JOINT DISTRIBUTION IN RESIDUE CLASSES OF FAMILIES OF MULTIPLICATIVE FUNCTIONS I

AKASH SINGHA ROY

Abstract. We study the joint distribution of families of multiplicative functions in residue classes, allowing the moduli to vary within a wide range and assuming some natural control on the average behavior of the functions at (any fixed power of) the primes. As an application, we obtain essentially best possible analogues of the Siegel–Walfisz theorem for families of multiplicative functions that can be controlled by polynomials at the first few powers of all primes. (This class includes several interesting arithmetic functions such as Euler's totient $\varphi(n)$, the sum-of-divisors $\sigma(n)$, and more generally, the sum-of-divisor-powers $\sigma_r(n) := \sum_{d|n} d^r$, and so on.) Our results extend (and give essentially optimal uniform analogues of) works of Narkiewicz, Rayner, Śliwa, Dobrowolski, Fomenko and others. One of the main ideas behind our arguments is the detection of a certain "mixing"/"quantitative ergodicity" phenomenon via methods from sieve theory and the anatomy of integers. Additionally, we also need several ideas and machinery from classical analytic number theory, character sums, linear algebra over rings, as well as arithmetic and algebraic geometry.

1. Introduction

We say that an arithmetic function $f : \mathbb{N} \to \mathbb{Z}$ is equidistributed modulo $q \in \mathbb{N}$ if $\#\{n \le x : f(n) \equiv a \pmod q\} \sim x/q$ as $x \to \infty$. This notion has been studied for additive functions: See results of Delange [8, 9], which have been partially extended in [27, 28, 1, 34]. However, for multiplicative functions, it turns out that this notion is not the correct one to work with: For instance, using classical results (such as those implicit in [17]), it can be shown that Euler's totient function $\varphi(n)$ is almost always divisible by any fixed integer q, and hence is **not** equidistributed modulo any fixed $q > 1$. Motivated by this, Narkiewicz defines a function $f : \mathbb{N} \to \mathbb{Z}$ to be weakly equidistributed (or WUD) modulo $q \in \mathbb{N}$ if there are infinitely many n for which $(f(n), q) = 1$, and if

$$\#\{n \le x : f(n) \equiv a \pmod q\} \sim \frac{1}{\varphi(q)} \#\{n \le x : (f(n), q) = 1\} \quad \text{as } x \to \infty,$$

for every coprime residue a mod q.[1] Hence, our sample space is $\{n : (f(n), q) = 1\}$, and every coprime residue mod q gets its fair share of the sample space. For example, if f is WUD mod 6, then the residues 1 and 5 mod 6 each asymptotically receive 50% of the sample space $\{n : (f(n), 6) = 1\}$. The notion of weak equidistribution extends naturally to a family of arithmetic functions $f_1, \dots, f_K : \mathbb{N} \to \mathbb{Z}$: We say that $(f_1, \dots, f_K)$ are WUD mod q if there are infinitely many n for which $(f_1(n) \dots f_K(n), q) = 1$, and if for all coprime residues $a_1, \dots, a_K$ mod q, we have

(1.1)
$$\#\{n \le x : (\forall i)\ f_i(n) \equiv a_i \pmod q\} \sim \frac{1}{\varphi(q)^K} \#\{n \le x : (f_1(n) \dots f_K(n), q) = 1\} \quad \text{as } x \to \infty.$$

2020 *Mathematics Subject Classification.* Primary 11A25; Secondary 11N36, 11N37, 11N64, 11N69.

Key words and phrases. multiplicative function, uniform distribution, equidistribution, weak uniform distribution, joint distribution, distribution in residue classes.

[1] Here, (a, b) denotes the gcd of a and b, unless stated otherwise.

The notion of weak equidistribution has been studied by several authors. Narkiewicz gave a general criterion [20, Theorem 1] deciding weak equidistribution of a "polynomially–defined" multiplicative function; here, $f : \mathbb{N} \to \mathbb{Z}$ is polynomially–defined if there exist $V \in \mathbb{N}$ and polynomials $\{F_v\}_{1 \le v \le V} \subset \mathbb{Z}[T]$ satisfying $f(p^v) = F_v(p)$ for all primes p and all $v \in [V]$. (Hence, F_v controls the behavior of f at the v-th powers of primes.) Using his criterion, he was able to show in [20] that $\varphi(n)$ is WUD mod q iff $(q, 6) = 1$. (Related work of Dence and Pomerance [10] studies the distribution of $\varphi(n)$ in residue classes mod 3 and 12, while Banks and Shparlisnki [2, Theorem 3.1] give bounds on $\#\{n \le x : \varphi(n) \equiv a \pmod p\}$ for a prime modulus p.) Śliwa [36] used Narkiewicz's criterion to show that the function $\sigma(n)$ is WUD mod q iff $6 \nmid q$. This was extended to the functions $\sigma_r(n) = \sum_{d|n} d^r$ (which also appear as Fourier coefficients of Eisenstein series) by Narkiewicz, Rayner, Śliwa, Doborowolski and Fomenko [36, 13, 25, 23, 24, 29, 30]. In his monograph [24], Narkiewicz gives algorithms to list the moduli to which a given polynomially–defined multiplicative function (satisfying additional restrictions) is weakly equidistributed. Finally, the most general results on this subject were also obtained by Narkiewicz [22, 21], where he gave an exact characterization (as well as other explicit sufficient conditions) deciding when an arbitrary *family* of polynomially–defined multiplicative functions is weakly equidistributed to a fixed modulus.

In all these results, the modulus q is always assumed to be fixed. It is then natural to ask for parallels of these results when q is allowed to vary with x, in analogy with the Siegel–Walfisz theorem from prime number theory. A general question in number theory is to find analogues of the Siegel–Walfisz theorem in other contexts, and this question has been studied in the contexts of smooth numbers and mean values of multiplicative functions, among others. In our context, the corresponding question is: Can we find analogues of the Siegel–Walfisz theorem for the joint distribution of a general family of multiplicative functions? To formalize this, given $f_1, \dots, f_K : \mathbb{N} \to \mathbb{Z}$ and a set $\mathcal{Q} \subset \mathbb{N}$, we shall say that $(f_1, \dots, f_K)$ are WUD mod q, uniformly for $q \in \mathcal{Q}$, if:

- For every $q \in \mathcal{Q}$, we have $(\prod_{i=1}^K f_i(n), q) = 1$ for infinitely many n, and
- (1.1) holds as $x \to \infty$, uniformly in $q \in \mathcal{Q}$ and in coprime residues $a_1, \dots, a_K$ mod q.

The residue–class distribution of multiplicative functions to varying moduli seems to have been first (and only) studied in [18, 26, 28], which have made partial progress towards uniformizing Narkiewicz's criteria in [20, 22]. However, while [18] restricts only to $\varphi(n)$ and to prime moduli q, [26] requires q to vary within a *really* small range and be *"almost prime"* in a certain sense, while several of the results and arguments in [28] are limited to a single function (and do not generalize to joint distributions of families). These works also require the defining polynomials to be separable and for q to be supported on large primes only. Moreover, while Narkiewicz's criteria in [20, 22] allow additional generality with prime powers (and the possibility of sparse input sets), [18, 26, 28] are unable to account for this generality. As such, these works do not give satisfactory uniform analogues of the works of Śliwa, Rayner and others alluded to above, either.

In this manuscript, we remove *all* these limitations and give uniform analogues of Narkiewicz's most general criterion in [22], that are essentially optimal in almost every aspect. These results are thus also new for a single function. Our modulus q will be allowed to vary either within optimal ranges or up to any fixed power of $\log x$, hence we will also obtain essentially best possible qualitative analogues of the Siegel–Walfisz theorem for a general family of polynomially–defined multiplicative functions. As consequences, we can extend all the aforementioned works of Narkiewicz, Rayner, Śliwa, Doborowolski and Fomenko to varying moduli with optimal arithmetic restrictions.

Since Narkiewicz's criterion in [22] requires some technical set-up, we relegate our most general results to the last section. For most of the paper, we focus on obtaining uniform analogues of a special case of his criterion so as to illustrate most of the important ideas. It is in section 7 that we state the optimal uniform extensions of [22] for a *single* varying modulus, and mention the additional ingredients required for them. Some key inputs in our arguments are the two flexible and uniform bounds in Theorems 1.1 and 1.2 below, that measure the failure of weak equidistribution for a general family of multiplicative functions.

These two results are motivated from a certain "mixing" phenomenon that played a central role in the latest work [28]. We generalize, refine and quantify this "mixing" for an arbitrary family of multiplicative functions $f_1, \ldots, f_K$, assuming only some natural control on the *average* (joint) behavior of the f_i at the primes, more precisely the sums $\sum_{y<p\le Y} \prod_{i=1}^{K} \chi_i(f_i(p))$ (for Dirichlet characters $(\chi_1, \ldots, \chi_K)$ mod q). Owing to this generality, we also expect these results to have applications to other (weak) equidistribution problems for families of multiplicative functions outside of the focus of this paper (a potential future target might be Ramanujan's τ-function and Fourier coefficients of modular forms). In fact, Theorem 1.1 shows that to study the weak equidistribution of the f_i among inputs n with J large prime factors (J being a parameter we are allowed to choose), we basically just need to control the J-th moments $\sum_{\widehat{\chi} \bmod q} \left|\sum_{y<p\le Y} \prod_{i=1}^{K} \chi_i(f_i(p))\right|^J$. Theorem 1.2 also removes this restriction on n. Every parameter in these results will be useful in our application.

Let $q \in \mathbb{N}$. In what follows, we use $\boldsymbol{\widehat{\chi}}$ for the tuple $(\chi_1, \ldots, \chi_K)$ of Dirichlet characters mod q, $\boldsymbol{\mathfrak{f}(\widehat{\chi})}$ for the lcm of the conductors of χ_i, and χ_0 for the trivial (or principal) character mod q. For a tuple of integer-valued functions $\boldsymbol{\widehat{f} := (f_1, \ldots, f_K)}$, let $\boldsymbol{f := f_1 \cdots f_K}$ and let $\boldsymbol{\widehat{\chi}(\widehat{f}(n)) := \chi_1(f_1(n)) \cdots \chi_K(f_K(n))}$. Let $\boldsymbol{P_R(n)}$ denote the R-th largest prime factor of n counted with multiplicity (with the convention that $P_R(n) := 1$ if $\Omega(n) < R$). [2]

Our first main result studies weak equidistribution among inputs with several large prime factors.

Theorem 1.1. *Let $f_1, \ldots, f_K : \mathbb{N} \to \mathbb{Z}$ be multiplicative functions, $q \in \mathbb{N}$, and $y \ge q$ a parameter such that for all $\widehat{\chi}$ mod q and all $Y \ge y$, we have*

$$\sum_{y<p\le Y} \widehat{\chi}(\widehat{f}(p)) \;=\; \alpha_{\widehat{\chi}} \sum_{y<p\le Y} \chi_0(f(p)) + O(MY\mathcal{E}(y)) \tag{1.2}$$

for some $M \ge 1$, some $\alpha_{\widehat{\chi}} \in \mathbb{C}$ (depending on $\widehat{\chi}$) in the unit disk, and some decreasing function $\mathcal{E} : \mathbb{R}^+ \to \mathbb{R}^+$. Then for all d dividing q, all $x \ge y$ and all $J \in \mathbb{N}$, we have

$$\sum_{\substack{n\le x:\ P_J(n)>y \\ (\forall i)\ f_i(n)\equiv a_i \pmod q}} 1 \;-\; \left(\frac{\varphi(d)}{\varphi(q)}\right)^K \sum_{\substack{n\le x:\ P_J(n)>y,\ (f(n),q)=1 \\ (\forall i)\ f_i(n)\equiv a_i \pmod d}} 1 \tag{1.3}$$
$$\ll\; \frac{1}{\varphi(q)^K} \sum_{\substack{\widehat{\chi} \bmod q \\ \mathfrak{f}(\widehat{\chi})\,\nmid\, d}} |\alpha_{\widehat{\chi}}|^J \sum_{\substack{n\le x \\ (f(n),q)=1}} 1 \;+\; \frac{Jx}{y} \;+\; Mx\mathcal{E}(y)\log x.$$

The implied constant here depends only on the implied constant in (1.2).

---

[2] In this paper, **bold** is used only for emphasis (for ease of reference), not as part of the notation itself.

A few remarks are in order. First, note that (1.2) is in general weaker than a hypothesis of the form $\sum_{y<p\le Y}\widehat{\chi}(\widehat{f}(p)) = \beta_{\widehat{\chi}}\sum_{y<p\le Y} 1 + O(MY\mathcal{E}(y))$. Also, to get (1.3) for a certain d dividing q, we need only assume (1.2) for all $\widehat{\chi}$ mod q with $\mathfrak{f}(\widehat{\chi}) \nmid d$. Taking $J = d = 1$ in (1.3) bounds the failure of weak equidistribution, however, to make the right side of (1.3) negligible in applications, we often need J to be large. We thus see what happens if we include the inputs n having $P_J(n) \le y$.

For each $r \ge 1$ and $z > 0$, **define $\boldsymbol{\xi_r(q;z)} > 0$** to be any parameter satisfying, for all $x \ge z$ and all coprime residues $a_1, \dots, a_K$ mod q, the bound

$$(K_r) \quad \begin{aligned} &\#\{P_1\cdots P_r \le x : P_i \text{ primes},\ P_1 > z,\ q < P_r < \cdots < P_1,\ (\forall i)\ f_i(P_1)\cdots f_i(P_r) \equiv a_i \pmod q\} \\ &\qquad\qquad \le\ \xi_r(q;z)\cdot x(\log\log x)^r \Big/ \log(z/q). \end{aligned}$$

Thus, $\xi_r(q;z)$ controls the number of squarefree, q-rough, non-z-smooth A with r prime factors, that satisfy $f_i(A) \equiv a_i \pmod q$ for all i. (Note that $\xi_r(q;z) = 1$ satisfies (K_r) by Chebyshev's estimates.) Let $\Psi(x,z)$ be the number of z-smooth numbers up to x, a well–studied quantity.

Theorem 1.2. *Assume the setting of Theorem 1.1. For all d dividing q, all $x, z > 0$, and all $J, R \in \mathbb{N}$ satisfying $y \le z \le x$ and $J \ge R$, we have*

$$(1.4) \quad \sum_{\substack{n\le x:\ P_R(n)>q \\ (\forall i)\ f_i(n)\equiv a_i \ (\mathrm{mod}\ q)}} 1 \;-\; \left(\frac{\varphi(d)}{\varphi(q)}\right)^K \sum_{\substack{n\le x:\ (f(n),q)=1 \\ (\forall i)\ f_i(n)\equiv a_i \ (\mathrm{mod}\ d)}} 1$$
$$\ll\ \frac{1}{\varphi(q)^K}\sum_{\substack{\widehat{\chi} \bmod q \\ \mathfrak{f}(\widehat{\chi})\,\nmid\, d}} |\alpha_{\widehat{\chi}}|^J \sum_{\substack{n\le x \\ (f(n),q)=1}} 1 \ +\ \Psi(x,z)\ +\ \frac{Jx}{y}\ +\ Mx\mathcal{E}(y)\log x$$
$$+\ \left\{\left(\frac{\varphi(d)}{\varphi(q)}\right)^K + \xi_R(q;z) + R^K \sum_{\substack{1\le r<R \\ 1\le s<K}} \frac{\xi_r(q;z)}{q^{\max\{s,R-r-s\}}}\right\}\frac{x(2\log\log x)^{R+J}}{\log(z/q)}\exp\Bigg(\sum_{\substack{p\le y \\ (f(p),q)=1}}\frac{1}{p}\Bigg).$$

The implied constant here depends only on K and on the implied constant in (1.2).

The flexibility in choosing the parameters d, y, z, J, R (especially the fact that z and J now appear only on the right of (1.4), unlike in (1.3)) are quite useful in applications: For instance, by taking $d = R = 1$ and a large J (as we'll do in an upcoming application), we can make the right side negligible while including *all* inputs n on the left. (Having a general R will also be useful.) The strength of Theorem 1.2 basically depends on the bounds available on $\sum_{\mathfrak{f}(\widehat{\chi})\,\nmid\, d} |\alpha_{\widehat{\chi}}|^J$, and on $\xi_r(q;z)$ for $r \in [R]$. As discussed in Remark 2, Theorems 1.1 and 1.2 can be extended by replacing the condition "$\mathfrak{f}(\widehat{\chi}) \nmid d$" by the more general "$\widehat{\chi} \notin \mathcal{S}$", where $\mathcal{S}$ is an arbitrary subset of the set of K-tuples of characters mod q; the corresponding variant of (1.4) would be uniform in $\mathcal{S}$.

We now come to the relevant special case of Narkiewicz's general criterion in [22]. Consider multiplicative functions $f_1, \dots, f_K : \mathbb{N} \to \mathbb{Z}$ for which there exist polynomials $F_1, \dots, F_K \in \mathbb{Z}[T]$ satisfying $f_i(p) = F_i(p)$ for all primes p and all $i \in [K]$. For $q \in \mathbb{N}$ we define

$$\alpha(q) := \frac{1}{\varphi(q)}\#\{u \in U_q : F_1(u)\cdots F_K(u) \in U_q\},\ \text{ where } U_q := (\mathbb{Z}/q\mathbb{Z})^\times.$$

Let $\boldsymbol{\mathcal{Q}(f_1, \dots, f_K)}$ denote the set of all $q \in \mathbb{N}$ with $\boldsymbol{\alpha(q) \neq 0}$ that satisfy the following property: **For all characters $(\chi_1, \dots, \chi_K) \neq (\chi_0, \dots, \chi_0)$ mod q for which $\widehat{\chi}(\widehat{F}(u)) = 1$ on its "unit support" $\{u \in U_q : F_1(u) \cdots F_K(u) \in U_q\}$, we have $\widehat{\chi}(\widehat{f}(2^j)) = -1$ for all $j \geq 1$.**

Narkiewicz's criterion in this special setting is stated as follows.

Theorem 1.3. [22] *Assume $\alpha(q) \neq 0$. Then $(f_1 \dots, f_K)$ are WUD mod q iff $q \in \mathcal{Q}(f_1, \dots, f_K)$.*

To state the uniform analogues of this, we assume that $F_1, \dots, F_K$ are **multiplicatively independent**, i.e. that $\prod_{i=1}^K F_i^{c_i}$ is nonconstant in $\mathbb{Q}(T)$ for all integers $(c_1, \dots, c_K) \neq (0, \dots, 0)$.[3] Factor each F_i as $r_i \prod_{j=1}^M G_j^{\mu_{i,j}}$ with $r_i \in \mathbb{Z}$, $\mu_{i,j} \in \mathbb{N} \cup \{0\}$, and with $\boldsymbol{\{G_j\}_{j=1}^M} \subset \mathbb{Z}[T]$ being pairwise coprime primitive irreducibles, such that each G_j appears with a positive exponent $\mu_{i,j}$ in some F_i. With $\boldsymbol{E_0}$ denoting the $M \times K$ integer matrix $\left((\boldsymbol{\mu_{i,j}})_{\substack{1 \leq i \leq K \\ 1 \leq j \leq M}} \right)^{\intercal}$, let $\boldsymbol{\beta := \beta(F_1, \dots, F_K)} \in \mathbb{N}$ be the largest invariant factor of E_0. Fixing any $B_0 > 0$, we say that q satisfies **Hypothesis $\boldsymbol{IFH(F_1, \dots, F_K; B_0)}$** if all prime divisors ℓ of q exceeding B_0 satisfy $\gcd(\ell - 1, \beta) = 1$. Note that if $\prod_{i=1}^K F_i$ is separable (as is often the case in applications), then $\beta = 1$ and any q satisfies $IFH(F_1, \dots, F_K; B_0)$ for any $B_0 > 0$. Our first uniform analogue of Theorem 1.3 is the following.

Theorem 1.4. *Fix $K_0 > 0$ and $\epsilon \in (0, 1)$, and assume that $F_1, \dots, F_K$ are multiplicatively independent. Then $(f_1, \dots, f_K)$ are WUD mod $q \leq (\log x)^{K_0}$ lying in $\mathcal{Q}(f_1, \dots, f_K)$ and satisfying $IFH(F_1, \dots, F_K; B_0)$, provided at least one of the following two conditions holds:*

(i) *$q \leq (\log x)^{(1-\epsilon)\alpha(q)(K-1/D_{\min})^{-1}}$, where $D_{\min} := \min\{\deg F_i : 1 \leq i \leq K\}$,* ***or***

(ii) *q is squarefree and $q^{K-1} D_{\min}^{\omega(q)} \leq (\log x)^{(1-\epsilon)\alpha(q)}$.*

The conditions in (i) and (ii) are ignored if $K = D_{\min} = 1$ (i.e. a single multiplicative function controlled by a linear polynomial at the primes); so in this case, we have uniformity up to any fixed power of $\log x$. As an application of Theorem 1.4, note that while Narkiewicz [20] (respectively, Śliwa [36]) show that $\varphi(n)$ (resp. $\sigma(n)$) is WUD modulo any fixed q coprime to 6 (resp. odd q), Theorem 1.4 shows that in either case, weak equidistribution holds *uniformly* modulo all respective $q \leq (\log x)^{K_0}$. Moreover, while Narkiewicz [21, Theorem 1] shows that (φ, σ) are (jointly) WUD modulo any *fixed* q coprime to 6, Theorem 1.4 shows that (φ, σ) are WUD *uniformly* modulo $q \leq (\log x)^{(1-\epsilon)\alpha_0(q)}$ coprime to 6, where $\alpha_0(q) = \prod_{\ell \mid q \text{ prime}} (\ell - 3)/(\ell - 1)$.

In section 6, we show that (except when $K = D_{\min} = 1$), both ranges in (i) and (ii) are essentially optimal, in that "$1 - \epsilon$" cannot be replaced by "$1 + \epsilon$" in either. In fact, if $K > 1$ and $D_{\min} = 1$, then this optimality holds in a much stronger sense: The ranges in (i) and (ii) are optimal no matter how many of the F_i we assume to be linear; for instance, the range $q \leq (\log x)^{(1-\epsilon)\alpha_0(q)}$ above is optimal even for the particular example of (φ, σ) above.

As our constructions in section 6 show, obstructions to uniformity come from prime inputs. We can modify those constructions to produce more obstructions coming from inputs n that have too few large prime factors. Hence to restore uniformity in the full Siegel-Walfisz range, we need to restrict our n to those with sufficiently many large prime factors.

[3]This hypothesis is easily satisfied in applications: For instance, it is satisfied if $\prod_{i=1}^K F_i$ is separable, or (more generally), if each F_i has an irreducible factor that is not present in the other F_j (for $j \neq i$).

Theorem 1.5. *Fix $K_0 > 0$ and assume that $F_1, \dots, F_K$ are multiplicatively independent. Then*

$$\sum_{\substack{n \le x:\ P_R(n) > q \\ (\forall i)\ f_i(n) \equiv a_i \pmod q}} 1 \sim \frac{1}{\varphi(q)^K} \sum_{\substack{n \le x:\ P_R(n) > q \\ (f(n), q) = 1}} 1 \sim \frac{1}{\varphi(q)^K} \sum_{\substack{n \le x \\ (f(n), q) = 1}} 1 \quad \text{as } x \to \infty, \tag{1.5}$$

uniformly in $q \le (\log x)^{K_0}$ lying in $\mathcal{Q}(f_1, \dots, f_K)$ and satisfying $IFH(F_1, \dots, F_K; B_0)$, as well as uniformly in coprime residues $a_1, \dots, a_K$ modulo q. Here

$$R = \begin{cases} KD + 1, & \text{in general, where } D := \sum_{i=1}^K \deg F_i, \\ 2K + 1, & \text{if } q \text{ is squarefree}, \\ 2, & \text{if } q \text{ is squarefree}, K = 1, \text{ and } F_1 \text{ is } \textbf{not squarefull} \text{ (has a simple root in } \mathbb{C}). \end{cases}$$

As an application, (φ, σ) are WUD modulo any $q \le (\log x)^{K_0}$ coprime to 6 if we restrict to inputs n with $P_5(n) > q$. (Without this restriction, we only have the aforementioned optimal range $q \le (\log x)^{(1-\epsilon)\alpha_0(q)}$.) Another application: By Narkiewicz [21, 22], $(\varphi, \sigma, \sigma_2)$ are WUD modulo any fixed q having $P^-(q) > 23$. By Theorem 1.4, this weak equidistribution holds uniformly modulo any such $q \le (\log x)^{(1/2-\epsilon)c_q}$ where $c_q := \prod_{\ell | q:\ \ell \equiv -1 \pmod 4} (\ell-3)/(\ell-1) \cdot \prod_{\ell | q:\ \ell \equiv 1 \pmod 4} (\ell-5)/(\ell-1)$. Uniformity is restored modulo *all* $q \le (\log x)^{K_0}$ satisfying $P^-(q) > 23$, provided we restrict to n having $P_{13}(n) > q$. This can be weakened to $P_7(n) > q$, for squarefree q.

