0}} h(n)\overline{\chi}(n)\Bigr|.\end{aligned} \tag{4-4}$$

Observe that the upper bound on n in the inner sum satisfies $y/(Dd_0) \in [y^{1/H-\delta}, x]$. By choosing $\delta = 1/(4H)$, this interval is contained in $[x^{1/(2H)-\delta/2}, x] = [x^{3/(8H)}, x]$. An application of the triangle inequality shows that the inner sum is bounded by

$$\frac{Dd_0}{y}\Bigl|\sum_{\chi\in\mathscr{C}}\chi(A)\sum_{n\leqslant y/(d_0D)} h(n)\overline{\chi}(n)\Bigr| + \frac{Dd_0}{y}\Bigl|\sum_{\chi\in\mathscr{C}}\chi(A)\sum_{n\leqslant y'} h(n)\overline{\chi}(n)\Bigr|,$$

where $y' = \min(y/(d_0D), (y/d_0)^{1-1/H}D^{-1}\max(d_1, \dots, d_{i-1}))$. We are now in a position to apply (4-2) to bound the first of these terms by

$$\ll_{C,H,\alpha_f} \frac{1}{(\log x)^{1+\alpha_f/(3H)}} \exp\biggl(\sum_{p\leqslant x,\, p\nmid Q} \frac{|h(p)|}{p}\biggr).$$

If $y' > x^{1/(4H)}$, then the same bound applies to the second term. If, on the other hand, $y' \leqslant x^{1/(4H)} \leqslant y^{1/(2H)}$, then the second term may trivially be bounded by

$$\phi(Q) y^{1/(4H)-1} d_0 D \leqslant \phi(Q) y^{1/(4H)-1+1/(4H)+1-1/H} \leqslant (\log x)^C x^{-1/(4H)}.$$

Inserting these bounds into (4-4) and completing the outer sums, we deduce that (4-4) is bounded by

$$\ll_{C,H,\alpha_f} \frac{1}{(\log x)^{1+\alpha_f/(3H)}} \exp\biggl(\sum_{\substack{p\leqslant x,\\ p\nmid Q}} \frac{|h(p)|}{p}\biggr) \sum_{\substack{d_0:p\mid d_0\\ \Rightarrow p>w(x)}} \frac{|h'(d_0)|}{d_0} \biggl(\sum_{d\leqslant x} \frac{|h(d)\chi(d)|}{d}\biggr)^{H-1}.$$

The sum over d in this bound satisfies

$$\biggl(\sum_{d\leqslant x} \frac{|h(d)\chi(d)|}{d}\biggr)^{H-1} \leqslant \prod_{p\leqslant x} \biggl(1+\frac{|h(p)\chi(p)|}{p}\biggr)^{H-1} \leqslant \exp\biggl((H-1) \sum_{p\leqslant x, p\nmid Q} \frac{|h(p)|}{p}\biggr),$$

and the sum over d_0 converges by (4-1), applied with $y=1$, provided x is sufficiently large for $w(x) \geqslant (2H)^{16}$ to hold. Collecting all information together, it follows from (4-3) that

$$\frac{1}{x}\frac{Q}{\phi(Q)} \sum_{\chi\in\mathscr{C}} \chi(A) \sum_{n\leqslant x} f(n)\bar{\chi}(n) \ll_{C,H,\alpha_f} \frac{1}{(\log x)^{1+\frac{\alpha_f}{(3H)}}} \frac{Q}{\phi(Q)} \exp\biggl(\sum_{\substack{p\leqslant x,\\ p\nmid Q}} \frac{|f(p)|}{p}\biggr),$$

which competes the proof. □

With Corollary 4.2 in place, we obtain the following sufficient condition for $f \in \mathscr{M}_H$ to belong to $\mathscr{F}_H$:

Proposition 4.4 (sufficient condition). *Suppose that $f \in \mathscr{M}_H$. Then $f \in \mathscr{F}_H$ if the following holds. For every $C>0$, there exists a function $\psi_C : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$, with the property that $\psi_C(x) \to 0$ as $x\to\infty$, such that*

$$S_{f\chi}(x') = S_{f\chi}(x) + O\biggl(\frac{\psi_C(x)}{\log x} \exp\biggl(\sum_{p\leqslant x,\, p\nmid Q} \frac{|f(p)|}{p}\biggr)\biggr), \quad (x\geqslant 2), \tag{4-5}$$

uniformly for all $x' \in (x(\log x)^{-C}, x]$ and all characters $\chi \pmod Q$ with $1 < Q \leqslant (\log x)^C$ and $W(x) \mid Q$.

Proof. Recall from Definition 1.4 that we have to show that there exists $\varphi_C = o(1)$ such that

$$|S_f(x';Q,A) - S_f(x;Q,A)| = O\biggl(\frac{\varphi_C(x)}{\log x}\frac{Q}{\phi(Q)} \prod_{p\leqslant x,\, p\nmid Q}\biggl(1+\frac{|f(p)|}{p}\biggr)\biggr) \tag{4-6}$$

uniformly for all $x' \in (x(\log x)^{-C}, x]$, all $1 \leqslant Q \leqslant (\log x)^C$ with $W(x) \mid Q$ and all reduced $A \pmod Q$. This will be a straightforward consequence of the fact that by Corollary 4.2 there are only finitely many characters in the character sum expansions of $S_f(x';Q,A)$ and $S_f(x;Q,A)$ that matter. Using the notation from the corollary, let $\mathscr{E}(Q)$ denote the set of characters modulo Q that are induced by the elements of $\mathscr{E}_0 \cup \cdots \cup \mathscr{E}_{k'}$. Then

$$S_f(x';\, Q,\, A) = \frac{Q}{\phi(Q)} \sum_{\substack{\chi \pmod{Q} \\ \chi \in \mathscr{E}(Q)}} \chi(A) S_{f\overline{\chi}}(x') + O\biggl(\frac{\psi(x)}{\log x} \frac{Q}{\phi(Q)} \exp\biggl(\sum_{\substack{p \leqslant x \\ p \nmid Q}} \frac{|f(p)|}{p} \biggr) \biggr),$$

where $\psi(x) = O_{C,H,\alpha_f}((\log x)^{-\alpha_f/(3H)})$, uniformly in x', Q and A as above. Thus, (4-6) follows from our assumptions with $\psi_C(x) = \psi(x) + \#\mathscr{E} \cdot \varphi_C(x)$. □

Example 4.5 (applications using Selberg–Delange-type arguments). The conditions required by Proposition 4.4 are of a type that can usually be checked by means of the Selberg–Delange method (see, e.g., [Tenenbaum 1995, Section II.5]) provided the function f is closely related to a ζ- or L- function. The range of the modulus Q of the characters χ that appear is small enough to ensure that exceptional characters can be handled. Examples of functions suitable for this approach include:

(i) the function $\frac{1}{4}r(n) = \frac{1}{4}\#\{(x,\, y) \in \mathbb{Z}^2 : x^2 + y^2 = n\}$,

(ii) the indicator function of the set of sums of two squares,

(iii) the characteristic function of set of numbers composed of primes that split completely in a given Galois extension $K/\mathbb{Q}$ of finite degree.

In the following subsection, we will further analyze the Lipschitz condition (4-5) and prove another sufficient condition, in this case for an element $f \in \mathscr{M}_H$ to belong to $\mathscr{F}_{H,n^{it}}$.

4B. *Lipschitz estimates for elements of $\mathscr{M}_H$ and another sufficient condition.* For applications of Proposition 4.4 or Corollary 4.2, the following four lemmas, which we all prove in Section 4C, are very useful. The first lemma is a slight generalization of the Lipschitz estimate for bounded multiplicative functions and a related decay estimate that Granville and Soundararajan established in Theorems 3 and 4 of [Granville and Soundararajan 2003].

Lemma 4.6 (Lipschitz estimates). *Let $f_0 \in \mathscr{M}_1$ and let $x \geqslant 3$. Suppose that $f : \mathbb{N} \to \mathbb{C}$ is multiplicative, bounded in absolute value by* 1 *and satisfies $|f(p^k)| = |f_0(p^k)|$ for all primes $p \geqslant \exp((\log\log x)^2)$ and $k \geqslant 1$. Define*

$$F(s) = \prod_{p \leqslant x} \Bigl(1 + \frac{f(p)}{p^s} + \frac{f(p^2)}{p^{2s}} + \cdots \Bigr).$$

If the maximum of $\max_{|y| \leqslant 2\log x} |F(1+iy)|$ *is attained at $y = t_{x,f}$, then, uniformly in x and f as above, we have*

$$\biggl| \frac{1}{x} \sum_{n \leqslant x} f(n) n^{-it_{x,f}} - \frac{1}{x'} \sum_{n \leqslant x'} f(n) n^{-it_{x,f}} \biggr| \ll_{f_0} \frac{1}{(\log x)^{1+C_0}} \exp\biggl(\sum_{p \leqslant x} \frac{|f(p)|}{p} \biggr) \tag{4-7}$$

for all $x' \in [x\exp(-(\log\log x)^{-4}), x]$, *where* $C_0 \in \left(0, \frac{1}{2}\alpha_{f_0}\right)$ *is a positive constant that only depends on* α_{f_0}. *Furthermore, we have, for any* $t_{x,f}$ *as above,*

$$\left|\frac{1}{x}\sum_{n\leqslant x} f(n)\right| \ll_{f_0} \frac{1}{|t_{x,f}|+1} + \frac{\log\log x}{\log x} + \frac{1}{(\log x)^{1+C_0}}\exp\left(\sum_{p\leqslant x}\frac{|f(p)|}{p}\right). \tag{4-8}$$

The conditions on f above will allow us to apply the lemma to twists $h\chi$ where $h \in \mathscr{M}_1$ and χ is a character modulo Q with $Q \leqslant (\log x)^C$ for any given constant $C > 0$. In order to extend this to twists $f\chi$ for $f \in \mathscr{M}_H$ and $H > 1$, we note that the function h associated to f via (1-6) belongs to $\mathscr{M}_1$. The following lemma will enable us to employ Lemma 4.6 in general.

Lemma 4.7. *Let* $f \in \mathscr{M}_H$ *and let* h *be as in* (1-6). *Let* $C > 0$ *be a fixed constant and let* $\psi : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$ *be a function that satisfies* $\psi(x) \to 0$ *as* $x \to \infty$. *Let* $x > 1$ *and suppose that* $\chi \pmod Q$, *with* $Q \leqslant (\log x)^C$ *and* $W(x) \mid Q$, *is a character such that*

$$|S_{h\chi}(y) - S_{h\chi}(y')| \leqslant \frac{\psi(x)}{\log y}\exp\left(\sum_{p\leqslant y,\, p\nmid Q}\frac{|h(p)|}{p}\right)$$

for all $y \in (x^{1/(2H)}, x]$ *and* $y' \in (y(\log x)^{-C}, y]$. *Then, for all* $x' \in (x(\log x)^{-C}, x]$, *we have*

$$|S_{f\chi}(x) - S_{f\chi}(x')| \leqslant \frac{\psi'(x)}{\log x}\exp\left(\sum_{p\leqslant x,\, p\nmid Q}\frac{|f(p)|}{p}\right),$$

where ψ' *is independent of* χ *and* Q *and satisfies* $\psi'(x) \to 0$ *as* $x \to \infty$; *more precisely,*

$$\psi'(x) = O_{H,C}\left(\psi(x) + (\log x)^{-\min(1,\alpha_f/(2H))} + x^{-1/8}(\log x)^{O(H)}\right).$$

The next lemma shows that if χ is a character that is negligible in the application of Proposition 4.4 or Corollary 4.2 to the function h, then χ is also negligible in an application of the result to f.

Lemma 4.8. *Let* $f \in \mathscr{M}_H$, *let* h *be as in* (1-6), *and let* $C > 0$. *Let* $\psi : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$ *be a function that satisfies* $\psi(x) \to 0$ *as* $x \to \infty$. *Let* $x > 1$ *and suppose that* $\chi \pmod Q$, *with* $Q \leqslant (\log x)^C$ *and* $W(x) \mid Q$, *is any character such that*

$$|S_{h\chi}(x')| \leqslant \frac{\psi(x)}{\log y}\exp\left(\sum_{p\leqslant y,\, p\nmid Q}\frac{|h(p)|}{p}\right)$$

for all $x' \in [x^{1/(4H)}, x]$. *Then*

$$|S_{f\chi}(y)| \leqslant \frac{\psi'(x)}{\log x}\exp\left(\sum_{p\leqslant x,\, p\nmid Q}\frac{|f(p)|}{p}\right), \quad (y \in [x^{1/2}, x]),$$

where ψ' *is independent of* χ *and* Q *and satisfies*

$$\psi'(x) = O(\psi(x) + x^{-1/(32H)}(\log x)^{O(H)}).$$

Finally, we observe that (4-7) holds *uniformly* in f for some $t = t_x$ that only depends on x and f_0. This proves particularly valuable when dealing with families of induced characters.

Lemma 4.9. *Let x, f_0 and f be as in Lemma 4.6, and let $x' \in \big[x\exp\big(-\frac{1}{2}(\log\log x)^4\big), x\big]$. Then there exists $|t_x| \leqslant 2\log x$, only dependent on x and f_0, but not on f or x', such that, uniformly in x, x' and f as before,*

$$\left|\frac{1}{x}\sum_{n\leqslant x} f(n)n^{-it_x} - \frac{1}{x'}\sum_{n\leqslant x'} f(n)n^{-it_x}\right| \ll \frac{1}{(\log x)^{1+C_0}}\exp\Big(\sum_{p\leqslant x}\frac{|f(p)|}{p}\Big), \tag{4-9}$$

where $C_0 > 0$ is a positive constant that only depends on α_{f_0}.

As a consequence of Lemmas 4.6–4.9 and Corollary 4.2, we obtain the following sufficient condition for testing whether a function belongs to $\mathscr{F}_{H,n^{it}}$.

Proposition 4.10 (another sufficient condition). *Let $f \in \mathscr{M}_H$ and h as in (1-6). For every $C > 0$, let $\psi_C : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$ be a function that satisfies $\psi_C(x) \to 0$ as $x \to \infty$. Suppose that for every sufficiently large x there exists $\tau_x \in \mathbb{R}$ with $|\tau_x| \leqslant 2\log x$ such that the following holds: If $1 \leqslant Q \leqslant (\log x)^C$ with $W(x) \mid Q$ and if $\chi \pmod Q$ is a character then either the bound*

$$|S_{g_x\chi}(x')| \leqslant \frac{\psi_C(x')}{\log x'}\exp\Big(\sum_{p\leqslant x',\, p\nmid Q}\frac{|g(p)|}{p}\Big), \qquad (x^{1/(8H)} \leqslant x' \leqslant x), \tag{4-10}$$

holds for either $g = h$ or $g = f$ and for $g_x : n \mapsto g(n)n^{-i\tau_x}$, or else we have $t_{x,f} = \tau_x$ or $t_x = \tau_x$ in the statement of Lemmas 4.6 or 4.9 when applied with $f_0 = h$ and with f replaced by $h\chi$.

Then the function $n \mapsto f(n)n^{-i\tau_x}$ satisfies (1-5) and $f \in \mathscr{F}_{H,n^{it}}$.

Example 4.11. The above proposition applies to:

(i) the Möbius function $f = \mu$. In this case we may take $\tau_x = 0$ for all x since $S_{\mu\chi}(x) \ll_B q^{1/2}(\log x)^{-B}$ for all $B > 0$ and all $\chi \pmod q$.[2] Indeed, if χ is a trivial character this estimate follows from prime number theorem-type bounds on $S_\mu(x)$; see Example 4.18(i) for details. If $\chi \pmod q$ is nontrivial, then its conductor, q' say, is at least 2 and one may deduce the estimate from [Iwaniec and Kowalski 2004, Corollary 5.29], which proves the claimed bound for nontrivial primitive characters. In fact, if $q \leqslant x$, then it follows from [loc. cit., (5.79)] that

$$\sum_{p\leqslant x}\chi(p) \ll_B q'^{1/2}x(\log x)^{-B} + \omega(q) \ll_B q^{1/2}x(\log x)^{-B},$$

since $\omega(q) \ll \log x$. If $q > x$, then

$$\sum_{p\leqslant x}\chi(p) \ll_B q^{1/2}x(\log x)^{-B}$$

holds trivially. Thus, [loc. cit., (5.79)] generalizes to all nontrivial χ and, by following the original proof from [loc. cit.], so does [loc. cit., (5.80)].

[2]This is the same information about μ as was used in [Green and Tao 2012a, Proposition A.1] to handle the "major arcs".

(ii) every multiplicative function f that takes values on the unit circle, i.e., for which $|f(n)| = 1$ for all n. This, in turn, follows from [Balog et al. 2013, Theorem 2], which provides the bound

$$|S_{f\chi}(x)| \ll \big((\log\log x)^2/\log x\big)^{1/20}, \tag{4-11}$$

valid for all characters χ of conductor $Q \leqslant \exp((\log\log x)^2)$, except perhaps for those induced by one exceptional character, ξ say. By Lemma 4.9, there exists $|t_x| \leqslant 2\log x$ such that (4-9) holds for all $\chi \pmod Q$ induced by ξ. Suppose now that $|t_x| > (\log x)^{1/100}$. Then Lemma 4.9, combined with the bound (4-8), implies that the above bound on $|S_{f\chi}(x)|$ also holds for characters induced by ξ. In this case, we may take $\tau_x = 0$. If, however, $|t_x| \leqslant (\log x)^{1/100}$, then we may use partial summation to deduce from (4-11) that

$$|S_{f_x\chi}(x)| \ll \big((\log\log x)^2/\log x\big)^{3/100}$$

for all χ not induced by ξ. In this case, we may set $\tau_x = t_x$.

Proof of Proposition 4.10. This result follows from Corollary 4.2 in a similar way as Proposition 4.4 does. To show that $n \mapsto f_x(n) := f(n)n^{-i\tau_x}$ satisfies (1-5), let $1 \leqslant Q \leqslant (\log x)^C$ be such that $W(x) \mid Q$ and let $\mathscr{E}(Q)$ denote the set of characters modulo Q that are induced by the elements of $\mathscr{E}_0 \cup \dots \cup \mathscr{E}_{k'}$ from Corollary 4.2, when applied to the function f_x. Then

$$\begin{aligned} S_{f_x}(y; Q, A) - S_{f_x}(x; Q, A) &= \frac{Q}{\phi(Q)} \sum_{\substack{\chi \ (\mathrm{mod}\ Q) \\ \chi \in \mathscr{E}(Q)}} \chi(A)(S_{f_x\bar\chi}(y) - S_{f_x\bar\chi}(x)) \\ &\quad + O_{C,H,\alpha_f}\bigg((\log x)^{-\alpha_f/(3H)} \frac{Q}{\phi(Q)} \frac{1}{\log x} \exp\bigg(\sum_{p \leqslant x,\, p\nmid Q} \frac{|f(p)|}{p}\bigg)\bigg), \end{aligned} \tag{4-12}$$

whenever $x^{1/2} \leqslant y \leqslant x$.

We begin with the contribution from those characters $\bar\chi$ to which the first alternative from the statement applies. Observe that (4-10) implies that

$$|S_{g_x\chi}(x')| \leqslant \frac{\psi'_C(x)}{\log x} \exp\bigg(\sum_{p \leqslant x,\, p\nmid Q} \frac{|g(p)|}{p}\bigg), \qquad (x^{1/(8H)} \leqslant x' \leqslant x),$$

where $\psi'_C(x) = O_H(1) \max_{x^{1/(8H)} \leqslant x' \leqslant x} \psi_C(x')$. To see this, note that for all x' as above,

$$\begin{aligned} \exp\bigg(\sum_{p \leqslant x,\, p\nmid Q} \frac{|g(p)|}{p}\bigg) \exp\bigg(-\sum_{p \leqslant x',\, p\nmid Q} \frac{|g(p)|}{p}\bigg) &\leqslant \exp\bigg(\sum_{x^{1/(8H)} < p \leqslant x} \frac{H}{p}\bigg) \\ &\leqslant \exp\big(H(\log\log x + \log(8H) - \log\log x + o(1))\big) \\ &\leqslant (8H)^{H(1+o(1))} \ll_H 1. \end{aligned}$$

Thus, by Lemma 4.8 it follows that all characters $\overline{\chi}$ with $\chi \in \mathscr{E}(Q)$ to which the first alternative from the statement applies satisfy

$$|S_{f_x\overline{\chi}}(y)| \leqslant \frac{\psi''_C(x)}{\log x} \exp\Bigl(\sum_{p\leqslant x,\, p\nmid Q} \frac{|f(p)|}{p}\Bigr), \qquad (x^{1/2} \leqslant y \leqslant x), \tag{4-13}$$

for a suitable function $\psi''_C = o(1)$.

For all remaining $\chi \in \mathscr{E}(Q)$, Lemma 4.6 or 4.9 provides the Lipschitz estimate

$$S_{f_x\overline{\chi}}(y) = S_{f_x\overline{\chi}}(x) + O\biggl(\frac{\psi'''(x)}{\log x} \exp\Bigl(\sum_{p\leqslant x,\, p\nmid Q} \frac{|f(p)|}{p}\Bigr)\biggr),$$

for $y \in [x\exp(-(\log\log x)^4/2), x]$, with $\psi'''(x) = (\log x)^{-C_0}$. Thus, the result follows from (4-12). □

4C. *Proofs of Lemmas 4.6–4.9.* We prove Lemmas 4.6, 4.9, 4.7 and 4.8, in this order.

Proof of Lemma 4.6. The proof of this lemma is almost identical to the proofs of the original results of Granville and Soundararajan [2003, Theorem 3 and 4], except for one ingredient: their Lemma 2.3 needs to be replaced by Lemma 4.12 below. The estimate (4-8) follows immediately from [loc. cit., §5] and Lemma 4.12. Concerning the Lipschitz estimate (4-7), we replace the application of [loc. cit., Theorem 3] at the beginning of [loc. cit., §6] by the estimate (4-8). The bound in [loc. cit., equation (6.2)] continues to apply. The first term in this bound is acceptable since in our case $w \leqslant \exp((\log\log x)^4)$, and since $C_0 < \alpha_{f_0}$. To bound the integrand in the second term, we use the bound [loc. cit., equation (6.5)] if α is large, which in our situation means that $\alpha > \exp\bigl(\sum_{p\leqslant x} |f(p)|/p\bigr)^{-1}(\log x)^{C_0}$ with $C_0 = C_0(\alpha_{f_0}, 1)$ as in the lemma below. If $\alpha \leqslant \exp\bigl(\sum_{p\leqslant x} |f(p)|/p\bigr)^{-1}(\log x)^{C_0}$, we proceed as in the small-α-case from the original proof but, again, apply our Lemma 4.12 instead of [loc. cit., Lemma 2.3]. □

Lemma 4.12 ("new Lemma 2.3"). *Let $x \geqslant 3$, $f_0 \in \mathscr{M}_H$ and let $f : \mathbb{N} \to \mathbb{C}$ be a multiplicative function such that $|f(p^k)| \leqslant |f_0(p^k)|$ at all prime powers p^k, and such that $|f(p^k)| = |f_0(p^k)|$ whenever $p \geqslant \exp((\log\log x)^2)$ and $k \in \mathbb{N}$. Let*

$$F(s) = \prod_{p\leqslant x} \Bigl(1 + \frac{f(p)}{p^s} + \frac{f(p^2)}{p^{2s}} + \cdots\Bigr).$$

Then there exists a positive constant $C_0 = C_0(\alpha_{f_0}, H) \in (0, \alpha_f/2)$ such that for all real numbers y and $1/\log x \leqslant |\beta| \leqslant \log x$, we have

$$|F(1+iy)F(1+i(y+\beta))| \ll \exp\Bigl(2\sum_{p\leqslant x} \frac{|f(p)|}{p}\Bigr)(\log x)^{-2C_0}.$$

Remark. Observe that we actually only use this lemma in the case where $H = 1$.

Proof. Suppose $|f(p)| = g(p) + h(p)$ for two nonnegative functions g and h. Then

$$\begin{aligned} |F(1+iy)F(1+i(y+\beta))| &\ll \exp\biggl(\Re \sum_{p\leqslant x} \frac{f(p)p^{-iy} + f(p)p^{-i(y+\beta)}}{p}\biggr) \\ &\ll \exp\biggl(\sum_{p\leqslant x} \frac{(g(p)+h(p))|1+p^{-i\beta}|}{p}\biggr) \\ &\ll \exp\biggl(2\sum_{p\leqslant x} \frac{g(p)|\cos\bigl(\frac{1}{2}|\beta|\log p\bigr)|}{p}\biggr)\exp\biggl(2\sum_{p\leqslant x}\frac{h(p)}{p}\biggr). \end{aligned} \tag{4-14}$$

The aim is to exploit the fact that $|\cos|$ is not the constant function 1 in order to bound this expression. We begin by decomposing the set of primes less than x into subsets on which $\bigl|\cos\bigl(\frac{1}{2}|\beta|\log p\bigr)\bigr|$ is almost constant. For this purpose, let $\delta = 1/(\log x)^3$ and consider the decomposition of $[0, 2\pi)$ into intervals of the form $\bigl(\frac{1}{2}(n-1)\log(1+\delta)|\beta|, \frac{1}{2}n\log(1+\delta)|\beta|\bigr]$. Thus, in order to cover the interval $[0, 2\pi)$, the parameter n runs over the range $1 \leqslant n \leqslant N$, for some $N \asymp (\delta|\beta|)^{-1}$, and, in particular, we have $(\log x)^2 \ll N \ll (\log x)^4$. By changing δ slightly, we can insure that $\frac{1}{2}N\log(1+\delta)|\beta| = 2\pi$ so that the decomposition of $[0, 2\pi)$ has exactly N full intervals and no smaller or larger ones. Next, we set $Y = \exp((\log\log x)^2)$ and decompose the set of primes in the interval $[Y, x]$ into N sets of the form

$$P_n(x) = \bigcup_{m\equiv n \pmod N} \{p \in (Y(1+\delta)^{m-1}, Y(1+\delta)^m] \cap (Y, x]\}, \quad (1 \leqslant n \leqslant N).$$

If $M = \log(x/Y)/\log(1+\delta)$, the Brun–Titchmarsh inequality implies that for each $n \leqslant N$:

$$\begin{aligned} \sum_{p\in P_n(x)} \frac{1}{p} &\leqslant \sum_{\substack{0\leqslant m\leqslant M \\ m\equiv n \pmod N}} \frac{\pi(Y(1+\delta)^m) - \pi(Y(1+\delta)^{m-1})}{Y(1+\delta)^{m-1}} \\ &\ll \sum_{m\equiv n \pmod N} \frac{\delta}{\log(Y\delta(1+\delta)^{m-1})} \\ &\ll \delta \sum_{0\leqslant k\leqslant M/N} \frac{1}{\log(x(1+\delta)^{-kN})} \\ &\ll \delta \sum_{0\leqslant k\leqslant M/N} \frac{1}{\log x - kN\log(1+\delta)} \\ &\ll \frac{\delta}{N\log(1+\delta)} \log\biggl(\frac{\log x}{N\log(1+\delta)}\biggr) \ll \frac{1}{N}\log\log x. \end{aligned} \tag{4-15}$$

Now suppose that g satisfies

$$\sum_{p\leqslant x} \frac{g(p)}{p} \sim \alpha \log\log x \tag{4-16}$$

for some $\alpha > 0$ and let $S \subset \{1, \dots, N\}$ denote the set of all indices n such that

$$\sum_{p \in P_n(x)} \frac{g(p)}{p} \geqslant \frac{\alpha}{2} \sum_{p \in P_n(x)} \frac{1}{p}. \tag{4-17}$$

Then, taking our choice of Y into account and since $g(p) \leqslant |f(p)| \leqslant H$, we have

$$\sum_{n \in S} \sum_{p \in P_n(x)} \frac{1}{p} \geqslant \frac{1}{H} \sum_{n \in S} \sum_{p \in P_n(x)} \frac{g(p)}{p} \geqslant \frac{1}{H} \biggl(\sum_{Y \leqslant p \leqslant x} \frac{g(p)}{p} - \frac{\alpha}{2} \sum_{p \leqslant x} \frac{1}{p} \biggr) \sim \frac{\alpha}{2H} \log\log x. \tag{4-18}$$

Comparing this bound with (4-15) shows that S contains a positive proportion of the integers up to N.

Our next aim is to find a subset $T \subset S$ that satisfies

$$\sum_{n \in T} \sum_{p \in P_n(x)} \frac{1}{p} \geqslant \frac{1}{2} \sum_{n \in S} \sum_{p \in P_n(x)} \frac{1}{p}, \tag{4-19}$$

and for which $\bigl|\cos\bigl(\frac{1}{2}|\beta| \log p\bigr)\bigr|$ is bounded away from 1 as p ranges over $\bigcup_{n \in T} P_n$.

By (4-15) and (4-18), we can choose a positive proportion of all $n \leqslant N$ not to belong to T. In particular, we can exclude all n from T for which

$$\bigl(\tfrac{1}{2}(n-1)\log(1+\delta)|\beta|, \tfrac{1}{2} n \log(1+\delta)|\beta|\bigr]$$

intersects $[0, c) \cup (\pi - c, \pi + c) \cup (2\pi - c, 2\pi)$ for some small constant $c > 0$ that only depends on α, H and on the implied constant in (4-15). By doing so, we ensure that

$$\bigl|\cos\bigl(\tfrac{1}{2}|\beta| \log p\bigr)\bigr| < \cos c < 1$$

for all $p \in P_n(x)$ with $n \in T$. Writing $c' := \cos c$ and considering the cosine sum in the final expression of (4-14), the above yields

$$\sum_{p \in P_n(x)} \frac{g(p)\bigl|\cos\bigl(\frac{1}{2}|\beta|\log p\bigr)\bigr|}{p} \leqslant c' \sum_{p \in P_n(x)} \frac{g(p)}{p} \leqslant \sum_{p \in P_n(x)} \frac{g(p)}{p} - (1-c') \sum_{p \in P_n(x)} \frac{g(p)}{p}$$

for $n \in T$. If $n \notin T$, we have the trivial bound

$$\sum_{p \in P_n(x)} \frac{g(p)\bigl|\cos\bigl(\frac{1}{2}|\beta| \log p\bigr)\bigr|}{p} \leqslant \sum_{p \in P_n(x)} \frac{g(p)}{p}.$$

By combining these two bounds with (4-17), (4-18) and (4-19), it follows that

$$\begin{aligned}
\sum_{p \leqslant x} \frac{g(p)\bigl|\cos\bigl(\frac{1}{2}|\beta| \log p\bigr)\bigr|}{p} &\leqslant \sum_{p \leqslant x} \frac{g(p)}{p} - (1-c') \sum_{n \in T} \sum_{p \in P_n(x)} \frac{g(p)}{p} \\
&\leqslant \sum_{p \leqslant x} \frac{g(p)}{p} - \frac{\alpha(1-c')}{2} \sum_{n \in T} \sum_{p \in P_n(x)} \frac{1}{p} \\
&\leqslant \sum_{p \leqslant x} \frac{g(p)}{p} - (C_0 + o(1)) \log\log x
\end{aligned}$$

for some constant $C_0 > 0$ that only depends on α and H. By (4-14), we thus deduce that

$$|F(1+iy)F(1+i(y+\beta))| \ll \exp\biggl(2\sum_{p\leqslant x}\frac{|f(p)|}{p}\biggr)(\log x)^{-2C_0}.$$

It remains to show that there exists a decomposition of $|f(p)|$ into nonnegative functions g and h such that (4-16) holds. This will follow from [Elliott 2017, Lemma 5]. To apply this result, we observe that the two conditions from Definition 1.3 and partial summation show that every $f_0 \in \mathscr{M}_H$ has the property that

$$\liminf_{x\to\infty}\frac{1}{\varepsilon\log x}\sum_{x^{1-\varepsilon}<p\leqslant x}\frac{|f_0(p)|\log p}{p} \geqslant \alpha_{f_0}. \tag{4-20}$$

for every $\varepsilon \in (0,1)$. Thus, the assumptions of [Elliott 2017, Lemma 5] are met and the lemma implies that there exists a nonnegative completely multiplicative function $g_0 \leqslant |f_0|$, which satisfies

$$\lim_{x\to\infty}(\log x)^{-1}\sum_{p\leqslant x}\frac{g_0(p)\log p}{p} = \frac{\alpha_{f_0}}{2}.$$

The function g_0 arises from $|f_0|$ as the result of a simple greedy-type argument that decides one by one for each prime p if $g_0(p)=0$ or $g_0(p)=|f_0(p)|$. Partial summation yields

$$\sum_{p\leqslant z}\frac{g_0(p)}{p} \sim \frac{\alpha_{f_0}}{2}\log\log z.$$

If we let $g(p)=g_0(p)$ for all $p\geqslant Y$ and $g(p)=0$ otherwise, then

$$\sum_{p\leqslant x}\frac{g(p)}{p} = \sum_{p\leqslant x}\frac{g_0(p)}{p} + O(H\log\log Y) \sim \frac{\alpha_{f_0}}{2}\log\log x + O(H\log\log\log x),$$

as required. Thus, we may set $\alpha = \frac{1}{2}\alpha_{f_0}$ in the first part of the proof and, hence, C_0 only depends on α_{f_0} and H. □

Proof of Lemma 4.9. Let f^* denote the multiplicative function that satisfies $f^*(p^k)=0$ whenever $k\geqslant 1$ and $p\leqslant\exp((\log\log x)^2)$, and $f^*(p^k)=f_0(p^k)$ whenever $k\geqslant 1$ and $p>\exp((\log\log x)^2)$. Then, by applying Lemma 4.6 twice, we have

$$\biggl|\frac{1}{y}\sum_{n\leqslant y}f^*(n)n^{-it} - \frac{1}{y'}\sum_{n\leqslant y'}f^*(n)n^{-it}\biggr| \ll \frac{1}{(\log x)^{1+C_0}}\exp\biggl(\sum_{p\leqslant x}\frac{|f(p)|}{p}\biggr),$$

for all $y, y' \in [x\exp(-(\log\log x)^{-4}), x]$ and some $t = t_{x,f^*}$ with $|t|\leqslant 2\log x$.

