# JOINT DISTRIBUTION IN RESIDUE CLASSES OF FAMILIES OF POLYNOMIALLY-DEFINED MULTIPLICATIVE FUNCTIONS I

AKASH SINGHA ROY

Abstract. We obtain best possible analogues of the Siegel–Walfisz Theorem for the value distributions of large classes of multiplicative functions and for the joint distributions of families consisting of such functions. We extend a criterion of Narkiewicz for such families to give new uniform results that are essentially optimal in the range and arithmetic restrictions on the modulus as well as in most parameters and hypotheses. This also significantly generalizes and improves upon previous work done for a single such function in specialized settings. Furthermore, we reveal some surprising phenomena leading to failure of equidistribution. Our results have applications to large classes of interesting (integer-valued) multiplicative functions, such as Euler's totient $\varphi(n)$, the sum-of-divisors $\sigma(n)$, the coefficients of the Eisenstein series, etc., and to the joint distribution of collections/families consisting of such functions. For instance, an application of our results shows that for any fixed $\epsilon > 0$, the functions $\varphi(n)$ and $\sigma(n)$ are jointly asymptotically equidistributed among the reduced residue classes to moduli $q$ coprime to 6 varying uniformly up to $(\log x)^{(1-\epsilon)\alpha(q)}$, where $\alpha(q) := \prod_{\ell \mid q}(\ell-3)/(\ell-1)$: This is the best possible result for the joint distribution of $(\varphi, \sigma)$ to a single varying modulus. Our results also give interesting consequences for the families $(\sigma, \sigma_3)$, $(\varphi, \sigma, \sigma_2)$, $(\varphi, \sigma, \sigma_2, \sigma_3)$ and so on.

One of the central themes behind our arguments is a certain "mixing" idea that can be interpreted as a Markov chain mixing phenomenon, however we carry out this idea using methods from the "anatomy of integers" in conjunction with classical as well as "pretentious" analytic arguments. In addition to these, our arguments involve counting solutions to simultaneous polynomial congruences in a large number of variables (that can be thought of as multiplicative analogues of the circle method), and our methods blend character sum estimates with linear algebra and module theory, commutative algebra, algebraic number theory, as well as arithmetic and algebraic geometry. These ideas have been useful in other problems as well.

## 1. Introduction

We say that an integer-valued arithmetic function $g$ is uniformly distributed (or equidistributed) modulo $q$ if $\#\{n \le x : g(n) \equiv b \pmod q\} \sim x/q$ as $x \to \infty$, for each residue class $b$ mod $q$. This definition generalizes naturally to families of arithmetic functions, and has been well-studied for (integral-valued) additive functions, – with work of Delange [10], [11] characterizing when a family of such functions is equidistributed to a fixed modulus $q$. These results have also been partially extended in [37], [38], [1] and [46], where the modulus $q$ itself has been allowed to vary up to a certain threshold depending on the stopping point $x$ of inputs.

However, for multiplicative functions, there are indications that uniform distribution is not the correct notion to consider. For instance, it can be shown that the Euler totient function $\varphi(n)$ is almost always divisible by any fixed integer $q$, and hence is not equidistributed modulo

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.

any $q > 1$. Motivated by this, Narkiewicz in [27] introduces the notion of weak uniform distribution: Given an integer-valued arithmetic function $f$ and a positive integer $q$, we say that $f$ is weakly uniformly distributed (or weakly equidistributed or WUD) modulo $q$ if there are infinitely many positive integers $n$ for which $\gcd(f(n), q) = 1$, and if

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

for each coprime residue class $a$ mod $q$. This definition extends naturally to families of arithmetic functions: we say that the integer-valued arithmetic functions $f_1, \dots, f_K$ are jointly weakly equidistributed (or jointly WUD) modulo $q$ if there are infinitely many $n$ for which $\gcd(f_1(n) \cdots f_K(n), q) = 1$, and if for all coprime residue classes $a_1, \dots, a_K$ mod $q$, we have

(1.1)
$$\#\{n \le x : \forall i \in [K],\ f_i(n) \equiv a_i \pmod q\} \sim \frac{1}{\varphi(q)^K} \#\{n \le x : \gcd(f_1(n) \cdots f_K(n), q) = 1\}$$

as $x \to \infty$. (Here and below, $[K]$ denotes the set $\{1, \dots, K\}$.)

The phenomenon of weak equidistribution has drawn a lot of attention for specific as well as for general classes of multiplicative functions. Narkiewicz [27] shows that $\varphi(n)$ is weakly equidistributed precisely to those moduli $q$ that are coprime to 6, while Dence and Pomerance [12] study the distribution of $\varphi(n)$ in residue classes modulo 3 and 12). Śliwa [49] shows that the sum of divisors function $\sigma(n) = \sum_{d \mid n} d$ is weakly equidistributed mod $q$ exactly when $q$ is not a multiple of 6. Generalizations of Śliwa's result to Fourier coefficients of Eisenstein series (more generally, the functions $\sigma_r(n) := \sum_{d \mid n} d^r$), as well as to families of such functions, has been studied in great depth by Narkiewicz, Rayner, Śliwa, Dobrowolski, Fomenko, and others; see [27], [49], [15], [28], [29], [32], [30], [31, Theorem 6.12], [40], [41]. In fact in [27, Theorem 1], Narkiewicz gives a general criterion for deciding weak equidistribution for a single "polynomially-defined" multiplicative function $f$, one that can be controlled by the values of polynomials at the first few powers of all primes. While the exact statement requires some set-up, the general flavor of the criterion is that such a function $f$ is weakly equidistributed modulo a fixed $q$ precisely when for every nontrivial Dirichlet character mod $q$ that acts trivially on a special subgroup of the unit group mod $q$, a certain "local factor" (or Euler factor) associated to this Dirichlet character vanishes. Narkiewicz dedicates a significant portion of his monograph [31] to give more explicit sufficient conditions that guarantee weak uniform distribution, and to obtain algorithms characterizing all the moduli to which a given "polynomially-defined" multiplicative function is weakly equidistributed.

In all these results, the modulus $q$ has been assumed to be fixed. A natural question of some interest is whether one can allow $q$ to **vary** with our stopping point $x$. This general problem of investigating equidistribution in residue classes to varying moduli has been ardently studied in various contexts, such as for smooth numbers and mean values of multiplicative functions. In our context, a concrete starting model is the celebrated Siegel–Walfisz Theorem, according to which for any fixed $K_0 > 0$, the primes up to any $x$ are weakly equidistributed uniformly to moduli $q \le (\log x)^{K_0}$. So one might ask: Can we find analogues of the Siegel–Walfisz theorem for the value distributions of multiplicative functions or (more generally) for the joint distributions of a family of multiplicative functions? To this end, given a constant $K_0 > 0$, we shall say that the functions $f_1, \dots, f_K : \mathbb{N} \to \mathbb{Z}$ are jointly weakly equidistributed (or jointly WUD) mod $q$, uniformly for $q \le (\log x)^{K_0}$, if:

(i) For every such $q$, $\prod_{i=1}^K f_i(n)$ is coprime to $q$ for infinitely many $n$, and

(ii) The relation (1.1) holds as $x \to \infty$, uniformly in moduli $q \le (\log x)^{K_0}$ and in coprime residue classes $a_1, \dots, a_K$ mod $q$. Explicitly, this means that for any $\epsilon > 0$, there exists $X(\epsilon) > 0$ such that the ratio of the left hand side of (1.1) to the right hand side lies in $(1-\epsilon, 1+\epsilon)$ for all $x > X(\epsilon)$, $q \le (\log x)^{K_0}$ and coprime residues $a_1, \dots, a_K$ mod $q$.

If $K = 1$ and $f_1 = f$, we shall simply say that $f$ is weakly equidistributed (or WUD) mod $q$, uniformly for $q \le (\log x)^{K_0}$.

The question of weak equidistribution to varying moduli seems to have been first studied in [23], [36] and [38], which made some partial progress towards obtaining a uniform analogue of Narkiewicz's aforementioned criterion for a single "polynomially-defined" multiplicative function. However, the settings in these papers were highly special instances of the setting in Narkiewicz's original criterion in [27] (in the sense that they imposed several additional restrictions), so much so that they could not be used to obtain satisfactory uniform analogues of the weak equidistribution results on $\sigma_r(n)$ alluded to above.

As a special case of our results in this manuscript, we are able to extend Narkiewicz's criterion in [27] in its **full generality**, in the sense that our results do not require *any* additional restrictions beyond those which can be proven to be necessary. Applications of our main theorems also extend the aforementioned works of Narkiewicz, Rayner, Śliwa, Dobrowolski, Fomenko and others in the best possible manner (see the discussion following the statement of Theorem 2.5). For instance, we get all the following uniform analogues of Śliwa's result in [49]: The sum of divisors function $\sigma(n)$ is weakly equidistributed uniformly to all odd moduli $q \le (\log x)^{K_0}$ as well as to all even $q$ not divisible by 3 that are either no more than a small power of $\log x$ or are squarefree without too many distinct prime factors. In addition, uniformity is restored to *all* (resp. to *squarefree*) even $q \le (\log x)^{K_0}$ that are not multiples of 3, provided we restrict to inputs $n$ having *six* (resp. *four*) large prime factors counted with multiplicity. By examples constructed in [47], all these restrictions are optimal.

All of these results and improvements are only for a single multiplicative function. In [29], Narkiewicz generalizes his criterion from [27] to decide joint weak equidistribution for *families* of "polynomially defined" multiplicative functions to a fixed modulus $q$; he uses this generalized criterion in [28] to characterize those fixed $q$ to which the Euler totient $\varphi(n)$ and sum of divisors $\sigma(n)$ are jointly weakly equidistributed. However, several arguments in the aforementioned papers [23, 36, 38] investigating varying-modulus analogues of his previous criterion are all strictly constrained to a single multiplicative function and do not generalize to families. In this manuscript, we extend Narkiewicz's general criterion in [29] for families of multiplicative functions to a single varying modulus $q$, and give new results that are **best possible** in the range of uniformity and arithmetic restrictions on $q$.

The qualitative summary of our main results is as follows. Under certain (provably) unavoidable conditions, a given family $f_1, \dots, f_K$ of polynomially-defined multiplicative functions is jointly weakly equidistributed *exactly* to those moduli $q$ that satisfy Narkiewicz's criterion, uniformly in $q$ varying up to small powers of $\log x$, where these powers are all essentially optimal as well. In addition, weak equidistribution is restored in the full "Siegel-Walfisz range" $q \le (\log x)^{K_0}$ provided we restrict to inputs $n$ having sufficiently many large prime factors. The threshold for "sufficiently many" can be reduced and optimized (thus ensuring equidistribution among larger sample spaces of inputs) whenever $q$ is squarefree.

The intuitive explanation for such constraints on our inputs $n$ comes from a certain 'mixing' phenomenon in the unit group mod $q$. To illustrate this, let $q$ be an odd positive integer. From the set of units $u$ mod $q$ for which $u+1$ is also a unit, choose uniformly at random $u_1, u_2, u_3, \dots$, and construct the sequence of partial products $u_1+1, (u_1+1)(u_2+1), (u_1+1)(u_2+1)(u_3+1), \dots$. Then as we go further into the sequence, each unit mod $q$ is roughly equally likely to appear as one of the products $(u_1+1)\cdots(u_J+1)$. This particular example lies at the core of the weak equidistribution of $\sigma(n)$ to odd moduli. The phenomenon for $\sigma(n)$ to even moduli not divisible by 3 is analogous, except that we work with units $u$ mod $q$ for which $u^2+u+1$ is also a unit mod $q$.

Interestingly, although this mixing phenomenon can be interpreted as a quantitative ergodicity phenomenon for random walks on multiplicative groups, no actual Markov chains are harmed in the production of our arguments. Instead, we detect this mixing quantitatively using methods from the "anatomy of integers", supplemented by character sum machinery in conjunction with linear algebraic arguments over residue rings. But this only takes us partway: To get the desired main terms, we crucially need arguments from both the classical and "pretentious" schools of analytic number theory. Note that the anatomical part of our arguments cannot be substituted by purely analytic arguments either, since the latter do *not* give us the desired asymptotic in the full range of uniformity. Furthermore, to bound the contributions of certain "bad" inputs, we need to understand the rational points of certain affine varieties over finite fields using tools from arithmetic and algebraic geometry. A more detailed summary of the arguments is given towards the end of the next section.

## 2. The setting and the main results

### 2.1. Narkiewicz's general criterion and shortcomings of previous work.

We say that an arithmetic function $f$ is polynomially-defined if there exists $V \ge 1$ and polynomials $\{W_v\}_{1 \le v \le V}$ with integer coefficients satisfying $f(p^v) = W_v(p)$ for all primes $p$ and all $v \in [V]$. Narkiewicz's criterion requires the following set-up:

- Consider $K, V \ge 1$ and polynomially-defined multiplicative functions $f_1, \dots, f_K \colon \mathbb{N} \to \mathbb{Z}$, with defining polynomials $\{W_{i,v}\}_{\substack{1 \le i \le K \\ 1 \le v \le V}} \subset \mathbb{Z}[T]$ satisfying $f_i(p^v) = W_{i,v}(p)$ for any prime $p$, and any $i \in [K], v \in [V]$.
- For any $q$ and $v \in [V]$, define $R_v(q) := \{u \in U_q : \prod_{i=1}^K W_{i,v}(u) \in U_q\}$; here $U_q := (\mathbb{Z}/q\mathbb{Z})^\times$ denotes the multiplicative group mod $q$. [1]
- Fix $k \in [V]$ and assume that $\{W_{i,k}\}_{1 \le i \le K}$ are all nonconstant. We say that $q \in \mathbb{N}$ is $k$-admissible (with respect to the family $(W_{i,v})_{\substack{1 \le i \le K \\ 1 \le v \le V}}$) if the set $R_k(q)$ is nonempty but the sets $R_v(q)$ are empty for all $v < k$.

[1] Hence, saying "$r \in U_q$" for an integer $r$ is synonymous with saying that "$\gcd(r, q) = 1$".

- Define $\mathcal{Q}(k; f_1, \cdots, f_K)$ to be the set of all $k$-admissible integers $q$ such that for every tuple $(\chi_1, \ldots, \chi_K) \neq (\chi_0, \ldots, \chi_0)$ of Dirichlet characters[2] mod $q$ for which the product $\prod_{i=1}^{K} \chi_i \circ W_{i,k}$ acts trivially on $R_k(q)$ [3], there exists a prime $p$ satisfying

$$\sum_{j \geq 0} \frac{\chi_1(f_1(p^j)) \cdots \chi_K(f_K(p^j))}{p^{j/k}} = 0. \tag{2.1}$$

Narkiewicz's criterion [29, Theorem 1] in this setting is then stated as follows; a precursor to this result is his older criterion [27, Theorem 1] for a single multiple function.

**Theorem N.** *Fix a $k$-admissible integer $q$. The functions $f_1, \ldots, f_K$ are jointly weakly equidistributed modulo $q$ if and only if $q \in \mathcal{Q}(k; f_1, \cdots, f_K)$.*

It is worth pointing out that the above result and the older [27, Theorem 1] are capable of dealing with really sparse input sets. For instance, for a fixed $k \in [V]$, if $q$ is $k$-admissible, then by Lemma 3.3, the sample space of relevant inputs $\{n \leq x : \gcd(f_1(n) \ldots f_K(n), q) = 1\}$ consists of those integers which are "almost" $k$-full, hence this sample space has size only $O(x^{1/k})$. In general, sparse sets like this can often present difficulties while studying arithmetic questions about them.

As mentioned in the introduction, the first steps towards obtaining uniform analogues of Narkiewicz's forerunning criterion [27, Theorem 1] to Theorem N were taken in [23], [36] and [38]. However the arguments in these papers are very much limited to the case of a single multiplicative function (i.e. $K = 1$), and even in that special case, they are still far from giving best possible analogues of [27, Theorem 1] because they crucially need $q$ to be 1-admissible (i.e. $k = 1$) and have sufficiently large prime factors, and also crucially need the defining polynomial $W_{1,1}$ to be separable. In particular, the results in [23], [36] and [38] are unable to deal with sparse input sets and hence also unable to give satisfactory uniform analogues of most of the previously-mentioned results of Narkiewicz, Rayner, Śliwa, Dobrowolski, Fomenko, and others in [27], [49], [15], [28], [29], [32], [30], [31, Theorem 6.12], [40], [41].

In this paper, we remove *all* these limitations, and obtain best possible uniform analogues of Theorem N, which are thus also best possible analogues of the Siegel–Walfisz theorem for families of polynomially–defined multiplicative functions. Our results will not impose any additional restrictions, beyond those that can be *proven* to be necessary and essentially optimal. These results are thus also new for a single multiplicative function as they address all the aforementioned shortcomings of [23], [36] and [38]. Special cases of our main results thus also give completely uniform analogues of *all* the works mentioned in the previous paragraph.

### 2.2. Multiplicative independence and the Invariant Factor Hypothesis.

For concrete and provably unavoidable reasons (see Theorems 2.4 and 2.5 below), we are going to need two additional hypotheses. We first define the relevant notation and terminology.

[2] Here $\chi_0$ or $\chi_{0,q}$ denotes, as usual, the trivial or principal character mod $q$.

[3] i.e. $\prod_{i=1}^{K} \chi_i(W_{i,k}(u)) = 1$ for all $u \in R_k(q)$; note that $R_k(q)$ is precisely the support of the product $\prod_{i=1}^{K} \chi_i \circ W_{i,k}$ (i.e. the set of $u$ where it is nonzero)

1. We say that the polynomials $\{F_i\}_{1\le i\le K} \subset \mathbb{Z}[T]$ are multiplicatively independent (over $\mathbb{Z}$) if there is no tuple of integers $(c_1, \dots, c_K) \neq (0, \dots, 0)$ for which the product $\prod_{i=1}^{K} F_i^{c_i}$ is identically constant in $\mathbb{Q}(T)$. This hypothesis is very easy to satisfy, for example if $\prod_{i=1}^{K} F_i$ is separable, then $\{F_i\}_{1\le i\le K} \subset \mathbb{Z}[T]$ are multiplicatively independent.

2. Assume that $\{F_i\}_{i=1}^{K} \subset \mathbb{Z}[T]$ are multiplicatively independent. Factor $F_i = r_i \prod_{j=1}^{M} G_j^{\mu_{ij}}$ with $r_i \in \mathbb{Z}$, $\{G_j\}_{j=1}^{M} \subset \mathbb{Z}[T]$ being pairwise coprime primitive[4] irreducibles and with $\mu_{ij} \geq 0$ being integers, such that each $G_j$ appears with a positive exponent $\mu_{ij}$ in some $F_i$. Let $\omega(F_1 \cdots F_K) \coloneqq M$ and define the exponent matrix of $(F_i)_{i=1}^{K}$ to be the $M \times K$ matrix

$$E_0 \coloneqq E_0(F_1, \dots, F_K) \coloneqq \begin{pmatrix} \mu_{11} & \cdots & \mu_{K1} \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ \mu_{1M} & \cdots & \mu_{KM} \end{pmatrix} \in \mathbb{M}_{M\times K}(\mathbb{Z}),$$

so that $E_0$ has a positive entry in each row. Since $\{F_i\}_{i=1}^{K} \subset \mathbb{Z}[T]$ are multiplicatively independent, the columns of $E_0$ are $\mathbb{Q}$-linearly independent and $\omega(F_1 \cdots F_K) = M \geq K$.

3. Continuing from above, $E_0$ has a Smith Normal Form given by the $M \times K$ diagonal matrix $\operatorname{diag}(\beta_1, \dots, \beta_K)$, where $\beta_1, \dots, \beta_K \in \mathbb{Z}$ are the invariant factors of $E_0$ satisfying $\beta_1 \mid \cdots \mid \beta_K$; since the columns of $E_0$ are $\mathbb{Q}$-linearly independent, it follows that $\beta_i$ are all nonzero. (Here we fixed some ordering of the $G_j$ to define $E_0$ but the invariant factors are independent of this ordering.) We shall use $\beta\left(F_1, \dots, F_K\right)$ to denote the last invariant factor $\beta_K$. We define the

Invariant Factor Hypothesis: Given $B_0 > 0$, we shall say that a positive integer $q$ satisfies $IFH(F_1, \dots, F_K; B_0)$ if $\gcd(\ell - 1, \beta(F_1, \dots, F_K)) = 1$ for any prime $\ell \mid q$ satisfying $\ell > B_0$.

*Example:* Often in applications, $\prod_{i=1}^{K} F_i$ is separable over $\mathbb{Q}$ (or more generally, the exponent matrix $E_0(F_1, \dots, F_K)$ is equivalent to the diagonal matrix $\operatorname{diag}(1, \dots, 1)$); when this happens, $\beta(F_1, \dots, F_K) = 1$, so *any* integer satisfies $IFH(F_1, \dots, F_K; B_0)$ for any $B_0 > 0$.

2.3. **Set-up for the main results.** Most of the set-up for the main results has already been done before Theorem N, however there is some additional notation, so for convenience of the reader, we state the complete set-up below:

- Consider multiplicative functions $f_1, \dots, f_K \colon \mathbb{N} \to \mathbb{Z}$ and polynomials $\{W_{i,v}\}_{\substack{1\le i\le K\\ 1\le v\le V}} \subset \mathbb{Z}[T]$ satisfying $f_i(p^v) = W_{i,v}(p)$ for any prime $p$, any $i \in [K]$ and $v \in [V]$.
- Let $f \coloneqq \prod_{i=1}^{K} f_i$ and $W_v \coloneqq \prod_{i=1}^{K} W_{i,v}$, so $f(p^v) = W_v(p)$ for all primes $p$ and all $v$.
- For each $v \in [V]$, define $D_v \coloneqq \deg W_v = \sum_{i=1}^{K} \deg W_{i,v}$. Also let $D \coloneqq D_k$, and $D_{\min} \coloneqq \min_{1\le i\le K} \deg W_{i,k}$.
- For any $q$ and $v \in [V]$, define $R_v(q) = \{u \in U_q : W_v(u) \in U_q\}$ and $\alpha_v(q) \coloneqq \frac{1}{\varphi(q)} \# R_v(q)$.
- Fix $k \in [V]$, and say that $q$ is $k$-admissible if $R_k(q) = \emptyset$ but $R_v(q) \neq \emptyset$ for all $v < k$.

  Note that if $q$ is $k$-admissible, then $\alpha_v(q) = 0$ for $v < k$, while $\alpha_k(q) \gg_{W_k} (\log\log(3q))^{-D}$ by the Chinese Remainder Theorem and a standard argument using Mertens' Theorem.

[4] We say that a polynomial in $\mathbb{Z}[T]$ is primitive when the greatest common divisor of its coefficients is 1.

- Assume that $\{W_{i,k}\}_{1\le i\le K}$ are multiplicatively independent.
- Define $\mathcal{Q}(k; f_1, \cdots, f_K)$ exactly as before the statement of Theorem N.

### 2.4. The Main Results.

In Theorems 2.1 to 2.3 below, we fix $K_0, B_0 > 0$. Our implied constants depend only on $K_0, B_0$ and the polynomials $\{W_{i,v}\}_{\substack{1\le i\le K\\ 1\le v\le k}}$, and are in particular independent of $V$ and of $\{W_{i,v}\}_{\substack{1\le i\le K\\ k<v\le V}}$.

**Theorem 2.1.** *Fix $\epsilon \in (0,1)$. The functions $f_1, \dots, f_K$ are jointly weakly equidistributed, uniformly to all moduli $q \le (\log x)^{K_0}$ lying in $\mathcal{Q}(k; f_1, \cdots, f_K)$ and satisfying $IFH(W_{1,k}, \dots, W_{K,k}; B_0)$, provided any one of the following holds.*

(i) *Either $K = 1$ and $W_{1,k} = W_k$ is linear, or $K \ge 2$, $q \le (\log x)^{(1-\epsilon)\alpha_k(q)/(K-1)}$ and at least one of $\{W_{i,k}\}_{1\le i\le K}$ is linear (i.e., $D_{\min} = 1$).*

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

(iii) *$D_{\min} > 1$ and $q \le (\log x)^{(1-\epsilon)\alpha_k(q)(K-1/D_{\min})^{-1}}$.*

**A concrete application:** By [28, Theorem 1], $\varphi(n)$ and $\sigma(n)$ are jointly WUD modulo a fixed integer $q$ precisely when $q$ is coprime to 6; in fact, $\mathcal{Q}(1; \varphi, \sigma) = \{q : (q,6) = 1\}$. Theorem 2.1 shows that $(\varphi, \sigma)$ are jointly WUD uniformly modulo $q \le (\log x)^{(1-\epsilon)\alpha(q)}$ coprime to 6, where $\alpha(q) := \alpha_1(q) = \prod_{\ell | q}(\ell - 3)/(\ell - 1)$ and $\epsilon > 0$ is fixed but arbitrary.

**Optimality of the conditions in Theorem 2.1:** In subsection § 8.1, we will show that except in the very first case when $K = 1$ and $W_k = W_{1,k}$ is linear, the ranges of $q$ in (i)–(iii) above are all essentially optimal. We will also show that for $K \ge 2$, the range of $q$ in (i) is essentially optimal, even if $q$ is squarefree and $\{W_{i,k}\}_{1\le i\le K}$ are *all* linear, for *any* choice of (pairwise coprime) linear functions! In particular, this means that the aforementioned range $(\log x)^{(1-\epsilon)\alpha(q)}$ is basically optimal for the joint weak equidistribution of $\varphi$ and $\sigma$, even if we restrict to squarefree $q$. Thus the special case of Theorem 2.1(i) for the family $(\varphi, \sigma)$ is the best possible uniform analogue of Narkiewicz's result in [28] for a single varying modulus.

**Restoring uniformity in the Siegel–Walfisz range:**
Our constructions in § 8.1 will reveal that obstructions to uniformity in $q$ come from inputs $n$ of the form $P^k$ for primes $P$. Modifying those constructions, we can produce more obstructions of the form $mP^k$ with $m$ fixed or growing slowly with $x$. It turns out that the problematic inputs in general are those with too few large prime factors. More precisely, uniformity in the full Siegel-Walfisz range $q \le (\log x)^{K_0}$ is restored if we restrict attention to those $n$ that are divisible by sufficiently many primes exceeding $q$.

To make this precise, let $P_1(n) := P(n)$ denote the largest prime divisor of $n$, with the convention that $P(1) := 1$. Inductively define $P_m(n) := P_{m-1}(n/P(n))$, so that $P_m(n)$ is the $m$-th largest prime factor of $n$ (counted with multiplicity), with $P_m(n) = 1$ if $\Omega(n) < m$. Since $D = 1$ forces $K = 1$ and $W_k = W_{1,k}$ to be linear (a case in which Theorem 2.1(i) already gives complete uniformity in $q \le (\log x)^{K_0}$), we assume in Theorems 2.2 and 2.3 below that $D \ge 2$.

**Theorem 2.2.** *As $x \to \infty$ and uniformly in coprime residues $a_1, \dots, a_K$ to moduli $q \le (\log x)^{K_0}$ lying in $\mathcal{Q}(k; f_1, \cdots, f_K)$ and satisfying $IFH(W_{1,k}, \dots, W_{K,k}; B_0)$, we have*

$$\begin{aligned} &\#\{n \le x : P_R(n) > q,\ (\forall i)\ f_i(n) \equiv a_i \pmod q\} \\ &\quad \sim \frac{1}{\varphi(q)^K}\#\{n \le x : \gcd(f(n), q) = 1\} \sim \frac{1}{\varphi(q)^K}\#\{n \le x : P_R(n) > q, \gcd(f(n), q) = 1\}, \end{aligned} \tag{2.2}$$

*where*

$$\begin{cases} R = k(KD+1), & \text{if } k < D \\ R \text{ is the least integer exceeding } k\left(1 + (k+1)\left(K - 1/D\right)\right), & \text{if } k \ge D. \end{cases}$$

Even in the special case $k = K = 1$, this theorem improves over Theorem 1.4(a) in [38]. The value of $R$ is optimal for $K = 1$ and $f_1(n) = \sigma(n)$ modulo even $q$; see the discussion on applications in subsection § 2.6. For squarefree $q$, it suffices to have much weaker restrictions on $n$ (that are also exactly or nearly optimal) to restore uniformity in the Siegel–Walfisz range.

**Theorem 2.3.** *The formulae* (2.2) *hold as $x \to \infty$, uniformly in coprime residues $a_1, \dots, a_K$ modulo $\underline{\text{squarefree}}$ $q \le (\log x)^{K_0}$ lying in $\mathcal{Q}(k; f_1, \cdots, f_K)$ and satisfying $IFH(W_{1,k}, \dots, W_{K,k}; B_0)$, with*

$$R := \begin{cases} 2, & \text{if } K = k = 1 \text{ and } W_{1,1} \text{ is not squarefull.} \\ k(Kk + K - k) + 1, & \text{if } k > 1 \text{ and at least one of } \{W_{i,k}\}_{1 \le i \le K} \text{ is not squarefull.} \\ k(Kk + K - k + 1) + 1, & \text{in general.} \end{cases}$$

Here we write a polynomial $F \in \mathbb{Z}[T]$ as $F = r\prod_{j=1}^M H_j^{\nu_j}$ for some $\nu_j \in \mathbb{N}$ and pairwise coprime primitive irreducibles $H_j \in \mathbb{Z}[T]$, and we say that $F$ is "squarefull" (in $\mathbb{Z}[T]$) if $(\prod_{j=1}^M H_j)^2 \mid F$. Note that this is equivalent to saying that $\prod_{\substack{\theta \in \mathbb{C} \\ F(\theta)=0}} (T - \theta)^2 \mid F(T)$ in $\mathbb{C}[T]$, i.e., that every root of $F$ in $\mathbb{C}$ has multiplicity at least 2.

It is worthwhile to strive for the optimality of $R$ above since doing so ensures weak equidistribution among the largest possible set of inputs $n$. In subsection § 11.1, we show that the first two values of $R$ in Theorem 2.3 are exactly optimal, in the sense that in order to have uniformity in $q \le (\log x)^{K_0}$, it is not possible to reduce the "2" to "1" or the "$k(Kk+K-k)+1$" to "$k(Kk + K - k)$". In these examples, $\{W_{i,k}\}_{i=1}^K$ will be pairwise coprime irreducibles, making $\prod_{i=1}^K W_{i,k}$ separable over $\mathbb{Q}$.

### 2.5. Necessity of the multiplicative independence and invariant factor hypotheses.

We now explain the necessity of these two hypotheses that we have been assuming in our results so far. It turns out that even if one of them is violated, then uniformity would fail in the above theorems in some of the worst possible ways: Not only would uniformity fail modulo arbitrarily large $q \le (\log x)^{K_0}$, but also would be *unrecoverable* no matter how much we restrict our set of inputs $n$ to those having many large prime factors! This substantiates our previous comment on Theorems 2.1 through 2.3 being essentially best possible qualitative analogues of the Siegel–Walfisz theorem for families of polynomially-defined multiplicative functions.

For instance, without the multiplicative independence condition, the $K$ congruences $f_i(n) \equiv a_i \pmod q$ (for $1 \le i \le K$) may degenerate to fewer congruences for sufficiently many inputs $n$, making weak equidistribution fail uniformly to *all* sufficiently large $q$, *no matter how much* we restrict the set of inputs $n$ to those having sufficiently many large prime factors.

**Theorem 2.4.** *Fix $R \geq 1$, $K > 1$ and assume that $\{W_{i,k}\}_{1 \leq i \leq K-1} \subset \mathbb{Z}[T]$ are multiplicatively independent, with $\sum_{i=1}^{K-1} \deg W_{i,k} > 1$. Suppose $W_{K,k} = \prod_{i=1}^{K-1} W_{i,k}^{\lambda_i}$ for some nonnegative integers $(\lambda_i)_{i=1}^{K-1} \neq (0, \ldots, 0)$. There exists a constant $C := C(W_{1,k}, \ldots, W_{K-1,k}) > 0$ such that*

$$\#\{n \leq x : P_{Rk}(n) > q,\ (\forall i \in [K])\ f_i(n) \equiv a_i \pmod{q}\} \gg \frac{1}{\varphi(q)^{K-1}} \cdot \frac{x^{1/k}(\log\log x)^{R-2}}{\log x}$$

*as $x \to \infty$, uniformly in $k$-admissible $q \leq (\log x)^{K_0}$ supported on primes $\ell > C$ satisfying $\gcd(\ell - 1, \beta(W_{1,k}, \ldots, W_{K-1,k})) = 1$, and in $a_i \in U_q$ with $a_K \equiv \prod_{i=1}^{K-1} a_i^{\lambda_i} \pmod{q}$.*

The compatibility of the relations in $\{W_{i,k}\}_{1 \leq i \leq K}$ and $(a_i)_{i=1}^{K}$ suggests why the $K$ congruences degenerate to $K-1$ congruences. Turning to the invariant factor hypothesis, we claim that the failure of this condition incurs an additional factor over the expected proportion of $n \leq x$ satisfying $\gcd(f(n), q) = 1$. For certain choices of $q$ and $\{W_{i,k}\}_{1 \leq i \leq K}$, this factor can be made too large, once again leading to an overrepresentation of the tuple $(a_i \bmod q)_{i=1}^{K}$ by the multiplicative functions $f_1, \ldots, f_K$. In what follows, $P^-(q)$ denotes the least prime dividing $q$.

**Theorem 2.5.** *Fix $R \geq 1$ and assume that $\{W_{i,k}\}_{1 \leq i \leq K} \subset \mathbb{Z}[T]$ are nonconstant, monic and multiplicatively independent, so that $\beta = \beta(W_{1,k}, \ldots, W_{K,k}) \in \mathbb{Z} \setminus \{0\}$. There exists a constant $C := C(W_{1,k}, \ldots, W_{K,k}) > 0$ such that*

$$\#\{n \leq x : P_{Rk}(n) > q,\ (\forall i \in [K])\ f_i(n) \equiv a_i \pmod{q}\} \gg \frac{2^{\#\{\ell \mid q:\ \gcd(\ell-1,\beta) \neq 1\}}}{\varphi(q)^K} \cdot \frac{x^{1/k}(\log\log x)^{R-2}}{\log x} \tag{2.3}$$

*as $x \to \infty$, uniformly in $k$-admissible $q \leq (\log x)^{K_0}$ having $P^-(q) > C$, and in coprime residues $(a_i)_{i=1}^{K}$ mod $q$ which are all congruent to $1$ modulo the largest squarefree divisor of $q$.*

We shall formally establish Theorems 2.4 and 2.5 in our sequel note [48].