The smaller the value of R in (1.5), the larger our set wherein equidistribution occurs, and the better our result. In section 6, we show that the third value of R is optimal, while the second is at most "one step away" from being optimal (in the sense that it cannot be reduced to $2K-1$). Note that even the $K = 1$ special case of Theorem 1.5 improves over [28, Theorem 1.4(a)]. Moreover, the multiplicative independence hypothesis and hypothesis $IFH(F_1, \dots, F_K; B_0)$ are also both necessary in Theorems 1.4 and 1.5: We shall establish this in the sequel [35].

We remark that Narkiewicz's general weak equidistribution criteria in [20, 22] are much more general than Theorem 1.3, in that they allow the possibility that the behavior of our f_i at the ν-th powers of primes is most significant (in a certain sense), where $\nu \in \mathbb{N}$ is fixed but arbitrary; this can also handle the possibility of sparse input sets. However, the results and methods from previous work on varying moduli [18, 26, 28] are unable to handle any $\nu > 1$. Adapting our methods for Theorems 1.4 and 1.5 and giving some additional arguments, we can also obtain optimal uniform results in this greater generality considered by Narkiewicz. This generality also has several new applications; we discuss them in section 7. Along the way, we also give extensions of Theorems 1.1 and 1.2 that apply for certain sparse input sets.

Our arguments comprise several themes. Theorems 1.1 and 1.2 can be intuitively thought of as being governed by a "mixing" phenomenon, which we detect via methods from the anatomy of integers. In Theorems 1.4 and 1.5, we also require inputs from classical analytic number theory, both from a general bound of Halász (which involves a careful estimation of "pretentious distances") and from a modification of the Landau–Selberg–Delange method (inspired from work of Scourfield). Further, we need machinery from character sums (extensions of Weil bounds), linear algebra over rings (via Smith normal forms and invariant factors), and arithmetic geometry/algebraic geometry (by constructing regular sequences and counting rational points on varieties over finite fields).

In a future paper, we shall study the joint weak equidistribution of multiplicative functions $(f_1, \dots, f_K)$ to moduli $(q_1, \dots, q_K)$, defined appropriately. To conclude the introduction, we mention the following estimate, which forms a key input in the proofs of Theorems 1.4 and 1.5, and may be of independent interest or application to other problems.

Proposition 1.6. *Under the hypotheses of Theorem 1.1, we have for any $\widehat{\chi}$ mod q,*

$$\sum_{n \le x} \widehat{\chi}(\widehat{f}(n)) \ll \frac{x}{\log x} \exp\left(\sum_{\substack{p \le y \\ (f(p),q)=1}} \frac{1}{p} + |\alpha_{\widehat{\chi}}| \sum_{\substack{y < p \le x \\ (f(p),q)=1}} \frac{1}{p} + O(M (\log x)^3 \mathcal{E}(y)) \right).$$

The implied constant depends only on the implied constant in (1.2).

2. The General Equidistribution Results: Proof of Theorems 1.1 and 1.2

We say that $n \le x$ is "convenient" (abbreviated as "conv") if the J largest prime divisors of n exceed y and are distinct. Thus any convenient n can be uniquely written as $mP_J \cdots P_1$, where

$$L_m := \max\{y, P(m)\} < P_J < \cdots < P_1. \tag{2.1}$$

Now, the number of $n \le x$ having a repeated prime factor exceeding y is at most $\sum_{p>y} \sum_{m \le x/p^2} 1 \ll x/y$, which is absorbed in the right of (1.4). Hence, to complete the proof of Theorem 1.2, it suffices to show (1.3) with the condition "$P_J(n) > y$" replaced by the condition that n is convenient.

By the orthogonality of Dirichlet characters, we detect the congruences $f_i(n) \equiv a_i \pmod q$.

$$\sum_{\substack{n \le x \text{ conv} \\ (\forall i)\ f_i(n) \equiv a_i \pmod q}} 1 = \frac{1}{\varphi(q)^K} \sum_{\widehat{\chi} \bmod q} \overline{\chi}_1(a_1) \cdots \overline{\chi}_K(a_K) \sum_{n \le x \text{ conv}} \widehat{\chi}(\widehat{f}(n)). \tag{2.2}$$

For any $\widehat{\chi}$ mod q for which $\mathfrak{f}(\widehat{\chi}) \mid d$, there exists a unique tuple of characters $\widehat{\psi} := (\psi_1, \dots, \psi_K)$ mod d such that $\chi_i = \chi_0 \psi_i$ for all i. Thus the total of contribution of all such $\widehat{\chi}$ in (2.2) equals

$$\frac{1}{\varphi(q)^K} \sum_{\widehat{\psi} \bmod d} \overline{\psi}_1(a_1) \cdots \overline{\psi}_K(a_K) \sum_{\substack{n \le x \text{ conv} \\ (f(n),q)=1}} \widehat{\psi}(\widehat{f}(n)) = \left(\frac{\varphi(d)}{\varphi(q)}\right)^K \sum_{\substack{n \le x \text{ conv: } (f(n),q)=1 \\ (\forall i)\ f_i(n) \equiv a_i \pmod d}} 1. \tag{2.3}$$

Now consider any $\widehat{\chi}$ mod q. Splitting any convenient n uniquely as $mP_J \cdots P_1$ above, we obtain

$$\sum_{n \le x \text{ conv}} \widehat{\chi}(\widehat{f}(n)) = \sum_{m \le x} \widehat{\chi}(\widehat{f}(m)) \cdot \frac{1}{J!} \sum_{\substack{P_1, \dots, P_J > L_m \\ P_1, \dots, P_J \text{ distinct} \\ P_1 \cdots P_J \le x/m}} \widehat{\chi}(\widehat{f}(P_1)) \cdots \widehat{\chi}(\widehat{f}(P_J)), \tag{2.4}$$

where we have have replaced the ordering on $P_1, \dots, P_J$ by the weaker condition "$P_1, \dots, P_J$ distinct" at the cost of $J!$. For each $i \in [J]$, we see that

$$\begin{aligned} &\sum_{\substack{P_i \le x/mP_1 \cdots P_{i-1} P_{i+1} \cdots P_J \\ P_i > L_m,\ r \ne i \implies P_r \ne P_i}} \widehat{\chi}(\widehat{f}(P_i)) \\ &\qquad = \alpha_{\widehat{\chi}} \sum_{\substack{P_i \le x/mP_1 \cdots P_{i-1} P_{i+1} \cdots P_J \\ P_i > L_m,\ r \ne i \implies P_r \ne P_i}} \chi_0(f(P_i)) + O\left(J + \frac{Mx\mathcal{E}(y)}{mP_1 \cdots P_{i-1} P_{i+1} \cdots P_J}\right), \end{aligned} \tag{2.5}$$

where we have used (1.2), and removed and restored the "$\neq$" condition with error $O(J)$. The total O-term in (2.5) summed over all m and all the other P_r in (2.4) is $\ll x/y + J^{-1}Mx\mathcal{E}(y)\sum_{n\le x/y} 1/n \ll x/y + J^{-1}Mx\mathcal{E}(y)\log x$. (Here we used the $(J-1)!$ from (2.4) to restore the ordering on $P_1,\dots,P_{i-1},P_{i+1},\dots,P_J$, and then noted that $n = mP_1\cdots P_{i-1}P_{i+1}\cdots P_J \le x/y$.) Using this observation along with (2.5) for each $i=1,\dots,J$ successively, we get from (2.4),

$$\sum_{n\le x \text{ conv}} \widehat{\chi}(\widehat{f}(n)) \;=\; (\alpha_{\widehat{\chi}})^J \sum_{m\le x} \widehat{\chi}(\widehat{f}(m)) \sum_{\substack{P_1,\dots,P_J \\ L_m<P_J<\dots<P_1 \\ P_1\cdots P_J\le x/m}} \chi_0(f(P_1\cdots P_J)) \;+\; O\left(\frac{Jx}{y} + Mx\mathcal{E}(y)\log x\right). \tag{2.6}$$

The first sum on the right side is $\ll |\alpha_{\widehat{\chi}}|^J \#\{n\le x \text{ conv} : (f(n),q)=1\}$. Using this for all $\widehat{\chi}$ mod q for which $\mathfrak{f}(\widehat{\chi}) \nmid d$, we obtain Theorem 1.1.

We now deduce Theorem 1.2 from Theorem 1.1. To this end, we start by observing that

$$\sum_{\substack{n\le x:\ P_J(n)\le y \\ (f(n),q)=1}} 1 \;\ll\; \frac{x(2\log_2 x)^J}{\log z}\ \exp\left(\sum_{p\le y} \frac{\mathbb{1}_{(f(p),q)=1}}{p}\right) \;+\; \Psi(x,z) \;+\; \frac{x}{y}. \tag{2.7}$$

Up to an error of $O(\Psi(x,z)+x/z)$, we may also assume that $P(n)>z$ and that n has no repeated prime factor exceeding z. Write any such n as $n = BAP$, where $P(B)\le y < P^-(A)$, and $P := P(n) \in (z, x/AB]$. Then B,A,P are pairwise coprime and $\Omega(A)\le J$ (since $P_J(n)\le y$). By Chebyshev's estimates, the number of P is $\ll x/AB\log z$. Moreover, $\sum_{A\le x:\ \Omega(A)\le J} 1/A \le (1+\sum_{p\le x} 1/p)^J \le (2\log_2 x)^J$, while $\sum_{\substack{B:\ P(B)\le y \\ (f(B),q)=1}} 1/B \le \prod_{p\le y}\left(1+\sum_{r\ge1} \mathbb{1}_{(f(p^r),q)=1}/p^r\right) \ll \exp(\sum_{p\le y} \mathbb{1}_{f(p),q)=1}/p)$. Collecting all the estimates in this paragraph proves (2.7).

To complete the proof of Theorem 1.2, we thus only need to bound the contribution of $n\le x$ satisfying $P_R(n)>q$, $P_J(n)\le y$, and $f_i(n)\equiv a_i \pmod q$ for all i. Let $\sum^*_{n\le x}$ denote any sum over $n\le x$ having a squarefree y-rough part, and satisfying $P_R(n)>q$, $P_J(n)\le y$, $P(n)>z$, and $f_i(n)\equiv a_i \pmod q$ for all i (additional conditions may be imposed). By the reductions in the previous paragraph, it suffices to bound $\sum^*_{n\le x} 1$. Let $E_{r,s}$ be the number of n counted in $\sum^*_{n\le x} 1$ having $\#\{p>q:\ p\,\|\,n\}=r$ and $\#\{p>q:\ p^2\mid n\}=s$. Then

$$\sum\nolimits^*_{n\le x} 1 \;\le\; \sideset{}{^*}\sum_{\substack{n\le x \\ \#\{p>q:\ p\|n\}\ge R}} 1 \;+\; \sideset{}{^*}\sum_{\substack{n\le x \\ \#\{p>q:\ p^2|n\}\ge K}} 1 \;+\; \sum_{\substack{1\le r<R \\ 1\le s<K}} E_{r,s}\ . \tag{2.8}$$

To bound $E_{r,s}$, note that any n counted in it can be decomposed as $mp_1^{c_1}\cdots p_s^{c_s}A$, where $P(m)\le q$, all $p_j>q$ are primes, all $c_j\ge 2$, where A is squarefree and q-rough with $\Omega(A)=r$, and where the factors are all pairwise coprime. Then $(f(m),q)=1$, $f_i(A)\equiv a_i f_i(mp_1^{c_1}\cdots p_s^{c_s})^{-1} \bmod q$ for all i, and $c_1+\dots+c_s \ge R-r$ (as $P_R(n)>q$). By (K_r), the number of possible A is $\ll \xi_r(q;z)x(\log_2 x)^r/mp_1^{c_1}\cdots p_s^{c_s}\log(z/q)$. Next, $\sum_{\substack{m:\ P(m)\le q \\ (f(m),q)=1}} 1/m \ll \exp(\sum_{p\le q} \mathbb{1}_{(f(p),q)=1}/p)$, and $\sum_{p_1,\dots,p_s>q} p_1^{-c_1}\cdots p_s^{-c_s} \ll q^{-(c_1+\dots+c_s-s)}$. The sum of $q^{-(c_1+\dots+c_s-s)}$ over all possible $(c_1,\dots,c_s)$ can be bounded in two ways: On the one hand, it is $\ll q^s\prod_{j=1}^s(\sum_{c_j\ge2} q^{-c_j}) \ll q^{-s}$, and on the other (writing $t := c_1 + \cdots + c_s$), it is $\ll \sum_{t \ge R-r} t^s/q^{t-s} \ll R^K/q^{R-r-s}$. Collecting everything, we get

$$E_{r,s} \ll R^K \frac{\xi_r(q;z)}{q^{\max\{s,R-r-s\}}} \cdot \frac{x(\log_2 x)^R}{\log(z/q)} \exp\left(\sum_{p \le q} \frac{\mathbb{1}_{(f(p),q)=1}}{p}\right),$$

which shows that the term $\sum_{\substack{1 \le r < R \\ 1 \le s < K}} E_{r,s}$ in (2.8) is absorbed in the right of (1.4). The other two sums on the right of (2.8) can be bounded with analogous (and simpler) arguments. (For example, we write the n counted in the second sum as $mp_1^{c_1} \cdots p_s^{c_s} P$, where $P_J(m) \le y$, $P > z$, all $p_j > q$, and all $c_j \ge 2$.) Hence $\sum_{n \le x}^{*}$ is bounded by the right of (1.4), establishing Theorem 1.2. □

Remark. Given an arbitrary subset $\mathcal{S}$ of the set of K-tuples of characters mod q, applying (2.6) and its ensuing estimate for $\widehat{\chi} \notin \mathcal{S}$, we see that the more general version of Theorem 1.2 holds with the condition "$\mathfrak{f}(\widehat{\chi}) \nmid d$" on the right of (1.4) replaced by "$\widehat{\chi} \notin \mathcal{S}$", and with the second term on the left replaced by $\varphi(q)^{-K} \sum_{\widehat{\chi} \in \mathcal{S}} \overline{\chi}_1(a_1) \cdots \overline{\chi}_K(a_K) \sum_{n \le x} \widehat{\chi}(\widehat{f}(n))$. This more general estimate doesn't involve d and is uniform in all $\mathcal{S}$. (The analogous comment holds for Theorem 1.1.)

3. A General Character Sum Bound: Proof of Proposition 1.6

Consider any $\widehat{\chi}$ mod q. By Halász's Theorem (as stated in [37, Corollary III.4.12]), we have

$$\sum_{n \le x} \widehat{\chi}(\widehat{f}(n)) \ll \frac{x}{L} + x \exp\left(- \min_{-L \le t \le L} \mathcal{D}(x,t)\right), \tag{3.1}$$

uniformly in $x, L \ge 2$, where by Mertens' Theorem,

$$\mathcal{D}(x,t) := \sum_{p \le x} \frac{1 - \mathrm{Re}(\widehat{\chi}(\widehat{f}(p))p^{-it})}{p} \ge \log_2 x - \sum_{\substack{p \le y \\ (f(p),q)=1}} \frac{1}{p} - \left|\sum_{y < p \le x} \frac{\widehat{\chi}(\widehat{f}(p))}{p^{1+it}}\right| + O(1). \tag{3.2}$$

Set $\boldsymbol{L := \log x}$, and cover the interval $(y, x]$ with disjoint "multiplicatively narrow" subintervals of the form $\boldsymbol{I_\eta := (\eta, \eta(1 + 1/\log^2 x)]}$, such that the rightmost of these juts out slightly past x but not past $x(1 + 1/\log^2 x)$. Note that $\sum_{x < p \le x(1+1/\log^2 x)} 1/p \ll 1/\log^3 x$, and that for any $p \in I_\eta$, we have $p = \eta(1 + O(1/\log^2 x))$ so that $p^{-(1+it)} = \eta^{-(1+it)}(1 + O(1/\log x))$ uniformly in $t \in [-L, L]$. Combining these and noting that $\sum_\eta \sum_{p \in I_\eta} 1/\eta \ll \sum_{p \le 2x} 1/p \ll \log_2 x$, we get

$$\sum_{y < p \le x} \frac{\widehat{\chi}(\widehat{f}(p))}{p^{1+it}} = \sum_\eta \frac{1}{\eta^{1+it}} \sum_{p \in I_\eta} \widehat{\chi}(\widehat{f}(p)) + O\left(\frac{\log_2 x}{\log x}\right). \tag{3.3}$$

Using (1.2) to estimate $\sum_{p \in I_\eta} \widehat{\chi}(\widehat{f}(p))$ and noting that there are $O(\log^3 x)$ many I_η's, we obtain

$$\sum_{y < p \le x} \frac{\widehat{\chi}(\widehat{f}(p))}{p^{1+it}} = \sum_\eta \frac{\alpha_{\widehat{\chi}}}{\eta^{1+it}} \sum_{p \in I_\eta} \chi_0(f(p)) + O\left(\frac{\log_2 x}{\log x} + M\mathcal{E}(y)(\log x)^3\right).$$

Finally, we use the two observations before (3.3) to replace the main term above by the expression $\alpha_{\widehat{\chi}} \sum_{y < p \le x} \chi_0(f(p))/p^{1+it}$ (upto error $\ll \log_2 x/\log x$). Inserting the resulting estimate into (3.2),

$$\mathcal{D}(x,t) \ge \log_2 x - \sum_{\substack{p \le y \\ (f(p),q)=1}} \frac{1}{p} - |\alpha_{\widehat{\chi}}| \sum_{\substack{y < p \le x \\ (f(p),q)=1}} \frac{1}{p} + O(1 + M\mathcal{E}(y)(\log x)^3). \tag{3.4}$$

Inserting this into (3.1) completes the proof of Proposition 1.6. □

4. Setting the stage for theorems 1.4 and 1.5: Character sums, module theory and a touch of arithmetic geometry

We remind the reader of the set-up in Theorems 1.4 and 1.5, which we will be assuming throughout until section 6: $f_1, \dots, f_K : \mathbb{N} \to \mathbb{Z}$ are multiplicative functions and $F_1, \dots, F_K \in \mathbb{Z}[T]$ satisfy $f_i(p) = F_i(p)$ for all primes p and all i. We define $f := \prod_{i=1}^K f_i$ and $F := \prod_{i=1}^K F_i$. We assume that $(F_i)_{i=1}^K$ are **multiplicatively independent**, hence the matrix $\boldsymbol{E_0} = \left((\mu_{i,j})_{\substack{1\le i\le K\\ 1\le j\le M}} \right)^\intercal$ has rank K. Its Smith normal form is $\mathrm{diag}(\beta_1, \dots, \beta_K)$ where $\beta_i \in \mathbb{N}$ and $\beta_i \mid \beta_{i+1}$ for all i. Recall that $\boldsymbol{\beta = \beta_K}$. Further, we are assuming that $\boldsymbol{\alpha(q)} = \frac{1}{\varphi(q)}\#\{u \in U_q : F(u) \in U_q\} \neq 0$; as such, the Chinese Remainder Theorem and Mertens' Theorem, it is easy to see that

$$\alpha(q) \gg (\log_2(3q))^{-D} \text{ uniformly in } q \ge 1, \text{ where } D = \deg F = \sum_{i=1}^K \deg F_i. \tag{4.1}$$

In what follows, for an odd prime power ℓ^e, let ψ_{ℓ^e} denote a character mod ℓ^e which generates the character group of U_{ℓ^e}. We will need two bounds on character sums. The first is a version of the Weil bounds, a special case of [38, Corollary 2.3] (see also [7], [39] and [31] for older results).

Proposition 4.1. *Let ℓ be a prime. Consider $H \in \mathbb{Z}[T]$ not of the form $c \cdot G^{\ell-1}$ in $\mathbb{F}_\ell[T]$. Then $|\sum_{u \bmod \ell} \psi_\ell(H(u))| \le (r-1)\sqrt{\ell}$ where r is the degree of the largest squarefree divisor of H.*

The second is an extension of this to higher prime powers, which comes from Theorems 1.2 and 7.1 and eqn. (1.15) in work of Cochrane [5] (see [6] for related results).

Proposition 4.2. *Let ℓ be a prime, $H \in \mathbb{Z}[T]$ nonconstant with $\mathrm{ord}_\ell(H) = 0$. Define $t := \mathrm{ord}_\ell(H')$, $\mathcal{C}_H := \ell^{-t}H'$ and $M := \max\{\mathrm{mult}_\theta(\mathcal{C}_H) : \theta \in \mathbb{F}_\ell, H(\theta) \neq 0\}$. Let $e \ge t+2$ be an integer.*

(i) *If $\ell > 2$, then* $\left|\sum_{u \bmod \ell^e} \psi_{\ell^e}(H(u))\right| \le \left(\sum_{\substack{\theta\in\mathbb{F}_\ell\\ H(\theta)\neq 0}} \mathrm{mult}_\theta(\mathcal{C}_H) \right) \ell^{t/(M+1)}\ \ell^{e(1-1/(M+1))}$.

(ii) *Assume that $e \ge t+3$ and let ψ be the character mod 2^e defined by $\psi(5) := \exp(2\pi i/2^{e-2})$ and $\psi(-1) := 1$. Then* $\left|\sum_{u \bmod 2^e} \psi(H(u))\right| \le (12.5)2^{t/(M+1)}\ 2^{e(1-1/(M+1))}$.

Here "$\mathrm{ord}_\ell(H) = 0$" guarantees that t is finite. Note that $\mathcal{C}_H$ is being viewed as a nonzero element of $\mathbb{F}_\ell[T]$, and M is the maximum multiplicity of a zero of $\mathcal{C}_H$ (in $\mathbb{F}_\ell$) that is not a zero of H. To make use of these bounds, we need the following crucial inputs from linear algebra over rings.

Lemma 4.3. *There exists a constant $C := C(F_1, \dots, F_K) > 0$ satisfying all the following:*

(1) *For any prime $\ell > C$ satisfying $(\ell-1, \beta) = 1$, the only tuple $(A_1, \dots, A_K) \in [\ell-1]^K$ for which $\prod_{i=1}^K F_i^{A_i}$ is of the form $c \cdot G^{\ell-1}$ in $\mathbb{F}_\ell[T]$ is $(A_1, \dots, A_K) = (\ell-1, \dots, \ell-1)$.*

(2) *Consider any integer $r \ge 2$, prime ℓ, and $(A_1, \dots, A_K) \in \mathbb{N}^K$ with $\ell \nmid (A_1, \dots, A_K)$. Let $\tau_\ell := \mathrm{ord}_\ell\Big((T^{\varphi(\ell^r)} \prod_{i=1}^K F_i(T)^{A_i})'\Big)$ and $\widetilde{F} := \sum_{i=1}^K A_i F_i' \prod_{j \ne i} F_j$. Then*

$$\tau_\ell \begin{cases} = \mathrm{ord}_\ell(\widetilde{F}) = 0, & \text{if } \ell > C, r \ge 2 \\ = \mathrm{ord}_\ell(\widetilde{F}) \le C, & \text{if } \ell \le C,\ \mathrm{ord}_\ell(F) = 0, \text{ and } r \ge C+2. \end{cases}$$

In both cases above, for any $\theta \in \mathbb{F}_\ell$ for which $\theta F(\theta) \ne 0$, the multiplicity of θ in the polynomials $\ell^{-\tau_\ell}\widetilde{F}$ and $\ell^{-\tau_\ell}(T^{\varphi(\ell^r)} \prod_{i=1}^K F_i(T)^{A_i})'$ are equal.