Observe that any f may be decomposed as $f = f' * f^*$, where $f'(p^k) = 0$ for all $p > \exp((\log\log x)^2)$. Thus, if $w := \exp\bigl(\frac{1}{2}(\log\log x)^{-4}\bigr)$ and $x' \in [x/w, x]$, then

$$\begin{aligned}\left|\frac{1}{x}\sum_{n\leqslant x} f(n)n^{-it} - \frac{1}{x'}\sum_{n\leqslant x'} f(n)n^{-it}\right| &\ll \sum_{d\leqslant w}\frac{|f'(d)|}{d}\left|\frac{d}{x}\sum_{n\leqslant x/d} f^*(n)n^{-it} - \frac{d}{x'}\sum_{n\leqslant x'/d} f^*(n)n^{-it}\right| \\ &\quad + \frac{1}{x}\sum_{d>w}\sum_{m<x/d}|f'(d)f^*(m)| + \frac{1}{x'}\sum_{d>w}\sum_{m<x'/d}|f'(d)f^*(m)| \\ &\ll \sum_{d\leqslant w}\frac{|f'(d)|}{d}\frac{1}{(\log x)^{1+C_0}}\exp\biggl(\sum_{p\leqslant x}\frac{|f^*(p)|}{p}\biggr) \\ &\quad + \sum_{m\leqslant x}\frac{1}{m}\frac{m}{x}\sum_{w<d\leqslant x/m}|f'(d)| + \sum_{m\leqslant x'}\frac{1}{m}\frac{m}{x'}\sum_{m\leqslant x'}\sum_{w<d\leqslant x'/m}|f'(d)|.\end{aligned}$$

To bound the last two terms, recall that $|f'(n)| \leqslant 1$ and that, see, e.g., [Tenenbaum 1995, Theorem III.5.1],

$$\Psi(z,y) := \#\{n \leqslant z : p \mid n \Rightarrow p \leqslant y\} \ll ze^{-u/2} \quad (z \geqslant y \geqslant 2), \tag{4-21}$$

where $u = \log z/\log y$. In particular

$$\Psi\bigl(z, \exp((\log\log x)^2)\bigr) \ll ze^{-(\log\log x)^2/4} = z(\log x)^{-(\log\log x)/4}$$

whenever $z \geqslant \exp\bigl(\frac{1}{2}(\log\log x)^4\bigr)$. Thus, the above is bounded by

$$\begin{aligned}&\ll \frac{1}{(\log x)^{1+C_0}}\exp\biggl(\sum_{p\leqslant x}\frac{|f(p)|}{p} + \sum_{p\leqslant d}\frac{1}{p^2(1-p^{-1})}\biggr) + \sum_{m\leqslant x}\frac{1}{m}(\log x)^{-(\log\log x)/2} \\ &\ll \frac{1}{(\log x)^{1+C_0}}\exp\biggl(\sum_{p\leqslant x}\frac{|f(p)|}{p}\biggr),\end{aligned}$$

which completes the proof. □

Proof of Lemma 4.8. By Lemma 4.3 and (4-4) it follows that

$$\begin{aligned}|S_{f\chi}(y)| \leqslant x^{-\delta/8}(\log x)^{O(H)} &+ \sum_{d_0\leqslant y^\delta}\ \sum_{D\leqslant\left(\frac{y}{d_0}\right)^{1-\frac{1}{H}}}\ \sum_{d_1\cdots d_{H-1}=D}\frac{|h'(d_0)h(d_1)\cdots h(d_{H-1})||\chi(d_0D)|}{d_0D} \\ &\times\sum_{i=1}^{H}\frac{Dd_0}{y}\Biggl(\Biggl|\sum_{n\leqslant\frac{y}{(Dd_0)}} h(n)\chi(n)\Biggr| + \Biggl|\sum_{n<\left(\frac{y}{d_0}\right)^{1-\frac{1}{H}}\frac{\max(d_1,\ldots,d_{i-1})}{D}} h(n)\chi(n)\Biggr|\Biggr).\end{aligned}$$

As in the proof of Corollary 4.2, we set $\delta = 1/(4H)$. Then the inner sums may be bounded using either the assumption or, if $(y/d_0)^{1-1/H}\max(d_1,\ldots,d_{i-1})/D < x^{1/(4H)}$, by the trivial estimate

$$\frac{Dd_0}{y}x^{1/(4H)} \leqslant y^{1-1/H+1/(4H)-1}x^{1/(4H)} = y^{-3/(4H)}x^{1/(4H)} \leqslant x^{-1/(8H)}.$$

The lemma follows by bounding the sums over D and d_0 as in the proof of Corollary 4.2. □

Proof of Lemma 4.7. We first use Lemma 4.3 to remove the contribution of the function h' defined in Lemma 1.8. Given $\delta \in (0,1)$, let δ' be such that $x^{\delta} = {x'}^{\delta'}$. Then

$$|S_{f\chi}(x) - S_{f\chi}(x')| \leqslant x^{-\delta/4}(\log x)^{O(H)} + \sum_{d_0 \leqslant x^{\delta}} \frac{|h'(d_0)\chi(d_0)|}{d_0} \biggl| \frac{d_0}{x} \sum_{n \leqslant x/d_0} h^{*H}(n)\chi(n) - \frac{d_0}{x'} \sum_{n \leqslant x'/d_0} h^{*H}(n)\chi(n) \biggr|. \tag{4-22}$$

To analyze the difference above, we seek to decompose h^{*H} using $H-1$ applications of the hyperbola trick[3],

$$\sum_{nm \leqslant Y} = \sum_{n \leqslant X} \sum_{m \leqslant Y/n} + \sum_{m \leqslant Y/X} \sum_{X \leqslant n \leqslant Y/m}.$$

Fix d_0 and let $X = (x'/d_0)^{1/H}$. If $y \in \{x'/d_0, x/d_0\}$, then applying the hyperbola trick with the chosen cutoff X and with $Y = y$, $n = d_1$ and $m = d_2 \cdots d_H$, we obtain

$$\sum_{d_1 \cdots d_H \leqslant y} = \sum_{d_1 \leqslant X} \sum_{d_2 \cdots d_H \leqslant y/d_1} + \sum_{d_2 \cdots d_H \leqslant y/X} \sum_{X \leqslant d_1 \leqslant y/(d_2 \cdots d_H)}.$$

We keep the second term and decompose the first term again, using the same cutoff X, and $Y = y/d_1$, $n = d_2$ and $m = d_3 \cdots d_H$. This leads to

$$\sum_{d_1 \cdots d_H \leqslant y} = \sum_{d_1 \leqslant X} \sum_{d_2 \leqslant X} \sum_{d_3 \cdots d_H \leqslant y/(d_1 d_2)} + \sum_{d_1 \leqslant X} \sum_{d_3 \cdots d_H \leqslant y/(d_1 X)} \sum_{X \leqslant d_2 \leqslant y/(d_1 d_3 \cdots d_H)} + \sum_{d_2 \cdots d_H \leqslant y/X} \sum_{X \leqslant d_1 \leqslant y/(d_2 \cdots d_H)}.$$

Continuing in this manner, i.e., keeping every time the second new term and further decomposing the first, we arrive at

$$\sum_{d_1 \cdots d_H \leqslant y} = \sum_{d_1, d_2, \ldots, d_{H-1} \leqslant X} \sum_{d_H \leqslant y/(d_1 d_2 \cdots d_{H-1})} + \sum_{i=1}^{H-1} \sum_{d_1, \ldots, d_{i-1} \leqslant X} \sum_{\substack{d_{i+1}, \ldots, d_H: \\ d_1 \cdots \hat{d}_i \cdots d_{i-1} \leqslant y/X}} \sum_{X \leqslant d_i \leqslant y/(d_1 \cdots \hat{d}_i \cdots d_H)}. \tag{4-23}$$

In order to apply this decomposition to (4-22), let us consider the difference of the normalized sums (4-23) for $y = x/d_0$ and $y = x'/d_0$. Recall that $x \geqslant x'$. By splitting the third sum of the second term of the decomposition into two sums when $y = x/d_0$, we obtain the following:

[3] This proof requires a different decomposition from the one used in (4-4) and Section 9D in order to be able to collect together terms in (4-24) below.

$$\frac{d_0}{x}\sum_{d_1\cdots d_H\leqslant x/d_0} - \frac{d_0}{x'}\sum_{d_1\cdots d_H\leqslant x'/d_0}$$
$$= \sum_{d_1,d_2,\dots,d_{H-1}\leqslant X}\Bigg(\frac{d_0}{x}\sum_{d_H\leqslant x/(d_0d_1\cdots d_{H-1})} - \frac{d_0}{x'}\sum_{d_H\leqslant x'/(d_0d_1\cdots d_H-1)}\Bigg)$$
$$+\sum_{i=1}^{H-1}\sum_{d_1,\dots,d_{i-1}\leqslant X}\sum_{\substack{d_{i+1},\dots,d_H:\\ d_0\cdots\hat{d}_i\cdots d_{i-1}\leqslant x'/X}}\Bigg(\frac{d_0}{x}\sum_{d_i\leqslant x/(d_1\cdots\hat{d}_i\cdots d_H)} - \frac{d_0}{x'}\sum_{d_i\leqslant x'/(d_1\cdots\hat{d}_i\cdots d_H)} + \frac{d_0}{x'}\sum_{d_i\leqslant X} - \frac{d_0}{x}\sum_{d_i\leqslant X}\Bigg)$$
$$+\sum_{i=1}^{H-1}\sum_{d_1,\dots,d_{i-1}\leqslant X}\sum_{\substack{d_{i+1},\dots,d_H:\\ x'/X\leqslant d_0\cdots\hat{d}_i\cdots d_H\leqslant x/X}}\frac{d_0}{x}\sum_{X\leqslant d_i\leqslant x/(d_0\cdots\hat{d}_i\cdots d_H)}. \tag{4-24}$$

When we apply this decomposition with the summation argument $g(d_1)\cdots g(d_H)$, where $g(n)=h(n)\chi(n)$, then the first and the second term above contain expressions of the form $S_g(z)-S_g(z')$ for suitable z and z'. These will be estimated using the assumptions of the lemma. Before turning towards these, let us consider the remaining terms.

The second term contains two short sums up to X that will be estimated using Shiu's bound (3-1) in the following form. For every fixed $j\in\mathbb{N}$, every $q\in\mathbb{N}$ for which the interval $((W(x)q)^2, x]$ is nonempty and every $y\in((W(x)q)^2, x]$, we have

$$\sum_{n\leqslant y}|g^{*j}(n)| = \sum_{A\in(\mathbb{Z}/qW(x)\mathbb{Z})^*}\sum_{\substack{n\leqslant y\\ n\equiv A \pmod{qW(x)}}}|h^{*j}(n)| \ll \frac{x}{\log x}\exp\Bigg(j\sum_{\substack{p\leqslant y\\ p\nmid qW(x)}}\frac{|h(p)|}{p}\Bigg), \tag{4-25}$$

since $W(x)q\leqslant y^{1/2}$, and thus $W(x)q = W(y)q'$ for some $q'\leqslant y^{1/2}$.

By applying this bound twice with $W(x)q=Q$, we obtain

$$\Bigg|\frac{d_0}{x'}\sum_{d_1,\dots,d_{i-1}\leqslant X} g(d_1)\cdots g(d_{i-1})\sum_{\substack{d_{i+1}\cdots d_H\leqslant\\ x'/(d_0\cdots d_{i-1}X)}} g(d_{i+1})\cdots g(d_H)\sum_{d_i\leqslant X} g(d_i)\Bigg|$$
$$\leqslant \frac{d_0X}{x'}\frac{1}{\log X}\exp\Bigg(\sum_{\substack{p\leqslant X\\ p\nmid Q}}\frac{|h(p)|}{p}\Bigg)\sum_{d_1,\dots,d_{i-1}\leqslant X}|g(d_1)\cdots g(d_{i-1})|\sum_{d\leqslant x'/(d_0\cdots d_{i-1}X)}|g^{*(H-i)}(d)|$$
$$\leqslant \frac{1}{(\log X)^2}\exp\Bigg((H-i+1)\sum_{\substack{p\leqslant x'\\ p\nmid Q}}\frac{|h(p)|}{p}\Bigg)\sum_{d_1,\dots,d_{i-1}\leqslant X}\frac{|g(d_1)\cdots g(d_{i-1})|}{d_1\cdots d_{i-1}}$$
$$\ll_{H,C}\frac{1}{(\log x)^2}\exp\Bigg(\sum_{\substack{p\leqslant x'\\ p\nmid Q}}\frac{|f(p)|}{p}\Bigg),$$

which saves $(\log x)^{-1}$.

In the final term of (4-24), we will take advantage of the fact that the third sum is short. Starting off with another application of (4-25), we get

$$\sum_{d_1,\dots,d_{i-1}\leqslant X} g(d_1)\cdots g(d_{i-1}) \sum_{\substack{d_{i+1},\dots,d_H:\\ x'/X\leqslant d_0\cdots\hat{d}_i\cdots d_H\leqslant x/X}} g(d_{i+1})\cdots g(d_H)\frac{d_0}{x} \sum_{X\leqslant d_i\leqslant x/(d_0\cdots\hat{d}_i\cdots d_H)} g(d_i)$$

$$\leqslant \frac{1}{\log x}\exp\biggl(\sum_{\substack{p\leqslant x'\\ p\nmid Q}}\frac{|h(p)|}{p}\biggr)\sum_{d_1,\dots,d_{i-1}\leqslant X}\frac{|g(d_1)\cdots g(d_{i-1})|}{d_1\cdots d_{i-1}}\sum_{\substack{d_{i+1},\dots,d_H:\\ x'/X\leqslant d_0\cdots\hat{d}_i\cdots d_H\leqslant x/X}}\frac{|g(d_{i+1})\cdots g(d_H)|}{d_{i+1}\cdots d_H}$$

$$\leqslant \frac{1}{\log x}\exp\biggl(\sum_{\substack{p\leqslant x'\\ p\nmid Q}}\frac{|h(p)|}{p}\biggr)\sum_{\substack{d_1,\dots,d_{i-1}\leqslant X\\ d_{i+1},\dots d_{H-1}\leqslant x}}\frac{|g(d_1)\cdots\widehat{g(d_i)}\cdots g(d_{H-1})|}{d_1\cdots\hat{d}_i\cdots d_{H-1}}\sum_{\substack{d_H:\\ x'/X\leqslant d_0\cdots\hat{d}_i\cdots d_H\leqslant x/X}}\frac{1}{d_H}$$

$$\leqslant \frac{1}{\log x}\exp\biggl(\sum_{\substack{p\leqslant x'\\ p\nmid Q}}\frac{|h(p)|}{p}\biggr)\sum_{\substack{d_1,\dots,d_{i-1}\leqslant X\\ d_{i+1},\dots,d_{H-1}\leqslant x}}\frac{|g(d_1)\cdots\widehat{g(d_i)}\cdots g(d_{H-1})|}{d_1\cdots\hat{d}_i\cdots d_{H-1}}\Bigl(\log\frac{x}{x'}+O(1)\Bigr)$$

$$\leqslant \frac{\log\log X+\log C+O(1)}{\log x}\exp\biggl((H-1)\sum_{\substack{p\leqslant x'\\ p\nmid Q}}\frac{|h(p)|}{p}\biggr),$$

which saves a factor $(\log x)^{-\alpha_f/H+\varepsilon}$.

To summarize our progress so far, note that the decomposition (4-24) and the previous two bounds yield

$$\biggl|\frac{d_0}{x}\sum_{n\leqslant x/d_0}h^{*H}(n)\chi(n)-\frac{d_0}{x'}\sum_{n\leqslant x'/d_0}h^{*H}(n)\chi(n)\biggr|$$

$$= \sum_{\substack{d_1,\dots,d_{H-1}\leqslant\\ (x'/d_0)^{1/H}}}\frac{|g(d_1)\cdots g(d_{H-1})|}{d_1\cdots d_{H-1}}\biggl|S_g\Bigl(\frac{x}{d_0\cdots d_{H-1}}\Bigr)-S_g\Bigl(\frac{x'}{d_0\cdots d_{H-1}}\Bigr)\biggr|$$

$$+\sum_{i=1}^{H-1}\sum_{\substack{d_1,\dots,d_{i-1}\leqslant\\ (x'/d_0)^{1/H}}}\frac{|g(d_1)\cdots g(d_{i-1})|}{d_1\cdots d_{i-1}}\sum_{\substack{d_{i+1},\dots,d_H:\\ d_1\cdots\hat{d}_i\cdots d_H\leqslant\\ (x'/d_0)^{1-1/H}}}\frac{|g(d_{i+1})\cdots g(d_H)|}{d_{i+1}\cdots d_H}$$

$$\times\biggl|S_g\Bigl(\frac{x}{d_0\cdots\hat{d}_i\cdots d_H}\Bigr)-S_g\Bigl(\frac{x'}{d_0\cdots\hat{d}_i\cdots d_H}\Bigr)\biggr|$$

$$+O_{H,C}\biggl((\log x)^{-\min(1,\alpha_f/(2H))}\frac{1}{\log x}\exp\biggl(\sum_{p\leqslant x',\,p\nmid Q}\frac{|f(p)|}{p}\biggr)\biggr).$$

Choosing $\delta=\frac{1}{2}$ to ensure that $d_0\leqslant x^{1/2}$, it follows that the terms $x/(d_0\cdots d_{H-1})$ and $x/(d_0\cdots\hat{d}_i\cdots d_H)$ in the above expression are at least as large as $x^{1/(2H)}$. Since $g=h\chi$, we may thus apply the assumptions of the lemma to deduce that the above is bounded by:

$$\ll \frac{\psi(x)}{\log((x'/d_0)^{1/H})} \exp\biggl(\sum_{p\leqslant x',\, p\nmid Q} \frac{|h(p)|}{p}\biggr)\biggl(\prod_{d\leqslant x} \frac{g(d)}{d}\biggr)^{H-1} + (\log x)^{-\min(1,\alpha_f/(2H))} \frac{1}{\log x} \exp\biggl(\sum_{p\leqslant x',\, p\nmid Q} \frac{|f(p)|}{p}\biggr)$$

$$\ll (\psi(x) + (\log x)^{-\min(1,\alpha_f/(2H))}) \frac{1}{\log x} \exp\biggl(\sum_{p\leqslant x',\, p\nmid Q} \frac{|f(p)|}{p}\biggr)$$

The lemma then follows from (4-22) since by (4-1), applied with $y = 1$ and $w = w(x)$, the completed outer sum over d_0 converges, i.e., $\sum_{d_0=1}^{\infty} |h'(d_0)\chi(d_0)|/d_0 < \infty$, provided x is sufficiently large for $w(x) \geqslant (2H)^{16}$ to hold. □

4D. ***Applications to functions bounded away from zero at primes.*** In this subsection, we will discuss a concrete example of an element of $\mathscr{F}_H$ and prove a criterion for real-valued f to belong to $\mathscr{F}_H$ that is just based on the values of f at primes. Let us begin by stating a special case of Proposition 4.10 for nonnegative $f \in \mathscr{M}_H$.

Lemma 4.13 (sufficient condition for nonnegative functions). *Let $f \in \mathscr{M}_H$ be a nonnegative function. Then there exists a constant $c > 0$, only depending on f, such that the following holds: If $x > 3$, if $1 < Q \leqslant \exp((\log\log x)^2)$ is a multiple of $W(x)$, and if $\chi_0 \pmod{Q}$ denotes the trivial character, then*

$$S_{f\chi_0}(x) = S_{f\chi_0}(x') + O\biggl((\log x)^{-c} \frac{1}{\log x} \prod_{p\leqslant x,\, p\nmid Q} \biggl(1 + \frac{|f(p)|}{p}\biggr)\biggr)$$

uniformly for all $x > 3$, $x' \in [x\exp(-(\log\log x)^2, x]$ and all Q as above. If, furthermore, for either $g = h$ or $g = f$ and for any $C > 0$, we have a uniform bound of the form

$$S_{g\chi}(x) = O\biggl(\frac{\psi_C(x)}{\log x} \prod_{p\leqslant x,\, p\nmid Q} \biggl(1 + \frac{|g(p)|}{p}\biggr)\biggr), \tag{4-26}$$

valid for all $x > 3$, all nontrivial $\chi \pmod{Q}$ and all $1 \leqslant Q \leqslant (\log x)^C$ with $W(x) \mid Q$, and where $\psi_C = o(1)$ may depend on C but is otherwise independent of χ and Q, then $f \in \mathscr{F}_H$.

Remark 4.14. Note that in the context of this corollary, the main term in the character sum expansion of $S_f(x; Q, A)$ always comes from the trivial character.

Proof. The first part follows from Lemmas 4.6 and 4.7 provided we can show that for all sufficiently large x we have $t_{x,h\chi_0} = 0$ in the statement of Lemma 4.6 when applied with f replaced by $h\chi_0$. This, however, is immediate since h is nonnegative. The second part is a consequence of Proposition 4.10. □

The following three lemmas all arise as (nontrivial) applications of Lemma 4.13.

Lemma 4.15 (coefficients of cusp forms). *Let f be a primitive holomorphic cusp form*[4] *of weight $k \in 2\mathbb{N}$ and level $N \in \mathbb{N}$ and let*

$$f(z) = \sum_{n=1}^{\infty} \lambda_f(n) n^{(k-1)/2} e(nz),$$

be its Fourier expansion, where the $\lambda_f(n)$ are the normalized Fourier coefficients. Then the function $n \mapsto |\lambda_f(n)|$ belongs to $\mathscr{F}_2$.

Lemma 4.16 (nonnegative f). *For every $H \geqslant 1$ and $\alpha > 0$, there exists $c = c(H, \alpha) > 0$ such that the following holds. If $f \in \mathscr{M}_H$ is nonnegative with $\alpha_f \geqslant \alpha$, and if there exists $\delta > 0$ such that*

$$\#\{p \leqslant x : f(p) > \delta\} \geqslant \frac{(1-c)x}{\log x}$$

for all sufficiently large x, then $f \in \mathscr{F}_H$.

Remark. As a special case, Lemma 4.16 yields the following simple criterion, which also proves one part of Proposition 1.5:

A nonnegative function $f \in \mathscr{M}_H$ belongs to $\mathscr{F}_H$ if it is bounded away from zero on the primes, i.e., if there exists $\delta > 0$ such that $f(p) > \delta$ for all p. The same holds true if the latter condition is replaced by $\#\{p \leqslant x : f(p) > \delta\} \geqslant (1 + o(1))x/\log x$ as $x \to \infty$.

The following variant of Lemma 4.16 will follow with minor changes in the proof.

Lemma 4.17 (real-valued f). *For every $H \geqslant 1$ and $\alpha > 0$, there exists $c = c(H, \alpha) > 0$ such that the following holds. If $f \in \mathscr{M}_H$ is a real-valued function with $\alpha_f \geqslant \alpha$, and if there exists $\delta > 0$ and a sign $\epsilon \in \{+, -\}$ such that*

$$\#\{p \leqslant x : \epsilon f(p) > \delta\} \geqslant \frac{(1-c)x}{\log x}$$

for all sufficiently large x, then $f \in \mathscr{F}_{H,n^{it}}$. If, furthermore, for every $C > 0$ there exists a function $\psi_C = o(1)$ such that

$$S_{f\chi_0}(x) = O\biggl(\frac{\psi_C(x)}{\log x} \prod_{p \leqslant x,\, p \nmid Q} \biggl(1 + \frac{|f(p)|}{p}\biggr)\biggr),$$

whenever χ_0 is the trivial character modulo Q for any $Q \in (1, (\log x)^C)$ with $W(x) \mid Q$, then $f \in \mathscr{F}_H$.

Remark. As a particular consequence, we deduce that $f \in \mathscr{F}_{H,n^{it}}$ for any function $f \in \mathscr{M}_H$ for which there exists $\delta > 0$ such that $f(p) < -\delta < 0$ at all primes p.

Example 4.18. Examples of functions the above results apply to include:

(i) The Möbius function $f(n) = \mu(n)$. Here, the full statement of Lemma 4.17 applies. We may deduce this from the estimate $S_\mu(x) \ll_B (\log x)^{-B}$ for $B > 0$ and $x \geqslant 2$. In fact, writing $d \mid Q^\infty$ to indicate that

[4]See [Iwaniec and Kowalski 2004, §14.1 and §14.7] for definitions.

$p \mid d$ implies $p \mid Q$, it follows via repeated Möbius inversion that

$$\sum_{\substack{n \leqslant x \\ (n,Q)=1}} \mu(n) = \sum_{d \mid Q^\infty} \sum_{n \leqslant x/d} \mu(n)$$

Recalling (4-21), the above is seen to be bounded by

$$\begin{aligned}
&\ll \sum_{\substack{d \mid Q^\infty \\ d \leqslant x^{1/2}}} \sum_{n \leqslant x/d} \mu(n) + \sum_{x^{1/2} < 2^k \leqslant x} \Psi(2^k, (\log x)^C) x 2^{-k} \\
&\ll_B \sum_{\substack{d \mid Q^\infty \\ d \leqslant x^{1/2}}} \frac{x}{d} (\log x)^{-B} + \sum_{x^{1/2} < 2^k \leqslant x} x \exp\Big(-\frac{\log x}{4C \log\log x}\Big) \\
&\ll_B x (\log x)^{-B} \prod_{p \mid Q} (1 - p^{-1})^{-1} \ll_B x (\log x)^{-B+1},
\end{aligned}$$

which yields the required decay estimate.

(ii) The function $f(n) = \delta^{\omega(n)}$ for any nonzero real number δ and where $\omega(n)$ counts the number of distinct prime factors of n. If $\delta > 0$, then Lemma 4.16 applies. For $\delta < 0$ we will now show that the full statement of Lemma 4.17 applies. Since $W(x) \mid Q$, we may simplify our task by removing finitely many primes from consideration to start with: let $A \geqslant |\delta|$ be a constant to be chosen later, let $Q_0 = \prod_{p > A} p^{v_p(Q)}$ and let $h(n) = \delta^{\omega(n)} \mathbf{1}_{\gcd(n, \prod_{p \leqslant A} p)=1}$ denote the restriciton of f to integers free from primes factors $p \leqslant A$. For this function, the Selberg–Delange method as stated in [Montgomery and Vaughan 2006, Theorem 7.18] implies $S_h(x) \ll (\log x)^{\delta-1}$ and $S_{|h|}(x) \asymp (\log x)^{|\delta|-1}$ for all $x \geqslant 2$, while Lemma 1.7 and Shiu's lemma in its original form [Shiu 1980, Theorem 1] yield $S_{|h|}(x) \asymp (1/\log x) \prod_{p \leqslant x} (1 + |\delta|/p)$. Proceeding in a similar way as in (i), repeated Möbius inversion shows that

$$\begin{aligned}
\sum_{\substack{n \leqslant x \\ (n,Q)=1}} \delta^{\omega(n)} &= \sum_{k \geqslant 1} \sum_{\substack{d_1 \mid Q_0^\infty \\ d_1 \geqslant 1}} \sum_{\substack{d_2 \mid d_1^\infty \\ d_2 > 1}} \cdots \sum_{\substack{d_k \mid d_{k-1}^\infty \\ d_k > 1}} |\delta|^{\omega(d_1)+\cdots+\omega(d_k)} \sum_{\substack{n \leqslant x/d_k \\ p \mid n \Rightarrow p > A}} \delta^{\omega(n)} \\
&\leqslant \sum_{d \mid Q_0^\infty} |\delta|^{\varpi(d)} \wp_m(d) \Big| \sum_{n \leqslant x/d} h(n) \Big|,
\end{aligned} \tag{4-27}$$

where $\varpi(d) = \omega(d)$ if $|\delta| < 1$ and $\varpi(d) = \Omega(d)$ if $|\delta| \geqslant 1$, and where $\wp_m$ counts factorizations of the following form:

$$\wp_m(d) = \#\big\{d = d_1 \cdots d_k,\ k \geqslant 1 : d_j > 1 \text{ and } (p \mid d_j \Rightarrow p \mid d_i \text{ for all } i < j) \text{ for all } 1 \leqslant j \leqslant k\big\}.$$

If $\wp(n)$ denotes the partition number of n as defined in [Hardy and Ramanujan 1918], then

$$\wp_m(d) \leqslant \prod_{p \mid d} \wp(v_p(d)).$$

To bound (4-27), we will use the fact that there exists a constant $B \geqslant 1$ such that $\wp(n) \leqslant B^{\sqrt{n}}$, as proved in [loc. cit., §2]. Further, we require a bound corresponding to the one recalled in (4-21) but for sums over $|\delta|^{\varpi(n)}\wp_m(n)$ restricted to smooth numbers that are coprime to all $p \leqslant A$. For such sums we have

$$\Psi^*(x,y) := \sum_{\substack{n\leqslant x\\ p|n\Rightarrow p\in(A,y]}} |\delta|^{\varpi(n)}\wp_m(n) \leqslant C_\varepsilon x^{1/2+\varepsilon} + \sum_{\substack{n\leqslant x\\ p|n\Rightarrow p\in(A,y]}} (nx^{-1/2})^\alpha \prod_{p|n} |\delta|^{\varpi(p^{v_p(n)})} B^{\sqrt{v_p(n)}}$$

for any $\alpha > 0$. Let $\alpha = (\log y)^{-1}$ and suppose that $y \geqslant 2$. By [Tenenbaum 1995, Corollary III.3.5.1], the final sum in the expression above is bounded by

$$\begin{aligned} &\ll x^{1-\alpha/2} \prod_{A<p\leqslant y} (1-p^{-1}) \sum_{k\geqslant 0} \frac{p^{\alpha k}|\delta|^{\varpi(p^k)}B^k}{p^k} \\ &\ll x^{1-\alpha/2} \prod_{A<p\leqslant y} (1-p^{-1})(1-e|\delta|Bp^{-1})^{-1} \ll x^{1-\alpha/2}(\log y)^{O(1)}, \end{aligned}$$

provided $A \geqslant eB\max(1,|\delta|)$. Thus, in total, we obtain

$$\Psi^*(x,y) \ll x^{1-\alpha/2}(\log y)^{O(1)} = x\exp\biggl(-\frac{\log x}{2\log y} + O(1)\log\log y\biggr) \ll x\exp\biggl(-\frac{\log x}{4\log y}\biggr)$$

for all $2 \leqslant y \leqslant x$, provided $A \geqslant eB\max(1,|\delta|)$.

Returning to (4-27), we choose $A = eB\max(1,|\delta|)$ in the definition of h and Q_0, and recall that $Q_0 \leqslant Q \leqslant (\log x)^C$. With the above bound on $\Psi^*(x,y)$ in place, the expression (4-27) can now be bounded by

$$\begin{aligned} &\ll \sum_{\substack{d|Q_0^\infty\\ d\leqslant x^{1/2}}} |\delta|^{\varpi(d)}\wp_m(d) \sum_{n\leqslant x/d} \delta^{\omega(n)} + \sum_{x^{1/2}<2^k\leqslant x} \Psi^*(2^k,(\log x)^C)x2^{-k}(\log(x2^{-k}))^{|\delta|-1} \\ &\ll \sum_{\substack{d|Q_0^\infty\\ d\leqslant x^{1/2}}} x\frac{|\delta|^{\varpi(d)}\wp_m(d)}{d}(\log x)^{\delta-1} + \sum_{x^{1/2}<2^k\leqslant x} x\exp\biggl(-\frac{\log x}{4C\log\log x}\biggr)(\log x)^{|\delta|} \\ &\ll x(\log x)^{\delta-1} \prod_{p|Q_0} \sum_{k\geqslant 0} \frac{|\delta|^{\varpi(p^k)}B^{\sqrt{k}}}{p^k} + O_E(x(\log x)^{-E}), \end{aligned}$$

which is further bounded by

$$\begin{aligned} &\ll x(\log x)^{\delta-1} \prod_{\substack{p|Q\\ p>\max(|\delta|,B)}} \sum_{k\geqslant 0} \frac{|\delta|^{\varpi(p^k)}B^k}{p^k} + O_E(x(\log x)^{-E}) \\ &\ll x(\log x)^{\delta-1}(\log Q)^{B|\delta|} \ll \frac{(C\log\log x)^{O(|\delta|)}}{(\log x)^{2|\delta|}} \frac{x}{\log x} \prod_{p\leqslant x,\ p\nmid Q} \biggl(1+\frac{|\delta|}{p}\biggr). \end{aligned}$$

Thus, the required decay estimate holds.

(iii) The general divisor functions $d_k(n) = \mathbf{1}^{(*k)}(n)$ for $k \geqslant 2$, i.e., the k-fold convolution of $\mathbf{1}$ with itself. In this case Lemma 4.16 applies.

The remainder of this subsection contains the proofs of Lemmas 4.15–4.17. We begin with the proof of Lemma 4.15, which is the least technical case. Lemmas 4.16 and 4.17 will follow with small modifications from the same proof.

Proof of Lemma 4.15. The function λ_f that describes the normalized Fourier coefficients of f is a multiplicative function and satisfies Deligne's bound

$$|\lambda_f(n)| \leqslant d(n),$$

where d is the divisor function. This shows that part (1) of Definition 1.3 holds with $H = 2$. Condition (2) of the definition follows from [Rankin 1973, Theorem 2], since

$$\sum_{p \leqslant x} |\lambda_f(p)| \log p \geqslant \frac{1}{2} \sum_{p \leqslant x} \lambda_f(p)^2 \log p \sim \frac{x}{2},$$

which allows us to take $\alpha_{\lambda_f} = \frac{1}{2} - \varepsilon$ for any $\varepsilon > 0$. Hence, $g = |\lambda_f|$ belongs to $\mathscr{M}_2$.

To show that $g \in \mathscr{F}_2$, let h be the bounded multiplicative functions defined, as in Lemma 1.8, by

$$h(p^k) = \begin{cases} \frac{1}{2}|\lambda_f(p)| & \text{if } k = 1, \\ 0 & \text{if } k > 1, \end{cases}$$

and note that, by Lemma 4.13, it suffices to show that

$$|S_{h\chi}(x)| = o\left(\frac{1}{\log x} \prod_{\substack{p \leqslant x \\ p \nmid q W(x)}} \left(1 + \frac{|h(p)|}{p} \right) \right) \tag{4-28}$$

for all nontrivial $\chi \pmod{Q}$ with $Q \leqslant (\log x)^C$ and $W(x) \mid Q$. We begin this task by invoking Halász's theorem. Since g is bounded, the Halász–Granville–Soundararajan bound [Granville and Soundararajan 2003, Corollary 1] implies that

$$|S_{h\chi}(x)| = \frac{1}{x} \left| \sum_{n \leqslant x} \chi(n) h(n) \right| \ll (M+1)e^{-M} + \frac{1}{Y} + \frac{\log \log x}{\log x}, \tag{4-29}$$

where

$$M = M(x, Y) = \min_{|y| \leqslant 2Y} \sum_{p \leqslant x} \frac{1 - \Re(h(p)\chi(p)p^{iy})}{p}.$$

Note that

$$\begin{aligned} M(x, Y) &= \min_{|y| \leqslant 2Y} \sum_{p \leqslant x} \frac{1 - h(p) + h(p) - \Re(h(p)\chi(p)p^{iy})}{p} \\ &= \sum_{p \leqslant x} \frac{1 - h(p)}{p} + \min_{|y| \leqslant 2Y} \sum_{p \leqslant x} \frac{h(p)(1 - \Re(\chi(p)p^{iy}))}{p}; \end{aligned} \tag{4-30}$$

we abbreviate the second term in this expression as

$$M_{h\chi}(x, Y) := \min_{|y| \leqslant 2Y} \sum_{p \leqslant x} \frac{h(p)(1 - \Re(\chi(p)p^{iy}))}{p}.$$

Observe that the product in the bound (4-28) satisfies

$$\prod_{p\leqslant x p\nmid qW(x)}\left(1+\frac{|h(p)|}{p}\right) \gg \exp\left(\sum_{(\log x)^{C+2}<p\leqslant x}\frac{|h(p)|}{p}\right) \gg_\varepsilon (\log x)^{-\varepsilon}\exp\left(\sum_{p\leqslant x}\frac{|h(p)|}{p}\right) \gg_\varepsilon (\log x)^{\alpha_h-\varepsilon} \tag{4-31}$$

with $\alpha_h=\alpha_g/H=\alpha_g/2$. Thus, if we let $Y=(\log x)^{1-\alpha_h/2}$, then the last two terms in (4-29) are negligible compared with the bound (4-28). Combining (4-29), (4-30) and (4-31) it follows that

$$\begin{aligned}|S_{h\chi}(x)| &\ll (1+M)e^{-M_{h\chi}(x,Y)}\exp\left(\sum_{p\leqslant x}\frac{|h(p)|-1}{p}\right)+(\log x)^{-1+\alpha_h/2}\\ &\ll \frac{\log\log x}{\log x}e^{-M_{h\chi}(x,Y)}\exp\left(\sum_{p\leqslant x}\frac{|h(p)|}{p}\right)+(\log x)^{-1+\alpha_h/2}\\ &\ll_\varepsilon \frac{(\log x)^{\varepsilon}e^{-M_{h\chi}(x,Y)}}{\log x}\prod_{\substack{p\leqslant x\\ p\nmid qW(x)}}\left(1+\frac{|h(p)|}{p}\right)+(\log x)^{-1+\alpha_h/2}.\end{aligned} \tag{4-32}$$

This reduces our task to that of finding a sufficiently good lower bound on $M_{h\chi}(x,Y)$. To achieve this, we aim to show that there are positive constants $\delta_0,\delta_1,\delta_2>0$ such that for all nontrivial $\chi \pmod Q$ with $Q\leqslant(\log x)^C$ and $W(x)\mid Q$, for all $0\leqslant t\leqslant 2Y$ and for all $y\in(\exp((\log x)^{1-\alpha_h/4}),x]$, the set

$$\mathscr{P}_{\delta_1,\delta_2}(y)=\{p\leqslant y: h(p)>\delta_1\}\cap\{p\leqslant y: 1-\Re(\chi(p)p^{it})>\delta_2\} \tag{4-33}$$

has positive relative density at least δ_0 in the set of primes up to y, i.e.,

$$\#\mathscr{P}_{\delta_1,\delta_2}(y)\geqslant\frac{\delta_0 y}{\log y}. \tag{4-34}$$

The restriction to nonnegative t is justified here since we consider together with every nontrivial $\chi \pmod{qW(x)}$ also its conjugate character $\bar{\chi}$.

Assuming (4-34) for the moment, we then have

$$\sum_{p\in\mathscr{P}_{\delta_1,\delta_2}(x)}\frac{1}{p}\geqslant\frac{\#\mathscr{P}_{\delta_1,\delta_2}(x)}{x}+\int_2^x\frac{\#\mathscr{P}_{\delta_1,\delta_2}(t)}{t^2}\,dt\geqslant\frac{\delta_0}{\log x}+\delta_0\int_{\exp((\log x)^{1-\alpha_h/4})}^x\frac{dt}{t\log t}\,dt\geqslant\frac{\delta_0\alpha_h}{4}\log\log x,$$

and, hence,

$$e^{M_{h\chi}(x,Y)}\gg\exp\left(\delta_1\delta_2\sum_{p\in\mathscr{P}_{\delta_1,\delta_2}(x)}\frac{1}{p}\right)\gg(\log x)^{\delta_0\delta_1\delta_2\alpha_h/4}.$$

Combined with (4-32), this shows, in particular, that

$$|S_{h\chi}(x)|\ll_\varepsilon(\log x)^{-\delta_0\delta_1\delta_2\alpha_h/4+\varepsilon}\frac{1}{\log x}\prod_{\substack{p\leqslant x\\ p\nmid Q}}\left(1+\frac{|h(p)|}{p}\right)+(\log x)^{-1+\alpha_h/2},$$

and, hence, that (4-28) holds. Thus, it remains to establish (4-34).