### 2.6. Some more concrete applications of our main results.

We give several applications of our main results to arithmetic functions of common interest. Recall Śliwa's [49] result that $\sigma(n)$ is weakly equidistributed precisely to moduli that are not multiples of 6; in fact, his result shows that $\mathcal{Q}(1; \sigma) = \{q : \gcd(q, 2) = 1\}$ and $\mathcal{Q}(2; \sigma) = \{q : \gcd(q, 6) = 2\}$. By Theorem 2.1(i), $\sigma(n)$ is WUD uniformly to all odd moduli $q \leq (\log x)^{K_0}$. Calling the members of the set $\mathcal{Q}(2; \sigma)$ "special", Theorem 2.1(ii) and (iii) show that $\sigma(n)$ is WUD uniformly to all special $q \leq (\log x)^{(2-\delta)\widetilde{\alpha}(q)}$ and also to all squarefree special $q \leq (\log x)^{K_0}$ satisfying $2^{\omega(q)} \leq (\log x)^{(1-\epsilon)\widetilde{\alpha}(q)}$, where $\widetilde{\alpha}(q) := \alpha_2(q) = \prod_{\substack{\ell \mid q \\ \ell \equiv 1 \pmod{3}}} (1 - 2/(\ell - 1))$. By the example constructed in [47, subsection 7.1], the latter restriction is optimal. Furthermore, by Theorem 2.2 (resp. 2.3), uniformity is restored to *all* (resp. to squarefree) special $q \leq (\log x)^{K_0}$ by restricting to inputs $n$ with $P_6(n) > q$ (resp. $P_4(n) > q$). [5] By the examples constructed in [47], both of these restrictions are optimal as well.

[5] Here we have noted that the condition $P_3(n) > q$ forces $P_4(n) > q$ since for $\sigma(n)$ to be coprime to the even number $q$, it is necessary for $n$ to be of the form $m^2$ or $2m^2$.

Another example: we saw using Theorem 2.1 that $\varphi(n)$ and $\sigma(n)$ are jointly WUD modulo $q \le (\log x)^{(1-\epsilon)\alpha(q)}$ coprime to 6, and that these two restrictions on $q$ are necessary and essentially optimal. By Theorem 2.2, complete uniformity is restored to all moduli $q \le (\log x)^{K_0}$ coprime to 6 by restricting to inputs $n$ with $P_5(n) > q$. Likewise, we can get interesting consequences of Theorems 2.1, 2.2 and 2.3 for the families $(\varphi, \sigma_3)$, $(\varphi, \sigma, \sigma_2)$, $(\varphi, \sigma, \sigma_2, \sigma_3)$ and so on.

We can give more applications of our main results to study the weak equidistribution of the Fourier coefficients of Eisenstein series; more generally, the functions $\sigma_r(n) := \sum_{d|n} d^r$ (for $r > 1$). An easy check shows that the polynomial $\sum_{0 \le j \le v} T^{rj} = \frac{T^{r(v+1)}-1}{T^r-1}$ shares no roots with its derivative, hence is separable. Calling the $q \in \mathcal{Q}(k; \sigma_r)$ as "$k$-special", Theorem 2.1 thus shows that $\sigma_r$ is WUD uniformly modulo all $k$-special $q \le (\log x)^{(1-\epsilon)\alpha_k(q)(1-1/kr)^{-1}}$, and modulo all squarefree $k$-special $q \le (\log x)^{K_0}$ having $\omega(q) \le (1-\epsilon)\alpha_k(q) \log\log x / \log(kr)$. Further, by Theorems 2.2 and 2.3, weak equidistribution is restored modulo all $k$-special (resp. squarefree $k$-special) $q \le (\log x)^{K_0}$ by restricting to $n$ with $P_{k(kr+1)}(n) > q$ (resp. $P_{k+1}(n) > q$).

An explicit characterization of the moduli $q \le (\log x)^{K_0}$ to which a given $\sigma_r$ is weakly equidistributed thus reduces to an understanding of the possible $k$ and of the set $\mathcal{Q}(k; \sigma_r)$ for a given (fixed) $r$; both of these are problems of fixed moduli that (as mentioned in the introduction) have been studied in great depth in [49], [15], [32], [30], [31], [40] and [41]. In fact, the sets $\mathcal{Q}(k; \sigma_r)$ have been explicitly characterized for all odd $r \le 200$ and all even $r \le 50$, and partial results are known for general $r \ge 4$. For example, the only two possible $k$'s for $\sigma_3$ are $k = 1, 2$, and $\mathcal{Q}(1; \sigma_3) = \{q : \gcd(q, 14) = 1\}$ while $\mathcal{Q}(2; \sigma_3) = \{q : \gcd(q, 6) = 2\}$.

For a general family $(f_1, \dots, f_K)$ of polynomially–defined multiplicative functions, Narkiewicz [28, 31] gives algorithms to determine the sets $\mathcal{Q}(k; f_1, \cdots, f_K)$ for a fixed $k$. He shows (among other results) that in some of the most commonly occurring cases (which includes the cases of $\sigma_r$ for all $r > 2$), the set of possible $k$ is finite, and that for each such $k$, we can describe $\mathcal{Q}(k; f_1, \dots, f_K)$ via (finitely many) coprimality restrictions that can be determined effectively.

We conclude this section with the remark that although for the sake of simplicity of statements, we have been assuming that our multiplicative functions $\{f_i\}_{i=1}^K$ and polynomials $\{W_{i,v}\}_{\substack{1 \le i \le K \\ 1 \le v \le V}}$ are both fixed, our proofs will reveal that these results are also uniform in the $\{f_i\}_{i=1}^K$ as long as they are defined by the fixed polynomials $\{W_{i,v}\}_{\substack{1 \le i \le K \\ 1 \le v \le V}}$.

### 2.7. **Notation and conventions.**

- We do not consider the zero function as multiplicative, so if $f$ is multiplicative, then $f(1) = 1$.
- Given $z > 0$, we say that a positive integer $n$ is $z$-smooth if $P(n) \le z$, and $z$-rough if $P^-(n) > z$. By the $z$-smooth part (resp. $z$-rough part) of $n$, we shall mean the largest $z$-smooth (resp. $z$-rough) positive integer dividing $n$.
- For a ring $R$, $R^\times$ denotes the multiplicative group of units of $R$. Write $U_q := (\mathbb{Z}/q\mathbb{Z})^\times$.
- We denote the number of primes dividing $q$ counted with and without multiplicity by $\Omega(q)$ and $\omega(q)$ respectively.
- For a Dirichlet character $\chi$ mod $q$, we use $\mathfrak{f}(\chi)$ to denote the conductor of $\chi$.

- When there is no danger of confusion, we shall write $(a_1, \dots, a_k)$ in place of $\gcd(a_1, \dots, a_k)$.
- Throughout, the letters $p$ and $\ell$ are reserved for primes.
- For nonzero $H \in \mathbb{Z}[T]$, we use $\mathrm{ord}_\ell(H)$ to denote the highest power of $\ell$ dividing all the coefficients of $H$; for an integer $m \neq 0$, we may use $v_\ell(m)$ in place of $\mathrm{ord}_\ell(m)$.
- Let $\mathbb{M}_{A\times B}(\mathbb{Z})$ denote the ring of $A \times B$ matrices with integer entries, while $GL_{A\times B}(\mathbb{Z})$ refer to the group of units of $\mathbb{M}_{A\times B}(\mathbb{Z})$, i.e. the matrices with determinant $\pm 1$.
- Implied constants in $\ll$ and $O$-notation, as well as implicit constants in qualifiers like "sufficiently large", may always depend on any parameters declared as "fixed"; in particular, they will always depend on the polynomials $\{W_{i,v}\}_{\substack{1\le i\le K\\ 1\le v\le k}}$. Other dependence will be noted explicitly (for example, with parentheses or subscripts): Notably, we shall use $C(F_1, \dots, F_K)$, $C'(F_1, \dots, F_K)$ and so on, to denote constants depending on the fixed polynomials $F_1, \dots, F_K$.
- We write $\log_k$ for the $k$-th iterate of the natural logarithm.

## 3. Technical preparation: The number of $n \le x$ for which $\gcd(f(n), q) = 1$

In this section, we shall provide a rough estimate on the count of $n \le x$ for which $f(n) = \prod_{i=1}^K f_i(n)$ is coprime to the modulus $q$, uniformly in $q \le (\log x)^{K_0}$. We will show the following estimate, which generalizes Proposition 2.1 in [38]. In the rest of the paper, we abbreviate $\alpha_v(q)$ to $\alpha_v$ for each $v \in [V]$.

**Proposition 3.1.** *For all sufficiently large $x$ and uniformly in $k$-admissible $q \le (\log x)^{K_0}$,*

$$\sum_{\substack{n\le x\\ (f(n),q)=1}} 1 \;=\; \sum_{\substack{n\le x\\ \text{each } (f_i(n),q)=1}} 1 = \frac{x^{1/k}}{(\log x)^{1-\alpha_k}} \exp(O((\log_2(3q))^{O(1)})). \tag{3.1}$$

3.1. **Proof of the lower bound.** Any $m \le x^{1/k}$ satisfying $\gcd(f(m^k), q) = 1$ is certainly counted in the left hand side of (3.1). To estimate the number of such $m$, we apply [38, Proposition 2.1], with $f(n^k)$ and $x^{1/k}$ playing the roles of "$f(n)$" and "$x$" in the quoted proposition. This shows that the sum in (3.1) is bounded below by the right hand side.

3.2. **Proof of the upper bound.** We start by giving an upper bound on the count of $r$-full smooth numbers; here we consider any $n \in \mathbb{N}$ to be 1-full (and we consider 1 as being $r$-full for any $r \ge 1$). The case $r = 1$ of the lemma below is a classical bound on smooth numbers.

**Lemma 3.2.** *Fix $r \in \mathbb{N}$. We have as $X, Z \to \infty$,*

$$\#\{n \le X : P(n) \le Z,\ n \text{ is } r\text{-full}\} \ll X^{1/r}(\log Z) \exp\left(-\frac{U}{r}\log U + O(U \log_2(3U))\right),$$

*uniformly for $(\log X)^{\max\{3,2r\}} \le Z \le X^{1/2}$, where $U := \log X/\log Z$.*

*Proof of Lemma 3.2.* The lemma is a classical application of Rankin's trick. We start by letting $\eta \le \min\{1/3, 1/2r\}$ be a positive parameter to be chosen later, and observe that

$$\sum_{\substack{n \le X:\ P(n) \le Z \\ n \text{ is } r\text{-full}}} 1 \le \sum_{\substack{n \text{ is } r\text{-full} \\ P(n) \le Z}} \left(\frac{X}{n}\right)^{(1-\eta)/r} \ll X^{(1-\eta)/r} \exp\left(\sum_{p \le Z} \frac{1}{p^{1-\eta}}\right), \tag{3.2}$$

where we have used the Euler product and noted that $\sum_p \sum_{v \ge r+1} p^{-v(1-\eta)/r} \ll \sum_p p^{-(1-\eta)(1+1/r)} \ll_r 1$ since $(1-\eta)(1+1/r) \ge (1+1/r)(1-\min\{1/3, 1/2r\}) > 1$.

Now set $\eta := \frac{\log U}{\log Z} \le \min\left\{\frac{1}{3}, \frac{1}{2r}\right\}$. We write $\sum_{p \le Z} 1/p^{1-\eta} = \log_2 Z + \sum_{p \le Z}(\exp(\eta \log p) - 1)/p + O(1)$. Since $\eta \log p \le \log 2 \ll 1$ for all $p \le 2^{1/\eta}$, we find that the contribution of $p \le 2^{1/\eta}$ to the last sum above is $\sum_{p \le 2^{1/\eta}}(\exp(\eta \log p) - 1)/p \ll \eta \sum_{p \le 2^{1/\eta}} \log p/p \ll 1$, while the contribution of $p \in (2^{1/\eta}, Z]$ is at most $(\exp(\eta \log Z) - 1)\sum_{2^{1/\eta} < p \le Z} 1/p \ \le\ U(\log_2 U + O(1))$. Collecting estimates, we obtain $\sum_{p \le Z} 1/p^{1-\eta} = \log_2 Z + O(U \log_2(3U))$, which from (3.2) completes the proof of the lemma. □

The following important observation will be useful throughout the paper.

**Lemma 3.3.** *If $q$ is $k$-admissible, then the $k$-free part of any positive integer $n$ satisfying $\gcd(f(n), q) = 1$ is bounded. More precisely, it is of size $O(1)$, where the implied constant depends only on the polynomials $\{W_{i,v}\}_{\substack{1 \le i \le K \\ 1 \le v \le k}}$.*

*Proof.* Let $S_v := \{\ell \text{ prime } : \alpha_v(\ell) = 0\}$. (Recall $\alpha_v$ and $W_v$ from § 2.3.) Note the following:

**Observation 1. For each $1 \le v < k$, the set $S_v$ consists only of primes of size $O(1)$, with the implied constant depending only on the polynomials $W_{1,v}, \dots, W_{K,v}$:**
This is because for any prime $\ell$, we have $\alpha_v(\ell) = \frac{1}{\varphi(\ell)} \#\{u \in U_\ell : W_v(u) \in U_\ell\} \ge 1 - D_v/(\ell - 1)$. Thus, $\alpha_v(\ell) > 0$ for all $\ell > 1 + D_v = 1 + \sum_{i=1}^K \deg W_{i,v}$.

**Observation 2. For any positive integer $n$ satisfying $\gcd(f(n), q) = 1$, the $k$-free part of $n$ must only be divisible by primes from $\bigcup_{1 \le v < k} S_v$:**
Assume by way of contradiction, that there exists some $n$ satisfying $\gcd(f(n), q) = 1$ and some prime $p \notin \bigcup_{1 \le v < k} S_v$ satisfying $p^r \parallel n$ for some $r < k$. Then $W_r(p) = f(p^r)$ divides $f(n)$. Since $q$ is $k$-admissible and $r < k$, we must have $\alpha_r(q) = 0$. But since $\alpha_r(q) = \prod_{\ell \mid q} \alpha_r(\ell)$ by the Chinese Remainder Theorem, it follows that there must be some prime $\ell_0 \mid q$ for which $\alpha_r(\ell_0) = 0$. By definition of $\alpha_r$, this means that for any unit $u \in U_{\ell_0}$, we must have $\ell_0 \mid W_r(u)$. In particular, since the prime $p$ above does not lie in $S_r$ while $\ell_0$ does, it follows that $p \ne \ell_0$, so that $\ell_0 \mid W_r(p) \mid f(n)$, contradicting the requirement that $\gcd(f(n), q) = 1$.

Lemma 3.3 follows immediately Observations 1 and 2. □

We will also need the following estimate, which is a restatement of [38, Lemma 2.4].

**Lemma 3.4.** *Let $G \in \mathbb{Z}[T]$ be a fixed nonconstant polynomial. For each positive integer $q$, let $\alpha_G(q) := \frac{1}{\varphi(q)}\#\{u \in U_q : G(u) \in U_q\}$. We have, uniformly in $q$ and $x \geq 3q$,*

$$\sum_{p \leq x} \frac{\mathbb{1}_{(G(p),q)=1}}{p} = \alpha_G(q)\log_2 x + O((\log_2 (3q))^{O(1)}).$$

Coming to the proof of the upper bound implied in (3.1), we define $y := \exp(\sqrt{\log x})$ and start by removing those $n$ which are divisible by the $(k+1)$-th power of a prime exceeding $y$. Writing any such $n$ as $AB$ for some $k$-free $B$ and $k$-full $A$, Lemma 3.3 shows that $B \ll 1$ so that the contribution of such $n$ to (3.1) is

(3.3)
$$\sum_{\substack{n\leq x:\ (f(n),q)=1\\ \exists\ p>y:\ p^{k+1}\mid n}} 1 \ll \sum_{\substack{A\leq x\\ A \text{ is } k\text{-full}\\ \exists\ p>y:\ p^{k+1}\mid n}} 1 \leq \sum_{p>y}\sum_{\substack{v\geq k+1\\ p^v\leq x}}\sum_{\substack{m\leq x/p^v\\ m \text{ is } k\text{-full}}} 1 \ll \sum_{p>y}\sum_{v\geq k+1}\left(\frac{x}{p^v}\right)^{1/k} \ll \left(\frac{x}{y}\right)^{1/k},$$

where we have used the fact that the number of $k$-full integers up to $X$ is $O(X^{1/k})$ (see [14]). The last expression above is negligible in comparison to the right hand side of (3.1). Hence, it remains to bound the number of $n$ satisfying $(f(n), q) = 1$ that are not divisible by the $(k+1)$-th power of any prime exceeding $y$.

We write any such $n$ in the form $BMN$, where $N$ is $y$-rough, $BM$ is $y$-smooth, $B$ is $k$-free, $M$ is $k$-full, and $B, M, N$ are pairwise coprime. By Lemma 3.3, we see that $B = O(1)$ and that $N$ is $k$-full. But also since $n$ is not divisible by the $(k+1)$-th power of any prime exceeding $y$, we must have $N = A^k$ for some squarefree $y$-rough integer $A$. Consequently,

$$\sum_{\substack{n\leq x:\ (f(n),q)=1\\ p>y \implies p^{k+1}\nmid n}} 1 \leq \sum_{\substack{B\leq x\\ (f(B),q)=1\\ B \text{ is } k\text{-free}}}\ \sum_{\substack{M\leq x/B:\ M \text{ is } k\text{-full}\\ P(M)\leq y,\ (f(M),q)=1}}\ \sum_{\substack{A\leq (x/BM)^{1/k}\\ P^-(A)>y:\ (f(A^k),q)=1\\ A \text{ squarefree}}} 1. \tag{3.4}$$

We now write the right hand side of the above inequality as $\Sigma_1 + \Sigma_2$, where $\Sigma_1$ and $\Sigma_2$ count the contribution of $(B, M, A)$ with $M \leq x^{1/2}$ and $M > x^{1/2}$, respectively.

*Bounding* $\Sigma_2$*:* Any $A$ counted in $\Sigma_2$ satisfies $A \leq (x/BM)^{1/k} \leq x^{1/2k}/B^{1/k}$, so that

$$\Sigma_2 \leq \sum_{\substack{B\leq x\\ (f(B),q)=1\\ B \text{ is } k\text{-free}}}\ \sum_{\substack{A\leq x^{1/2k}/B^{1/k}\\ P^-(A)>y:\ (f(A^k),q)=1\\ A \text{ squarefree}}}\ \sum_{\substack{M\leq x/BA^k:\ P(M)\leq y\\ M \text{ is } k\text{-full},\ (f(M),q)=1}} 1.$$

To bound the innermost sum, we invoke Lemma 3.2; here $U = \frac{\log(x/BA^k)}{\log y} \geq \frac{1}{2}\sqrt{\log x}$. This yields

$$\Sigma_2 \ll \sum_{\substack{B\leq x\\ (f(B),q)=1\\ B \text{ is } k\text{-free}}}\ \sum_{\substack{A\leq x^{1/2k}/B^{1/k}\\ P^-(A)>y:\ (f(A^k),q)=1\\ A \text{ squarefree}}} \frac{x^{1/k}}{B^{1/k}A}\exp\left(-\frac{1}{6k}\sqrt{\log x}\cdot\log_2 x\right).$$

Recalling that $B = O(1)$ and bounding the sum on $A$ trivially by $2\log x$, we deduce that $\Sigma_2 \ll x^{1/k}\exp\left(-\sqrt{\log x}\right)$, which is negligible compared to the right hand side of (3.1).

*Bounding* $\Sigma_1$*:* To bound the (innermost) sum on $A$ in $\Sigma_1$, we invoke [18, Theorem 01, p. 2] on the multiplicative function $g(A) := \mu(A)^2 \mathbb{1}_{P^-(A)>y} \mathbb{1}_{(f(A^k),q)=1}$, with $\mu$ denoting the Möbius function. Since $M \le x^{1/2}$ and $B \ll 1$, this gives

$$\Sigma_1 \ll \frac{x^{1/k}}{\log x} \exp\left( \sum_{y<p\le x} \frac{\mathbb{1}_{(W_k(p),q)=1}}{p} \right) \sum_{\substack{M\le x^{1/2}:\ M \text{ is } k\text{-full} \\ P(M)\le y,\ (f(M),q)=1}} \frac{1}{M^{1/k}}.$$

But since the sum on $M$ above is no more than

$$\sum_{\substack{M \text{ is } k\text{-full} \\ P(M)\le y,\ (f(M),q)=1}} \frac{1}{M^{1/k}} \le \prod_{p\le y} \left( 1 + \frac{\mathbb{1}_{(f(p^k),q)=1}}{p} + O\left( \frac{1}{p^{1+1/k}} \right) \right) \ll \exp\left( \sum_{p\le y} \frac{\mathbb{1}_{(W_k(p),q)=1}}{p} \right), \tag{3.5}$$

it follows by an estimation of $\sum_{p\le y} \mathbb{1}_{(W_k(p),q)=1}/p$ via Lemma 3.4, that $\Sigma_1$ is absorbed in the right hand side of (3.1). This establishes Proposition 3.1.

## 4. The main term in Theorems 2.1 to 2.3: Contribution of "convenient" $n$

In what follows, we define

$$J := \lfloor \log_3 x \rfloor \text{ and } y := \exp((\log x)^{\epsilon/2}),$$

where $\epsilon$ is as in the statement of Theorem 2.1 and $\epsilon := 1$ for Theorems 2.2 and 2.3. We call $n \le x$ convenient if the largest $J$ *distinct* prime divisors of $n$ exceed $y$ and each appear to exactly the $k$-th power in $n$. In other words, $n$ is convenient iff it can be uniquely written in the form $n = m(P_J \cdots P_1)^k$ for $m \le x$ and primes $P_1, \dots, P_J$ satisfying

$$L_m := \max\{y, P(m)\} < P_J < \cdots < P_1. \tag{4.1}$$

Note that any $n$ having $P_{Jk}(n) \le y$ must be inconvenient; on the other hand, if $n$ is inconvenient and satisfies $\gcd(f(n), q) = 1$ then either $P_{Jk}(n) \le y$ or $n$ is divisible by the $(k+1)$-th power of a prime exceeding $y$. These observations will be helpful in the rest of the paper.

We start by showing that there are a negligible number of inconvenient $n \le x$ satisfying $\gcd(f(n), q) = 1$.

**Proposition 4.1.** *We have as* $x \to \infty$*,*

$$\sum_{\substack{n\le x:\ (f(n),q)=1 \\ n \text{ inconvenient}}} 1 \ = \ o\left( \sum_{\substack{n\le x \\ (f(n),q)=1}} 1 \right), \tag{4.2}$$

*uniformly in* $k$*-admissible* $q \le (\log x)^{K_0}$*.*

*Proof.* By (3.3) and (3.1), the contribution of the $n$'s that are divisible by the $(k+1)$-th power of a prime exceeding $y$ is negligible. Letting $z := x^{1/\log_2 x}$, we show that the contribution of $z$-smooth $n$ to the left side of (4.2) is also negligible compared to the right. Indeed, writing any such $n$ in the form $AB$ for some $k$-free $B$ and $k$-full $A$, we have $P(A) \le z$ whereas (by Lemma 3.3) $B = O(1)$. Hence the contribution of $z$-smooth $n$ is, by Lemma 3.2,

$$\sum_{\substack{n\le x:\ P(n)\le z\\ (f(n),q)=1}} 1 \ll \sum_{\substack{A\le x:\ P(A)\le z\\ A \text{ is } k\text{-full}}} 1 \ll x^{1/k}\exp\left(-\left(\frac{1}{k}+o(1)\right)\log_2 x\log_3 x\right), \tag{4.3}$$

which is indeed negligible compared to the right hand side of (4.2).

It remains to consider the contribution of those $n$ which are neither $z$-smooth nor divisible by the $(k+1)$-th power of a prime exceeding $y$. Since $n$ is inconvenient, we must have $P_{Jk}(n) \le y$ (see the discussion just preceding the statement of this proposition). Hence, $n$ can be written in the form $mP^k$ where $P \coloneqq P(n) > z$ and $m = n/P^k$, so that $P_{Jk}(m) \le y$, $\gcd(m, P) = 1$ and $f(n) = f(m)f(P^k)$. Given $m$, there are at most $\sum_{z<P\le(x/m)^{1/k}} 1 \ll x^{1/k}/m^{1/k}\log z$ many possibilities for $P$. Consequently,

$$\sum_{\substack{n\le x \text{ inconvenient}\\ P(n)>z,\ (f(n),q)=1\\ p>y \implies p^{k+1}\nmid n}} 1 \le \sum_{\substack{n\le x:\ P_{Jk}(n)\le y\\ P(n)>z,\ (f(n),q)=1\\ p>y \implies p^{k+1}\nmid n}} 1 \ll \frac{x^{1/k}\log_2 x}{\log x}\sum_{\substack{m\le x\\ P_{Jk}(m)\le y,\ (f(m),q)=1\\ p>y \implies p^{k+1}\nmid m}} \frac{1}{m^{1/k}}. \tag{4.4}$$

As in the argument preceding (3.4), we write any $m$ occurring in the above sum (uniquely) in the form $BMA^k$, where $B$ is $k$-free, $M$ is $k$-full, $A$ is squarefree, $P(BM) \le y < P^-(A)$, and $\Omega(A) \le J$ (since $P_{Jk}(n) \le y$). Since $B = O(1)$, we deduce that

$$\sum_{\substack{m\le x\\ P_{Jk}(m)\le y,\ (f(m),q)=1\\ p>y \implies p^{k+1}\nmid m}} \frac{1}{m^{1/k}} \ll \sum_{\substack{M\ k\text{-full}\\ P(M)\le y,\ (f(M),q)=1}} \frac{1}{M^{1/k}} \sum_{\substack{A\le x\\ \Omega(A)\le J}} \frac{1}{A}.$$

The sum on $A$ is no more than $(1+\sum_{p\le x} 1/p)^J \le (2\log_2 x)^J \le \exp(O((\log_3 x)^2))$, while the sum on $M$ is $\ll \exp(\alpha_k \log_2 y + O((\log_2(3q))^{O(1)}))$ by (3.5) and Lemma 3.4. Altogether,

$$\sum_{\substack{m\le x\\ P_{Jk}(m)\le y,\ (f(m),q)=1\\ p>y \implies p^{k+1}\nmid m}} \frac{1}{m^{1/k}} \ll (\log x)^{\alpha_k\epsilon/2}\exp\left(O((\log_3 x)^2 + (\log_2(3q))^{O(1)})\right). \tag{4.5}$$

Inserting this into (4.4) and comparing with (3.1) completes the proof. □

It turns out that the convenient $n$ give the dominant contributions in our asymptotics, in the sense that it is these $n$ that give the desired main term.

**Theorem 4.2.** *Fix $K_0, B_0 > 0$ and assume that $\{W_{i,k}\}_{1\le i\le K} \subset \mathbb{Z}[T]$ are nonconstant and multiplicatively independent. As $x \to \infty$, we have*

$$\sum_{\substack{n\le x \text{ convenient}\\ (\forall i)\ f_i(n)\equiv a_i \pmod q}} 1 \sim \frac{1}{\varphi(q)^K}\sum_{\substack{n\le x\\ (f(n),q)=1}} 1,$$

*uniformly in coprime residues $a_1, \dots, a_K$ to moduli $q \le (\log x)^{K_0}$ lying in $\mathcal{Q}(k; f_1, \cdots, f_K)$ and satisfying $IFH(W_{1,k}, \dots, W_{K,k}; B_0)$.*

We shall prove this theorem in the next few sections. In this section and the next, we take the first step by showing a weaker version of this result, where we reduce the congruences $f_i(n) \equiv a_i$ from modulus $q$ to a bounded divisor of $q$.

**Proposition 4.3.** *Fix $K_0, B_0 > 0$ and assume that $\{W_{i,k}\}_{1\le i\le K} \subset \mathbb{Z}[T]$ are nonconstant and multiplicatively independent. There exists a constant $\lambda := \lambda(W_{1,k}, \ldots, W_{K,k}; B_0) > 0$ depending only on $\{W_{i,k}\}_{1\le i\le K} \subset \mathbb{Z}[T]$ and $B_0$, such that as $x \to \infty$, we have*

$$\sum_{\substack{n\le x \text{ convenient}\\ (\forall i)\ f_i(n)\equiv a_i \pmod q}} 1 \;=\; \left(\frac{\varphi(Q_0)}{\varphi(q)}\right)^K \sum_{\substack{n\le x:\ (f(n),q)=1\\ (\forall i)\ f_i(n)\equiv a_i \pmod{Q_0}}} 1 \;+\; o\!\left(\frac{1}{\varphi(q)^K}\sum_{\substack{n\le x\\ (f(n),q)=1}} 1\right), \tag{4.6}$$

*uniformly in coprime residues $a_1, \ldots, a_K$ to $k$-admissible moduli $q \le (\log x)^{K_0}$ satisfying $IFH(W_{1,k}, \ldots, W_{K,k}; B_0)$. Here $Q_0$ is some divisor of $q$ satisfying $Q_0 \le \lambda$.*

*Proof.* For any $N \ge 1$ and $(w_i)_{i=1}^K \in U_q^K$, we define

$$\mathcal{V}^{(k)}_{N,K}\left(q; (w_i)_{i=1}^K\right) := \left\{(v_1, \ldots, v_N) \in (U_q)^N :\ (\forall i \in [K]) \prod_{j=1}^N W_{i,k}(v_j) \equiv w_i \pmod q\right\}.$$

We write each convenient $n$ uniquely in the form $m(P_J \cdots P_1)^k$, where $m, P_J, \ldots, P_1$ satisfy (4.1). Then $f_i(n) = f_i(m)\prod_{j=1}^J W_{i,k}(P_j)$, so that the conditions $f_i(n) \equiv a_i \pmod q$ amount to $\gcd(f(m), q) = 1$ and $(P_1, \ldots, P_J) \bmod q \in \mathcal{V}'_{q,m} := \mathcal{V}^{(k)}_{J,K}\left(q; (a_i f_i(m)^{-1})_{i=1}^K\right)$. Noting that the conditions $P_1 \cdots P_J \le (x/m)^{1/k}$ and $(P_1, \ldots, P_J) \bmod q \in \mathcal{V}'_{q,m}$ are both independent of the ordering of $P_1, \ldots, P_J$, we obtain

$$\sum_{\substack{n\le x \text{ convenient}\\ (\forall i)\ f_i(n)\equiv a_i \pmod q}} 1 \;=\; \sum_{\substack{m\le x\\ (f(m),q)=1}} \; \sum_{(v_1,\ldots,v_J)\in\mathcal{V}'_{q,m}} \frac{1}{J!} \sum_{\substack{P_1,\ldots,P_J > L_m\\ P_1\cdots P_J \le (x/m)^{1/k}\\ P_1,\ldots,P_J \text{ distinct}\\ (\forall j)\ P_j\equiv v_j \pmod q}} 1.$$

Proceeding exactly as in [38] to remove the congruence conditions on $P_1, \ldots, P_J$ by successive applications of the Siegel–Walfisz Theorem, we deduce that

$$\sum_{\substack{P_1,\ldots,P_J > L_m\\ P_1\cdots P_J \le (x/m)^{1/k}\\ P_1,\ldots,P_J \text{ distinct}\\ (\forall j)\ P_j\equiv v_j \pmod q}} 1 \;=\; \frac{1}{\varphi(q)^J} \sum_{\substack{P_1,\ldots,P_J > L_m\\ P_1\cdots P_J \le (x/m)^{1/k}\\ P_1,\ldots,P_J \text{ distinct}}} 1 \;+\; O\left(\frac{x^{1/k}}{m^{1/k}} \exp\left(-K_1(\log x)^{\epsilon/4}\right)\right) \tag{4.7}$$

for some constant $K_1 := K_1(K_0) > 0$. Collecting estimates and noting that $\#\mathcal{V}'_{q,m} \le \varphi(q)^J \le (\log x)^{K_0 J}$, we obtain

$$\sum_{\substack{n\le x \text{ convenient}\\ (\forall i)\ f_i(n)\equiv a_i \pmod q}} 1 = \sum_{\substack{m\le x\\ (f(m),q)=1}} \frac{\#\mathcal{V}'_{q,m}}{\varphi(q)^J}\left(\frac{1}{J!}\sum_{\substack{P_1,\ldots,P_J > L_m\\ P_1\cdots P_J \le (x/m)^{1/k}\\ P_1,\ldots,P_J \text{ distinct}}} 1\right) + O\left(x^{1/k}\exp\left(-\frac{K_1}{2}(\log x)^{\epsilon/4}\right)\right). \tag{4.8}$$

Here in the last step we have crudely bounded the sum $\sum_{\substack{m\le x\\ (f(m),q)=1}} m^{-1/k}$ by writing each $m$ as $AB$ for some $k$-full $A$ and $k$-free $B$ satisfying $\gcd(A,B)=1$, and recalling that $B=O(1)$ while $\sum 1/A \le \prod_{p\le x}\left(1+1/p+O\left(1/p^{1+1/k}\right)\right)$. The following proposition estimates $\#\mathcal{V}'_{q,m}$. Note that it actually involves *only* $B_0$ and the polynomials $\{W_{i,k}\}_{1\le i\le K}$, nothing else.

**Proposition 4.4.** *Assume that $\{W_{i,k}\}_{1\le i\le K}$ are multiplicatively independent. Define the quantities $D=\sum_{i=1}^K \deg W_{i,k}$ and $\alpha_k(q)=\frac{1}{\varphi(q)}\#\{u\in U_q : \prod_{i=1}^K W_{i,k}(u)\in U_q\}$ as before.*

*There exists a constant $C_0 \coloneqq C_0(W_{1,k},\dots,W_{K,k};B_0) > (8D)^{2D+2}$ depending only on $\{W_{i,k}\}_{1\le i\le K}$ and $B_0$, such that for any constant $C>C_0$, the following two estimates hold uniformly in coprime residues $(w_i)_{i=1}^K$ to moduli $q$ satisfying $\alpha_k(q)\neq 0$ and $IFH(W_{1,k},\dots,W_{K,k};B_0)$:*

$$\begin{aligned}(4.9)\quad &\frac{\#\mathcal{V}^{(k)}_{N,K}\left(q;(w_i)_{i=1}^K\right)}{\varphi(q)^N}\\ &= \frac{\alpha_k(q)^N}{\alpha_k(Q_0)^N}\left(\frac{\varphi(Q_0)}{\varphi(q)}\right)^K \left\{\frac{\#\mathcal{V}^{(k)}_{N,K}\left(Q_0;(w_i)_{i=1}^K\right)}{\varphi(Q_0)^N}+O\left(\frac{1}{C^N}\right)\right\}\prod_{\substack{\ell\mid q\\ \ell>C_0}}\left(1+O\left(\frac{(4D)^N}{\ell^{N/D-K}}\right)\right),\end{aligned}$$

*uniformly for $N\ge KD+1$, where $Q_0$ is a $C_0$-smooth divisor of $q$ of size $O_C(1)$. Moreover*

$$(4.10)\qquad \frac{\#\mathcal{V}^{(k)}_{N,K}\left(q;(w_i)_{i=1}^K\right)}{\varphi(q)^N} \le \frac{\left(\prod_{\ell^e\|q} e\right)^{\mathbb{1}_{N=KD}}}{q^{N/D}}\ \exp\left(O(\omega(q))\right),\quad \textit{for each } 1\le N\le KD.$$

Applying (4.9) with $N\coloneqq J=\lfloor \log_3 x\rfloor \ge KD+1$, and with $C$ fixed to be a constant exceeding $2C_0^{C_0}$, we see that

$$\frac{\#\mathcal{V}'_{q,m}}{\varphi(q)^J} = (1+o(1))\frac{\alpha_k(q)^J}{\alpha_k(Q_0)^J}\left(\frac{\varphi(Q_0)}{\varphi(q)}\right)^K\left\{\frac{\#\mathcal{V}'_{Q_0,m}}{\varphi(Q_0)^J}+O\left(\frac{1}{C^J}\right)\right\},$$

where $\mathcal{V}'_{Q_0,m} \coloneqq \mathcal{V}^{(k)}_{J,K}\big(Q_0;(a_if_i(m)^{-1})_{i=1}^K\big)$ and we have noted that $\sum_{\substack{\ell\mid q\\ \ell>C_0}}(4D)^J/\ell^{J/D-K} \le \left(4D/C_0^{1/(2D+2)}\right)^J = o(1)$. We insert this into (4.8), and observe that since $\alpha_k(q)\neq 0$, since $Q_0\mid q$ and since $Q_0$ is $C_0$-smooth, we have $\alpha_k(Q_0)C \ge C\prod_{\ell\le C_0}\left(1-\frac{\ell-2}{\ell-1}\right) \ge \frac{C}{C_0^{C_0}} \ge 2$. Thus

$$\begin{aligned}(4.11)\quad &\sum_{\substack{n\le x \text{ convenient}\\ (\forall i)\ f_i(n)\equiv a_i \pmod q}} 1\\ &= (1+o(1))\left(\frac{\varphi(Q_0)}{\varphi(q)}\right)^K\frac{\alpha_k(q)^J}{\alpha_k(Q_0)^J}\sum_{\substack{m\le x\\ (f(m),q)=1}}\frac{\#\mathcal{V}'_{Q_0,m}}{\varphi(Q_0)^J}\Bigg(\frac{1}{J!}\sum_{\substack{P_1,\dots,P_J>L_m\\ P_1\cdots P_J\le (x/m)^{1/k}\\ P_1,\dots,P_J \text{ distinct}}} 1\Bigg)+o\Bigg(\frac{1}{\varphi(q)^K}\sum_{\substack{n\le x\\ (f(n),q)=1}} 1\Bigg),\end{aligned}$$

where we have noted that
(4.12)

$$\sum_{\substack{n\le x \text{ convenient}\\ \gcd(f(n),q)=1}} 1 \;= \alpha_k(q)^J \sum_{\substack{m\le x\\ (f(m),q)=1}} \left(\frac{1}{J!}\sum_{\substack{P_1,\dots,P_J>L_m\\ P_1\cdots P_J\le (x/m)^{1/k}\\ P_1,\dots,P_J \text{ distinct}}} 1\right) \;+\; O\left(x^{1/k}\exp\left(-\frac{K_1}{2}(\log x)^{\epsilon/4}\right)\right);$$

the above estimate can be proven by replicating the arguments leading to (4.8) and observing that $\#\{(v_1,\dots,v_J)\in U_q^J:\ \prod_{j=1}^J W_k(v_j)\in U_q\} = (\alpha_k(q)\varphi(q))^J$.