Proof. Writing $F_i' \prod_{j \ne i} F_j = \sum_{r=0}^{D-1} c_{i,r} T^r$ in $\mathbb{Z}[T]$ for each $i \in [K]$, let $\widetilde{\boldsymbol{\beta}} \in \mathbb{N}$ denote the largest invariant factor of the matrix $\boldsymbol{M_0} := \Big((c_{i,r})_{\substack{1 \le i \le K \\ 0 \le r < D}}\Big)^\intercal$. Fix $\boldsymbol{C}$ **exceeding** $\boldsymbol{2, \widetilde{\beta}}$, the product $\prod_{\boldsymbol{j=1}}^{\boldsymbol{M}} |\mathrm{disc}(\boldsymbol{G_j})| \boldsymbol{\cdot} \prod_{\boldsymbol{1 \le j \ne j' \le M}} |\mathrm{Res}(\boldsymbol{G_j, G_{j'}})|$ (> 0), and (the sizes of) the leading coefficients of all F_i and G_j. (Here $\{G_j\}_{j=1}^M$ are pairwise coprime primitive irreducibles and $F_i = \prod_{j=1}^M G_j^{\mu_{ij}}$, as in the introduction.) We show that any such C satisfies all the assertions in this proposition. Note that for all $\ell > C$, the polynomial $\prod_{j=1}^M G_j$ is separable over $\mathbb{F}_\ell$, i.e. squarefree in $\overline{\mathbb{F}}_\ell[T]$.

Proof of (1). Consider any prime $\ell > C$ and any $(A_1, \dots, A_K) \in [\ell-1]^K$ satisfying $(\ell-1, \beta) = 1$ and $\prod_{i=1}^K F_i^{A_i} = c \cdot G^{\ell-1}$ for some $c \in U_\ell$ and $G \in \mathbb{F}_\ell[T]$. Since $F_i = r_i \prod_{j=1}^M G_j^{\mu_{i,j}}$, we have $c \cdot G^{\ell-1} = \rho \prod_{j=1}^M G_j^{\sum_{i=1}^K \mu_{i,j} A_i}$ for some $\rho \in U_\ell$. Factoring out all leading coefficients and comparing exponents, we get $\sum_{i=1}^K \mu_{i,j} A_i \equiv 0 \pmod{\ell-1}$ for all j. (Here we used the last sentence of the previous paragraph.) This gives the matrix congruence $E_0(A_1 \dots A_K)^\intercal \equiv 0 \pmod{\ell-1}$.

Now considering invertible integer matrices P_0, R_0 for which $P_0 E_0 R_0$ is the Smith normal form $\mathrm{diag}(\beta_1, \dots, \beta_K)$, the last congruence above yields $\mathrm{diag}(\beta_1, \dots, \beta_K) R_0^{-1}(A_1 \dots A_K)^\intercal \equiv 0 \pmod{\ell-1}$. Since $(\ell-1, \beta_K) = 1$, we have $(\ell-1, \beta_1 \cdots \beta_K) = 1$, so that the last congruence forces $R_0^{-1}(A_1 \dots A_K)^\intercal \equiv 0 \pmod{\ell-1} \implies (A_1 \dots A_K)^\intercal \equiv 0 \pmod{\ell-1}$. Hence each $A_i = \ell - 1$.

Proof of (2). Since $(\prod_{i=1}^K F_i^{b_i})' = \sum_{i=1}^K b_i F_i' \prod_{j \ne i} F_j$ for any $(b_i)_{i=1}^K \in \mathbb{Z}^K$, the multiplicative independence of the $(F_i)_{i=1}^K$ forces the polynomials $\Big(F_i' \prod_{j \ne i} F_j\Big)_{i=1}^K$ to be $\mathbb{Q}$–linearly independent. Hence the $D \times K$ matrix $M_0 = \Big((c_{i,r})_{\substack{1 \le i \le K \\ 0 \le r < D}}\Big)^\intercal$ defined above has rank K, its Smith Normal Form has K positive diagonal entries, and its largest invariant factor $\widetilde{\beta}$ is the last diagonal entry.

Now with $d := \mathrm{ord}_\ell(\widetilde{F})$, since ℓ^d divides all the coefficients of $\widetilde{F}(T) = \sum_{i=1}^K A_i F_i'(T) \prod_{j \ne i} F_j(T) = \sum_{r=0}^{D-1} (\sum_{i=1}^K c_{i,r} A_i) T^r$, it follows that $M_0 (A_1 \dots A_K)^\intercal \equiv 0 \pmod{\ell^d}$. An argument analogous to that given for subpart (1) shows that if $\ell^d \nmid \widetilde{\beta}$, then $\ell \mid (A_1, \dots, A_K)$, a contradiction. Thus $\ell^d \mid \widetilde{\beta}$, forcing $\mathrm{ord}_\ell(\widetilde{F}) = d \le v_\ell(\widetilde{\beta}) \le \mathbb{1}_{\ell \le C} \cdot C$. All assertions in subpart (2) now follow easily from the fact that $\Big(T^{\varphi(\ell^r)} \prod_{i=1}^K F_i(T)^{A_i}\Big)' = \varphi(\ell^r) T^{\varphi(\ell^r)-1} \prod_{i=1}^K F_i(T)^{A_i} + T^{\varphi(\ell^r)} \Big(\prod_{i=1}^K F_i(T)^{A_i - 1}\Big) \widetilde{F}(T)$. $\square$

Let $\boldsymbol{\alpha_{\widehat{\chi}} := (\alpha(q)\varphi(q))^{-1} \sum_{u \in U_q} \widehat{\chi}(\widehat{F}(u)) = (\alpha(q)\varphi(q))^{-1} \sum_{u \bmod q} \chi_0(u) \prod_{i=1}^K \chi_i(F_i(u))}$. With B_0 and $D = \sum_{i=1}^K \deg F_i$ in Theorems 1.4 and 1.5, and with C as in the proof of Lemma 4.3,

(4.2) **Fix** $\boldsymbol{C_0 > \max\{B_0, C, (32D)^{2D+2}\}}$, and fix an **integer** $\boldsymbol{\kappa > 100D(DC_0^{2C_0})^{4C_0}}$.

Let $\boldsymbol{Q_0}$ be the largest C_0-smooth $(\kappa+1)$-free divisor of q, i.e. $\boldsymbol{Q_0 = \prod_{\ell \le C_0} \ell^{\min\{v_\ell(q),\kappa\}} \ll 1}$.

Proposition 4.4. *Fix any $\varepsilon > 0$.*
(1) *We have* $\sum_{\substack{\widehat{\chi} \bmod q\\ \mathfrak{f}(\widehat{\chi}) \nmid Q_0}} |\alpha_{\widehat{\chi}}|^N = o(1)$ *as $N \to \infty$, uniformly in $q \in \mathbb{N}$. Moreover, uniformly in $q \in \mathbb{N}$,*

$$\sum_{\widehat{\chi} \bmod q} |\alpha_{\widehat{\chi}}|^N \ll_N \begin{cases} \exp(O((\log q)^{1-1/D})), & \text{for each fixed } N \ge KD+1. \\ q^{K-N/D+\varepsilon}, & \text{for each fixed } N \le KD. \end{cases} \tag{4.3}$$

(2) *The bound (K_r) holds, where uniformly in all $q \in \mathbb{N}$ and $z \in (q, x]$, we have*

$$\xi_r(q;z) \ll \begin{cases} \varphi(q)^{-K} \exp(O((\log q)^{1-1/D})), & \text{for each fixed } r \ge KD+1. \\ q^{-r/D+\varepsilon}, & \text{for each fixed } r \le KD. \\ q^{-1/D_{\min}} \cdot \log_2(3q), & \text{if } r = 1, \text{ where } D_{\min} = \min_i \deg(F_i). \end{cases} \tag{4.4}$$

(3) *The bound (K_r) holds, where uniformly in all* ***squarefree*** *q and $z \in (q, x]$, we have*

$$\xi_r(q;z) \ll \begin{cases} \varphi(q)^{-K} \exp(O(\sqrt{\log q})), & \text{for each fixed } r \ge 2K+1. \\ q^{-r/2+\varepsilon}, & \text{for each fixed } r \le 2K. \\ D_{\min}^{\omega(q)} \cdot \varphi(q)^{-1}, & \text{if } r = 1. \\ \varphi(q)^{-1} \exp(O(\sqrt{\log q})), & \text{if } r = 2,\ K = 1, \text{ and } F_1 \text{ is not squarefull.} \end{cases} \tag{4.5}$$

Proof. We give the argument for $D > 1$; the case $D = 1$ (i.e. when $K = 1$ and F_1 is linear) is much simpler. Define $\boldsymbol{\omega_{>C_0}(m)} := \#\{\ell \mid m : \ell > C_0\}$ and $\boldsymbol{\widetilde{\omega}(m)} := \#\{\ell \mid m :\ \ell \le C_0,\ v_\ell(m) \ge \kappa+1\}$. We claim that for any tuple $\widehat{\chi} = (\chi_1, \dots, \chi_K)$ of Dirichlet characters mod q, we have

$$|\alpha_{\widehat{\chi}}| \ \le\ (DC_0^{C_0+1})^{\widetilde{\omega}(\mathfrak{f}(\widehat{\chi}))} \cdot (4D)^{\omega_{>C_0}(\mathfrak{f}(\widehat{\chi}))} \cdot \mathfrak{f}(\widehat{\chi})^{-1/D}. \tag{4.6}$$

To show this, we start by writing each $\chi_i = \prod_{\ell^e \| q} \boldsymbol{\chi_{i,\ell}}$ for some character $\chi_{i,\ell}$ mod ℓ^e. Then with $\boldsymbol{\alpha_{\widehat{\chi},\ell} := (\alpha(\ell)\varphi(\ell^e))^{-1} \sum_{u \bmod \ell^e} \chi_{0,\ell}(u) \prod_{i=1}^K \chi_{i,\ell}(F_i(u))}$, we have $\alpha_{\widehat{\chi}} = \prod_{\ell^e \| q} \alpha_{\widehat{\chi},\ell}$ by the Chinese Remainder Theorem. (Here $\chi_{0,\ell}$ is the trivial character mod ℓ.) To bound each $\alpha_{\widehat{\chi},\ell}$, let $\boldsymbol{r_\ell := v_\ell(\mathfrak{f}(\widehat{\chi}))}$ so that $\boldsymbol{\ell^{r_\ell}} = \mathrm{lcm}[\boldsymbol{\mathfrak{f}(\chi_{1,\ell}), \dots, \mathfrak{f}(\chi_{K,\ell})}]$. Let $\boldsymbol{\psi_{i,\ell}}$ denote the character mod ℓ^{r_ℓ} inducing $\chi_{i,\ell}$, so that **at least one of $(\boldsymbol{\psi_{1,\ell}, \dots, \psi_{K,\ell}})$ is a primitive character mod** ℓ^{r_ℓ}. It is easy to see that $\alpha_{\widehat{\chi},\ell} = (\alpha(\ell)\varphi(\ell^{r_\ell}))^{-1} \sum_{u \bmod \ell^{r_\ell}} \chi_{0,\ell}(u) \prod_{i=1}^K \psi_{i,\ell}(F_i(u))$.

We first consider the case $\boldsymbol{\ell > 2}$. With ψ_ℓ being a character generating the character group mod ℓ^{r_ℓ}, note that each $\boldsymbol{\psi_{i,\ell} = \psi_\ell^{A_i}}$ for some $\boldsymbol{A_i \in [\varphi(\ell^{r_\ell})]}$. Since some $\psi_{i,\ell}$ is primitive mod ℓ^{r_ℓ},

$$(A_1, \dots, A_K) \not\equiv \begin{cases} 0 \pmod{\ell - 1}, & \text{if } r_\ell = 1 \\ 0 \pmod{\ell}, & \text{if } r_\ell > 1. \end{cases}$$

Writing $\alpha_{\widehat{\chi},\ell} = (\alpha(\ell)\varphi(\ell^{r_\ell}))^{-1} \ \sum_{u \bmod \ell^{r_\ell}} \ \psi_\ell\left(u^{\varphi(\ell^{r_\ell})} \prod_{i=1}^K F_i(u)^{A_i}\right)$, we observe that

$$|\alpha_{\widehat{\chi},\ell}| \le \begin{cases} (4D) \cdot \ell^{-v_\ell(\mathfrak{f}(\widehat{\chi}))/D} & \text{if } \ell > C_0 \\ (DC_0^{C_0+1}) \cdot \ell^{-v_\ell(\mathfrak{f}(\widehat{\chi}))/D}, & \text{if } 2 < \ell \le C_0,\ v_\ell(\mathfrak{f}(\widehat{\chi})) \ge \kappa+1. \end{cases} \tag{4.7}$$

To see this, we apply Proposition 4.1 and Lemma 4.3(1) in the case $\ell > C_0$ and $r_\ell = 1$, recalling that the prime divisors $\ell > C_0 > B_0$ satisfy $\gcd(\ell-1, \beta) = 1$ by hypothesis $IFH(F_1, \dots, F_K; B_0)$. In all other cases, we apply Proposition 4.2(i) on the polynomial $H(T) := T^{\varphi(\ell^{r_\ell})} \prod_{i=1}^K F_i(T)^{A_i}$, noting that Lemma 4.3(2) gives both $\mathrm{ord}_\ell(H') \le r_\ell - 2$ and $\sum_{\theta \in \mathbb{F}_\ell:\ H(\theta)\neq 0} \mathrm{mult}_\theta(\mathcal{C}_H) \le D-1$. Finally, we note that since $\alpha(\ell) \neq 0$, we have $\alpha(\ell) = 1 - \frac{1}{\ell-1}\#\{u \in U_\ell : F_1(u)\cdots F_K(u) \equiv 0 \pmod{\ell}\} \ge 1 - \frac{\min\{\ell-2, D\}}{\ell-1}$, which gives $\alpha(\ell) \ge 1/2$ for $\ell > C_0$, and $\alpha(\ell) \ge 1/C_0$ for $\ell \le C_0$.

Now consider the case $\boldsymbol{\ell = 2}$. Hence we are assuming that $2 \mid q$ (which forces $\alpha(2) = 1$), and that $r_2 = v_2(\mathfrak{f}(\widehat{\chi})) \ge \kappa + 1$. The character group mod 2^{r_2} is generated by the characters ψ, η mod 2^{r_2} defined by $\psi(5) = \exp(2\pi i/2^{r_2-2})$, $\psi(-1) = 1$, $\eta(5) = 1$ and $\eta(-1) = -1$. Since at least one of the characters $(\psi_{i,2})_{i=1}^K$ is primitive mod 2^{r_2}, we can write each $\psi_{i,2} = \psi^{A_i}\eta^{B_i}$, where $A_i \in [2^{r_2-2}]$, $B_i \in [2]$, and $2 \nmid (A_1, \dots, A_K)$. Then $\alpha_{\widehat{\chi},2} = \varphi(2^{r_2})^{-1} \sum_{u \bmod 2^{r_2}} \psi\left(\prod_{i=1}^K F_i(u)^{A_i}\right) \eta\left(u^2 \prod_{i=1}^K F_i(u)^{B_i}\right)$. Defining $g_\lambda(T) = \prod_{i=1}^K F_i(4T+\lambda)^{A_i}$ for $\lambda = \pm 1$, and noting that η has conductor 4, we write

$$\alpha_{\widehat{\chi},2} = \frac{1}{2^{r_2-1}} \sum_{\lambda = \pm 1} \eta\left(\prod_{i=1}^K F_i(\lambda)^{B_i}\right) \cdot \frac{1}{4} \sum_{v \bmod 2^{r_2}} \psi(g_\lambda(v)).$$

If $\eta\left(\prod_{i=1}^K F_i(\lambda)^{B_i}\right) \neq 0$, then $\prod_{i=1}^K F_i(4T+\lambda) \equiv \prod_{i=1}^K F_i(\lambda) \equiv 1 \pmod 2$, which means that

$$t_{2,\lambda} := \mathrm{ord}_2(g'_\lambda) = 2 + \mathrm{ord}_2(\widetilde{F}(4T+\lambda)) \le 2D + \mathrm{ord}_2(\widetilde{F}) \le 2D + C_0 \le \kappa - 3 \le r_2 - 3. \tag{4.8}$$

Here the second inequality uses Lemma 4.3(2) (with $\widetilde{F} = \sum_{i=1}^K A_i F_i' \prod_{j \neq i} F_j$ as defined there), and the first inequality uses the general fact that $\mathrm{ord}_2(W(4T+\lambda)) \le \mathrm{ord}_2(W) + 2\deg(W)$ for any $W \in \mathbb{Z}[T]$. (This follows from a straightforward divisibility argument.) The identity $g'_\lambda(T) = 4\widetilde{F}(4T+\lambda) \prod_{i=1}^K F_i(4T+\lambda)^{A_i - 1}$ gives $\mathrm{mult}_\theta(2^{-t_{2,\lambda}} g'_\lambda) = \mathrm{mult}_\theta(2^{-(t_{2,\lambda}-2)}\widetilde{F}(4T+\lambda)) \le D-1$ for any $\theta \in \mathbb{F}_2$ for which $\theta F(\theta) \neq 0$. An application of Proposition 4.2(ii) and (4.8) now yields

$$|\alpha_{\widehat{\chi},2}| \le (D C_0^{C_0+1}) \cdot 2^{-v_2(\mathfrak{f}(\widehat{\chi}))/D} \quad \text{if} \quad 2 \mid q \quad \text{and} \quad v_2(\mathfrak{f}(\widehat{\chi})) \ge \kappa + 1.$$

Combining this with (4.7) in the factorization $|\alpha_{\widehat{\chi}}| = \prod_{\ell^e \| q} |\alpha_{\widehat{\chi},\ell}|$ establishes our claim (4.6).

Proof of (1). Since $\#\{\widehat{\chi} \bmod q : \mathfrak{f}(\widehat{\chi}) \mid m\} \le m^K$, (4.6) yields for any fixed N,

$$\sum_{\widehat{\chi} \bmod q} |\alpha_{\widehat{\chi}}|^N \ \le\ (D C_0^{C_0+1})^{N C_0} \sum_{m \mid q} \frac{(4D)^{N\omega_{>C_0}(m)}}{m^{N/D-K}} \ \ll_N \prod_{\ell^e \| q} \left(1 + (4D)^{N \mathbb{1}_{\ell > C_0}} \sum_{1 \le v \le e} \frac{1}{\ell^{v(N/D-K)}}\right),$$

where we have noted that $\omega_{>C_0}$ is an additive function. Now if $N \ge KD+1$, the product above is $\ll \exp(O(\sum_{\ell \mid q} \ell^{-1/D})) \le \exp(O(\sum_{\ell \le \omega(q)} \ell^{-1/D})) \le \exp(O((\log q)^{1-1/D}))$. If $N \le KD$, then it is $\ll \prod_{\ell^e \| q}\left(2(4D)^N e^{\mathbb{1}_{N=KD}} \ell^{e(K-N/D)}\right) \le \left(\prod_{\ell^e \| q} e\right)^{\mathbb{1}_{N=KD}} \cdot q^{K-N/D} \exp(O(\omega(q))) \ll_\varepsilon\ q^{K-N/D+\varepsilon}$.

To finish subpart (1), it thus only remains to show its very first assertion. As $N \to \infty$, (4.6) gives

$$\sum_{\substack{\widehat{\chi} \bmod q \\ \mathfrak{f}(\widehat{\chi}) \nmid Q_0}} |\alpha_{\widehat{\chi}}|^N \ \le\ \sum_{\substack{m \mid q \\ m \nmid Q_0}} \frac{(D C_0^{C_0+1})^{N\widetilde{\omega}(m)} \cdot (4D)^{N\omega_{>C_0}(m)}}{m^{N/D-K}} \le \sum_{\substack{m \mid q \\ m \nmid Q_0}} \frac{(D C_0^{C_0+1})^{N\widetilde{\omega}(m)} \cdot (4D)^{N\omega_{>C_0}(m)}}{m^{N/2D}}. \tag{4.9}$$

Since Q_0 is defined to be the largest C_0-smooth $(\kappa+1)$-free divisor of q, any m counted above either: (i) has $P(m) > C_0$ or (ii) has $P(m) \le C_0$ but is not $(\kappa+1)$-free. Let E_1 denote the contribution of all m of type (i) to the rightmost sum in (4.9), and E_2 be that of type (ii). Then

$$E_1 \le \sum_{\ell > C_0} \frac{(4D)^N}{\ell^{N/2D}} \sum_{M|q} \frac{(DC_0^{C_0+1})^{N\widetilde{\omega}(M)} \cdot (4D)^{N\omega_{>C_0}(M)}}{M^{N/2D}}. \tag{4.10}$$

Now $\sum_{\ell>C_0} (4D)^N \ell^{-N/2D} = (4D)^N \sum_{\ell>C_0} \ell^{-N/4D} \cdot \ell^{-N/4D} \le (4DC_0^{-1/4D})^N \sum_{\ell>C_0} \ell^{-5} = o(1)$ as $N \to \infty$. Since $\widetilde{\omega}$ and $\omega_{>C_0}$ are additive functions, the second sum in (4.10) is at most

$$\prod_{\ell|q\,:\,\ell\le C_0} \left(1 + \sum_{1\le v\le\kappa} \ell^{-vN/2D} + (DC_0^{C_0+1})^N \sum_{v>\kappa} \ell^{-vN/2D}\right) \cdot \prod_{\ell|q\,:\,\ell>C_0} \left(1 + (4D)^N \sum_{v\ge1} \ell^{-vN/2D}\right)$$

$$\le \exp\left(\sum_{\ell\le C_0}\sum_{v\le\kappa} \ell^{-5v} + (DC_0^{C_0+1}2^{-\kappa/4D})^N \sum_{\ell\le C_0}\sum_{v\ge0} \ell^{-5v}\right) \cdot \exp\left((4DC_0^{-1/4D})^N \sum_{\ell>C_0}\sum_{v\ge1} \ell^{-5v}\right) \ll 1,$$

where we have recalled (4.2). Collecting all these estimates gives $E_1 = o(1)$ as $N \to \infty$.

On the other hand any m counted in E_2 is C_0-smooth but not $(\kappa+1)$-free. Any such m can be written as $m = Am'$ where $A > 1$ is $(\kappa+1)$-full and m' is $(\kappa+1)$-free. Then $m' \ll_{C_0,\kappa} 1$, giving

$$E_2 \ll (DC_0^{C_0+1})^{NC_0} \sum_{\substack{A|q\,:\,P(A)\le C_0 \\ A>1 \text{ is } (\kappa+1)\text{-full}}} \frac{1}{A^{N/2D}} \le (DC_0^{C_0+1})^{NC_0} \left\{ \prod_{\substack{\ell^e\|q \\ \ell\le C_0}} \left(1 + \sum_{\kappa+1\le v\le e} \frac{1}{\ell^{vN/2D}}\right) - 1 \right\}$$

As $N \to \infty$, this is $\ll (DC_0^{C_0+1})^{NC_0}\{\exp(O(2^{-\kappa N/2D})) - 1\} \ll \left((DC_0^{C_0+1})^{C_0} 2^{-\kappa/2D}\right)^N \ll 50^{-N} = o(1)$. Inserting this and the conclusion on E_1 into (4.9), we get the first assertion of subpart (1).

Proof of (2). Fix any $r \in \mathbb{N}$ and consider any $z \in (q, x]$. We may write

$$\sum_{\substack{P_1,\dots,P_r:\ P_1\cdots P_r\le x \\ P_1>z,\ q<P_r<\cdots<P_1 \\ (\forall i)\ f_i(P_1)\cdots f_i(P_r)\equiv a_i \pmod q}} 1 \le \sum_{(v_1,\dots,v_r)\in\mathcal{V}_{r,K}(q;\widehat{a})} \sum_{\substack{q<P_r<\cdots<P_2\le x \\ (\forall j\ge2)\ P_j\equiv v_j \pmod q}} \sum_{\substack{q<P_1\le x/P_2\cdots P_r \\ P_1\equiv v_1 \pmod q}} 1,$$

with $\mathcal{V}_{r,K}(q;\widehat{a}) := \{(v_1,\dots,v_r) \in U_q^r :\ (\forall i)\ \prod_{j=1}^r F_i(v_j) \equiv a_i \pmod q\}$. Brun–Titchmarsh shows that the innermost sum above is $\ll x/\varphi(q)P_2\cdots P_r \log(z/q)$; combined with partial summation, it also shows that $\sum_{\substack{q<P_j\le x \\ P_j\equiv v_j \pmod q}} 1/P_j \ll \log_2 x/\varphi(q)$ for each $j \in \{2,\dots,r\}$. Consequently

$$\sum_{\substack{P_1,\dots,P_r:\ P_1\cdots P_r\le x \\ P_1>z,\ q<P_r<\cdots<P_1 \\ (\forall i)\ f_i(P_1)\cdots f_i(P_r)\equiv a_i \pmod q}} 1 \ll \frac{\#\mathcal{V}_{r,K}(q;\widehat{a})}{\varphi(q)^r} \cdot \frac{x(\log_2 x)^{r-1}}{\log(z/q)}. \tag{4.11}$$

Hereafter, the first two bounds in (4.4) follow immediately by detecting the congruences $\prod_{j=1}^r F_i(v_j) \equiv a_i \pmod q$ via orthogonality to get

$$\frac{\#\mathcal{V}_{r,K}(q;\widehat{a})}{\varphi(q)^r} = \frac{\alpha(q)^r}{\varphi(q)^K} \sum_{\widehat{\chi} \bmod q} \overline{\chi}_1(a_1)\cdots\overline{\chi}_K(a_K)\ (\alpha_{\widehat{\chi}})^r, \tag{4.12}$$

and then using (4.3). The last bound in (4.4) follows from (4.11), upon noting that $\#\mathcal{V}_{1,K}(q;\widehat{a}) \ll q^{1-1/D_{\min}} \ll q^{-1/D_{\min}}\ \varphi(q)\ \log_2(3q)$ by a result of Konyagin [15, 16].