The set of primes $\mathscr{P}_{\delta_1,\delta_2}(y)$ is determined by two conditions involving the behavior of h, χ and n^{it} at these primes. To find a lower bound on the cardinality of $\mathscr{P}_{\delta_1,\delta_2}(y)$, our first step is to remove the condition that $h(p) > \delta_1$ from consideration. To do so, recall that the Sato–Tate law [Barnet-Lamb et al. 2011] implies that

$$\#\{p \leqslant y : 0 \leqslant |\lambda_p| \leqslant \alpha\} \sim \frac{\mu(\alpha) y}{\log y}$$

for every $\alpha \in [0, 2]$, where

$$\mu(\alpha) = \frac{2\arcsin\bigl(\frac{1}{2}\alpha\bigr) + \sin\bigl(2\arcsin\bigl(\frac{1}{2}\alpha\bigr)\bigr)}{\pi}.$$

This shows, in particular, that for every $c_1 \in (0, 1)$ there exists a $\delta(c_1) > 0$ such that

$$\#\{p \leqslant y : g(p) > \delta(c_1)\} \geqslant \frac{c_1 y}{\log y} \tag{4-35}$$

for all sufficiently large y. Thus, to prove (4-34) for $\delta_2 = \frac{1}{12}$, say, it suffices to show that for every $0 \leqslant t \leqslant 2Y$ and every $y \in \bigl(\exp((\log x)^{1-\alpha_h/4}), x\bigr]$, the set

$$\mathscr{P}_{\chi,t}(y) := \bigl\{p \leqslant y : \Re(\chi(p)p^{it}) < \tfrac{11}{12}\bigr\} \tag{4-36}$$

has positive relative density in the set of primes up to y. Indeed, if

$$\#\mathscr{P}_{\chi,t}(y) \geqslant \frac{c_2 y}{\log y} \tag{4-37}$$

for some $c_2 > 0$, then, setting $c_1 = 1 - \frac{1}{2}c_2$ in (4-35) and letting $\delta_1 = \delta(c_1)$, we find that $\#\mathscr{P}_{\delta_1,\delta_2}(y)$ is at least $c_2 y/(2\log y)$, i.e., that (4-34) holds with $\delta_0 = \frac{1}{2}c_2 > 0$, as required.[5]

Having simplified our problem to that of establishing (4-37) for a set of primes only defined by the behavior of $\chi(p)$ and p^{it}, our next step is to also remove χ from consideration and to essentially turn the problem into a question about the distribution of $(t\log p/(2\pi))_{p\leqslant y}$ modulo one. Let us begin by decomposing the set of primes into classes on which $\chi(p)$ is constant and consider the primes in each progression $A \pmod Q$ for $\gcd(A, Q) = 1$ separately. Let $\{z\} = z - \lfloor z \rfloor$ denote the fractional part of a real number z, let $T = t/(2\pi)$ and consider for each A as above the set

$$\mathscr{N}_A(y) = \{p < y : \{T\log p\} \in I_{T\log y} \text{ and } p \equiv A \pmod Q\} \tag{4-38}$$

where $I_{T\log y} = \bigl[T\log y - \frac{1}{9}, T\log y\bigr] \pmod 1$ is an interval of fixed length $\frac{1}{9}$, the position of which only depends on the parameters y and t, but not on the residue class A. Our aim is to show that there exists a constant $c_3 > 0$ such that for every reduced residue class $A \pmod Q$ and every $y \in (\exp((\log x)^{1-\alpha_h/4}), x]$, we have

$$\#\mathscr{N}_A(y) \geqslant \frac{c_3 y}{\phi(Q)\log y}. \tag{4-39}$$

[5] In view of the reduction to (4-37), it becomes clear that we will only require (4-35) to hold for one specific value of c_1 in the end. This will later allow us to deduce Lemma 4.16 from this proof and, with some further modifications, also Lemma 4.17. For this reason we will track the information gathered on c_2 throughout the rest of the proof.

Since this bound clearly holds for all invertible residue classes if $t = T = 0$ and if $c_3 = 1 - \varepsilon,\ \varepsilon > 0$, we may restrict attention to the case $t \in (0, 2Y]$ below.

Assuming (4-39) for the moment, let us first show how to deduce the claimed bound (4-37). In view of (4-39), it suffices to show that for a positive proportion of the reduced residues $A \pmod Q$ we have $\mathcal{N}_A(y) \subset \mathcal{P}_{\chi,t}(y)$ for all $y \in \big(\exp((\log x)^{1-\alpha_h/4}), x\big]$.

If χ is a nontrivial real character, then each of the preimages $\chi^{-1}(1)$ and $\chi^{-1}(-1)$ contains $\frac{1}{2}\phi(Q)$ residue classes $A \pmod Q$. If the distance of $T \log y$ to the closest integer satisfies $\|T \log y\| > \frac{1}{6}$, then we have $\Re e(z) < \cos\big(\frac{2\pi}{6}\big) + \frac{1}{9} = \frac{1}{2} + \frac{1}{9} < \frac{3}{4}$ for every $z \in I_{T \log y}$, and, hence,

$$\Re(\chi(p)p^{it}) = \Re(e^{it \log p}) < \tfrac{3}{4} < \tfrac{11}{12}$$

for all $p \in \mathcal{N}_A(y)$ and all $\frac{1}{2}\phi(Q)$ classes $A \in \chi^{-1}(1)$. If, on the other hand, $\|T \log y\| \leqslant \frac{1}{6}$, then we have $\Re\, e(z) \geqslant \cos \frac{2\pi}{6} - \frac{1}{9} = \frac{1}{2} - \frac{1}{9} > 0$ for every $z \in I_{T \log y}$, and, thus,

$$\Re(\chi(p)p^{it}) = -\Re(e^{it \log p}) < 0 < \tfrac{11}{12}$$

for all $p \in \mathcal{N}_A(y)$ and all $\frac{1}{2}\phi(Q)$ classes $A \in \chi^{-1}(-1)$. Thus, (4-37) holds with $c_2 = \frac{1}{2}c_3$ if χ is nontrivial and real.

Turning towards the case where χ is not real, recall that the nonzero values of any Dirichlet character $\chi \pmod Q$ are the k-th roots of unity if χ has order k in the group of characters modulo Q and recall also that $\chi(A)$ assumes each k-th root of unity equally often as A runs over the reduced residue classes modulo Q. Thus, if $\chi \pmod Q$ is not a real character, then each of the four sets

$$\mathcal{R}_> = \{A \pmod Q : \Re(\chi(A)) > 0\}, \quad \mathcal{R}_< = \{A \pmod Q : \Re(\chi(A)) < 0\},$$
$$\mathcal{I}_> = \{A \pmod Q : \Im(\chi(A)) > 0\}, \quad \mathcal{I}_< = \{A \pmod Q : \Im(\chi(A)) < 0\}$$

is nonempty and contains a positive proportion of the reduced residues $A \pmod Q$. To see this, note that $k \geqslant 3$, since χ is not real. By the symmetry of the set of k-th roots of unity, we have $\#\mathcal{I}_< = \#\mathcal{I}_>$ and, if i is a k-th root of unity, then $\#\mathcal{R}_< = \#\mathcal{R}_>$ as well. If i is not a k-th root of unity, then $|\#\mathcal{R}_< - \#\mathcal{R}_>| \leqslant \phi(Q)/k$. Since $\mathcal{I}_< \cup \mathcal{I}_>$ and $\mathcal{R}_< \cup \mathcal{R}_>$ both exclude at most two of the k k-th roots and since the latter set excludes none if i is not a k-th root, we have

$$\#\mathcal{S} \geqslant \frac{k-2}{2}\frac{\phi(Q)}{k} = \left(\frac{1}{2} - \frac{1}{k}\right)\phi(Q) \geqslant \frac{\phi(Q)}{6}$$

for each set $\mathcal{S} \in \{\mathcal{R}_>, \mathcal{R}_<, \mathcal{I}_>, \mathcal{I}_<\}$. This proves the claim.

For each of the above sets $\mathcal{S}$, the product set

$$\big\{\chi(A)e^{2\pi i \tau} : A \in \mathcal{S}, \tau \in \big[T \log y - \tfrac{1}{9}, T \log y\big]\big\}$$

is contained in an arc of length $2\pi\big(\frac{1}{2} + \frac{1}{9}\big)$ on the unit circle and a rotation by $\frac{\pi}{2}$ maps each of these four arcs onto another one of them. This configuration has the property that no arc of length at most $\frac{\pi}{4}$ meets more than three of the product sets. Thus, for each choice of the endpoint $T \log y$ there is one set $\mathcal{S}$ for which the above product set avoids $\{e^{iz} : \|z\| < \pi/8\}$, and for that particular set $\mathscr{S}$, we then have

$$\Re(\chi(p)p^{it}) = \Re(\chi(A)e^{it\log p}) < \cos\frac{\pi}{8} = \frac{1}{2}\sqrt{2+\sqrt{2}} < \frac{11}{12}$$

for all $A \in \mathscr{S}$ and all $p \in \mathscr{N}_A(y)$. Thus, (4-37) holds with $c_2 = \frac{1}{6}c_3$. This completes the proof of the claim that (4-39) implies (4-37).

It finally remains to analyze the set $\mathscr{N}_A(y)$ that was defined in (4-38) and we will do this by borrowing an approach from Wintner's work [1935] on the distribution of $(\log p_n)_{n\leqslant x}$ modulo one.

Let us fix a reduced residue class $A \pmod Q$ and let $(p_n^{(A)})_{n\in\mathbb{N}}$ denote the sequence of primes congruent to $A \pmod Q$, ordered in increasing order. Adapting Wintner's notation to our setting, let $N_A(\tau)$ denote the largest index m for which $\log p_m^{(A)} < \tau$, if such an m exists, and let $N_A(\tau) = 0$ otherwise. By the prime number theorem in arithmetic progressions, we then have

$$N_A(\tau) = \frac{e^\tau}{\phi(Q)\tau}(1 + O(\tau^{-1})), \quad (\tau > 0). \tag{4-40}$$

Observe that $N_A(\tau/T)$ counts the number of $m > 0$ such that $T\log p_m \leqslant \tau$. Thus, if we set $\xi := \{T\log y\}$, so that $T\log y = [T\log y] + \xi$, then, in analogy to [Wintner 1935, equation (3)], we may express the quantity $\#\mathscr{N}_A(y)$ as

$$\begin{aligned}\#\mathscr{N}_A(y) &= \sum_{n=1}^{[T\log y]}\Big(N_A\Big(\frac{n+\xi}{T}\Big) - N_A\Big(\frac{n+\xi-1/9}{T}\Big)\Big)\\ &= \sum_{n=T}^{[T\log y]} N_A\Big(\frac{n+\xi}{T}\Big) - \sum_{n=T}^{[T\log y]} N_A\Big(\frac{n+\xi-1/9}{T}\Big).\end{aligned} \tag{4-41}$$

If $T \in (0, C']$ for any fixed constant $C' \geqslant 1$, then

$$\begin{aligned}\#\mathscr{N}_A(y) &> N_A\Big(\frac{[T\log y]+\xi}{T}\Big) - N_A\Big(\frac{[T\log y]+\xi-1/9}{T}\Big)\\ &= N_A(\log y) - N_A\Big(\log y - \frac{1}{9T}\Big)\\ &= \pi(y; Q, A) - \pi(ye^{-1/(9T)}; Q, A)\\ &> \pi(y; Q, A) - \pi(ye^{-1/(9C')}; Q, A)\\ &\gg_{C'} \pi(y; Q, A),\end{aligned} \tag{4-42}$$

and $c_3 \gg_{C'} 1$ in (4-39). This leaves us to establish (4-39) for $T \in \big(C', \frac{1}{\pi}(\log x)^{1-\alpha_h/2}\big]$.

To bound (4-41) below, note that the prime number theorem (4-40) implies that

$$\sum_{n=T}^{[\tau]} N_A\Big(\frac{n+\xi}{T}\Big) = \frac{T}{\phi(Q)}\sum_{n=T}^{[\tau]}\frac{e^{(n+\xi)/T}}{n+\xi} + O\Big(\frac{T^2}{\phi(Q)}\sum_{n=T}^{[\tau]}\frac{e^{(n+\xi)/T}}{(n+\xi)^2}\Big). \tag{4-43}$$

A corresponding expansion for the second sum in (4-41) is obtained on replacing ξ by $\xi-\frac{1}{9}$. The sum in the main term above may be asymptotically evaluated, using induction:

$$(e^{1/T}-1)\sum_{n=T}^{N-1}\frac{e^{n/T}}{n+\xi}=\frac{e^{N/T}}{N}+O\biggl(\frac{1}{T}+\sum_{n=T}^{N}\frac{e^{n/T}}{(n+1)^2}\biggr),\qquad (N\geqslant T+1). \tag{4-44}$$

Indeed, if $N=T+1$, then

$$\biggl|\frac{e}{T+\xi}-\frac{e^{1+1/T}}{T+1}\biggr|=\biggl|e^{1+1/T}\frac{1-\xi}{(T+1)(T+\xi)}-\frac{e}{T+\xi}\biggr|=O\biggl(\frac{e^{1+1/T}}{(T+1)^2}+\frac{1}{T}\biggr),$$

and if we assume that (4-44) holds for $N=M$, then it follows for $N=M+1$, since

$$\begin{aligned}\frac{e^{M/T}}{M}+\frac{(e^{1/T}-1)e^{M/T}}{M+\xi}&=\frac{e^{M/T}}{M}+\frac{(e^{1/T}-1)e^{M/T}}{M}+\frac{\xi e^{M/T}(e^{1/T}-1)}{M(M+\xi)}\\&=\frac{e^{(M+1)/T}}{M}+O\biggl(\frac{e^{(M+1)/T}}{M^2}\biggr)=\frac{e^{(M+1)/T}}{M+1}+O\biggl(\frac{e^{(M+1)/T}}{M^2}\biggr).\end{aligned}$$

Thus, evaluating main term in (4-43) by means of (4-44), we obtain

$$\sum_{n=T}^{[\tau]}N_A\Bigl(\frac{n+\xi}{T}\Bigr)=\frac{T}{\phi(Q)}\frac{e^{\xi/T}e^{([\tau]+1)/T}}{(e^{1/T}-1)([\tau]+1)}+O\biggl(\frac{T^2}{\phi(Q)}\sum_{n=T}^{[\tau]}\frac{e^{(n+1)/T}}{(n+1)^2}+\frac{1}{\phi(Q)}\biggr). \tag{4-45}$$

Since $\int 1/(\log x)^2\,\mathrm{d}x=\mathrm{li}(x)-x/\log x\ll x/(\log x)^2$, the sum in the error term satisfies

$$\sum_{n=T}^{[\tau]}\frac{e^{(n+1)/T}}{(n+1)^2}\leqslant\int_T^{\tau+1}\frac{e^{t/T}}{t^2}\,\mathrm{d}t=\frac{1}{T}\int_e^{e^{(\tau+1)/T}}\frac{\mathrm{d}u}{(\log u)^2}=O\biggl(\frac{Te^{(\tau+1)/T}}{(\tau+1)^2}+1\biggr).$$

Inserting this information, (4-45) and the analogues expression with ξ replaced by $\xi-\frac{1}{9}$ into (4-41), we obtain

$$\#\mathcal{N}_A(y)=\frac{T}{\phi(Q)}\frac{e^{\frac{[T\log y]+1}{T}}e^{\frac{\xi}{T}}}{[T\log y]+1}\frac{1-e^{-\frac{1}{9T}}}{e^{\frac{1}{T}}-1}+O\biggl(\frac{T}{\phi(Q)}\frac{e^{\frac{T\log y+1}{T}}}{(T\log y+1)^2/T^2}+\frac{T^2}{\phi(Q)}\biggr).$$

Recalling that $1<C'<T\leqslant(\log x)^{1-\alpha_h/2}/\pi$ and that $(1-\alpha_h/4)\log x<\log y\leqslant\log x$, this yields

$$\begin{aligned}\#\mathcal{N}_A(y)&=\frac{T}{\phi(Q)}\frac{ey}{[T\log y]+1}\frac{1-e^{-1/9T}}{e^{1/T}-1}+O\biggl(\frac{T}{\phi(Q)}\frac{y}{(\log y)^2}\biggr)\\&=\frac{T}{\phi(Q)}\frac{ey}{[T\log y]+1}\frac{1-e^{-1/9T}}{e^{1/T}-1}+O_{\alpha_h}\bigl(y(\log y)^{-1-\alpha_h/2}/\phi(Q)\bigr)\\&\gg_{\alpha_h}\frac{1-e^{-1/9T}}{e^{1/T}-1}\frac{1}{\phi(Q)}\frac{y}{\log y}.\end{aligned} \tag{4-46}$$

Thus, it remains to bound below the leading fraction in this bound. To this end, note that

$$e^{-\tau}=1-\frac{\tau}{2}-\frac{\tau-\tau^2}{2}-\sum_{k=1}^{\infty}\frac{\tau^{2k+1}}{(2k+1)!}\Bigl(1-\frac{\tau}{2k+2}\Bigr)\leqslant 1-\frac{\tau}{2}$$

for every $\tau \in [0, 1]$, and that

$$e^{\tau} \leqslant 1+\tau+\frac{\tau^2}{2}\sum_{k=0}^{\infty} 2^{-k} = 1+\tau+\tau^2 \leqslant 1+2\tau$$

for all $\tau \in \left[0, \frac{1}{2}\right]$. Thus, if $T \geqslant 2$, then the leading factor in the lower bound (4-46) satisfies

$$\frac{1-e^{-1/9T}}{e^{1/T}-1} > \frac{1-1+1/18T}{1+2/T-1} = \frac{1}{36},$$

and it follows that $c_3 \gg_{\alpha_h} 1$ in this case. Choosing $C' = 2$ in (4-42), this completes the proof of (4-39) and of the lemma. □

Proof of Lemma 4.16. This lemma follows from the proof above, observing that the information (4-35) gained from the Sato–Tate law is now included as an assumption in the statement of the lemma. More precisely, (4-35) is only required for $c_1 = 1 - \frac{1}{2}c_2$, with $c_2 = \frac{1}{6}c_3 \gg_{\alpha_h} 1$. Thus, $c_1 = 1 - c$ for some $c > 0$ only depending on $\alpha_h = \alpha/H$. □

Proof of Lemma 4.17. To deduce this lemma, we need to apply Proposition 4.10 instead of the special case recorded in Lemma 4.13. We restrict attention to the first part of this lemma, the second being a simplification. Let h be the function associated to f via (1-6). Let $x > 1$ and let t_x be a real number as in Lemma 4.9, applied with $f_0 = h$. By arguing as in the proof of Lemma 4.9, it follows from (4-8) that

$$|S_{h\chi_0}(x)| \ll \frac{1}{|t_x|+1} + \frac{\log\log x}{\log x} + \frac{1}{(\log x)^{1+C_0}} \exp\biggl(\sum_{p \leqslant x} \frac{|h(p)\chi_0(p)|}{p}\biggr)$$

whenever $\chi_0 \pmod{Q}$, $Q \leqslant \exp((\log\log x)^2)$, is a trivial character. If $|t_x| > (\log x)^{1-\alpha_h/2}$, then $|S_{h\chi_0}(x)|$ is small, and we set $\tau_x = 0$. If $|t_x| \leqslant (\log x)^{1-\alpha_h/2}$, we instead set $\tau_x = t_x$.

The rest of the proof proceeds almost exactly as that of Lemma 4.15, but with the following changes. Instead $|S_{h\chi}(x)|$, we now seek to bound $|S_{h(n)\chi(n)n^{-it_x}}(x)|$, or even

$$\max_{|t| \leqslant (\log x)^{1-\alpha_h/2}} |S_{h(n)\chi(n)n^{-it}}(x)|. \tag{4-47}$$

Since the parameter Y is chosen as $Y = (\log x)^{1-\alpha_h/2}$ in the proof of Lemma 4.15, we may readily turn the bound (4-29) into one on (4-47) by redefining M as $M = M(x, Y')$ with $Y' = 2Y$, a change which does not affect the rest of the argument. Continuing from here, we replace the decomposition (4-30) by

$$\begin{aligned} M(x, Y') &= \min_{|y| \leqslant 2Y'} \sum_{p \leqslant x} \frac{1 - |h(p)| + |h(p)| - \Re(h(p)\chi(p)p^{iy})}{p} \\ &= \sum_{p \leqslant x} \frac{1-|h(p)|}{p} + \min_{|y| \leqslant 2Y'} \sum_{p \leqslant x} \frac{|h(p)|(1 - \operatorname{sgn}(h(p))\Re(\chi(p)p^{iy}))}{p}, \end{aligned}$$

and let

$$M_{h\chi}(x, Y') = \min_{|y| \leqslant 2Y'} \sum_{p \leqslant x} \frac{|h(p)|(1 - \operatorname{sgn}(h(p))\Re(\chi(p)p^{iy}))}{p}$$

denote the second term from this new expression. As in the proof of Lemma 4.16, we need to replace (4-35) by our new assumptions, which will also allow us to fix $\operatorname{sgn}(h(p)) = \epsilon$. The set of primes in (4-36) now takes the form

$$\mathscr{P}_{\chi,t}(y) = \{p \leqslant y : \epsilon\Re(\chi(p)p^{iy})) < \tfrac{11}{12}\}.$$

The deduction of (4-37) from (4-39) remains, apart from obvious changes taking into account the additional sign ϵ, unchanged. □

5. The W-trick

Generalising the fact that the bound (1-1) only applies to Fourier coefficients $(1/x)\sum_{n\leqslant x} f(n)e(\alpha n)$ at an irrational phase α, it is the case that an arbitrary multiplicative function f may *correlate* with a given nilsequence, unless this sequence itself is sufficiently equidistributed. Thus, statements of the form

$$\frac{1}{N}\sum_{n\leqslant N} h(n)F(g(n)\Gamma) = o_{G/\Gamma}(1)$$

with $h = f$ or $h = f - S_f(N; 1, 1)$ cannot be expected to hold in general. On the other hand, it turns out to be sufficient to ensure that h is equidistributed in progressions to small moduli in order to resolve this problem. For arithmetic applications such as establishing a result of the form (1-9), this can be achieved with the help of the W-trick from [Green and Tao 2008]. The basic idea is to decompose f into a sum of functions that are equidistributed in progressions to small moduli. This is achieved by decomposing the range $\{1, \dots, N\}$ into subprogressions modulo a product $W(N)$ of small primes, which has the effect of fixing or eliminating the contribution from small primes on each of the subprogressions.

For multiplicative functions some minor modifications are necessary. Our aim is to decompose the interval $\{1, \dots, N\}$ into subprogressions $r \pmod q$ in such a way that

$$S_f(N; q, r) = (1 + o(1))S_f(N; qq', r + qr') \tag{5-1}$$

for small q' and $0 \leqslant r' < q$. Thus, f should essentially have a constant average value when decomposing one of the given subprogressions into further subprogressions of small moduli q'. The example of the characteristic function of sums of two squares shows that we cannot in general choose q to be a product of small primes (consider the case where $r \equiv 1 \pmod 2$, $q' = 2$ and $r + qr' \equiv 3 \pmod 4$), but rather need to allow q to be a product of small prime powers. Further, if f is a function for which Shiu's bound on $S_f(N; q, r)$ is correct in the sense that

$$S_f(N; q, r) \sim \frac{q}{\phi(q)}\frac{1}{\log N}\prod_{\substack{p\leqslant N\\ p\nmid q}}\Big(1 + \frac{f(p)}{p}\Big),$$

then, in order for (5-1) to hold, we must have $p \mid q$ whenever $p \mid q'$ and p is small.

Our aim in this section is to show that for every $f \in \mathscr{F}_H$ we may, instead of $q = W(N)$ as in [Green and Tao 2008], take $q = \widetilde{W}(N) := q^*(N)W(N)$ for some integer-valued function $q^* : \mathbb{N} \to \mathbb{N}$ that satisfies the bound $q^*(x) \leqslant (\log x)^{O(1)}$. For comparison, recall that $W(x) = \prod_{p \leqslant w(x)} p$, with w as in Definition 1.2. Thus,

$$\log W(x) = \sum_{p \leqslant w(x)} \log p \sim w(x) \quad \text{and} \quad W(x) \leqslant (\log x)^{1+o(1)}.$$

For such a function $\widetilde{W}$, we may decompose the range $[1, N]$ into subprogressions of the form

$$\{1 \leqslant m \leqslant N : m \equiv w_1 A \pmod{w_1 \widetilde{W}(N)}\},$$

where $A \in (\mathbb{Z}/\widetilde{W}(N)\mathbb{Z})^*$ and where $w_1 \geqslant 1$ is composed entirely of primes dividing $\widetilde{W}(N)$. Abbreviating $\widetilde{W} = \widetilde{W}(N)$, we have $\gcd(w_1, \widetilde{W}n + A) = 1$ and hence $f(w_1(\widetilde{W}n + A)) = f(w_1) f(\widetilde{W}n + A)$. Thus, it suffices to study the family of functions

$$\{n \mapsto f(\widetilde{W}n + A) : 0 < A < \widetilde{W}(N),\ \gcd(A, \widetilde{W}) = 1\}.$$

Our first concern is to discard the set of large values of w_1 from consideration, as by doing so we can insure that the range on which each function $n \mapsto f(\widetilde{W}n + A)$ needs to be considered is always large. Since large values of w_1 form a sparse set, their contribution in any arithmetic application can usually be bounded by just using the Cauchy–Schwarz inequality and a bound on the second moment of f as in [Browning and Matthiesen 2017, Lemma 7.9]. More precisely, one can show that if, for $C_1 > 1$,

$$\mathscr{S}_{C_1}(N) = \{w_1 \in \mathbb{N} : w_1 > (\log N)^{C_1},\ p \mid w_1 \Rightarrow p \mid \widetilde{W}(N)\},$$

then

$$\frac{1}{N} \sum_{n \leqslant N} \sum_{w_1 \in \mathscr{S}_{C_1}(N)} \mathbf{1}_{w_1 \mid n} |f(n)| \ll (\log N)^{-C_1/3},$$

provided $q^*(N) < (\log N)^{C_1/3}$ and C_1 is sufficiently large with respect to H; see [Matthiesen 2016, §5] for details. By choosing $C_1 > 3\alpha_f$, we can for instance ensure that this bound is $o\big(\frac{1}{N} \sum_{n \leqslant N} |f(n)|\big)$. As shown in [loc. cit., §5], the contribution of $\mathscr{S}_{C_1}(N)$ to correlations of the form (1-9) is negligible.

Thus, for the purpose of arithmetic applications, it suffices to consider $n \mapsto f(\widetilde{W}n + A)$ for $n \in \{1, \dots, T\}$ with

$$T = \frac{N - Aw_1}{w_1 \widetilde{W}(N)} \gg \frac{N}{(\log N)^{C_1} \widetilde{W}(N)}.$$

The next proposition shows that every function $f \in \mathscr{F}_H$ admits a W-trick. More precisely, any finite collection $f_1, \dots, f_r$ of elements from $\mathscr{F}_H$ simultaneously admits a W-trick and we moreover have control over the size of $\widetilde{W}(= q)$ and over the level of q' up to which (a weakened form of) the relation (5-1) holds. Below, $\widetilde{W}$ plays the role of q and q plays the role of q'.

Proposition 5.1 (the elements of $\mathscr{F}_{H,n^{it}}$ admit a W-trick). *Let $E, H \geqslant 1$ be constants and let $f_1, \dots, f_r \in \mathscr{F}_{H,n^{it}}$. Then there exists a constant κ, depending on E, H, r and $\alpha = \min_{1 \leqslant j \leqslant r} \alpha_{f_j}$, and functions $\varphi' : \mathbb{N} \to \mathbb{R}$ and $q^* : \mathbb{N} \to \mathbb{N}$ such that the following holds*:

(1) $\varphi'(x) \to 0$ *as* $x \to \infty$.

(2) *$q^*(x) \leqslant (\log x)^\kappa$ for all sufficiently large $x \in \mathbb{N}$.*

(3) *If $x \in \mathbb{N}$ is sufficiently large, if we set $\widetilde{W}(x) := q^*(x)W(x)$, and if we define*

$$f_x : n \mapsto f(n)n^{-it_x} \quad \textit{for any } f \in \{f_1, \dots, f_r\}$$

and with t_x as in Definition 1.6 with $C = 2E + \kappa + 4$, then the estimate

$$\frac{q\widetilde{W}(x)}{|I|} \sum_{\substack{m \in I \\ m \equiv A \ (q\widetilde{W}(x))}} f_x(m) - S_{f_x}(x; \widetilde{W}(x), A) = O_{E,H,\kappa}\Bigg(\varphi'(x) \frac{1}{\log x} \frac{\widetilde{W}(x)}{\phi(\widetilde{W}(x))} \prod_{\substack{p<x \\ p \nmid \widetilde{W}(x)}} \Bigg(1 + \frac{|f(p)|}{p}\Bigg)\Bigg) \tag{5-2}$$

holds uniformly for all intervals $I \subseteq \{1, \dots, x\}$ with $|I| > x(\log x)^{-E}$, for all integers $0 < q \leqslant (\log x)^E$ and for all $A \in (\mathbb{Z}/q\widetilde{W}(x)\mathbb{Z})^*$.

Remarks 5.2. (1) If $f \in \mathscr{F}_H$, then $f_x = f$.

(2) We will show that (5-2) holds with $\varphi'(x) = \varphi_C(x) + (\log w(x))^{-1} + (\log x)^{-\alpha_f/(3H)} + (\log x)^{-E}$, where φ_C is as in Definition 1.4 with $C = 2E + \kappa + 4$.

The rest of this section is devoted to a proof of Proposition 5.1. Our strategy is to first relate the left-hand side of (5-2) to a restricted character sum, which we will then attempt to bound by means of the "pretentious large sieve"-consequence recorded in Corollary 4.2.

We begin with a technical lemma that will at various points in the argument allow us to control the contribution of the prime divisors $p \mid \widetilde{W}(N)$ that are larger than $w(N)$.

Lemma 5.3. *Let $1 \leqslant a \leqslant (\log N)^E$ be an integer that is free from prime factors $p < w(N)$ and suppose that $0 \leqslant g(p) \leqslant H$ for all p. Then*

$$\prod_{p \mid a}\Bigg(1 + \frac{g(p)}{p}\Bigg) = 1 + O_{E,H}\Bigg(\frac{1}{\log w(N)}\Bigg).$$

Proof. The assumptions on a imply the bound $\Omega(a) \leqslant E \log\log N / \log w(N)$ on the total number of prime factors of a. Let $m = [w(N)/\log w(N) + \Omega(a)]$ and recall that the n-th prime p_n satisfies $p_n \sim n \log n$. Then,

$$p_m \sim m \log m \leqslant \frac{w(N) + E \log\log N}{\log w(N)} \log m \sim w(N) + E \log\log N.$$

Using the bounds on $w(N)$ from Definition 1.2 and Mertens' estimate, we obtain

$$\begin{aligned}\prod_{p \mid a}\Bigg(1 + \frac{g(p)}{p}\Bigg) &\leqslant \prod_{p \mid a}(1 + p^{-1})^H \leqslant \prod_{w(N) < p < p_m}(1 + p^{-1})^H \\ &\leqslant \Bigg(\frac{\log(w(N) + E \log\log N) + O(1)}{\log w(N) + O(1)}\Bigg)^H \\ &= \Bigg(1 + O_E\Bigg(\frac{1}{\log w(N)}\Bigg)\Bigg)^H = 1 + O_{E,H}\Bigg(\frac{1}{\log w(N)}\Bigg),\end{aligned}$$

as claimed. ∎

Corollary 5.4. *If x and q are as in Proposition 5.1, then*

$$\frac{q\widetilde{W}(x)}{\phi(q\widetilde{W}(x))} \leqslant \left(1 + O_{E,H}\left(\frac{1}{\log w(N)}\right)\right)\frac{\widetilde{W}(x)}{\phi(\widetilde{W}(x))}.$$

Proof. Let $a = \prod_{p|q,\, p\nmid \widetilde{W}(x)} p$. Then

$$\begin{aligned}\prod_{p|a}(1-p^{-1})^{-1} &= \prod_{p|a}(1+p^{-1})(1-p^{-2})^{-1}\\ &\leqslant \exp\left(\sum_{p|a}\frac{2}{p^2}\right)\prod_{p|a}(1+p^{-1}) = \left(1+O\left(\frac{1}{w(x)}\right)\right)\prod_{p|a}(1+p^{-1}). \qquad \square\end{aligned}$$

The next lemma replaces the general interval I from (5-2) by one of the form $\{1, \dots, y\}$.

Lemma 5.5. *If E, H, x, f and f_x are as in Proposition 5.1, if $\kappa \geqslant 0$ is a given constant and if $\widetilde{W}(x) \leqslant (\log x)^{\kappa+2}$ is a multiple of $W(x)$, then (5-2) follows if there exists a function $\varphi'' = o(1)$ such that*

$$S_{f_x}(x; q\widetilde{W}(x), A) = S_{f_x}(x; \widetilde{W}(x), A) + O_{E,H,\kappa}\left(\frac{\varphi''(x)}{\log x}\frac{\widetilde{W}q}{\phi(\widetilde{W}q)}\prod_{\substack{p<x\\ p\nmid \widetilde{W}q}}\left(1+\frac{|f(p)|}{p}\right)\right) \tag{5-3}$$

for all $q \in (0, (\log x)^{-E}]$ and $A \in (\mathbb{Z}/q\widetilde{W}(x)\mathbb{Z})^*$. More precisely, we may take

$$\varphi'(x) = \varphi_C(x) + \varphi''(x) + (\log w(x))^{-1}$$

in (5-2), where φ_C is as in Definition 1.4 with $C = 2E + \kappa + 4$.

Proof. In view of (5-3), it suffices to relate the first term in (5-2) to $S_{f_x}(x; q\widetilde{W}(x), A)$. Let $y_1, y_2 \in \mathbb{Z}_{\geqslant 0}$ and suppose that $I = (y_1, y_1 + y_2] \subset [1, x]$ with $y_2 \geqslant x(\log x)^{-E}$. Writing $\widetilde{W} = \widetilde{W}(x)$, an application of (1-5) with $C := 2E + \kappa + 4 > E$ shows that the first term in (5-2) satisfies

$$\begin{aligned}\frac{q\widetilde{W}}{|I|}\sum_{\substack{m\in I\\ m\equiv A\ (q\widetilde{W})}} f_x(m) = \frac{q\widetilde{W}}{y_2}\sum_{\substack{y_1<m\leqslant y_1+y_2\\ m\equiv A\ (q\widetilde{W})}} f_x(m) &= \frac{y_1+y_2}{y_2}S_{f_x}(y_1+y_2; q\widetilde{W}, A) - \frac{y_1}{y_2}S_{f_x}(y_1; q\widetilde{W}, A)\\ &= \frac{y_1+y_2}{y_2}S_{f_x}(x; q\widetilde{W}, A) - \frac{y_1}{y_2}S_{f_x}(y_1; q\widetilde{W}, A)\\ &\quad + O\left(\frac{\varphi_C(x)}{\log x}\frac{\widetilde{W}q}{\phi(\widetilde{W}q)}\prod_{\substack{p<x\\ p\nmid \widetilde{W}q}}\left(1+\frac{|f(p)|}{p}\right)\right).\end{aligned} \tag{5-4}$$

We now split into two cases. If, on the one hand,

$$x > y_1 > y_2(\log x)^{-E-\kappa-4} > x(\log x)^{-2E-\kappa-4},$$

then (1-5) shows that $S_{f_x}(y_1; q\widetilde{W}, A)$ can be replaced by $S_{f_x}(x; q\widetilde{W}, A)$ in the final expression in (5-4), so that (5-4) is seen to equal

$$S_{f_x}(x; q\widetilde{W}, A) + O\biggl(\frac{\varphi_C(x)}{\log x} \frac{\widetilde{W}q}{\phi(\widetilde{W}q)} \prod_{\substack{p<x\\ p\nmid \widetilde{W}q}} \biggl(1 + \frac{|f(p)|}{p}\biggr)\biggr).$$

In this case, (5-2) follows with $\varphi' = \varphi_C + \varphi'' + (\log w(x))^{-1}$ from (5-3) and Corollary 5.4.

If, on the other hand, $y_1 \leqslant y_2(\log x)^{-E-\kappa-4}$, then

$$\frac{y_1 + y_2}{y_2} = (1 + O((\log x)^{-E-\kappa-4})).$$

Since $\phi(q\widetilde{W}) \leqslant q\widetilde{W} \leqslant (\log x)^{E+\kappa+2}$, we further have

$$\begin{aligned} S_{f_x}(y_1; q\widetilde{W}, A) &= \frac{q\widetilde{W}}{y_1} \sum_{\substack{n\leqslant y_1\\ n\equiv A \pmod{q\widetilde{W}}}} f(n) \leqslant q\widetilde{W} \sum_{\substack{n\leqslant y_1\\ n\equiv A \pmod{q\widetilde{W}}}} \frac{f(n)}{n} \\ &\leqslant q\widetilde{W} \prod_{\substack{p\leqslant y\\ p\nmid q\widetilde{W}}} \biggl(1 + \frac{|f(p)|}{p} + \frac{H^2}{p^2(1-H/p)}\biggr) \ll q\widetilde{W} \prod_{\substack{p\leqslant y\\ p\nmid q\widetilde{W}}} \biggl(1 + \frac{|f(p)|}{p}\biggr) \\ &\ll (\log x)^{E+\kappa+2} \frac{q\widetilde{W}}{\phi(q\widetilde{W})} \prod_{\substack{p\leqslant y\\ p\nmid q\widetilde{W}}} \biggl(1 + \frac{|f(p)|}{p}\biggr), \end{aligned}$$

which implies that

$$\frac{y_1}{y_2} S_{f_x}(y_1; q\widetilde{W}, A) \leqslant (\log x)^{-E-\kappa-4} S_{f_x}(y_1; q\widetilde{W}, A) \ll (\log x)^{-2} \frac{\widetilde{W}q}{\phi(\widetilde{W}q)} \prod_{p<x,\, p\nmid \widetilde{W}q} \biggl(1 + \frac{|f(p)|}{p}\biggr).$$

Thus, in this case, (5-4) equals

$$S_{f_x}(x; q\widetilde{W}, A) + O\biggl(\frac{\varphi_C(x) + (\log x)^{-1}}{\log x} \frac{\widetilde{W}q}{\phi(\widetilde{W}q)} \prod_{\substack{p<x\\ p\nmid \widetilde{W}q}} \biggl(1 + \frac{|f(p)|}{p}\biggr)\biggr),$$

and an application of (5-3) yields (5-2) with $\varphi'(x) = \varphi_C(x) + \varphi''(x) + (\log w(x))^{-1}$, when taking into account Corollary 5.4. □

Following the above reduction, we now proceed to analyze the difference of the two mean values that appear in (5-3).

