

# Equidistribution of high traces of random matrices over finite fields and cancellation in character sums of high conductor

Ofir Gorodetsky, Valeriya Kovaleva

## Abstract

Let  $g$  be a random matrix distributed according to uniform probability measure on the finite general linear group  $\mathrm{GL}_n(\mathbb{F}_q)$ . We show that  $\mathrm{Tr}(g^k)$  equidistributes on  $\mathbb{F}_q$  as  $n \rightarrow \infty$  as long as  $\log k = o(n^2)$  and that this range is sharp. We also show that nontrivial linear combinations of  $\mathrm{Tr}(g^1), \dots, \mathrm{Tr}(g^k)$  equidistribute as long as  $\log k = o(n)$  and this range is sharp as well. Previously equidistribution of either a single trace or a linear combination of traces was only known for  $k \leq c_q n$ , where  $c_q$  depends on  $q$ , due to work of the first author and Rodgers.

We reduce the problem to exhibiting cancellation in certain short character sums in function fields. For the equidistribution of  $\mathrm{Tr}(g^k)$  we end up showing that certain *explicit* character sums modulo  $T^{k+1}$  exhibit cancellation when averaged over monic polynomials of degree  $n$  in  $\mathbb{F}_q[T]$  as long as  $\log k = o(n^2)$ . This goes far beyond the classical range  $\log k = o(n)$  due to Montgomery and Vaughan. To study these sums we build on the argument of Montgomery and Vaughan but exploit additional symmetry present in the considered sums.

## 1 Introduction

Fix  $\mathbb{F}_q$ , the finite field of  $q$  elements. We denote its characteristic by  $\mathrm{char}(\mathbb{F}_q)$ . Let  $g \in \mathrm{GL}_n(\mathbb{F}_q)$  be an invertible  $n \times n$  matrix over  $\mathbb{F}_q$  chosen according to the uniform probability measure. The first author and Rodgers [12, Thm. 1.1] showed that  $\mathrm{Tr}(g^k)$  equidistributes in  $\mathbb{F}_q$  as  $n \rightarrow \infty$ , uniformly for  $k \leq c_q n$  for some sufficiently small  $c_q$  depending on  $q$ , and the rate of convergence is superexponential. However, nothing beyond  $k = O(n)$  was known. We describe our three main results.

**Theorem 1.1.** *Let  $g \in \mathrm{GL}_n(\mathbb{F}_q)$  be chosen uniformly at random. Let  $k = k(n)$  be a positive integer such that  $\log k = o(n^2)$ . The distribution of  $\mathrm{Tr}(g^k)$  tends to the uniform distribution on  $\mathbb{F}_q$  as  $n$  tends to  $\infty$ .*

The range  $\log k = o(n^2)$  in Theorem 1.1 is optimal in the sense that we cannot replace it with  $\log k = O(n^2)$ . Indeed, if we take  $k = |\mathrm{GL}_n(\mathbb{F}_q)| = \prod_{i=0}^{n-1} (q^n - q^i)$  then  $\log k \asymp_q n^2$  and, by Lagrange's theorem,  $g^k = I_n$  for  $g \in \mathrm{GL}_n(\mathbb{F}_q)$  and so  $\mathrm{Tr}(g^k) \equiv n$  does not equidistribute. (This shows  $\mathrm{Tr}(g^k)$  is periodic in  $k$ .) Theorem 1.1 leaves open the question, not explored in this paper, whether the range can be extended to  $\log k \leq c_q n^2$  for some constant  $c_q > 0$  depending on  $q$ ; we believe the answer is negative. We also prove a theorem for combination of traces.

**Theorem 1.2.** *Let  $g \in \mathrm{GL}_n(\mathbb{F}_q)$  be chosen uniformly at random. Let  $k = k(n)$  be a positive integer such that  $\mathrm{char}(\mathbb{F}_q) \nmid k$  and  $\log k = o(n)$ . Let  $a_j \in \mathbb{F}_q$  for  $j = 1, \dots, k$  be arbitrary constants with  $a_k \neq 0$ . Then the distribution of  $\sum_{1 \leq i \leq k} a_i \mathrm{Tr}(g^i)$  tends to the uniform distribution on  $\mathbb{F}_q$  as  $n$  tends to  $\infty$ .*

Over  $\mathbb{F}_q$  we have  $\mathrm{Tr}(g^{\mathrm{char}(\mathbb{F}_q)}) = \mathrm{Tr}(g)$  and the condition  $\mathrm{char}(\mathbb{F}_q) \nmid k$  is necessary to avoid trivial linear combinations. The range  $\log k = o(n)$  in Theorem 1.2 is optimal in the sense that it cannot be replaced by  $\log k = O(n)$ , as we now demonstrate by giving an example where  $\sum_{1 \leq i \leq k} a_i \mathrm{Tr}(g^i)$  ( $\mathrm{char}(\mathbb{F}_q) \nmid k$ ,  $a_k \neq 0$ ) is not equidistributed and  $\log k \asymp_q n$ . Let  $\overline{\mathbb{F}_q}$  denote an algebraic closure of  $\mathbb{F}_q$ , and let  $f \in \mathbb{F}_q[T]$  be a monic polynomial. Given its factorization  $f(T) = \prod_{j=1}^{\deg f} (T - \lambda_j)$  over  $\overline{\mathbb{F}_q}$ , we define its  $i$ th ( $i \geq 0$ ) power sum symmetric polynomial as

$$p_i(f) := \sum_{j=1}^{\deg f} \lambda_j^i \in \overline{\mathbb{F}_q}. \quad (1.1)$$

By Newton's identities  $p_i(f)$  is an integral multivariate polynomial in coefficients of  $f$ , hence  $p_i(f)$  is in fact in  $\mathbb{F}_q$ . Moreover, we have  $p_i(fg) = p_i(f) + p_i(g)$  for any monic  $f, g \in \mathbb{F}_q[T]$ . If  $f(T) = \det(I_n T - g)$  is the characteristic polynomial of a matrix  $g$ , then  $p_i(f) = \text{Tr}(g^i)$ .

Let  $F(T) := \sum_{i=0}^k a_i T^i$  be the product of all monic irreducible polynomials in  $\mathbb{F}_q[T]$  of degree at most  $n$  (so  $a_k = 1, a_0 = 0$ ). Then, for any  $\lambda \in \cup_{j=1}^n \mathbb{F}_{q^j}$ ,

$$F(\lambda) = \sum_{i=1}^k a_i \lambda^i = 0.$$

Hence  $\sum_{i=1}^k a_i p_i(f) = 0$  for any  $f$  with  $\deg f \leq n$  and

$$\sum_{i=1}^k a_i \text{Tr}(g^i) = \sum_{i=1}^k a_i p_i(\det(I_n T - g)) = 0$$

for all  $g \in \text{GL}_n(\mathbb{F}_q)$ . Finally,  $k = \deg F \asymp q^n$  by the Prime Polynomial Theorem (see Lemma 3.1). If  $k$  happens to be divisible by  $\text{char}(\mathbb{F}_q)$ , we can replace  $F$  by  $TF$ .

The question whether the range of Theorem 1.2 can be extended to  $\log k \leq c_q n$  for some constant  $c_q > 0$ , for which we believe the answer is negative, remains open and is not explored here.

For certain values of  $k$ , for example, primes, we can go beyond the range  $\log k = o(n^2)$  of Theorem 1.1 using the following arithmetic criterion.

**Theorem 1.3.** *Let  $g \in \text{GL}_n(\mathbb{F}_q)$  be chosen uniformly at random. Let  $k = k(n)$  be a positive integer. Suppose*

$$\sum_{\substack{12 \log_q n < d \leq n \\ \gcd(k, q^d - 1) < q^{d/3}}} d^{-1} \rightarrow \infty \quad (1.2)$$

*as  $n \rightarrow \infty$ . Then, the distribution of  $\text{Tr}(g^k)$  tends to the uniform distribution on  $\mathbb{F}_q$  as  $n$  tends to  $\infty$ .*

As we shall see later, Theorem 1.1 is a consequence of Theorem 1.3.