Proof of (3). A simpler version of the argument for (4.6) yields $|\alpha_{\widehat{\chi}}| \ll (4D)^{\omega_{>C_0}(\mathfrak{f}(\widehat{\chi}))} \cdot \mathfrak{f}(\widehat{\chi})^{-1/2}$ for any $\widehat{\chi}$ modulo squarefree q. Proceeding analogously to the proof of (4.3) to give suitable bounds on $\sum_{\widehat{\chi} \bmod q} |\alpha_{\widehat{\chi}}|^r$ (for any fixed r), and then invoking (4.12) and (4.11), we obtain the first two bounds in (4.5). The third bound in (4.5) follows from the fact that $\#\mathcal{V}_{1,K}(q;\widehat{a}) = \prod_{\ell | q} \#\mathcal{V}_{1,K}(\ell;\widehat{a}) \le (D_{\min})^{\omega(q)}$. (Here we use the Chinese Remainder Theorem for the first equality.)

It only remains to show the fourth and final bound. In the rest of the proof, we thus assume that $r = 2$, $K = 1$ and that $F_1 \in \mathbb{Z}[T]$ is not squarefull. Since $\#\mathcal{V}_{2,1}(q;a) = \prod_{\ell|q} \#\mathcal{V}_{2,1}(\ell;a)$ for any $a \in U_q$, the fourth bound would follow once we show that $\#\mathcal{V}_{2,1}(\ell;a) \le \varphi(\ell)(1 + O(\ell^{-1/2}))$ uniformly in all primes $\ell \gg 1$ and all $a \in U_\ell$. Since F_1 is not squarefull, proceeding as in the first paragraph of the proof of Lemma 4.3, we are guaranteed that F_1 has at least one simple root in $\overline{\mathbb{F}}_\ell$ for all $\ell > C_1$, for some constant $C_1 = C_1(F_1) > 0$.

We claim that $F_1(X)F_1(Y) - w$ is irreducible in $\overline{\mathbb{F}}_\ell[X,Y]$ for all $\ell > C_1$ and all $w \in U_\ell$. Indeed, if $F_1(X)F_1(Y) - w = U(X,Y)V(X,Y)$ with $U, V \in \overline{\mathbb{F}}_\ell[X,Y]$, then for any root $\theta \in \overline{\mathbb{F}}_\ell$ of F_1, we have $-w = U(X,\theta)V(X,\theta)$, forcing $U(X,\theta)$ and $V(X,\theta)$ to be constant. Writing $U(X,Y) = \sum_{0 \le j \le r} u_j(Y)X^j$ and $V(X,Y) = \sum_{0 \le j \le s} v_j(Y)X^j$, we get $u_j(\theta) = v_j(\theta) = 0$ for all $j > 0$, so that $\prod_{\theta \in \overline{\mathbb{F}}_\ell \,:\, F_1(\theta)=0} (Y - \theta) \mid (\gcd_{j>0} u_j(Y), \gcd_{j>0} v_j(Y))$. Now if both $r, s > 0$, then comparing the leading (highest degree) terms in X on both sides of $F_1(X)F_1(Y) - w = U(X,Y)V(X,Y)$ yields $\prod_{\theta \in \overline{\mathbb{F}}_\ell \,:\, F_1(\theta)=0} (Y - \theta)^2 \mid F_1(Y)$ in $\overline{\mathbb{F}}_\ell[Y]$, contradicting that F_1 has a simple root in $\overline{\mathbb{F}}_\ell$. Hence, either $r = 0$ or $s = 0$. Assuming the latter (wlog), we have $V(X,Y) = v_0(Y)$ with $F_1(X)F_1(Y) - w = U(X,Y)v_0(Y)$. Substituting a root of F for X forces $V(X,Y) = v_0(Y)$ to be constant, proving our irreducibility claim.

Finally by the version of the Hasse–Weil bound in [19, Corollary 2b], we get $\#\mathcal{V}_{2,1}(\ell;a) \le \varphi(\ell)(1+O(\ell^{-1/2}))$ uniformly in $\ell > C_1$ and $a \in U_\ell$, as desired. ☐

4.1. **Setting the stage for Theorems 1.4 and 1.5.** With C_0 and κ as in (4.2), define the constant $\boldsymbol{\delta_1 := \delta_1(C_0,\kappa)} < 1$ to be the maximum of $(\alpha(M)\varphi(M))^{-1}|\sum_{u \in U_M} \widehat{\eta}(\widehat{F}(u))|$, taken over all integers $M \le C_0^{C_0\kappa}$ and over all tuples of characters $\widehat{\eta} := (\eta_1, \dots, \eta_K) \bmod M$ for which $\widehat{\eta}(\widehat{F}(u)) = \prod_{i=1}^K \eta_i(F_i(u))$ is not constant on its unit support $\{u \in U_M : F(u) \in U_M\}$. Note that $\delta_1 < 1$.

Let $\boldsymbol{y := \exp((\log x)^{\min\{\epsilon/2, (1-\delta_1)/2\}})}$, where $\boldsymbol{\epsilon}$ **is as in the statement of Theorem 1.4** (with $\boldsymbol{\epsilon := 1}$ **for Theorem 1.5**). For any $\widehat{\chi} \bmod q$ and $Y \ge y$, writing $\sum_{y<p\le Y} \widehat{\chi}(\widehat{f}(p)) = \sum_{u \in U_q} \widehat{\chi}(\widehat{F}(u)) \sum_{\substack{y<p\le Y \\ p \equiv u \pmod q}} 1$ and using Siegel–Walfisz to estimate the inner sum shows that (1.2) holds uniformly in $q \le (\log x)^{K_0}$, with $\boldsymbol{M := \varphi(q)}$, $\boldsymbol{\alpha_{\widehat{\chi}} = (\alpha(q)\varphi(q))^{-1} \sum_{u \in U_q} \widehat{\chi}(\widehat{F}(u))}$ as above, and $\boldsymbol{\mathcal{E}(y) := \exp(-c\sqrt{\log y})}$. ($c > 0$ being a constant coming from Siegel–Walfisz.)

We thus apply Theorem 1.2 with $\boldsymbol{d := Q_0}$ (here Q_0 was defined after (4.2)), $\boldsymbol{z := x^{1/\log_2 x}}$ and $\boldsymbol{J := \lfloor \log_3 x \rfloor}$. Take $\boldsymbol{R}$ **as in Theorem 1.5, and for Theorem 1.4, we take** $\boldsymbol{R := 1}$, noting that the condition "$P_1(n) > q$" (on the left of (1.4)) can be ignored with an error at most $\Psi(x,q) \le \Psi(x,z)$. To bound the right of (1.4), note that $\sum_{\substack{\widehat{\chi} \bmod q \\ \mathfrak{f}(\widehat{\chi}) \nmid Q_0}} |\alpha_{\widehat{\chi}}|^J = o(1)$ by Proposition 4.4(1), and $\Psi(x,z) \le x/(\log x)^{(1+o(1))\log_3 x}$ by known estimates on smooth numbers (see [4, p. 15] or [37, Theorem 5.13 and Corollary 5.19, Chapter III.5]). We use the relevant subparts of Proposition

4.4 to bound the $\xi_r(q;z)$. Finally, note that by [28, Proposition 2.1 and Lemma 2.4], we have

(4.13)

$$\sum_{\substack{n\le X\\ (f(n),q)=1}} 1 \;=\; \frac{X}{(\log X)^{1-\alpha(q)}}\exp(O((\log_2(3q))^{O(1)})), \qquad \sum_{\substack{p\le X\\ (f(p),q)=1}} \frac{1}{p} = \alpha(q)\log_2 X + O((\log_2(3q))^{O(1)}),$$

while $\alpha(q) \gg (\log_2(3q))^{-D}$ by (4.1). Combining all these observations, we deduce that (in the settings of both Theorems 1.4 and 1.5), the right of (1.4) is negligible compared to $o(\varphi(q)^{-K}\#\{n \le x : (f(n),q)=1\})$. Further, the second asymptotic in (1.5) follows from [26, Lemma 2.3] and the first estimate in (4.13). Theorems 1.4 and 1.5 would be thus be proven once we show that

$$\#\{n\le x:\ (f(n),q)=1,\ (\forall i)\ f_i(n)\equiv a_i \pmod{Q_0}\} \sim \varphi(Q_0)^{-K}\#\{n\le x : (f(n),q)=1\}.$$

Now if the radical (product of distinct prime divisors) of q is of size $O(1)$, then this is immediate by applying Theorem 1.3 to the integer $Q^* := \mathrm{lcm}[Q_0, \prod_{\ell\mid q}\ell] \ll 1$. Hence we may assume that q has sufficiently large radical. By orthogonality, (4.1) and the first estimate in (4.13), Theorems 1.4 and 1.5 would follow once we show that: There exists a constant $\delta_0 := \delta_0(C_0,\kappa) < 1$ such that

$$\sum_{n\le x} \mathbb{1}_{(f(n),q)=1}\ \widehat{\psi}(\widehat{f}(n)) \;\ll\; \frac{x}{(\log x)^{1-\delta_0\alpha(q)}} \quad \text{for all nontrivial character tuples } \widehat{\psi} \bmod Q_0,[4] \tag{4.14}$$

uniformly in $\boldsymbol{q \le (\log x)^{K_0}}$ lying in $\boldsymbol{\mathcal{Q}(f_1,\dots,f_K)}$, **having $\boldsymbol{Q_0 > 1}$ and having sufficiently large radical**. (Here for a given q, its "Q_0" is defined as after (4.2).)

Now if $\widehat{\psi}(\widehat{F}(u))$ is **not** constant on $\{u \in U_{Q_0} : F(u) \in U_{Q_0}\}$, then (4.14) follows by applying Proposition 1.6 to the character tuple $\widehat{\chi}$ mod q induced by $\widehat{\psi}$, for which we have $|\alpha_{\widehat{\chi}}| = (\alpha(Q_0)\varphi(Q_0))^{-1}|\sum_{u\in U_{Q_0}}\widehat{\psi}(\widehat{F}(u))| \le \delta_1$ by definition of δ_1. It thus only remains to show (4.14) if $\boldsymbol{\widehat{\psi}(\widehat{F}(u))}$ **is constant on** $\boldsymbol{\{u\in U_{Q_0} : F(u)\in U_{Q_0}\}}$. We do this in the next section.

5. Completing the proof of Theorems 1.4 and 1.5: Modifying the Landau–Selberg–Delange method

5.1. **Perron's formula and analytic continuations.** Recall we are assuming that $\boldsymbol{Q_0 > 1}$, that $\widehat{\psi}(\widehat{F}(u))$ takes a constant value $\boldsymbol{c_{\widehat{\psi}}}$ on its unit support $\{u\in U_{Q_0} : F(u)\in U_{Q_0}\}$, and that the radical $\boldsymbol{Q = \prod_{\ell\mid q}\ell}$ **is sufficiently large**. It suffices to show (4.14) for $x\in\mathbb{N}$, hence we assume this throughout. We will modify the Landau–Selberg–Delange method, borrowing ideas from work of Scourfield [32]. We start by considering the Dirichlet series

$$\mathcal{F}(s) \;=\; \sum_{n\ge1}\frac{\mathbb{1}_{(f(n),q)=1}\ \widehat{\psi}(\widehat{f}(n))}{n^s} \;=\; \sum_{n\ge1}\frac{\mathbb{1}_{(f(n),Q)=1}\ \widehat{\psi}(\widehat{f}(n))}{n^s},$$

and applying Perron's Formula (as stated in [37, Theorem II.2.3]) to write

(5.1)

$$\sum_{n\le x}\mathbb{1}_{(f(n),q)=1}\ \widehat{\psi}(\widehat{f}(n)) \;=\; \frac{1}{2\pi i}\int_{1+\frac{1}{\log x}-iT}^{1+\frac{1}{\log x}+iT}\frac{\mathcal{F}(s)x^s}{s}\,\mathrm{d}s \;+\; O\left(x\sum_{n\ge1}\frac{\mathbb{1}_{(f(n),q)=1}}{n^{1+\frac{1}{\log x}}(1+T|\log(x/n)|)}\right).$$

[4] i.e., for all tuples of characters $\widehat{\psi} = (\psi_1,\dots,\psi_K)$ mod Q_0, except for the tuple of trivial characters mod Q_0

uniformly in $T \geq 1$. The error term here is estimated as in the prime number theorem: First, note that the total contribution of all $n \leq 3x/4$ and $n \geq 5x/4$ to the O-term above is

(5.2)
$$\ll \frac{x}{T}\sum_{n\geq 1}\frac{1}{n^{1+1/\log x}} \ll \frac{x}{T}\exp\left(\sum_{p}\frac{1}{p^{1+1/\log x}}\right) \ll \frac{x}{T}\exp\left(\sum_{p\leq x}\frac{1}{p} + \sum_{j\geq 0}\exp(-2^j)\sum_{x^{2^j}<p\leq x^{2^{j+1}}}\frac{1}{p}\right),$$

which is $\ll x(\log x)/T$. Second, note that for any $n \in (3x/4, x-1]$, we can write $n = x - v$ for some integer $v \in [1, x/4)$, so that $|\log(x/n)| = -\log(1 - v/x) \gg v/x$. As such, the contribution of all such n to the O-term in (5.1) is $\ll T^{-1}\sum_{1\leq v<x/4} x/v \ll x(\log x)/T$, and likewise, so is the contribution of all $n \in [x+1, 5x/4)$. Collecting all estimates, we obtain

$$\sum_{n\leq x}\mathbb{1}_{(f(n),q)=1}\,\widehat{\psi}(\widehat{f}(n)) \;=\; \frac{1}{2\pi i}\int_{1+\frac{1}{\log x}-iT}^{1+\frac{1}{\log x}+iT}\frac{\mathcal{F}(s)x^s}{s}\,\mathrm{d}s \;+\; O\left(\frac{x\log x}{T}\right). \tag{5.3}$$

In what follows, we write a complex number s as $\sigma + it$, where $\sigma = \mathrm{Re}(s)$ and $t = \mathrm{Im}(s)$. Set $\boldsymbol{\mathcal{L}_Q(t) := \log(Q(|t|+1))}$. There exists an absolute constant $\boldsymbol{c_1} > 0$ such that $\prod_{\eta \bmod Q} L(s,\eta)$ has at most one zero $\boldsymbol{\beta_e}$ (the "Siegel zero") in the region $\sigma > 1 - c_1/\mathcal{L}_Q(t)$; moreover, β_e is real and simple and is associated to a real character $\boldsymbol{\eta_e}$ mod Q (the "exceptional character").

We make branch cuts of the complex plane: If $(\alpha(Q), c_{\widehat{\psi}}) \neq (1,1)$, we cut the complex plane along the real line up to $\sigma \leq 1$. If $(\alpha(Q), c_{\widehat{\psi}}) = (1,1)$ and β_e exists then we cut the complex plane along the real line up to $\sigma \leq \beta_e$. In the remaining case, we make no branch cut. The necessary analytic properties of our Dirichlet series $\mathcal{F}(s)$ are given below.

Lemma 5.1. *For $\sigma > 1$, we have $\mathcal{F}(s) = F_{\widehat{\psi}}(s)\cdot G_{\widehat{\psi}}(s)$, where*
(i) $F_{\widehat{\psi}}(s) := \left(\prod_{\eta \bmod Q} L(s,\eta)^{\gamma(\eta)}\right)^{\alpha(Q)c_{\widehat{\psi}}}$, $\gamma(\eta) := (\alpha(Q)\varphi(Q))^{-1}\sum_{u\in U_Q:\ F(u)\in U_Q}\overline{\eta}(u)$,
(ii) $G_{\widehat{\psi}}(s)$ *is analytic on $\sigma > 1 - 1/\log Q$ and satisfies $G_{\widehat{\psi}}(s) \ll (\log_2 Q)^3$ uniformly therein.*
As such, with the aforementioned branch cut conventions, $\mathcal{F}(s)$ analytically continues into the region $\{\sigma + it:\ \sigma > 1 - c_1/\mathcal{L}_Q(t)\} - [1-c_1, 1]$. (Here the real segment $[1-c_1,1]$ has been omitted.)

Proof. We start by using the Euler product of the Dirichlet series $\mathcal{F}(s)$ to write
(5.4)
$$\mathcal{F}(s) = \left(\prod_{p:\ pF(p)\in U_Q}\left(1 - \frac{c_{\widehat{\psi}}}{p^s}\right)^{-1}\right)\left(\prod_p\left(1 + \sum_{v\geq 1}\frac{\mathbb{1}_{(f(p^v),q)=1}\,\widehat{\psi}(\widehat{f}(p^v))}{p^{vs}}\right)\left(1 - \frac{c_{\widehat{\psi}}\,\mathbb{1}_{pF(p)\in U_Q}}{p^s}\right)\right).$$

Expand the logarithm of the first product into Taylor series and split the sum $\sum_{p:\ pF(p)\in U_Q} p^{-s}$ as $\sum_{b\in U_Q:\ F(b)\in U_Q}\sum_{p\equiv b\ (\mathrm{mod}\ Q)} p^{-s}$. By orthogonality and $\log L(s,\eta) = \sum_{p,v\geq 1}\eta(p^v)/vp^{vs}$, we have

$$\sum_{p\equiv b\ (\mathrm{mod}\ Q)}\frac{1}{p^s} \;=\; \frac{1}{\varphi(Q)}\sum_{\eta \bmod Q}\overline{\eta}(b)\log L(s,\eta) \;-\; \sum_{\substack{p,v\geq 2\\ p^v\equiv b\ (\mathrm{mod}\ Q)}}\frac{1}{vp^{vs}}.$$

Hence, the logarithm of the first product in (5.4) equals

$$\alpha(Q)c_{\widehat{\psi}} \sum_{\eta \bmod Q} \gamma(\eta) \log L(s,\eta) \ + \ \left(\sum_{\substack{p, v\geq 2 \\ pF(p)\in U_Q}} \frac{(c_{\widehat{\psi}})^v}{vp^{vs}} - \sum_{\substack{p,v\geq 2 \\ pF(p^v)\in U_Q}} \frac{c_{\widehat{\psi}}}{vp^{vs}} \right). \tag{5.5}$$

The product formula for $\mathcal{F}(s)$ now follows immediately from (5.4).

Next, note that for $\sigma > 2/3$, we have $\sum_{p,v\geq 2} |vp^{vs}|^{-1} \ \leq \ \sum_{p,v\geq 2} p^{-2v/3} \ \ll \ 1$, which also shows that the expression in parentheses in (5.5) defines an analytic function on the region $\sigma > 2/3$. Moreover, the total contribution of all primes $p \nmid Q$ to the second infinite product on the right of (5.4) is equal to $\prod_{p \nmid Q} \left(1 + \sum_{v\geq 2} p^{-vs} \left\{\mathbb{1}_{(f(p^v),Q)=1} \cdot \widehat{\psi}(\widehat{f}(p^v)) - c_{\widehat{\psi}} \cdot \mathbb{1}_{(F(p)f(p^{v-1}),Q)=1} \cdot \widehat{\psi}(\widehat{f}(p^{v-1}))\right\}\right)$, which defines an analytic function of size $O(1)$ for $\sigma > 2/3$. Finally for $\sigma > 1 - 1/\log Q$, the total contribution of all primes $p \mid Q$ to the second infinite product on the right of (5.4) has absolute value at most $\exp(\sum_{p|Q} \sum_{v\geq 1} p^{-v\sigma}) \ll \exp(\sum_{p|Q} p^{-1} \exp(\log p/\log Q)) \ll \exp(3 \sum_{p\leq \omega(Q)} p^{-1}) \ll (\log_2 Q)^3$. Collecting all the observations in this paragraph proves the assertions in (ii). ☐

5.2. **The contour shift.** We assume that $\boldsymbol{\beta_e}$ **exists and that** $\boldsymbol{(\alpha(Q), c_{\widehat{\psi}}) \neq (1,1)}$, otherwise much simpler versions of the argument below suffice. (We summarize the modifications for the other possibilities at the end of this section.) We may thus also assume that

$$\beta_e > 1 - \frac{5}{24} \cdot \frac{c_1}{\log Q}, \tag{5.6}$$

otherwise scaling down c_1 by a constant factor, we are in the simpler case when β_e doesn't exist. By scaling down c_1, we may also assume that the **conductor of** η_e (which is squarefree) **is large enough that it is divisible by a prime exceeding** $D + 2 = \sum_{i=1}^{K} \deg F_i + 2$**.**

Define

$$\boldsymbol{T = \exp(\sqrt{\log x}),\ \beta^* = \frac{2}{3} + \frac{\beta_e}{3}} \ \text{ and } \ \boldsymbol{\vartheta(t) = 1 - \frac{c_1}{4\mathcal{L}_Q(t)}}.$$

Let $\boldsymbol{\delta_e, \delta > 0}$ be any parameters satisfying $\vartheta(0) < \beta_e - 2\delta_e$ and $\beta_e + 2\delta_e < \beta^* < 1 - 2\delta$. (These constraints are only imposed to ensure that Figure 1 below makes sense. We will eventually let $\delta, \delta_e \to 0+$.) We consider the contour $\boldsymbol{\Gamma_1}$ consisting of the following components. (See Figure 1.)

- $\boldsymbol{\Gamma_2}$, the horizontal segment traversed from $\vartheta(T) + iT$ to $(1 + 1/\log x) + iT$.
- $\boldsymbol{\Gamma_3}$, the part of the curve $\vartheta(t) + it$ traversed upwards from $t = 0$ to $t = T$.
- $\boldsymbol{\Gamma_4}$, the horizontal segment traversed from $\beta_e - \delta_e$ to $\vartheta(0)$ **above** the branch cut.
- $\boldsymbol{\Gamma_5}$, the semicircle in the upper half plane with center β_e, radius δ_e, traversed anticlockwise.
- $\boldsymbol{\Gamma_6}$, the horizontal segment traversed from β^* to $\beta_e + \delta_e$ **above** the branch cut.
- $\boldsymbol{\Gamma_7}$, the horizontal segment traversed from $1 - \delta$ to β^* **above** the branch cut.
- $\boldsymbol{\Gamma_8}$, the circle with center at 1, radius δ, traversed anticlockwise as shown in Figure 1.
- $\boldsymbol{\overline{\Gamma}_j}$ **(for** $\boldsymbol{2 \leq j \leq 7}$**)**, the reflection of Γ_j about the real line, directed as in Figure 1.

By Cauchy's integral formula and the final assertion of Lemma 5.1, we may thus rewrite (5.3) as

$$\sum_{n\le x} \mathbb{1}_{(f(n),q)=1}\ \widehat{\psi}(\widehat{f}(n)) \ =\ \frac{1}{2\pi i}\int_{\Gamma_1} \frac{\mathcal{F}(s)x^s}{s}\,\mathrm{d}s \ +\ O\left(\frac{x\log x}{T}\right). \tag{5.7}$$

We define the function

$$H_{\widehat{\psi}}(s) := F_{\widehat{\psi}}(s)\ (s-1)^{\alpha(Q)c_{\widehat{\psi}}}\ (s-\beta_e)^{-\alpha(Q)\gamma(\eta_e)c_{\widehat{\psi}}}, \tag{5.8}$$

which analytically continues into the region $\sigma > 1 - c_1/\mathcal{L}_Q(t)$.