Lemma 5.6 (restricted character sum). *Let $g : \mathbb{N} \to \mathbb{C}$ be an arithmetic function, not necessarily multiplicative, let $\widetilde{W}, q, A \geqslant 1$ be integers, and suppose that* $\gcd(A, q\widetilde{W}) = 1$. *If $y \geqslant 1$, then*

$$S_g(y; \widetilde{W}, A) - S_g(y; q\widetilde{W}, A) = \frac{q\widetilde{W}}{y} \frac{1}{\phi(q\widetilde{W})} \sideset{}{^*}\sum_{\chi \pmod{q\widetilde{W}}} \chi(A) \sum_{n\leqslant y} g(n)\overline{\chi}(n), \tag{5-5}$$

where $\sum^*$ indicates the restriction of the sum to characters that are not induced from characters mod $\widetilde{W}$.

Proof. We have

$$\begin{aligned}
S_g(y;\widetilde{W},A)-S_g(y;q\widetilde{W},A) &= \frac{\widetilde{W}}{y}\biggl(\sum_{\substack{n\leqslant y\\ n\equiv A \pmod{\widetilde{W}}}} g(n) - q\sum_{\substack{n\leqslant y\\ n\equiv A \pmod{q\widetilde{W}}}} g(n)\biggr)\\
&= \frac{1}{y}\frac{\widetilde{W}}{\phi(q\widetilde{W})}\sum_{\chi \pmod{q\widetilde{W}}}\biggl(\sum_{\substack{A' \pmod{q\widetilde{W}}\\ A\equiv A' \pmod{\widetilde{W}}}} \chi(A') - q\chi(A)\biggr)\sum_{n\leqslant y} g(n)\bar{\chi}(n)\\
&= \frac{1}{y}\frac{\widetilde{W}}{\phi(q\widetilde{W})}\sideset{}{^*}\sum_{\chi \pmod{q\widetilde{W}}}\biggl(\sum_{\substack{A' \pmod{q\widetilde{W}}\\ A\equiv A' \pmod{\widetilde{W}}}} \chi(A') - q\chi(A)\biggr)\sum_{n\leqslant y} g(n)\bar{\chi}(n), \qquad (5\text{-}6)
\end{aligned}$$

where $\sum^*$ indicates the restriction of the sum to characters that are not induced from characters mod $\widetilde{W}$; for all other characters we have $\chi(A')=\chi(A)$ and the difference in the brackets above is zero. It remains to show that the sum over A' in (5-6) vanishes. However,

$$\sum_{\substack{A' \pmod{q\widetilde{W}}\\ A\equiv A' \pmod{\widetilde{W}}}} \chi(A') = \frac{1}{\phi(\widetilde{W})}\sum_{\chi' \pmod{\widetilde{W}}} \overline{\chi'}(A)\sum_{A' \pmod{q\widetilde{W}}} \chi(A')\chi'(A') = 0,$$

since $\chi\chi'$ is a nontrivial character modulo $q\widetilde{W}$. Thus the lemma follows. □

Finally, we aim to exploit the fact that the character sum on the right-hand side of (5-5) is restricted by invoking Corollary 4.2.

Proof of Proposition 5.1. Let $\varepsilon := \frac{1}{2}\min(1,\alpha/(2H))$, $k := \lceil \varepsilon^{-2}\rceil$ and $k' = k\lceil \log_2(4H)\rceil$, as in the statement of Corollary 4.2. Setting $C' = (E+1)3^{rk'+1}$, we let $\mathscr{E}$ denote the union of the sets of characters defined by Corollary 4.2 when applied with $C = C'$ to each of the r functions $f_x \in \mathscr{M}_H$ for $f \in \{f_1,\dots,f_r\}$.

Our aim is to find a suitable integer $\widetilde{W}(x)$ so that, if $\widetilde{W} = \widetilde{W}(x)$ and $q \leqslant (\log x)^E$, then none of the characters that appear in the restricted character sum (5-5) is induced by a character from the set $\mathscr{E}$. To do so, we construct a finite sequence of integers $W_0(x), W_1(x), \dots$ with the property

$$W_i(x) \leqslant W(x)^{2^i}(\log x)^{3^i E}$$

as follows. Let $W_0(x) = W(x)$ and suppose we have already defined $W_i(x)$ for all $0 \leqslant i \leqslant j$. Consider the set of integers in the interval $I_j = [W_j(x), W_j(x)(\log x)^E]$. If there exists a character $\chi \in \mathscr{E}$ whose conductor c_χ satisfies $c_\chi \nmid W_j(x)$ but $c_\chi < W_j(x)(\log x)^E$, then we choose one such character χ and define $W_{j+1}(x) := c_\chi W_j(x)$. Note that

$$W_{j+1}(x) < W_j(x)^2(\log x)^E < W(x)^{2\cdot 2^j}(\log x)^{(2\cdot 3^j+1)E} < W(x)^{2^{j+1}}(\log x)^{3^{j+1}E}.$$

If there is no such $\chi \in \mathscr{E}$, then we stop and set $\widetilde{W}(x) = W_j(x)$. Since $\#\mathscr{E} \leqslant rk'$, this process stops after at most rk' steps and, thus, $\widetilde{W}(x) \leqslant W(x)^{2^{rk'}}(\log x)^{3^{rk'}E} \leqslant (\log x)^{2^{rk'+1}+3^{rk'}E}$ and

$$\widetilde{W}(x)q < (\log x^{1/(8H)})^C$$

for all $q \leqslant (\log x)^E$ and sufficiently large x.

Our construction ensures that there exists no character $\chi \pmod{q\widetilde{W}(x)}$ with $q \leqslant (\log x)^E$ that is induced by an element from $\mathscr{E}$ but *not* induced from a character $\pmod{\widetilde{W}(x)}$. Since the sum (5-5) is restricted to those characters modulo $q\widetilde{W}(x)$ that are not induced from characters modulo $\widetilde{W}(x)$, we may apply Corollary 4.2 with $\mathscr{C}$ given by this restricted set of characters and with $Q = q\widetilde{W}(x)$. This application shows that whenever $1 \leqslant q \leqslant (\log x)^E$ and $x^{1/2} < y \leqslant x$, then

$$\frac{1}{y} \sideset{}{^*}\sum_{\chi \ (\mathrm{mod}\ q\widetilde{W})} \chi(A) \sum_{n \leqslant y} f_x(n)\overline{\chi}(n) \ll_{C,H,\alpha} \frac{1}{(\log x)^{1+\alpha/(3H)}} \exp\Biggl(\sum_{\substack{p \leqslant x \\ p \nmid q\widetilde{W}}} \frac{|f(p)|}{p}\Biggr).$$

In combination with Lemma 5.6 for $g = f_x$, this yields (5-3) for $\kappa = C' - 2$ and with $\varphi''(x) = (\log x)^{-\alpha/(3H)}$. Hence, Lemma 5.5 implies the result with $\kappa = C' - 2 \ll_{E,H,r,\alpha} 1$. □

We will refer to (5-2) as *the major arc estimate*. We will show in Section 6B that despite the restriction to invertible residues $A \in (\mathbb{Z}/q\widetilde{W}\mathbb{Z})^*$, the estimate (5-2) implies that $f(\widetilde{W}n + A) - S_f(x; \widetilde{W}, A)$ is orthogonal to periodic sequences of period at most $(\log x)^E$, for every $A \in (\mathbb{Z}/\widetilde{W}\mathbb{Z})^*$. This information will be used in combination with a factorization theorem to reduce the task of proving noncorrelation for $(f(\widetilde{W}n + A) - S_f(x; \widetilde{W}, A))$ with general nilsequences to the case where the nilsequence enjoys certain equidistribution properties and the Lipschitz function satisfies, in particular, $\int_{G/\Gamma} F = 0$.

6. The noncorrelation result

This section contains a precise statement of the main result, which, informally speaking, shows the following. Given $E \geqslant 1$ and a multiplicative function $f \in \mathscr{F}_H$, let $\widetilde{W}(x)$ be the function from Proposition 5.1. Then for every residue $A \in (\mathbb{Z}/\widetilde{W}(N)\mathbb{Z})^*$ and for parameters N and T such that $N^{1-o(1)} \ll T \ll N$, the sequence $\bigl(f(\widetilde{W}n + A) - S_f(N; \widetilde{W}, A)\bigr)_{n \leqslant T}$ is orthogonal to any given polynomial nilsequence, provided E is sufficiently large with respect to H, α_f and data related to the nilsequence. In Section 6B we carry out a standard reduction of the main result to an equidistributed version, modeled on [Green and Tao 2012a, §2].

6A. ***Statement of the main result.*** We begin by recalling the definition of a polynomial nilsequence and related notions from [Green and Tao 2012b]. Let G be a connected, simply connected, nilpotent Lie group. By definition, a filtration $G_\bullet$ on G is a finite sequence of closed connected subgroups

$$G = G_0 = G_1 \geqslant G_2 \geqslant \cdots \geqslant G_d \geqslant G_{d+1} = \{\mathrm{id}_G\}$$

with the property that for all pairs (i, j) with $0 \leqslant i, j \leqslant d$, the commutator group $[G_i, G_j]$ is a subgroup of G_{i+j}, where we set $G_{i+j} = \{\mathrm{id}_G\}$ if $i + j > d + 1$. The degree of $G_\bullet$ is defined to be the largest index j for which G_j is nontrivial. Since G is nilpotent, the lower central series, defined by $G_1 = G$ and $G_{i+1} = [G, G_i]$ for $i \geqslant 1$, terminates after finitely many steps. Setting $G_0 = G$, this series defines a filtration. If s denotes the degree of this filtration, then the Lie group G is called s-step nilpotent. One can show that s is the smallest possible degree that a filtration of G can have.

Let $g : \mathbb{Z} \to G$ be a sequence with values in G and define for every $h \in \mathbb{Z}$, the discrete derivative $\partial_h g(n) = g(n+h)g(n)^{-1}$. Then following [Green and Tao 2012b, Definition 1.8], the set $\mathrm{poly}(\mathbb{Z}, G_\bullet)$ of polynomial sequences with coefficients in $G_\bullet$ is defined to be the set of all sequences $g : \mathbb{Z} \to G$ for which every i-th derivative takes values in G_i, i.e., for which $\partial_{h_i} \cdots \partial_{h_1} g(n) \in G_i$ for all $i \in \{0, \dots, d+1\}$ and for all $n, h_1, \dots, h_i \in \mathbb{Z}$.

To define polynomial *nil*sequences, let $\Gamma < G$ be a discrete cocompact subgroup. Then the compact quotient G/Γ is called a nilmanifold. Any Malcev basis $\mathscr{X}$ (see [loc. cit., §2] for a definition) for G/Γ gives rise to a metric $d_{\mathscr{X}}$ on G/Γ as described in [loc. cit., Definition 2.2]. This metric allows us to define Lipschitz functions on G/Γ as the set of functions $F : G/\Gamma \to \mathbb{C}$ for which the Lipschitz norm (see [loc. cit., Definition 1.2])

$$\|F\|_{\mathrm{Lip}} = \|F\|_\infty + \sup_{x,y \in G/\Gamma} \frac{|F(x) - F(y)|}{d_{\mathscr{X}}(x, y)}$$

is finite. If F is a 1-bounded Lipschitz function, then $(F(g(n)\Gamma))_{n\in\mathbb{Z}}$ is called a (polynomial) nilsequence.

We are now ready to state the main result:

Theorem 6.1. *Let $E, H, d, m_G \geqslant 1$ be integers and let $f \in \mathscr{F}_{H,n^{it}}$. Let N be a positive integer parameter and let $\widetilde{W} = \widetilde{W}(N)$ be the integer produced by Proposition 5.1 for the function f when applied with the given values of E, H and with $x = N$. Let $A \in \mathbb{N}$ be such that $0 < A < \widetilde{W}$ and $\gcd(\widetilde{W}, A) = 1$. Suppose further that T satisfies $N/(\log N)^{E/2} \ll T \ll N$ and that $T, N > e^e$. Let G/Γ be a nilmanifold of dimension m_G together with a filtration $G_\bullet$ of G of degree d and let $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$ a polynomial sequence. Suppose that G/Γ has a M_0-rational Malcev basis adapted to $G_\bullet$ for some $M_0 \geqslant 2$ and let G/Γ be equipped with the metric defined by this basis. Let $F : G/\Gamma \to \mathbb{C}$ be a 1-bounded Lipschitz function. Then, provided $E \geqslant 1$ is sufficiently large with respect to d, m_G, α_f and H, we have*

$$\left| \frac{\widetilde{W}}{T} \sum_{n \leqslant T/\widetilde{W}} \bigl(f(\widetilde{W}n + A) - (\widetilde{W}n + A)^{it_N} S_{f(n)n^{-it_N}}(N; \widetilde{W}, A) \bigr) F(g(n)\Gamma) \right| \ll_{d, m_G, \alpha_f, H}$$
$$\left\{ \varphi'(N) + \frac{1}{\log w(N)} + \frac{M_0^{O_{d,m_G}(1)}}{(\log\log T)^{1/(4^{d+1} \dim G)}} \right\} \frac{1 + \|F\|_{\mathrm{Lip}}}{\log T} \frac{\widetilde{W}}{\phi(\widetilde{W})} \prod_{\substack{p \leqslant N \\ p \nmid \widetilde{W}(N)}} \left(1 + \frac{|f(p)|}{p} \right), \qquad (6\text{-}1)$$

where $t_N \in [-2\log N, 2\log N]$ is, as in Proposition 5.1, given by Definition 1.6 with $C = 2E + \kappa + 4$ (in particular, $t_N = 0$ if $f \in \mathscr{F}_H$), and where φ' is given by (5-2).

Remark. Partial summation, when combined with the estimate (1-5), which holds with the same value of C as above for the function $n \mapsto f(n)n^{-it_N}$, shows that

$$S_{f(n)n^{-it_N}}(N;\widetilde{W},A)=(1+it_N)N^{-it_N}S_f(N;\widetilde{W},A)$$
$$+O\biggl(|t_N|(1+|t_N|)\biggl((\log N)^{-E+O(H)}+\frac{\varphi_C(N)}{\log N}\frac{\widetilde{W}}{\phi(\widetilde{W})}\prod_{p<N,\,p\nmid\widetilde{W}}\Bigl(1+\frac{|f(p)|}{p}\Bigr)\biggr)\biggr),$$

where φ_C is as in (1-5). Thus, if $E\gg_{H,\alpha_f}1$ is sufficiently large and $|t_N|^2\varphi_C(N)=o(1)$, we may replace the term $(\widetilde{W}n+A)^{it_N}S_{f(n)n^{-it_N}}(N;\widetilde{W},A)$ in the statement above by $(1+it_N)((\widetilde{W}n+A)/N)^{it_N}S_f(N;\widetilde{W},A)$.

6B. *Reduction of Theorem 6.1 to the equidistributed case.* Proceeding similarly as in §2 of [Green and Tao 2012a], we will reduce Theorem 6.1 to a special case that involves only equidistributed polynomial sequences. Let us begin by recalling the quantitative notion of equidistribution and total equidistribution for polynomial sequences that was introduced in [Green and Tao 2012b, Definition 1.2].

Definition 6.2. Let G/Γ be a nilmanifold equipped with Haar measure, let $\delta>0$ and let $N\in\mathbb{N}$. A finite sequence $g:\{1,\dots,N\}\to G$ is called δ-equidistributed in G/Γ if

$$\biggl|\frac{1}{N}\sum_{n\leqslant N}F(g(n)\Gamma)-\int_{G/\Gamma}F\biggr|\leqslant\delta\|F\|_{\mathrm{Lip}}$$

for all Lipschitz functions $F:G/\Gamma\to\mathbb{C}$. It is called totally δ-equidistributed if, moreover,

$$\biggl|\frac{1}{\#P}\sum_{n\in P}F(g(n)\Gamma)-\int_{G/\Gamma}F\biggr|\leqslant\delta\|F\|_{\mathrm{Lip}}$$

for all Lipschitz functions $F:G/\Gamma\to\mathbb{C}$ and progressions $P\subset\{1,\dots,N\}$ of length $\#P\geqslant\delta N$.

The tool that makes a reduction to equidistributed polynomial sequences work is the Green and Tao's factorisation theorem [2012b, Theorem 1.19], which we recall for completeness:

Lemma 6.3 (factorization lemma). *Let m and d be positive integers, and let $M_0,N,B>1$ be real numbers. Let G/Γ be an m-dimensional nilmanifold together with a filtration $G_\bullet$ of degree d. Suppose that $\mathscr{X}$ is an M_0-rational Malcev basis adapted to $G_\bullet$ and let $g\in\mathrm{poly}(\mathbb{Z},G_\bullet)$ be a polynomial sequence. Then there is an integer M with $M_0\leqslant M\ll M_0^{O_{B,m,d}(1)}$, a rational subgroup $G'\subseteq G$, a Malcev basis $\mathscr{X}'$ for G'/Γ' in which each element is an M-rational combination of the elements of $\mathscr{X}$, and a decomposition $g=\varepsilon g'\gamma$ into polynomial sequences $\varepsilon,g',\gamma\in\mathrm{poly}(\mathbb{Z},G_\bullet)$ with the following properties*:

(1) *$\varepsilon:\mathbb{Z}\to G$ is (M,N)-smooth.*[6]

(2) *$g':\mathbb{Z}\to G'$ takes values in G' and the finite sequence $(g'(n)\Gamma')_{n\leqslant T}$ is totally M^{-B}-equidistributed in $G'\Gamma/\Gamma$ using the metric $d_{\mathscr{X}'}$ on $G'\Gamma/\Gamma$.*

(3) *$\gamma:\mathbb{Z}\to G$ is an M-rational sequence*[7] *and the sequence $(\gamma(n)\Gamma)_{n\in\mathbb{Z}}$ is periodic with period at most M.*

[6]The notion of smoothness was defined in [Green and Tao 2012b, Definition 1.18]. A sequence $(\varepsilon(n))_{n\in\mathbb{Z}}$ is said to be (M,N)-*smooth* if both $d_{\mathscr{X}}(\varepsilon(n),\mathrm{id}_G)\leqslant N$ and $d_{\mathscr{X}}(\varepsilon(n),\varepsilon(n-1))\leqslant M/N$ hold for all $1\leqslant n\leqslant N$.

[7]A sequence $\gamma:\mathbb{Z}\to G$ is said to be M-*rational* if for each n there is $0<r_n\leqslant M$ such that $(\gamma(n))^{r_n}\in\Gamma$; see [Green and Tao 2012b, Definition 1.17].

The following proposition handles the special case of Theorem 6.1 where the polynomial sequence is equidistributed.

Proposition 6.4 (noncorrelation, equidistributed case). *Let $E, H, m_G, d \geqslant 1$ be integers and suppose that $f \in \mathscr{M}_H$. Let N and T be integer parameters satisfying $N^{1-o(1)} \ll T \ll N$ and let $\delta = \delta(N) \in \big(0, \frac{1}{2}\big)$ depend on N in such a way that*

$$\log N \leqslant \delta(N)^{-1} \leqslant (\log N)^E.$$

Let G/Γ be a nilmanifold of dimension m_G together with a filtration $G_\bullet$ of degree d, and suppose that $\mathscr{X}$ is a $1/\delta(N)$-rational Malcev basis adapted to $G_\bullet$. This basis gives rise to the metric $d_{\mathscr{X}}$. Let $Q = Q(N) \leqslant (\log N)^E$ be an integer that is divisible by $W(N)$ and let $0 \leqslant A < Q$ be an integer such that $A \in (\mathbb{Z}/Q\mathbb{Z})^*$.

Then there is $E_0 \geqslant 1$, depending on d, m_G and H, such that the following holds provided E is sufficiently large with respect to d, m_G and H:

Let $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$ be any polynomial sequence such that the finite sequence

$$(g(n)\Gamma)_{n \leqslant T/Q}$$

is totally $\delta(N)^{E_0}$-equidistributed. Let $F : G/\Gamma \to \mathbb{C}$ be any 1-bounded Lipschitz function such that $\int_{G/\Gamma} F = 0$, and let $I \subset \{1, \ldots, T/Q\}$ be any discrete interval of length at least $T/(Q(\log N)^E)$. Then

$$\left| \frac{Q}{T} \sum_{n \in I} f(Qn + A) F(g(n)\Gamma) \right| \ll_{d, m_G, \alpha_f, H, E} \left\{ (\log\log T)^{-1/(2^{2d+3} \dim G)} + \frac{\delta(N)^{-10^d \dim G}}{(\log\log T)^{1/2^{d+2}}} \right\} \frac{1 + \|F\|_{\mathrm{Lip}}}{\log N} \frac{Q}{\phi(Q)} \prod_{\substack{p \leqslant N \\ p \nmid Q}} \left(1 + \frac{|f(p)|}{p}\right). \quad (6\text{-}2)$$

Proof of Theorem 6.1 assuming Proposition 6.4. We loosely follow the strategy of [Green and Tao 2012a, §2]. In view of the final error term in (6-1), we may assume that $M_0 \leqslant \log N$, as the theorem holds trivially otherwise. This implies that $\mathscr{X}$ is a $(\log N)$-rational Malcev basis. Applying the factorization lemma from above with T replaced by $T/\widetilde{W}$, with $M_0 = \log N$, and with a parameter $B > 1$ that will be determined in course of the proof (as parameter E_0 in an application of Proposition 6.4), we obtain a factorization of g as $\varepsilon g' \gamma$ with properties (1)–(3) from Lemma 6.3. In particular, there is M such that $\log N \leqslant M \leqslant (\log N)^{O_{B, m_G, d}(1)}$ and such that g' takes values in a M-rational subgroup G' of G and is M^{-B}-equidistributed in $G'\Gamma/\Gamma$. Our first aim is to decompose the summation range of n in (6-1) into subprogressions on which the three functions γ, ε and $(\widetilde{W}n + A)^{it_N}$ are all almost constant.

Since γ is periodic with some period $a \leqslant M$, the function $n \mapsto \gamma(an + b)$ is constant for every b, that is, γ is constant on every progression

$$P_{a,b} := \{n \in [1, T/\widetilde{W}] : n \equiv b \pmod a\},$$

where $0 \leqslant b < a$. Let γ_b denote the value that γ takes on $P_{a,b}$ and note that

$$|P_{a,b}| \geqslant T/(2a\widetilde{W}) \geqslant T/(2M\widetilde{W}).$$

Let $g'_{a,b} : \mathbb{Z} \to G'$ be defined via

$$g'_{a,b}(n) = g'(an+b).$$

Since $(g'(n)\Gamma)_{n \leqslant T/\widetilde{W}}$ is totally M^{-B}-equidistributed in $G'\Gamma/\Gamma$, it is clear that every finite subsequence $(g'_{a,b}(n)\Gamma)_{n \leqslant T/(Ca\widetilde{W})}$ is $M^{-B/2}$-equidistributed if a, b and C are such that both $0 \leqslant b < a \leqslant M$ and $C > 0$ and, furthermore, $M^{B/2} > Ca$ hold.

Let $R \geqslant 1$ be an integer that will be chosen later depending on d and $\dim G$. By splitting each progression $P_{a,b}$ into $\ll M(\log\log N)^{1/R}$ pieces $P^{(j)}_{a,b}$ of diameter bounded by $\ll T/(M\widetilde{W}(\log\log N)^{1/R})$, we may also arrange for ε and, simultaneously, for $(\widetilde{W}n + A)^{it_N}$ to be almost constant. More precisely, the fact that ε is $(M, T/\widetilde{W})$-smooth implies that

$$d_{\mathscr{X}}(\varepsilon(n), \varepsilon(n')) \leqslant |n-n'| M\widetilde{W}T^{-1} \ll (\log\log N)^{-1/R}$$

for all $n, n' \leqslant T/\widetilde{W}$ with $|n-n'| \ll T/(M\widetilde{W}(\log\log N)^{1/R})$. By choosing B sufficiently large, we may ensure that $M^{B/2} \geqslant M \log\log N$ and, hence, that the equidistribution properties of $g'_{a,b}$ are preserved on the new bounded diameter pieces of $P_{a,b}$. Let $\mathscr{P}$ denote the collection of all progressions $P^{(j)}_{a,b}$ in our decomposition.

Since F is a Lipschitz function and since $d_{\mathscr{X}}$ is right-invariant (see [Green and Tao 2012b, Appendix A]), we deduce that

$$\begin{aligned}\bigl|F(\varepsilon(n)g'(n)\gamma(n)) - F(\varepsilon(n')g'(n)\gamma(n))\bigr| &\leqslant (1+\|F\|)d(\varepsilon(n), \varepsilon(n')) \\ &\ll (1+\|F\|)(\log\log N)^{-1/R} \end{aligned} \tag{6-3}$$

for all $n, n' \in P_{a,b}$ with $|n-n'| \ll \dfrac{T}{M\widetilde{W}(\log\log N)^{1/R}}$. Thus, this bound holds in particular for any $n, n' \in P^{(j)}_{a,b}$.

To ensure that $(\widetilde{W}n + A)^{it_N}$ is almost constant on the bounded parameter progressions $P^{(j)}_{a,b}$ that we consider, let $\mathscr{P}' \subset \mathscr{P}$ denote the subset of progressions $P^{(j)}_{a,b}$ that are completely contained in the interval $[T/(\widetilde{W}(\log\log N)^{1/(2R)}), T/\widetilde{W}]$. Observe that the contribution of all other progressions $P^{(j)}_{a,b} \in \mathscr{P} \setminus \mathscr{P}'$ to (6-1) may be bounded by

$$(\log\log N)^{-1/(2R)} \frac{\widetilde{W}(N)}{\phi(\widetilde{W}(N))} \frac{2}{\log T} \prod_{\substack{p \leqslant N \\ p \nmid \widetilde{W}(N)}} \Bigl(1 + \frac{|f(p)|}{p}\Bigr),$$

where we used Shiu's bound (3-1) together with fact that we are only summing over $n \leqslant \dfrac{T}{\widetilde{W}(\log\log N)^{1/(2R)}}$. Since

$$\frac{\log\log N}{\log\log\log N} < w(N) \leqslant \log\log N$$

and since $1 \leqslant R \ll_{d,\dim G} 1$, we have

$$(\log\log N)^{-1/(2R)} \ll_{d,\dim G} (\log w(N))^{-1}, \tag{6-4}$$

which implies that the above contribution is negligible when compared to the bound in (6-1). For every remaining progression $P_{a,b}^{(j)} \in \mathscr{P}'$, the diameter is now short compared to the size of the endpoints and we have

$$\begin{aligned}\log(\widetilde{W}n + A) &= \log(\widetilde{W}n' + A) + \log \frac{\widetilde{W}(n'+n-n')+A}{\widetilde{W}n'+A} \\ &= \log(\widetilde{W}n' + A) + \log\biggl(1 + O\biggl(\frac{1}{M(\log\log N)^{1/(2R)}}\biggr)\biggr) \\ &= \log(\widetilde{W}n' + A) + O\biggl(\frac{1}{M(\log\log N)^{1/(2R)}}\biggr)\end{aligned}$$

for all $n, n' \in P_{a,b}^{(j)}$. Since $|t_N| \leqslant 2\log N$ and $M \geqslant \log N$, we deduce that

$$\begin{aligned}(\widetilde{W}n + A)^{it_N} &= (\widetilde{W}n' + A)^{it_N} \exp\biggl(O\biggl(\frac{\log N}{M(\log\log N)^{1/(2R)}}\biggr)\biggr) \\ &= (\widetilde{W}n' + A)^{it_N}(1 + O((\log\log N)^{-1/(2R)})) \\ &= (\widetilde{W}n' + A)^{it_N} + O((\log\log N)^{-1/(2R)})\end{aligned} \tag{6-5}$$

for all $n, n' \in P_{a,b}^{(j)}$.

Let us fix one element $n_{b,j}$ for each progression $P_{a,b}^{(j)} \in \mathscr{P}'$. As we will show next, it will be sufficient to bound the correlation

$$\begin{aligned}&\biggl|\sum_{n \in P_{a,b}^{(j)}} \bigl(f(\widetilde{W}n + A) - (\widetilde{W}n_{b,j} + A)^{it_N} S_{f(n)n^{-it_N}}(N;\, \widetilde{W},\, A)\bigr) F(\varepsilon(n_{b,j})g'(n)\gamma_b\Gamma)\biggr| \\ &= \biggl|\sum_{\substack{n: \\ an+b \in P_{a,b}^{(j)}}} \bigl(f(\widetilde{W}(an+b) + A) - (\widetilde{W}n_{b,j} + A)^{it_N} S_{f(n)n^{-it_N}}(N;\, \widetilde{W},\, A)\bigr) F(\varepsilon(n_{b,j})g'_{a,b}(n)\gamma_b\Gamma)\biggr|\end{aligned} \tag{6-6}$$

for each bounded diameter piece $P_{a,b}^{(j)} \in \mathscr{P}'$. Indeed, the estimates (6-3) and (6-5) applied with $n' = n_{b,j}$ to each such progression, show that the error term incurred from this reduction satisfies

$$\begin{aligned}&\biggl|\sum_{P_{a,b}^{(j)} \in \mathscr{P}'} \sum_{n \in P_{a,b}^{(j)}} \Bigl\{\bigl(f(\widetilde{W}n + A) - (\widetilde{W}n + A)^{it_N} S_{f(n)n^{-it_N}}(N;\, \widetilde{W},\, A)\bigr) F(\varepsilon(n)g'(n)\gamma(n)\Gamma) \\ &\qquad\qquad - \bigl(f(\widetilde{W}n + A) - (\widetilde{W}n_{b,j} + A)^{it_N} S_{f(n)n^{-it_N}}(N;\, \widetilde{W},\, A)\bigr) F(\varepsilon(n_{b,j})g'(n)\gamma_b\Gamma)\Bigr\}\biggr| \\ &\quad \leqslant \sum_{P_{a,b}^{(j)} \in \mathscr{P}'} \sum_{n \in P_{a,b}^{(j)}} |f(\widetilde{W}n + A)| \bigl|F(\varepsilon(n)g'(n)\gamma(n)\Gamma) - F(\varepsilon(n_{b,j})g'(n)\gamma_b\Gamma)\bigr| \\ &\qquad + \sum_{P_{a,b}^{(j)} \in \mathscr{P}'} \sum_{n \in P_{a,b}^{(j)}} |(\widetilde{W}n + A)^{it_N}| \bigl|F(\varepsilon(n)g'(n)\gamma(n)\Gamma) - F(\varepsilon(n_{b,j})g'(n)\gamma_b\Gamma)\bigr| S_{|f|}(N;\, \widetilde{W},\, A) \\ &\qquad + \sum_{P_{a,b}^{(j)} \in \mathscr{P}'} \sum_{n \in P_{a,b}^{(j)}} |(\widetilde{W}n + A)^{it_N} - (\widetilde{W}n_{b,j} + A)^{it_N}| \bigl|F(\varepsilon(n_{b,j})g'(n)\gamma_b\Gamma)\bigr| S_{|f|}(N;\, \widetilde{W},\, A)\end{aligned}$$

$$\ll \frac{T}{\widetilde{W}(N)} \frac{(1+\|F\|)S_{|f|}(N;\widetilde{W},A)}{(\log\log N)^{1/R}}.$$

By Shiu's bound (3-1), this in turn is bounded above by

$$\ll \frac{T}{\widetilde{W}(N)} \frac{(1+\|F\|)}{(\log\log N)^{1/R}} \frac{\widetilde{W}(N)}{\phi(\widetilde{W}(N))} \frac{1}{\log N} \exp\biggl(\sum_{w(N)<p\leqslant N} \frac{|f(p)|}{p}\biggr). \tag{6-7}$$

Taking into account (6-4), the error term (6-7) is acceptable in view of the bound in (6-1).

We aim to estimate the correlation (6-6) with the help of Proposition 6.4. This task will be carried out in four steps, the first of which will be to bound the contribution from noninvertible residues $\widetilde{W}b+A \pmod{\widetilde{W}a}$ to which Proposition 6.4 does not apply. The two subsequent steps consist of checking the various assumptions of Proposition 6.4, while the fourth step contains the actual application of the proposition.

Before we start, we record a final estimate that will be used throughout the rest of the proof. Note that the common difference of $P^{(j)}_{a,b}$ satisfies $a \leqslant M \ll (\log N)^{O_{d,m_G,B}(1)}$, which is bounded above by $(\log N)^E$, provided E is sufficiently large in terms of d, m_G and B.

Step 1: *Noninvertible residues.* We seek to bound the contribution to (6-1) of all progressions $P^{(j)}_{a,b} \in \mathscr{P}'$ with $\gcd(\widetilde{W}b+A, \widetilde{W}a) > 1$. Let $a' = \prod_{p\nmid \widetilde{W}(N)} p^{v_p(a)}$, so that $\widetilde{W}(N)$ is invertible modulo a'. Since $\gcd(A,\widetilde{W}) = 1$, it suffices to check whether b satisfies $\gcd(\widetilde{W}b+A, a') > 1$. Thus, the contribution we seek to bound takes the form

$$\frac{\widetilde{W}}{T} \sum_{d|a',\, d>1} \sum_{\substack{b<a:\\ \gcd(\widetilde{W}b+A,a')=d}} \sum_{\substack{n<T/\widetilde{W}\\ n\equiv b \pmod a}} \{|f(\widetilde{W}n+A)| + S_{|f|}(N;\widetilde{W},A)\}.$$

The contribution from the terms involving $S_{|f|}(N;\widetilde{W},A)$ is bounded by

$$\begin{aligned} &\ll S_{|f|}(N;\widetilde{W},A) \sum_{d|a',\, d>1} \sum_{\substack{b<a:\\ \gcd(\widetilde{W}b+A,a')=d}} \frac{1}{a} \\ &\ll S_{|f|}(N;\widetilde{W},A) \sum_{d|a',\, d>1} \frac{1}{a}\frac{a}{a'}\phi\Bigl(\frac{a'}{d}\Bigr) \ll S_{|f|}(N;\widetilde{W},A) \sum_{d|a',\, d>1} \frac{1}{d}, \end{aligned} \tag{6-8}$$

where we used the fact that $\widetilde{W}(N)$ is invertible modulo a'. In a similar fashion, we may bound the contribution from those terms involving $|f(\widetilde{W}n+A)|$ as follows:

$$\frac{\widetilde{W}}{T} \sum_{d|a',\, d>1} \sum_{\substack{b<a:\\ \gcd(\widetilde{W}b+A,a')=d}} \sum_{\substack{n<T/\widetilde{W}\\ n\equiv b \pmod a}} |f(\widetilde{W}n+A)| \leqslant \sum_{d|a',\, d>1} \sum_{\substack{b<a:\\ \gcd(\widetilde{W}b+A,a')=d}} \frac{|f(d)|}{a} S_{|f|}\Bigl(\frac{T}{d}; \frac{\widetilde{W}a}{d}, \frac{\widetilde{W}b+A}{d}\Bigr)$$

$$\begin{aligned}
&\leqslant \sum_{d|a',\, d>1} \frac{|f(d)|}{a} \frac{a}{a'} \phi\Big(\frac{a'}{d}\Big)\, S_{|f|}\Big(\frac{T}{d};\, \frac{\widetilde{W}a}{d},\, \frac{\widetilde{W}b+A}{d}\Big) \\
&\ll \sum_{d|a',\, d>1} \frac{|f(d)|}{a'} \phi\Big(\frac{a'}{d}\Big) \frac{\widetilde{W}a/d}{\phi(\widetilde{W}a/d)} \frac{1}{\log(T/d)} \exp\Big(\sum_{\substack{p<T/d\\ p\nmid \widetilde{W}a'/d}} \frac{|f(p)|}{p}\Big) \\
&\leqslant \sum_{d|a',\, d>1} \frac{|f(d)|}{a'} \phi\Big(\frac{a'}{d}\Big) \frac{a'/d}{\phi(a'/d)} \frac{\widetilde{W}}{\phi(\widetilde{W})} \frac{1}{\log(T/d)} \exp\Big(\sum_{\substack{p<T\\ p\nmid \widetilde{W}}} \frac{|f(p)|}{p}\Big) \\
&\ll \frac{\widetilde{W}}{\phi(\widetilde{W})} \frac{1}{\log T} \exp\Big(\sum_{\substack{p<T\\ p\nmid \widetilde{W}}} \frac{|f(p)|}{p}\Big) \sum_{d|a',\, d>1} \frac{|f(d)|}{d},
\end{aligned}$$

where we made use of (3-1) and of the fact that $d \leqslant a \leqslant (\log N)^E$ so that $\log(T/d) \geqslant \frac{1}{2}\log T$ once N and, hence, T are sufficiently large. Observe that the final sums in each of the two bounds above are similar. We restrict attention to bounding the latter of them. Assuming that the lower bound $w(N)$ on prime divisors $p \mid a'$ is sufficiently large with respect to H, we have

$$\begin{aligned}
\sum_{d|a',\, d>1} \frac{|f(d)|}{d} &\leqslant \prod_{p|a'} \Big(1 + \frac{H}{p} + \frac{H^2}{p^2} + \cdots\Big) - 1 \leqslant \prod_{p|a'} \Big(1+\frac{H}{p}\Big)\Big(1 + \frac{H^2}{p^2(1-\frac{H}{p})}\Big) - 1 \\
&\leqslant \exp\Big(\sum_{p|a'} \frac{2H^2}{p^2}\Big) \prod_{p|a'}\Big(1+\frac{H}{p}\Big) - 1 \leqslant \Big(1 + \frac{4H^2}{w(N)}\Big) \prod_{p|a'}\Big(1+\frac{H}{p}\Big) - 1 \\
&\ll_{E,H} \frac{4H^2}{w(N)} + \frac{1}{\log w(N)} \ll_{E,H} \frac{1}{\log w(N)},
\end{aligned}$$

where we applied Lemma 5.3 with a replaced by a' to estimate the product over $p \mid a'$.