## 1.1 Comparison with random matrix theory

Our investigation was motivated by results in random matrix theory, although we do not use any techniques from this area. Let  $U_n(\mathbb{C})$  be the group of  $n \times n$  unitary matrices over complex numbers, endowed with Haar measure of total mass 1. A classical result of Diaconis and Shahshahani [9] states that for any  $k \geq 1$ , the vector  $X_k = (\text{Tr}(U), \text{Tr}(U^2)/\sqrt{2}, \dots, \text{Tr}(U^k)/\sqrt{k})$  converges in distribution to  $Y_k = (Z_1, Z_2, \dots, Z_k)$ , where  $(Z_j)_{j=1}^k$  are independent standard complex Gaussians. Johansson [18, Thm. 2.6] showed that the rate of convergence of a linear combination of  $\text{Tr}(U^i)$  to a suitable Gaussian is superexponential in total variation distance. In [19], Johansson and Lambert extended [18] to the total variation distance of  $X_k$  from  $Y_k$  uniformly for  $k \ll n^{2/3-\varepsilon}$ , and in [6], Courteaut and Johansson established similar results for orthogonal and symplectic groups. In a recent work, Courteaut, Johansson, and Lambert [7] studied the convergence of  $\text{Tr}(U^k)/\sqrt{k}$  to  $Z_k$  as  $k$  varies, obtaining, among other results, that the distance goes to 0 for any  $k$  in the range  $1 \leq k < n$ . As for the complementary range, Rains [24] proved that there is a stabilizing phenomenon once  $k \geq n$ : The eigenvalues of  $U^k$  become distributed as  $n$  independent uniform random variables on the unit circle, and in particular  $\text{Tr}(U^k)/\sqrt{n}$  tends in distribution to the standard complex Gaussian.

## 1.2 Symmetric function perspective

Let us also formulate a natural problem in symmetric functions that will turn out to share strong similarities with the equidistribution of  $\text{Tr}(g^k)$ . Let  $e_k(t_1, \dots, t_n)$  be the  $k$ th elementary symmetric polynomial

$$e_k(t_1, \dots, t_n) := \sum_{1 \leq j_1 < j_2 < \dots < j_k \leq n} t_{j_1} t_{j_2} \dots t_{j_k}$$

and  $p_k(t_1, \dots, t_n)$  be the  $k$ th power sum symmetric polynomial

$$p_k(t_1, \dots, t_n) := \sum_{i=1}^n t_i^k.$$

Let  $\mathbf{X} = (X_i)_{i=1}^n$  be  $n$   $\overline{\mathbb{F}_q}$ -valued random variables such that  $e_1(\mathbf{X}), \dots, e_n(\mathbf{X})$  are independent and uniformly distributed on  $\mathbb{F}_q$ . By Newton's identities,  $p_k(\mathbf{X})$  must also be  $\mathbb{F}_q$ -valued. We ask, is  $p_k(\mathbf{X})$  close to uniform?

Here is one way to construct such a sequence  $\mathbf{X}$ . Taking  $a_1, \dots, a_n$  to be independent uniform random variables on  $\mathbb{F}_q$ , the polynomial  $T^n - a_1 T^{n-1} + a_2 T^{n-2} \mp \dots + (-1)^n a_n$  is distributed uniformly among  $\mathcal{M}_{n,q} \subseteq \mathbb{F}_q[T]$ , the subset of monic polynomials of degree  $n$ . Setting  $(X_i)_{i=1}^n$  to be its  $n$  roots in some order, we have that  $e_j(\mathbf{X}) = a_j$  satisfy the conditions above, and  $p_i(f)$  defined in (1.1) coincides with  $p_i(\mathbf{X})$ . We prove the following.

**Theorem 1.4.** *Let  $n \geq 1$ . Let  $k = k(n)$  be a positive integer. Suppose  $\mathbf{X} = (X_i)_{i=1}^n$  are  $n$  random variables such that  $(e_1(\mathbf{X}), \dots, e_n(\mathbf{X}))$  has the uniform distribution on  $(\mathbb{F}_q)^n$ . Then:*

1. *If  $\log k = o(n^2)$  then the distribution of  $p_k(\mathbf{X})$  tends to the uniform distribution on  $\mathbb{F}_q$ .*
2. *Let  $a_1, \dots, a_k \in \mathbb{F}_q$ . If  $\log k = o(n)$ ,  $a_k \neq 0$  and  $\text{char}(\mathbb{F}_q) \nmid k$  then the distribution of  $\sum_{i=1}^k a_i p_i(\mathbf{X})$  tends to the uniform distribution on  $\mathbb{F}_q$ .*
3. *If the sum in (1.2) diverges then the distribution of  $p_k(\mathbf{X})$  tends to the uniform distribution on  $\mathbb{F}_q$ .*

Theorem 1.4 is about  $p_k(f)$  when we choose  $f$  uniformly at random from  $\mathcal{M}_{n,q}$ , while Theorems 1.1–1.3 are about  $p_k(f)$  for a polynomial  $f$  drawn from the space of possible characteristic polynomials of a matrix from  $\text{GL}_n(\mathbb{F}_q)$  endowed with the uniform measure. As proved in [12, Thm. 1.4], the total variation distance of the law of the first  $k$  next-to-leading coefficients of  $\det(I_n T - g)$  from the uniform distribution on  $\mathbb{F}_q^k$  tends to 0 as  $n \rightarrow \infty$  for  $k$  as large as  $n - o(\log n)$ , so the setup of Theorem 1.4 is not that different from the setup of Theorems 1.1–1.3.

### 1.3 Cancellation in character sums of high conductor

Given  $n \geq 0$  we denote by  $\mathcal{M}_{n,q} \subseteq \mathbb{F}_q[T]$  the subset of monic polynomials of degree  $n$ . We denote by  $\mathcal{M}_q = \cup_{n \geq 0} \mathcal{M}_{n,q}$  the set of monic polynomials in  $\mathbb{F}_q[T]$ . We denote by  $\mathcal{P}_{n,q} \subseteq \mathcal{M}_{n,q}$  the set of irreducible polynomials of degree  $n$  and let  $\mathcal{P}_q = \cup_{n \geq 0} \mathcal{P}_{n,q}$ . Throughout, the letter  $P$  is reserved for elements of  $\mathcal{P}_q$ .

Given  $k \geq 1$  and a nontrivial additive character  $\psi: \mathbb{F}_q \rightarrow \mathbb{C}$ , we define a function  $\chi_{k,\psi}: \mathcal{M}_q \rightarrow \mathbb{C}$  by

$$\chi_{k,\psi}(f) := \begin{cases} \psi(p_{-k}(f)), & \text{if } (f, T) = 1, \\ 0, & \text{otherwise.} \end{cases}$$

*Remark 1.1.* If  $i$  is a negative integer, then  $p_i(f)$ , as defined in (1.1), is well defined as long as  $T \nmid f$ . Moreover, the usual properties are preserved:  $p_i(fg) = p_i(f) + p_i(g)$  if  $T \nmid fg$ ; if  $f(T) = \deg(I_n T - g)$  then  $p_i(f) = \text{Tr}(g^i)$ . When  $T \nmid f$ ,  $p_i(f) = p_{-i}(f(1/T)T^{\deg f}/f(0))$ , where  $f(1/T)T^{\deg f}/f(0)$  is a monic polynomial, proving  $p_i(f) \in \mathbb{F}_q$  as well.

We show in Lemma 2.1 that the function  $\chi_{k,\psi}$  is a primitive Dirichlet character modulo  $T^{k+1}$ . The following theorem is the main component behind the proof of Theorem 1.1.