Now for each $(w_i)_{i=1}^K \in U_q^K$, we define $\mathcal{U}_{J,K}\big(q,Q_0;(w_i)_{i=1}^K\big)$ to be the set of tuples $(v_1,\dots,v_J)\in U_q^J$ satisfying $\prod_{j=1}^J W_{i,k}(v_j)\in U_q$ and $\prod_{j=1}^J W_{i,k}(v_j)\equiv w_i \pmod{Q_0}$ for each $i\in[K]$. Observe that any convenient $n$ satisfying $\gcd(f(n),q)=1$ and $f_i(n)\equiv a_i \pmod{Q_0}$ for all $i\in[K]$, can be uniquely written in the form $n=m(P_J\cdots P_1)^k$, where $P_J,\dots,P_1$ are primes satisfying (4.1), and where $\gcd(f(m),q)=1$ and $(P_1,\dots,P_J) \bmod q \in \mathcal{U}_m \coloneqq \mathcal{U}_{J,K}\big(q,Q_0;(a_if_i(m)^{-1})_{i=1}^K\big)$. As such, by the arguments leading to (4.8), we obtain

$$\sum_{\substack{n\le x \text{ convenient}\\ \gcd(f(n),q)=1\\ (\forall i)\ f_i(n)\equiv a_i \ (\mathrm{mod}\ Q_0)}} 1 \;=\; \sum_{\substack{m\le x\\ (f(m),q)=1}} \frac{\#\mathcal{U}_m}{\varphi(q)^J}\left(\frac{1}{J!}\sum_{\substack{P_1,\dots,P_J>L_m\\ P_1\cdots P_J\le (x/m)^{1/k}\\ P_1,\dots,P_J \text{ distinct}}} 1\right) + o\left(\frac{1}{\varphi(q)^K}\sum_{\substack{n\le x\\ (f(n),q)=1}} 1\right). \tag{4.13}$$

A simple counting argument shows the following general observation: Let $F\in\mathbb{Z}[T]$ be nonconstant, and let $d,d'$ be positive integers such that $d'\mid d$ and $\alpha_F(d)\coloneqq \frac{1}{\varphi(d)}\#\{u\in U_d: F(u)\in U_d\}$ is nonzero (hence so is $\alpha_F(d')$). Then for any $u\in U_{d'}$ for which $F(u)\in U_{d'}$, we have

$$\#\{U\in U_d:\ U\equiv u \pmod{d'},\ F(U)\in U_d\} = \frac{\alpha_F(d)\varphi(d)}{\alpha_F(d')\varphi(d')}. \tag{4.14}$$

Using this general observation for $F\coloneqq W_k=\prod_{i=1}^K W_{i,k}$ (so that $\alpha_F=\alpha_k$), we obtain

$$\#\mathcal{U}_{J,K}\big(q,Q_0;(w_i)_{i=1}^K\big) = \left(\frac{\alpha_k(q)\varphi(q)}{\alpha_k(Q_0)\varphi(Q_0)}\right)^J \#\mathcal{V}_{J,K}^{(k)}\big(Q_0,(w_i)_{i=1}^K\big)$$

for all $(w_i)_{i=1}^K\in U_q^K$. Applying this with $w_i\coloneqq a_if_i(m)^{-1}$ and recalling that $\mathcal{V}'_{Q_0,m}=\mathcal{V}_{J,K}^{(k)}\big(Q_0;(a_if_i(m)^{-1})_{i=1}^K\big)$, we get from (4.13),

$$\sum_{\substack{n\le x \text{ convenient}\\ \gcd(f(n),q)=1\\ (\forall i)\ f_i(n)\equiv a_i \ (\mathrm{mod}\ Q_0)}} 1 \;=\frac{\alpha_k(q)^J}{\alpha_k(Q_0)^J}\sum_{\substack{m\le x\\ (f(m),q)=1}} \frac{\mathcal{V}'_{Q_0,m}}{\varphi(Q_0)^J}\left(\frac{1}{J!}\sum_{\substack{P_1,\dots,P_J>L_m\\ P_1\cdots P_J\le (x/m)^{1/k}\\ P_1,\dots,P_J \text{ distinct}}} 1\right) + o\left(\frac{1}{\varphi(q)^K}\sum_{\substack{n\le x\\ (f(n),q)=1}} 1\right).$$

Comparing this with (4.11), we obtain

$$\sum_{\substack{n\le x \text{ convenient}\\ (\forall i)\ f_i(n)\equiv a_i \ (\mathrm{mod}\ q)}} 1 \;=(1+o(1))\left(\frac{\varphi(Q_0)}{\varphi(q)}\right)^K \sum_{\substack{n\le x \text{ convenient}\\ \gcd(f(n),q)=1\\ (\forall i)\ f_i(n)\equiv a_i \ (\mathrm{mod}\ Q_0)}} 1 + o\left(\frac{1}{\varphi(q)^K}\sum_{\substack{n\le x\\ (f(n),q)=1}} 1\right).$$

Finally, an application of Proposition 4.1 allows us to remove the condition of $n$ being convenient from the main term on the right hand side above. This completes the proof of Proposition 4.3, up to the proof of Proposition 4.4, which we take up in the next section. □

## 5. Counting solutions to congruences: Proof of Proposition 4.4

### 5.1. Preparation for the proof of Proposition 4.4.

We temporarily abandon all the previous set-up just for this subsection. We shall make use of two character sum bounds, which we state in the next two propositions.

**Proposition 5.1.** *Let $\ell$ be a prime, $\chi$ a Dirichlet character mod $\ell$. Let $F \in \mathbb{Z}[T]$ be a nonconstant polynomial which is **not** congruent mod $\ell$ to a polynomial of the form $c \cdot G(T)^{\mathrm{ord}(\chi)}$ for some $c \in \mathbb{F}_\ell$ and $G \in \mathbb{F}_\ell[T]$, where $\mathrm{ord}(\chi)$ denotes the order of the character $\chi$. Then*

$$\left|\sum_{u \bmod \ell} \chi(F(u))\right| \le (d-1)\sqrt{\ell},$$

*where $d$ is the degree of the largest squarefree divisor of $F$.*

This is a version of the Weil bounds and is a special case of [51, Corollary 2.3] (see also [9], [52] and [42] for older results). We will also need an analogue of the above result for character sums to higher prime power moduli, and this input is provided by the following consequences of Theorems 1.2 and 7.1 and eqn. (1.15) in work of Cochrane [6] (see [7] for related results).

In what follows, for a polynomial $H \in \mathbb{Z}[T]$, we denote by $H'$ or $H'(T)$ the formal derivative of $H$. Let $\ell$ be a prime such that $\mathrm{ord}_\ell(H) = 0$, so that $H$ is not identically zero in $\mathbb{F}_\ell[T]$ (see § 2.7 for definition of $\mathrm{ord}_\ell$). By the $\ell$-critical polynomial associated to $H$ we shall mean the polynomial $\mathcal{C}_H := \ell^{-\mathrm{ord}_\ell(H')} H'$, which has integer coefficients and can be considered as a nonzero element of the ring $\mathbb{F}_\ell[T]$. By the $\ell$-critical points of $H$, we shall mean the set $\mathcal{A}(H;\ell) \subset \mathbb{F}_\ell$ of zeros of the polynomial $\mathcal{C}_H$ which are not zeros of $H$ (both polynomials considered mod $\ell$). Finally, for any $\theta \in \mathbb{F}_\ell$, we use $\mu_\theta(H)$ to denote the multiplicity of $\theta$ as a zero of $H$.

**Proposition 5.2.** *Let $\ell$ be a prime, $g \in \mathbb{Z}[T]$ a nonconstant polynomial, and $t := \mathrm{ord}_\ell(g')$. Consider an integer $e \ge t+2$ and a primitive character $\chi$ mod $\ell^e$. Let $M := \max_{\theta \in \mathcal{A}(g;\ell)} \mu_\theta(\mathcal{C}_g)$ be the maximum multiplicity of an $\ell$-critical point of $g$.*

(i) *For odd $\ell$, we have $|\sum_{u \bmod \ell^e} \chi(g(u))| \le \left(\sum_{\theta\in\mathcal{A}(g;\ell)} \mu_\theta(\mathcal{C}_g)\right) \ell^{t/(M+1)}\ \ell^{e(1-1/(M+1))}$.*

(ii) *For $\ell = 2$ and $e \ge t+3$, we have $|\sum_{u \bmod 2^e} \chi(g(u))| \le (12.5)2^{t/(M+1)}\ 2^{e(1-1/(M+1))}$. In fact, the sum is zero if $g$ has no 2-critical points.*

In order to make use of the aforementioned bounds, we will need to understand the quantities that appear when we apply them. The following observations enable us to do this.

**Proposition 5.3.** *Let $\{F_i\}_{i=1}^K \subset \mathbb{Z}[T]$ be nonconstant and multiplicatively independent. There exists a constant $C_1 := C_1(F_1, \dots, F_K) \in \mathbb{N}$ such that all of the following hold:*

- **(a)** *For any prime $\ell$, there are $O(1)$ many tuples $(A_1, \dots, A_K) \in [\ell - 1]^K$ for which $F_1^{A_1} \cdots F_K^{A_K}$ is of the form $c \cdot G^{\ell-1}$ in $\mathbb{F}_\ell[T]$ for some $c \in \mathbb{F}_\ell$ and $G \in \mathbb{F}_\ell[T]$; here, the implied constant depends at most on $\{F_i\}_{i=1}^K$. In fact, if $\ell > C_1$ and $\gcd(\ell - 1, \beta(F_1, \dots, F_K)) = 1$, then the only such tuple is $(A_1, \dots, A_K) = (\ell - 1, \dots, \ell - 1)$.*
- **(b)** *For any prime $\ell$ and any $(A_1, \dots, A_K) \in \mathbb{N}^K$ satisfying $\ell \nmid \gcd(A_1, \dots, A_K)$, we have in the two cases below,*

$$\tau(\ell) := \operatorname{ord}_\ell\left((T^{\varphi(\ell^r)} F_1(T)^{A_1} \cdots F_K(T)^{A_K})'\right) = \operatorname{ord}_\ell(\widetilde{F}(T)) \begin{cases} = 0, & \text{if } \ell > C_1, r \geq 2 \\ \leq C_1, & \text{if } \ell \leq C_1, \operatorname{ord}_\ell\left(\prod_{i=1}^K F_i\right) = 0, r \geq C_1 + 2, \end{cases} \tag{5.1}$$

*where $\widetilde{F}(T) := \sum_{i=1}^K A_i F_i'(T) \prod_{\substack{1 \leq j \leq K \\ j \neq i}} F_j(T)$. In either of the two cases above, any root $\theta \in \mathbb{F}_\ell$ of the polynomial $\mathcal{C}_\ell(T) := \ell^{-\tau(\ell)}(T^{\varphi(\ell^r)} F_1(T)^{A_1} \cdots F_K(T)^{A_K})'$ which is not a root of $T \prod_{i=1}^K F_i(T)$, must be a root of the polynomial $\ell^{-\tau(\ell)} \widetilde{F}(T)$ of the same multiplicity.*[6]

*Proof.* We start by writing $F_i =: r_i \prod_{j=1}^M G_j^{\mu_{ij}}$ as in subsection § 2.2, so that $r_i \in \mathbb{Z}$ and $G_1, \dots, G_M \in \mathbb{Z}[T]$ are irreducible, primitive and pairwise coprime, and $M = \omega(F_1 \cdots F_K)$. Recall that $M \geq K$ and that the exponent matrix $E_0(F_1, \dots, F_K)$ has $\mathbb{Q}$-linearly independent columns, making $\beta(F_1, \dots, F_K)$ a nonzero integer. Further, since $G_j$ are pairwise coprime irreducibles, the resultants $\operatorname{Res}(G_j, G_{j'})$ and discriminants $\operatorname{disc}(G_j)$ are nonzero integers for all $j \neq j' \in [M]$. Note that for any prime $\ell$ not dividing the leading coefficient of any $G_j$ and not dividing $\prod_{1 \leq j \leq M} \operatorname{disc}(G_j) \cdot \prod_{1 \leq j \neq j' \leq M} \operatorname{Res}(G_j, G_{j'})$, the product $\prod_{j=1}^M G_j$ is separable in $\mathbb{F}_\ell[T]$.

Also since $(F_1^{c_1} \cdots F_K^{c_K})' = \left(\prod_{i=1}^K F_i^{c_i - 1}\right) \sum_{i=1}^K c_i F_i' \prod_{\substack{1 \leq j \leq K \\ j \neq i}} F_j$, the multiplicative independence of the polynomials $\{F_i\}_{i=1}^K$ forces the polynomials $\left\{F_i' \prod_{\substack{1 \leq j \leq K \\ j \neq i}} F_j\right\}_{i=1}^K \subset \mathbb{Z}[T]$ to be $\mathbb{Q}$-linearly independent. Writing $D := \deg(F_1 \cdots F_K)$ and writing each $F_i'(T) \prod_{\substack{1 \leq j \leq K \\ j \neq i}} F_j(T) = \sum_{j=0}^{D-1} c_{i,j} T^j$ for some $\{c_{i,j}\}_{0 \leq j \leq D-1} \subset \mathbb{Z}$, we thus deduce that the columns of the matrix

$$M_1 := M_1(F_1, \dots, F_K) := \begin{pmatrix} c_{1,0} & \cdots & c_{K,0} \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ c_{1,D-1} & \cdots & c_{K,D-1} \end{pmatrix} \in \mathbb{M}_{D \times K}(\mathbb{Z}) \tag{5.2}$$

must be $\mathbb{Q}$-linearly independent. Consequently, $D \geq K$ and the last diagonal entry $\widetilde{\beta} \in \mathbb{Z} \setminus \{0\}$ of the Smith Normal form of $M_1$ is also the largest invariant factor of $M_1$ (in size).

Fix $C_1 := C_1(F_1, \dots, F_K)$ to be any positive integer exceeding all of the following:

- $\max\left\{2,\ |\widetilde{\beta}|,\ \prod_{j=1}^M |\operatorname{disc}(G_j)| \cdot \prod_{1 \leq j \neq j' \leq M} |\operatorname{Res}(G_j, G_{j'})|\right\}$ (recall that these are all nonzero),
- the sizes of the leading coefficients of $F_1, \dots, F_K, G_1, \dots, G_M$.

[6] Once again, the last three polynomials are being considered as nonzero elements of $\mathbb{F}_\ell[T]$.

We claim that any such $C_1$ satisfies the properties in the statement of the proposition.

*Proof of (a).* We may assume that $\ell > C_1$. Let $\beta := \beta(F_1, \dots, F_K)$, as defined in § 2.2. By the discussion at the start of the proof, the conditions defining $C_1$ force $G_1, \dots, G_M$ to be pairwise coprime in $\mathbb{F}_\ell[T]$. Let $(A_1, \dots, A_K) \neq (0, \dots, 0)$ be any tuple of nonnegative integers for which $F_1^{A_1} \cdots F_K^{A_K}$ is of the form $c \cdot G^{\ell-1}$ in $\mathbb{F}_\ell[T]$ for some $c \in \mathbb{F}_\ell$ and $G \in \mathbb{F}_\ell[T]$. We claim that $A_1, \dots, A_K$ must all be divisible by $(\ell-1)/d_1$ where $d_1 := \gcd(\ell-1, \beta)$. This will be enough to complete the proof of (a), since there are no more than $d_1^K \le |\beta|^K \ll 1$ many tuples $(A_1, \dots, A_K) \in [\ell-1]^K$ with each $A_i$ divisible by $(\ell-1)/d_1$.

To establish the above claim, we may assume without loss of generality that $G$ is monic, and note that $c \in \mathbb{F}_\ell^\times$ since $\mathrm{ord}_\ell(F_1 \cdots F_K) = 0$ by definition of $C_1$. Write each $G_j$ as $\lambda_j H_j$ in the ring $\mathbb{F}_\ell[T]$, for some $\lambda_j \in \mathbb{F}_\ell^\times$ and nonconstant monic $H_j \in \mathbb{F}_\ell[T]$ (which can be done since $\ell$ doesn't divide the leading coefficient of any $G_j$). Then $F_i = r_i \prod_{j=1}^M G_j^{\mu_{ij}} = \rho_i \prod_{j=1}^M H_j^{\mu_{ij}}$ for some $\rho_i \in \mathbb{F}_\ell^\times$. Since $c \cdot G^{\ell-1} = \prod_{i=1}^K F_i^{A_i} = \left(\prod_{i=1}^K \rho_i^{A_i}\right) \cdot \prod_{1 \le j \le M} H_j^{\sum_{i=1}^K \mu_{ij} A_i}$ in $\mathbb{F}_\ell[T]$, and $G, H_1, \dots, H_M$ are all monic, we find that $G^{\ell-1} = \prod_{1 \le j \le M} H_j^{\sum_{i=1}^K \mu_{ij} A_i}$. But now since $\prod_{1 \le j \le M} G_j$ is separable in $\mathbb{F}_\ell[T]$, so is $\prod_{1 \le j \le M} H_j$, and we deduce that $\sum_{i=1}^K \mu_{ij} A_i \equiv 0 \pmod{\ell-1}$ for each $j \in [M]$. This can be rewritten as the matrix congruence $(0 \cdots\cdots 0)^\top \equiv E_0 (A_1 \cdots A_K)^\top \pmod{\ell-1}$; each side of this congruence is an $M \times 1$ matrix, $Y^\top$ denotes the transpose of a matrix $Y$ and $E_0$ is the exponent matrix defined in § 2.2.

Now since $M \ge K$ and $E_0$ has full rank, there exist $P_0 \in GL_{M \times M}(\mathbb{Z})$ and $R_0 \in GL_{K \times K}(\mathbb{Z})$ for which $P_0 E_0 R_0$ is the Smith Normal Form $\mathrm{diag}(\beta_1, \dots, \beta_K)$ of $E_0$, with $\beta_1, \dots, \beta_K \in \mathbb{Z} \setminus \{0\}$ being the invariant factors of $E_0$. Thus $\beta_i \mid \beta_{i+1}$ for all $1 \le i < K$ and $\beta = \beta(F_1, \dots, F_K) = \beta_K$. This means that $P_0 E_0 = \mathrm{diag}(\beta_1, \dots, \beta_K) R_0^{-1}$ and writing $(q_{ij})_{1 \le i,j \le K} := R_0^{-1}$, we find that

$$\begin{pmatrix} 0 \\ \cdots \\ \cdots \\ 0 \end{pmatrix}_{M \times 1} \equiv P_0 E_0 \begin{pmatrix} A_1 \\ \cdots \\ A_K \end{pmatrix}_{K \times 1} \equiv \begin{pmatrix} \beta_1 (q_{11} A_1 + \cdots + q_{1K} A_K) \\ \cdots \\ \beta_K (q_{K1} A_1 + \cdots + q_{KK} A_K) \\ 0 \\ \cdots \\ 0 \end{pmatrix}_{M \times 1} \pmod{\ell - 1}.$$

Hence for each $i \in [K]$, $\beta_i (q_{i1} A_1 + \cdots + q_{iK} A_K) \equiv 0 \pmod{\ell-1}$, so that $(\ell-1)/\gcd(\ell-1, \beta_i)$ divides $q_{i1} A_1 + \cdots + q_{iK} A_K$. But since $\beta_i \mid \beta_K$, it follows that $(\ell-1)/\gcd(\ell-1, \beta_K) = (\ell-1)/d_1$ also divides $q_{i1} A_1 + \cdots + q_{iK} A_K$ for each $i \in [K]$. We obtain

$$\begin{pmatrix} 0 \\ \cdots \\ 0 \end{pmatrix}_{K \times 1} \equiv \begin{pmatrix} q_{11} A_1 + \cdots + q_{1K} A_K \\ \cdots \\ q_{K1} A_1 + \cdots + q_{KK} A_K \end{pmatrix}_{K \times 1} \equiv R_0^{-1} \begin{pmatrix} A_1 \\ \cdots \\ A_K \end{pmatrix}_{K \times 1} \left(\mathrm{mod}\ \frac{\ell-1}{d_1}\right), \tag{5.3}$$

establishing the desired claim that $(A_1, \dots, A_K) \equiv (0, \dots, 0) \left(\mathrm{mod}\ \frac{\ell-1}{d_1}\right)$.

*Proof of (b).* We start by noting that

$$(T^{\varphi(\ell^r)} F_1(T)^{A_1} \cdots F_K(T)^{A_K})' = \varphi(\ell^r) T^{\varphi(\ell^r)-1} \prod_{i=1}^K F_i(T)^{A_i} + T^{\varphi(\ell^r)} \left( \prod_{i=1}^K F_i(T)^{A_i - 1} \right) \widetilde{F}(T), \tag{5.4}$$

where $\widetilde{F}(T)$ is as in the statement of the proposition. We claim that $\mathrm{ord}_\ell(\widetilde{F}) \le \mathbb{1}_{\ell \le C_1} C_1$ for all primes $\ell$ satisfying $\mathrm{ord}_\ell(F_1 \cdots F_K) = 0$ and for all nonnegative integers $A_1, \dots, A_K$ satisfying $(A_1, \dots, A_K) \not\equiv (0, \dots, 0) \bmod \ell$. To show this, we proceed as in the proof of (a), but working with the matrix $M_1$ defined in (5.2) in place of the exponent matrix $E_0$. Observe that $\widetilde{F}(T) = \sum_{j=0}^{D-1} \left(\sum_{i=1}^{K} c_{i,j} A_i\right) T^j$, hence if $\kappa(\ell) \coloneqq \mathrm{ord}_\ell(\widetilde{F})$, then $\ell^{\kappa(\ell)}$ divides all the entries of the matrix $M_1 (A_1 \cdots A_K)^\top$. Since $M_1$ has full rank and $D = \sum_{i=1}^K \deg F_i \ge K$ many rows, and since $(A_1, \dots, A_K) \not\equiv (0, \dots, 0) \bmod \ell$, an argument entirely analogous to the one leading to (5.3) shows that $\ell^{\kappa(\ell)}$ divides the last invariant factor $\widetilde{\beta}$ of $M_1$. Hence $\mathrm{ord}_\ell(\widetilde{F}) = \kappa(\ell) \le v_\ell(\widetilde{\beta})$ and our claim follows as $|\widetilde{\beta}| < C_1$.

As a consequence, we find that $\mathrm{ord}_\ell \left(T^{\varphi(\ell^r)} \left(\prod_{i=1}^K F_i(T)^{A_i - 1}\right) \widetilde{F}(T)\right) = \mathrm{ord}_\ell(\widetilde{F}) \le \mathbb{1}_{\ell \le C_1} C_1$ for all primes $\ell \le C_1$ satisfying $\mathrm{ord}_\ell(F_1 \cdots F_K) = 0$, and also for all primes $\ell > C_1$ (for which the condition $\mathrm{ord}_\ell(F_1 \cdots F_K) = 0$ is automatic by definition of $C_1$). But now since $\mathrm{ord}_\ell(\varphi(\ell^r)) \ge 1$ for $r \ge 2$ and $\mathrm{ord}_\ell(\varphi(\ell^r)) \ge C_1 + 1$ for $r \ge C_1 + 2$, (5.4) shows that $\tau(\ell) = \mathrm{ord}_\ell \left(T^{\varphi(\ell^r)} \left(\prod_{i=1}^K F_i(T)^{A_i - 1}\right) \widetilde{F}(T)\right)$, establishing subpart (b) of the proposition.

Finally, since in both the cases of (5.1), we have $\tau(\ell) < r - 1$, the identity (5.4) reveals that

$$\mathcal{C}_\ell(T) \equiv \ell^{-\tau(\ell)} \left(T^{\varphi(\ell^r)} \prod_{i=1}^K F_i(T)^{A_i}\right)' \equiv T^{\varphi(\ell^r)} \left(\prod_{i=1}^K F_i(T)^{A_i - 1}\right) \left(\ell^{-\tau(\ell)} \widetilde{F}(T)\right) \text{ in the ring } \mathbb{F}_\ell[T].$$

As such, any root of the polynomial $\theta \in \mathbb{F}_\ell$ of $\mathcal{C}_\ell(T)$ (considered as a nonzero element of $\mathbb{F}_\ell[T]$) which is not a root of $T \prod_{i=1}^K F_i(T)$, must be a root of $\ell^{-\tau(\ell)} \widetilde{F}(T)$, and $\theta$ must have the same multiplicity in $\mathcal{C}_\ell(T)$ and $\ell^{-\tau(\ell)} \widetilde{F}(T)$. This completes the proof of Proposition 5.3. □

### 5.2. **Proof of Proposition 4.4.**

We return to the set-up in Proposition 4.4. Since $\alpha_k(q) \neq 0$, we have $\mathrm{ord}_\ell(\prod_{i=1}^K W_{i,k}) = 0$ for each prime $\ell \mid q$. Fix $C_0 \coloneqq C_0\left(\{W_{i,k}\}_{1 \le i \le K}; B_0\right)$ to be any constant exceeding $B_0$, $(32D)^{2D+2}$, the sizes of the leading and constant coefficients of $\{W_{i,k}\}_{1 \le i \le K}$, as well as the constant $C_1(W_{1,k}, \dots, W_{K,k})$ coming from an application of Proposition 5.3 to the family $\{W_{i,k}\}_{1 \le i \le K}$ of multiplicatively independent polynomials. We will show that any such choice of $C_0$ suffices.

**We first consider the case $D > 1$**; the case $D = 1$ will be dealt with towards the end of the argument. For an arbitrary positive integer $Q$ and coprime residues $w_1, \dots, w_K$ mod $Q$, an application of the orthogonality of Dirichlet characters yields

$$\#\mathcal{V}_{N,K}^{(k)}\left(Q; (w_i)_{i=1}^K\right) = \frac{1}{\varphi(Q)^K} \sum_{\chi_1, \dots, \chi_K \bmod Q} \overline{\chi}_1(w_1) \cdots \overline{\chi}_K(w_K) (Z_{Q; \chi_1, \dots, \chi_K})^N, \tag{5.5}$$

with $Z_{Q; \chi_1, \dots, \chi_K} \coloneqq \sum_{v \bmod Q} \chi_{0,Q}(v) \prod_{i=1}^K \chi_i(W_{i,k}(v))$ and $\chi_{0,Q}$ the trivial character mod $Q$.

**Dealing with the large primes dividing $q$:** We first show that there exists a constant $K' = K'(\{W_{i,k}\}_{1 \le i \le K})$ such that uniformly in primes $\ell > C_0$ dividing $q$, we have

$$\frac{\#\mathcal{V}_{N,K}^{(k)}\left(\ell^e; (w_i)_{i=1}^K\right)}{\varphi(\ell^e)^N} \begin{cases} = \dfrac{\alpha_k(\ell)^N}{\varphi(\ell^e)^K}\left(1 + O\left(\dfrac{(4D)^N}{\ell^{N/D-K}}\right)\right), \text{ uniformly in } N \ge KD+1 \\ \le K'\ e^{\mathbb{1}_{N=KD}} \ell^{-eN/D}, \text{ for each } 1 \le N \le KD. \end{cases} \tag{5.6}$$

To show these, we start by applying (5.5) to get

$$\begin{aligned} &\frac{\#\mathcal{V}_{N,K}^{(k)}(\ell^e; (w_i)_{i=1}^K)}{\varphi(\ell^e)^N} \\ &= \frac{\alpha_k(\ell)^N}{\varphi(\ell^e)^K}\left\{1 + \frac{1}{(\alpha_k(\ell)\varphi(\ell^e))^N} \sum_{(\chi_1,\dots,\chi_K) \ne (\chi_{0,\ell},\dots,\chi_{0,\ell}) \bmod \ell^e} \left(\prod_{i=1}^K \overline{\chi}_i(w_i)\right)(Z_{\ell^e;\chi_1,\dots,\chi_K})^N\right\}, \end{aligned} \tag{5.7}$$

where we have recalled that $\alpha_k(\ell) \ne 0$ since $\alpha_k(q) \ne 0$. For any tuple $(\chi_1, \dots, \chi_K) \ne (\chi_{0,\ell}, \dots, \chi_{0,\ell}) \bmod \ell^e$, let $\ell^{e_0} := \mathrm{lcm}[\mathfrak{f}(\chi_1), \dots, \mathfrak{f}(\chi_K)] \in \{\ell, \dots, \ell^e\}$. Using $\chi_1, \dots, \chi_K$ to also denote the characters mod $\ell^{e_0}$ inducing $\chi_1, \dots, \chi_K$ respectively, we see that $Z_{\ell^e;\chi_1,\dots,\chi_K} = \ell^{e-e_0}\ Z_{\ell^{e_0};\ \chi_1,\dots,\chi_K}$. Moreover $U_{\ell^{e_0}}$ is cyclic since $\ell > C_0 > 2$. Letting $\gamma$ denote a generator of $U_{\ell^{e_0}}$, we see that the character group mod $\ell^{e_0}$ is generated by the character $\psi_{e_0}$ given by $\psi_{e_0}(\gamma) := \exp(2\pi i/\varphi(\ell^{e_0}))$. Hence, there exists a tuple $(A_1, \dots, A_K) \in [\varphi(\ell^{e_0})]$ satisfying $\chi_i = \psi_{e_0}^{A_i}$ for each $i$, and since at least one of $\chi_1, \dots, \chi_K$ is primitive mod $\ell^{e_0}$, we also have

$$(A_1, \dots, A_K) \not\equiv \begin{cases} (0, \dots, 0) \pmod{\ell}, & \text{if } e_0 > 1, \\ (0, \dots, 0) \pmod{\ell - 1}, & \text{if } e_0 = 1. \end{cases} \tag{5.8}$$

We can now write

$$Z_{\ell^e;\ \chi_1,\dots,\chi_K} = \ell^{e-e_0}\ Z_{\ell^{e_0};\ \chi_1,\dots,\chi_K} = \ell^{e-e_0} \sum_{v \bmod \ell^{e_0}} \psi_{e_0}\left(v^{\varphi(\ell^{e_0})} \prod_{i=1}^K W_{i,k}(v)^{A_i}\right). \tag{5.9}$$

<u>Case 1</u>: If $e_0 = 1$, then since $\ell > C_0 > B_0$, we have $\gcd(\ell - 1, \beta(W_{1,k}, \dots, W_{K,k})) = 1$. Further, since $\ell > C_0 > C_1(W_{1,k}, \dots, W_{K,k})$, we see by (5.8) and Proposition 5.3(a) that $\prod_{i=1}^K W_{i,k}^{A_i}$ cannot be of the form $c \cdot G^{\ell-1}$ in $\mathbb{F}_\ell[T]$. As such, (5.9) and Proposition 5.1 show that $|Z_{\ell^e;\ \chi_1,\dots,\chi_K}| \le D\ell^{e-1/2}$ for any tuple $(\chi_1, \dots, \chi_K)$ mod $\ell^e$ having $e_0 = 1$.

<u>Case 2</u>: If $e_0 \ge 2$, then since $\mathrm{ord}_\ell(\prod_{i=1}^K W_{i,k}) = 0$ and $\ell > C_0 > C_1(W_{1,k}, \dots, W_{K,k})$, Proposition 5.3 and (5.8) show that $\tau(\ell) := \mathrm{ord}_\ell((T^{\varphi(\ell^{e_0})} \prod_{i=1}^K W_{i,k}(T)^{A_i})') = 0 \le e_0 - 2$. Thus (5.9) and Proposition 5.2(i) yield $|Z_{\ell^e;\ \chi_1,\dots,\chi_K}| \le \left(\sum_{\theta \in \mathcal{A}_\ell} \mu_\theta(\mathcal{C}_\ell)\right) \ell^{e-e_0/(M_\ell+1)}$, where $\mathcal{A}_\ell \subset \mathbb{F}_\ell$ denotes the set of $\ell$-critical points of the polynomial $T^{\varphi(\ell^{e_0})} \prod_{i=1}^K W_{i,k}(T)^{A_i}$, namely the roots of $\mathcal{C}_\ell(T) = (T^{\varphi(\ell^{e_0})} \prod_{i=1}^K W_{i,k}(T)^{A_i})'$ in $\mathbb{F}_\ell$ that are not roots of $T^{\varphi(\ell^{e_0})} \prod_{i=1}^K W_{i,k}(T)^{A_i}$. But by the last assertion in Proposition 5.3, we see that $M_\ell \le \sum_{\theta \in \mathcal{A}_\ell} \mu_\theta(\mathcal{C}_\ell) \le \deg(\sum_{i=1}^K A_i W_{i,k}' \prod_{\substack{1 \le j \le K \\ j \ne i}} W_{j,k})$ $\le D - 1$. This yields $|Z_{\ell^e;\ \chi_1,\dots,\chi_K}| \le D\ell^{e-e_0/D}$ for any tuple $(\chi_1, \dots, \chi_K)$ mod $\ell^e$ having $e_0 > 1$.