Bounding the inner sum in (6-8) in a similar fashion and applying (3-1) to estimate $S_{|f|}(N;\, \widetilde{W}, A)$, we deduce that the total contribution of noninvertible residues $\widetilde{W}b + A \pmod{\widetilde{W}a}$ to (6-1) is at most

$$O_{d,m_G,B,H}\bigg(\frac{1}{\log w(N)} \frac{\widetilde{W}}{\phi(\widetilde{W})} \frac{1}{\log T} \exp\Big(\sum_{w(N)<p<T} \frac{|f(p)|}{p}\Big)\bigg),$$

which has been taken care of in (6-1). This leaves us to considering the case where the value of b does not impose an obstruction to applying Proposition 6.4.

Step 2: *Checking the initial conditions of Proposition 6.4.* The central assumption of Proposition 6.4 concerns the equidistribution of the polynomial sequence it is applied to. To verify this assumption for the sequence that appears in (6-6), it is necessary to show that the conjugated sequence $h^* : n \mapsto \gamma_b^{-1} g'_{a,b}(n) \gamma_b$ is, in fact, a polynomial sequence and that it inherits the equidistribution properties of $g'_{a,b}(n)$. Both these questions have been addressed in [Green and Tao 2012a, §2] in a way we can directly build on:

Let $H=\gamma_b^{-1}G'\gamma_b$ and define $H_\bullet=\gamma_b^{-1}(G')_\bullet\gamma_b$. Let $\Lambda=\Gamma\cap H$ and define $F_{b,j}:H/\Lambda\to\mathbb{R}$ via

$$F_{b,j}(x\Lambda)=F(\varepsilon(n_{b,j})\gamma_b x\Gamma).$$

Then $h^*\in\mathrm{poly}(\mathbb{Z},H_\bullet)$ and the correlation (6-6) that we seek to bound takes the form

$$\left|\sum_n \bigl(f(\widetilde{W}(an+b)+A)-(\widetilde{W}n_{b,j}+A)^{it_N}S_{f(n)n^{-it_N}}(N;\widetilde{W},A)\bigr)F_{b,j}(h^*(n)\Lambda)\right|, \tag{6-9}$$

where the sum is over the n such that $(an+b)\in P_{a,b}^{(j)}$. The Claim from the end of [Green and Tao 2012a, §2] guarantees the existence of a Malcev basis $\mathscr{Y}$ for H/Λ adapted to $H_\bullet$ such that each basis element Y_i is a $M^{O(1)}$-rational combination of basis elements X_i. Thus, there is $C'=O(1)$ such that $\mathscr{Y}$ is $M^{C'}$-rational. Furthermore, it implies that there is $c'>0$, depending only on the dimension of G and the degree of $G_\bullet$, such that whenever B is sufficiently large the sequence

$$(h^*(n)\Lambda)_{n\leqslant T/(a\widetilde{W})} \tag{6-10}$$

is totally $M^{-c'B/2+O(1)}$-equidistributed in H/Λ, equipped with the metric $d_{\mathscr{Y}}$ induced by $\mathscr{Y}$. Taking B sufficiently large, we may assume that the sequence (6-10) is totally $M^{-c'B/4}$-equidistributed. Finally, the "Claim" also provides the bound $\|F_{b,j}\|_{\mathrm{Lip}}\leqslant M^{C''}\|F\|_{\mathrm{Lip}}$ for some $C''=O(1)$. This shows that all conditions of Proposition 6.4 are satisfied except for $\int_{H/\Lambda}F_{b,j}=0$.

Step 3: *The final condition.* The final condition that needs to be arranged for before we can apply Proposition 6.4 to (6-9) is that $\int_{H/\Lambda}F_{b,j}=0$. This is where the major arc condition (5-2) is needed, which in turn requires that $\gcd(\widetilde{W}b+A,\widetilde{W}a)=1$. To ensure that the integral over the test function is zero, we decompose $F_{b,j}(x\Lambda)$ as $(F_{b,j}(x\Lambda)-\mu_{b,j})+\mu_{b,j}$, where $\mu_{b,j}:=\int_{H/\Lambda}F_{b,j}$. The expression in parentheses represents a new test function that we can apply the proposition with, and we will show next that the contribution from the constant term $\mu_{b,j}$ is small provided $f(\widetilde{W}n+A)-(\widetilde{W}n_{b,j}+A)^{it_N}S_{f(n)n^{-it_N}}(N;\widetilde{W},A)$ does not correlate with the characteristic function $\mathbf{1}_{P_{a,b}^{(j)}}$ of the corresponding progression $P_{a,b}^{(j)}$.

To start with, recall that $T\geqslant N/(\log N)^{E/2}$, that the common difference of $P_{a,b}^{(j)}$ satisfies $a\leqslant(\log N)^E$, and that the length of $P_{a,b}^{(j)}$ is bounded below by

$$\begin{aligned}|P_{a,b}^{(j)}|\geqslant T/(2aM\widetilde{W}(\log\log N)^{1/R})&\gg T/(a\widetilde{W}(\log N)^{E/2})\\&\gg N/(a\widetilde{W}(\log N)^{E}),\end{aligned}$$

provided E is sufficiently large in terms of d, m_G and B. Observe that condition (5-2) applies to the function $n\mapsto f(n)n^{-it_N}$ and to all discrete intervals $I\subset\{1,\dots,T/\widetilde{W}\}$ of length $|I|\gg T/(\log T)^E$. In particular, we may choose $q=a$, $r=b$ and let I be a discrete interval of length $a\widetilde{W}(N)|P_{a,b}^{(j)}|$ that contains the set $\{\widetilde{W}(N)m+A:m\in P_{a,b}^{(j)}\}$. To relate $f(n)$ to $f(n)n^{-it_N}$, we observe that (6-5) and (6-4) imply that

$$f(\widetilde{W}n+A)=(\widetilde{W}n_{b,j}+A)^{it_N}f(\widetilde{W}n+A)(\widetilde{W}n+A)^{-it_N}+O\biggl(\frac{|f(\widetilde{W}n+A)|}{\log w(N)}\biggr)$$

for all $n,n_{b,j}\in P_{a,b}^{(j)}\in\mathscr{P}'$.

By applying condition (5-2) to the main term below and Shiu's bound (3-1) in combination with Corollary 5.4 to the error term, we obtain the following uniform estimate valid for all $P_{a,b}^{(j)} \in \mathscr{P}'$:

$$\begin{aligned}
&\frac{1}{|P_{a,b}^{(j)}|} \sum_{m \in P_{a,b}^{(j)}} f(\widetilde{W}m + A) \\
&= (\widetilde{W}m_{b,j} + A)^{it_N} \frac{a\widetilde{W}}{|I|} \sum_{\substack{m \in I \\ m \equiv \widetilde{W}b + A \ (a\widetilde{W})}} f(m) m^{-it_N} + O\biggl(\frac{1}{\log w(N)} \frac{a\widetilde{W}}{|I|} \sum_{\substack{m \in I \\ m \equiv \widetilde{W}b + A \ (a\widetilde{W})}} |f(m)|\biggr) \\
&= (\widetilde{W}m_{b,j} + A)^{it_N} S_{f(n)n^{-it_N}}(N; \widetilde{W}, A) + O\biggl(\biggl(\varphi'(N) + \frac{1}{\log w(N)}\biggr) \frac{1}{\log N} \frac{\widetilde{W}}{\phi(\widetilde{W})} \prod_{\substack{p < N \\ p \nmid \widetilde{W}}} \biggl(1 + \frac{|f(p)|}{p}\biggr)\biggr).
\end{aligned}$$

Let, as above, $\mu_{b,j} = \int_{H/\Lambda} F_{b,j}$, and note that $\mu_{b,j} \ll 1$. Thus, the error term incurred by replacing for each $P_{a,b}^{(j)}$ with $\gcd(\widetilde{W}b + A, \widetilde{W}a) = 1$ the factor $F_{b,j}(h(n)\Lambda)$ in (6-9) by $(F_{b,j}(h(n)\Lambda) - \mu_{b,j})$ is bounded as follows:

$$\begin{aligned}
&\Biggl| \frac{\widetilde{W}}{T} \sum_{\substack{P_{a,b}^{(j)} \in \mathscr{P}' \\ \gcd(\widetilde{W}b + A, a) = 1}} \mu_{b,j} \sum_{n \in P_{a,b}^{(j)}} \bigl(f(\widetilde{W}n + A) - (\widetilde{W}m_{b,j} + A)^{it_N} S_{f(n)n^{-it_N}}(N; \widetilde{W}, A)\bigr) \Biggr| \\
&\qquad \ll \frac{\widetilde{W}}{T} \sum_{P_{a,b}^{(j)} \in \mathscr{P}'} |P_{a,b}^{(j)}| \biggl(\varphi'(N) + \frac{1}{\log w(N)}\biggr) \frac{1}{\log N} \frac{\widetilde{W}}{\phi(\widetilde{W})} \prod_{\substack{p < N \\ p \nmid \widetilde{W}}} \biggl(1 + \frac{|f(p)|}{p}\biggr) \\
&\qquad \ll \biggl(\varphi'(N) + \frac{1}{\log w(N)}\biggr) \frac{1}{\log N} \frac{\widetilde{W}}{\phi(\widetilde{W})} \prod_{\substack{p < N \\ p \nmid \widetilde{W}}} \biggl(1 + \frac{|f(p)|}{p}\biggr),
\end{aligned}$$

where φ' is the function defined in Remarks 5.2. This error term has been taken care of in the bound (6-1).

Step 4: *Application of Proposition 6.4* The application of Proposition 6.4 to (6-9) will give rise to the third error term in (6-1). In view of the work carried out in Steps 1–3, we may now assume that $\gcd(\widetilde{W}b + A, \widetilde{W}a) = 1$ and that $\int_{H/\Lambda} F_{b,j} = 0$ holds, and apply Proposition 6.4 with

- $g = h,\ Q = \widetilde{W}a,\ I = \{n : an + b \in P_{a,b}^{(j)}\}$,
- a function $\delta : \mathbb{N} \to \mathbb{R}$ such that $\delta(N) = M^{-C'} (= M_0^{O_{d,m_G,B}(1)})$, which ensures that $\mathscr{Y}$ is $1/\delta(N)$-rational,
- E sufficiently large to ensure that $M^{C'} < (\log N)^E$, which in particular means that E depends on B, and
- $E_0 = \frac{1}{4} c' B = O_{d,m_G}(B)$ for some value of B that is sufficiently large to ensure that (6-10) is totally $M^{-c'B/4}$-equidistributed in H/Λ (see Step 2) and that is also sufficiently large for Proposition 6.4 to apply with the above choice of E_0.

Since there are $\ll aM(\log\log N)^{1/R}$ intervals $P_{a,b}^{(j)}$ in the decomposition $\mathscr{P}' \subset \mathscr{P}$, this yields the bound

$$\sum_{P_{a,b}^{(j)} \in \mathscr{P}'} \biggl| \sum_{\substack{n:\,(an+b)\\ \in P_{a,b}^{(j)}}} \bigl(f(\widetilde{W}(an+b)+A) - (\widetilde{W}n_{b,j}+A)^{it_N} S_{f(n)n^{-it_N}}(N;\widetilde{W},A)\bigr) F_{b,j}(h^*(n)\Lambda) \biggr|$$
$$\ll aM(\log\log N)^{1/R} \frac{1+M^{O(1)}\|F\|}{\log T} \frac{T}{\widetilde{W}a} \frac{\widetilde{W}a}{\phi(\widetilde{W}a)} \prod_{\substack{p\leqslant N\\ p\nmid \widetilde{W}a}} \Bigl(1+\frac{|f(p)|}{p}\Bigr)\mathscr{N}$$
$$\ll M^{O(1)}(\log\log N)^{1/R} \frac{1+\|F\|}{\log T} \frac{T}{\widetilde{W}} \frac{\widetilde{W}a}{\phi(\widetilde{W}a)} \prod_{\substack{p\leqslant N\\ p\nmid \widetilde{W}}} \Bigl(1+\frac{|f(p)|}{p}\Bigr)\mathscr{N}, \qquad (6\text{-}11)$$

where the implied constant depends on d, m_G, α_f, H and B, and where

$$\mathscr{N} = (\log\log T)^{-1/(2^{2d+3}\dim G)} + \frac{M^{10^d \dim G}}{(\log\log T)^{1/2^{d+2}}} \ll \frac{M^{10^d \dim G}}{(\log\log T)^{1/(2^{2d+3}\dim G)}}.$$

Finally, we invoke Corollary 5.4 to remove the dependence on a from (6-11). We complete the deduction of Theorem 6.1 by setting $R = 2^{2d+3}\dim G$ and comparing the bound arising from (6-11) with the third term in (6-1). □

It remains to establish Proposition 6.4.

7. Linear subsequences of equidistributed nilsequences

Our aim in this section is to study the equidistribution properties of families

$$\{(g(Dn+D')\Gamma)_{n\leqslant T/D} : D \in [K,2K)\}$$

of linear subsequences of an equidistributed sequence $(g(n)\Gamma)_{n\leqslant T}$, where D runs through dyadic intervals $[K,2K)$ for $K \leqslant T^{1-1/H}$. This result will only be needed in the case of unbounded multiplicative functions, which allows us to assume that $H > 1$ in this section.

We begin by recalling some essential definitions and notation. Let $P:\mathbb{Z}\to\mathbb{R}/\mathbb{Z}$ be a polynomial of degree at most d and let $\alpha_0,\ldots,\alpha_d \in \mathbb{R}/\mathbb{Z}$ be defined via

$$P(n) = \alpha_0 + \alpha_1\binom{n}{1} + \cdots + \alpha_d\binom{n}{d}.$$

Then the *smoothness norm* of g with respect to T is defined (cf. [Green and Tao 2012b, Definition 2.7]) as

$$\|P\|_{C^\infty[T]} = \sup_{1\leqslant j\leqslant d} T^j \|\alpha_j\|_{\mathbb{R}/\mathbb{Z}}.$$

If $\beta_0,\ldots,\beta_d \in \mathbb{R}/\mathbb{Z}$ are defined via

$$P(n) = \beta_d n^d + \cdots + \beta_1 n + \beta_0,$$

then, compare with [Matthiesen 2012, equation (14.3)], the smoothness norm is bounded above by a similar expression in terms of the β_i, namely

$$\|P\|_{C^\infty[T]} \ll_d \sup_{1\leqslant j\leqslant d} T^j \|j!\beta_j\|_{\mathbb{R}/\mathbb{Z}} \ll_d \sup_{1\leqslant j\leqslant d} T^j \|\beta_j\|_{\mathbb{R}/\mathbb{Z}}. \tag{7-1}$$

On the other hand, Lemma 3.2 of [Green and Tao 2012b] shows that there is a positive integer $q \ll_d 1$ such that

$$\|q\beta_j\|_{\mathbb{R}/\mathbb{Z}} \ll T^{-j}\|P\|_{C^\infty[T]}.$$

Apart from smoothness norms, we also require the notion of a horizontal character as given in [loc. cit., Definition 1.5]. A continuous additive homomorphism $\eta : G \to \mathbb{R}/\mathbb{Z}$ is called a *horizontal character* if it annihilates Γ. In order to formulate quantitative results, one defines a height function $|\eta|$ for these characters. A definition of this height, called the *modulus* of η, may be found in [Green and Tao 2012b, Definition 2.6]. All that we require to know about these heights is that there are at most $M^{O(1)}$ horizontal characters $\eta : G \to \mathbb{R}/\mathbb{Z}$ of modulus $|\eta| \leqslant M$.

The interest in smoothness norms and horizontal characters lies in Green and Tao's "quantitative Leibman Theorem":

Proposition 7.1 [Green and Tao 2012b, Theorem 2.9]. *Let m_G and d be nonnegative integers, let $0 < \delta < \frac{1}{2}$ and let $N \geqslant 1$. Suppose that G/Γ is an m_G-dimensional nilmanifold together with a filtration $G_\bullet$ of degree d and that $\mathscr{X}$ is a $1/\delta$-rational Malcev basis adapted to $G_\bullet$. Suppose that $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$. If $(g(n)\Gamma)_{n\leqslant N}$ is not δ-equidistributed, then there is a nontrivial horizontal character η with $0 < |\eta| \ll \delta^{-O_{d,m_G}(1)}$ such that*

$$\|\eta \circ g\|_{C^\infty[N]} \ll \delta^{-O_{d,m_G}(1)}.$$

The following lemma shows that for polynomial sequences the notions of equidistribution and total equidistribution are equivalent with a polynomial dependence in the equidistribution parameter.

Lemma 7.2. *Let N and A be positive integers and let $\delta : \mathbb{N} \to [0, 1]$ be a function that satisfies $\delta(x)^{-t} \ll_t x$ for all $t > 0$. Suppose that G has a $1/\delta(N)$-rational Malcev basis adapted to the filtration $G_\bullet$. Suppose that $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$ is a polynomial sequence such that $(g(n)\Gamma)_{n\leqslant N}$ is $\delta(N)^A$-equidistributed. Then there is $1 \leqslant B \ll_{d,m_G} 1$ such that $(g(n)\Gamma)_{n\leqslant N}$ is totally $\delta(N)^{A/B}$-equidistributed, provided $A/B > 1$ and provided N is sufficiently large.*

Remark 7.3. The Green–Tao factorization theorem (see property (2) of Lemma 6.3) usually allows one to arrange for $A > B$ to hold.

Proof. We allow all implied constants to depend on d and m_G. Let $B \geqslant 1$ and suppose that $(g(n)\Gamma)_{n\leqslant N}$ fails to be totally $\delta(N)^{A/B}$-equidistributed. Then there is a subprogression $P = \{\ell n + b : 0 \leqslant n \leqslant m - 1\}$ of $\{1, \dots, N\}$ of length $m > \delta(N)^{A/B}N$ such that the sequence $(\tilde{g}(n))_{0\leqslant n<m}$, where $\tilde{g}(n) = g(\ell n + b)$, fails to be $\delta(N)^{A/B}$-equidistributed. Provided $A > B$, Proposition 7.1 implies that there is a nontrivial horizontal character $\eta : G \to \mathbb{R}/\mathbb{Z}$ of modulus $|\eta| < \delta(N)^{-O(A/B)}$ such that

$$\|\eta \circ \tilde{g}\|_{C^\infty[m]} \ll \delta(N)^{-O(A/B)}.$$

The lower bound on m implies that this is equivalent to the assertion

$$\|\eta \circ \tilde{g}\|_{C^\infty[N]} \ll \delta(N)^{-O(A/B)},$$

where we recall that the implied constant may depend on d.

Observing that $\eta \circ g$ is a polynomial of degree at most d, let $\eta \circ g(n) = \beta_d n^d + \cdots + \beta_0$. Then

$$\eta \circ \tilde{g}(n) = \sum_{i=0}^{d} n^i \sum_{j=i}^{d} \beta_j \binom{j}{i} \ell^i b^{j-i},$$

and, hence,

$$\sup_{1 \leqslant i \leqslant d} N^i \left\| \sum_{j=i}^{d} \beta_j \binom{j}{i} \ell^i b^{j-i} \right\| \ll \delta(N)^{-O(A/B)}.$$

This yields the bound

$$\left\| \sum_{j=i}^{d} \beta_j \binom{j}{i} \ell^i b^{j-i} \right\| \ll N^{-i} \delta(N)^{-O(A/B)} \tag{7-2}$$

for $1 \leqslant i \leqslant d$. Note that the lower bound on m implies that $\ell < \delta(N)^{-A/B}$. Using a downwards induction argument, we aim to show that

$$\|\ell^d \beta_j\| \ll N^{-j} \delta(N)^{-O(A/B)} \tag{7-3}$$

for all $1 \leqslant j \leqslant d$. For $j = d$, this is clear from the above. Suppose (7-3) holds for all $j > i$. For each $i < j$ we then, in particular, have that

$$\left\| \ell^d \beta_j \binom{j}{i} b^{j-i} \right\| \ll_d \|\ell^d \beta_j\| b^{j-i} \ll_d N^{-j} \delta(N)^{-O(A/B)} b^{j-i} \ll_d N^{-i} \delta(N)^{-O(A/B)}.$$

Using the fact that $\delta(N)^{-t} \ll_t N$ for all $t > 0$, we deduce that (7-3) holds for $j = i$ from the above bounds and from (7-2). This shows that there is a nontrivial horizontal character, namely $\ell^d \eta$, of modulus at most $\delta(N)^{-O(A/B)}$, such that

$$\|\ell^d \eta \circ g\|_{C^\infty[N]} \ll \sup_{1 \leqslant i \leqslant d} N^i \|\ell^d \beta_i\|_{\mathbb{R}/\mathbb{Z}} \ll \delta(N)^{-O(A/B)},$$

where we made use of (7-1). Choosing B sufficiently large in terms of m and d, [Matthiesen 2012, Proposition 14.2(b)] implies that g is not $\delta(N)^A$-equidistributed, which is a contradiction. □

We are now ready to address the equidistribution properties of linear subsequences.

Proposition 7.4. *Let $H > 1$, let N and T be as before and let $E_1 \geqslant 1$. Let $(A_D)_{D \in \mathbb{N}}$ be a sequence of integers such that $|A_D| \leqslant D$ for every $D \in \mathbb{N}$. Further, let $\delta : \mathbb{N} \to (0, 1)$ be a function that satisfies $\delta(x)^{-t} \ll_t x$ for all $t > 0$. Suppose G/Γ has a $1/\delta(N)$-rational Malcev basis adapted to a filtration $G_\bullet$ of* *degree d. Let $g \in \mathrm{poly}(G_\bullet, \mathbb{Z})$ be a polynomial sequence and suppose that the finite sequence $(g(n)\Gamma)_{n\leqslant T}$ is totally $\delta(T)^{E_1}$-equidistributed in G/Γ. Then there is a constant $c_1 \in (0,1)$, depending only on d and $m_G := \dim G$, such that the following assertion holds for all integers $K \in [(\log T)^{\log\log T}, T^{1-1/H}]$, provided $c_1E_1 \geqslant 1$.*

Write $g_D(n) = g(Dn + A_D)$ and let $\mathscr{B}_K$ denote the set of integers $D \in [K, 2K)$ for which

$$(g_D(n)\Gamma)_{n\leqslant T/D}$$

fails to be totally $\delta(T)^{c_1E_1}$-equidistributed. Then

$$\#\mathscr{B}_K \ll K\delta(T)^{c_1E_1}.$$

Proof. Let $K \in [(\log T)^{\log\log T}, T^{1-1/H}]$ be a fixed integer and let $c_1 > 0$ to be determined in the course of the proof. Suppose that $E_1 > 1/c_1$. Lemma 7.2 implies that for every $D \in \mathscr{B}_K$, the sequence $(g_D(n)\Gamma)_{n\leqslant T/D}$ fails to be $\delta(T)^{c_1E_1B}$-equidistributed on G/Γ for some $B > 0$ only depending on d and m_G. We continue to allow implied constants to depend on d and m_G. By Proposition 7.1, there is a nontrivial horizontal character $\eta_D : G \to \mathbb{R}/\mathbb{Z}$ of modulus $|\eta_D| \ll \delta(T)^{-O(c_1E_1)}$ such that

$$\|\eta_D \circ g_D\|_{C^\infty[T/D]} \ll \delta(T)^{-O(c_1E_1)}. \tag{7-4}$$

For each nontrivial horizontal character $\eta : G \to \mathbb{R}/\mathbb{Z}$ we define the set

$$\mathscr{D}_\eta = \{D \in \mathscr{B}_K : \eta_D = \eta\}.$$

Note that this set is empty unless $|\eta| \ll \delta(T)^{-O(c_1E_1)}$. Suppose that

$$\#\mathscr{B}_K \geqslant K\delta(T)^{c_1E_1}.$$

By the pigeonhole principle, there is some η of modulus $|\eta| \ll \delta(T)^{-O(c_1E_1)}$ such that

$$\#\mathscr{D}_\eta \geqslant K\delta(T)^{O(c_1E_1)}.$$

Suppose

$$\eta \circ g(n) = \beta_d n^d + \dots \beta_1 n + \beta_0$$

and let

$$\eta \circ g_D(n) = \alpha_d^{(D)} n^d + \cdots + \alpha_1^{(D)} n + \alpha_0^{(D)}$$

for any $D \in \mathscr{B}_K$. The quantities $\alpha_j^{(D)}$ and β_j are linked through the relation

$$\alpha_j^{(D)} = D^j \sum_{i=j}^{d} \binom{i}{j} A_D^{i-j} \beta_i \tag{7-5}$$

for each $1 \leqslant j \leqslant d$. Thus, the bound (7-4) on the smoothness norm asserts that

$$\sup_{1\leqslant j\leqslant d} \frac{T^j}{K^j} \|\alpha_j^{(D)}\| \ll \delta(T)^{-O(c_1E_1)}. \tag{7-6}$$

With a downwards induction we deduce from (7-6) and (7-5) that

$$\sup_{1\leqslant j\leqslant d}\frac{T^j}{K^j}\|D^j\beta_j\|\ll\delta(T)^{-O(c_1E_1)}. \tag{7-7}$$

The bound (7-7) provides information on rational approximations of $D^j\beta_j$ for many values of D. Our next aim is to use this information in order to deduce information on rational approximations of the β_j themselves. To achieve this, we employ the Waring trick that appeared in the *Type I* sums analysis in [Green and Tao 2012a, §3], and begin by recalling the two lemmas that this trick rests upon. The first one is a recurrence result:

Lemma 7.5 [Green and Tao 2012b, Lemma 3.2]. *Let $\alpha\in\mathbb{R}$, $0<\delta<\frac{1}{2}$ and $0<\sigma<\frac{1}{2}\delta$, and let $I\subseteq\mathbb{R}/\mathbb{Z}$ be an interval of length σ such that $\alpha n\in I$ for at least δN values of n, $1\leqslant n\leqslant N$. Then there is some $k\in\mathbb{Z}$ with $0<|k|\ll\delta^{-O(1)}$ such that $\|k\alpha\|\ll\sigma\delta^{-O(1)}/N$.*

The second is a consequence of the asymptotic formula in Waring's problem:

Lemma 7.6 [Green and Tao 2012a, Lemma 3.3]. *Let $K\geqslant 1$ be an integer, and suppose that $S\subseteq\{1,\dots,K\}$ is a set of size αK. Suppose that $t\geqslant 2^j+1$. Then $\gg_{j,t}\alpha^{2t}K^j$ integers in the interval $[1,tK^j]$ can be written in the form $k_1^j+\cdots+k_t^j$, $k_1,\dots,k_t\in S$.*

Returning to the proof of Proposition 7.4, let us consider the set

$$\overline{\mathscr{D}}_j=\{m\leqslant s(2K)^j : m=D_1^j+\cdots+D_s^j,\ D_1,\dots,D_s\in\mathscr{D}_\eta\}$$

for some $s\geqslant 2^j+1$. Each element m of this set satisfies

$$\|\beta_j m\|\ll\delta(T)^{-O(c_1)}(K/T)^j,\quad 1\leqslant j\leqslant d, \tag{7-8}$$

in view of (7-7). Thus, Lemma 7.6 implies that there are

$$\#\overline{\mathscr{D}}_j\gg\delta(T)^{O(c_1E_1)}K^j$$

elements in this set. In view of the restrictions on K and the assumptions on the function $\delta(x)$, the conditions of Lemma 7.5 (on σ and δ) are satisfied provided T is sufficiently large. We conclude that there is an integer k_j such that

$$1\leqslant k_j\ll\delta(T)^{-O(c_1E_1)}$$

and such that

$$\|k_j\beta_j\|\ll\delta(T)^{-O(c_1E_1)}T^{-j}.$$

Thus

$$\beta_j=\frac{a_j}{\kappa_j}+\tilde{\beta}_j, \tag{7-9}$$

where $\kappa_j\mid k_j$, $\gcd(a_j,\kappa_j)=1$ and

$$0\leqslant\tilde{\beta}_j\ll\delta(T)^{-O(c_1E_1)}T^{-j}.$$

Hence,

$$\|\kappa_j \beta_j\| \ll \delta(T)^{-O(c_1 E_1)} T^{-j}. \tag{7-10}$$

Let $\kappa = \operatorname{lcm}(\kappa_1, \ldots, \kappa_d)$ and set $\tilde{\eta} = \kappa \eta$. We proceed as in [Green and Tao 2012a, §3]: The above implies that

$$\|\tilde{\eta} \circ g(n)\|_{\mathbb{R}/\mathbb{Z}} \ll \frac{\delta(T)^{-O(c_1 E_1)} n}{T},$$

which is small provided n is not too large. Indeed, if $T' = \delta(T)^{c_1 E_1 C} T$ for some sufficiently large constant $C \geqslant 1$, only depending on d and m_G, and if $n \in \{1, \ldots, T'\}$, then

$$\|\tilde{\eta} \circ g(n)\|_{\mathbb{R}/\mathbb{Z}} \leqslant \tfrac{1}{10}.$$

Let $\chi : \mathbb{R}/\mathbb{Z} \to [-1, 1]$ be a function of bounded Lipschitz norm that equals 1 on $\left[-\frac{1}{10}, \frac{1}{10}\right]$ and satisfies $\int_{\mathbb{R}/\mathbb{Z}} \chi(t)\, \mathrm{d}t = 0$. Then, by setting $F := \chi \circ \tilde{\eta}$, we obtain a Lipschitz function $F : G/\Gamma \to [-1, 1]$ that satisfies $\int_{G/\Gamma} F = 0$ and $\|F\|_{\mathrm{Lip}} \ll \delta(T)^{-O(c_1 E_1)}$. Finally, choosing c_1 sufficiently small, only depending on d and m_G, we may ensure that

$$\|F\|_{\mathrm{Lip}} < \delta(T)^{-E_1}$$

and, moreover, that

$$T' > \delta(T)^{E_1} T.$$

This choice of T', F and c_1 implies that

$$\left| \frac{1}{T'} \sum_{1 \leqslant n \leqslant T'} F(g(n)\Gamma) \right| = 1 > \delta(T)^{E_1} \|F\|_{\mathrm{Lip}},$$

which contradicts the fact that $(g(n)\Gamma)_{n \leqslant T}$ is totally $\delta(T)^{E_1}$-equidistributed. This completes the proof of the proposition. □

8. Equidistribution of product nilsequences

In this section we prove, building on material and techniques from [Green and Tao 2012a, §3], a result on the equidistribution of products of nilsequences which will allow us to perform applications of the Cauchy–Schwarz inequality in Section 9. The specific form of the result is adjusted to the requirements of Section 9.

We begin by introducing the product sequences we shall be interested in. Suppose $g \in \operatorname{poly}(G_\bullet, \mathbb{Z})$ is a polynomial sequence. This is equivalent to the assertion that there exists an integer k, elements $a_1, \ldots, a_k$ of G, and integral polynomials $P_1, \ldots, P_k \in \mathbb{Z}[X]$ such that

$$g(n) = a_1^{P_1(n)} a_2^{P_2(n)} \cdots a_k^{P_k(n)}.$$

Then, for any pair of integers (m, m'), the sequence $n \mapsto (g(mn), g(m'n)^{-1})$ is a polynomial sequence on $G \times G$ that may be represented by

$$(g(mn), g(m'n)^{-1}) = \left(\prod_{i=1}^{k}(a_i, 1)^{P_i(mn)}\right)\left(\prod_{i=1}^{k}(1, a_i)^{P_i(m'n)}\right)^{-1}.$$

The horizontal torus of $G \times G$ arises as the direct product $G/\Gamma[G, G] \times G/\Gamma[G, G]$ of horizontal tori for G. Let $\pi : G \to G/\Gamma[G, G]$ be the natural projection map. Any horizontal character on $G \times G$ restricts to a horizontal character on each of its factors. Thus, it takes the form $\eta \oplus \eta'(g_1, g_2) := \eta(g_1) + \eta'(g_2)$ for horizontal characters η, η' of G. The following proposition will be applied in the proof of Proposition 6.4 to sequences $g = g_D$ for unexceptional D in the sense of Proposition 7.4.

Proposition 8.1. *Let N and T be as before and let $E_2 \geqslant 1$. Let $(\tilde{D}_m)_{m\in\mathbb{N}}$ be a sequence of integers satisfying $|\tilde{D}_m| < m$ for every $m \in \mathbb{N}$. Further, let $\delta : \mathbb{N} \to (0, 1)$ be a function that satisfies $\delta(x)^{-t} \ll_t x$ for all $t > 0$. Suppose G/Γ has a $1/\delta(T)$-rational Malcev basis adapted to a filtration $G_\bullet$ of degree d. Let $P \subset \{1, \dots, T\}$ be a discrete interval. Suppose $F : G/\Gamma \to \mathbb{C}$ is a 1-bounded function of bounded Lipschitz norm $\|F\|_{\mathrm{Lip}}$ and suppose that $\int_{G/\Gamma} F = 0$. Let $g \in \mathrm{poly}(G_\bullet, \mathbb{Z})$ and suppose that the finite sequence $(g(n)\Gamma)_{n\leqslant T}$ is totally $\delta(T)^{E_2}$-equidistributed in G/Γ. Then there is a constant $c_2 \in (0, 1)$, only depending on d and $m_G := \dim G$, such that the following assertion holds for all integers K satisfying*

$$\exp((\log\log T)^2) \le K \le \exp\left(\frac{1}{H}(\log T - (\log T)^{1/U})\right),$$

where $1 < U \ll 1$, provided $c_2 E_2 \geqslant 1$.

Let $\mathscr{E}_K$ denote the set of integer pairs $(m, m') \in (K, 2K]^2$ such that the discrete interval

$$I_{m,m'} = \{n \in \mathbb{N} : nm + \tilde{D}_m \in P,\ nm' + \tilde{D}_{m'} \in P\}$$

has length at least

$$\#I_{m,m'} > \delta(N)^{c_2 E_2} T/K,$$

and such that

$$\left|\sum_{n\in I_{m,m'}} F(g(mn + \tilde{D}_m)\Gamma)\overline{F(g(m'n + \tilde{D}_{m'})\Gamma)}\right| > (1 + \|F\|_{\mathrm{Lip}})\delta(T)^{c_2 E_2}\, \#I_{m,m'}$$

holds. Then,

$$\#\mathscr{E}_K < K^2 \delta(T)^{O(c_2 E_2)},$$

uniformly for all K as above.

Remark 8.2. Using a trivial bound when $\#I_{m,m'} \leqslant \delta(N)^{c_2 E_2} T/K$, we deduce that

$$\left|\sum_{n\in I_{m,m'}} F(g(mn + \tilde{D}_m)\Gamma)\overline{F(g(m'n + \tilde{D}_{m'})\Gamma)}\right| < \frac{(1 + \|F\|_{\mathrm{Lip}})\delta(T)^{c_2 E_2} T}{K}$$

for all $(m, m') \in (K, 2K]^2 \setminus \mathscr{E}_K$.

Remark 8.3. Proposition 8.1 essentially continues to hold when the variables (m, m') are restricted to pairs of primes. Thanks to a suitable choice of a cutoff parameter X that appears in Section 9C, we will not need this variant of the proposition (cf. Section 9G) and only provide a very brief account of it at the very end of this section.