**Theorem 1.5.** *We have, uniformly for  $n \geq 1$  and  $k \geq 1$ ,*

$$q^{-n} \sum_{f \in \mathcal{M}_{n,q}} \chi_{k,\psi}(f) \ll \frac{1 + \log_q n + \sqrt{\log_q k}}{n}.$$

*In particular, we have cancellation when  $\log k = o(n^2)$  as  $n \rightarrow \infty$ .*

Note that the range  $\log k = o(n^2)$  is optimal because for  $k = \prod_{i=1}^n (q^i - 1)$  we have  $\gamma^{-k} = 1$  for every  $\gamma \in \cup_{i=1}^n \mathbb{F}_{q^i}^\times$  and thus  $p_{-k}(f) \equiv n$  on  $\{f \in \mathcal{M}_{n,q} : T \nmid f\}$ . Hence, there is no cancellation for such  $k$ . We do not know if cancellation for  $\log k = o(n^2)$  persists if we restrict to  $\mathcal{P}_{n,q}$  instead of  $\mathcal{M}_{n,q}$ ; the proof of Theorem 1.5 relies on summing over *all* degree- $n$  polynomials.

It is instructive to compare Theorem 1.5 to results about character sums for integers and polynomials. We switch temporarily to the integer setting. Let  $\chi$  be a nonprincipal Dirichlet character modulo  $m$  and consider the sum  $S_{x,\chi} := \sum_{n \leq x} \chi(n)$  as  $x \rightarrow \infty$ . Montgomery and Vaughan [22, Lem. 2] proved under the generalized Riemann hypothesis for  $L(s, \chi)$  that if  $m \geq x$  and  $y \in [(\log m)^4, x]$  is a parameter, then

$$S_{x,\chi} = \sum_{\substack{n \leq x \\ p|n \implies p \leq y}} \chi(n) + O(xy^{-1/2}(\log m)^4). \quad (1.3)$$

Recall  $\sum_{n \leq x: p|n \implies p \leq y} 1$  is  $o(x)$  if  $\log x / \log y \rightarrow \infty$  [8, eq. (1.9)]. Taking  $y = (\log m)^9$ , we see that  $S_{x,\chi}$  exhibits cancellation if  $\log \log m = o(\log x)$ . Granville and Soundararajan improved the error term in (1.3) and also showed that this range is optimal [13, Cor. A], in the sense that for any given  $A > 0$  and for any prime  $m$ , there exists a nonprincipal character  $\chi \bmod m$  with  $|S_{x,\chi}| \gg_A x$ , where  $x = \log^A m$ .

Much less is known unconditionally. Burgess proved  $S_{x,\chi}$  exhibits cancellation when  $m \leq x^{3-\varepsilon}$ , and this can be extended to  $m \leq x^{4-\varepsilon}$  if  $m$  is cubefree [4, 5]. When the conductor is smooth, better results exist [17, Ch. 12]. In particular, if  $m = p^r$ , then Banks and Shparlinski [1] showed one has cancellation in the range  $\log m = o((\log x)^{3/2})$  (if  $r \geq C$  and  $p \leq x^c$  for some absolute  $C > 0$  and  $c > 0$ ); this improved earlier work of Postnikov [17, Thm. 12.16]. It is worthwhile to recall  $\sum_{n \leq x} n^{it}$  exhibits cancellation when  $\log(|t| + 2) = o((\log x)^{3/2})$ , a result due to Vinogradov [17, Cor. 8.26]. The bound we prove for the sum of  $f \mapsto \chi_{k,\psi}(f)$  also holds for the function  $f \mapsto \psi(p_k(f))$ , which is an analogue of  $n \mapsto n^{it}$ , see Remark 2.1 and [12, Lem. 2.2].

Now let us return to the polynomial setting. The generalized Riemann hypothesis in  $\mathbb{F}_q[T]$  is a seminal theorem due to A. Weil [28]. Let  $\chi$  be a nonprincipal Dirichlet character modulo a polynomial  $Q \in \mathbb{F}_q[T]$ . In [2, Thm. 3], Bhowmick and L  adapted (1.3) to the polynomial setting, proving unconditionally that

$$\sum_{f \in \mathcal{M}_{n,q}} \chi(f) = \sum_{\substack{f \in \mathcal{M}_{n,q} \\ P|f \implies \deg P \leq m}} \chi(f) + O(q^{n-\frac{m}{2}} \deg Q) \quad (1.4)$$

holds for  $m \in [2 \log_q \deg Q, n]$ . Taking  $m = \lceil 4 \log_q \deg Q \rceil$  shows that, as  $n \rightarrow \infty$ ,  $\sum_{f \in \mathcal{M}_{n,q}} \chi(f)$  exhibits cancellation when  $\log_q \deg Q = o(n)$ . (For a self-contained bound on the sum in the right-hand side of (1.4), see Lemma 3.5.)

The range  $\log k = o(n^2)$  in Theorem 1.5 is far beyond the ranges that the generalized Riemann hypothesis implies, and heavily exploits a special symmetry satisfied by  $\chi_{k,\psi}$ , which is shown in Lemma 3.8. We are not aware of any other *explicit* family of characters, in either integers or polynomials, where the generalized Riemann hypothesis implies cancellation when the conductor exceeds the Montgomery–Vaughan range  $\log \log m = o(\log x)$  in  $\mathbb{Z}$  or  $\log \deg Q = o(n)$  in  $\mathbb{F}_q[T]$ .

In §A we prove a new bound on general character sums in function fields which is not used in the paper.

## 1.4 Structure of the paper

In Section 2, we discuss the connection between the distribution of traces and character sums in more detail, and bound the total variation distance between our respective distributions and the uniform distribution by corresponding character sums. By doing so, we reduce Theorem 1.1 and the first part of Theorem 1.4 to Theorem 1.5, and Theorem 1.2 and the second part of Theorem 1.4 to obtaining bounds for the character sum in (1.4). The latter is quite straightforward as shown in Lemma 2.2, while the former requires more careful treatment. We prove Theorem 1.5 in Section 3; we consider this to be the technical part of the paper. We start with observing that there is underlying symmetry in sums of  $\chi_{k,\psi}$  against primes. To see this, in Lemma 3.8, we prove

$$\sum_{f \in \mathcal{M}_{n,q}} \Lambda(f) \chi_{k,\psi}(f) = \sum_{f \in \mathcal{M}_{n,q}} \Lambda(f) \chi_{k',\psi}(f),$$

where  $\Lambda(f)$  is the function field von Mangoldt function, and  $k' = \gcd(k, q^n - 1)$ . This identity allows us to improve the Weil bound

$$\left| \sum_{f \in \mathcal{M}_{n,q}} \Lambda(f) \chi_{k,\psi}(f) \right| \leq q^{\frac{n}{2}} k$$

to

$$\left| \sum_{f \in \mathcal{M}_{n,q}} \Lambda(f) \chi_{k,\psi}(f) \right| \leq q^{\frac{n}{2}} k' = q^{\frac{n}{2}} \gcd(k, q^n - 1) \quad (1.5)$$

as shown in Corollary 3.9. This is helpful when  $\gcd(k, q^n - 1)$  is small. Motivated by Montgomery and Vaughan [22, Lem. 2] given  $S \subseteq \{1, \dots, n\}$ , we write

$$q^{-n} \sum_{f \in \mathcal{M}_{n,q}} \chi_{k,\psi}(f) = q^{-n} \sum_{\substack{f \in \mathcal{M}_{n,q} \\ P|f \implies \deg P \notin S}} \chi_{k,\psi}(f) + q^{-n} \sum_{\substack{f \in \mathcal{M}_{n,q} \\ \exists P|f \text{ such that } \deg P \in S}} \chi_{k,\psi}(f). \quad (1.6)$$

In our case, choosing

$$S = \{ \lceil 12 \log_q n \rceil \leq d \leq n : \gcd(k, q^d - 1) < q^{d/3} \}$$

allows us to bound the second sum on the right-hand side of (1.6) using (1.5) (see Lemma 3.4) while a sieve bound (Lemma 3.5) bounds the first sum on the right-hand side of (1.6). This strategy ends up proving the following criterion.

**Proposition 1.6.** *Let  $n \geq 1$  and  $k \geq 1$ . Let  $\psi: \mathbb{F}_q \rightarrow \mathbb{C}^\times$  be a nontrivial additive character. We have*

$$q^{-n} \sum_{f \in \mathcal{M}_{n,q}} \chi_{k,\psi}(f) \ll n^{-1} + \exp \left( - \sum_{\substack{12 \log_q n < d \leq n \\ \gcd(k, q^d - 1) < q^{d/3}}} d^{-1} \right). \quad (1.7)$$

In particular, a sufficient criterion for cancellation is that the sum in (1.2) diverges as  $n \rightarrow \infty$ .