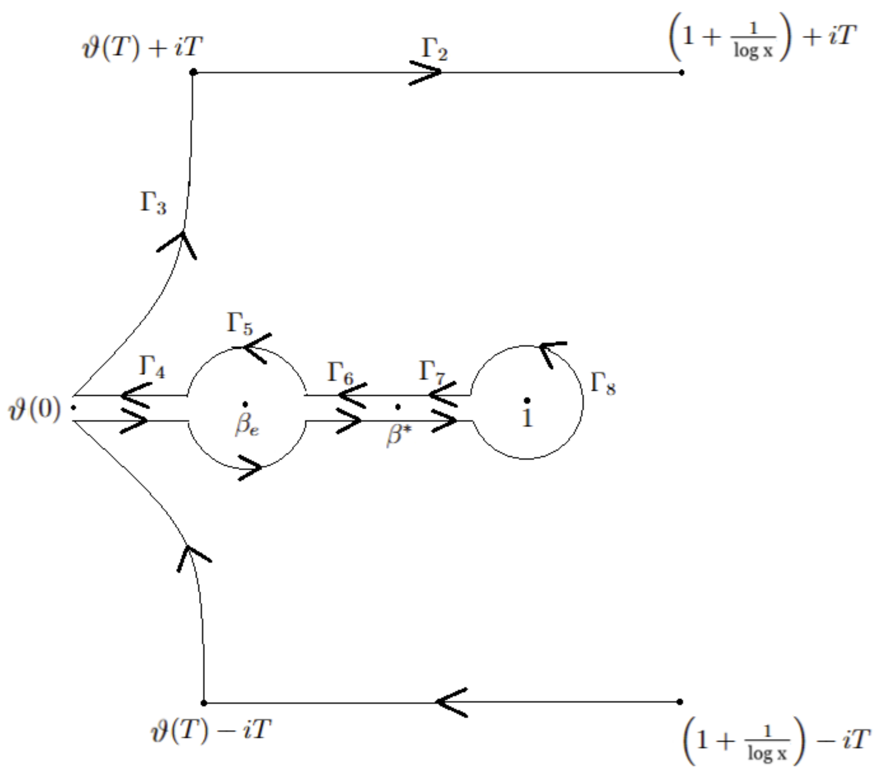

FIGURE 1. The Contour Γ_1 in the case when the Siegel–zero exists and $(\alpha(Q), c_{\widehat{\psi}}) \neq (1,1)$

We now proceed to show that the contributions of all parts of Γ_1, except for Γ_7 and $\overline{\Gamma}_7$, are negligible. The following bounds will be useful for this.

Proposition 5.2.
(1) *We have* $|\mathcal{F}(s)| \ll \log x$ *uniformly in* $s \in \Gamma_2 \cup \Gamma_3 \cup \overline{\Gamma}_2 \cup \overline{\Gamma}_3$.
(2) *For any fixed* $\epsilon > 0$, *we have* $|H_{\widehat{\psi}}(s)| \ll (\log x)^{\epsilon\alpha(Q)/5}$ *uniformly in real* $s \in [1 - c_1/4\log Q, 1]$.

Proof. We start with the following useful observation:

(5.9) We have $H_{\widehat{\psi}}(s) \asymp H_{\widehat{\psi}}(w)$, uniformly in $s, w \in \mathbb{C}$ having $\mathrm{Im}(s) = \mathrm{Im}(w) =: t$ and $1 - c_1/2\mathcal{L}_Q(t) \le \mathrm{Re}(s) \le \mathrm{Re}(w) \le 1 + 100/\mathcal{L}_Q(t)$.

Indeed by Lemmas 15(i) and 9(ii) in [32],[5] we have $|H'_{\widehat{\psi}}(z)/H_{\widehat{\psi}}(z)| = |F'_{\widehat{\psi}}(z)/F_{\widehat{\psi}}(z)+\alpha(Q)c_{\widehat{\psi}}/(z-1) -\alpha(Q)\gamma(\eta_e)c_{\widehat{\psi}}/(z-\beta_e)| \ll \mathcal{L}_Q(\mathrm{Im}(z))$ uniformly in $z \in \mathbb{C}$ having $\mathrm{Re}(z) > 1 - c_1/2\mathcal{L}_Q(\mathrm{Im}(z))$. (To be in the setting of the two lemmas, we take "ξ" there to be $\exp(6\mathcal{L}_Q(t))$, and use the facts that $L(z,\eta) = L(z,\eta^*)\prod_{\ell|Q}(1-\eta^*(\ell)/\ell^z)$ and $\gamma(\eta) = \gamma(\eta^*)$ when $\eta \bmod Q$ is induced by the primitive character η^*.) Now (5.9) follows from $\log|H_{\widehat{\psi}}(w)/H_{\widehat{\psi}}(s)| \le \int_{\mathrm{Re}(s)}^{\mathrm{Re}(w)} |H'_{\widehat{\psi}}(u+it)/H_{\widehat{\psi}}(u+it)|\,\mathrm{d}u$.

Proof of (1). Let $\boldsymbol{\mu(t) := 1 + c_1/7\mathcal{L}_Q(t)}$. Consider any $s \in \Gamma_2 \cup \Gamma_3 \cup \overline{\Gamma}_2 \cup \overline{\Gamma}_3$. By Lemma 5.1(ii), $\mathcal{F}(s) \ll (\log_2 Q)^3|F_{\widehat{\psi}}(s)|$. By (5.9), $|H_{\widehat{\psi}}(s)| \ll |H_{\widehat{\psi}}(\mu(t)+it)|$. Using these with (5.8),

$$\mathcal{F}(s) \ll (\log_2 Q)^3 \cdot |H_{\widehat{\psi}}(\mu(t)+it)| \cdot |s-1|^{-\alpha(Q)\mathrm{Re}(c_{\widehat{\psi}})}\ |s-\beta_e|^{\alpha(Q)\mathrm{Re}(\gamma(\eta_e)c_{\widehat{\psi}})}. \tag{5.10}$$

Since $\mathrm{Re}(\mu(t)+it) > 1$, we can use the Dirichlet series of $\log L(\mu(t)+it,\eta)$ and interchange sums to write $|F_{\widehat{\psi}}(\mu(t)+it)| \le \exp(\sum_{p,r\ge1} p^{-r\mu(t)}|\sum_{\eta \bmod Q} \alpha(Q)\gamma(\eta)\eta(p^r)|)$. By definition of $\gamma(\eta)$,

$$\sum_{\eta \bmod Q} \alpha(Q)\gamma(\eta)\eta(p^r) = \frac{1}{\varphi(Q)} \sum_{\substack{u\in U_Q\\ F(u)\in U_Q}} \sum_{\eta \bmod Q} \overline{\eta}(u)\eta(p^r) = \sum_{\substack{u\in U_Q\\ F(u)\in U_Q}} \mathbb{1}_{p^r\equiv u \bmod Q} = \mathbb{1}_{p^rF(p^r)\in U_Q}.$$

Hence $|F_{\widehat{\psi}}(\mu(t)+it)| \le \exp(\sum_{p,r\ge1} p^{-r\mu(t)}) \ll \exp(\sum_p p^{-\mu(t)}) \ll \zeta(\mu(t)) = 1/(\mu(t)-1)+O(1) \ll \mathcal{L}_Q(t)$. By (5.8), this gives the following bound, which holds uniformly in all $t \in \mathbb{R}$.

$$|H_{\widehat{\psi}}(\mu(t)+it)| \ll \mathcal{L}_Q(t)\ |\mu(t)+it-1|^{\alpha(Q)\mathrm{Re}(c_{\widehat{\psi}})}\ |\mu(t)+it-\beta_e|^{-\alpha(Q)\mathrm{Re}(\gamma(\eta_e)c_{\widehat{\psi}})}. \tag{5.11}$$

Finally, we observe that for both $\theta \in \{1,\beta_e\}$, we have $|s-\theta| \ge c_1/48\mathcal{L}_Q(t)$ for all $s = \sigma+it$ lying on $\Gamma_2\cup\Gamma_3\cup\overline{\Gamma}_2\cup\overline{\Gamma}_3$. (This is immediate for $|t| > 1/100$, and also if $\theta = 1$ and $|t| \le 1/100$. If $\theta = \beta_e$ and $|t| \le 1/100$, this follows from (5.6).) This observation easily yields $|s-\theta| \asymp |\mu(t)+it-\theta|$. Inserting this estimate and (5.11) into (5.10), we get $\mathcal{F}(s) \ll (\log_2 Q)^3\mathcal{L}_Q(t) \ll (\log_3 x)^3\log(QT) \ll \log x$.

Proof of (2). For any real $s \in [1-c_1/4\log Q, 1]$, (5.9) and (5.11) yield $|H_{\widehat{\psi}}(s)| \ll |H_{\widehat{\psi}}(\mu(0))| \ll (\log Q)^3(1-\beta_e)^{-\alpha(Q)}$. To complete the proof, note that $1-\beta_e \gg_\epsilon Q^{-\epsilon/20K_0} \gg (\log x)^{-\epsilon/20}$ for any fixed $\epsilon > 0$ (by Siegel's Theorem), and that $\alpha(Q) \gg (\log_3 x)^{-D}$ by (4.1). ☐

Proposition 5.3. *Set* $I_j := \int_{\Gamma_j} \mathcal{F}(s)x^s/s\,\mathrm{d}s$ *and* $\overline{I}_j := \int_{\overline{\Gamma}_j} \mathcal{F}(s)x^s/s\,\mathrm{d}s$.
(1) $|I_2| + |I_3| + |\overline{I}_2| + |\overline{I}_3| \ll x\exp(-(\log x)^{1/3})$.
(2) $|I_4+\overline{I}_4| + |I_6+\overline{I}_6| \ll x\exp(-(\log x)^{99/100})$.
(3) $\lim_{\delta_e\to0+} I_5 = \lim_{\delta_e\to0+} \overline{I}_5 = 0$. *Moreover,* $\lim_{\delta\to0+} I_8 = 0$.

Proof. **(1)** By Proposition 5.2, we have $|I_2| \ll T^{-1}(\log x)\int_{\vartheta(T)}^{1+1/\log x} x^u\,\mathrm{d}u \ll x/T$. For $s \in \Gamma_3$, note that $|s| \gg t+1$, which gives $|I_3| \ll x^{\vartheta(T)}(\log x)\int_0^T \mathrm{d}t/(t+1) \ll x(\log x)^{3/2}\exp(-c_1\log x/4\mathcal{L}_Q(T)) \ll x\exp(-(\log x)^{1/3})$. The same arguments go through for $\overline{I}_2$ and $\overline{I}_3$.

(2) We only show the bound for $I_4+\overline{I_4}$; the argument for $I_6+\overline{I_6}$ is analogous. Note that for $s \in \Gamma_4$, we have $(s-1)^{-\alpha(Q)c_{\widehat{\psi}}}\cdot(s-\beta_e)^{\alpha(Q)\gamma(\eta_e)c_{\widehat{\psi}}} = |1-s|^{-\alpha(Q)c_{\widehat{\psi}}}\cdot|\beta_e-s|^{\alpha(Q)\gamma(\eta_e)c_{\widehat{\psi}}}\cdot e^{-i\pi\alpha(Q)c_{\widehat{\psi}} + i\pi\alpha(Q)\gamma(\eta_e)c_{\widehat{\psi}}}$,

[5] These two lemmas form an extremely crucial input in our argument, and it is to apply them that we require the full strength of our hypothesis that $\widehat{\psi}(\widehat{F}(u))$ is constant on its unit support.

while for $s \in \overline{\Gamma}_4$, the same equality holds but with the opposite sign of the power of "e". As such, by (5.8), Lemma 5.1 and Proposition 5.2(2), we obtain

$$|I_4 + \overline{I}_4| \ll \int_{\vartheta(0)}^{\beta_e - \delta_e} \frac{|H_{\widehat{\psi}}(u) G_{\widehat{\psi}}(u)| x^u}{u} (1-u)^{-\alpha(Q)\mathrm{Re}(c_{\widehat{\psi}})} \cdot (\beta_e - u)^{\alpha(Q)\mathrm{Re}(\gamma(\eta_e) c_{\widehat{\psi}})} \,\mathrm{d}u$$

$$\ll x^{\beta_e} (\log x)^{\alpha(Q)/200} \cdot (\log_2 Q)^3 \cdot (1-\beta_e)^{-\alpha(Q)} \int_{\vartheta(0)}^{\beta_e - \delta_e} (\beta_e - u)^{\alpha(Q)\mathrm{Re}(\gamma(\eta_e) c_{\widehat{\psi}})} \,\mathrm{d}u. \tag{5.12}$$

Now by our hypothesis in the paragraph following (5.6), the conductor $\mathfrak{f}(\eta_e)$ has a prime factor $\ell_0 > D+2$. As such, factoring $\eta_e = \prod_{\ell | Q} \eta_{e,\ell}$ with $\eta_{e,\ell}$ a character mod ℓ, we observe that

$$|\alpha(Q)\gamma(\eta_e)| = \prod_{\ell | Q} \frac{1}{\varphi(\ell)} \left| \sum_{\substack{u \in U_\ell \\ F(u) \in U_\ell}} \eta_{e,\ell}(u) \right| \le \frac{1}{\ell_0 - 1} \left| - \sum_{\substack{u \in U_{\ell_0} \\ F(u) \equiv 0 \ (\mathrm{mod}\ \ell_0)}} \eta_{e,\ell_0}(u) \right| \le \frac{D}{D+1}. \tag{5.13}$$

Hence the integral in (5.12) is at most $(\beta_e - \vartheta(0))^{1+\alpha(Q)\mathrm{Re}(\gamma(\eta_e)c_{\widehat{\psi}})}/(1+\alpha(Q)\mathrm{Re}(\gamma(\eta_e)c_{\widehat{\psi}})) \ll 1$. The bound on $|I_4 + \overline{I}_4|$ now follows from $1 - \beta_e \gg (\log x)^{-1/200}$ (Siegel's Theorem) and (4.1).

(3) Let $M_{\widehat{\psi}}$ be the maximum of $|H_{\widehat{\psi}}(s)|$ on a fixed small closed disk centered at β_e that is contained in the region $\sigma > 1 - c_1/4\mathcal{L}_Q(t)$. Note that $M_{\widehat{\psi}}$ is finite as $H_{\widehat{\psi}}(s)$ is holomorphic on the region $\sigma > 1 - c_1/\mathcal{L}_Q(t)$. Parametrize the points on the semicircle Γ_5 as $s = \beta_e + \delta_e \cdot e^{i\theta}$ for $-\pi \le \theta \le \pi$. Invoking Lemma 5.1 and (5.8), and arguing as above, we find that

$$|I_5| \ll M_{\widehat{\psi}} (\log_2 Q)^3 \, x^{\beta_e + \delta_e} \cdot (1 - \beta_e - \delta_e)^{-\alpha(Q)} \cdot \delta_e^{1+\alpha(Q)\mathrm{Re}(\gamma(\eta_e)c_{\widehat{\psi}})}$$

By (5.13), the power of δ_e above is at least $1/(D+1)$, hence the above display shows that $\lim_{\delta_e \to 0+} I_5 = 0$. The arguments for $\overline{I}_5$ and I_8 are analogous; for I_8, we only need the fact that $\lim_{\delta \to 0+} \delta^{1-\alpha(Q)\mathrm{Re}(c_{\widehat{\psi}})} = 0$ since $(\alpha(Q), c_{\widehat{\psi}}) \ne (1,1)$ as assumed in the paragraph before (5.6). $\square$

Letting $\delta, \delta_e \to 0+$ and invoking Proposition 5.3, we thus obtain from (5.7),

$$\sum_{n \le x} \mathbb{1}_{(f(n),q)=1} \, \widehat{\psi}(\widehat{f}(n)) = \frac{1}{2\pi i} \int_{\Gamma_0} \frac{\mathcal{F}(s) x^s}{s} \,\mathrm{d}s + O\left(x \exp(-(\log x)^{1/3})\right), \tag{5.14}$$

with Γ_0 being the contour consisting of the two horizontal lines joining the points 1 and β^* above and below the branch cut, directed like Γ_7. (It is easily seen that the integral in (5.14) converges.)

5.3. **Bounding the remaining integral in** (5.14)**.** We start as in the proof of Proposition 5.3(2), by using Lemma 5.1 and (5.8), and then rewriting as a real integral. We get, for any fixed $\epsilon > 0$,

$$\left| \int_{\Gamma_0} \frac{\mathcal{F}(s) x^s}{s} \,\mathrm{d}s \right| \ll \int_{\beta^*}^{1} \frac{|H_{\widehat{\psi}}(u) G_{\widehat{\psi}}(u)| x^u}{u} (1-u)^{-\alpha(Q)\mathrm{Re}(c_{\widehat{\psi}})} \cdot (u - \beta_e)^{\alpha(Q)\mathrm{Re}(\gamma(\eta_e)c_{\widehat{\psi}})} \,\mathrm{d}u$$

$$\ll (\log x)^{\epsilon\alpha(Q)/4} \int_{\beta^*}^{1} |G_{\widehat{\psi}}(u)| \cdot x^u (1-u)^{-\alpha(Q)\mathrm{Re}(c_{\widehat{\psi}})} \,\mathrm{d}u; \tag{5.15}$$

the last line uses Proposition 5.2(2) and $(u - \beta_e)^{\alpha(Q)\mathrm{Re}(\gamma(\eta_e)c_{\widehat{\psi}})} \ll (1-\beta_e)^{-\alpha(Q)} \ll_\epsilon (\log x)^{\epsilon\alpha(Q)/20}$.

Set $\boldsymbol{\epsilon_0 := \epsilon_0(C_0, \kappa) := \max\{\cos(2\pi/r) : 1 < r \le C_0^{C_0\kappa}\}} \in [-1, 1)$ and note that since $1 < Q_0 \le C_0^{C_0\kappa}$ and $c_{\widehat{\psi}}$ is a $\varphi(Q_0)$-th root of unity, we must either have $\mathrm{Re}(c_{\widehat{\psi}}) \le \epsilon_0$ or $c_{\widehat{\psi}} = 1$. In the former case, using Lemma 5.1(ii) and writing $u =: 1 - v/\log x$ yields, from (5.15),

$$\left|\int_{\Gamma_0} \frac{\mathcal{F}(s)x^s}{s}\,\mathrm{d}s\right| \ll \frac{x(\log_3 x)^3}{(\log x)^{1-\alpha(Q)(\epsilon_0+\epsilon/4)}} \cdot \Gamma(1-\alpha(Q)\mathrm{Re}(c_{\widehat{\psi}})) \ll \frac{x}{(\log x)^{1-\alpha(Q)(\epsilon_0+\epsilon/3)}},$$

where Γ is the Gamma-function, and we have noted that Γ is of size $O(1)$ on the interval $[1-\epsilon_0, 1]$. Inserting the above bound into (5.14) proves (4.14) when $c_{\widehat{\psi}} \neq 1$ (by taking, say, $\delta_0 \leq 1-\epsilon_0$.)

Now assume that $c_{\widehat{\psi}} = 1$, so that $\widehat{\psi}(\widehat{F}(u)) = 1$ for all $u \in U_{Q_0}$ satisfying $F(u) \in U_{Q_0}$. Since $q \in \mathcal{Q}(f_1, \dots, f_K)$, applying the definition of $\mathcal{Q}(f_1, \dots, f_K)$ to the characters mod q induced by $\widehat{\psi} = (\psi_1, \dots, \psi_K)$, we find that $\mathbb{1}_{(f(2^j),q)=1}\ \widehat{\psi}(\widehat{f}(2^j)) = -1$ for all $j \geq 1$. Isolating the contribution of $p = 2$ from the second infinite product in (5.4), we can thus write $G_{\widehat{\psi}}(s) = G_{\widehat{\psi},1}(s)G_{\widehat{\psi},2}(s)$, where $G_{\widehat{\psi},2}(s) := (1 + \sum_{j\geq 1} \mathbb{1}_{(f(2^j),q)=1}\ \widehat{\psi}(\widehat{f}(2^j))/2^{js})\ (1 - c_{\widehat{\psi}}\ \mathbb{1}_{(2F(2),Q)=1}/2^s)$ is analytic on $\sigma > 0$ and satisfies $G_{\widehat{\psi},2}(1) = 0$, and where the function $G_{\widehat{\psi},1}(s)$ is also analytic on the region $\sigma > 1-1/\log Q$ and satisfies $|G_{\widehat{\psi},1}(s)| \ll (\log_2 Q)^3$ uniformly therein. (The assertions on $G_{\widehat{\psi},1}$ can be shown by following the arguments in Lemma 5.1.) Hence for all $u \in [\beta^*, 1]$, we have

(5.16)

$$|G_{\widehat{\psi}}(u)| \ll (\log_2 Q)^3 \cdot |G_{\widehat{\psi},2}(1) - G_{\widehat{\psi},2}(u)| \ll (\log_3 x)^3 \int_u^1 |G'_{\widehat{\psi},2}(w)|\,\mathrm{d}w \ll (\log_3 x)^3(1-u).$$

Inserting this bound into (5.15) and making the substitution $u =: 1 - v/\log x$, we obtain

$$\left|\int_{\Gamma_0} \frac{\mathcal{F}(s)x^s}{s}\,\mathrm{d}s\right| \ll \frac{x(\log_3 x)^3}{(\log x)^{2-\alpha(Q)(1+\epsilon/4)}} \cdot \Gamma(2-\alpha(Q)) \ll \frac{x}{(\log x)^{1-\alpha(Q)\epsilon/3}}, \tag{5.17}$$

establishing (4.14) in the remaining case $c_{\widehat{\psi}} = 1$.

To conclude, we summarize the changes if either β_e doesn't exist or if $(\alpha(Q), c_{\widehat{\psi}}) = (1, 1)$. If β_e doesn't exist and $(\alpha(Q), c_{\widehat{\psi}}) \neq (1, 1)$, then we replace the semicircles Γ_5 and $\overline{\Gamma}_5$ by straight segments above and below the branch cut (there is no δ_e). If $(\alpha(Q), c_{\widehat{\psi}}) = (1, 1)$ and β_e exists, then (as mentioned in the paragraph preceding Lemma 5.1), we only make a branch cut up to $\sigma \leq \beta_e$, so we can replace all of Γ_j and $\overline{\Gamma}_j$ (for $j \geq 5$) by a circle of radius δ_e centered at β_e, going from below to above the branch cut; we eventually let $\delta_e \to 0+$. If $(\alpha(Q), c_{\widehat{\psi}}) = (1, 1)$ and β_e doesn't exist either, then there is no branch cut, and our contour Γ_1 just consists of Γ_2, Γ_3, $\overline{\Gamma}_2$ and $\overline{\Gamma}_3$. By the discussion preceding (4.14), this completes the proofs of Theorems 1.4 and 1.5. ☐

6. Optimalities in Theorems 1.4 and 1.5

Lemma 6.1. *Let $f_1, \dots, f_K : \mathbb{N} \to \mathbb{Z}$ be multiplicative functions and $\{F_i\}_i \subset \mathbb{Z}[T]$ be nonconstant such that $\prod_{i=1}^K F_i$ is squarefree, and $f_i(p) = F_i(p)$ for all i and all primes p. There exists a constant $C_{\widehat{F}} > 0$ depending only on $\{F_i\}_i$ such that any $q \in \mathbb{N}$ with $P^-(q) > C_{\widehat{F}}$ lies in $\mathcal{Q}(f_1, \dots, f_K)$.*

Proof. If $P^-(q) > D+1$, then $\alpha(q) = \prod_{\ell \mid q}(1 - (\ell-1)^{-1}\#\{u \in U_\ell : \prod_{i=1}^K F_i(u) \equiv 0 \pmod{\ell}\}) \geq \prod_{\ell \mid q}(1 - D/(\ell-1)) > 0$. Moreover, if $\mathcal{T}(q) := \{(F_1(u), \dots, F_K(u)) : u\prod_{i=1}^K F_i(u) \in U_q\}$ generates U_q^K, then $q \in \mathcal{Q}(f_1, \dots, f_K)$ vacuously, i.e. there is no tuple of characters $\widehat{\chi} \neq (\chi_0, \dots, \chi_0)$ mod q for which $\prod_{i=1}^K \chi_i(F_i(u)) = 1$ on $\{u \in U_q : \prod_{i=1}^K F_i(u) \in U_q\}$. Hence, it suffices to show that there exists a constant $C_{\widehat{F}}$ such that $\mathcal{T}(q)$ generates U_q^K for all q with $P^-(q) > C_{\widehat{F}}$.