Proof. To begin with, we endow $G/\Gamma \times G/\Gamma$ with a metric by setting

$$d((x, y), (x', y')) = d_{G/\Gamma}(x, x') + d_{G/\Gamma}(y, y').$$

Let $\tilde{F} : G/\Gamma \times G/\Gamma \to \mathbb{C}$ be defined via $\tilde{F}(\gamma, \gamma') = F(\gamma)\overline{F(\gamma')}$. This is a Lipschitz function. Indeed, the fact that F and $\bar{F}$ are 1-bounded Lipschitz functions allows us to deduce that $\|\tilde{F}\|_{\mathrm{Lip}} \leqslant \|F\|_{\mathrm{Lip}}$. Let $g_{m,m'} : \mathbb{N} \to G \times G$ be the polynomial sequence defined by

$$g_{m,m'}(n) = (g(nm + \tilde{D}_m), g(nm' + \tilde{D}_{m'})).$$

Furthermore, we write $\Gamma' = \Gamma \times \Gamma$. Then $\tilde{F}$ satisfies

$$\int_{G/\Gamma \times G/\Gamma} \tilde{F}(\gamma, \gamma')\, \mathrm{d}(\gamma, \gamma') = \int_{G/\Gamma} F(\gamma) \int_{G/\Gamma} \overline{F(\gamma')}\, \mathrm{d}\gamma'\, \mathrm{d}\gamma = 0.$$

Now, suppose that

$$K \in \bigl[\exp\bigl((\log\log T)^2\bigr), \exp\bigl(H^{-1}(\log T - (\log T)^{1/U})\bigr)\bigr]$$

and that

$$\mathscr{E}_K \geqslant K^2 \delta(T)^{c_2 E_2}.$$

For each pair $(m, m') \in \mathscr{E}_K$, we have

$$\Biggl|\sum_{n \in I_{m,m'}} \tilde{F}(g_{m,m'}(n)\Gamma)\Biggr| > (1 + \|F\|_{\mathrm{Lip}})\delta(T)^{c_2 E_2} \# I_{m,m'}. \tag{8-1}$$

Thus, for every pair $(m, m') \in \mathscr{E}_K$, the corresponding sequence

$$(\tilde{F}(g_{m,m'}(n)\Gamma))_{n \leqslant T/\max(m,m')}$$

fails to be totally $\delta(T)^{c_2 E_2}$-equidistributed. Lemma 7.2 implies that this finite sequence also fails to be $\delta(T)^{c_2 E_2 B}$-equidistributed for some $B \geqslant 1$ that only depends on d and m_G. All implied constants in the sequel will be allowed to depend on d and m_G, without explicit mentioning. By [Green and Tao 2012b, Theorem 2.9][8], there is for each pair $(m, m') \in \mathscr{E}_K$ a nontrivial horizontal character

$$\tilde{\eta}_{m,m'} = \eta_{m,m'} \oplus \eta'_{m,m'} : G \times G \to \mathbb{R}/\mathbb{Z}$$

of modulus $\ll \delta(T)^{-O(c_2 E_2)}$ such that

$$\|\tilde{\eta}_{m,m'} \circ \tilde{g}_{m,m'}\|_{C^\infty[T/\max(m,m')]} \ll \delta(T)^{-O(c_2 E_2)}. \tag{8-2}$$

[8]The $1/\delta(T)$-rational Malcev basis for G/Γ induces one for $G/\Gamma \times G/\Gamma$. Thus [Green and Tao 2012b, Theorem 2.9] is applicable.

Given any nontrivial horizontal character $\tilde{\eta} : G \times G \to \mathbb{R}/\mathbb{Z}$, we define the set

$$\mathscr{M}_\eta = \{(m, m') \in \mathscr{E}_K \mid \tilde{\eta}_{m,m'} = \tilde{\eta}\}.$$

This set is empty unless $|\tilde{\eta}| \ll \delta(T)^{-O(c_2 E_2)}$. Pigeonholing over all nontrivial $\tilde{\eta}$ of modulus bounded by $\delta(T)^{-O(c_2 E_2)}$, we find that there is some $\tilde{\eta}$ amongst them for which

$$\#\mathscr{M}_{\tilde{\eta}} > K^2 \delta(T)^{O(c_2 E_2)}.$$

Let us fix such a character $\tilde{\eta} = \eta \oplus \eta'$ and suppose without loss of generality that the component η is nontrivial. Suppose

$$\tilde{\eta} \circ (g(n), g(n')) = (\alpha_d n^d + \alpha'_d n'^d) + \cdots + (\alpha_1 n + \alpha'_1 n') + (\alpha_0 + \alpha'_0)$$

and define for $(m, m') \in \mathscr{E}_K$ the coefficients $\alpha_j(m, m')$, $1 \leqslant j \leqslant d$, via

$$\tilde{\eta} \circ g_{m,m'}(n) = \alpha_d(m, m') n^d + \cdots + \alpha_1(m, m') n + \alpha_0(m, m').$$

Then the bound (8-2) on the smoothness norm asserts that

$$\sup_{1 \leqslant j \leqslant d} \frac{T^j}{K^j} \|\alpha_j(m, m')\| \ll \delta(T)^{-O(c_2 E_2)}. \tag{8-3}$$

Observe that each $\alpha_j(m, m')$, $1 \leqslant j \leqslant d$, satisfies

$$\alpha_j(m, m') = \sum_{i=j}^{d} \binom{i}{j} (\tilde{D}_m^{i-j} \alpha_i m^j + \tilde{D}_{m'}^{i-j} \alpha'_i m'^j). \tag{8-4}$$

We now aim to show with a downwards induction starting from $j = d$ that

$$\alpha_j = \frac{a_j}{\kappa_j} + \tilde{\alpha}_j, \tag{8-5}$$

where $1 \leqslant \kappa_j \ll \delta(T)^{-O(c_2 E_2)}$, $\gcd(a_j, \kappa_j) = 1$, and

$$\tilde{\alpha}_j \ll \delta(T)^{-O(c_2 E_2)} T^{-j}. \tag{8-6}$$

Suppose $j_0 \leqslant d$ and that the above holds for all $j > j_0$. Set $k_{j_0} = \operatorname{lcm}(\kappa_{j_0+1}, \ldots, \kappa_d)$ if $j_0 < d$, and $k_{j_0} = 1$ when $j_0 = d$. Note that $k_{j_0} \ll \delta(T)^{-O(c_2 E_2)}$.

Pigeonholing, we find that there is $\tilde{m}'$ such that $m' = \tilde{m}'$ for $\gg K\delta(T)^{O(c_2 E_2)}$ pairs $(m, m') \in \mathscr{M}_{\tilde{\eta}}$. Amongst these there are furthermore $\gg K\delta(T)^{O(c_2 E_2)}$ values of m that belong to the same fixed residue class modulo k_{j_0}. Denote this set of integers m by $\mathscr{M}'$. Suppose $m = k_{j_0} m_1 + m_0 \in \mathscr{M}'$. Letting $\{x\}$ denote the fractional part of $x \in \mathbb{R}$, we then have

$$\{\tilde{D}_m^{i-j_0} \alpha_i m^{j_0}\} = \left\{ \tilde{D}_m^{i-j_0} \tilde{\alpha}_i m^{j_0} + \frac{a_i m_0^{j_0}}{\kappa_i} \right\}, \quad (i \geqslant j_0),$$

where, in view of (8-6),

$$\tilde{D}_m^{i-j_0} \tilde{\alpha}_i m^{j_0} \ll \delta(T)^{-O(c_2 E_2)} K^i T^{-i}.$$

Since m_0 is fixed, it thus follows from (8-3), (8-4), (8-5) and the above bound that as m varies over $\mathscr{M}'$, the value of

$$\|\alpha_{j_0} m^{j_0}\|$$

lies in a fixed interval of length $\ll \delta(T)^{-O(c_2E_2)} K^{j_0} T^{-j_0}$.

We aim to make use of this information in combination with the Waring trick from [Green and Tao 2012a, §3] that was already employed in Section 7. For this purpose, we consider the set of integers

$$\mathscr{M}^* = \{m \leqslant s(2K)^{j_0} : m = m_1^{j_0} + \cdots + m_s^{j_0},\ m_1, \ldots, m_s \in \mathscr{M}'\}$$

with $s \geqslant 2^{j_0} + 1$. For each element $m \in \mathscr{M}^*$ of this set, $\|\alpha_{j_0} m\|$ lies in an interval of length $\ll_s \delta(T)^{-O(c_2E_2)} K^{j_0} T^{-j_0}$. Furthermore, Lemma 7.6 implies that $\#\mathscr{M}^* \gg \delta(T)^{O(c_2E_2)} K^{j_0}$. The restrictions on the size of K and the assumptions on the function δ imply that the conditions of Lemma 7.5 are satisfied once T is sufficiently large. Thus, assuming T is sufficiently large, there is an integer $1 \leqslant \kappa'_{j_0} \ll \delta(T)^{-O(c_2E_2)}$ such that

$$\|\kappa'_{j_0} \alpha_{j_0}\| \ll \delta(T)^{-O(c_2E_2)} T^{-j_0},$$

i.e.,

$$\alpha_{j_0} = \frac{a_{j_0}}{\kappa_{j_0}} + \tilde{\alpha}_{j_0},$$

where $\kappa_{j_0} \mid \kappa'_{j_0}$, $\gcd(a_{j_0}, \kappa_{j_0}) = 1$ and $\tilde{\alpha}_{j_0} \ll \delta(T)^{-O(c_2E_2)} T^{-j_0}$, as claimed.

In particular, we have

$$\|\kappa_j \alpha_j\| \ll \delta(T)^{-O(c_2E_2)} T^{-j} \tag{8-7}$$

for $1 \leqslant j \leqslant d$. Proceeding as in [Green and Tao 2012a, §3], let $\kappa = \operatorname{lcm}(\kappa_1, \ldots, \kappa_d)$ and set $\tilde{\eta} = \kappa\eta$. Then (8-7) implies that

$$\|\tilde{\eta} \circ g\|_{C^\infty[T]} = \sup_{1 \leqslant j \leqslant d} T^j \|\kappa \alpha_j\| \ll \delta(T)^{-O(c_2E_2)},$$

which in turn shows that

$$\|\tilde{\eta} \circ g(n)\|_{\mathbb{R}/\mathbb{Z}} \ll \delta(T)^{-O(c_2E_2)} n/T$$

for every $n \in \{1, \ldots, T\}$. Note that the latter bound can be controlled by restricting n to a smaller range. For this, set $T' = \delta(T)^{c_2E_2C} T$ for some constant $C \geqslant 1$ depending only on d and m_G, chosen sufficiently large to guarantee that

$$\|\tilde{\eta} \circ g(n)\|_{\mathbb{R}/\mathbb{Z}} \leqslant 1/10,$$

whenever $n \in \{1, \ldots, T'\}$. Let $\chi : \mathbb{R}/\mathbb{Z} \to [-1, 1]$ be a function of bounded Lipschitz norm that equals 1 on $\left[-\frac{1}{10}, \frac{1}{10}\right]$ and satisfies $\int_{\mathbb{R}/\mathbb{Z}} \chi(t)\, dt = 0$. Then, by setting $F := \chi \circ \tilde{\eta}$, we obtain a function $F : G/\Gamma \to [-1, 1]$ such that $\int_{G/\Gamma} F = 0$ and $\|F\|_{\mathrm{Lip}} \ll \delta(T)^{-O(c_2E_2)}$. Choosing c_2 sufficiently small, we may ensure that

$$\|F\|_{\mathrm{Lip}} < \delta(T)^{-E_2}$$

and, moreover, that

$$T' > \delta(T)^{E_2} T.$$

The quantities T', F and c_2 are chosen in such a way that

$$\left| \frac{1}{T'} \sum_{1 \leqslant n \leqslant T'} F(g(n)\Gamma) \right| = 1 > \delta(T)^{E_2} \|F\|_{\mathrm{Lip}},$$

This contradicts the fact that $(g(n)\Gamma)_{n \leqslant T}$ is totally $\delta(T)^{E_2}$-equidistributed and completes the proof of the proposition. □

8A. ***Restriction of Proposition 8.1 to pairs of primes.*** We end this section by making the contents of Remark 8.3 more precise. The variables (m, m') in Proposition 8.1 can without much additional effort be restricted to range over pairs of primes. It is clear that in the above proof all applications of the pigeonhole principle that involve the parameters m and m' have to be restricted to the set of primes. The only true difference lies in the application of Waring's result: Lemma 7.6 needs to be replaced by the following one.

Lemma 8.4. *Let $K \geqslant 1$ be an integer and let $S \subset \{1, \ldots, K\} \cap \mathscr{P}$ be a subset of the primes less than K. Suppose $\#S = \alpha K / \log K$. Let $s \geqslant 2^k + 1$. Let $X \subset \{1, \ldots, sK^k\}$ denote the set of integers that are representable as $p_1^k + \cdots + p_s^k$ with $p_1, \ldots, p_s \in S$. Then*

$$|X| \gg_{k,s} \alpha^{2s} K^k,$$

as $K \to \infty$.

Proof. Let $I_s(N)$ denote the number of solutions to the equation

$$p_1^k + \cdots + p_s^k = N$$

in positive prime numbers $p_1, \ldots, p_s$. Hua's asymptotic formula [Hua 1965, Theorem 11] for the Waring–Goldbach problem implies that

$$I_s(N) \ll_{k,s} \frac{N^{s/k-1}}{(\log N)^s}.$$

Thus, for $1 \leqslant n \leqslant sK^k$, we have

$$I_s(n) \ll_{k,s} \frac{K^{s-k}}{(\log K)^s}.$$

Hence,

$$\alpha^{2s} \frac{K^{2s}}{(\log K)^{2s}} = \left(\sum_{n=1}^{sK^k} I_s(n) \right)^2 \leqslant |X| \sum_n I_s^2(n)$$
$$\ll_{k,s} |X| K^k \frac{K^{2s-2k}}{(\log K)^{2s}} \ll_{k,s} |X| \frac{K^{2s-k}}{(\log K)^{2s}}.$$

Rearranging completes the proof of the lemma. □

9. Proof of Proposition 6.4

In this section we prove Proposition 6.4 by invoking the possibly trivial Dirichlet decomposition from Lemma 1.8. Let $f \in \mathscr{M}_H$, let h, h' be as in Lemma 1.8 and let $f = f_1 * \cdots * f_H$ with $f_i = h$ for $i < H$ and $f_H = h * h'$. We are given integers Q and A such that $0 \leqslant A < Q \leqslant (\log N)^E$ and such that $A \in (\mathbb{Z}/Q\mathbb{Z})^*$. Recall that $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$ is a polynomial sequence with the property that $(g(n)\Gamma)_{n \leqslant T/Q}$ is totally $\delta(N)^{E_0}$-equidistributed in G/Γ. Let $I \subset \{1, \ldots, T/Q\}$ be a discrete interval of length at least $T/(Q(\log N)^E)$. Our aim is to bound above the expression

$$\left| \frac{Q}{T} \sum_{n \in I} f(Qn + A) F(g(n)\Gamma) \right|. \tag{9-1}$$

If $H = 1$, then we may write this expression as

$$\left| \frac{Q}{T} \sum_{n \in I} f(Qn + D') F(g(Dn + D'')\Gamma) \right|, \tag{9-2}$$

where $D = 1$, $D' = A$ and $D'' = 0$. The aim of the next two sections is to show that in the case where $H > 1$ and the Dirichlet decomposition is nontrivial, the task of bounding (9-1) can be reduced to that of bounding an expression similar to (9-2), but with f replaced by one of the f_i.

9A. *Reduction by hyperbola method.* Taking into account that $f = f_1 * \cdots * f_H$, the correlation from Proposition 6.4 may be written as

$$\begin{aligned} &\frac{Q}{T} \sum_{n \leqslant T/Q} \mathbf{1}_I(n) f(Qn + A) F(g(n)\Gamma) \\ &\qquad = \frac{Q}{T} \sum_{\substack{d_1 \cdots d_H \leqslant T \\ d_1 \cdots d_H \equiv A \\ (\mathrm{mod}\ Q)}} f_1(d_1) f_2(d_2) \cdots f_H(d_H) F\left(g\left(\frac{d_1 \cdots d_H - A}{Q} \right) \Gamma \right) \mathbf{1}_P(d_1 \cdots d_H), \end{aligned} \tag{9-3}$$

where P is the finite progression defined via $P = QI + A$. Our first step is to split this summation via inclusion-exclusion into a finite sum of weighted correlations of individual factors f_i with a nilsequence. To describe these weighted correlations, let $i \in \{1, \ldots, H\}$. For every $j \neq i$, let d_j be a fixed positive integer and write $D_i := \prod_{j \neq i} d_j$. Let $a_i \in [0, T/D_i)$ be an integer. Weighted correlations involving f_i will then take the form

$$\begin{aligned} &\frac{Q}{T} \left(\prod_{j \neq i} f_j(d_j) \right) \sum_{\substack{a_i < d_i \leqslant T/D_i \\ d_i D_i \equiv A\ (\mathrm{mod}\ Q)}} \mathbf{1}_P(d_i D_i) f_i(d_i) F\left(g\left(\frac{d_i D_i - A}{Q} \right) \Gamma \right) \\ &\qquad = \frac{Q}{T} \left(\prod_{j \neq i} f_j(d_j) \right) \sum_{\frac{a_i - D_i'}{Q} < n \leqslant \frac{T - D_i'}{D_i Q}} f_i(Qn + D_i') F(g(D_i n + D_i'')\Gamma) \mathbf{1}_I(D_i n + D_i''), \end{aligned} \tag{9-4}$$

for suitable integers D_i', D_i'', determined by the values of $D_i \pmod Q$ and A. In order to bound correlations of the form (9-4), we need to ensure that d_i runs over a sufficiently long range, which will be achieved by arranging for $D_i \leqslant T^{1-1/H}$ to hold.

Let $\tau = T^{1-1/H}$ and note that $D_i = D_j d_j / d_i$. Hence,

$$D_i > \tau \iff d_j > \frac{\tau d_i}{D_j}.$$

With the help of this equivalence, the function $\mathbf{1} : \mathbb{Z}^H \to 1$ can be decomposed as follows. Suppose $d_1 \cdots d_H \leqslant T$. Then

$$\begin{aligned}\mathbf{1}(d_1, \ldots, d_H) &= \mathbf{1}_{D_1 \leqslant \tau} + \mathbf{1}_{D_1 > \tau}\big(\mathbf{1}_{D_2 \leqslant \tau} + \mathbf{1}_{D_2 > \tau}(\mathbf{1}_{D_3 \leqslant \tau} + \cdots (\mathbf{1}_{D_H \leqslant \tau} + \mathbf{1}_{D_H > \tau}) \cdots)\big) \\ &= \mathbf{1}_{D_1 \leqslant \tau} + \mathbf{1}_{D_1 > \tau}\big(\mathbf{1}_{D_2 \leqslant \tau}\mathbf{1}_{d_2 > \tau d_1/D_2} + \mathbf{1}_{D_2 > \tau}(\mathbf{1}_{D_3 \leqslant \tau}\mathbf{1}_{d_3 > \tau \max(d_1,d_2)/D_3} + \cdots \\ &\qquad\qquad \cdots + \mathbf{1}_{D_{H-1} > \tau}(\mathbf{1}_{D_H \leqslant \tau} + \mathbf{1}_{D_H > \tau}) \cdots)\big) \\ &= \mathbf{1}_{D_1 \leqslant \tau} + \mathbf{1}_{D_2 \leqslant \tau}\mathbf{1}_{d_2 > \tau d_1/D_2} + \mathbf{1}_{D_3 \leqslant \tau}\mathbf{1}_{d_3 > \tau \max(d_1,d_2)/D_3} + \cdots + \mathbf{1}_{D_H \leqslant \tau}\mathbf{1}_{d_H > \tau \max(d_1,\ldots,d_{H-1})/D_H}.\end{aligned}$$

Thus,

$$\sum_{d_1 \cdots d_H < T} = \sum_{i=1}^{H} \sum_{D \leqslant T^{1-1/H}} \sum_{\substack{d_1,\ldots,\widehat{d_i},\ldots,d_H \\ D_i = D}} \sum_{\substack{d_i \leqslant T/D_i \\ d_i > \tau \max(d_1,\ldots,d_{i-1})/D_i}}.$$

This shows that the original summation (9-3) may be decomposed as a sum of summations of the shape (9-4) while only increasing the total number of terms by a factor of order $O(H)$. Expressing, if necessary, the summation range

$$\left(\frac{\tau \max(d_1, \ldots, d_{i-1})}{D_i}, \frac{T}{D_i}\right)$$

of d_i as the difference of two intervals starting from 1, we can ensure that d_i runs over an interval of length $\gg T/D_i \gg T^{1/H}$. The correlation now decomposes as

$$\begin{aligned}&\frac{Q}{T} \sum_{\substack{d_1 \cdots d_H \leqslant T \\ d_1 \cdots d_H \equiv A \\ (\mathrm{mod}\ Q)}} f_1(d_1) f_2(d_2) \cdots f_H(d_H) F\left(g\left(\frac{d_1 \cdots d_H - A}{Q}\right)\Gamma\right)\mathbf{1}_P(d_1 \cdots d_H) \\ &\leqslant \sum_{i=1}^{H} \sum_{k=0}^{\frac{1-1/H}{\log 2}\log T} \sum_{\substack{D \sim 2^k \\ (D,Q)=1}} \sum_{\substack{d_1,\ldots,\widehat{d_i},\ldots,d_H \\ D_i = D}} \left(\prod_{j \neq i} \frac{|f_j(d_j)|}{d_i}\right) \left|\frac{DQ}{T} \sum_{\substack{n \leqslant T/D \\ n > \tau \max(d_1,\ldots,d_{i-1})/D \\ Dn + D'' \in I}} f_i(Qn + D') F(g(Dn + D''))\right|. \qquad (9\text{-}5)\end{aligned}$$

Observe that (9-2) can be regarded as the special case $H = 1$ and $D = 1$ of this bound. Our next aim is to analyze the innermost sum of (9-5) as $D \sim 2^k$ varies. Setting $E_1 = E_0$, we deduce from Proposition 7.4 that whenever $2^k \in [\exp((\log\log T)^2), (\log T)^{1-1/H}]$ then there is a set $\mathscr{B}_{2^k}$ of cardinality at most $O(\delta(N)^{c_1 E_0} 2^k)$ such that for each $D \sim 2^k$ with $D \notin \mathscr{B}_{2^k}$ the sequence

$$(g_D(n)\Gamma)_{n \leqslant T/Q}, \quad g_D(n) := g(Dn + D''),$$

is totally $\delta(N)^{c_1 E_0}$-equidistributed. Before turning to the case of $D \notin \mathscr{B}_{2^k}$, we bound the total contribution from exceptional D, that is, from $D \in \mathscr{B}_{2^k}$ and from $D \leqslant \exp((\log\log T)^2)$.

9B. *Contribution from exceptional D.* Let $\mathscr{B}_{2^k}$ denote the exceptional set from the previous section.

Lemma 9.1. *Whenever E_0 is sufficiently large in terms of d, m_G and H, we have*

$$\sum_{\frac{(\log\log T)^2}{\log 2} < k \leqslant \frac{(1-1/H)\log T}{\log 2}} \sum_{D \in \mathscr{B}_{2^k}} \sum_{\substack{d_1 \cdots d_H \leqslant T \\ d_1 \cdots d_H \equiv A \pmod Q \\ D_i = D}} |f_1(d_1) f_2(d_2) \cdots f_H(d_H)| \mathbf{1}_P(d_1 \cdots d_H) \ll_t \frac{T}{Q} \frac{1}{(\log T)^2}$$

and

$$\sum_{\substack{D \leqslant \exp((\log\log T)^2) \\ \gcd(D,W)=1}} \sum_{\substack{d_1 \cdots d_H \leqslant T \\ d_1 \cdots d_H \equiv A \pmod Q \\ D_i = D}} |f_1(d_1) f_2(d_2) \cdots f_H(d_H)| \mathbf{1}_P(d_1 \cdots d_H)$$
$$\ll (\log\log T)^{2H} \frac{T}{Q} \frac{Q}{\phi(Q)} \frac{1}{\log T} \prod_{\substack{p \leqslant T \\ p \nmid Q}} \left(1 + \frac{|f(p)|}{Hp}\right).$$

Before we prove this lemma, let us consider its contribution to the bound in Proposition 6.4. The contribution from the first part is easily seen to be negligible. Regarding the second part, recall that $H > 1$ and note that by property (2) of Definition 1.3, we have

$$\prod_{Q < p \leqslant T} \left(1 + \frac{(H-1)|f(p)|}{Hp}\right) \gg \left(\frac{\log T}{E \log\log T}\right)^{(H-1)\alpha_f/H},$$

where the exponent is positive. Thus, the bound in the second part saves a power of $\log x$ when compared with the bound in (6-2) and is therefore also negligible.

Proof. Set

$$\bar{f}_i(n) := |f_1 * \dots \widehat{f_i} \dots * f_H(n)|.$$

Then $\bar{f}_i = |h^{*(H-1)} * h'|$ or $|h^{*(H-1)}|$ and it follows from (3-2) and the properties of h that $\bar{f}_i(n) \leqslant (CH)^{\Omega(n)}$ for some constant C. This implies a second moment bound of the form

$$\sum_{\substack{n \leqslant x \\ \gcd(n,Q)=1}} \bar{f}_i(n)^2 \leqslant x \sum_{\substack{n \leqslant x \\ \gcd(n,Q)=1}} \frac{\bar{f}_i(n)^2}{n} \leqslant x \prod_{w(N) < p \leqslant x} \left(1 - \frac{(CH)^2}{p}\right)^{-1} \leqslant x (\log x)^{O(H^2)}.$$

Similarly, we have

$$\sum_{\substack{n \leqslant x \\ \gcd(n,Q)=1}} |f_i(n)| \ll x (\log x)^{O(H)}.$$

Since Proposition 7.4 provides the bound $\#\mathscr{B}^*_{2^k} \ll \delta(N)^{c_1 E_0} 2^k$, a trivial application of the Cauchy–Schwarz inequality yields

$$\sum_{D\in\mathscr{B}^*_{2^k}} \bar{f}_i(D) \sum_{\substack{n\leqslant T/D\\ nD\equiv A \pmod Q\\ nD\in P}} |f_i(n)| \leqslant \sum_{\substack{n\leqslant T/2^k\\ \gcd(n,Q)=1}} |f_i(n)| \sum_{\substack{D\in\mathscr{B}^*_{2^k}\\ \gcd(n,Q)=1}} \bar{f}_i(D)$$
$$\leqslant \sum_{\substack{n\leqslant T/2^k\\ \gcd(n,Q)=1}} |f_i(n)| 2^k \delta(N)^{c_1E_0} k^{O(H^2)} \leqslant T(\log T)^{O(H)}\delta(N)^{c_1E_0}k^{O(H^2)}.$$

Recall that c_1 only depends on d and m_G, and that by the assumptions of Proposition 6.4 we have $\delta(N) \ll (\log T)^{-1}$. Since the summation in k has length at most $\log T$ and since $k^{O(H^2)} < (\log T)^{O_H(1)}$ for each k, the first part of the lemma follows by choosing E_0 sufficiently large in terms of d, m_G and H.

Concerning the second part, we have

$$\sum_{\substack{D\leqslant\exp((\log\log T)^2)\\ \gcd(D,Q)=1}} \bar{f}_i(D) \sum_{\substack{n\leqslant T/D\\ nD\equiv A \pmod Q\\ nD\in P}} |f_i(n)| \leqslant \sum_{\substack{D\leqslant\exp((\log\log T)^2)\\ \gcd(D,Q)=1}} \bar{f}_i(D) \sum_{\substack{n\leqslant T/D\\ n\equiv A\bar{D}\\ \pmod Q}} |f_i(n)|,$$

where $\bar{D}D \equiv 1 \pmod Q$. Since $\log(T/D) \asymp \log T$ and $T/D < T$, Shiu's bound (3-1) yields the upper bound

$$\ll \sum_{\substack{D\leqslant\exp((\log\log T)^2)\\ \gcd(D,Q)=1}} \frac{\bar{f}_i(D)}{D}\frac{T}{Q}\frac{Q}{\phi(Q)}\frac{1}{\log T}\prod_{\substack{p\leqslant T\\ p\nmid Q}}\left(1+\frac{|f_i(p)|}{p}\right). \tag{9-6}$$

The outer sum satisfies

$$\sum_{\substack{D\leqslant\exp((\log\log T)^2)\\ \gcd(D,Q)=1}} \frac{\bar{f}_i(D)}{D} \ll \prod_{w(N)<p\leqslant\exp((\log\log T)^2)}\left(1+\frac{|f(p)|}{Hp}\right)^{H-1}$$
$$\ll \exp\left((H-1)\sum_{p\leqslant\exp((\log\log T)^2)}\frac{1}{p}\right) \ll (\log\log T)^{2H},$$

which completes the proof of the second part. □

9C. ***Montgomery–Vaughan approach.*** Since $\mathscr{M}_1 \subset \mathscr{M}_H$ for all $H > 1$, it suffices to prove Proposition 6.4 for $H > 1$. Since the task of bounding (9-2) for $D = 1$ presents an easier special case of the task of bounding the inner sum of (9-5) for unexceptional D when $H > 1$, a proof for the $H = 1$ case may, however, be extracted from the argument below. In fact, most of the argument directly applies when setting $H = D = 1$. The main differences leading to simplifications are that

(1) if $H = D = 1$, one can, instead of later referring to the results from Section 7, directly work with the equidistribution properties of the given polynomial sequence g, and

(2) the extra work of handling the outer sums in (9-5) is not required when $H = D = 1$.

From now on we assume that $H > 1$ and that D is unexceptional, that is $D \sim 2^k$ for $k \geqslant (\log\log T)^2/\log 2$ and $D \notin \mathscr{B}_{2^k}$, where $\mathscr{B}_{2^k}$ is the exceptional set from Section 9A. To bound the inner sum of (9-5) for unexceptional D, we employ the strategy of Montgomery and Vaughan [1977] outlined in Section 2, and begin by introducing a factor $\log n$ into the average. This will later allow us to reduce matters to understanding equidistribution along sequences defined in terms of primes. We set $h = f_i$. We caution that this is *not* the function h from Lemma 1.8, but could either be h or $h * h'$ in the notation of the lemma.

Cauchy–Schwarz and several integral comparisons show that

$$\sum_{n \leqslant T/(DQ)} \mathbf{1}_I(Dn+D'')h(Qn+D')F(g(Dn+D'')\Gamma)\log\biggl(\frac{T/D}{Qn+D'}\biggr)$$
$$\leqslant \biggl(\sum_{n \leqslant T/(DQ)} \biggl(\log\biggl(\frac{T}{DQ}\biggr) - \log n\biggr)^2\biggr)^{1/2} \biggl(\sum_{n \leqslant T/(DQ)} h^2(Qn+D')\biggr)^{1/2}$$
$$\ll \frac{T}{DQ}\sqrt{\frac{DQ}{T}\sum_{n \leqslant T/(DQ)} h^2(Qn+D')},$$

and hence, invoking $D \leqslant T^{1-1/H}$,

$$\frac{DQ}{T}\sum_{\substack{n \leqslant T/(DQ) \\ Dn+D'' \in I}} h(Qn+D')F(g(Dn+D'')\Gamma)$$
$$\ll_H \frac{1}{\log T}\sqrt{\frac{DQ}{T}\sum_{n \leqslant T/(DQ)} h^2(Qn+D')} + \frac{1}{\log T}\biggl|\frac{DQ}{T}\sum_{\substack{n \leqslant T/(DQ) \\ Dn+D'' \in I}} h(Qn+D')F(g(Dn+D'')\Gamma)\log(Qn+D')\biggr|. \tag{9-7}$$

Lemma 1.8 shows that the contribution of the first term in this bound to (9-5) is at most

$$O_H\biggl(\frac{1}{(\log T)^{1/2}}\frac{Q}{\phi(Q)}\frac{1}{\log x}\prod_{\substack{p \leqslant x \\ p \nmid Q}}\biggl(1 + \frac{|f(p)|}{p}\biggr)\biggr),$$

which is negligible in view of the bound stated in Proposition 6.4. It remains to estimate the second term from (9-7). For this, it will be convenient to abbreviate

$$g_D(n) := g(Dn+D''),$$

and to introduce the two finite progressions

$$I_D = \{n : Dn+D'' \in I\} \quad \text{and} \quad P_D = \biggl\{n : \frac{n-D'}{Q} \in I_D\biggr\}. \tag{9-8}$$

Since $\log n = \sum_{m|n} \Lambda(m)$, our task is to bound

$$\frac{DQ}{T\log T}\biggl|\sum_{\substack{mn \leqslant T/D \\ mn \equiv D' \ (\mathrm{mod}\ Q)}} \mathbf{1}_{P_D}(nm)h(nm)\Lambda(m)F\biggl(g_D\biggl(\frac{nm-D'}{Q}\biggr)\Gamma\biggr)\biggr|. \tag{9-9}$$

To further simplify this expression we now show that one can, at the expense of a small error term, restrict the summation in (9-9) to pairs (m, n) of coprime integers for which $m = p$ is prime. To see this, recall that F is 1-bounded and observe that

$$\sum_{\substack{nm\leqslant T/D\\ \Omega(m)\geqslant 2 \text{ or } \gcd(n,m)>1\\ mn\equiv D' \ (\mathrm{mod}\ Q)}} |h(nm)|\Lambda(m) \leqslant 2\sum_{p}\sum_{k\geqslant 2} k\log p \sum_{\substack{n\leqslant T/D,\, p^k\| n\\ n\equiv D' \ (\mathrm{mod}\ Q)}} |h(n)|$$
$$\leqslant 2\sum_{p>w(N)}\sum_{k\geqslant 2} H^k k\log p \sum_{\substack{n\leqslant T/(Dp^k)\\ p^k n\equiv D' \ (\mathrm{mod}\ Q)}} |h(n)|.$$

If $p^k \leqslant (T/D)^{1/2}$, then Shiu's bound (3-1) implies for the inner sum:

$$\sum_{\substack{n\leqslant T/(Dp^k)\\ p^k n\equiv D' \ (\mathrm{mod}\ Q)}} |h(n)| \ll \frac{1}{p^k}\frac{T}{D}\frac{1}{\phi(Q)}\frac{1}{\log T}\prod_{\substack{p\leqslant T/D\\ p\nmid Q}}\left(1+\frac{|h(p)|}{p}\right).$$

If N is sufficiently large, then $H\log p \ll p^{1/4}$ for all $p > w(N)$ and thus

$$\sum_{p>w(N)}\sum_{\substack{k\geqslant 2\\ p^k\leqslant (T/D)^{1/2}}} \frac{H^k\log p^k}{p^k} \ll \sum_{p>w(N)}\frac{1}{p^{2-1/2}} \ll \frac{1}{w(N)^{1/2}}.$$

Combining the last three steps, the contribution to (9-9) from the terms $p^k \leqslant (T/D)^{1/2}$ is seen to be bounded by

$$\ll \frac{1}{w(N)^{1/2}\log T}\frac{Q}{\phi(Q)}\frac{1}{\log T}\prod_{\substack{p\leqslant T/D\\ p\nmid Q}}\left(1+\frac{|h(p)|}{p}\right).$$

Turning towards the case of $p^k > (T/D)^{1/2}$, note first that, provided N is large enough that $w(N) > H$, then

$$\sum_{\substack{n\leqslant T/(Dp^k)\\ p^k n\equiv D' \ (\mathrm{mod}\ Q)}} |h(n)| \leqslant \frac{T}{Dp^k}\sum_{\substack{n\leqslant T/(Dp^k)\\ \gcd(n,Q)=1}}\frac{|h(n)|}{n} \leqslant \frac{T}{Dp^k}\prod_{w(N)<p'\leqslant T/(Dp^k)}\left(1-\frac{H}{p'}\right)^{-1} \leqslant \frac{T}{Dp^k}\left(\log_+\frac{T}{Dp^k}\right)^{O(H)},$$

where $\log_+(x) = \max\{\log x, 0\}$ for $x > 0$, as usual. Assuming, again, that $H\log p \ll p^{1/4}$ for all $p > w(N)$, the remaining sum over $p^k > (T/D)^{1/2}$ therefore satisfies

$$\frac{DQ}{T\log T}\sum_{p>w(N)}\sum_{\substack{k\geqslant 2\\ p^k>(T/D)^{1/2}}} H^k\log p^k \sum_{\substack{n\leqslant T/(Dp^k)\\ p^k n\equiv D' \ (\mathrm{mod}\ Q)}} |h(n)|$$
$$\leqslant \frac{DQ}{T\log T}\sum_{p>w(N)}\sum_{\substack{k\geqslant 2\\ p^k>(T/D)^{1/2}}} H^k\log p^k\,\frac{T}{Dp^k}\left(\log_+\frac{T}{Dp^k}\right)^{O(H)},$$

which is further bounded by

$$\begin{aligned}
&\ll \frac{Q}{\log T}\Big(\log\frac{T}{D}\Big)^{O(H)} \sum_{p>w(N)} \sum_{\substack{k\geqslant 2\\ p^k>(T/D)^{1/2}}} \frac{H^k \log p^k}{p^k}\\
&\ll \frac{Q}{\log T}\Big(\log\frac{T}{D}\Big)^{O(H)} \sum_{p>w(N)} \sum_{\substack{k\geqslant 2\\ p^k>(T/D)^{1/2}}} p^{-k(1-1/4)}\\
&\ll \frac{Q}{\log T}\Big(\log\frac{T}{D}\Big)^{O(H)} \sum_{p>w(N)} p^{-2+1/2}p^{1/4}\Big(\frac{T}{D}\Big)^{-1/4}\\
&\ll \frac{Q}{\log T}\Big(\frac{T}{D}\Big)^{-1/4}\Big(\log\frac{T}{D}\Big)^{O(H)} \ll T^{-1/8H}.
\end{aligned}$$

This contribution is dominated by that of the smaller prime powers above.

Thus, the total contribution to (9-5) of pairs (m,n) that are not of the form (m,p), where p is prime and does not divide m, is bounded by

$$\frac{1}{\log T}\sum_{i=1}^{t}\sum_{k}\sum_{D\sim 2^k}\sum_{d_1\cdots\hat{d}_i\cdots d_t=D}\Big(\prod_{j\neq i}\frac{|f_j(d_j)|}{d_i}\Big)\frac{Q}{\phi(Q)}\frac{1}{\log T}\prod_{\substack{p\leqslant T/D\\ p\nmid Q}}\Big(1+\frac{|h(p)|}{p}\Big)$$
$$\leqslant \frac{1}{w(N)^{1/2}\log T}\frac{Q}{\phi(Q)}\frac{1}{\log T}\prod_{\substack{p\leqslant T\\ p\nmid Q}}\Big(1+\frac{|f(p)|}{p}\Big),$$

which is negligible in view of the bound claimed in Proposition 6.4.