In Section 2, we reduce Theorem 1.3 and the third part of Theorem 1.4 to Proposition 1.6. To deduce Theorem 1.5 from Proposition 1.6, we prove a sharp upper bound on  $\sum_{L < d \leq 2L} \log \gcd(k, q^d - 1)$  in Lemma 3.6, that may be of independent interest.

The criterion above is particularly interesting because it allows us to exhibit cancellation on the left-hand side of (1.7) for arbitrarily large  $k$ , for example, whenever  $k$  is a prime, or a product of a bounded number of primes, the sum in (1.2) diverges simply because the following shorter sum does:

$$\sum_{\substack{12 \log_q n < d \leq n \\ \gcd(k, q^d - 1) = 1}} d^{-1}.$$

*Remark 1.2.* Proposition 1.6, and hence Theorem 1.5, apply as is to  $\sum_{f \in \mathcal{M}_{n,q}} \mu(f) \chi_{k,\psi}(f)$  where  $\mu$  is the Möbius function, see Remark 3.3. Recall  $\mu$  is multiplicative with  $\mu(P) = -1$  and  $\mu(P^e) = 0$  for  $e \geq 2$ . However, we do not know that  $\log k = o(n^2)$  is optimal in this case.

*Remark 1.3.* For every  $\varepsilon > 0$ , there are examples where  $\log k \sim \varepsilon n^2$  for which Proposition 1.6 does not yield cancellation, for example,  $k = \prod_{i=1}^{\lfloor \varepsilon' n \rfloor} (q^i - 1)$  where taking logarithm shows  $\varepsilon' > 0$  is defined via  $\varepsilon = \varepsilon'^2 (\log q)/2$ .

*Remark 1.4.* It is natural to ask whether analogues of Theorems 1.1 and 1.2 hold for other groups. Consider for instance  $U(n, q)$ , the unitary group over  $\mathbb{F}_{q^2}$ . Then, as in Section 2, we are led to consider the sums  $\sum_{f \in \mathcal{M}_{n,q^2}} P_U(f) \chi(f)$  where  $P_U(f) := \mathbb{P}_{g \in U(n,q)}(\det(I_n T - g) = f)$  is a function studied in [12, §5] and  $\chi$  is a Dirichlet character depending on  $k$  (if we are in the situation of Theorem 1.1) or on  $a_1, \dots, a_k$  (if we are in the situation of Theorem 1.2). The support of  $P_U$  is restricted since the characteristic polynomial of any matrix from  $U(n, q)$  is *unitary self-reciprocal*, that is, it belongs to

$$\mathcal{M}_{n,q^2}^{usr} := \left\{ f \in \mathcal{M}_{n,q^2} : f(0) \neq 0, f(T) = T^{\deg f} \frac{\sigma(f(1/T))}{\sigma(f(0))} \right\}, \quad \sigma\left(\sum c_i T^i\right) := \sum c_i^q T^i.$$

Using the strategy in Section 2 together with [12, Thm. 5.10], the problem ultimately reduces to bounding character sums over  $\mathcal{M}_{n,q^2}^{usr}, \sum_{f \in \mathcal{M}_{n,q^2}^{usr}} \chi(f)$ . We expect that it is possible to adapt our proofs to bound such sums with additional work, although various complications arise, and this problem deserves further study. For a review of character sums over such sets, we refer the reader to [12, §5].

## 2 Reductions

### 2.1 An involution

Given  $f \in \mathcal{M}_q$  we use the notation  $p_i(f)$  introduced in (1.1). Given  $a_1, \dots, a_k \in \mathbb{F}_q$  and an additive character  $\psi: \mathbb{F}_q \rightarrow \mathbb{C}^\times$ , define a function  $\xi_{\mathbf{a},\psi}: \mathcal{M}_q \rightarrow \mathbb{C}$  by

$$\xi_{\mathbf{a},\psi}(f) := \psi\left(\sum_{i=1}^k p_i(f)a_i\right).$$

The function  $\xi_{\mathbf{a},\psi}$ , sometimes called a short interval character, is closely related to a certain Dirichlet character modulo  $T^{k+1}$  as observed by Hayes [16, pp. 115–116] and Keating and Rudnick [21, §4.1][20, p. 389]. Let us explain this idea. We let  $\mathcal{M}_q^{gl} := \{f \in \mathcal{M}_q : (f, T) = 1\}$  (this notation, borrowed from [12, §3.1], is motivated by the fact that the characteristic polynomial of any matrix from  $\cup_{n \geq 0} \mathrm{GL}_n(\mathbb{F}_q)$  lies in  $\mathcal{M}_q^{gl}$ ). We define an involution  $\iota$  on  $\mathcal{M}_q^{gl}$  by

$$\iota(f) := f(1/T)T^{\deg f}/f(0).$$

If we factorize  $f \in \mathcal{M}_q^{gl}$  as  $f(T) = \prod_{j=1}^{\deg f} (T - \lambda_j)$  over  $\overline{\mathbb{F}_q}$ , we find that

$$\iota(f) = \prod_{j=1}^{\deg f} (T - \lambda_j^{-1})$$

so  $p_k(f) = p_{-k}(\iota(f))$  (see Remark 1.1 for the definition of  $p_{-k}$ ). For any function  $\alpha: \mathcal{M}_q \rightarrow \mathbb{C}$ , we define  $\iota_\alpha: \mathcal{M}_q \rightarrow \mathbb{C}$  by

$$\iota_\alpha(f) := \begin{cases} \alpha(\iota(f)) & \text{if } f \in \mathcal{M}_q^{gl}, \\ 0 & \text{if } f \notin \mathcal{M}_q^{gl}. \end{cases}$$

We say that a function  $\alpha: \mathcal{M}_q \rightarrow \mathbb{C}$  is completely multiplicative if  $\alpha(fg) = \alpha(f)\alpha(g)$  holds for all  $f, g \in \mathcal{M}_q$  and  $\alpha(1) = 1$ . If  $\alpha$  is completely multiplicative, then so is  $\iota_\alpha$  because  $\iota$  is (when extended to  $\mathcal{M}_q$  via  $\iota(T) = 0$ ). Given an arithmetic function  $\alpha: \mathcal{M}_q \rightarrow \mathbb{C}$ , we use the notation

$$S(n, \alpha) := \sum_{f \in \mathcal{M}_{n,q}} \alpha(f).$$

Letting

$$\chi_0(f) := \mathbf{1}_{T \nmid f},$$

we have

$$S(n, \alpha \cdot \chi_0) = S(n, \iota_\alpha) \tag{2.1}$$

and, since  $S(n, \alpha \cdot \chi_0) = S(n, \alpha) - \alpha(T)S(n-1, \alpha)\mathbf{1}_{n \geq 1}$  holds for completely multiplicative  $\alpha$ ,

$$S(n, \alpha) = \sum_{i=0}^n S(i, \alpha \cdot \chi_0) \alpha(T)^{n-i} = \sum_{i=0}^n S(i, \iota_\alpha) \alpha(T)^{n-i}. \tag{2.2}$$

A variant of the following lemma was established in [12, Lem. 2.2].