Combining the conclusions of Cases 1 and 2, and using the fact that there are at most $\ell^{e_0 K}$ many tuples $(\chi_1, \dots, \chi_K)$ of characters mod $\ell^e$ having $\text{lcm}[\mathfrak{f}(\chi_1), \dots, \mathfrak{f}(\chi_K)] = \ell^{e_0}$, we get

$$\frac{1}{(\alpha_k(\ell)\varphi(\ell^e))^N} \sum_{(\chi_1,\dots,\chi_K)\neq(\chi_{0,\ell},\dots,\chi_{0,\ell}) \bmod \ell^e} |Z_{\ell^e;\,\chi_1,\dots,\chi_K}|^N \le (4D)^N \sum_{1\le e_0\le e} \ell^{e_0(K-N/D)}, \tag{5.10}$$

where in the last inequality above, we have used the facts that $D \ge 2$ and $\alpha_k(\ell) \ge 1 - D/(\ell - 1) \ge 1 - D/(C_0 - 1) \ge 1/2$. Now if $N \ge KD + 1$, then $\ell^{K-N/D} \le C_0^{-1/D} \le 1/2$, so that the last sum in (5.10) is at most $2(4D)^N \ell^{K-N/D}$. On the other hand, if $N \le KD$, then the same sum is $\ll e^{\mathbb{1}_{N=KD}} \ell^{e(K-N/D)}$. Inserting these two bounds into (5.10) and (5.7) gives (5.6).

**Dealing with the small primes dividing** $q$**:** Now for an arbitrary $q$, we let $\widetilde{q} \coloneqq \prod_{\substack{\ell^e \| q \\ \ell \le C_0}} \ell^e$ denote the $C_0$-smooth part of $q$. By (5.5),

$$\#\mathcal{V}_{N,K}^{(k)}\left(\widetilde{q}; (w_i)_{i=1}^K\right) = \frac{1}{\varphi(\widetilde{q})^K} \sum_{\chi_1,\dots,\chi_K \bmod \widetilde{q}} \overline{\chi}_1(w_1) \cdots \overline{\chi}_K(w_K) (Z_{\widetilde{q};\,\chi_1,\dots,\chi_K})^N. \tag{5.11}$$

Given a constant $C > C_0$, we fix $\kappa$ to be any **integer** constant exceeding $C \cdot (30DC_0^{C_0})^{2C_0}$. Let $Q_0 \coloneqq \prod_{\ell^e \| \widetilde{q}} \ell^{\min\{e,\kappa\}} = \prod_{\ell \le C_0} \ell^{\min\{v_\ell(q),\kappa\}}$ denote the largest $(\kappa+1)$-free divisor of $\widetilde{q}$. Write the expression on the right hand side of (5.11) as $\mathcal{S}' + \mathcal{S}''$, where

$$\mathcal{S}' \coloneqq \frac{1}{\varphi(\widetilde{q})^K} \sum_{\substack{\chi_1,\dots,\chi_K \bmod \widetilde{q} \\ \text{lcm}[\mathfrak{f}(\chi_1),\dots,\mathfrak{f}(\chi_K)] \text{ is } (\kappa+1)\text{-free}}} \overline{\chi}_1(w_1) \cdots \overline{\chi}_K(w_K) (Z_{\widetilde{q};\,\chi_1,\dots,\chi_K})^N$$

denotes the contribution of those tuples $(\chi_1, \dots, \chi_K)$ mod $\widetilde{q}$ for which $\text{lcm}[\mathfrak{f}(\chi_1), \dots, \mathfrak{f}(\chi_K)]$ is $(\kappa+1)$-free, or equivalently, those $(\chi_1, \dots, \chi_K)$ for which $\text{lcm}[\mathfrak{f}(\chi_1), \dots, \mathfrak{f}(\chi_K)]$ divides $Q_0$.

For each tuple $(\chi_1, \dots, \chi_K)$ counted in $\mathcal{S}'$, there exists a unique tuple $(\psi_1, \dots, \psi_K)$ of characters mod $Q_0$ inducing $(\chi_1, \dots, \chi_K)$ mod $\widetilde{q}$, respectively. Noting that $\alpha_k(\widetilde{q}) = \alpha_k(Q_0)$, a straightforward calculation using (4.14) shows that

$$Z_{\widetilde{q};\,\chi_1,\dots,\chi_K} = \sum_{u \bmod Q_0} \chi_{0,Q_0}(u) \prod_{i=1}^K \psi_i(W_{i,k}(u)) \sum_{\substack{v \bmod \widetilde{q} \\ v \equiv u \bmod Q_0 \\ \gcd(v \prod_{i=1}^K W_{i,k}(v), \widetilde{q})=1}} 1 = \frac{\varphi(\widetilde{q})}{\varphi(Q_0)} Z_{Q_0;\,\psi_1,\dots,\psi_K}$$

Using this and invoking (5.5) with $Q \coloneqq Q_0$, we obtain

$$\frac{\mathcal{S}'}{\varphi(\widetilde{q})^N} = \frac{\varphi(\widetilde{q})^{-K}}{\varphi(Q_0)^N} \sum_{\psi_1,\dots,\psi_K \bmod Q_0} \left(\prod_{i=1}^K \overline{\psi}_i(w_i)\right) (Z_{Q_0;\,\psi_1,\dots,\psi_K})^N = \left(\frac{\varphi(Q_0)}{\varphi(\widetilde{q})}\right)^K \frac{\#\mathcal{V}_{N,K}^{(k)}\left(Q_0; (w_i)_{i=1}^K\right)}{\varphi(Q_0)^N} \tag{5.12}$$

We now deal with the remaining sum $\mathcal{S}''$ which is the contribution of those $(\chi_1, \dots, \chi_K)$ mod $\widetilde{q}$ for which $\text{lcm}[\mathfrak{f}(\chi_1), \dots, \mathfrak{f}(\chi_K)]$ is not $(\kappa+1)$-free. For each such $(\chi_1, \dots, \chi_K)$, we factor $\chi_i \eqqcolon \prod_{\ell^e \| \widetilde{q}} \chi_{i,\ell}$, where $\chi_{i,\ell}$ is a character mod $\ell^e$. With $e_\ell \coloneqq v_\ell\left(\text{lcm}[\mathfrak{f}(\chi_1), \dots, \mathfrak{f}(\chi_K)]\right)$, we observe that since $\mathfrak{f}(\chi_i) = \prod_{\ell^e \| q} \mathfrak{f}(\chi_{i,\ell})$ and each $\mathfrak{f}(\chi_{i,\ell})$ is a power of $\ell$, we must have $\text{lcm}[\mathfrak{f}(\chi_{1,\ell}), \dots, \mathfrak{f}(\chi_{K,\ell})] = \ell^{e_\ell}$. For each $\ell^e \| \widetilde{q}$, let $(\chi_{1,\ell}, \dots, \chi_{K,\ell})$ also denote the characters mod $\ell^{e_\ell}$ inducing $(\chi_{1,\ell}, \dots, \chi_{K,\ell})$ mod $\ell^e$ respectively. Then at least one of $\chi_{1,\ell}, \dots, \chi_{K,\ell}$ must be primitive mod $\ell^{e_\ell}$. The factorization $Z_{\widetilde{q};\chi_1,\dots,\chi_K} = \prod_{\ell^e \| \widetilde{q}} Z_{\ell^e;\chi_{1,\ell},\dots,\chi_{K,\ell}}$ now yields

$$|Z_{\widetilde{q};\chi_1,\dots,\chi_K}| \le \left( \prod_{\substack{\ell^e \| \widetilde{q} \\ e_\ell \le \kappa}} \varphi(\ell^e) \right) \prod_{\substack{\ell^e \| \widetilde{q} \\ e_\ell \ge \kappa+1}} \left( \ell^{e-e_\ell} |Z_{\ell^{e_\ell};\chi_{1,\ell},\dots,\chi_{K,\ell}}| \right). \tag{5.13}$$

We claim that for all prime powers $\ell^e \parallel \widetilde{q}$ with $e_\ell \ge \kappa+1$, we have

$$|Z_{\ell^{e_\ell};\chi_{1,\ell},\dots,\chi_{K,\ell}}| \le \left(DC_0^{C_0}\right) \ell^{e_\ell(1-1/D)}. \tag{5.14}$$

For odd $\ell$, this follows essentially by the same argument as that given to bound $Z_{\ell^e;\chi_1,\dots,\chi_K}$ in "Case 2" before: The only difference is that this time we use *both* the assertions in (5.1) since $e_\ell \ge \kappa + 1 > (30DC_0)^{2C_0} + 1 > C_0 + 2$. Now assume that $\ell = 2$, i.e. $e_2 = v_2(\mathrm{lcm}[\mathfrak{f}(\chi_1), \dots, \mathfrak{f}(\chi_K)]) \ge \kappa + 1 \ge 31$. We shall use Proposition 5.2(ii).

To do this, we observe that the characters $\psi, \eta$ mod $2^{e_2}$ defined by

$$\psi(5) := \exp(2\pi i/2^{e_2-2}),\ \psi(-1) := 1 \quad \text{and} \quad \eta(5) := 1, \eta(-1) := -1$$

generate the character group mod $2^{e_2}$. Hence for each $i \in [K]$, there exist $r_i \in [2^{e_2-2}]$ and $s_i \in [2]$ satisfying $\chi_{i,2} = \psi^{r_i}\eta^{s_i}$; also $2 \nmid \gcd(r_1, \dots, r_K)$ as $e_2 \ge 4$ and at least one of $\chi_{1,2}, \dots, \chi_{K,2}$ is primitive mod $2^{e_2}$. Thus $Z_{2^{e_2}} = \sum_{v \bmod 2^{e_2}} \psi\left(g(v)\right)\ \eta\left(v^2 \prod_{i=1}^K W_{i,k}(v)^{s_i}\right)$, where $g(T) := \prod_{i=1}^K W_{i,k}(T)^{r_i}$ and we have abbreviated $Z_{2^{e_2};\chi_{1,2},\dots,\chi_{K,2}}$ to $Z_{2^{e_2}}$. Since $\eta$ is induced by the nontrivial character mod 4, writing $v := 4u + \lambda$ and $h_\lambda(T) := g(4T + \lambda)$ gives

$$Z_{2^{e_2}} = \sum_{\lambda=\pm1} \eta\left(\prod_{i=1}^K W_{i,k}(\lambda)^{s_i}\right) \sum_{u \bmod 2^{e_2-2}} \psi\left(h_\lambda(u)\right) = \frac{1}{4} \sum_{\lambda=\pm1} \eta\left(\prod_{i=1}^K W_{i,k}(\lambda)^{s_i}\right) \sum_{u \bmod 2^{e_2}} \psi\left(h_\lambda(u)\right) \tag{5.15}$$

If $\eta\left(\prod_{i=1}^K W_{i,k}(\lambda)^{s_i}\right) \ne 0$, then $\prod_{i=1}^K W_{i,k}(\lambda)^{s_i} \equiv 1 \pmod 2$, so $\mathrm{ord}_2\left(\prod_{i=1}^K W_{i,k}(4T+\lambda)^{r_i-1}\right) = 0$. As such, with $\widetilde{G} := \sum_{i=1}^K r_i W_{i,k}' \prod_{j \ne i} W_{j,k}$, we see that

$$\tau_\lambda(2) := \mathrm{ord}_2(h_\lambda'(T)) = 2 + \mathrm{ord}_2(\widetilde{G}(4T+\lambda)) \le 2 + \mathrm{ord}_2(\widetilde{G}) + 2\deg(\widetilde{G}) \le C_0 + 2D; \tag{5.16}$$

here we have used (5.1) and the fact that $\mathrm{ord}_2(F(4T+\lambda)) \le \mathrm{ord}_2(F) + 2\deg(F)$ for any nonconstant polynomial $F$. [7]

Two consequences of (5.16) are that $2^{-(\tau_\lambda(2)-2)}\widetilde{G}(4T+\lambda) \in \mathbb{Z}[T]$ and that $\tau_\lambda(2) \le \kappa - 3 \le e_2 - 3$. Thus Proposition 5.2(ii) applies, yielding $\left|\sum_{u \bmod 2^{e_2}} \psi\left(h_\lambda(u)\right)\right| \le (12.5) \cdot 2^{C_0+2D} \cdot 2^{e_2(1-1/(M_\lambda+1))}$, where $M_\lambda$ is the maximum multiplicity of a 2-critical point of $h_\lambda$. Since $\prod_{i=1}^K W_{i,k}(4T+\lambda)^{r_i-1} \equiv 1 \pmod 2$, it follows that any such critical point $\theta \in \mathbb{F}_2$ is a root of the polynomial $2^{-(\tau_\lambda(2)-2)}\widetilde{G}(4T+\lambda)$, giving $M_\lambda \le \deg \widetilde{G}(4T+\lambda) \le D - 1$, so that $\left|\sum_{u \bmod 2^{e_2}} \psi\left(h_\lambda(u)\right)\right| \le (12.5) \cdot 2^{C_0+2D} \cdot 2^{e_2(1-1/D)} \le DC_0^{C_0} \cdot 2^{e_2(1-1/D)}$. Inserting this into (5.15) completes the proof of (5.14) in the remaining case $\ell = 2$.

[7] This can be seen by writing the coefficients of $F(4T+\lambda)$ in terms of those of $F$, and using a simple divisibility argument.

Combining (5.13) with (5.14), we find that for each $(\chi_1, \dots, \chi_K)$ counted in $\mathcal{S}''$, we have $|Z_{\widetilde{q};\, \chi_1,\dots,\chi_K}| \le (2D_0 C_0^{C_0})^{C_0} \varphi(\widetilde{q}) A^{-1/D_0}$, where $A := \prod_{\ell^e \| \widetilde{q}:\ e_\ell \ge \kappa+1} \ell^{e_\ell}$ denotes the $(\kappa+1)$-full part of $\operatorname{lcm}[\mathfrak{f}(\chi_1), \dots, \mathfrak{f}(\chi_K)]$, i.e, the largest $(\kappa+1)$-full divisor of $\operatorname{lcm}[\mathfrak{f}(\chi_1), \dots, \mathfrak{f}(\chi_K)]$. Now for a divisor $d$ of $\widetilde{q}$, there are at most $d^K$ many tuples $(\chi_1, \dots, \chi_K)$ of characters mod $\widetilde{q}$ for which $\operatorname{lcm}[\mathfrak{f}(\chi_1), \dots, \mathfrak{f}(\chi_K)] = d$. Hence, summing this last bound over all possible $(\chi_1, \dots, \chi_K)$ occurring in the sum $\mathcal{S}''$, we obtain

$$
\begin{aligned}
|\mathcal{S}''| &\le \frac{1}{\varphi(\widetilde{q})^K} \sum_{\substack{A \mid \widetilde{q}:\ A>1 \\ A \text{ is } (\kappa+1)\text{-full}}} \ \sum_{\substack{d \mid \widetilde{q} \\ (\kappa+1)\text{-full part} \\ \text{of } d \text{ is } A}} d^K \cdot \frac{(2D_0 C_0^{C_0})^{C_0 N} \varphi(\widetilde{q})^N}{A^{N/D}} \\
&\ll \frac{\varphi(\widetilde{q})^N}{\varphi(\widetilde{q})^K} \cdot (2D_0 C_0^{C_0})^{C_0 N} \sum_{\substack{A \mid \widetilde{q}:\ A>1 \\ A \text{ is } (\kappa+1)\text{-full}}} \frac{1}{A^{N/D-K}}.
\end{aligned}
$$

In the last step above, we have noted that for any $d$ dividing $\widetilde{q}$ whose $(\kappa+1)$-full part is $A$, we have $d \ll A$. Continuing,

$$
\frac{|\mathcal{S}''|}{\varphi(\widetilde{q})^N} \ll \frac{(2D_0 C_0^{C_0})^{C_0 N}}{\varphi(\widetilde{q})^K} \left\{ \prod_{\ell^e \| \widetilde{q}} \left( 1 + \sum_{\kappa+1 \le \nu \le e} \frac{1}{\ell^{\nu(N/D-K)}} \right) - 1 \right\}. \tag{5.17}
$$

Now if $N \ge KD+1$, then since $\kappa > C \cdot (30 D C_0^{C_0})^{2C_0} \ge D(D+3)$, we see that the sum on $\nu$ above is at most $2^{-\kappa(N/D-K)} \left(1 - 2^{-1/D}\right)^{-1} \le \frac{2^{D+2}}{2^{\kappa/D}} \le \frac{1}{2}$. Hence $\log(1 + \sum_{\kappa+1 \le \nu \le e} \ell^{-\nu(N/D-K)}) \ll 2^{-\kappa(N/D-K)} \ll 2^{-\kappa N/D}$. In addition, since $P(\widetilde{q}) \le C_0$, (5.17) gives

$$
\frac{|\mathcal{S}''|}{\varphi(\widetilde{q})^N} \ll \frac{(2D_0 C_0^{C_0})^{C_0 N}}{\varphi(\widetilde{q})^K} \left\{ \exp\left( O\left( \frac{1}{2^{\kappa N/D}} \right) \right) - 1 \right\} \ll \frac{1}{\varphi(\widetilde{q})^K} \cdot \left( \frac{(2D_0 C_0^{C_0})^{C_0}}{2^{\kappa/D}} \right)^N \ll \frac{C^{-N}}{\varphi(\widetilde{q})^K}, \tag{5.18}
$$

where in the last step, we have recalled that $\kappa/D > D^{-1} \cdot C \cdot (30 D C_0^{C_0})^{2C_0} > C \cdot (2C_1)^{C_0}$. Combining (5.18) with (5.12), we deduce that

$$
\frac{\#\mathcal{V}_{N,K}^{(k)}\left(\widetilde{q}; (w_i)_{i=1}^K\right)}{\varphi(\widetilde{q})^N} = \frac{\mathcal{S}' + \mathcal{S}''}{\varphi(\widetilde{q})^N} = \left( \frac{\varphi(Q_0)}{\varphi(\widetilde{q})} \right)^K \left\{ \frac{\#\mathcal{V}_{N,K}^{(k)}\left(Q_0; (w_i)_{i=1}^K\right)}{\varphi(Q_0)^N} + O\left( \frac{1}{C^N} \right) \right\}, \tag{5.19}
$$

uniformly for $N \ge KD+1$ and in coprime residues $w_1, \dots, w_K$ to any modulus $q$.

On the other hand, for each $N \in [KD]$, we have $1 + \sum_{\kappa+1 \le \nu \le e} \ell^{-\nu(N/D-K)} \ll e^{\mathbb{1}_{N=KD}} \ell^{e(K-N/D)}$, which from (5.17), yields the bound $|\mathcal{S}''|/\varphi(\widetilde{q})^N \ll \left( \prod_{\ell^e \| \widetilde{q}} e \right)^{\mathbb{1}_{N=KD}} \Big/ \widetilde{q}^{N/D}$. Combining this with the trivial bound $|\mathcal{S}'|/\varphi(\widetilde{q})^N \ll \varphi(\widetilde{q})^{-K} \ll \widetilde{q}^{-K} \ll \widetilde{q}^{-N/D}$ coming from (5.12), we find that for each $N \in [KD]$, we have

$$
\frac{\#\mathcal{V}_{N,K}^{(k)}\left(\widetilde{q}; (w_i)_{i=1}^K\right)}{\varphi(\widetilde{q})^N} \ll \frac{\left( \prod_{\ell^e \| \widetilde{q}} e \right)^{\mathbb{1}_{N=KD}}}{\widetilde{q}^{N/D}}, \quad \text{uniformly in } q \text{ and } (w_i)_{i=1}^K \in U_q^K. \tag{5.20}
$$

Proposition 4.4 now follows in the case $D > 1$ by combining (5.6) with (5.19) (for $N > KD$) or (5.20) (for $N \le KD$), and then noting that $\prod_{\ell \mid q:\ \ell > C_0} \alpha_k(\ell) = \alpha_k(q)/\alpha_k(Q_0)$.

<u>**Now assume that** $D = 1$</u>, so that $K = 1$ and $W_{1,k}(T) := RT + S$ for some integers $R$ and $S$ with $R \neq 0$. We first make the following general observation, which is immediate from Proposition 5.2: For any primitive character $\chi$ mod $\ell^b$, the sum $Z_{\ell^b;\chi} := \sum_{v \bmod \ell^b} \chi_{0,\ell}(v)\chi(Rv+S) = \sum_{v \bmod \ell^b} \chi(v^{\varphi(\ell^b)}(Rv+S))$ is zero for any odd prime $\ell$ and any integer $b \geq v_\ell(R)+2$, as well as for $\ell = 2$ and any $b \geq v_2(R)+3$. Indeed in both these cases, the polynomial $F(T) = T^{\varphi(\ell^b)}(RT+S)$ has no $\ell$-critical point, since $\mathrm{ord}_\ell(F') = v_\ell(R)$ which forces $\ell^{-\mathrm{ord}_\ell(F')}F'(T) = (\ell^{-v_\ell(R)}R)T^{\varphi(\ell^b)}$ in $\mathbb{F}_\ell[T]$.

By this observation, it follows that uniformly in $N \geq 1$ and in $\ell^e \parallel q$ with $\ell > C_0$ $(> |R|)$,

$$\frac{\#\mathcal{V}_{N,1}^{(k)}(\ell^e; w)}{\varphi(\ell^e)^N} = \frac{\alpha_k(\ell)^N}{\varphi(\ell^e)} \left(1 + O\left(\left(\frac{2}{\ell-1}\right)^{N-1}\right)\right). \tag{5.21}$$

Indeed, we simply invoke (5.7) and note that if $\mathfrak{f}(\chi) = \ell^{e_0}$ for some $e_0 \geq 2 = v_\ell(R)+2$, then $Z_{\ell^e;\chi} = 0$ as seen above. On the other hand, if $\mathfrak{f}(\chi) = \ell$ (and there are $\ell - 2$ many such characters mod $\ell^e$), then $|Z_{\ell^e;\chi}| = \ell^{e-1}\left|\sum_{v \bmod \ell} \chi(Rv+S) - \chi(S)\right| = \ell^{e-1}\left|\sum_{u \bmod \ell}\chi(u) - \chi(S)\right| \leq \ell^{e-1}$.

Letting $\widetilde{q} := \prod_{\substack{\ell^e \parallel q\\ \ell \leq C_0}} \ell^e$ as before, we fix an integer $\kappa > C_0 + 3$, and write $\#\mathcal{V}_{N,1}^{(k)}(\widetilde{q}; w) = \varphi(\widetilde{q})^{-1}\sum_{\chi \bmod \widetilde{q}} \overline{\chi}(w)(Z_{\widetilde{q};\chi})^N = \mathcal{S}' + \mathcal{S}''$, where $\mathcal{S}'$ again denotes the contribution of those $\chi$ mod $\widetilde{q}$ for which $\mathfrak{f}(\chi)$ is $(\kappa+1)$-free. Then (5.12) continues to hold, and $\mathcal{S}'' = 0$ by the general observation above. This yields $\#\mathcal{V}_{N,1}^{(k)}(\widetilde{q}; w)/\varphi(\widetilde{q})^N = (\varphi(Q_0)/\varphi(\widetilde{q})) \cdot (\#\mathcal{V}_{N,1}^{(k)}(Q_0; w)/\varphi(Q_0)^N)$, which along with (5.21), establishes Proposition 4.4 in the remaining case $D = 1$. □

With Proposition 4.4 established, the proof of Proposition 4.3 is now complete. We will eventually also need the following variant of Proposition 4.4, which follows from an argument that is a much simpler version of that given for (5.6).

**Corollary 5.4.** *Assume that $\{W_{i,k}\}_{1\leq i\leq K}$ are multiplicatively independent. Then*

$$\frac{\#\mathcal{V}_{N,K}^{(k)}\left(q; (w_i)_{i=1}^K\right)}{\varphi(q)^N} \ll \begin{cases} \varphi(q)^{-K}\exp\left(O(\sqrt{\log q})\right), & \text{for each fixed } N \geq 2K+1 \\ q^{-N/2}\exp\left(O(\omega(q))\right), & \text{for each fixed } N \leq 2K, \end{cases} \tag{5.22}$$

*uniformly in $w_i \in U_q$ modulo squarefree $q$ satisfying $\alpha_k(q) \neq 0$ and $IFH(W_{1,k}, \dots, W_{K,k}; B_0)$.*

**Towards Theorem 4.2.**

In order to deduce Theorem 4.2 from Proposition 4.3, we apply the orthogonality of Dirichlet characters to see that the main term in the right hand side of (4.6) is equal to

$$\left(\frac{\varphi(Q_0)}{\varphi(q)}\right)^K \sum_{\substack{n\leq x:\ (f(n),q)=1\\ (\forall i)\ f_i(n)\equiv a_i \pmod{Q_0}}} 1$$

$$= \frac{1}{\varphi(q)^K}\sum_{\substack{n\leq x\\ (f(n),q)=1}} 1 + \frac{1}{\varphi(q)^K} \sum_{(\chi_1,\dots,\chi_K)\neq(\chi_{0,Q_0},\dots,\chi_{0,Q_0}) \bmod Q_0} \left(\prod_{i=1}^K \overline{\chi}_i(a_i)\right) \sum_{n\leq x} \mathbb{1}_{(f(n),q)=1} \prod_{i=1}^K \chi_i(f_i(n)).$$

Henceforth, let $Q := \prod_{\ell \mid q} \ell$ denote the radical of $q$. To obtain Theorem 4.2, it remains to prove that each $\sum_{n \le x} \mathbb{1}_{(f(n),q)=1} \prod_{i=1}^K \chi_i(f_i(n)) = o\left( \sum_{\substack{n \le x \\ (f(n),q)=1}} 1 \right)$. For $Q \ll 1$, this follows by applying Theorem N to the divisor $Q^* := \mathrm{lcm}[Q, Q_0] \ll 1$ of $q$. (Note that as $q$ lies in $\mathcal{Q}(k; f_1, \cdots, f_K)$, so does $Q^*$, since $q$ and $Q^*$ have the same prime factors.) So we may assume that $Q$ is sufficiently large. Theorem 4.2 would follow once we show the result below. Here $\lambda$ and $Q_0$ are as in Proposition 4.3.

**Theorem 5.5.** *There exists a constant $\delta_0 := \delta_0(\lambda) > 0$ such that, uniformly in moduli $q \le (\log x)^{K_0}$ lying in $\mathcal{Q}(k; f_1, \cdots, f_K)$ and having sufficiently large radical, we have*

$$\sum_{n \le x} \chi_1(f_1(n)) \cdots \chi_K(f_K(n)) \mathbb{1}_{(f(n),q)=1} \ll \frac{x^{1/k}}{(\log x)^{1-(1-\delta_0)\alpha_k(Q)}}$$

*for all tuples of characters $(\chi_1, \ldots, \chi_K) \ne (\chi_{0,Q_0}, \ldots, \chi_{0,Q_0}) \bmod Q_0$.*

Let $\mathcal{C}_k(Q_0)$ denote the set of tuples of characters $(\psi_1, \ldots, \psi_K) \bmod Q_0$, not all trivial, such that $\prod_{i=1}^K \psi_i(W_{i,k}(u))$ is constant on its support, which is precisely the set $R_k(Q_0) = \{u \in U_{Q_0} : W_k(u) \in U_{Q_0}\}$. To prove Theorem 5.5, we separately consider the two cases when a tuple of characters mod $Q_0$ lies in $\mathcal{C}_k(Q_0)$ or not.

## 6. Proof of Theorem 5.5 for nontrivial tuples of characters not in $\mathcal{C}_k(Q_0)$

For any integer $d \ge 1$ and any nontrivial tuple $(\psi_1, \ldots, \psi_K)$ of characters mod $d$ not lying in $\mathcal{C}_k(d)$, we have $|\sum_{u \bmod d} \chi_{0,d}(u) \psi_1(W_{1,k}(u)) \cdots \psi_K(W_{K,k}(u))| < \alpha_k(d)\varphi(d)$. With $\lambda$ as in Proposition 4.3, we define the constant $\delta_1 := \delta_1(W_{1,k}, \ldots, W_{K,k}; B_0) \in (0,1)$ to be

$$\max_{\substack{d \le \lambda \\ \alpha_k(d) \ne 0}} \max_{\substack{(\psi_1,\ldots,\psi_K) \ne (\chi_{0,d},\ldots,\chi_{0,d}) \bmod d \\ (\psi_1,\ldots,\psi_K) \notin \mathcal{C}_k(d)}} \frac{1}{\alpha_k(d)\varphi(d)} \left| \sum_{u \bmod d} \chi_{0,d}(u) \psi_1(W_{1,k}(u)) \cdots \psi_K(W_{K,k}(u)) \right|.$$

Then since $Q_0 \le \lambda$, we have for any nontrivial tuple $(\chi_1, \ldots, \chi_K) \notin \mathcal{C}_k(Q_0)$,

$$\left| \sum_{u \bmod Q_0} \chi_{0,Q_0}(u) \chi_1(W_{1,k}(u)) \cdots \chi_K(W_{K,k}(u)) \right| \le \delta_1 \alpha_k(Q_0) \varphi(Q_0). \tag{6.1}$$

We set $\delta := (1-\delta_1)/2$ and $Y := \exp((\log x)^{\delta/3})$. To establish Theorem 5.5 for all $(\chi_1, \ldots, \chi_K) \notin \mathcal{C}_k(Q_0)$, it suffices to show that

$$\sum_{\substack{n \le x \\ p > Y \implies p^{k+1} \nmid n}} \chi_1(f_1(n)) \cdots \chi_K(f_K(n)) \, \mathbb{1}_{(f(n),q)=1} \ll \frac{x^{1/k}}{(\log x)^{1-(\delta_1+\delta)\alpha_k}}. \tag{6.2}$$

This is because by the arguments before (3.3), the contribution of the $n$'s not counted above is negligible. Writing any $n$ counted in (6.2) uniquely as $BMA^k$ (as done before (3.4)), we see that the sum in (6.2) equals

$$\sum_{\substack{B\le x\\ P(B)\le Y\\ B \text{ is } k\text{-free}}} \mathbb{1}_{(f(B),q)=1}\left(\prod_{i=1}^K \chi_i(f_i(B))\right) \sum_{\substack{M\le x/B\\ M \text{ is } k\text{-full}\\ P(M)\le Y}} \mathbb{1}_{(f(M),q)=1}\left(\prod_{i=1}^K \chi_i(f_i(M))\right) \tag{6.3}$$
$$\sum_{A\le (x/BM)^{1/k}} \mathbb{1}_{P^-(A)>Y}\ \mathbb{1}_{(f(A^k),q)=1}\ \mu(A)^2 \prod_{i=1}^K \chi_i(f_i(A^k))$$

Moreover, the arguments leading to the bound for $\Sigma_2$ towards the end of section 3 show that the tuples $(B, M, A)$ having $M > x^{1/2}$ give negligible contribution to the above sum. To prove (6.2), it thus only remains to bound the contribution of tuples $(B, M, A)$ with $M \le x^{1/2}$ to the triple sum in (6.3). To deal with such tuples, we will establish the following general upper bound uniformly for $X \ge \exp((\log Y)^2)$:

$$\sum_{A\le X} \mathbb{1}_{P^-(A)>Y}\ \mathbb{1}_{(f(A^k),q)=1}\ \mu(A)^2 \prod_{i=1}^K \chi_i(f_i(A^k)) \ll \frac{X}{(\log X)^{1-\alpha_k(\delta_1+\delta/2)}}. \tag{6.4}$$

We apply a quantitative version of Halász's Theorem [50, Corollary III.4.12] on the multiplicative function $F(A) := \mathbb{1}_{P^-(A)>Y}\ \mathbb{1}_{(f(A^k),q)=1}\ \mu(A)^2 \prod_{i=1}^K \chi_i(f_i(A^k))$, taking $T := \log X$. This requires us to put, for each $t \in [-T, T]$, a lower bound on the sum below (which is the square of a certain "pretentious distance"):

$$\mathcal{D}(X;t) := \sum_{p\le X} \frac{1}{p}\left(1 - \operatorname{Re}\left(\mathbb{1}_{p>Y}\ \mathbb{1}_{(f(p^k),q)=1}\ \mu(p)^2\ p^{-\mathrm{i}t} \prod_{i=1}^K \chi_i(f_i(p^k))\right)\right)$$
$$= (1-\alpha_k)\log_2 X + \alpha_k \log_2 Y + \sum_{\substack{Y<p\le X\\ (W_k(p),q)=1}} \frac{1}{p}\left(1 - \operatorname{Re}\left(p^{-\mathrm{i}t} \prod_{i=1}^K \chi_i(W_{i,k}(p))\right)\right) \tag{6.5}$$
$$+ O\big((\log_2(3q))^{O(1)}\big);$$

here the second line uses Lemma 3.4. To get this lower bound, we proceed analogously to the proof of [37, Lemma 3.3]. The key idea is to split the range of the last sum above into blocks of small multiplicative width, so that the complex number $p^{-\mathrm{i}t}$ is essentially constant for all $p$ lying in a given block. More precisely, we cover the interval $(Y, X]$ with finitely many disjoint intervals $\mathcal{I} := \big(\eta, \eta(1 + 1/\log^2 X)\big]$ for certain choices of $\eta \in (Y, X]$, choosing the smallest $\eta$ to be $Y$ and allowing the rightmost endpoint of such an interval to jut out slightly past $X$ but no more than $X(1 + 1/\log^2 X)$. Then the last sum in (6.5) equals

$$\sum_{\mathcal{I}} \sum_{\substack{p\in\mathcal{I}\\ (W_k(p),q)=1}} \frac{1}{p}\left(1 - \operatorname{Re}\left(p^{-\mathrm{i}t} \prod_{i=1}^K \chi_i(W_{i,k}(p))\right)\right) + O\left(\frac{1}{\log^3 X}\right) \tag{6.6}$$

Consider any $\mathcal{I}$ occurring in the sum above. For each $p \in \mathcal{I}$, we have

$$|p^{-\mathrm{i}t} - \eta^{-\mathrm{i}t}| \le \left|\int_{t\log\eta}^{t\log p} \exp(-i\varrho)\, d\varrho\right| \le |t\log p - t\log\eta| \le \frac{|t|}{\log^2 X} \le \frac{1}{\log X}.$$

This shows that uniformly in $\mathcal{I}$, the inner sum in (6.6) is equal to

$$\begin{aligned}(6.7)\qquad &\sum_{\substack{p\in\mathcal{I}\\(W_k(p),q)=1}}\frac{1}{p}\left(1-\mathrm{Re}\left(p^{-\mathrm{i}t}\prod_{i=1}^{K}\chi_i(W_{i,k}(p))\right)\right)\\ &=\sum_{\substack{u\in U_q\\(W_k(u),q)=1}}\left(1-\mathrm{Re}\left(\eta^{-\mathrm{i}t}\prod_{i=1}^{K}\chi_i(W_{i,k}(u))\right)\right)\sum_{\substack{p\in\mathcal{I}\\p\equiv u \ (\mathrm{mod}\ q)}}\frac{1}{p}+O\left(\frac{1}{\log X}\sum_{p\in\mathcal{I}}\frac{1}{p}\right)\end{aligned}$$

Note that $p=(1+o(1))\eta$ for all $p\in\mathcal{I}$. (Here and in what follows, the asymptotic notation refers to the behavior as $x\to\infty$, and is uniform in the choice of $\mathcal{I}$.) For parameters $Z,W$ depending on $X$, we write $Z\gtrsim W$ to mean $Z\geq(1+o(1))W$. By the Siegel Walfisz Theorem,

$$\sum_{\substack{p\in\mathcal{I}\\p\equiv u \ (\mathrm{mod}\ q)}}\frac{1}{p}\gtrsim\frac{1}{\eta}\sum_{\substack{p\in\mathcal{I}\\p\equiv u \ (\mathrm{mod}\ q)}}1\gtrsim\frac{1}{\varphi(q)}\cdot\frac{1}{\eta}\sum_{p\in\mathcal{I}}1\gtrsim\frac{1}{\varphi(q)}\sum_{p\in\mathcal{I}}\frac{1}{p}.$$

Hence the whole main term on the right hand side of (6.7) is

$$(6.8)\quad \gtrsim\frac{1}{\varphi(q)}\sum_{p\in\mathcal{I}}\frac{1}{p}\sum_{\substack{u\in U_q\\(W_k(u),q)=1}}\left(1-\mathrm{Re}\left(\eta^{-\mathrm{i}t}\prod_{i=1}^{K}\chi_i(W_{i,k}(u))\right)\right)\gtrsim(\alpha_k-\alpha_k\delta_1)\left(\sum_{p\in\mathcal{I}}\frac{1}{p}\right),$$

where in the last step above, we have used (4.14) and (6.1) to see that

$$\frac{1}{\varphi(q)}\left|\sum_{\substack{u\in U_q\\(W_k(u),q)=1}}\prod_{i=1}^{K}\chi_i(W_{i,k}(u))\right|=\frac{\alpha_k(q)}{\alpha_k(Q_0)\varphi(Q_0)}\left|\sum_{r \bmod Q_0}\chi_{0,Q_0}(r)\prod_{i=1}^{K}\chi_i(W_{i,k}(r))\right|\leq\alpha_k\delta_1.$$

Inserting the bound obtained in (6.8) into (6.7), we find that each inner sum in (6.6) is

$$\sum_{\substack{p\in\mathcal{I}\\(W_k(p),q)=1}}\frac{1}{p}\left(1-\mathrm{Re}\left(p^{-\mathrm{i}t}\prod_{i=1}^{K}\chi_i(W_{i,k}(p))\right)\right)\gtrsim\alpha_k(1-\delta_1)\sum_{p\in\mathcal{I}}\frac{1}{p}+O\left(\frac{1}{\log X}\sum_{p\in\mathcal{I}}\frac{1}{p}\right).$$

The $O$-term above when summed over all $\mathcal{I}$ is $\ll(\log X)^{-1}\sum_{p\leq 2X}p^{-1}\ll\log_2 X/\log X$. Thus, the whole main term in (6.6) is at least $\alpha_k\left(1-\delta_1-\frac{\delta}{2}\right)(\log_2 X-\log_2 Y)$. Using this fact along with (6.5) yields

$$\mathcal{D}(X;t)\geq\left(1-\alpha_k\left(\delta_1+\frac{\delta}{2}\right)\right)\log_2 X+\alpha_k\left(\delta_1+\frac{\delta}{2}\right)\log_2 Y+O\big((\log_2(3q))^{O(1)}\big),$$

uniformly for $t\in[-T,T]$. As such, [50, Corollary III.4.12] establishes the claimed bound (6.4).