This reduces the task of proving Proposition 6.4 to that of bounding the expression

$$\frac{DQ}{T}\Bigg|\sum_{\substack{mp\leqslant T/D\\ mp\equiv D' \pmod Q}} \mathbf{1}_{P_D}(mp)h(m)h(p)\Lambda(p)F\Big(g\Big(\frac{pm-D'}{Q}\Big)\Gamma\Big)\Bigg|. \tag{9-10}$$

9D. ***Decomposing the summation range.*** We prepare the analysis of (9-10) by first splitting the summation into large and small divisors with respect to the parameter

$$X = X(D) = \Big(\frac{T}{D}\Big)^{1-1/(\log\frac{T}{D})^{(U-1)/U}},$$

for a fixed integer $U\geqslant 4$. With this choice of X we obtain

$$\frac{QD}{T}\sum_{\substack{m<X\\ \gcd(m,Q)=1}}\sum_{\substack{p\leqslant T/(mD)\\ p\equiv D'\overline{m} \pmod Q}} \mathbf{1}_{P_D}(mp)h(m)h(p)\Lambda(p)F\Big(g\Big(\frac{pm-D'}{Q}\Big)\Gamma\Big)$$
$$+\frac{QD}{T}\sum_{\substack{m>X\\ \gcd(m,Q)=1}}\sum_{\substack{p\leqslant T/(mD)\\ p\equiv D'\overline{m} \pmod Q}} \mathbf{1}_{P_D}(mp)h(m)h(p)\Lambda(p)F\Big(g\Big(\frac{pm-D'}{Q}\Big)\Gamma\Big). \tag{9-11}$$

In order to analyze these expressions, we dyadically decompose in each of the two terms the sum with shorter summation range. The cutoff parameter X is chosen in such a way that one of the dyadic decompositions is of short length, depending on U. Indeed, we have $\log_2 X \sim \log_2(T/D)$, while

$$\log_2 \frac{T}{DX} = \frac{(\log \frac{T}{D})^{1/U}}{\log 2}.$$

Define

$$T_0 = \exp((\log\log T)^2).$$

Then the two sums from (9-11) decompose as

$$\frac{QD}{T} \sum_{\substack{m<T_0 \\ \gcd(m,Q)=1}} \sum_{\substack{p\leqslant \frac{T}{(mD)} \\ p\equiv D'\bar{m} \pmod Q}} \mathbf{1}_{P_D}(mp)h(m)h(p)\Lambda(p)F\biggl(g_D\Bigl(\frac{pm-D'}{Q}\Bigr)\Gamma\biggr)$$
$$+\frac{QD}{T}\sum_{j=1}^{\log_2 \frac{X}{T_0}} \sum_{\substack{m\sim 2^{-j}X \\ \gcd(m,Q)=1}} \sum_{\substack{p\leqslant \frac{T}{(mD)} \\ p\equiv D'\bar{m} \pmod Q}} \mathbf{1}_{P_D}(mp)h(m)h(p)\Lambda(p)F\biggl(g_D\Bigl(\frac{pm-D'}{Q}\Bigr)\Gamma\biggr) \tag{9-12}$$

and

$$\frac{QD}{T}\Biggl\{ \sum_{\substack{m>X \\ \gcd(m,Q)=1}} \sum_{\substack{p\leqslant \min(\frac{T}{mD},T_0) \\ p\equiv D'\bar{m} \pmod Q}} \mathbf{1}_{P_D}(mp)h(m)h(p)\Lambda(p)F\biggl(g_D\Bigl(\frac{pm-D'}{Q}\Bigr)\Gamma\biggr)$$
$$+\sum_{j=1}^{\log_2 \frac{T}{XDT_0}} \sum_{\substack{m>X \\ \gcd(m,Q)=1}} \sum_{\substack{p\sim 2^{-j}\frac{T}{XD} \\ p\equiv D'\bar{m} \pmod Q}} \mathbf{1}_{pm<\frac{T}{D}} \mathbf{1}_{P_D}(mp)h(m)h(p)\Lambda(p)F\biggl(g_D\Bigl(\frac{pm-D'}{Q}\Bigr)\Gamma\biggr)\Biggr\}. \tag{9-13}$$

We now analyze the contribution from these four sums to (9-5) in turn, beginning with the two short sums up to T_0, which are both straightforward to bound. The main work goes into handling the large primes case corresponding to the long sum in (9-12). Here we will make use of the results from Sections 7 and 8. The long sum from (9-13) will, again, be straightforward to handle due to the above choice of the parameter X.

9E. *Short sums.* The following lemma provides straightforward bounds on the contribution of the short sums in (9-12) and (9-13) to (9-5).

Lemma 9.2. *Writing $\bar{f}_i(n) = |f_1 * \cdots * \widehat{f_i} * \cdots * f_H(n)|$, we have*

$$\sum_{\substack{D\leqslant T^{1-1/H} \\ (D,Q)=1}} \frac{\bar{f}_i(D)}{\log T}\Biggl|\frac{Q}{T}\sum_{\substack{m<T_0 \\ \gcd(m,Q)=1}} \sum_{\substack{p\leqslant T/(mD) \\ p\equiv D'\bar{m} \\ \pmod Q}} \mathbf{1}_{P_D}(mp)h(m)h(p)\Lambda(p)F\biggl(g_D\Bigl(\frac{pm-D'}{Q}\Bigr)\Gamma\biggr)\Biggr|$$
$$\ll (\log\log T)^2 \frac{1}{\log T}\frac{Q}{\phi(Q)}\exp\biggl(\frac{H-1}{H}\sum_{\substack{p\leqslant T \\ p\nmid Q}} \frac{|f(p)|}{p}\biggr) \tag{9-14}$$

and

$$\sum_{\substack{D\leqslant T^{1-1/H}\\(D,Q)=1}}\frac{\bar f_i(D)}{\log T}\bigg|\frac{Q}{T}\sum_{\substack{m>X\\ \gcd(m,Q)=1}}\sum_{\substack{p\leqslant\min(\frac{T}{Dm},T_0)\\ p\equiv D\bar m \pmod Q}}\mathbf{1}_{P_D}(mp)h(m)h(p)\Lambda(p)F\Big(g_D\Big(\frac{pm-D'}{Q}\Big)\Gamma\Big)\bigg|$$
$$\ll\frac{(\log\log T)^2}{\log T}\frac{1}{\log T}\frac{Q}{\phi(Q)}\prod_{\substack{p\leqslant T\\ p\nmid Q}}\Big(1+\frac{|f(p)|}{p}\Big).\quad (9\text{-}15)$$

Remark. Both these bounds are negligible when compared to (6-2). In the first case this follows from property (2) of Definition 1.3.

Proof. The short sum in (9-12) satisfies

$$\bigg|\frac{QD}{T}\sum_{\substack{m<T_0\\ \gcd(m,Q)=1}}\sum_{\substack{p\leqslant T/(mD)\\ p\equiv D\bar m \pmod Q}}\mathbf{1}_{P_D}(mp)h(m)h(p)\Lambda(p)F\Big(g_D\Big(\frac{pm-D'}{Q}\Big)\Gamma\Big)\bigg|$$
$$\ll\sum_{\substack{m<T_0\\ \gcd(m,Q)=1}}|h(m)|\frac{QD}{T}\sum_{\substack{p\leqslant T/(mD)\\ p\equiv D\bar m \pmod Q}}\Lambda(p)\ll\frac{Q}{\phi(Q)}\sum_{\substack{m<T_0\\ \gcd(m,Q)=1}}\frac{|h(m)|}{m}$$
$$\ll\frac{Q}{\phi(Q)}\frac{1}{\log T}\exp\Big(\sum_{w(N)<p<T_0}\frac{1}{p}\Big)\ll(\log\log T)^2\frac{Q}{\phi(Q)}.$$

Thus, the left-hand side of (9-14) is bounded by

$$(\log\log T)^2\frac{Q}{\phi(Q)}\frac{1}{\log T}\sum_{\substack{D\leqslant T^{1-1/H}\\(D,Q)=1}}\frac{\bar f_i(D)}{D}.$$

The claimed bound now follows since

$$\sum_{D\leqslant T^{1-1/H},(D,Q)=1}\frac{\bar f_i(D)}{D}\ll\exp\Big(\sum_{p\leqslant T,p\nmid Q}\frac{|f_1(p)+\cdots+f_H(p)-f_i(p)|}{p}\Big)=\exp\Big(\frac{H-1}{H}\sum_{p\leqslant T,p\nmid Q}\frac{|f(p)|}{p}\Big),$$

recalling the definition of the functions f_j from (1-6).

The short sum in (9-13) is bounded by

$$\bigg|\frac{QD}{T}\sum_{\substack{m>X\\ \gcd(m,Q)=1}}\sum_{\substack{p\leqslant\min(T/(mD),T_0)\\ p\equiv D\bar m \pmod Q}}\mathbf{1}_{P_D}(mp)h(m)h(p)\Lambda(p)F\Big(g_D\Big(\frac{pm-D'}{Q}\Big)\Gamma\Big)\bigg|$$
$$\ll\sum_{w(N)<p<T_0}\frac{\Lambda(p)}{p}\max_{A'\in(\mathbb{Z}/Q\mathbb{Z})^*}S_{|h|}\Big(\frac{T}{pD};Q,A'\Big)$$
$$\ll\sum_{w(N)<p<T_0}\frac{\Lambda(p)}{p}\frac{1}{\log T}\frac{Q}{\phi(Q)}\prod_{\substack{p\leqslant T\\ p\nmid Q}}\Big(1+\frac{|h(p)|}{p}\Big)\ll\frac{\log T_0}{\log T}\frac{Q}{\phi(Q)}\prod_{\substack{p\leqslant T\\ p\nmid Q}}\Big(1+\frac{|h(p)|}{p}\Big),$$

where we used (3-1). This shows that left-hand side of (9-15) is bounded by

$$\frac{\log T_0}{(\log T)^2}\frac{Q}{\phi(Q)}\sum_{\substack{D\leqslant T^{1-1/H}\\(D,Q)=1}}\frac{\bar f_i(D)}{D}\prod_{\substack{p\leqslant T\\p\nmid Q}}\Big(1+\frac{|h(p)|}{p}\Big).$$

Recall from Section 9D that $\log T_0=(\log\log T)^2$. To finish the proof of (9-15), recall also that $\bar f_i(p)=(H-1)f(p)/H$ and $h(p)=f(p)/H$, that $|f(p)|\leqslant H$, and that $|\bar f_i(p^k)|\leqslant (CH)^k$ for some positive constant C. Assuming that N is sufficiently large to ensure that $w(N)>2CH$, we then have

$$\begin{aligned}\sum_{\substack{D\leqslant T^{1-1/H}\\(D,Q)=1}}\frac{\bar f_i(D)}{D}\prod_{\substack{p\leqslant T\\p\nmid Q}}\Big(1+\frac{|h(p)|}{p}\Big)&\leqslant\prod_{\substack{w(N)<p\leqslant T\\p\nmid q}}\Big(1+\frac{|f(p)|}{Hp}\Big)\Big(1+\frac{(H-1)|f(p)|}{Hp}+\frac{(CH)^2}{p^2}\Big(1-\frac{CH}{p}\Big)^{-1}\Big)\\&\leqslant\exp\Big(\sum_{w(N)<p\leqslant T}\frac{2(CH)^2}{p^2}+\frac{H-1}{p^2}\Big)\prod_{\substack{w(N)<p\leqslant T\\p\nmid q}}\Big(1+\frac{|f(p)|}{p}\Big)\\&\ll\prod_{\substack{p\leqslant T\\p\nmid Q}}\Big(1+\frac{|f(p)|}{p}\Big),\end{aligned}$$

which completes the proof. □

9F. *Large primes.* In this subsection we finally apply the results from Section 8 to bound the contribution of the dyadic parts of (9-12) to (9-5). More precisely, we prove:

Lemma 9.3 (contribution from large primes). *Keep the assumptions of Proposition 6.4. Let Q be as in Proposition 6.4, recall the definition of P_D from (9-8), and let $E_h^\sharp(T,D,j)$ denote the expression*

$$\Big|\frac{DQ}{T}\sum_{\substack{m\sim 2^{-j}X\\\gcd(m,Q)=1}}\sum_{\substack{p\leqslant T/(mD)\\p\equiv D'\overline{m}\ (\mathrm{mod}\ Q)}}\mathbf{1}_{P_D}(mp)h(m)h(p)\Lambda(p)F\Big(g_D\Big(\frac{pm-D'}{Q}\Big)\Gamma\Big)\Big|.$$

Then, provided the parameter E_0 from Proposition 6.4 is sufficiently large depending on d, m_G and H, we have

$$\begin{aligned}\sum_{i=1}^{H}\sum_{k=\frac{(\log\log T)^2}{\log 2}}^{(1-\frac{1}{H})\frac{\log T}{\log 2}}\sum_{\substack{D\sim 2^k\\(D,Q)=1}}\mathbf{1}_{D\notin\mathscr{B}_{2^k}}\sum_{\substack{d_1,\dots,\widehat{d_i},\dots,d_H\\D_i=D}}\Big(\prod_{i'\neq i}\frac{|f_{i'}(d_{i'})|}{d_{i'}}\Big)\sum_{j=0}^{\log_2\frac{X}{T_0}}\frac{E_{f_i}^\sharp(T,D,j)}{\log T}&\\\ll\Big((\log\log T)^{-1/(2^{s+2}\dim G)}+\frac{\delta(N)^{-10^s\dim G}}{(\log\log T)^{1/2}}\Big)\frac{1+\|F\|_{\mathrm{Lip}}}{\log T}\frac{Q}{\phi(Q)}\prod_{\substack{p\leqslant T\\p\nmid Q}}\Big(1+\frac{|f(p)|}{p}\Big),&\qquad(9\text{-}16)\end{aligned}$$

where the implied constant may depend on d, m_G, α_f, E and H.

Remark. This contribution agrees with the bound (6-2).

The remainder of this subsection is concerned with the proof of (9-16). Considering $E_h^{\sharp}(T, D, j)$ for a fixed value of j, $1 \leqslant j \leqslant \log_2(X/T_0)$, the Cauchy–Schwarz inequality yields

$$\sum_{\substack{m \sim 2^{-j} X \\ \gcd(m,Q)=1}} \sum_{\substack{p \leqslant 2^j \frac{T}{XD} \\ p \equiv D\overline{m} \pmod Q \\ mp \in P_D}} \mathbf{1}_{mp \leqslant N} h(m) h(p) \Lambda(p) F\biggl(g_D\Bigl(\frac{pm - D'}{Q}\Bigr)\Gamma\biggr)$$

$$\leqslant \biggl(\sum_{p \leqslant 2^j \frac{T}{XD}} |h(p)|^2 \Lambda(p)\biggr)^{\frac{1}{2}} \Biggl(\frac{Q}{\phi(Q)} \sum_{A' \in (\mathbb{Z}/Q\mathbb{Z})^*} \sum_{\substack{m,m' \sim 2^{-j} X \\ m \equiv m' \equiv A' \pmod Q \\ \gcd(Q,m)=1}} h(m) h(m')$$

$$\times \frac{\phi(Q)}{Q} \sum_{\substack{p \leqslant \frac{T}{D \max(m,m')} \\ pA' \equiv D' \pmod Q \\ pm,\, pm' \in P_D}} \Lambda(p) F\biggl(g_D\Bigl(\frac{pm - D'}{Q}\Bigr)\Gamma\biggr) \overline{F\biggl(g_D\Bigl(\frac{pm' - D'}{Q}\Bigr)\Gamma\biggr)}\Biggr)^{\frac{1}{2}}. \tag{9-17}$$

The first factor is easily seen to equal $O(2^j T/(XD))$, since $h(p) \ll_H 1$ at primes. To estimate the second factor, we seek to employ the orthogonality of the "W-tricked von Mangoldt function" with nilsequences, combined with the fact that for most pairs (m, m') the product nilsequence that appears in the above expression is equidistributed (see Proposition 8.1). For this purpose, let us make the change of variables $p = Qn + D'_m$ in the inner sum of the second factor, where D'_m is such that $D'_m \equiv D'\overline{m} \pmod Q$. This yields

$$\frac{\phi(Q)}{Q} \sum_{\substack{p \leqslant T/(D\max(m,m')) \\ pA' \equiv D' \pmod Q \\ pm,\, pm' \in P_D}} \Lambda(p) F\biggl(g_D\Bigl(\frac{pm - D'}{Q}\Bigr)\Gamma\biggr) \overline{F\biggl(g_D\Bigl(\frac{pm' - D'}{Q}\Bigr)\Gamma\biggr)}$$

$$= \sum_{\substack{n \leqslant T/(QD\max(m,m')) \\ nm + \widetilde{D}_m,\, nm' + \widetilde{D}_{m'} \in I_D}} \frac{\phi(Q)}{Q} \Lambda(Qn + D'_m) F(g_D(nm + \widetilde{D}_m)\Gamma) \overline{F(g_D(nm' + \widetilde{D}_{m'})\Gamma)}, \tag{9-18}$$

for suitable values of $0 \leqslant \widetilde{D}_m < m$, $0 \leqslant \widetilde{D}_{m'} < m'$ and with $I_D = \{n : Dn + D'' \in I\}$ as defined in (9-8) and I as in the statement of Proposition 6.4. Let us consider the summation range

$$I_{m,m'} = \{n \in \mathbb{N} : nm + \widetilde{D}_m \in I_D,\ nm' + \widetilde{D}_{m'} \in I_D\}$$

in the above expression more closely. Since I is a discrete interval, I_D is a discrete interval too and, for $m, m' \sim 2^{-j} X$, we have

$$\#\{n \in \mathbb{N} : nm + \widetilde{D}_m \in I_D\} \ll |I_D| 2^j / X \ll |I| 2^j / (DX) \leqslant T 2^j / (DXQ)$$

and, similarly, $\#\{n \in \mathbb{N} : nm' + \widetilde{D}_{m'} \in I_D\} \ll T 2^j / (DXQ)$. We will now split the set

$$\{(m, m') : m, m' \sim 2^{-j} X, m \equiv m' \equiv A' \pmod Q)\}$$

into two subsets, one containing all pairs (m, m') for which $\# I_{m,m'} \leqslant \delta(N) 2^j T/(DXQ)$, and one containing those pairs for which

$$\# I_{m,m'} > \delta(N) 2^j T/(DXQ). \tag{9-19}$$

In the former case, the trivial bound of (9-18) asserts that

$$\Bigg| \sum_{\substack{n \leqslant T/QD \max(m,m') \\ n \in I_{m,m'}}} \frac{\phi(Q)}{Q} \Lambda(Qn + D'_m) F(g_D(nm + \widetilde{D}_m)\Gamma) \overline{F(g_D(nm' + \widetilde{D}_{m'})\Gamma)} \Bigg| \leqslant \frac{\delta(N) T 2^j}{DXQ}.$$

This leaves us to bound (9-18) in the case where (9-19) holds.

To start with, recall our assumption from the start of Section 9C that all values of D are unexceptional in the sense that $D \sim 2^k$ for some $k \geqslant (\log\log T)^2/\log 2$ and $D \notin \mathscr{B}_{2^k}$, where $\mathscr{B}_K$ was defined in Proposition 7.4. Thus, for any fixed unexceptional value of D, the finite sequence

$$(g_D(n)\Gamma)_{n \leqslant T/(Dq)}$$

is totally $\delta(N)^{c_1 E_0}$-equidistributed. Thus, applying Proposition 8.1 with $g = g_D$ and with $E_2 = c_1 E_0$, we obtain for every integer

$$K \in [T_0, X]$$

an exceptional set $\mathscr{E}_K$ of size

$$\# \mathscr{E}_K \ll \delta(T)^{O(c_1 c_2 E_0)} K^2 \tag{9-20}$$

such that for all pairs of integers $(m, m') \in (K, 2K]^2 \setminus \mathscr{E}_K$ the following estimate holds:

$$\Bigg| \sum_{\substack{n \leqslant T/(QD \max(m,m')) \\ nm + \widetilde{D}_m, nm' + \widetilde{D}_{m'} \in I_D}} F(g_D(nm + \widetilde{D}_m)\Gamma) \overline{F(g_D(nm' + \widetilde{D}_{m'})\Gamma)} \Bigg| < \frac{(1 + \|F\|_{\mathrm{Lip}}) \delta(N)^{c_1 c_2 E_0} T}{KQD}.$$

Before we continue with the analysis of (9-18), we prove a quick lemma that will allow us to handle the contribution of exceptional sets $\mathscr{E}_K$ in the proof of Lemma 9.3.

Lemma 9.4. *Suppose $j \leqslant \log_2(X/T_0)$ and let $\mathscr{E}_K$ be the exceptional set obtained from Proposition 8.1 when applied with $g = g_D$. Then, provided E_0 is sufficiently large, we have*

$$\frac{1}{\phi(Q)} \sum_{A' \in (\mathbb{Z}/Q\mathbb{Z})^*} \sum_{\substack{m,m' \sim 2^{-j} X \\ m \equiv m' \equiv A' \ (\mathrm{mod}\ Q)}} |h(m)h(m')| \mathbf{1}_{(m,m') \in \mathscr{E}_{D,2^{-j}X}} \ll \delta(N)^{O(c_1 c_2 E_0)} \Bigg(\frac{2^{-j} X}{\phi(Q)} \frac{1}{\log(2^{-j}X)} \prod_{\substack{p \leqslant 2^{-j}X \\ p \nmid Q}} \Big(1 + \frac{|h(p)|}{p}\Big) \Bigg)^2,$$

where c_1 and c_2 are the constants defined in Proposition 7.4 and Proposition 8.1, respectively.

Proof. In view of (9-20), Cauchy–Schwarz yields

$$\frac{1}{\phi(Q)}\sum_{A'\in(\mathbb{Z}/Q\mathbb{Z})^*}\sum_{\substack{m,m'\sim 2^{-j}X\\ m\equiv m'\equiv A' \ (\mathrm{mod}\ Q)}}|h(m)h(m')|\mathbf{1}_{(m,m')\in\mathscr{E}_{D,2^{-j}X}}$$
$$\ll \frac{2^{-j}X}{\phi(Q)}\delta(N)^{O(c_1c_2E_0)}\Bigg(\sum_{\substack{m,m'\sim 2^{-j}X\\ m\equiv m'\equiv A' \ (\mathrm{mod}\ Q)}}|h(m)|^2|h(m')|^2\Bigg)^{1/2} \ll \frac{2^{-j}X}{\phi(Q)}\delta(N)^{O(c_1c_2E_0)}\sum_{\substack{m\sim 2^{-j}X\\ m\equiv A' \ (\mathrm{mod}\ Q)}}|h(m)|^2.$$

Since $2^{-j}X \geqslant T_0 = \exp((\log\log T)^2) \gg Q^2$, we may apply Shiu's bound (3-1) and the trivial inequality $h(p)^2 \leqslant |h(p)|$ to obtain the upper bound

$$\ll \left(\frac{2^{-j}X}{\phi(Q)}\right)^2\delta(N)^{O(c_1c_2E_0)}\frac{1}{\log(2^{-j}X)}\prod_{\substack{p\leqslant 2^{-j}X\\ p\nmid Q}}\left(1+\frac{|h(p)|}{p}\right)$$
$$\ll \delta(N)^{O(c_1c_2E_0)}\log(2^{-j}X)\Bigg(\frac{2^{-j}X}{\phi(Q)}\frac{1}{\log(2^{-j}X)}\prod_{\substack{p\leqslant 2^{-j}X\\ p\nmid Q}}\left(1+\frac{|h(p)|}{p}\right)\Bigg)^2.$$

Recall that X was defined in Section 9D and satisfies $X \leqslant T \ll N$. Since furthermore $\delta(N) \leqslant (\log N)^{-1}$, any sufficiently large choice of E_0 guarantees that

$$\delta(N)^{O(c_1c_2E_0)}\log(2^{-j}X) \leqslant \delta(N)^{O(c_1c_2E_0)}$$

holds. This completes the proof. □

As a final tool for the proof of Lemma 9.3, we require an explicit bound on the correlation of the "W-tricked von Mangoldt function" with nilsequences. The following lemma provides such bounds in our specific setting. We include a proof building on that of Green and Tao [2010, Proposition 10.2] in the Appendix.

Lemma 9.5. *Let G/Γ be an s-step nilmanifold, let $G_\bullet$ be a filtration of G of degree d and let $\mathscr{X}$ be an M-rational Malcev basis adapted to it. Let $\Lambda' : \mathbb{N} \to \mathbb{R}$ be the restriction of the ordinary von Mangoldt function to primes, that is, $\Lambda'(p^k) = 0$ whenever $k > 1$. Let $W = W(x)$, let q' and b' be integers such that $0 < b' < Wq' \leqslant (\log x)^E$ and $\gcd(Wq', b') = 1$ hold. Let $\alpha \in (0, 1)$. Then, for every $y \in [\exp((\log x)^\alpha), x]$ and for every polynomial sequence $g \in \mathrm{poly}(\mathbb{Z}, G_\bullet)$, the following estimate holds*:

$$\left|\sum_{n\leqslant y}\frac{\phi(Wq')}{Wq'}\Lambda'(Wq'n+b')F(g(n)\Gamma)\right| \ll_{\alpha,d,\dim G,E,\|F\|_{\mathrm{Lip}}} \left|\sum_{n\leqslant y}F(g(n)\Gamma)\right| + y\mathscr{E}(x),$$

where

$$\mathscr{E}(x) := (\log\log x)^{-1/(2^{2d+3}\dim G)} + \frac{M^{O(10^d\dim G)}}{(\log\log x)^{1/2^{d+2}}}.$$

Employing Lemma 9.5 for the upper endpoint of an interval $[y_0, y_1]$, and either a trivial estimate or the lemma for the lower endpoint, say, depending on whether or not $y_0 \leqslant y_1^{1/2}$, we obtain as an immediate consequence that

$$\Big|\sum_{y_0 \leqslant n \leqslant y_1} \frac{\phi(Wq')}{Wq'} \Lambda'(Wq'n+b')F(g(n)\Gamma)\Big| \\ \ll_{\alpha,s,E,\|F\|_{\mathrm{Lip}}} \Big|\sum_{n\leqslant y_0} F(g(n)\Gamma)\Big| + \Big|\sum_{n\leqslant y_1} F(g(n)\Gamma)\Big| + y_1^{1/2} + y_1 \mathscr{E}(x) \quad (9\text{-}21)$$

for any $0 < y_0 < y_1 \leqslant x$ such that $y_1^{1/2} \geqslant \exp((\log x)^\alpha)$.

This brings us back to the task of bounding (9-18) under the assumption of (9-19). We shall start by applying (9-21) with $[y_0, y_1] = I_{m,m'}$ and $x = N = T^{1+o(1)}$. To do so, note that (9-19) implies that

$$\begin{aligned} \frac{1}{2}\log y_1 &\geqslant \log \frac{\delta(N) T 2^j}{DXQ} \\ &\geqslant \log \frac{T}{DX} + j\log 2 + \log \delta(N) - \log Q \\ &\geqslant \Big(\log \frac{T}{D}\Big)^{1/U} + j \log 2 - \log\log N - 2E \log\log T \\ &\geqslant \Big(\frac{\log T}{H}\Big)^{1/4} - \log\log N - 2E\log\log T \\ &\gg_{E,H} (\log T)^{1/4}, \end{aligned}$$

where we used the definition of X and the assumptions that Proposition 6.4 makes on δ. Thus, choosing $\alpha = \frac{1}{5}$, say, the conditions of Lemma 9.5 are satisfied for every T that is sufficiently large with respect to E and H. Hence, (9-21) yields the following estimate for the interval $[y_0, y_1] = I_{m,m'}$:

$$\Big|\sum_{\substack{n \leqslant T/(QD\max(m,m')) \\ n \in I_{m,m'}}} \frac{\phi(Q)}{Q} \Lambda(Qn + D'_m) F(g_D(nm+\widetilde{D}_m)\Gamma)\overline{F(g_D(nm'+\widetilde{D}_{m'})\Gamma)}\Big| \\ \ll_{s,E,H,\|F\|_{\mathrm{Lip}}} \Big|\sum_{n\leqslant y_0} F(g_D(nm+\widetilde{D}_m)\Gamma)\overline{F(g_D(nm'+\widetilde{D}_{m'})\Gamma)}\Big| \\ + \Big|\sum_{n\leqslant y_1} F(g_D(nm+\widetilde{D}_m)\Gamma)\overline{F(g_D(nm'+\widetilde{D}_{m'})\Gamma)}\Big| + \frac{T2^j}{DXQ}\mathscr{E}(T).$$

Proposition 8.1 shows that the right-hand side is small for most pairs (m, m'). Indeed, together with Proposition 8.1, the above implies that (9-17) is bounded above by

$$\ll_{s,E,H,\|F\|_{\mathrm{Lip}}} \Big(\frac{T2^j}{DX}\Big)^{1/2} \\ \times \Big(\frac{Q}{\phi(Q)} \sum_{A' \in (\mathbb{Z}/Q\mathbb{Z})^*} \sum_{\substack{m,m' \sim 2^{-j}X \\ m\equiv m' \equiv A' \ (\mathrm{mod}\ Q)}} |h(m)h(m')| \frac{T2^j}{QDX}\big(\delta(N)^{O(c_1c_2E_0)} + \mathbf{1}_{(m,m')\in \mathscr{E}_{D,2^{-j}X}} + \mathscr{E}(T)\big)\Big)^{1/2}.$$

Treating the part of this expression to which Lemma 9.4 applies separately and rewriting in the remaining part the sum over m, m' as a square, we obtain after collecting together all the normalization factors:

$$\ll_{s,E,H,\|F\|_{\mathrm{Lip}}} \frac{T}{QD}\bigg\{\max_{A'\in(\mathbb{Z}/Q\mathbb{Z})^*}\bigg(\frac{Q2^j}{X}\sum_{\substack{m\sim 2^{-j}X\\ m\equiv A'(Q)}}|h(m)|\bigg)^2(\delta(N)^{O(c_1c_2E_0)}+\mathscr{E}(T)) \\ +\delta(N)^{O(c_1c_2E_0)}\bigg(\frac{Q}{\phi(Q)}\frac{1}{\log(2^{-j}X)}\prod_{\substack{p\leqslant 2^{-j}X\\ p\nmid Q}}\bigg(1+\frac{|h(p)|}{p}\bigg)\bigg)^2\bigg\}^{\frac{1}{2}}$$

$$\ll_{s,E,H,\|F\|_{\mathrm{Lip}}} \frac{T}{QD}(\delta(N)^{O(c_1c_2E_0)}+\mathscr{E}(T))\bigg(\frac{Q}{\phi(Q)}\frac{1}{\log(2^{-j}X)}\prod_{\substack{p\leqslant 2^{-j}X\\ p\nmid Q}}\bigg(1+\frac{|h(p)|}{p}\bigg)\bigg),$$

where we applied Shiu's bound in the last step. Summing the above expression over $j\leqslant \log_2(X/T_0)$ and taking into account the factor $(\log T)^{-1}$, we deduce that the inner sum in (9-16) is bounded by

$$\ll_{s,E,H,\|F\|_{\mathrm{Lip}}} (\delta(N)^{O(c_1c_2E_0)}+\mathscr{E}(T)) \\ \times\bigg(\frac{1}{\log T}\frac{Q}{\phi(Q)}\prod_{\substack{p\leqslant T\\ p\nmid Q}}\bigg(1+\frac{|h(p)|}{p}\bigg)\bigg)\sum_{j=1}^{\log_2(X/T_0)}\frac{1}{\log(2^{-j}X)}\prod_{2^{-j}X<p'<T}\bigg(1-\frac{|h(p')|}{p'}\bigg)$$

Since $\delta(N)\leqslant(\log N)^{-1}$, choosing E_0 sufficiently large in terms of d and m_0 ensures that

$$\delta(N)^{O(c_1c_2E_0)}+\mathscr{E}(T)\ll(\log N)^{-1}+\mathscr{E}(T)\ll\mathscr{E}(T).$$

To complete the proof of Lemma 9.3, it thus remains to show that the inner sum over j in the expression above is $O_{\alpha_f}(1)$. To see this, observe that property (2) of Definition 1.3 yields

$$\prod_{X2^{-j}<p\leqslant T}\bigg(1-\frac{|h(p)|}{p}\bigg)\ll\bigg(\frac{\log(2^{-j}X)}{\log T}\bigg)^{\alpha_f/H}.$$

Thus,

$$\sum_{j=1}^{\log_2(X/T_0)}\frac{1}{\log(2^{-j}X)}\prod_{2^{-j}X<p'<T}\bigg(1-\frac{|h(p')|}{p'}\bigg)\ll\frac{1}{(\log T)^{\alpha_f/H}}\sum_{j=1}^{\log_2(X/T_0)}\frac{1}{(\log X-j\log 2)^{1-\alpha_f/H}} \\ \ll_{\alpha_f}\frac{(\log X)^{\alpha_f/H}}{(\log T)^{\alpha_f/H}}\ll_{\alpha_f}1,$$

as required.

9G. ***Small primes.*** To complete the proof of Proposition 6.4, it remains to bound the contribution of the dyadic parts of (9-13) to (9-5). This is achieved by the following lemma, which will be proved by a combination of Cauchy–Schwarz, Lemma 1.8 and the choice of the parameter X from Section 9D.

Lemma 9.6 (contribution from small primes). *Let $E_h^{\flat}(T,D,j)$ denote the expression*

$$\left|\frac{DQ}{T}\sum_{\substack{m>X\\ \gcd(m,Q)=1}}\ \sum_{\substack{p\sim 2^{-j}\frac{T}{XD}\\ p\equiv D'\overline{m}\ (\mathrm{mod}\ Q)}}\mathbf{1}_{pm<\frac{T}{D}}\mathbf{1}_{P_D}(mp)h(m)h(p)\Lambda(p)F\left(g_D\left(\frac{pm-D'}{Q}\right)\Gamma\right)\right|.$$

Then

$$\sum_{i=1}^{H}\ \sum_{k=(\log\log T)^2/\log 2}^{(1-1/H)\log T/\log 2}\ \sum_{\substack{D\sim 2^k\\ (D,Q)=1}}\mathbf{1}_{D\notin\mathscr{B}_{2k}}\sum_{\substack{d_1,\dots,\widehat{d_i},\dots,d_H\\ D_i=D}}\left(\prod_{j\neq i}\frac{|f_j(d_j)|}{d_j}\right)\sum_{j=0}^{\log_2(T/(XDT_0))}\frac{E_{f_i}^{\flat}(T,D,j)}{\log T}$$
$$\ll(\log T)^{-1/4}\frac{1}{\log T}\frac{\phi(Q)}{Q}\prod_{\substack{p\leqslant T\\ p\nmid Q}}\left(1+\frac{|f(p)|}{p}\right).$$

Proof. Applying Cauchy–Schwarz to the expression $E_h^{\flat}(T,D,j)$ for a fixed value of j satisfying $0\leqslant j\leqslant\log_2(T/(XDT_0))$, we obtain

$$\left|\frac{QD}{T}\sum_{\substack{m>X\\ (m,Q)=1}}\ \sum_{\substack{p\sim 2^{-j}\frac{T}{XD}\\ p\equiv D'\overline{m}\ (\mathrm{mod}\ Q)}}\mathbf{1}_{pm<\frac{T}{D}}h(m)h(p)\Lambda(p)F\left(g\left(\frac{pm-D'}{Q}\right)\Gamma\right)\mathbf{1}_{P_D}(mp)\right|$$
$$\leqslant\left(\frac{Q}{\phi(Q)}\frac{1}{2^jX}\sum_{\substack{X<m<2^jX\\ \gcd(m,Q)=1}}|h(m)|^2\right)^{\frac12}\left(\phi(Q)\left(\frac{2^jXD}{T}\right)^2\sum_{\substack{p,p'\sim 2^{-j}\frac{T}{XD}\\ p\equiv p'\ (\mathrm{mod}\ Q)}}h(p)h(p')\Lambda(p)\Lambda(p')\right.$$
$$\left.\times\frac{Q}{X2^j}\sum_{\substack{X<m<T/(D\max(p,p'))\\ mp\equiv D'\ (\mathrm{mod}\ Q)\\ pm\in I_D}}F\left(g\left(\frac{pm-D'}{Q}\right)\Gamma\right)\overline{F\left(g\left(\frac{p'm-D'}{Q}\right)\Gamma\right)}\right)^{\frac12}. \tag{9-22}$$

We estimate the second factor trivially as $O(1)$ by using the bounds $|h(p)h(p')|\ll 1$ and $\|F\|_\infty=\|\overline{F}\|_\infty\leqslant 1$. Thus, (9-22) is bounded by

$$\ll\left(\frac{Q}{\phi(Q)}\frac{1}{2^jX}\sum_{\substack{X<m<2^jX\\ \gcd(m,Q)=1}}|h(m)|^2\right)^{1/2}.$$

This expression can be handled as that in Lemma 1.8: Note that $X\leqslant 2^jX\leqslant T/(DT_0)$, where

$$X=\left(\frac{T}{D}\right)^{1-1/(\log\frac{T}{D})^{(U-1)/U}}\gg(T/D)^{1/2}$$

and

$$\frac{T}{DT_0}=\left(\frac{T}{D}\right)^{1-(\log\log\frac{T}{D})^2/\log\frac{T}{D}}.$$

Thus, Shiu's bound (3-1) and the trivial inequality $|h(p)|^2 \leqslant |h(p)|$ imply that

$$\biggl(\frac{Q}{\phi(Q)}\frac{1}{2^j X}\sum_{\substack{X<m<2^j X\\ \gcd(m,Q)=1}}|h(m)|^2\biggr)^{1/2} \ll \biggl(\frac{1}{\log T}\frac{Q}{\phi(Q)}\prod_{\substack{p\leqslant T\\ p\nmid Q}}\biggl(1+\frac{|h(p)|}{p}\biggr)\biggr)^{1/2}.$$

The right-hand side is bounded below by $(\log T)^{-1/2}$, thus the above is bounded by

$$\ll (\log T)^{1/2}\biggl(\frac{1}{\log T}\frac{Q}{\phi(Q)}\prod_{\substack{p\leqslant T\\ p\nmid Q}}\biggl(1+\frac{|h(p)|}{p}\biggr)\biggr).$$

Finally, note that the summation range in j is short: it is bounded by

$$\log_2(T/(XDT_0)) \ll (\log T)^{1/U} \ll (\log T)^{1/4}.$$

This shows that

$$\sum_{i=1}^{H}\sum_{k=1}^{(1-1/H)\log T/\log 2}\sum_{\substack{D\sim 2^k\\ (D,Q)=1}}\mathbf{1}_{D\notin\mathscr{B}_{2^k}}\sum_{\substack{d_1,\dots,\widehat{d_i},\dots,d_H\\ D_i=D}}\biggl(\prod_{j\neq i}\frac{|f_j(d_j)|}{d_j}\biggr)\sum_{j=0}^{\log_2(T/(XDT_0))}\frac{E^{\flat}_{f_i}(T,D,j)}{\log T}$$
$$\ll (\log T)^{-1+\frac{1}{2}+\frac{1}{4}}\sum_{i=1}^{H}\sum_{\substack{D\leqslant T^{1-1/H}\\ (D,Q)=1}}\sum_{\substack{d_1,\dots,\widehat{d_i},\dots,d_t\\ D_i=D}}\biggl(\prod_{j\neq i}\frac{|f_j(d_j)|}{d_j}\biggr)\frac{1}{\log T}\frac{Q}{\phi(Q)}\prod_{\substack{p\leqslant T\\ p\nmid Q}}\biggl(1+\frac{|h(p)|}{p}\biggr)$$
$$\ll (\log T)^{-\frac{1}{4}}\frac{1}{\log T}\frac{Q}{\phi(Q)}\prod_{\substack{p\leqslant T\\ p\nmid Q}}\biggl(1+\frac{|f(p)|}{p}\biggr).$$

This completes the proof of Lemma 9.6 as well as the proof of Proposition 6.4. □

Appendix: Explicit bounds on the correlation of Λ with nilsequences

The aim of this appendix is to provide a proof of Lemma 9.5. This result is due to Green and Tao and we expect that a statement like Lemma 9.5 will eventually appear in [Green 2014]. The author is grateful to Ben Green for very helpful discussions.