**Lemma 2.1.** *Let  $a_1, \dots, a_k \in \mathbb{F}_q$  with  $a_k \neq 0$  and  $\mathrm{char}(\mathbb{F}_q) \nmid k$ . Let  $\psi: \mathbb{F}_q \rightarrow \mathbb{C}^\times$  be a nontrivial additive character, and let  $\alpha = \xi_{\mathbf{a},\psi}$ . Then,  $\iota_\alpha$  is a primitive Dirichlet character modulo  $T^{k+1}$ .*

*Proof.* We know  $\iota_\alpha$  is completely multiplicative, vanishes at multiples of  $T$  and  $\iota_\alpha(1) = 1$ . It remains to show that it only depends on the residue of the input modulo  $T^{k+1}$  and that it is primitive. Newton's identities yield that  $p_i(f)$  is a function of the  $i$  first next-to-leading coefficients of  $f$ , where the  $j$ th ( $j \geq 1$ ) next-to-leading coefficient of  $T^n + \sum_{i=0}^{n-1} a_i T^i$  is defined to be  $a_{n-j}$  if  $j \leq n$  and to be 0 otherwise. Hence,  $\alpha = \xi_{\mathbf{a},\psi}(f)$  only depends on the  $k$  first next-to-leading coefficients of  $f$ . Since by definition  $\iota(f)$  reverses the coefficients of  $f$  and normalizes by  $f(0)$ , it follows that  $\iota_\alpha(f)$  depends only on the last  $k+1$  coefficients of  $f$ , that is, on  $f \bmod T^{k+1}$ . Finally,  $\iota_\alpha(T^{k+1} - c) = 1$  for every  $c \in \mathbb{F}_q^\times$  while a short computation using Newton's identities shows  $\iota_\alpha(T^k - c) = \psi(ka_k/c)$  is not equal to 1 for suitable  $c$ .  $\square$

For  $k \geq 1$  and an additive character  $\psi: \mathbb{F}_q \rightarrow \mathbb{C}$ , with a slight abuse of notation, let

$$\xi_{k,\psi}(f) := \psi(p_k(f)),$$

which coincides with  $\xi_{\mathbf{a},\psi}$  for  $a_1 = \dots = a_{k-1} = 0$ ,  $a_k = 1$ . By definition, we have

$$\chi_{k,\psi} = \iota_{\xi_{k,\psi}} = \iota_{\xi_{k,\psi} \cdot \chi_0}.$$

*Remark 2.1.* We have  $S(n, \xi_{k,\psi}) = \sum_{i=0}^n S(i, \chi_{k,\psi})$  by (2.2). In particular,  $q^{-n}S(n, \chi_{k,\psi}) \ll (1 + \log_q n + \sqrt{\log_q k})/n$  implies  $q^{-n}S(n, \xi_{k,\psi}) \ll (1 + \log_q n + \sqrt{\log_q k})/n$  and vice versa.

## 2.2 Reduction of Theorem 1.4 to character sum estimates

Recall that  $\mathbf{X} = (X_i)_{i=1}^n \in \overline{\mathbb{F}_q}^n$  are  $n$  random variables such that  $e_1(\mathbf{X}), \dots, e_n(\mathbf{X})$  are independent and uniformly distributed on  $\mathbb{F}_q$ , and that one can take  $X_i$  to be the zeros of a polynomial  $f$  chosen uniformly at random from  $\mathcal{M}_{n,q}$ .

Let  $a_1, \dots, a_k \in \mathbb{F}_q$  and denote by  $\widehat{\mathbb{F}_q}$  the group of  $q$  additive characters from  $\mathbb{F}_q$  to  $\mathbb{C}^\times$ , with  $\psi_0$  denoting the trivial character. By Fourier analysis on  $\mathbb{F}_q$ , given  $x \in \mathbb{F}_q$ ,

$$\begin{aligned} \mathbb{P}\left(\sum_{i=1}^k a_i p_i(\mathbf{X}) = x\right) &= \mathbb{P}_{f \in \mathcal{M}_{n,q}}\left(\sum_{i=1}^k a_i p_i(f) = x\right) \\ &= q^{-1} \sum_{\psi \in \widehat{\mathbb{F}_q}} \overline{\psi}(x) \mathbb{E}_{f \in \mathcal{M}_{n,q}} \psi\left(\sum_{i=1}^k a_i p_i(f)\right) = q^{-1} \sum_{\psi \in \widehat{\mathbb{F}_q}} q^{-n} \overline{\psi}(x) S(n, \xi_{\mathbf{a},\psi}). \end{aligned}$$

Since  $S(n, \xi_{\mathbf{a},\psi_0}) = q^n$ , the triangle inequality implies

$$\sum_{x \in \mathbb{F}_q} \left| \mathbb{P}\left(\sum_{i=1}^k a_i p_i(\mathbf{X}) = x\right) - q^{-1} \right| \leq \sum_{\psi_0 \neq \psi \in \widehat{\mathbb{F}_q}} |q^{-n} S(n, \xi_{\mathbf{a},\psi})|. \quad (2.3)$$

By (2.3), (2.2), and the triangle inequality,

$$\sum_{x \in \mathbb{F}_q} \left| \mathbb{P}\left(\sum_{i=1}^k a_i p_i(\mathbf{X}) = x\right) - q^{-1} \right| \leq \sum_{j=0}^n q^{j-n} \sum_{\psi_0 \neq \psi \in \widehat{\mathbb{F}_q}} |q^{-j} S(j, \iota_{\xi_{\mathbf{a},\psi}})|. \quad (2.4)$$

A special case of (2.4) is

$$\sum_{x \in \mathbb{F}_q} \left| \mathbb{P}(p_k(\mathbf{X}) = x) - q^{-1} \right| \leq \sum_{j=0}^n q^{j-n} \sum_{\psi_0 \neq \psi \in \widehat{\mathbb{F}_q}} |q^{-j} S(j, \chi_{k,\psi})|. \quad (2.5)$$

The first part of Theorem 1.4 is immediate from (2.5) and Theorem 1.5. The third part follows from (2.5) and Proposition 1.6. For the second part of Theorem 1.4, we use (2.4) and observe  $\iota_{\xi_{\mathbf{a},\psi}}$  is a nonprincipal Dirichlet character modulo  $T^{k+1}$  by Lemma 2.1, and the result follows from applying the following lemma.

**Lemma 2.2** (Bhowmick–Lê). *Let  $\chi: \mathcal{M}_q \rightarrow \mathbb{C}$  be a nonprincipal Dirichlet character modulo  $Q$ . We have*

$$q^{-n} S(n, \chi) \ll \frac{1 + \log_q(1 + \deg Q)}{n + 1}.$$

Lemma 2.2 follows from (1.4) with  $m = \lceil 2 \log_q(n(1 + \deg Q)) \rceil$  (see Section 3.3 for details).

### 2.3 Reductions of Theorems 1.1–1.3 to character sum estimates

We define an arithmetic function  $P_{\text{GL}}: \mathcal{M}_q \rightarrow \mathbb{C}$  as follows. If  $f \in \mathcal{M}_q$  and  $n = \deg f$ , then

$$P_{\text{GL}}(f) := \mathbb{P}_{g \in \text{GL}_n(\mathbb{F}_q)}(\det(I_n T - g) = f),$$

where  $\text{GL}_n(\mathbb{F}_q)$  is endowed with the uniform measure. It follows from the work of Reiner [25] and Gerstenhaber [11] that  $P_{\text{GL}}$  is multiplicative (in the sense that  $P_{\text{GL}}(fg) = P_{\text{GL}}(f)P_{\text{GL}}(g)$  when  $\gcd(f, g) = 1$ ) and supported on  $\mathcal{M}_q^{gl}$ . Their works also show that on prime powers it is given by  $P_{\text{GL}}(T^e) = 0$  and

$$P_{\text{GL}}(P^e) = q^{-e \deg P} \prod_{i=1}^e (1 - q^{-i \deg P})^{-1}, \quad (2.6)$$

where  $P \in \mathcal{P}_q \setminus \{T\}$  and  $e \geq 1$  (cf. [12, Thm. 3.3]). The following identity is a quick consequence of (2.6) (it can also be derived directly from [12, Thm. 3.4]). Recall  $|f| = q^{\deg f}$ .