Now for each $M\leq x^{1/2}$, we have $(x/BM)^{1/k}\gg x^{1/2k}$. Applying (6.4) to each of the innermost sums in (6.3), we see that the total contribution of all tuples $(B,M,A)$ with $M\leq x^{1/2}$ to the triple sum in (6.3) is

$$\ll\sum_{B\ll 1}\sum_{\substack{M\leq x^{1/2}:\ M\text{ is }k\text{-full}\\P(M)\leq Y,\ (f(M),q)=1}}\frac{(x/BM)^{1/k}}{(\log x)^{1-\alpha_k(\delta_1+\delta/2)}}\ll\frac{x^{1/k}}{(\log x)^{1-\alpha_k(\delta_1+\delta)}};$$

here we have bounded the sum on $M$ using (3.5) (with "$Y$" playing the role of "$y$") and Lemma 3.4. This proves (6.2), and hence also Theorem 5.5 for all nontrivial tuples of characters $(\chi_1, \dots, \chi_K)$ mod $Q_0$ not in $\mathcal{C}_k(Q_0)$. □

## 7. Proof of Theorem 5.5 for tuples of characters in $\mathcal{C}_k(Q_0)$

It suffices to consider the case when $x$ is an integer, and we will do so in the rest of the section. Our argument consists of suitably modifying the Landau–Selberg–Delange method for mean values of multiplicative functions (see for instance [50, Chapter II.5]), and to study the behavior of a product of $L$-functions raised to complex powers by accounting for the presence of Siegel zeros modulo $q$. This is partly inspired from work of Scourfield [45] and will also need some results from her paper. We will denote complex numbers in the standard notation $s = \sigma + it$.[8]

Recall that $Q = \prod_{\ell | q} \ell$; since $q$ is $k$-admissible, so is $Q$. Consider any $\widehat{\chi} := (\chi_1, \dots, \chi_K) \in \mathcal{C}_k(Q_0)$, so that the product $\prod_{i=1}^K \chi_i(W_{i,k}(u))$ is constant on $R_k(Q_0)$; let $c_{\widehat{\chi}}$ denote this constant value. Consider the Dirichlet series

$$F_\chi(s) := \sum_{n \geq 1} \frac{\mathbb{1}_{(f(n),q)=1}}{n^s} \prod_{i=1}^K \chi_i(f_i(n)) = \sum_{n \geq 1} \frac{\mathbb{1}_{(f(n),Q)=1}}{n^s} \prod_{i=1}^K \chi_i(f_i(n))$$

which is absolutely convergent in the half-plane $\sigma > 1$.

In the rest of this section, we fix $\mu_0$ satisfying $\max\{0.7, k/(k+1)\} < \mu_0 < 1$.

### 7.1. Analysis of the Dirichlet series.

We start by giving a meromorphic continuation of $F_\chi(s)$ to a larger region. To do this, set $\mathcal{L}_Q(t) := \log(Q(|tk| + 1))$ and recall that there exists an absolute constant $c_1 > 0$ such that the product $\prod_{\psi \bmod Q} L(s, \psi)$ has at most one zero $\beta_e$ (counted with multiplicity) in the region $\sigma > 1 - c_1/\log(Q(|t| + 1))$, which is necessarily real and simple; $\beta_e$ is called the "Siegel zero". If $\beta_e$ exists, then it is a root of $L(s, \psi_e)$ for some real character $\psi_e$ mod $Q$, which we will be referring to as the "exceptional character". By reducing the constant $c_1$ if necessary, we may assume that $c_1 < 1 - \mu_0$, and that the conductor of $\psi_e$ (which is squarefree) is large enough that it is not $(D+2)$-smooth.

Let $\mathcal{D}_k(c_0)$ denote the region $\left\{\sigma + it : \sigma > \frac{1}{k}\left(1 - \frac{c_1}{\mathcal{L}_Q(t)}\right)\right\}$. Then $\prod_{\psi \bmod Q} L(sk, \psi)$ has at most one zero and exactly one pole in the region $\mathcal{D}_k(c_0)$, namely $\beta_e/k$ and $1/k$, respectively.

**Branch cuts and complex logarithms:** In the rest of the section, we assume that the complex plane has been cut along the line $\sigma \leq 1/k$ if $\alpha_k(Q)$ and $c_{\widehat{\chi}}$ are not both 1, whereas if $\alpha_k(Q) = c_{\widehat{\chi}} = 1$, then the complex plane is cut along the line $\sigma \leq \beta_e/k$. (If $\alpha_k(Q) = c_{\widehat{\chi}} = 1$ and if there is also no Siegel zero mod $q$, then there is no cut.)

**Lemma 7.1.** *The Dirichlet series $F_\chi(s)$ is absolutely convergent on the half-plane $\sigma > \frac{1}{k}$, where it satisfies*

$$F_\chi(s) = F_1(sk)^{c_{\widehat{\chi}}}\ g(sk)^{c_{\widehat{\chi}}}\ G_{\chi,1}(s)\ G_{\chi,2}(s) \tag{7.1}$$

[8] The parameters $\sigma$ and $\sigma_k$ (to be defined later) in this section have nothing to do with the divisor functions $\sigma_r(n) = \sum_{d|n} d^r$ mentioned in the introduction. We are not working with the divisor functions in this section.

*with*

$$F_1(sk) = \left( \prod_{Q_1 | Q} \prod_{\substack{\psi \bmod Q_1 \\ \psi \text{ primitive}}} L(sk, \psi)^{\gamma(\psi)} \right)^{\alpha_k(Q)}$$

$$g(sk) = \left( \prod_{Q_1 | Q} \prod_{\substack{\psi \bmod Q_1 \\ \psi \text{ primitive}}} \prod_{\ell | \frac{Q}{Q_1}} \left(1 - \frac{\psi(\ell)}{\ell^{ks}}\right)^{\gamma(\psi)} \right)^{\alpha_k(Q)}, \quad \gamma(\psi) = \frac{1}{\alpha_k(Q)\varphi(Q)} \sum_{\substack{v \in U_Q \\ W_k(v) \in U_Q}} \overline{\chi}(v).$$

*Here, the functions $F_1(sk)$, $g(sk)$, $G_{\chi,1}(s)$ and $G_{\chi,2}(s)$ satisfy the following properties:*

(i) *$F_1(sk)$ is holomorphic and nonvanishing in the region $\mathcal{D}_k(c_0) - (-\infty, 1/k]$.* [9] *In fact, if $\alpha_k(Q) = c_{\widehat{\chi}} = 1$ and if $\beta_e$ exists (resp. doesn't exist), then the same is true in the bigger region $\mathcal{D}_k(c_0) - (-\infty, \beta_e/k]$ (resp. $\mathcal{D}_k(c_0)$).*

(ii) *$g(sk)$ and $G_{\chi,1}(s)$ are holomorphic and nonvanishing in the half-plane $\sigma > \mu_0/k$, and we have, uniformly for all $s$ in this region,*

$$\max\left\{ \left| \frac{g'(sk)}{g(sk)} \right|, \left| \frac{G'_{\chi,1}(s)}{G_{\chi,1}(s)} \right| \right\} \ll \max\{1, (\log Q)^{1-\sigma k}\} \ \log\log Q. \tag{7.2}$$

(iii) *$G_{\chi,2}(s)$ is holomorphic in the half-plane $\sigma > \mu_0/k$, wherein $|G_{\chi,2}(s)|, |G'_{\chi,2}(s)| \ll 1$.*

*Proof.* **Absolute convergence of $F_\chi(s)$ on the region $\sigma > 1/k$:** To see this, we start by noting that $F_\chi(s)$ is tautologically absolutely convergent on $\sigma > 1$, and in this half plane, we have the Euler product

$$F_\chi(s) = \prod_p \left( 1 + \sum_{v \geq 1} \frac{\mathbb{1}_{(f(p^v), Q) = 1}}{p^{vs}} \prod_{i=1}^{K} \chi_i(f_i(p^v)) \right). \tag{7.3}$$

In the rest of the proof, we **fix** $B_k > 2^{k/\mu_0}$ such that $B_k$ exceeds any $k$-free integer $n$ satisfying $\gcd(f(n), q) = 1$; recall that by Lemma 3.3, $B_k$ can be chosen to depend only on $\{W_{i,v}\}_{\substack{1 \leq i \leq K \\ 1 \leq v \leq k}}$. Then the contribution of primes $p \leq B_k$ to the aforementioned Euler product is a finite product, each factor of which is absolutely convergent in the region $\sigma > 0$. On the other hand, by Lemma 3.3 and the facts that $Q$ is $k$-admissible and $(\chi_1, \ldots, \chi_K) \in \mathcal{C}_k(Q_0)$, the total contribution of all primes $p > B_k$ to the above Euler product (7.3) is

$$\prod_{p > B_k} \left( 1 + \frac{c_{\widehat{\chi}} \mathbb{1}_{(W_k(p), Q) = 1}}{p^{ks}} + O\left( \frac{1}{p^{(k+1)\sigma}} \right) \right), \tag{7.4}$$

which is absolutely convergent in the region $\sigma > 1/k$, since the sum $\sum_p c_{\widehat{\chi}} \mathbb{1}_{(W_k(p), Q) = 1} / p^{ks}$ is. This shows that $F_\chi(s)$ is absolutely convergent on the region $\sigma > 1/k$.

**The product decomposition** (7.1)**:** Thus (7.3) holds in the region $\sigma > 1/k$, and in this same region, we may write

[9] This region is obtained by omitting the ray $(-\infty, 1/k]$ from the region $\mathcal{D}_k(c_0)$.

$$F_\chi(s) = \left( \prod_{\substack{b \in U_Q \\ W_k(b) \in U_Q}} \prod_{p \equiv b \ (\mathrm{mod}\ Q)} \left(1 - \frac{1}{p^{ks}}\right)^{-c_{\widehat{\chi}}} \right) \cdot \left( \prod_{\substack{p \mid Q \\ W_k(p) \in U_Q}} \left(1 - \frac{1}{p^{ks}}\right)^{-c_{\widehat{\chi}}} \right) \\ \cdot \prod_p \left( 1 + \sum_{v \geq 1} \frac{\mathbb{1}_{(f(p^v), Q) = 1}}{p^{vs}} \prod_{i=1}^K \chi_i(f_i(p^v)) \right) \left( 1 - \frac{\mathbb{1}_{(W_k(p), Q) = 1}}{p^{ks}} \right)^{c_{\widehat{\chi}}} \tag{7.5}$$

Now for $\sigma > 1/k$, the orthogonality of Dirichlet characters mod $Q$ and the fact that $\log L(sk, \psi) = \sum_{p,v} \psi(p^v)/p^{vsk}$ show that the logarithm of the first double product in (7.5) is equal to

$$c_{\widehat{\chi}} \sum_{\substack{b \in U_Q \\ W_k(b) \in U_Q}} \sum_{\substack{p, v \geq 1 \\ p \equiv b \ (\mathrm{mod}\ Q)}} \frac{1}{vp^{vks}} \ = \ c_{\widehat{\chi}} \sum_{\substack{b \in U_Q \\ W_k(b) \in U_Q}} \left\{ \frac{1}{\varphi(Q)} \sum_{\psi \bmod Q} \overline{\psi}(b) \sum_p \frac{\psi(p)}{p^{ks}} + \sum_{\substack{p, v \geq 2 \\ p \equiv b \ (\mathrm{mod}\ Q)}} \frac{1}{vp^{vks}} \right\}$$

$$= \alpha_k(Q) c_{\widehat{\chi}} \sum_{\psi \bmod Q} \gamma(\psi) \log L(sk, \psi) + c_{\widehat{\chi}} \sum_{\substack{b \in U_Q \\ W_k(b) \in U_Q}} \sum_{v \geq 2} \left( \sum_{p \equiv b \ (\mathrm{mod}\ Q)} \frac{1}{vp^{vks}} - \sum_{p \colon p^v \equiv b \ (\mathrm{mod}\ q)} \frac{1}{vp^{vks}} \right).$$

We insert this into (7.5), noting that $L(sk, \psi) = L(sk, \psi^*) \prod_{\ell \mid \frac{Q}{Q_1}} (1 - \psi^*(\ell)/\ell^{sk})$ and that $\gamma(\psi) = \gamma(\psi^*)$ if the primitive character $\psi^*$ mod $Q_1$ induces $\psi$ mod $Q$. This yields (7.1), with

$$G_{\chi,2}(s) := \prod_{p \leq B_k} \left( 1 + \sum_{v \geq 1} \frac{\mathbb{1}_{(f(p^v), Q) = 1}}{p^{vs}} \prod_{i=1}^K \chi_i(f_i(p^v)) \right) \left( 1 - \frac{\mathbb{1}_{(W_k(p), Q) = 1}}{p^{ks}} \right)^{c_{\widehat{\chi}}}$$

and

$$G_{\chi,1}(s) := \prod_{p > B_k} \left( 1 + \sum_{v \geq 1} \frac{\mathbb{1}_{(f(p^v), Q) = 1}}{p^{vs}} \prod_{i=1}^K \chi_i(f_i(p^v)) \right) \left( 1 - \frac{\mathbb{1}_{(W_k(p), Q) = 1}}{p^{ks}} \right)^{c_{\widehat{\chi}}} \\ \cdot \prod_{\substack{p \mid Q \\ W_k(p) \in U_Q}} \left(1 - \frac{1}{p^{ks}}\right)^{-c_{\widehat{\chi}}} \cdot \exp\left( c_{\widehat{\chi}} \sum_{\substack{b \in U_Q \\ W_k(b) \in U_Q}} \sum_{v \geq 2} \left( \sum_{p \equiv b \ (\mathrm{mod}\ Q)} \frac{1}{vp^{vks}} - \sum_{p \colon p^v \equiv b \ (\mathrm{mod}\ q)} \frac{1}{vp^{vks}} \right) \right), \tag{7.6}$$

where $B_k$ was as defined after (7.3).

**Proving statements (i)–(iii) of the lemma:** To see (i), recall that $\prod_{\psi \bmod Q} L(sk, \psi)$ has is holomorphic and nonvanishing in the region $\mathcal{D}_k(c_0) - (-\infty, 1/k]$. In fact, if $\alpha_k(Q) = c_{\widehat{\chi}} = 1$, then $F_1(sk) = L(sk, \chi_0) \cdot \left( \prod_{\substack{Q_1 \mid Q \\ Q_1 > 1}} \prod_{\substack{\psi \bmod Q_1 \\ \psi \text{ primitive}}} L(sk, \psi)^{\gamma(\psi)} \right)^{\alpha_k(Q)}$, which shows the other assertions of (i). Also (iii) is immediate by a direct calculation using the definition of $G_{\chi,2}(s)$.

We thus focus on (ii). By the very definition of $g(sk)$, we see that it is holomorphic and nonvanishing in the half-plane $\sigma > 0$. Also the bound on $|g'(sk)/g(sk)|$ in (7.2) is an immediate consequence of [45, Lemma 9(ii)].

To show the assertions for $G_{\chi,1}(s)$, we recall that by the arguments preceding (7.4) the first product (over primes $p > B_k$) in (7.6) is equal to

$$\prod_{p>B_k}\left(1+\frac{c_{\widehat{\chi}}\mathbb{1}_{(W_k(p),Q)=1}}{p^{ks}}+O\left(\frac{1}{p^{(k+1)\sigma}}\right)\right)\left(1-\frac{\mathbb{1}_{(W_k(p),Q)=1}}{p^{ks}}\right)^{c_{\widehat{\chi}}} = \prod_{p>B_k}\left(1+O\left(\frac{1}{p^{(k+1)\sigma}}\right)\right),$$

which is absolutely convergent and defines a holomorphic function in the half plane $\sigma > \mu_0/k$. (Here is it important that $\mu_0/k > 1/(k+1)$.) Likewise the exponential factor in (7.6) defines a holomorphic function in the same half plane, hence so does $G_{\chi,1}(s)$. To see that $G_{\chi,1}(s)$ is also nonvanishing in this region, we need only see that the condition $p > B_k > 2^{k/\mu_0}$ guarantees that each of the factors $1+\sum_{v\geq 1}\frac{\mathbb{1}_{(f(p^v),Q)=1}}{p^{vs}}\prod_{i=1}^{K}\chi_i(f_i(p^v))$ in (7.6) has size at least $1-\sum_{v\geq k}p^{-v\sigma} > 1-2p^{-k\sigma} > 1-2B_k^{-\mu_0} > 0$. Finally, a straightforward computation using (7.6) shows that for $\sigma > \mu_0/k$, we have

$$\frac{G'_{\chi,1}(s)}{G_{\chi,1}(s)} = -c_{\widehat{\chi}}k\sum_{\substack{p\mid Q\\ W_k(p)\in U_Q}}\frac{\log p}{p^{ks}}+O(1) \ll \sum_{p\mid Q}\frac{\log p}{p^{k\sigma}},$$

completing the proof of (7.2) via [45, Lemma 3(i)(a)]. □

7.2. **Preparing for the contour shift: Auxiliary functions and intermediate bounds.** Our objective is to relate the sum in Theorem 5.5 to the Dirichlet series $F_\chi(s)$ by an effective version of Perron's formula, and shift the contour to the left of the line $\sigma = 1/k$. As such, we will need the following proposition in order to estimate the resulting integrals.

To set up, we choose $\epsilon_1 := \epsilon_1(\lambda)$ to be a constant (depending only on $\lambda$) satisfying $0 < \epsilon_1 < 1-\cos(2\pi/d)$ for any positive integer $d \leq \lambda$. Consider the functions

$$\widetilde{F}_\chi(s) := F_1(sk)^{c_{\widehat{\chi}}}\, g(sk)^{c_{\widehat{\chi}}}\, G_{\chi,1}(s)$$

$$\widetilde{H}_\chi(s) := \widetilde{F}_\chi(s)\left(s-\frac{1}{k}\right)^{\alpha_k(Q)c_{\widehat{\chi}}}\left(s-\frac{\beta_e}{k}\right)^{-\alpha_k(Q)c_{\widehat{\chi}}\gamma(\psi_e)},\quad H_\chi(s) := \frac{\widetilde{F}_\chi(s)}{s}\left(s-\frac{1}{k}\right)^{\alpha_k(Q)c_{\widehat{\chi}}},$$

where here and in what follows, any term or factor involving $\beta_e$ is to be understood as omitted if the Siegel zero doesn't exist. By Lemma 7.1(i) and (ii), we see that:

1. $\widetilde{F}_\chi(s)$ is holomorphic and nonvanishing on $\mathcal{D}_k(c_0) - (-\infty, 1/k]$. If $\alpha_k(Q) = c_{\widehat{\chi}} = 1$ and if $\beta_e$ exists (resp. doesn't exist), then the same is true on $\mathcal{D}_k(c_0) - (-\infty, \beta_e/k]$ (resp. $\mathcal{D}_k(c_0)$).
2. $H_\chi(s)$ analytically continues into and is nonvanishing on $\mathcal{D}_k(c_0) - (-\infty, \beta_e/k]$.
3. $\widetilde{H}_\chi(s)$ analytically continues into and is nonvanishing on $\mathcal{D}_k(c_0)$.

(Recall our branch cut conventions elucidated at the start of the section.)

In what follows, we set $T := \exp(\sqrt{\log x})$.

**Proposition 7.2.** *We have the following bounds:*

(i) $|H_\chi(1/k)| \ll (\log x)^{\alpha_k(Q)\epsilon_1/5}$.

(ii) $|\widetilde{H}_\chi(s)| \ll (\log x)^{\alpha_k(Q)\epsilon_1/4}$ *uniformly for real $s$ satisfying* $\frac{1}{k}\left(1-\frac{c_1}{4\log Q}\right) \leq s \leq \frac{1}{k}$.

(iii) $|F_\chi(s)| \ll (\log x)^{(1/2+\epsilon_1)\alpha_k(Q)}$ *uniformly for complex numbers* $s = \sigma + it$ *satisfying* $\sigma \geq \frac{1}{k}\left(1 - \frac{c_1}{2\mathcal{L}_Q(t)}\right)$, $|t| \leq T$ *and* $|s - \theta/k| \gg 1/\mathcal{L}_Q(t)$ *for* $\theta \in \{1, \beta_e\}$.

(iv) *Uniformly in real* $s \leq 1/k$ *satisfying* $s \geq \frac{1}{k}\left(\frac{2}{3} + \frac{\beta_e}{3}\right)$ *(if the Siegel zero exists) or* $s \geq \frac{1}{k}\left(1 - \frac{c_1}{4\log Q}\right)$ *(otherwise), we have*

$$\left|H_\chi\left(\frac{1}{k}\right)G_{\chi,2}\left(\frac{1}{k}\right) - H_\chi(s)G_{\chi,2}(s)\right| \ll (\log x)^{(1/20+\alpha_k(Q)/5)\epsilon_1}\left(\frac{1}{k} - s\right).$$

*Proof.*
**General observation:** We have $|\widetilde{H}_\chi(s)| \asymp |\widetilde{H}_\chi(w)|$ uniformly in complex numbers $s$ and $w$ satisfying $\mathrm{Im}(s) = \mathrm{Im}(w) =: t$, and $|s - w| \ll \mathcal{L}_Q(t)^{-1}$ and $\mathrm{Re}(w) \geq \mathrm{Re}(s) \geq \frac{1}{k}\left(1 - \frac{c_1}{2\mathcal{L}_Q(t)}\right)$.

Indeed by the definitions of $\widetilde{H}_\chi(s)$ and $\widetilde{F}_\chi(s)$, we have

$$\left|\frac{\widetilde{H}'_\chi(z)}{\widetilde{H}_\chi(z)}\right| = \left|c_{\widehat{\chi}}k\left(\frac{F_1'(kz)}{F_1(kz)} + \frac{\alpha_k(Q)}{kz-1} - \frac{\alpha_k(Q)\gamma(\psi_e)}{kz-\beta_e}\right) + c_{\widehat{\chi}}k\frac{g'(kz)}{g(kz)} + \frac{G'_{\chi,1}(z)}{G_{\chi,1}(z)}\right| \ll \mathcal{L}_Q(t) \tag{7.7}$$

uniformly for complex numbers $z = u + it$ satisfying $u \geq \frac{1}{k}\left(1 - \frac{c_1}{2\mathcal{L}_Q(t)}\right)$. In the last bound above, we have used (7.2) as well as [45, Lemma 15(i)] with "$\xi$" being $\exp(6\mathcal{L}_Q(t))$. The general observation now follows by writing $\log\left(\widetilde{H}_\chi(w)/\widetilde{H}_\chi(s)\right) = \int_{\mathrm{Re}(s)}^{\mathrm{Re}(w)} \widetilde{H}'_\chi(u+it)/\widetilde{H}_\chi(u+it)\,\mathrm{d}u$.

(i) Let $b_k(t) := \frac{1}{k}\left(1 + \frac{c_3}{\mathcal{L}_Q(t)}\right)$ for some absolute constant $c_3 > 0$. By the above observation and the definitions of $\widetilde{F}_\chi(s)$, $\widetilde{H}_\chi(s)$ and $H_\chi(s)$, we see that

(7.8)
$$\begin{aligned}\left|H_\chi\left(\frac{1}{k}\right)\right| &\ll \left|\widetilde{H}_\chi\left(\frac{1}{k}\right)\right|(1-\beta_e)^{-\alpha_k(Q)} \ll |\widetilde{H}_\chi(b_k(0))|\ (1-\beta_e)^{-\alpha_k(Q)}\\ &\ll |\widetilde{F}_\chi(b_k(0))|\ (\log Q)(1-\beta_e)^{-2\alpha_k(Q)} \ll |F_1(kb_k(0))g(kb_k(0))|^{\mathrm{Re}(c_{\widehat{\chi}})}\ (\log Q)^2(1-\beta_e)^{-2\alpha_k(Q)}.\end{aligned}$$

Here in the last bound, we have noted that $|G_{\chi,1}(b_k(0))| \ll \log_2 Q$, as is evident from the fact that $\prod_{\substack{p|Q\\ W_k(p)\in U_Q}}(1 - p^{-kb_k(0)})^{-1} \ll \exp(\sum_{p|Q} 1/p) \ll \exp(\sum_{p\leq\omega(Q)} 1/p) \ll \log\omega(Q) \ll \log_2 Q$.

Now proceeding as in [45, Lemma 8], we see that for all $s$ with $\sigma > 1/k$, we have

$$\sum_{n\geq 1}\frac{\mathbb{1}_{(f(n^k),Q)=1}}{n^{ks}} = F_1(ks)\ g(ks)\ \widetilde{G}(s), \tag{7.9}$$

where

$$\widetilde{G}(s) = \prod_p\left(1 + \sum_{v\geq 2}\frac{1}{p^{vks}}(\mathbb{1}_{(f(p^{kv}),Q)=1} - \mathbb{1}_{(W_k(p),Q)=1}\ \mathbb{1}_{(f(p^{k(v-1)}),Q)=1})\right).$$

$$\prod_{\substack{p \mid Q \\ W_k(p) \in U_Q}} \left(1 - \frac{1}{p^{ks}}\right)^{-1} \cdot \exp\left( \sum_{\substack{b \in U_Q \\ W_k(b) \in U_Q}} \sum_{v \ge 2} \left( \sum_{p \equiv b \pmod{Q}} \frac{1}{v p^{vks}} - \sum_{p\colon\ p^v \equiv b \pmod{Q}} \frac{1}{v p^{vks}} \right) \right).$$

Uniformly for $s$ with $\sigma \ge 1/k$, we observe that the infinite product above has size at least $1 - \sum_{p, v \ge 2} 1/p^v \gg 1$ and at most $\exp(\sum_{p, v \ge 2} 1/p^v) \ll 1$. Likewise, the exponential factor has size $\asymp 1$ in the same region. Moreover, for $\sigma \ge 1/k$, the product over $p \mid Q$ is $\asymp |\exp(\sum_{p \mid Q\colon\ (W_k(p), Q) = 1}\ p^{-ks})|$, which is $\gg 1$ and $\ll \exp(\sum_{p \mid Q}\ p^{-1}) \ll \log_2 Q$. Putting these observations together, we find that

$$1 \ll \widetilde{G}(s) \ll \log_2 Q, \quad \text{uniformly in complex numbers } s \text{ having } \sigma \ge 1/k. \tag{7.10}$$

Applying this lower bound on $\widetilde{G}(b_k(0))$, the equality (7.9) yields

$$|F_1(kb_k(0))\ g(kb_k(0))| \ll \sum_{n \ge 1} \frac{\mathbb{1}_{(f(n^k), Q) = 1}}{n^{kb_k(0)}} \le \zeta(kb_k(0)) = \frac{1}{kb_k(0) - 1} + O(1) \ll \log Q,$$

so that from (7.8), we obtain $|H_\chi(1/k)| \ll (\log Q)^3 (1 - \beta_e)^{-2\alpha_k(Q)}$. Subpart (i) now follows as $Q \le (\log x)^{K_0}$ and as $1 - \beta_e \gg_{\epsilon_1} Q^{-\epsilon_1/20K_0} \gg_{\epsilon_1} (\log x)^{-\epsilon_1/20}$ by Siegel's Theorem.

(ii) By the general observation at the start of the proof, we have $|\widetilde{H}_\chi(s)| \ll |\widetilde{H}_\chi(1/k)| \ll |H_\chi(1/k)|(1 - \beta_e)^{-\alpha_k(Q)} \ll |H_\chi(1/k)|(\log x)^{\alpha_k(Q)\epsilon_1/20}$. The result now follows from (i).

(iii) By the same general observation, we have $|\widetilde{H}_\chi(s)| \ll |\widetilde{H}_\chi(b_k(t) + it)|$, and since $|s - \theta/k| \gg 1/\mathcal{L}_Q(t)$, we have $b_k(t) + it - \theta/k \asymp s - \theta/k$ for $\theta \in \{1, \beta_e\}$. Thus $|\widetilde{F}_\chi(s)| \ll |\widetilde{F}_\chi(b_k(t) + it)|$. (Recall that $\widetilde{H}_\chi(s) := \widetilde{F}_\chi(s) \left(s - \frac{1}{k}\right)^{\alpha_k(Q) c_{\widehat{\chi}}} \left(s - \frac{\beta_e}{k}\right)^{-\alpha_k(Q) c_{\widehat{\chi}} \gamma(\psi_e)}$.) Using (7.6) and replicating the arguments that led to (7.10), we also obtain

$$(\log_2 Q)^{-1} \ll G_{\chi,1}(s) \ll \log_2 Q, \quad \text{uniformly in complex numbers } s \text{ having } \sigma \ge 1/k. \tag{7.11}$$

Thus uniformly for $s$ as in subpart (iii) of the proposition, we have

$$|\widetilde{F}_\chi(s)| \ \ll\ |\widetilde{F}_\chi(b_k(t) + it)| \ \ll\ (\log_2 Q) \cdot |F_1(k(b_k(t) + it)) g(k(b_k(t) + it))|^{\mathrm{Re}(c_{\widehat{\chi}})}.$$

(Recall that $\widetilde{F}_\chi(s) = F_1(sk)^{c_{\widehat{\chi}}}\ g(sk)^{c_{\widehat{\chi}}}\ G_{\chi,1}(s)$.) Next by (7.9) and (7.10), we get $|\widetilde{F}_\chi(s)| \ll (\log_2 Q) \left|\sum_{n \ge 1} \mathbb{1}_{(f(n^k), Q) = 1}/n^{k(b_k(t) + it)}\right|^{\mathrm{Re}(c_{\widehat{\chi}})} \ll (\log_2 Q) \left(\sum_{n \ge 1} \mathbb{1}_{(f(n^k), Q) = 1}/n^{kb_k(t)}\right)^{\mathrm{Re}(c_{\widehat{\chi}})}$. By (7.9), (7.10) and (7.11), we get $|\widetilde{F}_\chi(s)| \ll (\log_2 Q)^2 |F_1(kb_k(t)) g(kb_k(t))|^{\mathrm{Re}(c_{\widehat{\chi}})} \ll (\log_2 Q)^3 |\widetilde{F}_\chi(b_k(t))|$.

By definitions of $b_k(t)$ and $\widetilde{H}_\chi(b_k(t))$, the last bound gives

$$|\widetilde{F}_\chi(s)| \ll (\log_3 x)^3 |\widetilde{H}_\chi(b_k(t))| \mathcal{L}_Q(t)^{\alpha_k(Q)} (1 - \beta_e)^{-\alpha_k(Q)}.$$

Finally, recall that $|t| \le T = \exp(\sqrt{\log x})$, that $1 - \beta_e \gg_{\epsilon_1} (\log x)^{-\epsilon_1/20}$, and that $|\widetilde{H}_\chi(b_k(t))| \ll |\widetilde{H}_\chi(1/k)| \ll (\log x)^{\alpha_k(Q)\epsilon_1/4}$ (by subpart (ii) the general observation at the start of the proof). This yields $|\widetilde{F}_\chi(s)| \ll (\log x)^{\alpha_k(Q)(1/2 + \epsilon_1)}$. Lemma 7.1(iii) now proves the assertion.

(iv) It suffices to show that uniformly for $s$ satisfying the same conditions as in this subpart,

$$|H_\chi(s)| + |H'_\chi(s)| \ll (\log x)^{\alpha_k(Q)\epsilon_1/5} \left(\log Q + \frac{1}{1 - \beta_e}\right). \tag{7.12}$$

(Here as usual, the second term on the right is omitted if there is no Siegel zero, otherwise it dominates.) Indeed once we establish (7.12), then from the bound $1-\beta_e \gg_{\epsilon_1} (\log x)^{-\epsilon_1/20}$, it follows that $|H_\chi(s)|+|H'_\chi(s)| \ll (\log x)^{(1/20+\alpha_k(Q)/5)\epsilon_1}$, which combined with Lemma 7.1(iii) and the observation $|H_\chi(1/k)G_{\chi,2}\,(1/k) - H_\chi(s)G_{\chi,2}(s)| = \left|\int_s^{1/k}(H_\chi(u)G_{\chi,2}(u))'\,\mathrm{d}u\right|$ completes the proof of the subpart.