The proof of Lemma 9.5 rests upon the decomposition of Λ' that already appeared in the proof of the original result, [Green and Tao 2010, Proposition 10.2]. To be precise, let $\gamma \in (0,1)$ be a small positive real number that will later be chosen depending on the degree d of the given filtration $G_\bullet$. Further, let $\chi^\flat + \chi^\sharp = \mathrm{id}_{\mathbb{R}}$ be a smooth decomposition of the identity function $\mathrm{id}_{\mathbb{R}} : \mathbb{R}\to\mathbb{R}$, $\mathrm{id}_{\mathbb{R}}(t) := t$, that is such that $\mathrm{supp}(\chi^\sharp)\subset(-1,1)$ and $\mathrm{supp}(\chi^\flat)\subset\mathbb{R}\setminus\bigl[-\frac{1}{2},\frac{1}{2}\bigr]$. This decomposition of $\mathrm{id}_{\mathbb{R}}$ induces the following decomposition of Λ':

$$\frac{\phi(Wq')}{Wq'}\Lambda'(Wq'n+b')-1 = \frac{\phi(Wq')}{Wq'}\Lambda^\flat(Wq'n+b') + \Bigl(\frac{\phi(Wq')}{Wq'}\Lambda^\sharp(Wq'n+b')-1\Bigr),$$

where (cf. [Green and Tao 2010, (12.2)])

$$\Lambda^\sharp(n) = -\log x^\gamma \sum_{d|n} \mu(d)\chi^\sharp\left(\frac{\log d}{\log x^\gamma}\right) \quad \left(|t| \geqslant 1 \Rightarrow \chi^\sharp(t) = 0\right)$$

is a truncated divisor sum, and where

$$\Lambda^\flat(n) = -\log x^\gamma \sum_{d|n} \mu(d)\chi^\flat\left(\frac{\log d}{\log x^\gamma}\right) \quad \left(|t| \leqslant \tfrac{1}{2} \Rightarrow \chi^\flat(t) = 0\right)$$

is an average of $\mu(d)$ running over large divisors of n. This decomposition in turn splits the correlation from Lemma 9.5 into two correlations that shall be bounded separately.

The correlation estimate of the $\Lambda^\flat$ term with nilsequences follows as in [Green and Tao 2010, §12] from the noncorrelation of Möbius with nilsequences and inherits an error term which saves a factor $O_A(\log x)^{-A}$ for any given $A \geqslant 1$ when compared to the trivial bound. In [Green and Tao 2010, Conjecture 8.5], it was conjectured that the Möbius function is orthogonal to linear nilsequences. Since [Green and Tao 2012a, Theorem 1.1] proves this conjecture, not just for linear, but for polynomial nilsequences, it follows without any essential changes in the proof, that the correlation estimate [Green and Tao 2010, equation (12.10)] continues to hold for polynomial sequences. That is to say, we have an estimate of the form

$$\left|\sum_{n\leqslant N} \Lambda^\flat(n)F(g(n)\Gamma)\right| \ll_{\|F\|_{\mathrm{Lip}},G/\Gamma,s,A} N(\log N)^{-A}. \tag{A-1}$$

In our setting, we may express the congruence condition modulo Wq' as a character sum

$$\frac{\phi(Wq')}{Wq'}\Lambda^\flat(n)\mathbf{1}_{n\equiv b' \pmod{Wq'}} = \mathbb{E}_{\chi \pmod{Wq'}} \frac{\phi(Wq')}{Wq'}\Lambda^\flat(n)\chi(n)\overline{\chi}(b').$$

As with equation (12.8) of [Green and Tao 2010], the factor $F(g(n)\Gamma)$ from the statement of Lemma 9.5 may be reinterpreted as $F(g'(Wq'n+b')\Gamma)$ for a new polynomial sequence g'. Reinterpreting the product $\chi(n)F(g'(n)\Gamma)$ of a character χ with the given nilsequence as a nilsequence itself allows us to employ the correlation estimate (A-1) with N given by $xq'W \ll x(\log x)^E$ to handle the correlation for $\Lambda^\flat$. Thanks to the saving of an arbitrary power of $\log x$ in (A-1), we can compensate the factor of Wq', which is bounded above by $(\log x)^E$, that we lose when passing to the character sums. In total, we obtain

$$\frac{1}{y}\sum_{n\leqslant y}\frac{\phi(Wq')}{Wq'}\Lambda^\flat(Wq'n+b')F(g(n)\Gamma) \ll_{\|F\|_{\mathrm{Lip}},s,G/\Gamma,B} (\log y)^{-B} \ll_{\|F\|_{\mathrm{Lip}},s,G/\Gamma,B'} (\log x)^{-B'}.$$

It remains to analyze the contribution of the function $\lambda^\sharp : \mathbb{N} \to \mathbb{R}$, defined via

$$\lambda^\sharp(n) := \frac{\phi(Wq')}{Wq'}\Lambda^\sharp(Wq'n+b') - 1.$$

This contribution satisfies the general bound

$$\left|\frac{1}{y}\sum_{n\leqslant y}\left(\frac{\phi(Wq')}{Wq'}\Lambda^{\sharp}(Wq'n+b')-1\right)F(g(n)\Gamma)\right| \leqslant \|\lambda^{\sharp}\|_{U^{k+1}[y]}\|F(g(\cdot)\Gamma)\|_{U^{k+1}[y]^*}$$

for every $k\geqslant 1$, where the dual uniformity norm is defined via

$$\|F(g(\,\cdot\,)\Gamma)\|_{U^{k+1}[N]^*} := \sup\left\{\left|\frac{1}{N}\sum_{n\leqslant N} f(n)F(g(n)\Gamma)\right| : \|f\|_{U^k[N]}\leqslant 1\right\}.$$

The main task that remains is to obtain control on the above dual uniformity norm for at least one value of k. In [Green and Tao 2010], this is achieved through their Proposition 11.2, which decomposes a general nilsequence into an averaged nilsequence of bounded dual uniformity norm plus an error term that is small in the L^∞ norm. The proof of this decomposition uses a compactness argument and, as such, does not provide explicit error terms. Central ideas for a new approach not working with compactness were indirectly provided by work of Eisner and Zorin-Kranich [2013] on a different question. They replace in their work the Lipschitz function in the definition of a nilsequence by a smooth function and the Lipschitz norm by a Sobolev norm. Moreover, they show that certain constructions that play a central role in [Green and Tao 2012b] have counterparts in the Sobolev norm setting. Building on these observations, Green [2014] proves that in the Sobolev norm setting the dual U^{s+1} norm of an s-step nilsequence *is* in fact bounded. The statement of the latter result involves the following notion of Sobolev norms.

Definition A.1 [Green 2014]. Let G/Γ be an m-dimension nilmanifold together with a Malcev basis $\mathscr{X}=\{X_1,\dots,X_m\}$. For any $\psi\in C^\infty(G/\Gamma)$, set

$$\|\psi\|_{W^m,\mathscr{X}} = \sup_{m'\leqslant m}\ \sup_{1\leqslant i_1,\dots,i_{m'}\leqslant m}\|D_{X_{i_1}}\cdots D_{X_{i_{m'}}}\psi\|_\infty,$$

where $D_X\psi(g\Gamma)=\lim_{t\to 0}(\mathrm{d}/\mathrm{d}t)\psi(\exp(tX)g\Gamma)$.

Lemma A.2 [Green 2014, Theorem 5.3.1]. *Let G/Γ be a k-step nilmanifold together with a filtration $G_\bullet$ of degree $d\geqslant k$ and an M-rational Malcev basis adapted to it. Let $g\in\mathrm{poly}(\mathbb{Z},G_\bullet)$ and suppose $\tilde{F}\in C^\infty(G/\Gamma)$. Then*

$$\|\tilde{F}(g(\cdot)\Gamma)\|_{U^{d+1}[N]^*} := \sup\left\{\left|\frac{1}{N}\sum_{n\leqslant N} f(n)\tilde{F}(g(n)\Gamma)\right| : \|f\|_{U^{d+1}[N]}\leqslant 1\right\} \ll M^{10^d\dim G}\|\tilde{F}\|_{W^{2^d\dim G},\mathscr{X}}.$$

In order to apply Lemma A.2 in our situation, an auxiliary result is needed that allows one to pass from the Lipschitz setting to the Sobolev setting, i.e., to write any Lipschitz function on G/Γ as the sum of a smooth function, to which Lemma A.2 can be applied, and a small L^∞ error. This is the content of the following lemma which will be proved using a standard smoothing trick; the author thanks Ben Green for pointing out this approach.

Lemma A.3. *Suppose that $F: G/\Gamma\to\mathbb{C}$ is a Lipschitz function and let m be a positive integer. Then there is a constant $c\in(0,1)$, only depending on G, such that for every $\varepsilon\in(0,c)$ there exists a function*

$\psi_m \in C^\infty(G/\Gamma)$ *such that*

$$\|F - F * \psi_m\|_\infty \leqslant \varepsilon(1 + \|F\|_{\mathrm{Lip}}) \tag{A-2}$$

and

$$\|F * \psi_m\|_{W^m,\mathscr{X}} \ll (m/\varepsilon)^{2m} M^{O(m)}. \tag{A-3}$$

Taking Lemma A.3 on trust for the moment, we first complete the proof of Lemma 9.5 before providing that of Lemma A.3. Recall that the filtration $G_\bullet$ of the nilmanifold G/Γ from Lemma 9.5 is of degree d. The previous two lemmas allow us to reduce the proof of Lemma 9.5 to a bound on the U^{d+1}-norm of $\lambda^\sharp : \mathbb{N} \to \mathbb{R}$. More precisely, we have

$$\begin{aligned}\frac{1}{y}\sum_{n\leqslant y}&\Big(\frac{\phi(Wq')}{Wq'}\Lambda^\sharp(Wq'n+b')-1\Big)F(g(n)\Gamma)\\&\ll \varepsilon(1+\|F\|_{\mathrm{Lip}}) + \frac{1}{y}\sum_{n\leqslant y}\Big(\frac{\phi(Wq')}{Wq'}\Lambda^\sharp(Wq'n+b')-1\Big)(F*\psi_m)(g(n)\Gamma)\\&\ll \varepsilon(1+\|F\|_{\mathrm{Lip}}) + \|\lambda^\sharp\|_{U^{d+1}[y]}\|(F*\psi_m)(g(\cdot)\Gamma)\|_{U^{d+1}[y]^*}.\end{aligned} \tag{A-4}$$

Since $\Lambda^\sharp$ is a truncated divisor sum, one can analyze its U^{d+1}-norm with the help of Theorem D.3 in Appendix D of [Green and Tao 2010]. We will follow the final section of that appendix (“The correlation estimate for $\Lambda^\sharp$”) of closely.

For each nonempty subset $\mathscr{B} \subset \{0, 1\}^{d+1}$, let

$$\Psi_{\mathscr{B}}(n, \boldsymbol{h}) = (Wq'(n + \boldsymbol{\omega} \cdot \boldsymbol{h}) + b')_{\omega\in\mathscr{B}}\,, \quad (n, \boldsymbol{h}) \in \mathbb{Z} \times \mathbb{Z}^{d+1},$$

denote the relevant system of forms. The set of exceptional primes for this system, denoted by $\mathscr{P}_{\Psi_{\mathscr{B}}}$, is defined to be the set of all primes p such that the reduction modulo p of $\mathscr{P}_{\Psi_{\mathscr{B}}}$ contains two linearly dependent forms or a form that degenerates to a constant. It is clear that whenever x is sufficiently large, the set $\mathscr{P}_{\Psi_{\mathscr{B}}}$ consists of all prime factors of $W(x)q'$ and, in particular, it contains all primes up to $w(x)$. For each prime p, the local factor $\beta_p^{(\mathscr{B})}$ corresponding to $\Psi_{\mathscr{B}}$ is defined to be

$$\beta_p^{(\mathscr{B})} = \frac{1}{p^{d+2}} \sum_{(n,\boldsymbol{h})\in(\mathbb{Z}/p\mathbb{Z})^{d+2}} \prod_{\omega\in\mathscr{B}} \frac{p}{\phi(p)} \mathbf{1}_{p\nmid Wq'(n+\boldsymbol{\omega}\cdot\boldsymbol{h})+b'}.$$

By [Green and Tao 2010, Lemma 1.3], we have $\beta_p^{(\mathscr{B})} = 1 + O_d(1/p^2)$ for all $p \notin \mathscr{P}_{\Psi_{\mathscr{B}}}$, and hence

$$\prod_{p\notin\mathscr{P}_{\Psi_{\mathscr{B}}}} \beta_p^{(\mathscr{B})} = 1 + O_d\Big(\frac{1}{w(x)}\Big) = 1 + O_d\Big(\frac{1}{\log\log x}\Big),$$

while the product of exceptional local factors satisfies

$$\prod_{p\in\mathscr{P}_{\Psi}} \beta_p^{(\mathscr{B})} = \prod_{p\mid W(x)q'} \beta_p^{(\mathscr{B})} = \Big(\frac{W(x)q'}{\phi(W(x))q'}\Big)^{|\mathscr{B}|},$$

since $\gcd(W(x)q', b') = 1$.

Let K_y be a convex body that is contained in the hypercube $[-y, y]^{d+2}$. Then, Theorem D.3 of [Green and Tao 2010], applied with $a_i = 1$ and $\chi_i = \chi_\#$, implies that if $\gamma > 0$ is sufficiently small depending on d, then

$$\frac{1}{y^{d+2}} \sum_{(n,\boldsymbol{h})\in K_y} \prod_{\omega\in\mathscr{B}} \Lambda^\sharp(Wq'(n+\boldsymbol{\omega}\cdot\boldsymbol{h})+b') = \frac{\mathrm{vol}(K_y)}{y^{d+2}} \prod_p \beta_p^{(\mathscr{B})} + O_d\Bigl((\log y^\gamma)^{-1/20} \exp\Bigl(O_d\Bigl(\sum_{p\in P_{\Psi_{\mathscr{B}}}} p^{-1/2}\Bigr)\Bigr)\Bigr).$$

Since $Wq' \leqslant (\log x)^E$, we have $|\mathscr{P}_{\Psi_{\mathscr{B}}}| \ll w(x)/\log x + E\log\log x/\log w(x)$. Recall that $w(x) \leqslant \log\log x$ and that $\log y \in [(\log x)^\alpha, \log x]$. Thus,

$$\begin{aligned}(\log y^\gamma)^{-1/20} \exp\Bigl(O_d\Bigl(\sum_{p\in\mathscr{P}_{\Psi_{\mathscr{B}}}} p^{-1/2}\Bigr)\Bigr) &\ll (\gamma(\log x)^\alpha)^{-1/20}\exp(O_d(|P_{\Psi_{\mathscr{B}}}|)) \\ &\ll_d (\log x)^{-\alpha/20}(\log x)^{O_d(E)/\log w(x)},\end{aligned}$$

which is $o(1)$ as $x \to \infty$.

Choosing $K_y = \{(n,\boldsymbol{h}) : 0 < n + \boldsymbol{\omega}\cdot\boldsymbol{h} \leqslant y \text{ for all } \boldsymbol{\omega} \in \{0,1\}^{d+1}\}$, we obtain

$$\begin{aligned}\|\lambda^\sharp\|^{2^{d+1}}_{U^{d+1}[y]} &= \frac{\mathrm{vol}(K_y)}{y^{d+2}} \sum_{\mathscr{B}\subseteq\{0,1\}^{d+1}} (-1)^{|\mathscr{B}|} \prod_{p\notin P_{\Psi_{\mathscr{B}}}} \beta_p^{(\mathscr{B})} + O_d\bigl((\log x)^{-\alpha/20+O_d(E)/\log w(x)}\bigr) \\ &\ll_d \frac{\mathrm{vol}(K_y)}{y^{d+2}} \frac{1}{\log\log x} + (\log x)^{-\alpha/20+O(E)/\log w(x)} \ll_{d,\alpha,E} \frac{1}{\log\log x}.\end{aligned}$$

Returning to (A-4), it follows from the above bound, Lemma A.2 and an application of Lemma A.3 with $m = 2^d \dim G$ and $\varepsilon = (\log\log x)^{-1/(m2^{d+3})}$, that for $\exp((\log x)^\alpha) \leqslant y \leqslant x$

$$\begin{aligned}&\frac{1}{y}\sum_{n\leqslant y}\Bigl(\frac{\phi(Wq')}{Wq'}\Lambda'(Wq'n+b')-1\Bigr)F(g(n)\Gamma) \\ &\ll_{d,\alpha,E} \frac{1+\|F\|_{\mathrm{Lip}}}{(\log\log x)^{1/(2^{2d+3}\dim G)}} + \|\lambda_\sharp\|_{U^{d+1}[y]}\|(F*\psi_m)(g(\cdot)\Gamma)\|_{U^{d+1}[y]^*} \\ &\ll_{d,\alpha,E} \frac{1+\|F\|_{\mathrm{Lip}}}{(\log\log x)^{1/(2^{2d+3}\dim G)}} + \frac{M^{10^d\dim G}\|F*\psi_m\|_{W^{2^d\dim G},\mathscr{X}}}{(\log\log x)^{1/2^{d+1}}} \\ &\ll_{d,\dim G,\alpha,E} \frac{1+\|F\|_{\mathrm{Lip}}}{(\log\log x)^{1/(2^{2d+3}\dim G)}} + \frac{(\log\log x)^{1/2^{d+2}}M^{O(10^d\dim G)}}{(\log\log x)^{1/2^{d+1}}},\end{aligned}$$

which reduces the proof of Lemma 9.5 to that of Lemma A.3.

Proof of Lemma A.3. Let $d_{\mathscr{X}}$ denote the metric on G/Γ that was introduced in [Green and Tao 2012b, Definition 2.2] and define for every $\varepsilon' > 0$ the following ε'-neighborhood

$$\mathscr{B}_{\varepsilon'} = \{x \in G/\Gamma : d_{\mathscr{X}}(x, \mathrm{id}_G\,\Gamma) < \varepsilon'\}.$$

Let $\varepsilon \in (0,1)$. Since F is Lipschitz, we have $|F(x) - F(y)| \leqslant \varepsilon(1 + \|F\|_{\mathrm{Lip}})$ whenever both x and y belong to the neighborhood $\mathscr{B}_\varepsilon$ of $\mathrm{id}_G \Gamma$. To ensure that (A-2) holds, it thus suffices to ensure that ψ_m is nonnegative, supported in $\mathscr{B}_\varepsilon$ and that $\int_{G/\Gamma} \psi_m = 1$. Indeed, these assumptions imply that

$$\begin{aligned}\left|F(x) - \int_{G/\Gamma} F(y)\psi_m(x-y)\,\mathrm{d}y\right| &= \left|\int_{G/\Gamma} (F(y) - F(x))\psi_m(x-y)\,\mathrm{d}y\right| \\ &\leqslant \varepsilon(1 + \|F\|_{\mathrm{Lip}}) \int_{G/\Gamma} \psi_m(x-y)\,\mathrm{d}y \\ &= \varepsilon(1 + \|F\|_{\mathrm{Lip}}).\end{aligned}$$

The function ψ_m will be constructed as the m-fold convolution of a smooth bump function. For this purpose, observe that

$$m\mathscr{B}_{\varepsilon/m} \subseteq \mathscr{B}_\varepsilon.$$

If $g = \exp(s_1 X_1) \cdots \exp(s_{\dim G} X_{\dim G})$, then the (unique) coordinates

$$\psi(g) := (s_1, \ldots, s_{\dim G})$$

are called Malcev coordinates, while the unique coordinates

$$\psi_{\exp}(g) := (t_1, \ldots, t_{\dim G})$$

for which $g = \exp(t_1 X_1 + \cdots + t_{\dim G} X_{\dim G})$ are called exponential coordinates. Proceeding as in the proof of Lemma A.14 in [Green and Tao 2012b], one can identify G/Γ with the fundamental domain $\left\{g \in G : \psi(g) \in \left[-\frac{1}{2}, \frac{1}{2}\right)\right\} \subset G$. Furthermore, their Lemma A.2 shows that the change of coordinates between exponential and Malcev coordinates, i.e., $\psi \circ \psi_{\exp}^{-1}$ or $\psi_{\exp} \circ \psi^{-1}$, is in either direction a polynomial mapping with $M^{O(1)}$-rational coefficients. Thus, $\mathscr{B}_\varepsilon$ lies within the fundamental domain provided $\varepsilon < c_0$ for some sufficiently small constant c_0. This embedding of $\mathscr{B}_\varepsilon$ in G allows us to define log on $\mathscr{B}_\varepsilon$. Let us equip $\mathfrak{g}$ with the maximum norm associated to $\mathscr{X}$, that is $\|X\| := \max_i |t_i|$ for $X = \sum_i t_i X_i$. Then the definition of $d_{\mathscr{X}}$ and Green and Tao's Lemma A.2 imply that

$$\{X \in \mathfrak{g} : \|X\| < \delta\} \subseteq \log \mathscr{B}_{\varepsilon/m}$$

for some δ of the form $\delta = (\varepsilon/m)M^{-O(1)}$. Following the above preparation, we now choose a nonnegative smooth function $\chi_1 : \mathbb{R}^{\dim G} \to \mathbb{R}_{\geqslant 0}$ with support in $\{\boldsymbol{t} \in \mathbb{R}^{\dim G} : \|\boldsymbol{t}\|_\infty < 1\}$ that satisfies $\int_{\mathbb{R}^{\dim G}} \chi_1(\boldsymbol{t})\,\mathrm{d}\boldsymbol{t} = 1$. Then, by setting $\chi(\boldsymbol{t}) = \delta \cdot \chi_1(\delta \boldsymbol{t})$, we obtain a function $\chi : \mathbb{R}^{\dim G} \to \mathbb{R}_{\geqslant 0}$ that is supported on $\{\boldsymbol{t} \in \mathbb{R}^{\dim G} : \|\boldsymbol{t}\|_\infty < \delta\}$, satisfies $\int_{\mathbb{R}^{\dim G}} \chi(\boldsymbol{t})\,\mathrm{d}\boldsymbol{t} = 1$ and has furthermore the property that

$$\left\|\frac{\partial}{\partial t_i}\chi(t_1, \ldots, t_{\dim G})\right\|_\infty \ll (m/\varepsilon)^2 M^{O(1)} \tag{A-5}$$

for $1 \leqslant i \leqslant \dim G$. We may identify χ with a function defined on the vector space $\mathfrak{g}$ equipped with the basis $\{X_1, \ldots, X_{\dim G}\}$, by setting $\chi(t_1 X_1 + \cdots + t_{\dim G} X_{\dim G}) = \chi(t_1, \ldots, t_{\dim G})$.

To obtain a smooth bump function on G/Γ, we consider the composition $\chi \circ \log : G/\Gamma \to \mathbb{R}$, which is supported in $\mathscr{B}_{\varepsilon/m}$. Since the differential $\mathrm{d}\log_{\mathrm{id}_G} : \mathfrak{g} \to \mathfrak{g}$ is the identity, there are positive constants C_0, C_1 and c_1, such that

$$C_0 \leqslant \int_{G/\Gamma} \chi \circ \log \leqslant C_1,$$

provided $\varepsilon < c_1$. Hence there is a constant C such that $\int_{G/\Gamma} \psi = 1$ for $\psi = C\chi \circ \log$.

With this function ψ at hand, let $\psi_m = \psi^{*m}$ be the m-th convolution power of ψ. It is clear that for every $0 < k \leqslant m$, the function ψ^{*k} is supported in $\mathscr{B}_\varepsilon$ and that $\int_{G/\Gamma} \psi^{*k} = 1$. Setting $\psi^{*0} = \delta_0$, where δ_0 denotes the Kronecker δ-function with weight 1 at 0, we furthermore have

$$D_{X_{i_1}} \cdots D_{X_{i_k}} (F * \psi_m) = F * D_{X_{i_1}} \psi * \cdots * D_{X_{i_k}} \psi * \psi^{*(m-k)}$$

and, hence,

$$\| D_{X_{i_1}} \cdots D_{X_{i_k}} (F * \psi_m) \|_\infty \leqslant \|F\|_\infty \cdot \|D_{X_{i_1}} (C\chi \circ \log)\|_\infty \cdots \|D_{X_{i_k}} (C\chi \circ \log)\|_\infty$$

for any $k \leqslant m$. Our final task is to bound $\|D_{X_j}(C\chi \circ \log)\|_\infty$ for every $j \leqslant \dim G$. Writing $[\cdot]_i : \mathfrak{g} \to \mathbb{R}$ for the i-th co-ordinate map with respect to the basis $\mathscr{X}$, we have

$$D_{X_j}(\chi \circ \log)(g) = \sum_{i=1}^{\dim G} \frac{\partial \chi}{\partial X_i}(\log g) \cdot \lim_{t \to 0} \big[\log(\exp(t X_j) g)\big]_i. \tag{A-6}$$

Since the differential $\mathrm{d}\log_{\mathrm{id}_G} : \mathfrak{g} \to \mathfrak{g}$ is the identity, there are constants $C_2 > 0$ and $c_2 > 0$, such that for every $g \in \mathscr{B}_{c_2}$ and for $1 \leqslant i \leqslant m$, the derivative

$$\Big| \lim_{t \to 0} [\log(\exp(t X_j) g)]_i \Big|$$

is bounded by C_2. Choosing $c < \min(c_0, c_1, c_2)$, the bound (A-3) now follows from (A-6) and the bounds given in (A-5). □

Acknowledgements

I would like to thank Hedi Daboussi, Régis de la Bretèche, Nikos Frantzikinakis, Andrew Granville, Ben Green, Adam Harper, Dimitris Koukoulopoulos and Terence Tao for very helpful discussions and Nikos Frantzikinakis for valuable comments on a previous version of this paper. I am very grateful to the anonymous referee for extremely helpful and detailed comments, which led in particular to the development of Section 4.

References

[Balog et al. 2013] A. Balog, A. Granville, and K. Soundararajan, "Multiplicative functions in arithmetic progressions", *Ann. Math. Qué.* **37**:1 (2013), 3–30. MR Zbl

[Barnet-Lamb et al. 2011] T. Barnet-Lamb, D. Geraghty, M. Harris, and R. Taylor, "A family of Calabi–Yau varieties and potential automorphy II", *Publ. Res. Inst. Math. Sci.* **47**:1 (2011), 29–98. MR Zbl

[Bourgain et al. 2013] J. Bourgain, P. Sarnak, and T. Ziegler, "Disjointness of Moebius from horocycle flows", pp. 67–83 in *From Fourier analysis and number theory to Radon transforms and geometry*, edited by F. H. et al., Dev. Math. **28**, Springer, 2013. MR Zbl

[Browning and Matthiesen 2017] T. D. Browning and L. Matthiesen, "Norm forms for arbitrary number fields as products of linear polynomials", *Ann. Sci. Éc. Norm. Supér.* (4) **50**:6 (2017), 1383–1446. MR Zbl

[Daboussi and Delange 1974] H. Daboussi and H. Delange, "Quelques propriétés des fonctions multiplicatives de module au plus égal à 1", *C. R. Acad. Sci. Paris Sér. A* **278** (1974), 657–660. MR Zbl

[Daboussi and Delange 1982] H. Daboussi and H. Delange, "On multiplicative arithmetical functions whose modulus does not exceed one", *J. London Math. Soc.* (2) **26**:2 (1982), 245–264. MR Zbl

[Eisner and Zorin-Kranich 2013] T. Eisner and P. Zorin-Kranich, "Uniformity in the Wiener–Wintner theorem for nilsequences", *Discrete Contin. Dyn. Syst.* **33**:8 (2013), 3497–3516. MR Zbl

[Elliott 2017] P. D. T. A. Elliott, "Multiplicative function mean values: asymptotic estimates", *Funct. Approx. Comment. Math.* **56**:2 (2017), 217–238. MR

[Elliott and Kish 2016] P. D. T. A. Elliott and J. Kish, "Harmonic analysis on the positive rationals II: Multiplicative functions and Maass forms", *J. Math. Sci. Univ. Tokyo* **23**:3 (2016), 615–658. MR Zbl

[Frantzikinakis and Host 2017] N. Frantzikinakis and B. Host, "Higher order Fourier analysis of multiplicative functions and applications", *J. Amer. Math. Soc.* **30**:1 (2017), 67–157. MR Zbl

[Granville 2009] A. Granville, "The pretentious large sieve", unpublished notes, Université de Montréal, 2009, available at www.dms.umontreal.ca/~andrew/PDF/PretentiousLargeSieve.pdf.

[Granville and Soundararajan 2003] A. Granville and K. Soundararajan, "Decay of mean values of multiplicative functions", *Canad. J. Math.* **55**:6 (2003), 1191–1230. MR Zbl

[Granville and Soundararajan $\geq$ 2018] A. Granville and K. Soundararajan, *Multiplicative number theory*, Work in progress.

[Granville et al. 2017] A. Granville, A. J. Harper, and K. Soundararajan, "A new proof of Halász's theorem, and its consequences", 2017. arXiv

[Green 2014] B. Green, "Hadamard Lectures on Nilsequences", unpublished notes, IHES, 2014.

[Green and Tao 2008] B. Green and T. Tao, "The primes contain arbitrarily long arithmetic progressions", *Ann. of Math.* (2) **167**:2 (2008), 481–547. MR Zbl

[Green and Tao 2010] B. Green and T. Tao, "Linear equations in primes", *Ann. of Math.* (2) **171**:3 (2010), 1753–1850. MR Zbl

[Green and Tao 2012a] B. Green and T. Tao, "The Möbius function is strongly orthogonal to nilsequences", *Ann. of Math.* (2) **175**:2 (2012), 541–566. MR Zbl

[Green and Tao 2012b] B. Green and T. Tao, "The quantitative behaviour of polynomial orbits on nilmanifolds", *Ann. of Math.* (2) **175**:2 (2012), 465–540. MR Zbl

[Green et al. 2012] B. Green, T. Tao, and T. Ziegler, "An inverse theorem for the Gowers $U^{s+1}[N]$-norm", *Ann. of Math.* (2) **176**:2 (2012), 1231–1372. MR Zbl

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

[Hardy and Ramanujan 1918] G. H. Hardy and S. Ramanujan, "Asymptotic Formulaae in Combinatory Analysis", *Proc. London Math. Soc.* (2) **17** (1918), 75–115. MR

[Harper 2011] A. Harper, "A different proof of a finite version of Vinogradov's bilinear sums", unpublished notes, 2011, available at https://www.dpmms.cam.ac.uk/~ajh228/FiniteBilinearNotes.pdf.

[Hua 1965] L. K. Hua, *Additive theory of prime numbers*, Translations of Mathematical Monographs **13**, American Mathematical Society, Providence, RI, 1965. MR Zbl

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

[Kátai 1986] I. Kátai, "A remark on a theorem of H. Daboussi", *Acta Math. Hungar.* **47**:1-2 (1986), 223–225. MR Zbl

[Matthiesen 2012] L. Matthiesen, "Linear correlations amongst numbers represented by positive definite binary quadratic forms", *Acta Arith.* **154**:3 (2012), 235–306. MR Zbl

[Matthiesen 2016] L. Matthiesen, "Linear correlations of multiplicative functions", 2016. arXiv

[Montgomery and Vaughan 1977] H. L. Montgomery and R. C. Vaughan, "Exponential sums with multiplicative coefficients", *Invent. Math.* **43**:1 (1977), 69–82. MR Zbl

[Montgomery and Vaughan 2006] H. Montgomery and R. Vaughan, *Multiplicative Number Theory I: Classical Theory*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2006.

[Nair and Tenenbaum 1998] M. Nair and G. Tenenbaum, "Short sums of certain arithmetic functions", *Acta Math.* **180**:1 (1998), 119–144. MR

[Rankin 1973] R. A. Rankin, "An Ω-result for the coefficients of cusp forms", *Math. Ann.* **203** (1973), 239–250. MR Zbl

[Shiu 1980] P. Shiu, "A Brun–Titchmarsh theorem for multiplicative functions", *J. Reine Angew. Math.* **313** (1980), 161–170. MR Zbl

[Tao 2011] T. Tao, "The Katai–Bourgain–Sarnak–Ziegler orthogonality criterion", blog post, 2011, available at http://terrytao.wordpress.com/2011/11/21/the-bourgain-sarnak-ziegler-orthogonality.

[Tenenbaum 1995] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, Cambridge Studies in Advanced Mathematics **46**, Cambridge University Press, Cambridge, 1995. Translated from the second French edition (1995) by C. B. Thomas. MR Zbl

[Tenenbaum 2017] G. Tenenbaum, "Valeur moyennes effectives de fonctions multiplicatives complexes", *Ramanujan Journal* **44** (2017), 641–701.

[Wintner 1935] A. Wintner, "On the cyclical distribution of the logarithms of the prime numbers", *The Quarterly Journal of Mathematics* **6**:1 (1935), 65–68. Zbl

Communicated by Roger Heath-Brown
Received 2016-07-04 Revised 2017-09-18 Accepted 2017-10-30

lilian.matthiesen@math.kth.se *Department of Mathematics, KTH, Stockholm, Sweden*

Algebra & Number Theory

msp.org/ant

See inside back cover or msp.org/ant for submission instructions.

The subscription price for 2018 is US \$340/year for the electronic version, and \$535/year (+\$55, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Algebra & Number Theory (ISSN 1944-7833 electronic, 1937-0652 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

ANT peer review and production are managed by EditFLOW® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

Algebra & Number Theory

Volume 12 No. 6 2018

Generalized Fourier coefficients of multiplicative functions 1311
LILIAN MATTHIESEN

A blowup algebra for hyperplane arrangements 1401
MEHDI GARROUSIAN, ARON SIMIS and ŞTEFAN O. TOHĂNEANU

Torsion in the 0-cycle group with modulus 1431
AMALENDU KRISHNA

Continuity of the Green function in meromorphic families of polynomials 1471
CHARLES FAVRE and THOMAS GAUTHIER

Local root numbers and spectrum of the local descents for orthogonal groups: p-adic case 1489
DIHUA JIANG and LEI ZHANG

Bases for quasisimple linear groups 1537
MELISSA LEE and MARTIN W. LIEBECK