**Lemma 2.3.** *We have  $P_{\text{GL}} = \alpha_1 * \alpha_2$  where  $\alpha_1(f) = |f|^{-1} \cdot \mathbf{1}_{T \nmid f}$  and  $\alpha_2(f) = P_{\text{GL}}(f)/|f|$ .*

Let  $a_1, \dots, a_k \in \mathbb{F}_q$ . By Fourier analysis on  $\mathbb{F}_q$ , given  $x \in \mathbb{F}_q$ , we have

$$\mathbb{P}_{g \in \text{GL}_n(\mathbb{F}_q)}\left(\sum_{i=1}^k a_i \text{Tr}(g^i) = x\right) = q^{-1} \sum_{\psi \in \widehat{\mathbb{F}_q}} \mathbb{E}_{g \in \text{GL}_n(\mathbb{F}_q)} \bar{\psi}(x) \psi\left(\sum_{i=1}^k a_i \text{Tr}(g^i)\right).$$

The trivial character  $\psi_0$  contributes  $1/q$  to the right-hand side. Recall  $p_k(f)$  ( $f \in \mathcal{M}_q$ ) is defined in (1.1). Since  $p_i(\deg(I_n T - g)) = \text{Tr}(g^i)$ ,

$$\mathbb{P}_{g \in \text{GL}_n(\mathbb{F}_q)}\left(\sum_{i=1}^k a_i \text{Tr}(g^i) = x\right) - q^{-1} = q^{-1} \sum_{\psi_0 \neq \psi \in \widehat{\mathbb{F}_q}} \bar{\psi}(x) \sum_{f \in \mathcal{M}_{n,q}} P_{\text{GL}}(f) \xi_{\mathbf{a},\psi}(f).$$

Summing over  $x$  and using the triangle inequality gives

$$\sum_{x \in \mathbb{F}_q} \left| \mathbb{P}_{g \in \text{GL}_n(\mathbb{F}_q)}\left(\sum_{i=1}^k a_i \text{Tr}(g^i) = x\right) - q^{-1} \right| \leq \sum_{\psi_0 \neq \psi \in \widehat{\mathbb{F}_q}} |S(n, P_{\text{GL}} \cdot \xi_{\mathbf{a},\psi})|. \quad (2.7)$$

In [12, Thm. 3.6], the first author and Rodgers showed that

$$|S(n, P_{\text{GL}} \cdot \xi_{\mathbf{a},\psi})| \leq q^{-\frac{n^2}{2k} + O_q(n)}.$$

For fixed  $k$  this is essentially optimal up to constants, because  $|S(n, P_{\text{GL}} \cdot \xi_{\mathbf{a},\psi})|$  cannot decay faster than exponentially in  $n^2$  due to  $|\text{GL}_n(\mathbb{F}_q)| = q^{\Theta(n^2)}$ . In this paper, we focus on the range of cancellation rather than the rate of cancellation, and in this aspect, we can do better. Observe that from Lemma 2.3,

$$S(n, P_{\text{GL}} \cdot \xi_{\mathbf{a},\psi}) = q^{-n} \sum_{i+j=n} \sum_{f \in \mathcal{M}_{i,q}, T \nmid f} \xi_{\mathbf{a},\psi}(f) \sum_{g \in \mathcal{M}_{j,q}} P_{\text{GL}}(g) \xi_{\mathbf{a},\psi}(g).$$

Since  $P_{\text{GL}}$  is a probability measure on  $\mathcal{M}_{j,q}$ , the triangle inequality yields

$$|S(n, P_{\text{GL}} \cdot \xi_{\mathbf{a},\psi})| \leq q^{-n} \sum_{i=0}^n |S(i, \xi_{\mathbf{a},\psi} \cdot \chi_0)|. \quad (2.8)$$

From (2.8), (2.7), and (2.1),

$$\sum_{x \in \mathbb{F}_q} \left| \mathbb{P}_{g \in \text{GL}_n(\mathbb{F}_q)}\left(\sum_{i=1}^k a_i \text{Tr}(g^i) = x\right) - q^{-1} \right| \leq q^{-n} \sum_{\psi_0 \neq \psi \in \widehat{\mathbb{F}_q}} \sum_{i=0}^n |S(i, \iota_{\xi_{\mathbf{a},\psi}})|. \quad (2.9)$$



Similarly to the symmetric function case, we see that Theorem 1.2 follows from (2.9) and Lemma 2.2.

In the special case  $a_1 = \dots = a_{k-1}$  and  $a_k = 1$ , we have

$$\sum_{x \in \mathbb{F}_q} \left| \mathbb{P}_{g \in \text{GL}_n(\mathbb{F}_q)}(\text{Tr}(g^k) = x) - q^{-1} \right| \leq q^{-n} \sum_{\psi_0 \neq \psi \in \widehat{\mathbb{F}_q}} \sum_{i=0}^n |S(i, \chi_{k, \psi})|. \quad (2.10)$$

Theorem 1.1 follows from (2.10) and Theorem 1.5. Theorem 1.3 follows from (2.10) and Proposition 1.6.

### 3 Proof of Theorem 1.5

The von Mangoldt function  $\Lambda: \mathcal{M}_q \rightarrow \mathbb{C}$  is defined as

$$\Lambda(f) = \begin{cases} \deg P & \text{if } f = P^k, P \in \mathcal{P}_q, k \geq 1, \\ 0 & \text{otherwise.} \end{cases}$$

Gauss' identity [27, Thm. 2.2] states that

$$\sum_{f \in \mathcal{M}_{n,q}} \Lambda(f) = \sum_{d|n} |\mathcal{P}_{d,q}| d = q^n, \quad (3.1)$$

and is known to imply the following.

**Lemma 3.1.** [23, Lem. 4] *We have  $q^n/n - 2q^{n/2}/n \leq |\mathcal{P}_{n,q}| \leq q^n/n$  for  $n \geq 1$ .*

Given a Dirichlet character  $\chi$  its  $L$ -function is defined as

$$L(u, \chi) = \sum_{f \in \mathcal{M}_q} \chi(f) u^{\deg f} = \prod_{P \in \mathcal{P}_q} (1 - \chi(P) u^{\deg P})^{-1}, \quad (3.2)$$

which converges absolutely for  $|u| < 1/q$ . If  $\chi$  is a nonprincipal character modulo  $Q$ , then  $L(u, \chi)$  is a polynomial of degree at most  $\deg Q - 1$  as follows from the orthogonality relations for  $\chi$ . We set

$$d(\chi) := \deg L(u, \chi) < \deg Q.$$

**Theorem 3.2** (Weil's RH [28]). *Let  $\chi$  be a nonprincipal Dirichlet character. Factoring  $L(u, \chi)$  as*

$$L(u, \chi) = \prod_{i=1}^{d(\chi)} (1 - \gamma_i u), \quad (3.3)$$

*we have  $|\gamma_i| \in \{1, \sqrt{q}\}$  for all  $1 \leq i \leq d(\chi)$ .*

**Lemma 3.3.** *Let  $\chi$  be a nonprincipal Dirichlet character, then*

$$\left| \sum_{f \in \mathcal{M}_{n,q}} \chi(f) \Lambda(f) \right| \leq q^{\frac{n}{2}} d(\chi). \quad (3.4)$$

*Proof.* We take the logarithmic derivative of the Euler product (3.2) and (3.3) and compare coefficients to obtain

$$\sum_{f \in \mathcal{M}_{n,q}} \chi(f) \Lambda(f) = - \sum_{i=1}^{d(\chi)} \gamma_i^n$$

for all  $n \geq 1$ . Then, Theorem 3.2 implies the lemma via the triangle inequality.  $\square$

Let  $\alpha: \mathcal{M}_q \rightarrow \mathbb{C}$  be a completely multiplicative function. Given any subset  $T$  of the positive integers  $\mathbb{N}$ , we define

$$F_T(u, \alpha) := \sum_{\substack{f \in \mathcal{M}_q \\ P|f \Rightarrow \deg P \in T}} \alpha(f) u^{\deg f} = \prod_{P \in \mathcal{P}_q, \deg P \in T} (1 - \alpha(P) u^{\deg P})^{-1}.$$

Let us write  $[u^m]F$  for the  $m$ th coefficient of a power series  $F$ . Our starting point is the identity

$$\begin{aligned} \sum_{f \in \mathcal{M}_{n,q}} \alpha(f) &= [u^n] F_{\mathbb{N}}(u, \alpha) = [u^n] F_{T^c}(u, \alpha) F_T(u, \alpha) \\ &= [u^n] F_{T^c}(u, \alpha) + [u^n] F_{T^c}(u, \alpha) (F_T(u, \alpha) - 1) \\ &= q^{-n} \sum_{\substack{f \in \mathcal{M}_{n,q} \\ P|f \Rightarrow \deg P \notin T}} \alpha(f) + q^{-n} \sum_{\substack{f \in \mathcal{M}_{n,q} \\ \exists P|f \text{ such that } \deg P \in T}} \alpha(f), \end{aligned} \tag{3.5}$$

where  $T^c = \mathbb{N} \setminus T$  is the complement of  $T$ .