Now under the isomorphism $U_q^K \cong \prod_{\ell^e \| q} U_{\ell^e}^K$, the set $\mathcal{T}(q)$ maps to $\prod_{\ell^e \| q} \mathcal{T}(\ell^e)$. Thus if $\mathcal{T}(q)$ doesn't generate U_q^K, then by [24, Lemma 5.13], there is some $\ell^e \| q$ and characters $(\psi_1, \dots, \psi_K)$ mod ℓ^e, not all trivial, for which $\prod_{i=1}^K \psi_i(F_i(u))$ is constant on $\{u \in U_{\ell^e} : \prod_{i=1}^K F_i(u) \in U_{\ell^e}\}$. The lemma now follows from [22, Lemma 5]. □

We show that the ranges of q in Theorem 1.4 are essentially optimal, in that "$1-\epsilon$" cannot be replaced by "$1+\epsilon$" in either subpart. In all our examples below, $\{F_i\}_{i=1}^K$ will be nonconstant and $\prod_{i=1}^K F_i$ will be separable, – guaranteeing that $\{F_i\}_{i=1}^K$ are multiplicatively independent, and that any $q \in \mathbb{N}$ satisfies $IFH(F_1, \dots, F_K; B_0)$ for any $B_0 > 0$. Fix $r \in \mathbb{N}$. Our K_0 will be fixed large enough in terms of $F_1, \dots, F_K$, and we will always have $q \le (\log x)^{K_0}$.

Optimality in Theorem 1.4(i). Let $F_i(T) := (T-1)^r + i$, and q be a perfect r-th power with $P^-(q) > \max\{C_{\widehat{F}}, 2K\}$ where $C_{\widehat{F}}$ is as in Lemma 6.1. Then $q \in \mathcal{Q}(f_1, \dots, f_K)$ and any prime $P \equiv 1 \pmod{q^{1/r}}$ satisfies $f_i(P) = F_i(P) \equiv i \pmod q$. By Siegel–Walfisz, $\#\{n \le x : (\forall i)\ f_i(n) \equiv i \pmod q\} \gg x/(q^{1/r}\log x)$ uniformly in $q \le (\log x)^{K_0}$. The last expression grows faster than the expected main term $\varphi(q)^{-K}\#\{n \le x : (f(n), q) = 1\}$, once $q > (\log x)^{(1+\epsilon)\alpha(q)(K-1/r)^{-1}} = (\log x)^{(1+\epsilon)\alpha(q)(K-1/D_{\min})^{-1}}$ for any fixed ϵ. (To see this, use (4.1) and the first estimate in (4.13).)

Stronger optimality in Theorem 1.4(i) for $\boldsymbol{K > 1, D_{\min} = 1}$ (i.e., at least one of the F_i is linear). We show that in this case, the range of q in Theorem 1.4(i) is optimal in a much stronger sense: It is optimal even if **all** the F_i are **any** pairwise coprime linear polynomials. Indeed, if $F_i(T) = c_iT + b_i$, then fixing any $b \in \mathbb{Z} \setminus \{0\}$ with $\prod_{i=1}^K F_i(b) \ne 0$, and taking any q with $P^-(q) > \max\{C_{\widehat{F}}, 1 + |b\prod_{i=1}^K F_i(b)|\}$, we note that any prime $P \equiv b \pmod q$ satisfies $f_i(P) \equiv F_i(b) \pmod q$. By Siegel–Walfisz, $\#\{n \le x : (\forall i)\ f_i(n) \equiv F_i(b) \pmod q\} \gg x/\varphi(q)\log x$. This expression grows faster than the expected main term as soon as $q > (\log x)^{(1+\epsilon)\alpha(Q)/(K-1)}$.

Optimality in Theorem 1.4(ii). Let $F_i(T) := \prod_{j=1}^r (T-2j) + 2(2i-1)$. Each F_i is irreducible by Eisenstein's criterion at 2, hence $\prod_{i=1}^K F_i$ is separable. Now let $q \le (\log x)^{K_0}$ be a **squarefree** integer having $P^-(q) > \max\{C_{\widehat{F}}, 4Kr\}$. By the Chinese Remainder Theorem, the congruence $\prod_{j=1}^r (v-2j) \equiv 0 \pmod q$ has exactly $r^{\omega(q)}$ distinct solutions $v \in U_q$. Hence by Siegel–Walfisz,

$$\sum_{\substack{n \le x \\ (\forall i)\ f_i(n) \equiv 2(2i-1) \pmod q}} 1 \ \ge \sum_{\substack{n \le x:\ P(n) > q \\ (\forall i)\ f_i(n) \equiv 2(2i-1) \pmod q}} 1 \ \ge \sum_{\substack{q < P \le x \\ \prod_{j=1}^r (P-2j) \equiv 0 \pmod q}} 1 \ \gg \frac{r^{\omega(q)} x}{\varphi(q)\log x}. \tag{6.1}$$

By (4.13) and (4.1), the rightmost expression grows faster than $\varphi(q)^{-K}\#\{n \le x : (f(n), q) = 1\}$ as soon as $q^{K-1}D_{\min}^{\omega(q)} = q^{K-1}r^{\omega(q)} > (\log x)^{(1+\epsilon)\alpha(q)}$ for any fixed $\epsilon > 0$.

For completeness, we construct arbitrarily large squarefree q satisfying the last inequality; in fact, we ensure that $r^{\omega(q)} > (\log x)^{(1+\epsilon)\alpha(q)}$. Indeed, let $q := \prod_{\max\{C_{\widehat{F}}, 4Kr\} < \ell \le Y} \ell$, with $Y \le (K_0/2)\log_2 x$ a parameter to be chosen later. Then $q \le (\log x)^{K_0}$, $\omega(q) \ge Y/2\log Y$, and by the Chinese Remainder Theorem and the Prime Ideal Theorem, we have $\alpha(Q) \le c/\log Y$ for some constant $c > 0$ depending only on $F_1, \dots, F_K$. The inequality $r^{\omega(q)} > (\log x)^{(1+\epsilon)\alpha(q)}$ is then ensured once we choose any $Y \in (4c\log_2 x/\log r, (K_0/2)\log_2 x)$. (We can fix $K_0 > 16c$ at the start.)

Optimality in Theorem 1.5. Note that by the $K = 1$ case of (6.1), the third value of R in Theorem 1.5 is optimal, in that it cannot be reduced to 1. We now show that the second value of R is "nearly optimal", in that it cannot be reduced to $2K-1$. With $F_i(T) = \prod_{j=1}^r (T-2j) + 2(2i-1)$

as above, we consider multiplicative functions $f_1, \dots, f_K : \mathbb{N} \to \mathbb{Z}$ satisfying $f_i(p) = F_i(p)$ and $f_i(p^2) = 1$ for all i and all primes p. Consider any squarefree q as in the previous paragraph, and let $n = (p_1 \cdots p_{K-1})^2 P$ where p_i, P are primes satisfying $q < p_{K-1} < \cdots < p_1 < x^{1/4K} < x^{1/3} < P$ and $\prod_{j=1}^r (P - 2j) \equiv 0 \pmod q$. Then $f_i(n) \equiv 2(2i-1) \pmod q$, so that by Siegel–Walfisz,

$$\sum_{\substack{n \le x:\ P_{2K-1}(n) > q \\ (\forall i)\ f_i(n) \equiv 2(2i-1) \ (\mathrm{mod}\ q)}} 1 \gg \frac{r^{\omega(q)} x}{\varphi(q) \log x} \sum_{\substack{p_1,\ \dots,\ p_{K-1} \in (q, x^{1/4K}) \\ p_1,\ \dots,\ p_{K-1} \text{ distinct}}} \frac{1}{(p_1 \cdots p_{K-1})^2}, \tag{6.2}$$

where we have replaced the ordering condition on $p_1, \dots, p_{K-1}$ by a distinctness condition at the cost of $(K-1)!$. Note that $\sum_{p_1,\ \dots,\ p_{K-1} \in (q, x^{1/4K})} (p_1 \cdots p_{K-1})^{-2} \gg (\sum_{q < p \le x^{1/4K}} p^{-2})^{K-1} \gg (q \log q)^{-(K-1)}$. On the other hand, the sum of $1/p_1 \cdots p_{K-1}$ over all primes $p_1, \dots, p_{K-1} > q$ when $p_i = p_j$ for some $i \ne j$, is $\ll (\sum_{p > q} p^{-4}) \cdot (\sum_{p > q} p^{-2})^{K-3} \ll q^{-K}$. Combining, we get

$$\sum_{\substack{n \le x:\ P_{2K-1}(n) > q \\ (\forall i)\ f_i(n) \equiv 2(2i-1) \ (\mathrm{mod}\ q)}} 1 \gg \frac{r^{\omega(q)}}{\varphi(q)(q \log q)^{K-1}} \cdot \frac{x}{\log x} \gg \frac{r^{\omega(q)}}{\varphi(q)^K (\log_2 x)^K} \cdot \frac{x}{\log x}, \tag{6.3}$$

where we used $q \ll \varphi(q) \log_2 q$. The last expression above grows faster than $\varphi(q)^{-K} \#\{n \le x : (f(n), q) = 1\}$ once $r^{\omega(q)} > (\log x)^{(1+\epsilon)\alpha(q)}$ for any fixed $\epsilon > 0$, which is already satisfied by our q.

7. Generalizing Theorem 1.2 and uniformizing Narkiewicz's general criteria

The hypothesis "$\alpha(q) \ne 0$" played a key role in Theorem 1.3 and its uniform analogues Theorems 1.4 and 1.5: Not only did it prevent the input sets from being too sparse (via (4.13)), but also guaranteed that the "polynomial–type" control on the f_i at the *primes* was most significant in a certain sense. It turns out that if $\alpha(q) = 0$, then the behavior of the f_i at a *higher prime power* becomes more significant. Narkiewicz's criterion in [22] allows for this generality, and we obtain best possible uniform analogues of his general criterion as well. To this end, we start with

7.1. **A generalization of Theorems 1.1 and 1.2 for sparse input sets.** Fix $\boldsymbol{\nu} \in \mathbb{N}$ and $\boldsymbol{C > 0}$. We say that multiplicative functions $f_1, \dots, f_K : \mathbb{N} \to \mathbb{Z}$ are **$\boldsymbol{\nu}$-supported** if the ν-free part of any n satisfying $(f(n), q) = 1$ is at most C. Equivalently, any such $n = n_1 n_2$, where $n_1 \le C$ and n_2 is ν-free. Note that if $f_1, \dots, f_K$ are ν-supported, then they are also $(\nu - 1)$-supported. In particular, any $f_1, \dots, f_K$ are 1-supported. The general paradigm in this entire section is that the control of the f_i at the ν-th powers of primes is most significant: All hypotheses, results and arguments are modified accordingly. In particular, taking $\nu = 1$ here recovers everything before.

In what follows, let $\boldsymbol{\Psi_\nu(x, z)}$ denote the **number of $\boldsymbol{\nu}$-full $\boldsymbol{z}$-smooth numbers up to $\boldsymbol{x}$**; we give a bound for this in Lemma 7.2. We **redefine $\boldsymbol{\xi_r(q; z)}$** > 0 to be any parameter satisfying

$$\begin{aligned} &\#\{P_1 \cdots P_r \le x : P_i \text{ primes},\ P_1 > z,\ q < P_r < \cdots < P_1,\ (\forall i)\ f_i(\boldsymbol{P_1^\nu}) \cdots f_i(\boldsymbol{P_r^\nu}) \equiv a_i \pmod q\} \\ &\qquad\qquad \le \xi_r(q; z) \cdot x (\log\log x)^r \Big/ \log(z/q). \end{aligned} \tag{7.1}$$

The following theorem generalizes Theorems 1.1 and 1.2; the differences have been highlighted.

Theorem 7.1. *Let $f_1, \dots, f_K : \mathbb{N} \to \mathbb{Z}$ be* ***ν-supported*** *multiplicative functions, q a positive integer* ***greater than C****, and $y \ge q$ a parameter such that for all $\widehat{\chi}$ mod q and all $Y \ge y$, we have*

$$\sum_{y \le p \le Y} \widehat{\chi}(\widehat{f}(\boldsymbol{p^\nu})) \;=\; \alpha_{\widehat{\chi}} \sum_{y \le p \le Y} \chi_0(f(\boldsymbol{p^\nu})) + O(MY\mathcal{E}(y)) \tag{7.2}$$

for some $M \ge 1$, some $\alpha_{\widehat{\chi}}$ in the unit disk, and some decreasing function $\mathcal{E} : \mathbb{R}^+ \to \mathbb{R}^+$.

Then for all d dividing q, all $x, z > 0$, and all $J, R \in \mathbb{N}$ satisfying $y \le z \le x$ and $J \ge R$, we have

$$\sum_{\substack{n \le x:\ P_R(n) > q \\ (\forall i)\ f_i(n) \equiv a_i \pmod q}} 1 \;-\; \left(\frac{\varphi(d)}{\varphi(q)}\right)^K \sum_{\substack{n \le x:\ (f(n),q)=1 \\ (\forall i)\ f_i(n) \equiv a_i \pmod d}} 1 \tag{7.3}$$

$$\ll\; \frac{1}{\varphi(q)^K} \sum_{\substack{\widehat{\chi} \bmod q \\ \mathfrak{f}(\widehat{\chi}) \nmid d}} |\alpha_{\widehat{\chi}}|^J \sum_{\substack{n \le x \\ (f(n),q)=1}} 1 \;+\; \boldsymbol{\Psi_\nu(x,z)} \;+\; J\left(\frac{\boldsymbol{x}}{\boldsymbol{y}}\right)^{\boldsymbol{1/\nu}} \;+\; M\boldsymbol{x^{1/\nu}}\mathcal{E}(y)\log x$$

$$+ \left\{ \left(\frac{\varphi(d)}{\varphi(q)}\right)^K + \boldsymbol{\xi_{\lceil R/\nu \rceil}(q;z)} + R^K \sum_{\substack{\boldsymbol{1 \le r < R/\nu} \\ \boldsymbol{1 \le s < K\nu}}} \frac{\xi_r(q;z)}{q^{\boldsymbol{\max\{\frac{s}{\nu}, \frac{R}{\nu} - r - s\}}}} \right\} \frac{\boldsymbol{x^{1/\nu}}(2\log_2 x)^{R+J}}{\log(z/q)} \exp\left(\sum_{\substack{p \le y \\ \boldsymbol{(f(p^\nu),q)=1}}} \frac{1}{p} \right).$$

The analogue of (1.3) *also holds, with the relevant replacements on the right hand side, and with the condition "$P_J(n) > y$" replaced by "$\boldsymbol{P_{J\nu}(n) > y}$". The implied constants in both estimates depend only on $\boldsymbol{\nu, K, C}$ and the implied constant in* (7.2).

Proof. The proofs are essentially analogous, so we mention the main changes. First, we define n to be "convenient" if the J largest **distinct** prime factors of n all exceed y and each appear exactly to the ν-th power in n. Thus, any convenient n can be uniquely written as $m(P_J \cdots P_1)^\nu$ where (2.1) holds. Once again, to prove the analogue of (1.3), it suffices to show the corresponding estimate with the condition "$P_{J\nu}(n) > y$" replaced by "n convenient": This is because any inconvenient n having $P_{J\nu}(n) > y$ must be divisible by the $(\nu+1)$-th power of a prime exceeding y, and because

$$\sum_{\substack{n \le x:\ (f(n),q)=1 \\ \exists\ p > y \text{ s.t. } p^{\nu+1} \mid n}} 1 \;\le\; \sum_{\substack{p > y \\ r \ge \nu+1}} \sum_{\substack{m \le x/p^r \\ (f(m),q)=1}} 1 \;\le\; \sum_{\substack{p > y \\ r \ge \nu+1}} \sum_{m_1 \le C} \sum_{\substack{m_2 \le x/p^r \\ m_2 \text{ is } \nu\text{-full}}} 1 \;\ll\; x^{1/\nu} \sum_{\substack{p > y \\ r \ge \nu+1}} \frac{1}{p^{r/\nu}} \;\ll\; \left(\frac{x}{y}\right)^{1/\nu}, \tag{7.4}$$

where we've used the Erdős–Szekeres estimate [12] on ν-full numbers. Now (2.2) and (2.3) hold as stated. The analogues of (2.4) and (2.5) hold with all x/m and $\widehat{\chi}(\widehat{f}(P_i))$ replaced by $x^{1/\nu}/m^{1/\nu}$ and $\widehat{\chi}(\widehat{f}(P_i^\nu))$; the O-term in (2.5) is replaced by $O\left(J + Mx^{1/\nu}\mathcal{E}(y)/m^{1/\nu}P_1 \cdots P_{i-1}P_{i+1} \cdots P_J\right)$. Reversing the splitting of convenient n, we see that this O-term summed over m and all other P_i is $\ll \sum_{n \le x/y^\nu:\ (f(n),q)=1} 1 + J^{-1}Mx^{1/\nu} \sum_{n \le x/y^\nu:\ (f(n),q)=1} n^{-1/\nu}$. The first of the two sums is $\ll \sum_{n_1 \le C} \sum_{\substack{n_2 \le x/y^\nu \\ n_2 \text{ is } \nu\text{-full}}} 1 \ll x^{1/\nu}/y$. The second of the two sums is at most

$$\sum_{\substack{n \le x \\ (f(n),q)=1}} \frac{1}{n^{1/\nu}} \le \sum_{n_1 \le C} \frac{1}{n_1^{1/\nu}} \sum_{\substack{n_2 \le x \\ n_2 \text{ is } \nu\text{-full}}} \frac{1}{n_2^{1/\nu}} \ll \prod_{p \le x} \left(1 + \sum_{r \ge \nu} \frac{1}{p^{r/\nu}}\right) \ll \exp\left(\sum_{p \le x} \frac{1}{p}\right) \ll \log x. \tag{7.5}$$

We thus find that the analogue of (2.6) holds with all products $P_1 \cdots P_J$ replaced by $(P_1 \cdots P_J)^\nu$, and with all instances of x in the O-term replaced by $x^{1/\nu}$. This proves the analogue of (1.3).

Next to obtain (7.3), we show the analogue of (2.7), which is

$$\sum_{\substack{n\le x:\ P_{J\nu}(n)\le y\\ (f(n),q)=1}} 1 \ll \frac{x^{1/\nu}(2\log_2 x)^J}{\log z}\ \exp\left(\sum_{p\le y}\frac{\mathbb{1}_{f(p^\nu),q)=1}}{p}\right) + \Psi_\nu(x,z) + \left(\frac{x}{y}\right)^{1/\nu}. \tag{7.6}$$

By (7.4), we may assume that the y-rough part of n is $(\nu+1)$-free, so that (by definition of ν-supported and $y \ge q > C$) the y-rough part must be a perfect ν-th power. We may also assume that $P(n) > z$ since $\#\{n \le x : P(n) \le z,\ (f(n),q) = 1\} \le \sum_{n_1\le C}\sum_{\substack{n_2\le C:\ P(n_2)\le z\\ n_2 \text{ is } \nu \text{ full}}} 1 \ll \Psi_\nu(x,z)$. Under these two assumptions, $n = BA^\nu P^\nu$, with $P = P(n) > z$, $P(B) \le y < P^-(A)$, $\Omega(A) \le J$ and $(f(B),q) = 1$. Proceeding as in the proof of (2.7), and noting that $\sum_{B:\ P(B)\le y} \mathbb{1}_{(f(B),q)=1}/B^{1/\nu} \ll \exp(\sum_{p\le y} \mathbb{1}_{(f(p),q)=1}/p)$ by the method in (7.5), establishes (7.6).

Finally, we bound the n satisfying $P_R(n) > q$, $P_{J\nu}(n) \le y$ and $f_i(n) \equiv a_i \pmod q$. By the above reductions, it suffices to bound the total contribution $\sum^*_{n\le x} 1$ of all $n \le x$ having a $(\nu+1)$-free y-rough part and satisfying $P_R(n) > q$, $P_{J\nu}(n) \le y$, $P(n) > z$, and $f_i(n) \equiv a_i \pmod q$. Let $E_{r,s}$ be the number of n counted in $\sum^*_{n\le x} 1$ with $\#\{p > q : p^\nu \parallel n\} = r$ and $\#\{p > q : p^{\nu+1} \mid n\} = s$. The analogue of (2.8) holds with "R", "K", "$p \parallel n$" and "$p^2 \mid n$" replaced by "R/ν", "$K\nu$", "$p^\nu \parallel n$" and "$p^{\nu+1} \mid n$" respectively. (Here we note that the q-rough part of n is ν-full as $q > C$ and $f_1, \dots, f_K$ are ν-supported.) Any n counted in $E_{r,s}$ is of the form $mp_1^{c_1}\cdots p_s^{c_s}\boldsymbol{A^\nu}$, where $P(m) \le q$, $p_j > q$, $c_j \ge \nu+1$, $c_1+\dots+c_s \ge R-\nu r$, and A is squarefree with $P^-(A) > q$, $P(A) > z$, $\Omega(A) = r$. Now proceed as in the rest of section 2, using the method in (7.5) whenever appropriate. ☐

The following lemma gives a general uniform bound on $\Psi_\nu(x,z)$, the number of ν-full z-smooth numbers up to x. This generalizes some known bounds on smooth numbers, such as in [4, p. 15].

Lemma 7.2. *Fix $\nu \in \mathbb{N}$. As $x, z \to \infty$, we have*

$$\Psi_\nu(x,z) \ll x^{1/\nu}(\log z)\exp(-u\nu^{-1}\log u + O(u\log_2(3u))),$$

uniformly for $(\log x)^{\max\{3,2\nu\}} \le z \le x^{1/2}$. Here $u := \log x/\log z$.

Proof. This is a classic application of Rankin's trick. Let $\eta := (\log u)/(\log z) \le \min\{1/3, 1/2\nu\}$. Noting that $\sum_p \sum_{r\ge\nu+1} p^{-r(1-\eta)/\nu} \ll_\nu \sum_p p^{-(1-\eta)(1+1/\nu)} \ll_\nu 1$, we have

$$\Psi_\nu(x,z) \le \sum_{\substack{n:\ P(n)\le z\\ n \text{ is } \nu\text{-full}}} \left(\frac{x}{n}\right)^{(1-\eta)/\nu} \le x^{(1-\eta)/\nu}\prod_{p\le z}\left(1+\sum_{r\ge\nu}\frac{1}{p^{r(1-\eta)/\nu}}\right) \ll x^{(1-\eta)/\nu}\ \exp\left(\sum_{p\le z}\frac{1}{p^{1-\eta}}\right).$$

Now $\sum_{p\le z} p^{-(1-\eta)} = \log_2 z + \sum_{p\le z} (\exp(\eta\log p) - 1)/p + O(1)$. In the last sum, the contribution of all $p \le 2^{1/\eta}$ is $\ll \eta\sum_{p\le 2^{1/\eta}} \log p/p \ll 1$, while that of $p \in (2^{1/\eta}, z]$ is at most $(\exp(\eta\log z) - 1)\sum_{2^{1/\eta}<p\le z} 1/p \le u(\log_2(3u) + O(1))$. Collecting all these bounds completes the proof. ☐

We also mention the following generalization of Proposition 1.6 which will be useful to us soon.

Proposition 7.3. *In the setting of Theorem 7.1, we have for any $\widehat{\chi}$ mod q,*

$$\sum_{n\le x}\widehat{\chi}(\widehat{f}(n)) \ll \frac{x^{1/\nu}}{\log x}\exp\left(\sum_{\substack{p\le y\\(f(p^\nu),q)=1}}\frac{1}{p} + |\alpha_{\widehat{\chi}}|\sum_{\substack{y<p\le x\\(f(p^\nu),q)=1}}\frac{1}{p} + O(M(\log x)^3\mathcal{E}(y))\right) + \left(\frac{x}{y}\right)^{1/\nu} + x^{1/\nu}(\log x)^2\exp\left(-u(2\nu)^{-1}\log u + O(u\log_2(3u))\right). \tag{7.7}$$

The implied constant depends only on the implied constant in (1.2).

Proof. By ν-supportedness (and the fact that $y > C$), the y-rough part of any n on the left of (7.7) is ν-full. By (7.4), it also suffices to restrict the sum on the left of (7.7) to the n's whose y-rough part is $(\nu+1)$-free. These restricted n are thus of the form BMA^ν where $P(BM) \le Y < P^-(A)$, where A is squarefree, M is ν-full, B is ν-free with $B \ll 1$, and $(f(B)f(M)f(A^\nu),q) = 1$.