To show (7.12), we recall that $H_\chi(s)$ is non-vanishing for $s$ as in the subpart. Further (7.7) applies with $z=s$ for all $s$ considered in this subpart, yielding

$$\left|\frac{H'_\chi(s)}{H_\chi(s)}\right| = \left|\frac{\widetilde{H}'_\chi(s)}{\widetilde{H}_\chi(s)} - \frac{1}{s} + \frac{\alpha_k(Q)c_{\widehat{\chi}}\gamma(\psi_e)}{s-\beta_e/k}\right| \ll \mathcal{L}_Q(0)+1+\frac{1}{1-\beta_e} \ll \log Q + \frac{1}{1-\beta_e}.$$

As a consequence,

$$\left|\log\frac{H_\chi(1/k)}{H_\chi(s)}\right| = \left|\int_s^{1/k}\frac{H'_\chi(u)}{H_\chi(u)}\,\mathrm{d}u\right| \ll \left(\frac{1}{k}-s\right)\left(\log Q + \frac{1}{1-\beta_e}\right) \ll 1,$$

showing that $|H_\chi(s)| \asymp |H_\chi(1/k)|$ uniformly for all $s$ in the statement. Collecting these bounds, we obtain for all such $s$,

$$|H_\chi(s)|+|H'_\chi(s)| \ll \left|H_\chi\left(\frac{1}{k}\right)\right| + \left|\frac{H'_\chi(s)}{H_\chi(s)}\right|\cdot\left|\frac{H_\chi(s)}{H_\chi(1/k)}\right|\cdot\left|H_\chi\left(\frac{1}{k}\right)\right| \ll \left|H_\chi\left(\frac{1}{k}\right)\right|\left(\log Q+\frac{1}{1-\beta_e}\right),$$

so that the desired bound (7.12) now follows from subpart (i). This concludes the proof. □

### 7.3. **Perron's formula and the contour shifts.**

We first show that there is some $X$ sufficiently close to $x$ for which the error term arising from an effective Perron's formula is small.

**Lemma 7.3.** *Let $h := x/\log^2 x$. There exists a positive integer $X \in (x, x+h]$ satisfying*

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

*Proof.* This would follow once we show that

$$\sum_{x<X\le x+h}\ \sum_{\substack{3X/4<n<5X/4\\ n\neq X}} \frac{\mathbb{1}_{(f(n),Q)=1}}{|\log(X/n)|} \ll x^{1/k}h\log x, \tag{7.13}$$

with the outer sum being over integers $X \in (x, x+h]$. (Recall that $x\in\mathbb{Z}^+$ in this entire section.) To show this, we write the sum on the left hand side as $S_1+S_2$, where $S_1$ denotes the contribution of the case $3X/4 < n \le X-1$. Writing any $n$ contributing to $S_1$ as $X-v$ for some integer $v\in[1, X/4)$, we see that $|\log(X/n)| = -\log(1-v/X) \gg v/X \gg v/x$. Recalling that $n = Bm$ for some $k$-free $B$ of size $O(1)$ and some $k$-full $m$, we thus have

$$S_1 \le \sum_{3x/4<n<x+h}\ \sum_{\substack{x<X\le x+h\\ n+1\le X<4n/3}} \frac{\mathbb{1}_{(f(n),Q)=1}}{|\log(X/n)|} \ll x\sum_{B\ll 1}\ \sum_{\substack{\frac{3x}{4B}<m<\frac{x+h}{B}\\ m \text{ is } k\text{-full}}}\ \sum_{\substack{1\le v<\frac{x+h}{4}\\ x<v+Bm\le x+h}} \frac{1}{v}$$

$$\ll x \sum_{1 \le v \le \frac{x+h}{4}} \frac{1}{v} \sum_{B \ll 1} \sum_{\substack{\frac{x-v}{B} < m \le \frac{x-v+h}{B} \\ m \text{ is } k\text{-full}}} 1 \ll x \log x \left( x^{1/k} \frac{h}{x} + x^{1/(k+1)} \right) \ll x^{1/k} h \log x,$$

where we have bounded the last inner sum on $m$ using the Erdös-Szekeres estimate on the count of $k$-full integers (see [14]). This shows that the sum $S_1$ is bounded by the right hand expression in (7.13). Similarly so is the sum $S_2$, thus establishing (7.13). □

To complete the proof of Theorem 5.5, it suffices to establish the bound therein for $X$ in place of $x$, for once we do so, we may simply note that

$$\Big| \sum_{x < n \le X} \chi_1(f_1(n)) \cdots \chi_K(f_K(n)) \mathbb{1}_{(f(n),q)=1} \Big| \le \sum_{x < n \le X} \mathbb{1}_{(f(n),Q)=1} \le \sum_{B \ll 1} \sum_{\substack{\frac{x}{B} < m \le \frac{X}{B} \\ m \text{ is } k\text{-full}}} 1 \ll \frac{x^{1/k}}{\log^2 x}.$$

To show the bound in Theorem 5.5 for $X$, we start by applying an effective version of Perron's formula [50, Theorem II.2.3]. To bound the resulting error, we use Lemma 7.3 and note that

$$X^{\frac{1}{k}\left(1+\frac{1}{\log X}\right)} \left( \sum_{n \le 3X/4} + \sum_{n \ge 5X/4} \right) \frac{\mathbb{1}_{(f(n),Q)=1}}{T |\log(X/n)| n^{\frac{1}{k}\left(1+\frac{1}{\log X}\right)}} \ll \frac{X^{1/k}}{T} \sum_{B \ll 1} \sum_{\substack{m \ge 1 \\ m \text{ is } k\text{-full}}} \frac{1}{m^{\frac{1}{k}\left(1+\frac{1}{\log X}\right)}}$$

$$\ll \frac{X^{1/k}}{T} \prod_p \left( 1 + \frac{1}{p^{1+1/\log X}} + O\left( \frac{1}{p^{1+1/k}} \right) \right) \ll \frac{X^{1/k}}{T} \exp\left( \sum_p \frac{1}{p^{1+1/\log X}} \right) \ll \frac{X^{1/k} \log X}{T},$$

with the last bound above being a consequence of Mertens' Theorem along with the fact that

$$\sum_{p > X} \frac{1}{p^{1+1/\log X}} \le \sum_{j \ge 0} \sum_{X^{2^j} < p \le X^{2^{j+1}}} \frac{1}{p^{1+1/\log X}} \le \sum_{j \ge 0} \exp(-2^j) \sum_{X^{2^j} < p \le X^{2^{j+1}}} \frac{1}{p} \ll 1.$$

(Recall that $T = \exp(\sqrt{\log x}) \ge \exp\left(\frac{1}{2}\sqrt{\log X}\right)$.) As such, [50, Theorem II.2.3] yields (7.14)

$$\sum_{n \le X} \chi_1(f_1(n)) \cdots \chi_K(f_K(n)) \mathbb{1}_{(f(n),Q)=1} = \frac{1}{2\pi i} \int_{\frac{1}{k}\left(1+\frac{1}{\log X}\right)-iT}^{\frac{1}{k}\left(1+\frac{1}{\log X}\right)+iT} \frac{F_\chi(s) X^s}{s} \, \mathrm{d}s + O\left( \frac{X^{1/k} \log X}{T} \right).$$

Our arguments will be divided into three possibilities:
Case 1: When $(\alpha_k(Q), c_{\widehat{\chi}}) \neq (1,1)$ and there is a Seigel zero $\beta_e$ mod $Q$.
Case 2: When $(\alpha_k(Q), c_{\widehat{\chi}}) \neq (1,1)$ and there is no Seigel zero mod $Q$.
Case 3: When $(\alpha_k(Q), c_{\widehat{\chi}}) = (1,1)$.
In Case 1, we will be assuming henceforth that $\beta_e > 1 - \frac{5c_1}{24 \log Q}$; otherwise decreasing $c_1$ reduces to Case 2. Let $\beta^* := \frac{2}{3} + \frac{\beta_e}{3}$ and $\sigma_k(t) := \frac{1}{k}\left(1 - \frac{c_1}{4\mathcal{L}_Q(t)}\right)$, so that $\frac{\beta_e}{k} > \sigma_k(0)$. Let $\delta, \delta_1 \in (0, \beta_e/10k)$ satisfy $\sigma_k(0) < \frac{\beta_e}{k} - 2\delta_1 < \frac{\beta_e}{k} + 2\delta_1 < \frac{\beta^*}{k} < \frac{1}{k} - 2\delta$. Consider the contours

- $\Gamma_2$, the horizontal segment traversed from $\frac{1}{k}\left(1 + \frac{1}{\log X}\right) + iT$ to $\sigma_k(T) + iT$.
- $\Gamma_3$, the part of the curve $\sigma_k(t) + it$ traversed from $t = T$ to $t = 0$.
- $\Gamma_4 := \Gamma_4(\delta_1)$, the segment traversed from $\sigma_k(0)$ to $\beta_e/k - \delta_1$ **above** the branch cut.

- $\Gamma_5 \coloneqq \Gamma_5(\delta_1)$, the semicircle of radius $\delta_1$ centered at $\beta_e/k$, lying in the upper half plane and traversed clockwise.
- $\Gamma_6 \coloneqq \Gamma_6(\delta_1)$, the segment traversed from $\beta_e/k + \delta_1$ to $\beta^*/k$ **above** the branch cut.
- $\Gamma_7 \coloneqq \Gamma_7(\delta)$, the segment traversed from $\beta^*/k$ to $1/k - \delta$ **above** the branch cut.
- $\Gamma_8 \coloneqq \Gamma_8(\delta)$, the circle of radius $\delta$ centered at $1/k$, traversed clockwise from the point $1/k - \delta$ above the branch cut to its reflection below the branch cut.
- $\Gamma_4^* \coloneqq \Gamma_4^*(\delta)$, the segment traversed from $\sigma_k(0)$ to $1/k - \delta$ **above** the branch cut.
- $\Gamma_5^* \coloneqq \Gamma_5^*(\delta_1)$, the circle of radius $\delta_1$ centered at $\beta_e/k$, traversed clockwise from the point $\beta_e/k - \delta_1$ above the branch cut to its reflection below the branch cut.

Here $\Gamma_5^*(\delta_1)$ is relevant only when our branch cut is along $\sigma \le \beta_e/k$ (i.e., when $\alpha_k(Q) = c_{\widehat{\chi}} = 1$ and $\beta_e$ exists), while the rest of the contours are defined irrespective of the branch cut. For a contour $\Omega$, let $-\overline{\Omega}$ denote the contour given by the complex conjugate of $\Omega$ traversed in the opposite direction and **below** the respective branch cuts. (Note that $-\overline{\Gamma_5}$ is still traversed **clockwise** but below the branch cut.) We define the contour $\Gamma_1$ by

$$\Gamma_1 \coloneqq \begin{cases} \sum_{j=2}^{8} \Gamma_j + \sum_{j=2}^{7} (-\overline{\Gamma_j}), & \text{under Case 1} \\ \Gamma_2 + \Gamma_3 + \Gamma_4^* + \Gamma_8 + (-\overline{\Gamma_4^*}) + (-\overline{\Gamma_3}) + (-\overline{\Gamma_2}), & \text{under Case 2} \\ \sum_{j=2}^{4} \Gamma_j + \Gamma_5^* + \sum_{j=2}^{4} (-\overline{\Gamma_j}), & \text{under Case 3.} \end{cases}$$

In Case 3, if $\beta_e$ doesn't exist, then there is no branch cut and $\Gamma_4$, $\overline{\Gamma}_4$ and $\Gamma_5^*$ are excluded from $\Gamma_1$. In all three cases, the integrand in (7.14) is analytic in the region enclosed by $\Gamma_1$ and the segment joining $\frac{1}{k}\left(1 + \frac{1}{\log X}\right) - iT$ and $\frac{1}{k}\left(1 + \frac{1}{\log X}\right) + iT$. (Note that if $c_{\widehat{\chi}} = 1$, the definitions of $\mathcal{Q}(k; f_1, \cdots, f_K)$ and $G_{\chi,1}$, $G_{\chi,2}$ in Lemma 7.1 give $G_{\chi,2}(1/k) = 0$, canceling the simple pole of $F_1(sk)$ at $s = 1/k$. In particular, this happens in Case 3.) So

$$\sum_{n \le X} \chi_1(f_1(n)) \cdots \chi_K(f_K(n)) \mathbb{1}_{(f(n),Q)=1} = -\frac{1}{2\pi i} \int_{\Gamma_1} \frac{F_\chi(s) X^s}{s} \, \mathrm{d}s + O\left(\frac{X^{1/k} \log X}{T}\right). \tag{7.15}$$

We now proceed to estimate the integrals occurring on the right hand side above. In the following proposition, any result about an integral is valid whenever the corresponding contour is a part of $\Gamma_1$: so for instance, the assertion on $\Gamma_8$ (resp. $\Gamma_5^*$) holds under Cases 1 or 2 (resp. Case 3), those on $\Gamma_5$ and $\Gamma_6$ hold under Case 1, and the bound involving $\Gamma_4$ holds under Cases 1 and 3. Let $I_j$ (resp. $\overline{I_j}$, $I_j^*$) denote the corresponding integral along $\Gamma_j$ (resp. $-\overline{\Gamma_j}$, $\Gamma_j^*$).

**Proposition 7.4.** *We have the following bounds:*

(i) $|I_2| + |\overline{I_2}| + |I_3| + |\overline{I_3}| \ll X^{1/k} \exp(-\kappa_0 \sqrt{\log X})$ *for some constant* $\kappa_0 \coloneqq \kappa_0(c_1, k) > 0$.

(ii) $\max\{|I_4 + \overline{I_4}|, |I_6 + \overline{I_6}|\} \ll X^{1/k} \exp(-\sqrt{\log X})$ *uniformly in* $\delta, \delta_1$ *as above.*

(iii) $\lim_{\delta_1 \to 0+} |I_5| = \lim_{\delta_1 \to 0+} |\overline{I_5}| = \lim_{\delta_1 \to 0+} |I_5^*| = \lim_{\delta \to 0+} |I_8| = 0$.

*Proof.* To show subpart (i), we use the fact that since $\beta_e > 1 - 5c_1/24 \log Q$, any $s$ lying on $\Gamma_2$, $\Gamma_3$ or their conjugates satisfies the requirements of Proposition 7.2(iii). As such, (i) follows immediately from Proposition 7.2(iii) and the fact that $|s| \gg |t| + 1$ for all $s$.

For subpart (ii), we note that for all $s \in \Gamma_4$, we have $(s-1/k)^{-\alpha_k(Q)c_{\widehat{\chi}}} = (1/k-s)^{-\alpha_k(Q)c_{\widehat{\chi}}}\ e^{-i\pi\alpha_k(Q)c_{\widehat{\chi}}}$ and $(s-\beta_e/k)^{\alpha_k(Q)c_{\widehat{\chi}}\gamma(\psi_e)} = (\beta_e/k-s)^{\alpha_k(Q)c_{\widehat{\chi}}\gamma(\psi_e)}\ e^{i\pi\alpha_k(Q)c_{\widehat{\chi}}\gamma(\psi_e)}$. (This is clear if the branch cut is along $\sigma \le 1/k$, and also if the branch cut is along $\sigma \le \beta_e/k$ which is when $(\alpha_k(Q), c_{\widehat{\chi}}) = (1,1)$.) Likewise, for all $s \in \overline{\Gamma_4}$, we have $(s-1/k)^{-\alpha_k(Q)c_{\widehat{\chi}}} = (1/k-s)^{-\alpha_k(Q)c_{\widehat{\chi}}}\ e^{i\pi\alpha_k(Q)c_{\widehat{\chi}}}$ and $(s-\beta_e/k)^{\alpha_k(Q)c_{\widehat{\chi}}\gamma(\psi_e)} = (\beta_e/k-s)^{\alpha_k(Q)c_{\widehat{\chi}}\gamma(\psi_e)}\ e^{-i\pi\alpha_k(Q)c_{\widehat{\chi}}\gamma(\psi_e)}$. Since $e^{\pm i\pi\alpha_k(Q)c_{\widehat{\chi}}(\gamma(\psi_e)-1)} \ll 1$, the definitions of $\widetilde{F}_\chi(s)$ and $\widetilde{H}_\chi(s)$ show that

$$|I_4 + \overline{I_4}| \ll \left| \int_{\sigma_k(0)}^{\beta_e/k-\delta_1} \frac{\widetilde{H}_\chi(s) G_{\chi,2}(s) X^s}{s} \left(\frac{1}{k}-s\right)^{-\alpha_k(Q)c_{\widehat{\chi}}} \left(\frac{\beta_e}{k}-s\right)^{\alpha_k(Q)c_{\widehat{\chi}}\gamma(\psi_e)} \mathrm{d}s \right|.$$

But now by Lemma 7.1(iii) and Proposition 7.2(ii), we see that

$$\begin{aligned} |I_4 + \overline{I_4}| &\ll X^{\beta_e/k} (\log X)^{\alpha_k(Q)\epsilon_1/4} (1-\beta_e)^{-\alpha_k(Q)} \int_{\sigma_k(0)}^{\beta_e/k-\delta_1} \left(\frac{\beta_e}{k}-s\right)^{\alpha_k(Q)\mathrm{Re}(c_{\widehat{\chi}}\gamma(\psi_e))} \mathrm{d}s \\ &\ll X^{\beta_e/k} (\log X)^{3\alpha_k(Q)\epsilon_1/10} \cdot \left(\frac{\beta_e}{k} - \sigma_k(0)\right)^{1+\alpha_k(Q)\mathrm{Re}(c_{\widehat{\chi}}\gamma(\psi_e))} \ll X^{1/k} \exp(-\sqrt{\log X}). \end{aligned}$$

Here we have recalled that $\beta_e \le 1 - c(\epsilon_1)/Q^{\epsilon_1/20K_0} \le 1 - c(\epsilon_1)/(\log X)^{\epsilon_1/20}$ for some constant $c(\epsilon_1) > 0$, and (as argued before Lemma 7.1) that $Q_e := \mathfrak{f}(\psi_e)$ has a prime factor $\ell_e > D+2$, which upon factoring $\psi_e = \prod_{\ell \mid Q} \psi_{e,\ell}$ with $\psi_{e,\ell}$ being a character mod $\ell$, led to
(7.16)

$$\alpha_k(Q)|\gamma(\psi_e)| \le \alpha_k(Q) \prod_{\ell \mid Q_e} \left| \frac{\sum_{v:\ vW_k(v) \in U_\ell} \overline{\psi}_{e,\ell}(v)}{\alpha_k(\ell)(\ell-1)} \right| \le \frac{1}{\ell_e - 1} \left| \sum_{\substack{v \bmod \ell_e \\ W_k(v) \equiv 0 \ (\mathrm{mod}\ \ell_e)}} \overline{\psi}_{e,\ell}(v) \right| \le \frac{D}{D+1}.$$

This shows the desired bound on $I_4$ in (ii), and the assertion for $I_6$ is entirely analogous.

Coming to subpart (iii), we parametrize the points of $\Gamma_5$ by $s = \beta_e/k + \delta_1 e^{i\theta}$ where $\pi \ge \theta \ge 0$. Since $\widetilde{M} := \sup_{|s-\frac{\beta_e}{k}| \le \frac{1}{2}(\frac{\beta_e}{k} - \sigma_k(0))} |\widetilde{H}_\chi(s)|$ is finite, we have for all sufficiently small $\delta_1 > 0$,

$$|I_5| \ll \widetilde{M} \int_0^\pi X^{\beta_e/k+\delta_1} \left(\frac{1-\beta_e}{k} - \delta_1\right)^{-\alpha_k(Q)\mathrm{Re}(c_{\widehat{\chi}})} \delta_1^{1+\alpha_k(Q)\mathrm{Re}(c_{\widehat{\chi}}\gamma(\psi_e))} \mathrm{d}\theta \ll \frac{\widetilde{M} X^{\beta_e/k+\delta_1} \delta_1^{1/(D+1)}}{\left(\frac{1-\beta_e}{k} - \delta_1\right)^{\alpha_k(Q)}},$$

where we have again seen that $1 + \alpha_k(Q)\mathrm{Re}(c_{\widehat{\chi}}\gamma(\psi_e)) \ge 1/(D+1)$ by (7.16). The last expression shows that $\lim_{\delta_1 \to 0+} |I_5| = 0$, and the assertions on $|\overline{I_5}|$ and $|I_5^*|$ are proved similarly. The same argument also shows that $|I_8| \ll M^* X^{1/k+\delta} \delta^{1-\alpha_k(Q)\mathrm{Re}(c_{\widehat{\chi}})} \left(\frac{1-\beta_e}{k} - \delta\right)^{-\alpha_k(Q)}$ for all sufficiently small $\delta > 0$, where $M^* = \sup_{|s-\frac{1}{k}| \le \frac{1-\beta^*}{k}} |\widetilde{H}_\chi(s)|$. This yields $\lim_{\delta \to 0+} |I_8| = 0$, because $\alpha_k(Q)\mathrm{Re}(c_{\widehat{\chi}}) < 1$ whenever $(\alpha_k(Q), c_{\widehat{\chi}}) \ne (1,1)$. □

Now in case 3, we let $\delta_1 \downarrow 0$ in (7.15) and invoke the relevant assertions of Proposition 7.4 to obtain $\sum_{n \le X} \chi_1(f_1(n)) \cdots \chi_K(f_K(n)) \mathbb{1}_{(f(n),Q)=1} \ll X^{1/k} \exp(-\kappa_1\sqrt{\log X})$ for some constant $\kappa_1 > 0$. Hence to complete the proof of Theorem 5.5, it suffices to assume that $(\alpha_k(Q), c_{\widehat{\chi}}) \ne (1,1)$. In case 1, we obtain, by letting $\delta \downarrow 0$ and $\delta_1 \downarrow 0$ in (7.15),

$$\sum_{n \le X} \chi_1(f_1(n)) \cdots \chi_K(f_K(n)) \mathbb{1}_{(f(n),Q)=1} = -\lim_{\delta \to 0+} \frac{I_7 + \overline{I_7}}{2\pi i} + O(X^{1/k} \exp(-\kappa_1\sqrt{\log X})). \tag{7.17}$$

By an argument analogous to that given for Proposition 7.4(ii), it is easy to see that the above limit exists. Furthermore, writing $(s-1/k)^{-\alpha_k(Q)c_{\widehat{\chi}}} = (1/k-s)^{-\alpha_k(Q)c_{\widehat{\chi}}}\ e^{\pm i\pi\alpha_k(Q)c_{\widehat{\chi}}}$ as before, we see that the limit in (7.17) is equal to

$$\frac{\sin(\pi\alpha_k(Q)c_{\widehat{\chi}})}{\pi}\int_{\beta^*/k}^{1/k} H_\chi(s)G_{\chi,2}(s)X^s\left(\frac{1}{k}-s\right)^{-\alpha_k(Q)c_{\widehat{\chi}}}\mathrm{d}s,$$

We write the above integral as $H_\chi(1/k)G_{\chi,2}(1/k)I_1 - I_2$, where $I_1 := \int_{\beta^*/k}^{1/k} X^s(1/k-s)^{-\alpha_k(Q)c_{\widehat{\chi}}}\,\mathrm{d}s$. Letting $s = 1/k - u/\log X$, and using $\beta^* = 2/3+\beta_e/3 \le 1 - c(\epsilon_1)/3(\log X)^{\epsilon_1/20}$ along with a standard bound on the tail of the integral defining a Gamma function [45, Lemma 7], we get

$$I_1 = \frac{X^{1/k}}{(\log X)^{1-\alpha_k(Q)c_{\widehat{\chi}}}}\left\{\Gamma(1-\alpha_k(Q)c_{\widehat{\chi}}) + O(\exp(-\sqrt{\log X}))\right\}.$$

Now using Proposition 7.2(iv) and making the same change of variable, we find that

$$I_2 \ll (\log X)^{\left(\frac{1}{20}+\frac{\alpha_k(Q)}{5}\right)\epsilon_1}\int_{\beta^*/k}^{1/k} X^s\left(\frac{1}{k}-s\right)^{1-\alpha_k(Q)\mathrm{Re}(c_{\widehat{\chi}})}\mathrm{d}s \ll \frac{X^{1/k}}{(\log X)^{2-\alpha_k(Q)\mathrm{Re}(c_{\widehat{\chi}})-(1/20+\alpha_k(Q)/5)\epsilon_1}}$$

as $\Gamma(2-\alpha_k(Q)\mathrm{Re}(c_{\widehat{\chi}})) \ll 1$. Collecting estimates, we obtain from (7.17),

(7.18)
$$\sum_{n\le X}\mathbb{1}_{(f(n),Q)=1}\prod_{i=1}^{K}\chi_1(f_1(n)) = \frac{H_\chi(1/k)G_{\chi,2}\,(1/k)}{\Gamma(\alpha_k(Q)c_{\widehat{\chi}})}\cdot\frac{X^{1/k}}{(\log X)^{1-\alpha_k(Q)c_{\widehat{\chi}}}}\Big(1+O(\exp(-\sqrt{\log X}))\Big) + O\left(\frac{X^{1/k}}{(\log X)^{2-\alpha_k(Q)\mathrm{Re}(c_{\widehat{\chi}})-(1/20+\alpha_k(Q)/5)\epsilon_1}}\right),$$

by the reflection formula for the Gamma function and as $\Gamma(z) \gg 1$ for all $z$ with $|z|\le 2$.

If $c_{\widehat{\chi}} \ne 1$, then $\mathrm{Re}(c_{\widehat{\chi}}) \le \cos(2\pi/\varphi(Q_0)) < 1-\epsilon_1$. Lemma 7.1(iii) and Proposition 7.2(i) yield

$$\sum_{n\le X}\mathbb{1}_{(f(n),Q)=1}\prod_{i=1}^{K}\chi_1(f_1(n)) \ll \frac{X^{1/k}}{(\log X)^{1-\alpha_k(Q)(\mathrm{Re}(c_{\widehat{\chi}})+\epsilon_1/5)}} \ll \frac{X^{1/k}}{(\log X)^{1-\alpha_k(Q)(1-\delta_0)}},$$

with $\delta_0 := \delta_0(\lambda) := \min\{3\epsilon_1/4, 1-\epsilon_1/2\}$. On the other hand, if $c_{\widehat{\chi}} = 1$, then since $q \in \mathcal{Q}(k; f_1,\cdots,f_K)$, we must have $G_{\chi,2}(1/k) = 0$ (as observed before (7.15)). Hence, (7.18) yields

$$\sum_{n\le X}\chi_1(f_1(n))\cdots\chi_K(f_K(n))\mathbb{1}_{(f(n),Q)=1} \ll \frac{X^{1/k}}{(\log X)^{2-\alpha_k(Q)-(1/20+\alpha_k(Q)/5)\epsilon_1}} \ll \frac{X^{1/k}}{(\log X)^{1-\alpha_k(Q)(1-\delta_0)}},$$

completing the proof of Theorem 5.5 in case 1.

Finally in case 2, (7.15) and Proposition 7.4 lead to the following analogue of (7.17):

$$\sum_{n\le X}\chi_1(f_1(n))\cdots\chi_K(f_K(n))\mathbb{1}_{(f(n),Q)=1} = -\lim_{\delta\to 0+}\frac{I_4^* + \overline{I_4^*}}{2\pi i} + O(X^{1/k}\exp(-\kappa_0\sqrt{\log X})). \tag{7.19}$$

An argument entirely analogous to the one given above leads to the sharper variant of (7.18) with the $\exp(-\sqrt{\log X})$ replaced by $\exp\left(-\frac{c_1\log X}{8kK_0\log_2 X}\right)$, completing the proof of Theorem 5.5.

This finally concludes the proof of Theorem 4.2. In order to establish Theorems 2.1 to 2.3, we thus need to appropriately bound the contributions of inconvenient $n$'s considered in the respective theorems. We take this up in the next several sections.

## 8. Equidistribution to restricted moduli: Proof of Theorem 2.1

By Theorem 4.2, it remains to show that

$$\sum_{\substack{n \le x \text{ inconvenient} \\ (\forall i)\ f_i(n) \equiv a_i \pmod q}} 1 \;=\; o\left(\frac{1}{\varphi(q)^K} \sum_{\substack{n \le x \\ (f(n), q) = 1}} 1\right) \quad \text{as } x \to \infty, \tag{8.1}$$

uniformly in coprime residues $(a_i)_{i=1}^K$ to $k$-admissible moduli $q \le (\log x)^{K_0}$, under any one of the conditions (i)-(iii) of Theorem 2.1.

To show this, we set $z := x^{1/\log_2 x}$ and recall that, by (4.3), (3.3) and (3.1), the $n$'s that are either $z$-smooth or divisible by the $(k+1)$-th power of a prime exceeding $y$ give negligible contribution to the left hand side of (8.1) in comparison to the right hand side. The remaining $n$ can be written in the form $mP^k$, where $P := P(n) > z$, $P_{Jk}(m) \le y$, $m$ is not divisible by the $(k+1)$-th power of a prime exceeding $y$, and $\gcd(m, P) = 1$, so that $f_i(n) = f_i(m) W_{i,k}(P)$. Given $m$, the number of possible $P$ is, by the Brun-Titchmarsh inequality,

$$\ll \frac{V''_{1,q}}{\varphi(q)} \cdot \frac{(x/m)^{1/k}}{\log(z/q)} \ll \frac{V''_{1,q}}{\varphi(q)} \cdot \frac{x^{1/k} \log_2 x}{m^{1/k} \log x},$$

where $V''_{1,q} := \max\left\{\#\mathcal{V}^{(k)}_{1,K}\big(q; (w_i)_{i=1}^K\big) :\ (w_i)_{i=1}^K \in U_q^K\right\}$. Summing this over possible $m$, we get

$$\sum_{\substack{n \le x \text{ inconvenient} \\ P(n) > z;\ p > y \implies p^{k+1} \nmid n \\ (\forall i)\ f_i(n) \equiv a_i \pmod q}} 1 \;\ll\; \frac{V''_{1,q}}{\varphi(q)} \cdot \frac{x^{1/k}}{(\log x)^{1 - \alpha_k \epsilon/2}} \exp\left(O((\log_3 x)^2 + (\log_2(3q))^{O(1)})\right)$$

via (4.5). By Proposition 3.1, the quantity on the right hand side above is negligible compared to the right hand side of (8.1) whenever $q^{K-1} V''_{1,q} \ll (\log x)^{(1 - 2\epsilon/3)\alpha_k}$. But this does hold under any one of conditions (i)-(iii) in the statement of Theorem 2.1, because:

(i) $V''_{1,q} \ll 1$ if at least of one of $\{W_{i,k}\}_{1 \le i \le K}$ is linear.

(ii) $V''_{1,q} \ll D_{\min}^{\omega(q)}$ if $q$ is squarefree, since $\#\mathcal{V}^{(k)}_{1,K}(\ell; (w_i)_{i=1}^K) \le D_{\min}$ for all $\ell \gg 1$.

(iii) $V''_{1,q} \ll q^{1 - 1/D_{\min}}$ by work of Konyagin [19, 20]

This establishes (8.1), completing the proof of Theorem 2.1. □

### 8.1. Optimality in the ranges of $q$ in Theorem 2.1.

In all our examples below, $\{W_{i,k}\}_{i=1}^K \subset \mathbb{Z}[T]$ will be nonconstant with $\prod_{i=1}^K W_{i,k}$ separable over $\mathbb{Q}$. Then $\beta(W_{1,k}, \dots, W_{K,k}) = 1$, guaranteeing that any integer satisfies $IFH(W_{1,k}, \dots, W_{K,k}; 1)$. We claim that there exists a constant $\widetilde{C} := \widetilde{C}(W_{1,k}, \dots, W_{K,k})$ such that for *any* multiplicative functions $(f_1, \dots, f_K)$ satisfying $f_i(p^k) = W_{i,k}(p)$ for all primes $p$ and all $i \in [K]$, any $\widetilde{C}$-rough $k$-admissible integer $q$ lies in $\mathcal{Q}(k; f_1, \cdots, f_K)$; in other words, $(f_1, \dots, f_K)$ are jointly WUD modulo any fixed $\widetilde{C}$-rough $k$-admissible integer $q$. Indeed, viewing a character of $U_q^K$ as a tuple of characters mod

$q$,[10] the condition (2.1) becomes vacuously true whenever $\mathcal{T}_k(q) \coloneqq \{(W_{1,k}(u), \cdots, W_{K,k}(u)) \in U_q^K : u \in U_q\}$ generates the group $U_q^K$. Now under the canonical isomorphism $U_q^K \to \prod_{\ell^e \| q} U_{\ell^e}^K$, the set $\mathcal{T}_k(q)$ maps to $\prod_{\ell^e \| q} \mathcal{T}_k(\ell^e)$. Thus by [31, Lemma 5.13], if $\mathcal{T}_k(q)$ does not generate $U_q^K$, then there is some $\ell^e \parallel q$ and some tuple of characters $(\psi_1, \cdots, \psi_K) \neq (\chi_{0,\ell}, \ldots, \chi_{0,\ell})$ mod $\ell^e$ for which $\prod_{i=1}^K \psi_i(W_{i,k}(u))$ is constant on the set $R_k(\ell^e)$. Our claim now follows from [29, Lemma 5].

Fix any $k \in \mathbb{N}$. Let $\widetilde{C}_0 > \max\{\widetilde{C}, 4KD\}$ be any constant depending only on the polynomials $\{W_{i,k}\}_{1 \le i \le K}$, which also exceeds the size of the leading coefficient and (nonzero) discriminant of $\prod_{i=1}^K W_{i,k}$. Then by Theorem N, $f_1, \ldots, f_K$ are jointly weakly equidistributed modulo any (fixed) $\widetilde{C}_0$-rough $k$-admissible integer. Fix a prime $\ell_0 > \widetilde{C}_0$, and consider any nonconstant polynomials $\{W_{i,v}\}_{\substack{1 \le i \le K \\ 1 \le v \le k-1}} \subset \mathbb{Z}[T]$ all of whose coefficients are divisible by $\ell_0$, so that $\alpha_v(\ell_0) = 0$ for each $v < k$. Our moduli $q$ will have $P^-(q) = \ell_0$, so that $\alpha_v(q) = 0$ for all $v < k$. In each example below, we will show that $\alpha_k(q) \neq 0$, so that $q$ is $k$-admissible and lies in $\mathcal{Q}(k; f_1, \cdots, f_K)$ by definition of $\widetilde{C}_0$. The constant $K_0$ (in the assumption $q \le (\log x)^{K_0}$) is taken large enough in terms of $\{W_{i,k}\}_{i=1}^K$.