### 3.1 Montgomery and Vaughan's estimate

To bound the second term in (3.5), we prove the following lemma generalizing an estimate of Montgomery and Vaughan [22, Lem. 2] in the polynomial setting.

**Lemma 3.4.** *Let  $\alpha: \mathcal{M}_q \rightarrow \mathbb{C}$  be a 1-bounded completely multiplicative function, meaning  $|\alpha(f)| \leq 1$  for all  $f \in \mathcal{M}_q$ . Let  $n \geq 1$  and  $S \subseteq \{1, \dots, n\}$  be a set of positive integers, and  $S^c = \{1, \dots, n\} \setminus S$  be its complement. Let  $s_0 = \min_{s \in S} s$ . Then,*

$$q^{-n} \sum_{\substack{f \in \mathcal{M}_{n,q} \\ \exists P|f \text{ such that } \deg P \in S}} \alpha(f) \ll \exp(A_1)(\exp(A_2) - 1),$$

where

$$A_1 = \sum_{d \in S^c} \frac{1}{d}, \quad A_2 = \sum_{d \in S} \frac{q^{-d}}{d} \left| \sum_{f \in \mathcal{M}_{d,q}} \alpha(f) \Lambda(f) \right| + O(q^{-\frac{s_0}{2}}/s_0).$$

*Remark 3.1.* In [2] Lemma 3.4 is proved in the special case  $S = \{m+1, m+2, m+3, \dots, n\}$  and  $\alpha$  being a Dirichlet character. For this choice, one immediately recovers (1.4) from (3.4).

*Proof.* Using identity (3.5) and noticing that we can restrict the set of degrees to  $\{1, \dots, n\}$ , we have

$$X := q^{-n} \sum_{\substack{f \in \mathcal{M}_{n,q} \\ \exists P|f \text{ such that } \deg P \in S}} \alpha(f) = [u^n] F_{S^c}(u, \alpha) (F_S(u, \alpha) - 1).$$

Because  $\alpha$  is 1-bounded, the coefficients in the series of  $F_{S^c}(u, \alpha)$  are bounded in absolute value by the respective coefficients of

$$Z_{S^c}(u) := \sum_{\substack{f \in \mathcal{M}_q \\ P|f \Rightarrow \deg P \in S^c}} u^{\deg f} = \prod_{P \in \mathcal{P}_q, \deg P \in S^c} (1 - u^{\deg P})^{-1}.$$

Further, we may write  $F_S(u, \alpha)$  as

$$\begin{aligned}
F_S(u, \alpha) &= \prod_{P \in \mathcal{P}_q, \deg P \in S} (1 - \alpha(P)u^{\deg P})^{-1} = \exp \left( \sum_{d \in S} \sum_{i \geq 1} i^{-1} u^{di} \sum_{P \in \mathcal{P}_{d,q}} \alpha(P)^i \right) \\
&= \exp \left( \sum_{d \in S} u^d \sum_{P \in \mathcal{P}_{d,q}} \alpha(P) + \sum_{d \in S} \sum_{i \geq 2} i^{-1} u^{di} \sum_{P \in \mathcal{P}_{d,q}} \alpha(P)^i \right) \\
&= \exp \left( \sum_{d \in S} \frac{u^d}{d} \sum_{f \in \mathcal{M}_{d,q}} \alpha(f) \Lambda(f) + \sum_{d \in S} \sum_{i \geq 2} \frac{u^{di}}{i} \sum_{P \in \mathcal{P}_{d,q}} \alpha(P)^i - \sum_{d \in S} \frac{u^d}{d} \sum_{\substack{e|d \\ e \neq d}} e \sum_{P \in \mathcal{P}_{e,q}} \alpha(P)^{d/e} \right).
\end{aligned}$$

It follows that  $[u^j](F_S(u, \alpha) - 1)$  for  $0 \leq j \leq n$  are bounded in absolute value by the coefficients of  $u^j$  in

$$Z_{S,\alpha}(u) := \exp \left( \sum_{d \in S} \frac{u^d}{d} \left| \sum_{f \in \mathcal{M}_{d,q}} \alpha(f) \Lambda(f) \right| + \sum_{d \in S} \sum_{i \geq 2} \frac{u^{di}}{i} |\mathcal{P}_{d,q}| + \sum_{d \in S} \frac{u^d}{d} \sum_{\substack{e|d \\ e \neq d}} e |\mathcal{P}_{e,q}| \right) - 1. \quad (3.6)$$

Putting this together, we have

$$|X| = |[u^n] F_{S^c}(u, \alpha) (F_S(u, \alpha) - 1)| \leq [u^n] Z_{S^c}(u) (Z_{S,\alpha}(u) - 1).$$

We make a general observation. If the coefficients of a power series  $F(u) = \sum_{i \geq 0} f_i u^i$  are bounded in absolute value by the respective coefficients of  $G(u) = \sum_{i \geq 0} g_i u^i$ , then  $|f_n| \leq G(R) R^{-n}$  for every  $0 < R < C$  such that  $G(R)$  converges. Indeed,

$$G(R) R^{-n} = \sum_{i=0}^{\infty} g_i R^{i-n} \geq g_n R^{n-n} = g_n \geq |f_n|.$$

Applying this observation with  $R = 1/q$ , we get

$$|X| \leq q^n Z_{S^c}(1/q) (Z_{S,\alpha}(1/q) - 1). \quad (3.7)$$

We estimate  $Z_{S^c}(1/q)$  using Lemma 3.1 as

$$Z_{S^c}(1/q) = \exp \left( \sum_{\deg P \in S^c} (q^{-\deg P} + O(q^{-2 \deg P})) \right) = \exp \left( \sum_{d \in S^c} d^{-1} + O(1) \right). \quad (3.8)$$

Similarly,

$$\sum_{d \in S} \sum_{i \geq 2} \frac{q^{-di}}{i} |\mathcal{P}_{d,q}| + \sum_{d \in S} \frac{q^{-d}}{d} \sum_{\substack{e|d \\ e \neq d}} e |\mathcal{P}_{e,q}| \leq \sum_{d \in S} \sum_{i \geq 2} \frac{q^{-di}}{i} \frac{q^d}{d} + 2 \sum_{d \in S} \frac{q^{-\frac{d}{2}}}{d} \ll \frac{q^{-s_0/2}}{s_0}, \quad (3.9)$$

where  $s_0 = \min_{s \in S} s$ . The required estimate then follows from (3.7), (3.8), (3.6), and (3.9).  $\square$

We now move on to estimating the first term in (3.5).

## 3.2 Sieve estimate

**Lemma 3.5** (Sieve bound). *Let  $S \subseteq \{1, \dots, n\}$  be a set of positive integers, and  $S^c = \{1, \dots, n\} \setminus S$  be its complement. Then*

$$\sum_{\substack{f \in \mathcal{M}_{n,q} \\ P|f \implies \deg P \notin S}} 1 \ll \frac{q^n}{n} \exp \left( \sum_{d \in S^c} d^{-1} \right) \asymp q^n \exp \left( - \sum_{d \in S} d^{-1} \right).$$

*Remark 3.2.* In the integer setting, the estimate  $\sum_{n \leq x: (n,k)=1} 1 \ll x \prod_{p|k} (1 - 1/p)$  for any positive integer  $k$  whose prime factors do not exceed  $x$  is a classical consequence of Selberg's sieve (see [15] for a discussion of this and an alternative proof). A permutation analogue of Lemma 3.5 was established (in greater generality) by Ford [10, Thm. 1.5]. The proof we give is self-contained and is in the spirit of [15, 14, 10].