If $M > x^{1/2}$, then $A \le (x/BM)^{1/\nu} \le x^{1/2\nu}/B^{1/\nu}$. Given B and A, the number of M is $\ll (x^{1/\nu}/B^{1/\nu}A)\cdot(\log z)\exp(-u(2\nu)^{-1}\log u + O(u\log_2(3u)))$ by Lemma 7.2. Since $\sum 1/B \ll 1$ and $\sum 1/A \ll \log x$, the total contribution of n with $M > x^{1/2}$ is absorbed in the right of (7.7). On the other hand, the total contribution of all n having $M \le x^{1/2}$ to the left hand side of (7.7) equals

$$\sum_{\substack{B\le x:\ P(B)\le y\\ B\text{ is }\nu\text{-free}}}\widehat{\chi}(\widehat{f}(B))\sum_{\substack{M\le x^{1/2}:\ P(M)\le y\\ M\text{ is }\nu\text{-full}}}\widehat{\chi}(\widehat{f}(M))\sum_{A\le(x/BM)^{1/\nu}}\mu(A)^2\,\mathbb{1}_{P^-(A)>y}\,\widehat{\chi}(\widehat{f}(A^\nu)). \tag{7.8}$$

Now we observe that uniformly in all B and M in (7.8), we have

$$\sum_{A\le(x/BM)^{1/\nu}}\mu(A)^2\,\mathbb{1}_{P^-(A)>y}\,\widehat{\chi}(\widehat{f}(A^\nu)) \ll \frac{(x/BM)^{1/\nu}}{\log x}\exp\left(|\alpha_{\widehat{\chi}}|\sum_{\substack{y<p\le x\\(f(p^\nu),q)=1}}\frac{1}{p} + O(M(\log x)^3\mathcal{E}(y))\right). \tag{7.9}$$

To show this, we proceed as in section 3: Writing $X \coloneqq (x/BM)^{1/\nu}$, the analogue of (3.4) is

$$\mathcal{D}(X,t) = \sum_{p\le X}\frac{1-\mathrm{Re}(\mathbb{1}_{p>y}\widehat{\chi}(\widehat{f}(p^\nu))p^{-it})}{p} \ge \log_2 X - |\alpha_{\widehat{\chi}}|\sum_{y<p\le X}\frac{\mathbb{1}_{f(p^\nu),q)=1}}{p} + O(1+M\mathcal{E}(y)(\log x)^3).$$

Inserting (7.9) into (7.8), recalling that $B \ll 1$, and observing that $\sum_{\substack{M:\ P(M)\le y\\ M\text{ is }\nu\text{-full}}}\mathbb{1}_{(f(M),q)=1}/M^{1/\nu} \le \prod_{p\le y}\left(1+\sum_{r\ge\nu}\mathbb{1}_{(f(p^r),q)=1}/p^{r/\nu}\right) \ll \exp(\sum_{p\le y}\mathbb{1}_{(f(p^\nu),q)=1}/p)$ completes the proof. ∎

7.2. **Narkiewicz's general criterion and its uniform analogues.** As alluded to above, Narkiewicz's criterion allows for the possibility that the behavior of the f_i at the ν-th powers of primes is most significant, for any (fixed) $\nu \in \mathbb{N}$. We thus start with multiplicative functions $f_1,\dots,f_K : \mathbb{N}\to\mathbb{Z}$ for which there exist polynomials $\{F_{i,j}\}_{\substack{1\le i\le K\\1\le j\le V}} \subset \mathbb{Z}[T]$ satisfying $f_i(p^j) = F_{i,j}(p)$ for all $i\in[K]$, $j\in[V]$, and all primes p. (Thus, $F_{i,j}$ controls the behavior of f_i at the j-th powers of primes.) For each $j\in[V]$, define $\alpha_j(q) \coloneqq \varphi(q)^{-1}\#\{u\in U_q : \prod_{i=1}^K F_{i,j}(u)\in U_q\}$. For a fixed $\nu\in[V]$, we say that q is ***ν*-admissible** if $\alpha_\nu(q)\ne 0$ but $\alpha_j(q) = 0$ for all $j<\nu$. (In other words, ν is the minimal index j for which $\prod_{i=1}^K F_{i,j}$ carries some unit mod q to some unit mod q.) Finally, let $\mathcal{Q}(\nu; f_1,\cdots,f_K)$ be the set of all ***ν*-admissible moduli *q*** satisfying the following property

(7.10) For all $\widehat{\chi} \neq (\chi_0, \dots, \chi_0) \bmod q$ satisfying $\prod_{i=1}^{K} \chi_i(F_{i,\nu}(u)) = 1$ on its unit support

$$\{u \in U_q : \prod_{i=1}^{K} F_{i,\nu}(u) \in U_q\}, \text{ there exists a prime } p \text{ satisfying } \sum_{m \geq 0} \widehat{\chi}(\widehat{f}(p^m))p^{-m/\nu} = 0.$$

By the triangle inequality, such a prime $p \leq 2^\nu$; thus $\mathcal{Q}(1; f_1, \dots, f_K)$ is precisely the set $\mathcal{Q}(f_1, \dots, f_K)$ that we had been working with before. We now state Narkiewicz's general criterion.

Theorem 7.4. [22] *Fix ν-admissible q. Then $(f_1, \dots, f_K)$ are WUD mod q iff $q \in \mathcal{Q}(\nu; f_1, \cdots, f_K)$.*

As the reader may expect by now, the polynomials $(F_{1,\nu}, \dots, F_{K,\nu})$ play the role of $(F_1, \dots, F_K)$ from the previous sections. Our first complete uniform analogue of Theorem 7.4 is thus the following generalization of Theorem 1.3, giving uniformity without any restrictions on inputs. All implied constants in what follows depend only on the polynomials $\{F_{i,j}\}_{\substack{1 \leq i \leq K \\ \mathbf{1 \leq j \leq \nu}}}$.

Theorem 7.5. *Fix $K_0 > 0$ and $\epsilon \in (0, 1)$, and assume that $F_{1,\nu}, \dots, F_{K,\nu}$ are multiplicatively independent. Then $(f_1, \dots, f_K)$ are WUD mod $q \leq (\log x)^{K_0}$ lying in $\mathcal{Q}(\nu; f_1, \cdots, f_K)$ and satisfying $IFH(F_{1,\nu}, \dots, F_{K,\nu}; B_0)$, provided at least one of the following two conditions holds:*

(i) *$q \leq (\log x)^{(1-\epsilon)\alpha_\nu(q)(K-1/D_{\min})^{-1}}$, where $D_{\min} := \min\{\deg F_{i,\nu} : 1 \leq i \leq K\}$,* ***or***

(ii) *q is squarefree and $q^{K-1} D_{\min}^{\omega(q)} \leq (\log x)^{(1-\epsilon)\alpha_\nu(q)}$.*

Again, both these ranges are essentially optimal. If $K > 1$ and $D_{\min} = 1$, then optimality holds in the stronger sense, i.e. even if $F_{1,\nu}, \dots, F_{K,\nu}$ are **any** pairwise coprime linear polynomials. The examples are analogous to those constructed in section 6; we discuss the changes in subsection § 7.5. As in Theorem 1.5, we need to restrict our inputs to restore full "Siegel–Walfisz uniformity".

Theorem 7.6. *Fix $K_0 > 0$ and assume that $F_{1,\nu}, \dots, F_{K,\nu}$ are multiplicatively independent. Then (1.5) holds uniformly in moduli $q \leq (\log x)^{K_0}$ lying in $\mathcal{Q}(\nu; f_1, \cdots, f_K)$ and satisfying $IFH(F_{1,\nu}, \dots, F_{K,\nu};\, B_0)$, and uniformly in $a_i \in U_q$. Here for a general q, we can take*

$$R = \begin{cases} \nu(KD+1), & \text{if } \nu < D := \sum_{i=1}^{K} \deg F_{i,\nu}. \\ \text{the least integer exceeding } \nu(1 + (\nu+1)(K - 1/D)), & \text{if } \nu \geq D. \end{cases} \tag{7.11}$$

If q is squarefree, we can take $R = \nu(K\nu + K - \nu + 1) + 1$, and we can also improve this to

$$R = \begin{cases} 2, & \text{if } K = \nu = 1 \text{ and } F_{1,\nu} \text{ is not squarefull.} \\ \nu(K\nu + K - \nu) + 1, & \text{if } \nu > 1 \text{ and at least one of } \{F_{i,\nu}\}_{i=1}^{K} \text{ is not squarefull.} \end{cases} \tag{7.12}$$

Both values of R in (7.12) are optimal in general; we discuss this in subsection § 7.5. The *second* value of R in (7.11) is also optimal, as shown by the example of $\sigma(n)$ in the discussion below. Note that the second values of R in (7.11) and (7.12) had no analogues in Theorem 1.5.

We mention concrete examples which can only be addressed with this greater generality. Recall the definition of $\mathcal{Q}(\nu; f_1, \cdots, f_K)$ given in (7.10).

- Śliwa [36] showed that $\sigma(n)$ is WUD modulo *fixed* q iff $6 \nmid q$; in fact, $\mathcal{Q}(1;\sigma) = \{q : 2 \nmid q\}$ and $\mathcal{Q}(2;\sigma) = \{q : 2 \mid q, 3 \nmid q\}$. While Theorem 1.4 only dealt with odd q, Theorem 7.5 shows that $\sigma(n)$ is WUD modulo all $q \le (\log x)^{(2-\delta)\widetilde{\alpha}(q)}$ in $\mathcal{Q}(2;\sigma)$, as well as modulo all squarefree $q \le (\log x)^{K_0}$ in $\mathcal{Q}(2;\sigma)$ satisfying $2^{\omega(q)} \le (\log x)^{(1-\delta)\widetilde{\alpha}(q)}$, where $\widetilde{\alpha}(q) = \prod_{\ell \mid q:\ \ell \equiv 1 \pmod 3}(1 - 2/(\ell-1))$. (Here $\nu = D = 2$ and IFH is trivial as T^2+1 is separable.) The restriction on squarefree q is optimal by the example in [33, subsection 7.2].
- Moreover, Theorem 7.6 shows that $\sigma(n)$ is WUD modulo all $q \le (\log x)^{K_0}$ in $\mathcal{Q}(2;\sigma)$ if we restrict to inputs n with $P_5(n) > q$ (respectively, $P_3(n) > q$ for squarefree q). By the examples constructed in [33, subsections 6.1 and 7.1], both these restrictions are optimal. (Note that our inputs n must be of the form m^2 or $2m^2$ since $\sigma(n)$ is odd, q being even.)
- Narkiewicz [23] studied the weak equidistribution of $\sigma_3(n) = \sum_{d \mid n} d^3$. He showed that $\mathcal{Q}(1;\sigma_3) = \{q : (q,14) = 1\}$, $Q(2;\sigma_3) = \{q : (q,6) = 2\}$, and $Q(\nu;\sigma_3) = \emptyset$ for all $\nu > 2$. Our results extend this to show that $\sigma(n)$ is WUD modulo all $q \le (\log x)^{(3/2-\delta)\alpha(q)}$ in $\mathcal{Q}(1;\sigma_3)$, as well as modulo all squarefree $q \le (\log x)^{K_0}$ in $\mathcal{Q}(1;\sigma_3)$ that satisfy $3^{\omega(q)} \le (\log x)^{(1-\delta)\widetilde{\alpha}(q)}$, where $\widetilde{\alpha}(q) = \prod_{\substack{\ell \mid q \\ \ell \equiv 1 \pmod 3}} (1 - 3/(\ell-1)) \cdot \prod_{\substack{\ell \mid q \\ \ell \not\equiv 1 \pmod 3}} (1 - 1/(\ell-1))$. Uniformity is restored modulo all $q \le (\log x)^{K_0}$ in $\mathcal{Q}(1;\sigma_3)$ by restricting to inputs n with $P_4(n) > q$ (resp. $P_2(n) > q$ for squarefree $q \le (\log x)^{K_0}$ in $\mathcal{Q}(1;\sigma_3)$). Analogous results can be given for varying q in $\mathcal{Q}(2;\sigma_3)$; this genuinely requires the additional generality in this section.

Our results apply to any $\sigma_r(n)$ (for which all involved polynomials are separable), as well as to families like $(n, \varphi(n), \sigma(n))$ $(\sigma, \sigma_2, \sigma_3)$, $(\varphi, \sigma, \sigma_2, \sigma_3)$, etc. In general, once we have an explicit description of the sets $\{\mathcal{Q}(\nu; f_1, \dots, f_K)\}_{\nu=1}^\infty$ (which is a "fixed-modulus" problem for each fixed ν), we have explicit analogues of Siegel–Walfisz for $(f_1, \dots, f_K)$ with optimal arithmetic restrictions. However, this problem of giving explicit descriptions of $\{\mathcal{Q}(\nu; f_1, \dots, f_K)\}_{\nu=1}^\infty$ is not solved in general. In fact, even just for the single function σ_r (with $r > 1$ fixed), the sets $\{\mathcal{Q}(\nu;\sigma_r)\}_{\nu=1}^\infty$ have only been computed upto $r \le 200$ in works of Narkiewicz and Rayner [23, 29, 30]. Narkiewicz [21, 24] shows that for $(f_1, \dots, f_K)$ satisfying some natural additional constraints, we have $\mathcal{Q}(\nu; f_1, \dots, f_K) = \emptyset$ for all $\nu \gg 1$; he also gives algorithms to compute the nonempty $\mathcal{Q}(\nu; f_1, \dots, f_K)$.

7.3. **Additional ingredients for Theorems 7.5 and 7.6.** From now on, we are in the setting of Theorem 7.4 (as described before the statement of the theorem). Our first key observation:

$$\text{If } q \text{ is } \nu\text{-admissible, then } \boldsymbol{(f_1, \dots, f_K)} \textbf{ are } \boldsymbol{\nu}\textbf{-supported.}^6 \tag{7.13}$$

To see this, let $\mathcal{A}$ be the set of primes ℓ which satisfy $\alpha_j(\ell) = 0$ for some $j < \nu$. Note that since $\alpha_j(\ell) \ge 1 - (\ell-1)^{-1}\sum_{i=1}^K \deg F_{i,j}$, the primes in the set $\mathcal{A}$ are no more than $1 + \max_{j \in [\nu]} \sum_{i=1}^K \deg F_{i,j} \ll 1$. To show (7.13), it thus suffices to show that for any n satisfying $(f(n), q) = 1$, the primes dividing the ν-free part of n must lie in $\mathcal{A}$. So assume by way of contradiction that $p^j \parallel n$ for some $p \notin \mathcal{A}$ and $j < \nu$. Then $\prod_{i=1}^K F_{i,j}(p) = f(p^j)$ divides $f(n)$. Since $j < \nu$ and q is ν-admissible, we have $\alpha_j(q) = 0$, so that by the Chinese Remainder Theorem, $\alpha_j(\ell_0) = 0$ for some prime $\ell_0 \mid q$. Then $\ell_0 \in \mathcal{A}$, so that $p \in U_{\ell_0}$. This forces $\ell_0 \mid \prod_{i=1}^K F_{i,j}(p) = f(p^j)$ as $\alpha_j(\ell_0) = 0$. Hence $\ell_0 \mid f(n)$, violating the fact that $(f(n), q) = 1$. This establishes (7.13).

We intend to apply Theorem 7.1. To set the stage, we need two additional ingredients. The first is an estimate on the input sets that generalizes and plays the role of the first estimate in (4.13).

[6] where we can take the "C" in the definition of ν-supported to be some constant depending only on $\{F_{i,j}\}_{\substack{1 \le i \le K \\ 1 \le j \le \nu}}$.

Proposition 7.7. *As $x \to \infty$, we have uniformly in ν-admissible $q \le (\log x)^{K_0}$,*

$$\#\{n \le x : \ (f(n), q) = 1\} = \frac{x^{1/\nu}}{(\log x)^{1-\alpha_\nu(q)}} \exp(O((\log_2(3q))^{O(1)})). \tag{7.14}$$

Proof. The lower bound follows just by looking at the n of the form m^ν and applying the first estimate in (4.13) to the multiplicative function $m \mapsto f(m^\nu)$. To show the upper bound implied in (7.14), define $Y := \exp(\sqrt{\log x})$, and note that by the arguments leading to (7.8), it suffices to count the n of the form BMA^ν where $P(BM) \le Y < P^-(A)$, where A is squarefree, M is ν-full and $M \le x^{1/2}$, where B is ν-free with $B \ll 1$, and where $(f(B)f(M)f(A^\nu), q) = 1$. By [14, Theorem 01, p. 2], given B and M, the total number of A is

$$\sum_{A \le (x/BM)^{1/\nu}} \mu(A)^2 \mathbb{1}_{P^-(A) > Y} \mathbb{1}_{(f(A^\nu), q)=1} \ll \frac{(x/BM)^{1/\nu}}{\log x} \exp\left(\sum_{Y < p \le x} \frac{\mathbb{1}_{(f(p^\nu), q)=1}}{p} \right).$$

Bounding $\sum 1/M^{1/\nu}$ as at the end of the proof of Proposition 7.3, we see that the number of such n is $\ll (x^{1/\nu}/\log x) \exp(\sum_{p \le x} \mathbb{1}_{(f(p^\nu),q)=1}/p)$. This is absorbed in the right of (7.14) because

$$\sum_{\substack{p \le X \\ (f(p^\nu), q)=1}} \frac{1}{p} = \sum_{\substack{p \le X \\ (F_{1,\nu}(p) \cdots F_{K,\nu}(p), q)=1}} \frac{1}{p} = \alpha_\nu(q) \log_2 X + O((\log_2(3q))^{O(1)}) \text{ uniformly in } X \ge 3q, \tag{7.15}$$

by [28, Lemma 2.4]. This establishes the upper bound implied in (7.14). ∎

The second ingredient is a partial improvement of (4.5) in the case $r = 3$.

Lemma 7.8. *Fix $\varepsilon > 0$. If at least one of $\{F_{i,\nu}\}_{i=1}^K$ is not squarefull (in $\mathbb{C}$), then*

$$\xi_3(q; z) \ll q^{-2+\varepsilon}, \ \text{uniformly in squarefree } q \text{ and in } z \in (q, x]. \tag{7.16}$$

Proof. The analogue of (4.11) holds with $f_i(P_j)$ replaced by $f_i(P_j^\nu)$ (and with $\mathcal{V}_{r,K}(q, \widehat{a})$ redefined accordingly), so we need only show that $\#\mathcal{V}_{3,K}(q; \widehat{a}) \ll \varphi(q)^{1+\epsilon}$. As $\#\mathcal{V}_{3,K}(q; \widehat{a}) = \prod_{\ell | q} \#\mathcal{V}_{3,K}(\ell; \widehat{a})$ and $\omega(q) \ll \log q / \log_2 q$, it suffices to show (to complete the proof of the lemma) that

$$\#\mathcal{V}_{3,K}(\ell; \widehat{a}) \ll \varphi(\ell) \text{ uniformly in primes } \ell \gg 1 \text{ and } \widehat{a} = (a_1, a_2, a_3) \in U_\ell^3. \tag{7.17}$$

Assume w.l.o.g. that $F_{1,\nu}$ is not squarefull (in $\mathbb{C}$). As in the last bound in (4.5), our idea will be to embed $\mathcal{V}_{3,K}(\ell; \widehat{a})$ into the set $V_\ell(\mathbb{F}_\ell)$ of all $\mathbb{F}_\ell$-rational points of the variety

$$V_\ell := \{(X, Y, Z) \in \overline{\mathbb{F}}_\ell^3 : \ F_{1,\nu}(X)F_{1,\nu}(Y)F_{1,\nu}(Z) - a_1 = F_{2,\nu}(X)F_{2,\nu}(Y)F_{2,\nu}(Z) - a_2 = 0\}.$$

Now, any sufficiently large prime ℓ satisfies all the following properties:

(i) ℓ doesn't divide the leading coefficients of $F_{1,\nu}$ or $F_{2,\nu}$;

(ii) $F_{1,\nu}$ is not squarefull in $\overline{\mathbb{F}}_\ell$;

(iii) If $F_{1,\nu}^{c_1} \cdot F_{2,\nu}^{c_2}$ is constant in $\overline{\mathbb{F}}_\ell(T)$ for some $c_1, c_2 \in \mathbb{Z}$, then $\ell \mid (c_1, c_2)$.

We ensure (ii) by arguing as in the first paragraph of the proof of Lemma 4.3. To ensure (iii), we proceed as in the proof of Lemma 4.3(2) to see that the polynomials $(F_{1,\nu}' F_{2,\nu}, F_{1,\nu} F_{2,\nu}')$ must be $\mathbb{F}_\ell$-linearly independent for all $\ell \gg 1$. (Basically, ℓ shouldn't divide the invariant factors of the

2-column matrix listing the coefficients of $F_{1,\nu}'F_{2,\nu}$ and $F_{1,\nu}F_{2,\nu}'$.) Now if $F_{1,\nu}^{c_1}\cdot F_{2,\nu}^{c_2}$ is constant in $\overline{\mathbb{F}}_\ell(T)$ for any such ℓ, then from its derivative, we get $c_1F_{1,\nu}'F_{2,\nu}+c_2F_{1,\nu}F_{2,\nu}'=0$, forcing $\ell\mid(c_1,c_2)$.

Our next important observation is that if ℓ is large enough to satisfy (i)-(iii) above, then

$$\begin{aligned}(7.18)\quad F_1^*(X,Y,Z)&\coloneqq F_{1,\nu}(X)F_{1,\nu}(Y)F_{1,\nu}(Z)-a_1 \text{ is irreducible over } \overline{\mathbb{F}}_\ell\\ &\text{and does \textbf{not} divide } F_2^*(X,Y,Z)\coloneqq F_{2,\nu}(X)F_{2,\nu}(Y)F_{2,\nu}(Z)-a_2 \text{ in } \overline{\mathbb{F}}_\ell[X,Y,Z].\end{aligned}$$

The first assertion can be shown by simply replicating the arguments given for the irreducibility of "$F_1(X)F_1(Y)-w$" in the proof of Proposition 4.4(3). Next, assume for the sake of contradiction $F_2^*=F_1^*H_1$ for some $H_1\in\overline{\mathbb{F}}_\ell[T]$. Comparing the coefficients of the monomial $Y^{r_1}Z^{r_2}$ of maximum total degree r_1+r_2 on both sides of the last identity, we obtain $F_{1,\nu}(X)\mid F_{2,\nu}(X)$. Hence $F_{2,\nu}=F_{1,\nu}^m\cdot H$ for some $m\in\mathbb{N}$ and $H\in\overline{\mathbb{F}}_\ell[X]$ such that $F\nmid H$. We will show that H must be constant.

To see this, note that by an easy induction argument, we have for all $t\in\{0,1,\dots,m\}$,

$$(7.19)\quad F_1^* \text{ divides } G_t(X,Y,Z)\coloneqq F_{1,\nu}(X)^{m-t}F_{1,\nu}(Y)^{m-t}F_{1,\nu}(Z)^{m-t}H(X)H(Y)H(Z)-a_1^{-t}a_2 \text{ in } \overline{\mathbb{F}}_\ell[X,Y,Z].$$

Indeed, since $G_0=F_2^*$, this is tautological for $t=0$. If (7.19) holds for some $t\le m-1$, then writing $G_t=F_1^*Q_t$ shows that $F_{1,\nu}(X)F_{1,\nu}(Y)F_{1,\nu}(Z)\mid(Q_t(X,Y,Z)-a_1^{-(t+1)}a_2)$. Defining Q_{t+1} by the relation $Q_t(X,Y,Z)-a_1^{-(t+1)}a_2=F_{1,\nu}(X)F_{1,\nu}(Y)F_{1,\nu}(Z)Q_{t+1}(X,Y,Z)$, the identity $G_t=F_1^*Q_t$ leads to $G_{t+1}=F_1^*Q_{t+1}$, completing the induction and proving (7.19). Applying (7.19) for $t=m$, we get $H(X)H(Y)H(Z)-a_1^{-m}a_2=F_1^*(X,Y,Z)Q(X,Y,Z)$ for some $Q\in\overline{\mathbb{F}}_\ell[X,Y,Z]$. Now if H were not constant, then again comparing the coefficient of the monomial $Y^{r_1}Z^{r_2}$ with maximal r_1+r_2 on both sides of the last identity would show that $F\mid H$, a contradiction. Hence $F_{1,\nu}^{-m}\cdot F_{2,\nu}=H$ must be constant in $\overline{\mathbb{F}}_\ell(T)$, violating property (iii) from before. This proves (7.18).