**Optimality under condition (i)**. We show that for any $K \ge 2$, the range of $q$ in Theorem 2.1(i) is optimal, – even if *all* of $W_{1,k}, \ldots, W_{K,k}$ are assumed to be linear, for *any* choice of (pairwise coprime) linear functions. Indeed, consider $W_{i,k}(T) \coloneqq c_i T + b_i \in \mathbb{Z}[T]$ for nonzero integers $c_i$ and integers $b_i$ satisfying $b_i/c_i \neq b_j/c_j$ for all $i \neq j$. Then $\prod_{i=1}^K W_{i,k}$ is clearly separable in $\mathbb{Q}[T]$. Choose a nonzero integer $b$ such that $\prod_{i=1}^K (c_i b + b_i) \neq 0$. Let $\widetilde{C}_0 > \max\{|b|, |c_i b + b_i| : 1 \le i \le K\}$ be any constant satisfying the aforementioned requirements, so that any $q$ with $P^-(q) = \ell_0 > \widetilde{C}_0$ is coprime to $b$ and to $\prod_{i=1}^K W_{i,k}(b) = \prod_{i=1}^K (c_i b + b_i)$. Thus $\alpha_k(q) \neq 0$ and $q \in \mathcal{Q}(k; f_1, \cdots, f_K)$. Now any prime $P \le x^{1/k}$ satisfying $P \equiv b \pmod{q}$ also satisfies $f_i(P^k) = W_{i,k}(P) \equiv c_i b + b_i \pmod{q}$ for all $i \in [K]$. The Siegel–Walfisz Theorem thus shows that there are $\gg x^{1/k}/\varphi(q) \log x$ many $n \le x$ satisfying $f_i(n) \equiv c_i b + b_i \pmod{q}$ for all $i \in [K]$. By Proposition 3.1, this last expression grows strictly faster than $\varphi(q)^{-K} \#\{n \le x : (f(n), q) = 1\}$ as soon as $q \ge (\log x)^{(1+\epsilon)\alpha_k/(K-1)}$ for any fixed $\epsilon \in (0,1)$, showing that the range of $q$ in Theorem 2.1 under condition (i) is essentially optimal. Note that with $Y \in [2(1+\epsilon)\log_2 x/(K-1), (K_0/2)\log_2 x]$, the squarefree integer $q \coloneqq \prod_{\ell_0 \le \ell \le Y} \ell$ satisfies all desired conditions; in particular $(\log x)^{(1+\epsilon)/(K-1)} \le q \le (\log x)^{K_0}$ and $P^-(q) = \ell_0$.

**Optimality under condition (ii).** To show that the range of squarefree $q$ in Theorem 2.1(ii) is optimal, we define $W_{i,k}(T) \coloneqq \prod_{1 \le j \le d}(T - 2j) + 2(2i-1) \in \mathbb{Z}[T]$ for some fixed $d > 1$. Eisenstein's criterion at the prime 2 shows that each $W_{i,k}$ is irreducible in $\mathbb{Q}[T]$, and the distinct $W_{i,k}$'s differ by a constant, making $\prod_{i=1}^K W_{i,k}$ separable over $\mathbb{Q}$. Now $2 \in U_q$, and $W_{i,k}(2) = 2(2i-1) \le 2(2K-1) < 4KD < \widetilde{C}_0 < P^-(q)$ for each $i \in [K]$. Thus, $q \in \mathcal{Q}(k; f_1, \cdots, f_K)$ and $(2(2i-1))_{i=1}^K \in U_q^K$. Further, any prime $P$ satisfying $\prod_{1 \le j \le d}(P - 2j) \equiv 0 \pmod{q}$ also satisfies $f_i(P^k) = W_{i,k}(P) \equiv 2(2i-1) \pmod{q}$ for each $i$. Since $2d = 2\deg W_{i,k} < 4KD < P^-(q)$, we see that $2, 4, \ldots, 2d$ are all distinct coprime residues modulo each prime dividing $q$, whereupon it follows that the congruence $\prod_{1 \le j \le d}(v - 2j) \equiv 0 \pmod{q}$ has exactly

[10] Here $U_q^K$ is the direct product of $U_q$ taken $K$ times.

$d^{\omega(q)}$ distinct solutions $v \in U_q$ for squarefree $q$. Hence, there are $\gg \frac{d^{\omega(q)}}{\varphi(q)} \cdot \frac{x^{1/k}}{\log x}$ many primes $P \le x^{1/k}$ satisfying $f_i(P^k) \equiv 2(2i-1) \pmod q$ for all $i$, so there are also at least as many $n \le x$ for which all $f_i(n) \equiv 2(2i-1) \pmod q$. The last expression grows strictly faster than $\varphi(q)^{-K}\#\{n \le x : (f(n), q) = 1\}$ as soon as $q^{K-1}D_{\min}^{\omega(q)} = q^{K-1}d^{\omega(q)} > (\log x)^{(1+\epsilon)\alpha_k}$ for any fixed $\epsilon > 0$, showing that the range of $q$ in Theorem 2.1(ii) is essentially optimal.

Note that it is possible to construct squarefree $q \le (\log x)^{K_0}$ satisfying the much stronger requirement that $d^{\omega(q)} > (\log x)^{(1+\epsilon)\alpha_k}$ (and $P^-(q) = \ell_0$). Indeed, let $q := \prod_{\ell_0 \le \ell \le Y} \ell$ for some $Y \le (K_0/2)\log_2 x$. Then $\omega(q) = \sum_{\ell_0 \le \ell \le Y} 1 \ge Y/2\log Y$, while by the Chinese Remainder Theorem and the Prime Ideal Theorem, $\alpha_k(q) \le \kappa'/\log Y$ for some constant $\kappa' := \kappa'(W_{1,k}, \dots, W_{K,k}; \ell_0)$. So we need only choose $Y \in (4\kappa' \log_2 x/\log d, (K_0/2)\log_2 x)$ to have $q \le (\log x)^{K_0}$ and $d^{\omega(q)} > (\log x)^{(1+\epsilon)\alpha_k}$.

For future reference, we observe that any $n$ of the form $P^k$ with $P$ a prime exceeding $q$ satisfies $P_k(n) > q$. Hence in the above setting, we have shown the stronger lower bound

$$\sum_{\substack{n \le x:\ P_k(n) > q \\ (\forall i)\ f_i(n) \equiv 2(2i-1) \pmod q}} 1 \ \ge \sum_{\substack{q < P \le x^{1/k} \\ \prod_{1 \le j \le d}(P-2j) \equiv 0 \pmod q}} 1 \ \gg \frac{d^{\omega(q)}}{\varphi(q)} \cdot \frac{x^{1/k}}{\log x}. \tag{8.2}$$

**Optimality under condition (iii).** Fix $d > 1$ and define $W_{i,k}(T) := (T-1)^d + i \in \mathbb{Z}[T]$, so that $\prod_{i=1}^K W_{i,k}(T+1) = \prod_{i=1}^K (T^d + i)$ is clearly separable in $\mathbb{Q}[T]$, hence so is $\prod_{i=1}^K W_{i,k}(T)$. Let $q := Q^d$ for some $Q \le (\log x)^{K_0/d}$ satisfying $P^-(Q) = \ell_0$. Then $1 \in R_k(q)$, showing that $q \in \mathcal{Q}(k; f_1, \cdots, f_K)$. Moreover, $i \in U_q$ for each $i \in [K]$, and any prime $P \equiv 1 \pmod Q$ satisfies $f_i(P^k) = W_{i,k}(P) = (P-1)^d + i \equiv i \pmod q$. Consequently, there are $\gg x^{1/k}\big/q^{1/d}\log x$ many $n \le x$ satisfying $f_i(n) \equiv i \pmod q$ for all $i$, and this last expression grows strictly faster than $\varphi(q)^{-K}\#\{n \le x : (f(n), q) = 1\}$ as soon as $q^{K-1/D_{\min}} = q^{K-1/d} \ge (\log x)^{(1+\epsilon)\alpha_k}$ for some fixed $\epsilon \in (0,1)$. This establishes that the range of $q$ in condition (iii) of Theorem 2.1 is optimal, and concrete examples of moduli $q$ satisfying the conditions imposed so far, are those of the form $Q^d$, with $Q$ lying in $[(\log x)^{(1+\epsilon)(K-1/d)^{-1}/d}, (\log x)^{K_0/d}]$ and having least prime factor $\ell_0$.

## 9. Restricted inputs to general moduli: Proof of Theorem 2.2

Fix $T \in \mathbb{N}_{>1}$. By Proposition 4.1 and the fact that $P_{Jk}(n) \le P_T(n)$, it is immediate that

$$\sum_{\substack{n \le x:\ P_T(n) \le q \\ \gcd(f(n), q) = 1}} 1 = o\Bigg(\sum_{\substack{n \le x \\ \gcd(f(n), q) = 1}} 1\Bigg). \tag{9.1}$$

In Theorems 2.2 and 2.3, we may assume $q$ to be sufficiently large, for otherwise these results follow directly from Theorem N and (9.1). The latter formula also show the equality of the second and third expressions in (2.2), so it remains to show the first equality in either. Recall that for this theorem, we have $\epsilon := 1$ and $y = \exp(\sqrt{\log x})$ in the framework developed in section 4. Now any convenient $n$ has $P_{Jk}(n) > y$ and hence is counted in the left hand side of (2.2). By Theorem 4.2, it suffices to show that the contributions of the inconvenient $n$ to the left hand sides of (2.2) are negligible compared to $\varphi(q)^{-K}\#\{n \le x : (f(n), q) = 1\}$. In fact, by (4.3) and (3.3), it remains to show the bound (9.2)below to establish the theorem:

$$\textstyle\sum^{*}_{n:\ P_R(n)>q} 1 \ll \frac{x^{1/k}}{\varphi(q)^K(\log x)^{1-2\alpha_k/3}}. \tag{9.2}$$

Here and in the rest of the manuscript, any sum of the form $\sum^*_n$ denotes a sum over positive integers $n \le x$ that are not $z$-smooth, not divisible by the $(k+1)$-th power of a prime exceeding $y$, have $P_{Jk}(n) \le y$ and satisfy $f_i(n) \equiv a_i \pmod q$ for all $i \in [K]$. Other conditions imposed on this sum are additional to these.

Defining $\omega_{\|}(n) := \#\{p > q : p^k \parallel n\}$ and $\omega^*(n) := \#\{p > q : p^{k+1} \mid n\}$, we first show the following three bounds:

$$\textstyle\sum^{*}_{n:\ \omega_{\|}(n)\ge KD+1} 1,\ \sum^{*}_{\substack{n:\ \omega_{\|}(n)=KD\\ \omega^*(n)\ge 1}} 1,\ \sum_{\substack{n\le x:\ (f(n),q)=1\\ \omega^*(n)\ge Kk,\ P_{Jk}(n)\le y,\ P(n)>z\\ p>y\implies p^{k+1}\nmid n}} 1 \ll \frac{x^{1/k}}{\varphi(q)^K(\log x)^{1-2\alpha_k/3}}. \tag{9.3}$$

Any $n$ counted in the first sum is of the form $m(P_{KD+1}\cdots P_1)^k$, where $P_{Jk}(m) \le y$, where $P_1, \dots, P_{KD+1}$ are primes exceeding $q$ satisfying $P_1 := P(n) > z$ and $q < P_{KD+1} < \cdots < P_1$, and where $f_i(n) = f_i(m)\prod_{j=1}^{KD+1} f_i(P_j^k) = f_i(m)\prod_{j=1}^{KD+1} W_{i,k}(P_j)$. The conditions $f_i(n) \equiv a_i \pmod q$ can be rewritten as $(P_1, \dots, P_{KD+1}) \bmod q \in \mathcal{V}^{(k)}_{KD+1,K}\big(q; (a_i f_i(m)^{-1})_{i=1}^K\big)$. Given $m$, $(v_1, \dots, v_{KD+1}) \in \mathcal{V}^{(k)}_{KD+1,K}\big(q; (a_i f_i(m)^{-1})_{i=1}^K\big)$, and $P_2, \dots, P_{KD+1}$, the number of $P_1$ in $(q, x^{1/k}/m^{1/k}P_2\cdots P_{KD+1}]$ satisfying $P_1 \equiv v_1 \pmod q$ is $\ll x^{1/k}\log_2 x/m^{1/k}P_2\cdots P_{KD+1}\varphi(q)\log x$, by Brun-Titchmarsh. We sum this over all possible $P_2, \dots, P_{KD+1}$, making use of the bound $\sum_{\substack{q<p\le x\\ p\equiv v \pmod q}} 1/p \ll \log_2 x/\varphi(q)$ uniformly in $v \in U_q$ (this follows from Brun–Titchmarsh and partial summation). We deduce that the number of possible $(P_1, \dots, P_{KD+1})$ satisfying $P_j \equiv v_j \pmod q$ for each $j \in [KD+1]$ is no more than

$$\sum_{\substack{q<P_{KD+1}<\cdots<P_2\le x\\ (\forall j)\ P_j\equiv v_j \pmod q}} \ \sum_{\substack{z<P_1\le x^{1/k}/m^{1/k}P_2\cdots P_{KD+1}\\ P_1\equiv v_1 \pmod q}} 1 \ll \frac{1}{\varphi(q)^{KD+1}}\cdot\frac{x^{1/k}(\log_2 x)^{O(1)}}{m^{1/k}\log x}. \tag{9.4}$$

Define $V'_{r,K} := \max\Big\{\#\mathcal{V}^{(k)}_{r,K}\big(q; (w_i)_{i=1}^K\big) : w_1, \dots, w_K \in U_q\Big\}$. Summing (9.4) over all $(v_1, \dots, v_{KD+1}) \in \mathcal{V}^{(k)}_{KD+1,K}\big(q; (a_i f_i(m)^{-1})_{i=1}^K\big)$ and then over all $m$ via (4.5) shows that

$$\textstyle\sum^{*}_{n:\ \omega_{\|}(n)\ge KD+1} 1 \ll \frac{V'_{KD+1,K}}{\varphi(q)^{KD+1}}\cdot\frac{x^{1/k}}{(\log x)^{1-\alpha_k/2}}\cdot\exp\left(O\big((\log_3 x)^2 + (\log_2(3q))^{O(1)}\big)\right). \tag{9.5}$$

Applying (4.9) with $N := KD+1$, we get $V'_{KD+1,K}/\varphi(q)^{KD+1} \ll \varphi(q)^{-K}\prod_{\ell\mid q}(1+O(\ell^{-1/D})) \ll \varphi(q)^{-K}\exp\big(O((\log q)^{1-1/D})\big)$. This yields the first bound in (9.3).

Next, any $n$ counted in the second sum in (9.3) can be written in the form $mp^c(P_{KD}\cdots P_1)^k$ for some $m, c$ and distinct primes $p, P_1, \dots, P_{KD}$ exceeding $q$, which satisfy the conditions $P_1 = P(n) > z$, $q < P_{KD} < \cdots < P_1$, $P_{Jk}(m) \le y$, $c \ge k+1$ and $f_i(n) = f_i(m)f_i(p^c)\prod_{j=1}^{KD} W_{i,k}(P_j)$, so that $(P_1, \dots, P_{KD}) \bmod q \in \mathcal{V}^{(k)}_{KD,K}\big(q; (a_i f_i(mp^c)^{-1})_{i=1}^K\big)$. Given $m, p, c$ and $(v_1, \dots, v_{KD}) \in \mathcal{V}^{(k)}_{KD,K}\big(q; (a_i f_i(mp^c)^{-1})_{i=1}^K\big)$, the arguments leading to (9.4) show that the number of possible

$(P_1, \dots, P_{KD})$ satisfying $(P_j)_{i=1}^{KD} \equiv (v_j)_{i=1}^{KD} \pmod q$ is $\ll x^{1/k}(\log_2 x)^{O(1)} \Big/ \varphi(q)^{KD} m^{1/k} p^{c/k} \log x$. Summing this successively over all $(v_1, \dots, v_{KD})$, $c \geq k+1$, $p > q$ and all possible $m$, shows that the second of the three sums in (9.3) is $\ll \frac{V'_{KD,K}}{q^{1/k}\varphi(q)^{KD}} \cdot \frac{x^{1/k}}{(\log x)^{1-2\alpha_k/3}}$. (Here we have noted that $\sum_{p>q,\ c\geq k+1} p^{-c/k} \ll \sum_{p>q} p^{-1-1/k} \ll q^{-1/k}$.) By (4.10), we have $V'_{KD,K} \Big/ q^{1/k}\varphi(q)^{KD} \ll 1/q^K$, proving the second inequality in (9.3).

Lastly, any $n$ counted in the third sum in (9.3) still has $P(n) > z$ and $P(n)^k \parallel q$, and thus can be written in the form $mp_1^{c_1} \cdots p_{Kk}^{c_{Kk}} P^k$ for some distinct primes $p_1, \dots, p_{Kk}, P$ exceeding $q$ and some integers $m, c_1, \dots, c_{Kk}$, which satisfy $P = P(n) > z$, $P_{Jk}(m) \leq y$, $c_j \geq k+1$ for all $j \in [Kk]$, and $\gcd(f(m), q) = 1$. Given $m, p_1, \dots, p_{Kk}, c_1, \dots, c_{Kk}$, the number of possible $P > z$ satisfying $P^k \leq x/mp_1^{c_1} \cdots p_{Kk}^{c_{Kk}}$ is $\ll x^{1/k} \big/ (mp_1^{c_1} \cdots p_{Kk}^{c_{Kk}})^{1/k} \log z$. Summing this over all $c_1, \dots, c_{Kk} \geq k+1$, and then over all $p_1, \dots, p_{Kk}, m$, shows the third bound in (9.3).

In the rest of the argument, $R$ as in the statement of the theorem is the least integer exceeding

$$\max\left\{k(KD+1)-1, k\left(1+(k+1)\left(K-\frac{1}{D}\right)\right)\right\} = \begin{cases} k(KD+1)-1, & \text{if } k < D \\ k\left(1+(k+1)\left(K-1/D\right)\right) & \text{if } k \geq D. \end{cases}$$

Since $q$ is sufficiently large, the $q$-rough part of any $n$ satisfying $\gcd(f(n), q) = 1$ is $k$-full (by Lemma 3.3). As such, any $n$ with $\omega^*(n) = 0$ counted in (9.2) must have $\omega_\parallel(n) \geq \lfloor R/k \rfloor \geq KD+1$, and hence is counted in the first sum in (9.3). Moreover, any $n$ with $\omega_\parallel(n) = KD$ counted in (9.2) must also have $\omega^*(n) \geq R - k\omega_\parallel(n) \geq k(KD+1) - kKD \geq 1$, and hence is counted in the second sum in (9.3). By (9.3), it thus remains to show that the contribution of $n$ having $\omega_\parallel(n) \in [KD-1]$ and $\omega^*(n) \in [Kk-1]$ to the left hand side of (9.2) is absorbed in the right hand side. This would follow once we show that for any fixed $r \in [KD-1]$ and $s \in [Kk-1]$, the contribution $\Sigma_{r,s}$ of all $n$ with $\omega_\parallel(n) = r$ and $\omega^*(n) = s$ to the left hand side of (9.2) is absorbed in the right hand side.

Now any $n$ counted in $\Sigma_{r,s}$ is of the form $mp_1^{c_1} \cdots p_s^{c_s} P_1^k \cdots P_r^k$ for some distinct primes $p_1, \dots, p_s$, $P_1, \dots, P_r$ and integers $m, c_1, \dots, c_s$, which satisfy the following conditions: **(i)** $P(m) \leq q$; **(ii)** $P_1 := P(n) > z$; $q < P_r < \cdots < P_1$; **(iii)** $p_1, \dots, p_s > q$; **(iv)** $c_1, \dots, c_s \geq k+1$ and $c_1 + \cdots + c_s \geq R - kr$; **(v)** $m$, $p_1, \dots, p_s, P_1, \dots, P_r$ are all pairwise coprime, so that $f_i(n) = f_i(m) f(p_1^{c_1}) \cdots f(p_s^{c_s}) \prod_{j=1}^r W_{i,k}(P_j)$ for each $i \in [K]$. Here, property (i) holds because the $q$-rough part of any $n$ satisfying $\gcd(f(n), q) = 1$ is $k$-full, whereas $\omega_\parallel(n) = r$, $\omega^*(n) = s$ .

With $\tau_i := \min\{c_i, R - kr\}$, it is easy to see that the integers $\tau_1, \dots, \tau_s \in [k+1, R-kr]$ satisfy $\tau_1 \leq c_1, \dots, \tau_s \leq c_s$ and $\tau_1 + \cdots + \tau_s \geq R - kr$. (Here it is important that $R \geq k(KD+1)$, $r \leq KD - 1$ and $c_1 + \cdots + c_s \geq R - kr$.) Turning this around, we find that

$$\Sigma_{r,s} \leq \sum_{\substack{\tau_1, \dots, \tau_s \in [k+1, R-kr] \\ \tau_1 + \cdots + \tau_s \geq R - kr}} \mathcal{N}_{r,s}(\tau_1, \dots, \tau_s), \tag{9.6}$$

where $\mathcal{N}_{r,s}(\tau_1, \dots, \tau_s)$ denotes the contribution of all $n$ counted in (9.2) which can be written in the form $mp_1^{c_1} \cdots p_s^{c_s} P_1^k \cdots P_r^k$ for some distinct primes $p_1, \dots, p_s, P_1, \cdots, P_r$ and integers $m, c_1, \dots, c_s$ satisfying the conditions (i)-(v) above, along with the condition $c_1 \geq \tau_1, \dots, c_s \geq$

$\tau_s$. We will show that for each tuple $(\tau_1, \dots, \tau_s)$ occurring in (9.6), we have

$$\mathcal{N}_{r,s}(\tau_1, \dots, \tau_s) \ll \frac{x^{1/k}(\log_2 x)^{O(1)}}{q^K \log x}. \tag{9.7}$$

Consider an arbitrary such tuple $(\tau_1, \dots, \tau_s)$, and write $n$ in the form $mp_1^{c_1} \cdots p_s^{c_s} P_1^k \cdots P_r^k$ as above. The conditions $f_i(n) \equiv a_i \pmod q$ lead to $(P_1, \dots, P_r) \bmod q \in \mathcal{V}_{r,K}^{(k)}\big(q; (a_i f_i(mp_1^{c_1} \cdots p_s^{c_s})^{-1})_{i=1}^K\big)$. Given $m, p_1, \dots, p_s, c_1, \dots, c_s$ and $(v_1, \dots, v_r) \in \mathcal{V}_{r,K}^{(k)}\big(q; (a_i f_i(mp_1^{c_1} \cdots p_s^{c_s})^{-1})_{i=1}^K\big)$, the arguments leading to (9.4) show that the number of possible $P_1, \dots, P_r$ satisfying $P_j \equiv v_j \bmod q$ for each $j \in [r]$, is $\ll x^{1/k}(\log_2 x)^{O(1)} \Big/ \varphi(q)^r m^{1/k} p_1^{c_1/k} \cdots p_s^{c_s/k} \log x$. With $V'_{r,K} = \max_{(w_i)_i \in U_q^K} \#\mathcal{V}_{r,K}^{(k)}\big(q; (w_i)_{i=1}^K\big)$ as before, the bounds $\sum_{p_i > q:\ c_i \ge \tau_i} p_i^{-c_i/k} \ll q^{-(\tau_i/k-1)}$ yield

$$\mathcal{N}_{r,s}(\tau_1, \dots, \tau_s) \ll \frac{1}{q^{(\tau_1 + \dots + \tau_s)/k - s}} \frac{V'_{r,K}}{\varphi(q)^r} \cdot \frac{x^{1/k}(\log_2 x)^{O(1)}}{\log x} \sum_{\substack{m \le x:\ P(m) \le q \\ \gcd(f(m), q) = 1}} \frac{1}{m^{1/k}}. \tag{9.8}$$

Proceeding as in the argument for (4.5), we write any $m$ in the above sum as $BM$ where $B$ is $k$-free and $M$ is $k$-full, so that $B = O(1)$ and $P(M) \le q$. We find that

$$\sum_{\substack{m \le x:\ P(m) \le q \\ \gcd(f(m), q) = 1}} \frac{1}{m^{1/k}} \ll \sum_{\substack{M \le x:\ P(M) \le q \\ M \text{ is } k\text{-full}}} \frac{1}{M^{1/k}} \le \prod_{p \le q} \left(1 + \frac{1}{p} + O\left(\frac{1}{p^{1+1/k}}\right)\right) \ll \exp\left(\sum_{p \le q} \frac{1}{p}\right) \ll \log q. \tag{9.9}$$

Inserting this into (9.8), we obtain

$$\mathcal{N}_{r,s}(\tau_1, \dots, \tau_s) \ll \frac{1}{q^{(\tau_1 + \dots + \tau_s)/k - s}} \frac{V'_{r,K}}{\varphi(q)^r} \cdot \frac{x^{1/k}(\log_2 x)^{O(1)}}{\log x}. \tag{9.10}$$

Now since $1 \le r \le KD - 1$, an application of (4.10) with $N := r$ now yields

$$\mathcal{N}_{r,s}(\tau_1, \dots, \tau_s) \ll \frac{\exp\big(O(\omega(q))\big)}{q^{(\tau_1 + \dots + \tau_s)/k - s + r/D}} \cdot \frac{x^{1/k}(\log_2 x)^{O(1)}}{\log x} \ll \frac{\exp\big(O(\omega(q))\big)}{q^{\max\{s/k, R/k - r - s\} + r/D}} \cdot \frac{x^{1/k}(\log_2 x)^{O(1)}}{\log x}, \tag{9.11}$$

where in the last equality we have recalled that $\tau_1, \dots, \tau_s \ge k + 1$ and $\tau_1 + \dots + \tau_s \ge R - kr$. We claim that $\max\{s/k, R/k - r - s\} + r/D > K$. This is tautological if $s/k + r/D > K$, so suppose $s/k + r/D \le K$. Then $r \le D(K - s/k) \le DK - D/k$, and $s \le k(K - r/D)$ so that $R/k - r - s + r/D \ge R/k - Kk + ((k+1)/D - 1)r$. If $k < D$, then $(k+1)/D - 1 \le 0$, so for all $1 \le r \le DK - D/k$, we have $R/k - Kk + ((k+1)/D - 1)r \ge R/k - Kk + ((k+1)/D - 1)(DK - D/k)$ and this exceeds $K$ since $R \ge k(KD + 1)$. If on the other hand, we had $k \ge D$, then $k + 1 > D$ and the minimum value of $R/k - Kk + ((k+1)/D - 1)r$ is attained at $r = 1$, giving us $R/k - Kk + ((k+1)/D - 1)r \ge R/k - Kk + ((k+1)/D - 1)$ which also exceeds $K$ since $R > k\big(1 + (1 + k)(K - 1/D)\big)$. This shows our claim, so that (9.11) leads to (9.7). Summing (9.7) over the $O(1)$ many possible tuples $(\tau_1, \dots, \tau_s)$ occurring in the right hand side of (9.6) yields $\Sigma_{r,s} \ll x^{1/k}(\log_2 x)^{O(1)} \big/ q^K \log x$, which (as argued before) establishes Theorem 2.2.

## 10. Final preparatory step for Theorem 2.3: Counting points on varieties

To establish Theorem 2.3, we will need the following partial improvements of Corollary 5.4. In this section, we again deviate from the general notation set up for Theorems 2.1 to 2.3, so the notation set up in this section will be relevant in this section only.

**Proposition 10.1.** *Let $F \in \mathbb{Z}[T]$ be a fixed nonconstant polynomial which is not squarefull.*

**(a)** *Define $\mathcal{V}_{2,1}(\ell; w) := \{(v_1, v_2) \in U_\ell^2 : F(v_1)F(v_2) \equiv w \pmod \ell\}$. Then $\#\mathcal{V}_{2,1}(\ell; w) \le \varphi(\ell)\left(1 + O\left(\ell^{-1/2}\right)\right)$, uniformly for primes $\ell$ and coprime residues $w$ mod $\ell$.*

**(b)** *Let $G \in \mathbb{Z}[T]$ be any fixed polynomial such that $\{F, G\} \subset \mathbb{Z}[T]$ are multiplicatively independent. Let $\mathcal{V}_{3,2}(\ell; u, w)$ be the set of $(v_1, v_2, v_3) \in U_\ell^3$ satisfying the two congruences $F(v_1)F(v_2)F(v_3) \equiv u \pmod \ell$ and $G(v_1)G(v_2)G(v_3) \equiv w \pmod \ell$. Then $\#\mathcal{V}_{3,2}(\ell; u, w) \ll_{F,G} \varphi(\ell)$, uniformly in primes $\ell$ and coprime residues $u, w$ mod $\ell$.*

Our starting idea will be to look at $\mathcal{V}_{2,1}(\ell; w)$ and $\mathcal{V}_{3,2}(\ell; u, w)$ as subsets of the sets of $\mathbb{F}_\ell$-rational points of certain varieties over the algebraic closure $\overline{\mathbb{F}}_\ell$ of $\mathbb{F}_\ell$.

**Proposition 10.2.** *Let $V$ be a variety defined over $\mathbb{F}_\ell$ and $V(\mathbb{F}_\ell) := V \cap \mathbb{F}_\ell$.*

**(a)** *If $V$ is an absolutely irreducible affine plane curve, then $\#V(\mathbb{F}_\ell) \le \ell + O(\sqrt{\ell})$, where the implied constant depends only on the degree of $V$.*

**(b)** *Let $d$ be the positive integer such that $V \subset (\overline{\mathbb{F}}_\ell)^d$. We have $\#V(\mathbb{F}_\ell) \ll \ell^{\dim V}$, where $\dim V$ is the dimension of $V$ as a variety, and the implied constant depends at most on $d$ and on the number and degrees of the polynomials defining $V$.*

Subpart(a) is a consequence of [24, Corollary 2b], while subpart (b) is a weaker version of [13, Claim 7.2] but in fact goes back to work of Lang and Weil [22, Lemma 1]. To make use of the aforementioned results, we will also be needing the following observations.

**Lemma 10.3.** *Let $F, G \in \mathbb{Z}[T]$ be fixed multiplicatively independent polynomials such that $F$ is not squarefull. There exist constants $\kappa_0(F)$ and $\kappa_1(F, G)$ such that:*

**(a)** *For any $N \ge 2$, $\ell > \kappa_0(F)$ and $w \in \mathbb{F}_\ell^\times$, the polynomial $\prod_{i=1}^N F(X_i) - w$ is absolutely irreducible over $\mathbb{F}_\ell$, that is, it is irreducible in the ring $\overline{\mathbb{F}}_\ell[X_1, \ldots, X_N]$.*

**(b)** *For any $\ell > \kappa_1(F, G)$ and $u, w \in \mathbb{F}_\ell^\times$, the polynomial $F(X)F(Y)F(Z) - u$ is irreducible and doesn't divide the polynomial $G(X)G(Y)G(Z) - w$ in the ring $\overline{\mathbb{F}}_\ell[X, Y, Z]$.*

*Proof.* Write $F := r \prod_{j=1}^M G_j^{b_j}$ for some $r \in \mathbb{Z}$, $b_j \in \mathbb{N}$, and pairwise coprime irreducibles $G_j \in \mathbb{Z}[T]$, so that by the nonsquarefullness of $F$ in $\mathbb{Z}[T]$, we have $b_j = 1$ for some $j \in [M]$. By the observations at the start of the proof of Proposition 5.3, there exists a constant $\kappa_0(F)$ such that for any prime $\ell > \kappa_0(F)$, $\ell$ doesn't divide the leading coefficient of $F$ and $\prod_{j=1}^M G_j$ is separable in $\mathbb{F}_\ell[T]$. This forces $\prod_{\substack{\theta \in \overline{\mathbb{F}}_\ell \\ F(\theta)=0}} (T - \theta)^2 \nmid F(T)$ in $\overline{\mathbb{F}}_\ell[T]$.

*Proof of (a).* We will show that for any $\ell > \kappa_0(F)$ and $U, V \in \overline{\mathbb{F}}_\ell[X_1, \ldots, X_N]$ satisfying

$$\prod_{i=1}^{N} F(X_i) - w = U(X_1, \ldots, X_N)V(X_1, \ldots, X_N), \tag{10.1}$$

one of $U$ or $V$ must be constant. First note that for any root $\theta \in \overline{\mathbb{F}}_\ell$ of $F$, we have $-w = U(X_1, \ldots, X_{N-1}, \theta)V(X_1, \ldots, X_{N-1}, \theta)$, forcing $U(X_1, \ldots, X_{N-1}, \theta)$ and $V(X_1, \ldots, X_{N-1}, \theta)$ to be constant in the ring $\overline{\mathbb{F}}_\ell[X_1, \ldots, X_N]$. Writing $U(X_1, \ldots, X_N)$, $V(X_1, \ldots, X_N)$ as

$$\sum_{\substack{i_1,\ldots,i_{N-1}\geq 0\\ i_1\leq R_1,\ldots,i_{N-1}\leq R_{N-1}}} u_{i_1,\ldots,i_{N-1}}(X_N)\ X_1^{i_1}\cdots X_{N-1}^{i_{N-1}}, \qquad \sum_{\substack{j_1,\ldots,j_{N-1}\geq 0\\ j_1\leq T_1,\ldots,j_{N-1}\leq T_{N-1}}} v_{j_1,\ldots,j_{N-1}}(X_N)\ X_1^{j_1}\cdots X_{N-1}^{j_{N-1}}$$

respectively (where $u_{i_1,\ldots,i_{N-1}}, v_{j_1,\ldots,j_{N-1}} \in \overline{\mathbb{F}}_\ell[X_N]$ and neither $u_{R_1,\ldots,R_{N-1}}$ nor $v_{T_1,\ldots,T_{N-1}}$ is identically zero), we thus find that $u_{i_1,\ldots,i_{N-1}}(\theta) = v_{j_1,\ldots,j_{N-1}}(\theta) = 0$ for any $(i_1, \ldots, i_{N-1}) \neq (0, \ldots, 0)$, $(j_1, \ldots, j_{N-1}) \neq (0, \ldots, 0)$, and any $\theta$ as above. Thus, if the tuples $(R_1, \ldots, R_{N-1})$ and $(T_1, \ldots, T_{N-1})$ are both nonzero, then $\prod_{\substack{\theta\in\overline{\mathbb{F}}_\ell\\ F(\theta)=0}} (X_N - \theta)$ divides $u_{R_1,\ldots,R_{N-1}}(X_N)$ and $v_{T_1,\ldots,T_{N-1}}(X_N)$ in $\overline{\mathbb{F}}_\ell[X_N]$. But then, if $\alpha \in \mathbb{Z}$ is the leading coefficient of $F$, then comparing the monomials (in $X_1, \ldots, X_{N-1}$) with maximal total degree in (10.1), we find that $\alpha^{N-1}F(X_N) = u_{R_1,\ldots,R_{N-1}}(X_N)\, v_{T_1,\ldots,T_{N-1}}(X_N) \equiv 0 \pmod{\prod_{\substack{\theta\in\overline{\mathbb{F}}_\ell\\ F(\theta)=0}} (X_N-\theta)^2}$, which is impossible by the observations in the first paragraph of the proof. This forces one of $(R_1, \ldots, R_{N-1})$ or $(T_1, \ldots, T_{N-1})$ to be $(0, \ldots, 0)$, say the latter. Then $V(X_1, \ldots, X_N) = v_{0,\ldots,0}(X_N)$ and since $N \geq 2$, plugging $X_1 := \theta$ for some root $\theta \in \overline{\mathbb{F}}_\ell$ of $F$ into (10.1) yields $-w = U(\theta, X_2, \ldots, X_N)v_{0,\ldots,0}(X_N)$, forcing $V$ to be identically constant.

*Proof of (b).* We claim that for all primes $\ell \gg_{F,G} 1$, if the rational function $F^aG^b$ is constant in the ring $\overline{\mathbb{F}}_\ell(T)$ for some integers $a, b$, then $a \equiv b \equiv 0 \pmod{\ell}$.[11] The argument for this is a simple variant of that given for the inequality "$\mathrm{ord}_\ell(\widetilde{F}) \leq \mathbb{1}_{\ell\leq C_1}C_1$" in the proof of Proposition 5.3(b), so we only sketch the outline. Since $\{F, G\} \subset \mathbb{Z}[T]$ are multiplicatively independent, the polynomials $\{F'G, FG'\} \subset \mathbb{Z}[T]$ are $\mathbb{Q}$-linearly independent, hence so are the columns of the matrix $M_1$ listing the coefficients of $F'G$ and $FG'$ in two columns. Hence we can find invertible matrices $P_1$ and $Q_1$ (where $Q_1$ is a $2 \times 2$ matrix) such that $P_1M_1Q_1 = \mathrm{diag}(\beta_1, \beta_2)$ for some $\beta_1, \beta_2 \in \mathbb{Z} \setminus \{0\}$ satisfying $\beta_1 \mid \beta_2$. Let $\ell > |\beta_2|$ be any prime not dividing the leading coefficients of $F$, $G$, $F'G$ or $FG'$. If $F^aG^b$ is identically constant in $\mathbb{F}_\ell[T]$, then $aF'G+bFG' \equiv 0$ in $\mathbb{F}_\ell[T]$, so $M_1(a\ b)^\top \equiv 0 \pmod{\ell}$. Hereafter, familiar calculations yield $(a\ b)^\top \equiv 0 \pmod{\ell}$.

Collecting our observations, we have shown that there exists a constant $\kappa_1(F, G)$ such that for all primes $\ell > \kappa_1(F, G)$, the following three properties hold:

(i) $\ell > \kappa_0(F)$, so that $\prod_{\substack{\theta\in\overline{\mathbb{F}}_\ell\\ F(\theta)=0}} (T - \theta)^2 \nmid F(T)$ in $\overline{\mathbb{F}}_\ell[T]$;

(ii) $\ell$ doesn't divide the leading coefficient of $F$ or $G$; and,

(iii) For any $a, b \in \mathbb{Z}$ for which $F^aG^b$ is identically constant in $\overline{\mathbb{F}}_\ell(T)$, we have $\ell \mid a$ and $\ell \mid b$.

[11] It is not difficult to see that this also forces $a = b = 0$, but we won't need that.

We will now show that any such constant $\kappa_1(F,G)$ satisfies the property in subpart (b) of the lemma. By subpart (a), $F(X)F(Y)F(Z)-u$ is already irreducible in $\overline{\mathbb{F}}_\ell[X,Y,Z]$ for any $u \in \mathbb{F}_\ell^\times$. Assume by way of contradiction that for some $\ell > \kappa_1(F,G)$ and $u, w \in \mathbb{F}_\ell^\times$, we have

$$G(X)G(Y)G(Z)-w = H_0(X,Y,Z)\ (F(X)F(Y)F(Z)-u) \ \text{ for some } H_0 \in \overline{\mathbb{F}}_\ell[X,Y,Z]. \tag{10.2}$$

Write $H_0(X,Y,Z) =: \sum_{\substack{0\le i_1\le r_1\\ 0\le i_2\le r_2}} h_{i_1,i_2}(X)Y^{i_1}Z^{i_2}$ for some $h_{i_1,i_2} \in \overline{\mathbb{F}}_\ell[X]$ with $h_{r_1,r_2}$ not identically zero. If $(r_1,r_2)=(0,0)$, then substituting a root of $F$ and $G$ in place of $Y$ and $Z$ respectively, we see that $H_0$ must be a constant $\lambda_0 \in \overline{\mathbb{F}}_\ell \setminus \{0\}$ satisfying $w = \lambda_0 u$. Thus $G(X)G(Y)G(Z) = \lambda_0 F(X)F(Y)F(Z)$. Now substituting some $\eta \in \overline{\mathbb{F}}_\ell$ which is not a root of $FG$ in place of both $Y$ and $Z$ leads to $F(X)G(X)^{-1} = \lambda_0^{-1}F(\eta)^{-2}G(\eta)^2$, a nonzero constant. But since $(1,-1) \not\equiv (0,0) \pmod{\ell}$, this violates condition (iii) in the definition of $\kappa_1(F,G)$. Hence $(r_1,r_2) \neq (0,0)$.

Let $\alpha, \beta \in \mathbb{Z}$ denote the leading coefficients of $F$ and $G$ respectively. Comparing the monomials in $Y$ and $Z$ of maximal total degree in (10.2) yields $\beta^2 G(X) = \alpha^2 F(X) h_{r_1,r_2}(X)$ in $\overline{\mathbb{F}}_\ell[X]$, so that (since either side of this identity is nonzero), we get $F \mid G$ in $\overline{\mathbb{F}}_\ell[X]$. Write $G = F^m H$ for some $m \ge 1$ and $H \in \overline{\mathbb{F}}_\ell[X]$ such that $F \nmid H$ in $\overline{\mathbb{F}}_\ell[X]$. An easy finite induction shows that with $G_t(X,Y,Z) := F(X)^{m-t}F(Y)^{m-t}F(Z)^{m-t}H(X)H(Y)H(Z) - u^{-t}w$ and $\widehat{F}(X,Y,Z) := F(X)F(Y)F(Z)-u$, we have $\widehat{F} \mid G_t$ for each $t \in \{0,1,\dots,m\}$. Indeed, the case $t=0$ is just (10.2), and if $\widehat{F} \mid G_t$ for some $t \le m-1$, then writing $G_t = Q_t\widehat{F}$ shows that $F(X)F(Y)F(Z) \mid (Q_t(X,Y,Z) - u^{-(t+1)}w)$. With $Q_{t+1}$ defined by $Q_t(X,Y,Z) - u^{-(t+1)}w = F(X)F(Y)F(Z)Q_{t+1}(X,Y,Z)$, we obtain $G_{t+1} = Q_{t+1}\widehat{F}$ completing the induction.

Applying this last observation with $t := m$ shows that $\widehat{F}(X,Y,Z)$ divides $H(X)H(Y)H(Z) - u^{-m}w$ in $\overline{\mathbb{F}}_\ell[X,Y,Z]$. We claim that this forces $H$ to be constant. Indeed if not, then letting $\gamma \in \overline{\mathbb{F}}_\ell \setminus \{0\}$ be the leading coefficient of $H$, [12] writing $H(X)H(Y)H(Z) - u^{-m}w = (F(X)F(Y)F(Z)-u)\ \sum_{\substack{0\le i_1\le b_1\\ 0\le i_2\le b_2}} g_{i_1,i_2}(X)Y^{i_1}Z^{i_2}$ for some $g_{i_1,i_2} \in \overline{\mathbb{F}}_\ell[X]$ with $g_{b_1,b_2} \neq 0$, and comparing the monomials in $Y$ and $Z$ of maximal degree, we obtain $\gamma^2 H(X) = \alpha^2 F(X) g_{b_1,b_2}(X)$. This leads to $F \mid H$, contrary to hypothesis. Hence $H$ must be constant, so the identity $F^{-m}G = H$ in $\overline{\mathbb{F}}_\ell(X)$ violates condition (iii) in the definition of $\kappa_1(F,G)$, as $(-m,1) \not\equiv (0,0) \pmod{\ell}$. This shows that $\widehat{F}$ cannot divide $G(X)G(Y)G(Z) - w$, completing the proof. □

Given a commutative ring $R$ and an $R$-module $M$, we say that $x \in R$ is an $M$-**regular element** if $x$ is not a zero-divisor on $M$, that is, if $xz = 0$ for some $z \in M$ implies $z = 0$. A sequence $x_1, \dots, x_n$ of elements of $R$ is said to be $M$-**regular** if $x_1$ is an $M$-regular element, each $x_i$ is an $M/(x_1,\dots,x_{i-1})M$-regular element, *and* $M/(x_1,\dots,x_n)M \neq 0$. It is well-known (see [5, Proposition 1.2.14]) that for any proper ideal $I$ in a Noetherian ring $R$, the height of $I$ is at least the length of the longest $R$-regular sequence contained in $I$.

*Proof of Proposition 10.1.* With $\kappa_0(F)$ and $\kappa_1(F,G)$ as in Lemma 10.3, the affine plane curve $\{(X,Y) \in \overline{\mathbb{F}}_\ell^2 : \ F(X)F(Y) - w = 0\}$ is absolutely irreducible for any $\ell > \kappa_0(F)$, so that Proposition 10.2(a) yields Proposition 10.1(a). For (b), it suffices to show that for any prime $\ell > \kappa_1(F,G)$, the variety $V_\ell \subset \overline{\mathbb{F}}_\ell^3$ defined by the polynomials $\widehat{F}(X,Y,Z) := F(X)F(Y)F(Z) - u$

[12] Here $\gamma \neq 0$ in $\overline{\mathbb{F}}_\ell$ because $\ell$ doesn't divide the leading coefficient of $G = F^m H$.

and $\widehat{G}(X,Y,Z) \coloneqq G(X)G(Y)G(Z) - w$ has $\ll_{F,G} \ell$ many $\mathbb{F}_\ell$-rational points. Consider the ideal $I(V_\ell)$ of the ring $R \coloneqq \overline{\mathbb{F}}_\ell[X,Y,Z]$ consisting of all polynomials vanishing at all the points of $V_\ell$, so that $(\widehat{F},\widehat{G}) \subset I(V_\ell)$. If $I(V_\ell) = R$, then $V_\ell = \emptyset$, so suppose $I(V_\ell) \subsetneq R$. Lemma 10.3(b) shows that the sequence $\widehat{G}, \widehat{F} \in I(V_\ell)$ is $R$-regular, so by [5, Proposition 1.2.14], $I(V_\ell)$ has height at least 2. By [4, Chapter 11, Exercise 7], the Krull-dimension of $R$ is 3, whence that of $R/I(V_\ell)$ is at most $3-2=1$ (by, say, [25, p. 31]). Thus $\dim(V_\ell) \le 1$, and Proposition 10.2 completes the proof. □

## 11. Restricted inputs to squarefree moduli: Proof of Theorem 2.3

Returning to the notation set up in the introduction, we start with the same initial reductions as in section 9. As such, in order to establish the theorem, it suffices to show that

$$\sideset{}{^*}\sum_{n:\ P_R(n)>q} 1 \ll \frac{x^{1/k}}{\varphi(q)^K (\log x)^{1-2\alpha_k/3}}, \tag{11.1}$$

with the respective values of $R$ defined in the statement. Here we again have $\epsilon = 1$ and $y = \exp(\sqrt{\log x})$ in the framework developed in section 4. We retain the notation $\omega_\parallel(n) = \#\{p > q : p^k \parallel n\}$ and $\omega^*(n) = \#\{p > q : p^{k+1} \mid n\}$ from section 9.

**The case $K=1$, $W_{1,k}$ not squarefull.** In this case, (11.1) would follow once we show that

$$\sideset{}{^*}\sum_{n:\ P_{k+1}(n)>q} 1 \ll \frac{x^{1/k}}{\varphi(q)(\log x)^{1-2\alpha_k/3}}, \tag{11.2}$$

Indeed, any $n$ counted in (11.2) which is divisible by the $(k+1)$-th power of a prime exceeding $q$ can be written in the form $mp^cP^k$ for some positive integers $m, c$ and primes $p, P$, satisfying $P = P(n) > z$, $q < p < P$, $c \ge k+1$, $P_{Jk}(m) \le y$ and $f(n) = f(m)f(p^c)W_k(P)$. Recalling that $\#\{u \in U_q : W_k(u) \equiv b \pmod{q}\} \ll D^{\omega(q)}$ uniformly in $b \in \mathbb{Z}$, the argument given for the second bound in (9.3) shows that the contribution of such $n$ is $\ll \frac{D^{\omega(q)}}{q^{1/k}\varphi(q)} \cdot \frac{x^{1/k}}{(\log x)^{1-2\alpha_k/3}} \ll \frac{x^{1/k}}{\varphi(q)(\log x)^{1-2\alpha_k/3}}$. On the other hand, for any $n$ counted in (11.2) which is not divisible by the $(k+1)$-th power of any prime exceeding $q$, the condition $P_{k+1}(n) > q$ forces $\omega_\parallel(n) \ge 2$ (again since $q$ is sufficiently large and the $q$-rough part of $n$ is $k$-full). Thus $n = m(P_2P_1)^k$, for some $m$ and primes $P_1, P_2$ satisfying $P_1 \coloneqq P(n) > z$, $q < P_2 < P_1$, $P_{Jk}(m) \le y$ and $f(n) = f(m)W_k(P_1)W_k(P_2)$. The arguments before (9.5) show that the contribution of such $n$ is $\ll \frac{V'_{2,1}}{\varphi(q)^2} \cdot \frac{x^{1/k}}{(\log x)^{1-\alpha_k/2}} \exp((\log_3 x)^{O(1)})$, which is $\ll \frac{x^{1/k}}{\varphi(q)(\log x)^{1-2\alpha_k/3}}$ by Proposition 10.1(a).

**The remaining cases.** To complete the proof of Theorem 2.3, it thus remains to show that we may take:
(i) $R = k(Kk + K - k) + 1$ if $K, k \ge 2$ and at least one of $\{W_{i,k}\}_{1 \le i \le K}$ is not squarefull.
(ii) $R = k(Kk + K - k + 1) + 1$, in general.
We shall call (i) as "Subcase 1" and (ii) as "Subcase 2", and we shall denote $R = k(Kk + K - k + \mathbb{1}) + 1$ to mean the respective value of $R$ in the respective subcase.

We have the following analogues of the first two bounds in (9.3), which can be shown by replicating arguments and replacing the use of Proposition 4.4 by Corollary 5.4.

$$\sum\nolimits^*_{n:\ \omega_\|(n)\ge 2K+1} 1,\ \sum\nolimits^*_{\substack{n:\ \omega_\|(n)=2K\\ \omega^*(n)\ge 1}} 1 \ \ll\ \frac{x^{1/k}}{\varphi(q)^K(\log x)^{1-2\alpha_k/3}}, \tag{11.3}$$

If $\omega^*(n)=0$, then $k\omega_\|(n)\ge R\ge k(Kk+K-k+\mathbb{1})+1$, so that $\omega_\|(n)\ge Kk+K-k+\mathbb{1}+1\ge 2K+1$; hence, any $n$ with $\omega^*(n)=0$ counted in (11.1) is automatically counted in the first sum in (11.3). Likewise, the condition $\omega_\|(n)=2K$ forces $\sum_{p>q:\ p^{k+1}|n} v_p(n)\ge R-k\omega_\|(n)\ge k((K-1)(k-1)-1+\mathbb{1})+1\ge 1$, so that $\omega^*(n)\ge 1$; as such, any $n$ with $\omega_\|(n)=2K$ contributing to (11.1) is counted in the second sum in (11.3). Furthermore, by the third bound in (9.3), the contribution of all $n$ having $\omega^*(n)\ge Kk$ to the left hand side of (11.1) is absorbed in the right hand side. It thus suffices to show that for any $r\in[2K-1]$ and $s\in[Kk-1]$, the contribution $\Sigma_{r,s}$ of all $n$ with $\omega_\|(n)=r$ and $\omega^*(n)=s$ to the left hand side of (11.1) is absorbed in the right hand side.

Recall that any $n$ counted in $\Sigma_{r,s}$ is of the form $mp_1^{c_1}\cdots p_s^{c_s}P_1^k\cdots P_r^k$ for some distinct primes $p_1,\dots,p_s$, $P_1,\dots,P_r$ and integers $m,c_1,\dots,c_s$, which satisfy the conditions (i)–(v) in the proof of Theorem 2.2, but with the current values of $R$. Once again, the integers $\tau_1,\dots,\tau_s$ defined by $\tau_j:=\min\{c_j,R-kr\}$ satisfy $\tau_j\in[k+1,R-kr]$, $\tau_j\le c_j$ and $\tau_1+\cdots+\tau_s\ge R-kr$. (Here $R-kr\ge k+1$ follows from $r\le 2K-1$ and $R=k(Kk+K-k+\mathbb{1})+1$.) Thus,

$$\Sigma_{r,s}\le \sum_{\substack{\tau_1,\dots,\tau_s\in[k+1,R-kr]\\ \tau_1+\cdots+\tau_s\ge R-kr}} \mathcal{N}_{r,s}(\tau_1,\dots,\tau_s), \tag{11.4}$$

where $\mathcal{N}_{r,s}(\tau_1,\dots,\tau_s)$ denotes the contribution of all $n$ counted in the left hand side of (11.1) which can be written in the form $mp_1^{c_1}\cdots p_s^{c_s}\ P_1^k\cdots P_r^k$ for some distinct primes $p_1,\dots,p_s,P_1$, $\cdots,P_r$ and integers $m,c_1,\dots,c_s$ satisfying $c_1\ge\tau_1,\dots,c_s\ge\tau_s$ and the conditions (i)–(v) in the proof of Theorem 2.2 (but with the current values of $R$). We will show that for each tuple $(\tau_1,\dots,\tau_s)$ occurring in (11.4), we have

$$\mathcal{N}_{r,s}(\tau_1,\dots,\tau_s)\ \ll\ \frac{x^{1/k}(\log_2 x)^{O(1)}}{q^K\log x}\exp\big(O(\sqrt{\log q})\big). \tag{11.5}$$

Now the bound (9.10) continues to hold, so we have

$$\mathcal{N}_{r,s}(\tau_1,\dots,\tau_s)\ll\frac{1}{q^{(\tau_1+\cdots+\tau_s)/k-s}}\ \frac{V'_{r,K}}{\varphi(q)^r}\cdot\frac{x^{1/k}(\log_2 x)^{O(1)}}{\log x} \tag{11.6}$$

with the current values of $r,s,\tau_1,\dots,\tau_s$ and with $V'_{r,K}$ defined as before. By (5.22),

$$\mathcal{N}_{r,s}(\tau_1,\dots,\tau_s)\ll\frac{\exp\big(O(\omega(q))\big)}{q^{(\tau_1+\cdots+\tau_s)/k-s+r/2}}\cdot\frac{x^{1/k}(\log_2 x)^{O(1)}}{\log x}\ll\frac{\exp\big(O(\omega(q))\big)}{q^{\max\{s/k+r/2,\ R/k-r/2-s\}}}\cdot\frac{x^{1/k}(\log_2 x)^{O(1)}}{\log x}.$$

Now $\max\{s/k+r/2,R/k-r/2-s\}>K$ whenever one of the following holds:
**(a)** In Subcase 1, we have *either* $k\ge 3$, $r\ge 3$, *or* $k=2$, $r\ge 4$.
**(b)** In Subcase 2, we have $r\ge 2$.
Indeed, if $s/k+r/2\le K$, then $s\le k(K-r/2)$, so that $R/k-r/2-s\ge K+(k-1)(r/2-1)-1+\mathbb{1}+1/k$. This last quantity strictly exceeds $K$ precisely under (a) or (b) above, establishing (11.5) under one of these two conditions. It thus only remains to tackle:

**(i)** the possibility that $r = 1$ in both Subcases 1 and 2, and
**(ii)** the possibilities $r = 2$ and $k = 2, r = 3$ in Subcase 1.

The possibility $r = 1$ is easily handled (in both subcases) by inserting into (11.6) the trivial bound $V'_{r,K} = V'_{1,K} \ll D_{\min}^{\omega(q)}$. Now assume we are in Subcase 1 and either $r = 2$ or $k = 2, r = 3$. Suppose wlog that $W_{1,k}$ is not squarefull. If $r = 2$, then Proposition 10.1(a) yields $\#\mathcal{V}_{2,K}^{(k)}(q; (w_i)_{i=1}^K)/\varphi(q)^2 \le \#\mathcal{V}_{2,1}(q; w_1)/\varphi(q)^2 \ll \varphi(q)^{-1} \exp(O(\sqrt{\log q}))$, uniformly for $(w_i)_{i=1}^K \in U_q^K$. Inserting this bound into (11.6), we deduce that $\mathcal{N}_{2,s}(\tau_1, \dots, \tau_s) \ll q^{-\max\{s/k+1, R/k-1-s\}} \cdot \frac{x^{1/k}(\log_2 x)^{O(1)}}{\log x} \exp\left(O(\sqrt{\log q})\right)$. Since $\max\{s/k+1, R/k-1-s\} \ge K$, this shows (11.5) in Subcase 1 when $r = 2$.

For $k = 2, r = 3$, the multiplicative independence of $\{W_{1,k}, W_{2,k}\}$ allows us to use Proposition 10.1(b) to get $\#\mathcal{V}_{3,K}^{(k)}(q; (w_i)_{i=1}^K)/\varphi(q)^3 \ll \exp\left(O(\omega(q))\right)/\varphi(q)^2$ uniformly for $(w_i)_{i=1}^K$. By (11.6), $\mathcal{N}_{3,s}(\tau_1, \dots, \tau_s) \ll q^{-\max\{s/2+2,\ R/2-1-s\}} \cdot \frac{x^{1/k}(\log_2 x)^{O(1)}}{\log x} \exp\left(O(\omega(q))\right)$, and it is easily checked that $\max\{s/2+2,\ R/2-1-s\} > K$. This shows (11.5) in Subcase 1 when $k = 2, r = 3$, completing the proof of Theorem 2.3.

11.1. **Optimality in the conditions of Theorem 2.3.** We will now show that the first two values of $R$ given in Theorem 2.3 are optimal. We retain the setting in subsection § 8.1 we had used to show optimality in Theorem 2.1(ii). To recall: fix an arbitrary $k \in \mathbb{N}$ and $d > 1$, and define $W_{i,k}(T) := \prod_{j=1}^d (T - 2j) + 2(2i - 1)$, so that $\prod_{i=1}^K W_{i,k}$ is separable (over $\mathbb{Q}$). Let $\widetilde{C}_0 > 4KD$ be any constant (depending only on $\{W_{i,k}\}_{1 \le i \le K}$) exceeding the size of the (nonzero) discriminant of $\prod_{i=1}^K W_{i,k}$, and such that any $\widetilde{C}_0$-rough $k$-admissible integer lies in $\mathcal{Q}(k; f_1, \cdots, f_K)$. Fix a prime $\ell_0 > C_0$ and nonconstant polynomials $\{W_{i,v}\}_{\substack{1 \le i \le K \\ 1 \le v < k}} \subset \mathbb{Z}[T]$ with all coefficients divisible by $\ell_0$. Let $q \le (\log x)^{K_0}$ be any squarefree integer having $P^-(q) = \ell_0$, so that as before $q \in \mathcal{Q}(k; f_1, \cdots, f_K)$. Recall also that $(2(2i-1))_{i=1}^K \in U_q^K$, that any prime $P$ satisfying $\prod_{j=1}^d (P - 2j) \equiv 0 \pmod q$ also satisfies $f_i(P^k) \equiv 2(2i-1) \pmod q$, and that the congruence $\prod_{j=1}^d (v - 2j) \equiv 0 \pmod q$ has exactly $d^{\omega(q)}$ distinct solutions $v \in U_q$.

The first value $R = 2$ in Theorem 2.3 is optimal since the condition $P_2(n) > q$ cannot be replaced by the condition $P(n) > q$, as shown in (8.2). We now show that the condition "$R = k(Kk + K - k) + 1$" in Theorem 2.3 cannot be weakened to "$R = k(Kk + K - k)$" for **any** $K, k$. To this end, let $f_1, \dots, f_K \colon \mathbb{N} \to \mathbb{Z}$ be any multiplicative functions such that $f_i(p^v) := W_{i,v}(p)$ and $f_i(p^{k+1}) := 1$ for all primes $p$, all $i \in [K]$ and $v \in [k]$. Consider $n$ of the form $(p_1 \cdots p_{k(K-1)})^{k+1} P^k \le x$ where $P, p_1, \dots, p_{k(K-1)}$ are primes satisfying the conditions $P := P(n) > x^{1/3k}$, $q < p_{k(K-1)} < \cdots < p_1 < x^{1/4Kk^2}$, and $\prod_{1 \le j \le d}(P - 2j) \equiv 0 \pmod q$. Then $P_{k(Kk+K-k)}(n) = p_{k(K-1)} > q$ and $f_i(n) = f_i(P^k) \prod_{j=1}^{k(K-1)} f_i(p_j^{k+1}) \equiv 2(2i-1) \pmod q$ for each $i \in [K]$. Given $p_1, \dots, p_{k(K-1)}$, the number of primes $P$ satisfying $x^{1/3k} < P \le x^{1/k}/(p_1 \cdots p_{k(K-1)})^{1+1/k}$ is $\gg d^{\omega(q)} x^{1/k}/\varphi(q)(p_1 \cdots p_{k(K-1)})^{1+1/k} \log x$ by Siegel–Walfisz; here we have noted that $(p_1 \cdots p_{k(K-1)})^{1+1/k} \le x^{(K-1)(k+1)/4Kk^2} \le x^{1/2k}$. Dividing by $k!$ allows us to replace the condition $p_{k(K-1)} < \cdots < p_1$ by a distinctness condition, giving us

$$\sum_{\substack{n \le x:\, P_{k(Kk+K-k)}(n) > q \\ (\forall i)\, f_i(n) \equiv 2(2i-1) \pmod q}} 1 \;\gg\; \frac{d^{\omega(q)} x^{1/k}}{\varphi(q) \log x} \left( \mathcal{T}_1 - \mathcal{T}_2 \right), \tag{11.7}$$

where $\mathcal{T}_1$ denotes the sum ignoring the distinctness condition on the $p_1, \dots, p_{k(K-1)}$, and $\mathcal{T}_2$ denotes the sum over all the tuples $(p_1, \dots, p_{k(K-1)})$ for which $p_i = p_j$ for some $i \neq j \in [k(K-1)]$. Now $\mathcal{T}_1 = \prod_{1 \le j \le k(K-1)} \left( \sum_{q < p_j \le x^{1/4Kk^2}} p_j^{-(1+1/k)} \right) \gg 1/q^{K-1} (\log q)^{k(K-1)}$ while $\mathcal{T}_2 \ll \left( \sum_{p > q} p^{-(2+2/k)} \right) \left( \sum_{p > q} p^{-(1+1/k)} \right)^{k(K-1)-2} \ll 1/q^K$. Consequently, the expression on the right hand side of (11.7) is $\gg d^{\omega(q)} x^{1/k} / \varphi(q)^K (\log_2 x)^{k(K-1)+1} \log x$, which by Proposition 3.1, grows strictly faster than $\varphi(q)^{-K} \#\{n \le x : \gcd(f(n), q) = 1\}$ as soon as $d^{\omega(q)} > (\log x)^{(1+\epsilon)\alpha_k}$. We have already constructed such $q$ in subsection § 8.1. Hence, the condition $P_{k(Kk+K-k)+1}(n) > q$ in Theorem 2.3 is optimal for any values of $K$ and $k$.

As a remark, note that this example also shows that if $k = 1$, then for any $K$, the condition "$P_{2K+1}(n) > q$" coming from the third value of $R$ in Theorem 2.3 is "almost optimal" in the sense that it cannot be replaced by the condition "$P_{2K-1}(n) > q$".

## 12. Concluding Remarks

It is interesting to note that despite the extensive amount of 'multiplicative machinery' known in analytic number theory, there does not seem to be any estimate in the literature, a direct application of which can replace our arguments in section 7. For instance, Halász's Theorem only yields an upper bound on the character sums that is not precise enough, while a direct application of the (known forms of) the Landau-Selberg-Delange method, – one of the most precise estimates on the mean values of multiplicative functions known in literature, – seems to give an extremely small range of uniformity in $q$.

Theorem 2.3 suggests a few directions of improvement. First, we are still "one step away" from optimality in the $K \ge 2$, $k = 1$ case: we proved that "$2K + 1$" is sufficient while "$2K - 1$" is not, so the question is whether the optimal value is "$2K$" or "$2K + 1$". If it is the former, then we will need a sharper bound on $V'_{2K,K}$ than what comes from our methods in section 11. One can also ask whether it is possible to weaken the nonsquarefullness conditions in the theorem.

## Acknowledgements

This work was done in partial fulfillment of my PhD at the University of Georgia. As such, I would like to thank my advisor, Prof. Paul Pollack, for the past joint research and fruitful discussions that have led me to think about this question, as well as for his continued support and encouragement. I would also like to thank the Department of Mathematics at UGA for their support and hospitality.

**Data Availability** The manuscript has no associated data.

## Declarations

**Conflict of Interest** On behalf of all authors, the corresponding author states that there is no conflict of interest.

## References

[1] A. Akande, *Uniform distribution of polynomially-defined additive functions to varying moduli*, submitted.

[2] K. Alladi, *The distribution of $\nu(n)$ in the sieve of Eratosthenes*, Quart. J. Math. Oxford Ser. (2) **33** (1982), no. 130, 129–148.

[3] K. Alladi and P. Erdös, *On an additive arithmetic function*, Pacific J. Math. **71** (1977), no. 2, 275–294.

[4] M.F. Atiyah, and L.G. Macdonald, *Introduction to Commutative Algebra*, Addison-Wesley Publishing Company, 1969.

[5] W. Bruns, and J. Herzog, *Cohen-Macaulay Rings*, Cambridge Studies in Advanced Mathematics, vol. 39, Cambridge University Press, Cambridge, 1998.

[6] T. Cochrane, *Exponential sums modulo prime powers*, Acta Arith. **101** (2002), 131–149.

[7] 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.

[8] T. Cochrane and Z. Zheng., *Pure and mixed exponential sums.*, Acta Arith. **91** (1999), 249–278.

[9] H. Davenport, *On character sums in finite fields*, Acta Math. **71** (1939), 99–121.

[10] H. Delange, *On integral-valued additive functions*, J. Number Theory **1** (1969), 419–430.

[11] ———, *On integral-valued additive functions, II*, J. Number Theory **6** (1974), 161–170.

[12] T. Dence and C. Pomerance, *Euler's function in residue classes*, Ramanujan J. **2**(1998), 7–20.

[13] Z. Dvir, J. Kollár, and S. Lovett, *Variety Evasive Sets*, Comput. Complexity **23** (2014), 509–529, ISSN 1016-3328.

[14] P. Erdös and G. Szekeres, *Über die Anzahl der Abelschen Gruppen gegebener Ordnung und über ein verwandtes zahlentheoretisces Problem*, Acta Univ. Szeged, vol. **7** (1934-1935), pp. 95–102.

[15] O.M. Fomenko, *The distribution of values of multiplicative functions with respect to a prime modulus*, Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI), **93**, 1980, pp. 218-224. (Russian)

[16] D. Goldfeld, *On an additive prime divisor function of Alladi and Erdős*, Analytic number theory, modular forms and $q$-hypergeometric series, Springer Proc. Math. Stat., vol. 221, Springer, Cham, 2017, pp. 297–309.

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

[18] R.R. Hall and G. Tenenbaum, *Divisors*, Cambridge Tracts in Mathematics, vol. 90, Cambridge University Press, Cambridge, 1988.

[19] 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.

[20] ———, *The number of solutions of congruences of the nth degree with one unknown*, Mat. Sb. (N.S.) **109(151)** (1979), 171–187, 327.

[21] E. Landau, *Lösung des Lehmer'schen Problems*, American J. Math. **31** (1909), 86–102.

[22] S. Lang, and A. Weil. *Number of Points of Varieties in Finite Fields.*, American J. Math. **76**, no. 4 (1954), 819–827.

[23] 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., to appear.

[24] 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.

[25] H. Matsumura, *Commutative ring theory*, Cambridge Studies in Advanced Mathematics, vol. 8, Cambridge University Press, Cambridge, 2006.

[26] H.L. Montgomery and R.C. Vaughan, *Multiplicative number theory. I. Classical theory*, Cambridge Studies in Advanced Mathematics, vol. 97, Cambridge University Press, Cambridge, 2007.

[27] W. Narkiewicz, *On distribution of values of multiplicative functions in residue classes*, Acta Arith. **12** (1967), 269–279.

[28] ———, *Euler's function and the sum of divisors*, J. reine angew. Math. **323** (1981), 200–212.

[29] ———, *On a kind of uniform distribution for systems of multiplicative functions*, Litovsk. Mat. Sb. **22** (1982), 127–137.

[30] ———, *Distribution of coefficients of Eisenstein series in residue classes*, Acta Arith. **43** (1983), 83–92.

[31] ———, *Uniform distribution of sequences of integers in residue classes*, Lecture Notes in Mathematics, vol. 1087, Springer-Verlag, Berlin, 1984.

[32] W. Narkiewicz and F. Rayner, *Distribution of Values of $\sigma_2(n)$ in Residue Classes*, Monatsh. Math. **94** (1982), 133–141.

[33] K.K. Norton, *On the number of restricted prime factors of an integer. I*, Illinois J. Math. **20** (1976), 681–705.

[34] S.E. Payne, *A Second Semester of Linear Algebra*, University of Colorado Denver, 2009.

[35] S.S. Pillai, *Generalisation of a theorem of Mangoldt*, Proc. Indian Acad. Sci., Sect. A **11** (1940), 13–20.

[36] P. Pollack and A. Singha Roy, *Joint distribution in residue classes of polynomial-like multiplicative functions*, Acta Arith. **202** (2022), 89–104.

[37] ———, *Benford behavior and distribution in residue classes of large prime factors*, Canad. Math. Bull., **66** (2023), no. 2, 626–642.

[38] ———, *Distribution in coprime residue classes of polynomially-defined multiplicative functions*, Math. Z. **303** (2023), no. 4, Paper No. 93, 20. MR 4565094.

[39] C. Pomerance, *On the distribution of amicable numbers*, J. Reine Angew. Math. **293(294)** (1977), 217–222.

[40] F. Rayner, *Weak Uniform Distribution for Divisor Functions. I*, Math. Comp. **50** (1988), 335–342.

[41] ———, *Weak Uniform Distribution for Divisor Functions. II*, Math. Comp. **51** (1988), 331–337.

[42] W.M. Schmidt, *Equations over finite fields*, Lecture Notes in Mathematics, vol. 536, Springer-Verlag Berlin Heidelberg 1976.

[43] W. Schwarz and J. Spilker, *Arithmetical functions*, London Mathematical Society Lecture Note Series, vol. 184, Cambridge University Press, Cambridge, 1994, An introduction to elementary and analytic properties of arithmetic functions and to some of their almost-periodic properties.

[44] E.J. Scourfield, *Uniform estimates for certain multiplicative properties*, Monatsh. Math. **97** (1984), 233–247.

[45] ———, *A uniform coprimality result for some arithmetic functions*, J. Number Theory **20** (1985), 315–353

[46] A. Singha Roy, *Joint distribution in residue classes of families of polynomially-defined additive functions*, submitted.

[47] ———, *Mean values of multiplicative functions and applications to the distribution of the sum of divisors*, submitted.

[48] ———, *Joint distribution in reisude classes of families of polynomially-defined multiplicative functions II*, submitted.

[49] J. Śliwa, *On distribution of values of $\sigma(n)$ in residue classes*, Colloq. Math. **27** (1973), 283-291, 332.

[50] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, third ed., Graduate Studies in Mathematics, vol. 163, American Mathematical Society, Providence, RI, 2015.

[51] D. Wan, *Generators and irreducible polynomials over finite fields*, Math. Comp. **66** (1997), no. 219, 1195–1212.

[52] 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