The final stretch of the argument involves some key inputs from commutative algebra and algebraic geometry. Note that by (7.18), (F_2^*,F_1^*) form an $\overline{\mathbb{F}}_\ell[X,Y,Z]$-regular sequence[7]: This is because if F_1^*Q lies in the ideal (F_2^*), then (7.18) forces $Q\in(F_2^*)$. By [3, Proposition 1.2.14], it thus follows that the ideal (F_2^*,F_1^*) has height at least 2, so that the variety V_ℓ has Krull dimension at most 1 (since the ambient ring $\overline{\mathbb{F}}_\ell[X,Y,Z]$ has Krull dimension 3). Finally, the version of the Lang–Weil bound in [11, Claim 7.2] yields $\#V_\ell(\mathbb{F}_\ell)\ll\ell$ for all $\ell\gg1$, yielding the desired (7.17). ☐

We are now ready to apply Theorem 7.1 in the context of Theorems 7.5 and 7.6. **Defining $\boldsymbol{\alpha}_{\widehat{\mathcal{X}}}$ and $\boldsymbol{y}$ as in subsection § 4.1** (but with $F_{i,\nu}$ playing the role of "F_i" there), we see that (7.2) holds with same M and $\mathcal{E}(y)$ as in § 4.1. We thus apply Theorem 7.1 with the same J,z as in § 4.1, and with $d\coloneqq Q_0$ (with Q_0 defined as after (4.2).) Proposition 4.4 (which just required multiplicative independence) continues to hold as stated, with $F_{i,\nu}$ playing the role of F_i, and with the $\xi_r(q,z)$ (re)defined in (7.1). This proposition and Lemma 7.8 bound all the relevant quantities on the right of (7.3). Lemma 7.2 and (7.15) estimate $\Psi_\nu(x,z)$ and $\sum_{p\le y}\mathbb{1}_{(f(p^\nu),q)=1}/p$ respectively.

Collecting all these observations shows that the right of (7.3) is negligible compared to the main term $\varphi(q)^{-K}\#\{n\le x:(f(n),q)=1\}$. (For instance, to show that the values of R in (7.11) work, we just need to check that $\max\{s/\nu+r/D,R/\nu-r+r/D-s\}>K$. If $s/\nu+r/D\le K$, then $R/\nu-r+r/D-s\ge R/\nu-K\nu+((\nu+1)/D-1)r$, and this last quantity attains its minimum

[7]Following [3, Chapter 1], given a module M over a ring R, we say that $x_1,\dots,x_n\in R$ form an M-regular sequence if $M(x_1,\dots,x_n)\neq M$, and if for each $i\ge1$, the multiplication-by-x_i map is injective in the quotient module $M/(x_1,\dots,x_{i-1})M$. (For $i=1$, this means that the map $M\to M:\lambda\mapsto x_i\lambda$ is injective.)

over all $r \in [1, R/\nu]$, either at $r = 1$ or $r = R/\nu$, depending on whether $\nu \geq D$ or $\nu < D$. The corresponding check for squarefree q is just more mechanically tedious; one needs only use the best available bound out of (7.16) and the analogues of (4.5) for each $r \in [1, R/\nu)$.)

Hence, proving Theorems 7.5 and 7.6 once again comes down to showing the following analogue of (4.14): There exists a constant $\delta_0 := \delta_0(C_0, \kappa) < 1$ such that

$$\sum_{n \leq x} \mathbb{1}_{(f(n),q)=1}\ \widehat{\psi}(\widehat{f}(n)) \ \ll\ \frac{x^{1/\nu}}{(\log x)^{1-\delta_0 \alpha(q)}} \quad \text{for all nontrivial tuples } \widehat{\psi} \bmod Q_0, \tag{7.20}$$

uniformly in $\boldsymbol{q \leq (\log x)^{K_0}}$ lying in $\boldsymbol{\mathcal{Q}(f_1, \ldots, f_K)}$, **having $\boldsymbol{Q_0 > 1}$ and having sufficiently large radical**. Once again, Proposition 7.3 yields (7.20) when $\prod_{i=1}^K \psi_i(F_{i,\nu}(u))$ is not constant on its unit support $\{u \in U_{Q_0} : \prod_{i=1}^K F_{i,\nu}(u) \in U_{Q_0}\}$. Hence, we need only show (7.20) when $\prod_{i=1}^K \psi_i(F_{i,\nu}(u))$ takes a constant value $c_{\widehat{\psi}}$ on its unit support. To do this, we just need to modify the arguments in section 5 as follows.

7.4. **The additional analytic ingredients.** Our first (and most nontrivial) modification is that in order to get the analogue of (5.3), we need the following observation.

Lemma 7.9. *For any $x \geq 2$, there exists an integer $X \in [x, x + x/\log^2 x)$ such that*

$$\sum_{\substack{3X/4 < n < 5X/4 \\ n \neq X}} \frac{\mathbb{1}_{(f(n),q)=1}}{|\log(X/n)|} \ \ll\ X^{1/\nu} \log X.$$

Proof. The lemma would follow once we show that with $h := x/\log^2 x$, we have

$$\sum_{X\ \in\ \mathbb{Z}\cap[x,x+h)}\ \sum_{\substack{3X/4 < n < 5X/4 \\ n \neq X}} \frac{\mathbb{1}_{(f(n),q)=1}}{|\log(X/n)|} \ \ll\ x^{1/\nu} h \log x. \tag{7.21}$$

We write the total double sum on the left as $S_1 + S_2$, where S_1 denotes the contribution of $n \in (3X/4, X-1]$. Then for any n in S_1, we can write $n = X - r$, where $r \in [1, X/4) \subset [1, (x+h)/4]$. By (7.13), we have $n = n_1 n_2$ where $n_1 \ll 1$ and n_2 is ν-full. Hence, $n_2 = (X-r)/n_1 \in [(x-r)/n_1, (x+h-r)/n_1)$. Also $|\log(X/n)| = -\log(1 - v/X) \gg v/X \gg v/x$. Combining these,

$$S_1 \ \leq\ x \sum_{1 \leq r \leq \frac{x+h}{4}} \frac{1}{r} \sum_{n_1 \ll 1} \sum_{\substack{\frac{x-r}{n_1} \leq n_2 < \frac{x+h-r}{n_1} \\ n_2 \text{ is } \nu\text{-full}}} 1 \ \ll\ x \log x \left(x^{1/\nu} \cdot \frac{h}{x} + x^{1/(\nu+1)} \right) \ \ll\ x^{1/\nu} h \log x,$$

where we have noted that the sum on n_2 above is $\ll x^{1/\nu}\{(1 + O(h/x))^{1/\nu} - 1\} + x^{1/(\nu+1)}$ by the Erdős-Szekeres estimate [12]. Likewise, we have $S_2 \ll x^{1/\nu} h \log x$, establishing (7.21). ☐

Note that $\sum_{x < n \leq x + x/\log^2 x} \mathbb{1}_{(f(n),q)=1} \ll x^{1/\nu}/\log^2 x$ (write $n = n_1 n_2$ and proceed as above). Hence, it suffices to show (7.20) with "x" replaced by the "X" in Lemma 7.9. Now, our Dirichlet series $\mathcal{F}(s) = \sum_{n \geq 1} \mathbb{1}_{(f(n),q)=1}\ \widehat{\psi}(\widehat{f}(n))/n^s$ absolutely converges on $\mathrm{Re}(s) > 1/\nu$, since in the Euler product $\mathcal{F}(s) = \prod_p (1 + \sum_{r \geq 1} \mathbb{1}_{(f(p^r),q)=1}/p^{rs})$, all but finitely many of the factors are of the form $1 + \sum_{r \geq \nu} \mathbb{1}_{(f(p^r),q)=1}/p^{rs}$ by ν-admissibility. By Perron's formula [37, Theorem II.2.3]),

$$\sum_{n \leq X} \mathbb{1}_{(f(n),q)=1}\ \widehat{\psi}(\widehat{f}(n)) \ =\ \frac{1}{2\pi i} \int_{\frac{1}{\nu}(1+\frac{1}{\log X})-iT}^{\frac{1}{\nu}(1+\frac{1}{\log X})+iT} \frac{\mathcal{F}(s) X^s}{s}\, \mathrm{d}s \ +\ O\left(\frac{X^{1/\nu} \log X}{T} \right).$$

To bound the error term from Perron, we used Lemma 7.9, and the technique in (5.2) which gave

$$\sum_{\substack{n\geq 1\\ n\notin(3X/4,5X/4)}} \frac{\mathbb{1}_{(f(n),q)=1}}{n^{\frac{1}{\nu}(1+\frac{1}{\log X})}|\log(X/n)|} \ll \sum_{\substack{n_2\geq 1\\ n_2 \text{ is } \nu\text{-full}}} \frac{1}{n_2^{\frac{1}{\nu}(1+\frac{1}{\log X})}} \ll \prod_p \left(1+\sum_{r\geq \nu} \frac{1}{p^{\frac{r}{\nu}(1+\frac{1}{\log X})}}\right) \ll \log X.$$

The rest of the argument in section 5 goes through only by scaling things by ν appropriately. Branch cut conventions are analogous to those before the statement of Lemma 5.1: Branch cuts up to $\sigma \leq 1$ (respectively, $\sigma \leq \beta_e$) there are replaced by those up to $\sigma \leq 1/\nu$ (resp. $\sigma \leq \beta_e/\nu$). Redefining $\boldsymbol{\mathcal{L}_Q(t) = \log(Q(|t\nu|+1))}$, Lemma 5.1 holds with $F_{\widehat{\psi}}(s)$ replaced by

$$F_{\widehat{\psi}}(s\nu) := \left(\prod_{\eta \bmod Q} L(s\nu,\eta)^{\gamma(\eta)}\right)^{\boldsymbol{\alpha_\nu(Q)}c_{\widehat{\psi}}}, \quad \text{with } \gamma(\eta) := \frac{1}{\alpha_\nu(Q)\varphi(Q)} \sum_{\substack{u\in U_Q\\ \prod_{i=1}^K F_{i,\nu}(u)\in U_Q}} \overline{\eta}(u),$$

so that $G_{\widehat{\psi}}(s)$ (redefined accordingly) is analytic on $\sigma > \nu^{-1}(1-1/\log Q)$ and satisfies $G_{\widehat{\psi}}(s) \ll (\log_2 Q)^3$ uniformly therein. (The analogue of (5.4) holds with $1-c_{\widehat{\psi}}\cdot \mathbb{1}_{p\prod_{i=1}^K F_{i,\nu}(p)\in U_Q}/p^{s\nu}$ replacing $1-c_{\widehat{\psi}}\cdot \mathbb{1}_{pF(p)\in U_Q}/p^s$ there.) We redefine our contour Γ_1 (Figure 1) by replacing $1+1/\log x$, 1, β^*, β_e, by $\nu^{-1}(1+1/\log X)$, $1/\nu$, β^*/ν, β_e/ν, and by taking $\vartheta(t) := \nu^{-1}(1-c_1/4\mathcal{L}_Q(t))$, $T = \exp(\sqrt{\log x})$.

In place of (5.8), we work with $H_{\widehat{\psi}}(s) := F_{\widehat{\psi}}(s\nu)\ (s-1/\nu)^{\alpha_\nu(q)c_{\widehat{\psi}}}\ (s-\beta_e/\nu)^{-\alpha_\nu(q)\gamma(\eta_e)c_{\widehat{\psi}}}$ which analytically continues into $\sigma > \nu^{-1}(1-c_1/\mathcal{L}_Q(t))$. The analogue of Proposition 5.2 holds with s in subpart (2) lying in $[\nu^{-1}(1-c_1/4\log Q),\ \nu^{-1}]$. Its proof is analogous, the main changes being:

- (5.9) holds with "$1-c_1/2\mathcal{L}_Q(t) \leq \mathrm{Re}(s)$" replaced by "$\nu^{-1}(1-c_1/2\mathcal{L}_Q(t)) \leq \mathrm{Re}(s)$".
- In the proof of subpart (1), we redefine $\mu(t)$ to $\nu^{-1}(1+c_1/7\mathcal{L}_Q(t))$. Hence the mention of $\zeta(\mu(t))$ is replaced by $\zeta(\nu\mu(t))$.

Proposition 5.3 and everything thereafter goes through with analogous arguments; the analogue of Γ_0 in (5.14) is the union of segments joining β^*/ν and $1/\nu$ above and below the branch cut. At the very end (in the case when $c_{\widehat{\psi}} = 1$), we can write $G_{\widehat{\psi}}(s) = G_{\widehat{\psi},1}(s)G_{\widehat{\psi},2}(s)$, where $G_{\widehat{\psi},2}(s) := \prod_{p\leq 2^\nu}\left(1+\sum_{r\geq 1}\mathbb{1}_{(f(p^r),q)=1}/p^{rs}\right)$, and where $G_{\widehat{\psi},1}(s)$ is analytic on $\sigma > \nu^{-1}(1-1/\log Q)$ and satisfies $G_{\widehat{\psi},1}(s) \ll (\log_2 Q)^3$ uniformly therein. By definition of $\mathcal{Q}(\nu; f_1, \cdots, f_K)$ (i.e., (7.10) and its ensuing observation), we have $G_{\widehat{\psi},2}(1/\nu) = 0$, allowing us to adapt (5.16) and (5.17). This establishes the desired (7.20), completing the proofs of Theorems 7.5 and 7.6. $\square$

7.5. **Optimality in Theorems 7.5 and 7.6.** The arguments for Lemma 6.1 show that given $\widehat{F} := (F_{1,\nu}, \ldots, F_{K,\nu})$ with $\prod_{i=1}^K F_{i,\nu}$ squarefree, there exists a constant $C_{\widehat{F}} > 0$ such that any q with $P^-(q) > C_{\widehat{F}}$ has $\alpha_\nu(q) \neq 0$ and satisfies (7.10) vacuously (i.e., there is **no** tuple of characters $\widehat{\chi} \neq (\chi_0, \ldots, \chi_0) \bmod q$ for which $\prod_{i=1}^K \chi_i(F_{i,\nu}(u))$ is constant on $\{u \in U_q : \prod_{i=1}^K F_{i,\nu}(u) \in U_q\}$). Henceforth: We fix $\nu \in \mathbb{N}$, a sufficiently large prime $\ell_0 > C_{\widehat{F}}$, and polynomials $\{F_{i,j}\}_{\substack{1\leq i\leq K\\ 1\leq j\leq \nu-1}} \subset \mathbb{Z}[T]$ with **all** coefficients divisible by ℓ_0. Our $\{F_{i,\nu}\}_{i=1}^K$ play the role of the $\{F_i\}_{i=1}^K$ in section 6. Our moduli $q \leq (\log x)^{K_0}$ will have $P^-(q) = \ell_0$, so that by this discussion, $q \in \mathcal{Q}(\nu; f_1, \cdots, f_K)$. [8]

[8]The construction of $\{F_{i,j}\}_{\substack{1\leq i\leq K\\ 1\leq j\leq \nu-1}}$ guarantees that $\alpha_j(q) = 0$ for all $j < \nu$, and the definition of $C_{\widehat{F}}$ guarantees both that $\alpha_\nu(q) \neq 0$ and that q satisfies (7.10).

The ranges (i) and (ii) in Theorem 7.5 are both optimal, and if $D_{\min} = 1$, the range in (i) is optimal for $\{F_{i,\nu}\}_{i=1}^{K}$ being **any** pairwise coprime linear polynomials. All this can be shown by using the set-up in the previous paragraph, and having the prime powers P^ν play the role of the prime inputs "P" in section 6. (Example: To prove optimality of Theorem 7.5(i), take $F_{i,\nu}(T) := (T-1)^r + i$, and q a perfect r-th power with $P^-(q) = \ell_0 > \max\{C_{\widehat{F}}, 2K\}$. Then any prime $P \equiv 1 \pmod{q^{1/r}}$ satisfies $f_i(P^\nu) \equiv i \pmod q$, showing that $\#\{n \le x : (\forall i)\ f_i(n) \equiv i \pmod q\} \gg x^{1/\nu}/(q^{1/r}\log x)$.)

Finally, we show the optimality of the values of R in (7.12); hereafter q is **squarefree**. The value $R = 2$ is optimal by the constructions in section 6. To show that $R = \nu(K\nu + K - \nu) + 1$ is optimal in the setting of (7.12), we adapt the last construction in section 6 as follows:

Take $F_{i,\nu}(T) := \prod_{j=1}^{r}(T - 2j) + 2(2i-1)$, and continue with the set-up in the first paragraph of this subsection (with $P^-(q) = \ell_0 > \max\{C_{\widehat{F}}, 4Kr\}$). Let $f_1, \dots, f_K : \mathbb{N} \to \mathbb{Z}$ be any multiplicative functions satisfying $f_i(p^j) = F_{i,j}(p)$ for all $i \in [K]$ and all $j \in [\nu]$, as well as $f_i(p^{\nu+1}) := 1$ for all i. Let $n = (p_1 \dots p_{(K-1)\nu})^{\nu+1} \cdot P^\nu$, with $p_1, \dots, p_{(K-1)\nu}, P$ being distinct primes satisfying $q < p_{K-1} < \dots < p_1 < x^{1/8K\nu^2} < x^{1/3\nu} < P \le \left(x/(p_1 \dots p_{(K-1)\nu})^{\nu+1}\right)^{1/\nu}$ and $\prod_{j=1}^{r}(P - 2j) \equiv 0 \pmod q$. (Recall that there are exactly $r^{\omega(q)}$ many possible coprime residue classes mod q that P could lie in.) Then $n \le x$, and $f_i(n) = F_{i,\nu}(P) \equiv 2(2i-1) \pmod q$ for all i, so that a computation entirely analogous to that done in (6.2) and (6.3) shows that

$$\sum_{\substack{n \le x:\ P_{\nu(K\nu+K-\nu)}(n) > q \\ (\forall i)\ f_i(n) \equiv 2(2i-1) \pmod q}} 1 \gg \frac{r^{\omega(q)}}{\varphi(q)^K (\log_2 x)^{\nu(K-1)+1}} \cdot \frac{x^{1/\nu}}{\log x}$$

As in section 6, the right can be made to grow much faster than our expected main term.

Acknowledgements

The author would like to thank Paul Pollack for fruitful discussions and for his continued support and encouragement. He would also like to thank his Ph.D. committee members, Neil Lyall, Akos Magyar and Giorgis Petridis, for helpful comments on the exposition.

References

1. A. Akande, *Uniform distribution of polynomially-defined additive functions to varying moduli*, submitted.
2. W. D. Banks and I. E. Shparlinski, *Congruences and exponential sums with the Euler function*, High primes and misdemeanours: lectures in honour of the 60th birthday of Hugh Cowie Williams, Fields Inst. Commun., vol. 41, Amer. Math. Soc., Providence, RI, 2004, pp. 49–59.
3. W. Bruns and J. Herzog, *Cohen–Macaulay rings*, Cambridge Studies in Advanced Mathematics, vol. 39, Cambridge University Press, Cambridge, 1998.
4. E. R. Canfield, P. Erdős, and C. Pomerance, *On a problem of Oppenheim concerning "factorisatio numerorum"*, J. Number Theory **17** (1983), 1–28.
5. T. Cochrane, *Exponential sums modulo prime powers*, Acta Arith. **101** (2002), 131–149.
6. T. Cochrane, C. L. Liu, and Z. Y. Zheng, *Upper bounds on character sums with rational function entries*, Acta Math. Sin. (Engl. Ser.) **19** (2003), 327–338.
7. H. Davenport, *On character sums in finite fields*, Acta Math. **71** (1939), 99–121.
8. H. Delange, *On integral–valued additive functions*, J. Number Theory **1** (1969), 419–430.
9. ______, *On integral–valued additive functions, II*, J. Number Theory **6** (1974), 161–170.
10. T. Dence and C. Pomerance, *Euler's function in residue classes*, Ramanujan J. **2** (1998), 7–20.
11. Z. Dvir, J. Kollár, and S. Lovett, *Variety evasive sets*, Comput. Complexity **23** (2014).
12. P. Erdös and G. Szekeres, *Über die Anzahl der Abelschen Gruppen gegebener Ordnung und über ein verwandtes zahlentheoretisces Problem*, Acta Univ. Szeged **7** (1934–1935), 95–102.

13. O. M. Fomenko, *The distribution of values of multiplicative functions with respect to a prime modulus (Russian)*, Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **93** (1980), 218–224.
14. R. R. Hall and G. Tenenbaum, *Divisors*, Cambridge Tracts in Mathematics, vol. 90, Cambridge University Press, Cambridge, 1988.
15. S. Konyagin, *Letter to the editors: "The number of solutions of congruences of the nth degree with one unknown"*, Mat. Sb. (N.S.) **110(152)** (1979), 158.
16. ______, *The number of solutions of congruences of the nth degree with one unknown*, Mat. Sb. (N.S.) **109(151)** (1979), 171–187, 327.
17. E. Landau, *Lösung des Lehmer'schen Problems*, American J. Math. **31** (1909), 86–102.
18. N. Lebowitz-Lockard, P. Pollack, and A. Singha Roy, *Distribution mod p of Euler's totient and the sum of proper divisors*, Michigan Math. J. **74** (2024), 143–166.
19. D. B. Leep and C. C. Yeomans, *The number of points on a singular curve over a finite field*, Arch. Math. (Basel) **63** (1994), 420–426.
20. W. Narkiewicz, *On distribution of values of multiplicative functions in residue classes*, Acta Arith. **12** (1967), 269–279.
21. ______, *Euler's function and the sum of divisors*, J. Reine Angew. Math. **323** (1981), 200–212.
22. ______, *On a kind of uniform distribution for systems of multiplicative functions*, Litovsk. Mat. Sb. **22** (1982), no. 1, 127–137.
23. ______, *Distribution of coefficients of Eisenstein series in residue classes*, Acta Arith. **43** (1983), 83–92.
24. ______, *Uniform distribution of sequences of integers in residue classes*, Lecture Notes in Mathematics, vol. 1087, Springer-Verlag, Berlin, 1984.
25. W. Narkiewicz and F Rayner, *Distribution of values of $\sigma_2(n)$ in residue classes*, Monatsh. Math. **94** (1982), 133–141.
26. P. Pollack and A. Singha Roy, *Joint distribution in residue classes of polynomial-like multiplicative functions*, Acta Arith. **202** (2022), 89–104.
27. ______, *Benford behavior and distribution in residue classes of large prime factors*, Canad. Math. Bull. **66** (2023), no. 2, 626–642.
28. ______, *Distribution in coprime residue classes of polynomially-defined multiplicative functions*, Math. Z. **303** (2023), no. 4, 20 pages.
29. F. Rayner, *Weak uniform distribution for divisor functions. I*, Math. Comp. **50** (1988), 335–342.
30. ______, *Weak uniform distribution for divisor functions. II*, Math. Comp. **51** (1988), 331–337.
31. W. M. Schmidt, *Equations over finite fields*, Lecture Notes in Mathematics, vol. 536, Springer-Verlag Berlin Heidelberg, 1976.
32. E. J. Scourfield, *A uniform coprimality result for some arithmetic functions*, J. Number Theory **20** (1985), 315–353.
33. A. Singha Roy, *Anatomical mean value bounds on multiplicative functions and the distribution of the sum of divisors*, submitted.
34. ______, *Joint distribution in residue classes of families of polynomially-defined additive functions*, submitted.
35. ______, *Joint distribution in residue classes of families of polynomially-defined multiplicative functions II*, submitted.
36. J. Śliwa, *On distribution of values of $\sigma(n)$ in residue classes*, Colloq. Math. **27** (1973), 283–291, 332.
37. G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, third ed., Graduate Studies in Mathematics, vol. 163, American Mathematical Society, Providence, RI, 2015.
38. D. Wan, *Generators and irreducible polynomials over finite fields*, Math. Comp. **66** (1997), no. 219, 1195–1212.
39. A. Weil, *Sur les courbes algébriques et les variétes qui s'en déduisent*, Actual. Sci. Industr. **1041** (1948).

Department of Mathematics, University of Georgia, Athens, GA 30602

Email address: akash01s.roy@gmail.com