*Proof.* Let

$$A_i := \sum_{\substack{f \in \mathcal{M}_{i,q} \\ P|f \implies \deg P \notin S}} 1,$$

and define

$$F(u) := \prod_{\deg P \in S^c} (1 - u^{\deg P})^{-1} = \exp \left( \sum_{\deg P \in S^c} \sum_{k \geq 1} \frac{u^{k \deg P}}{k} \right) =: \sum_{i=0}^{\infty} \tilde{A}_i u^i. \quad (3.10)$$

The coefficients satisfy  $\tilde{A}_i = A_i$  for  $0 \leq i \leq n$ . Differentiating (3.10) formally, we see that

$$\sum_{i \geq 1} i \tilde{A}_i u^i = F(u) G(u), \quad (3.11)$$

where

$$G(u) = \sum_{i \geq 1} B_i u^i, \quad B_i := \sum_{\substack{k \geq 1, \deg P \in S^c \\ k \deg P = i}} \deg P = \sum_{\substack{d \in S^c \\ d|i}} d |\mathcal{P}_{d,q}|.$$

Comparing coefficients in (3.11), we obtain

$$A_n = n^{-1} \sum_{i=1}^n A_{n-i} B_i.$$

Lemma 3.1 implies

$$B_i \leq \mathbf{1}_{i \notin S} (i |\mathcal{P}_{i,q}|) + \sum_{d \leq i/2} d |\mathcal{P}_{d,q}| \leq \mathbf{1}_{i \notin S} q^i + 2q^{i/2},$$

and so

$$A_n \leq n^{-1} \sum_{i \in S^c} q^i A_{n-i} + 2n^{-1} \sum_{i=1}^n A_{n-i} q^{i/2}.$$

The trivial bound  $A_{n-i} \leq q^{n-i}$  implies

$$A_n \leq n^{-1} \sum_{i \in S^c} q^i A_{n-i} + 5n^{-1} q^n \leq n^{-1} q^n \left( \sum_{0 \leq i \leq n} q^{-i} A_i + 5 \right).$$

Since

$$F(1/q) = \sum_{i \geq 0} q^{-i} \tilde{A}_i \geq \sum_{0 \leq i \leq n} q^{-i} A_i,$$

we find that

$$A_n \leq n^{-1} q^n (F(1/q) + 5).$$

By Lemma 3.1,

$$\log F(1/q) = \sum_{d \in S^c} |\mathcal{P}_{d,q}| \sum_{k \geq 1} \frac{q^{-kd}}{k} \leq \sum_{d \in S^c} \frac{1 + q^{-d}}{d} \leq \sum_{d \in S^c} \frac{1}{d} + 1,$$

and recalling  $\log n = \sum_{d=1}^n d^{-1} + O(1)$ , we conclude that

$$A_n \ll n^{-1} q^n \exp \left( \sum_{d \in S^c} d^{-1} \right) \asymp q^n \exp \left( - \sum_{d \in S} d^{-1} \right)$$

as needed.  $\square$

### 3.3 Proof of Lemma 2.2

Due to the trivial estimate  $|S(n, \chi)| \leq q^n$ , we may assume  $n \geq 12(1 + \log_q(1 + \deg Q))$ . Since  $S(n, \chi) = 0$  for  $n \geq \deg Q$  we may also assume  $\deg Q > n$ . Applying Lemma 3.4 with  $\alpha = \chi$  and  $S = \{m+1, m+2, \dots, n\}$  for some  $m \in (2 \log_q \deg Q, n]$  (as indicated in Remark 3.1), we find that (1.4) indeed holds by appealing to Lemma 3.3. By Lemma 3.5,

$$q^{-n} \sum_{\substack{f \in \mathcal{M}_{n,q} \\ P|f \implies \deg P \leq m}} \chi(f) \ll q^{-n} \sum_{\substack{f \in \mathcal{M}_{n,q} \\ P|f \implies \deg P \leq m}} 1 \ll \frac{m}{n}.$$

Choosing  $m = \lceil 2 \log_q(n(1 + \deg Q)) \rceil$  in (1.4) yields the statement of the lemma.

### 3.4 Bound on a gcd sum

For  $L \geq 1$ , let

$$B_{L,k} := \sum_{L \leq d < 2L} \log \gcd(k, q^d - 1).$$

We trivially have  $B_{L,k} \ll L \min\{\log k, L \log q\}$ , and this bound is optimal when  $\log_q k \gg L^2$ . For example, take  $k = \prod_{1 \leq i < 2L} (q^i - 1) = q^{\Theta(L^2)}$ , then  $q^d - 1 \mid k$  for all  $L \leq d < 2L$  so that  $B_{L,k} \asymp L^2 \log q$ . For smaller  $k$ , however, this bound is too generous.

**Lemma 3.6.** *Suppose  $L \geq \sqrt{\log_q k}$ . Then  $B_{L,k} \ll L \sqrt{\log k \log q}$ .*

To prove this result we will need a couple of facts concerning cyclotomic polynomials. Recall that the cyclotomic polynomials  $(\phi_n(x))_{n \geq 1}$  are defined recursively by  $\prod_{d|n} \phi_d(x) = x^n - 1$ . They lie in  $\mathbb{Z}[x]$  and can be written as  $\phi_n(x) = \prod_{j \in (\mathbb{Z}/n\mathbb{Z})^\times} (x - e^{2\pi i j/n})$ . This last relation shows that if  $m$  and  $p$  are coprime where  $p$  is a prime then  $\phi_{mp^i}(x) = \phi_m(x^{p^i})/\phi_m(x^{p^{i-1}})$  for  $i \geq 1$ .

The following lemma is classical but we could not find it explicitly stated in the literature and so we give a proof based on an argument implicit in Roitman [26].

**Lemma 3.7.** *Fix an integer  $a$ . For a prime  $p$  define  $A_p := \{n \geq 1 : p \mid \phi_n(a)\}$ . If  $p \mid a$ , then  $A_p = \emptyset$ . Otherwise,  $A_p = \{p^i \text{ord}_p(a) : i \geq 0\}$  where  $\text{ord}_p(a)$  is the multiplicative order of  $a$  modulo  $p$ .*

*Proof.* If  $p \mid \phi_n(a)$ , then  $p \mid a^n - 1$  since  $\phi_n(a) \mid a^n - 1$ . If  $p \mid a$ , then  $a^n - 1 \equiv -1 \pmod{p}$ , which is a contradiction. Thus,  $A_p = \emptyset$  in this case.

Now suppose  $p \nmid a$ . By definition, we have  $p \mid a^{\text{ord}_p(a)} - 1$ . Suppose  $p \mid \phi_n(a) \mid a^n - 1$  for some positive integer  $n$ , then

$$p \mid \gcd(a^n - 1, a^{\text{ord}_p(a)} - 1) = a^{\gcd(n, \text{ord}_p(a))} - 1.$$

Since  $\text{ord}_p(a) = \min\{n \geq 1 : p \mid a^n - 1\}$ , we have  $\gcd(n, \text{ord}_p(a)) \geq \text{ord}_p(a)$  and thus  $\text{ord}_p(a) \mid n$ . This shows

$$A_p \subseteq \{m \cdot \text{ord}_p(a) : m \geq 1\}. \quad (3.12)$$

Next we show  $\text{ord}_p(a) \in A_p$ . Since  $p \mid a^{\text{ord}_p(a)} - 1 = \prod_{e \mid \text{ord}_p(a)} \phi_e(a)$ , it follows that  $A_p$  contains some divisor of  $\text{ord}_p(a)$ , which by (3.12) then has to be  $\text{ord}_p(a)$  itself. Next, we want to show that for  $n \in A_p$ ,  $n/\text{ord}_p(a)$  is a power of  $p$ . If  $p'$  is a prime dividing  $n/\text{ord}_p(a)$ , then  $\text{ord}_p(a) \mid n/p'$  and so

$$p \mid \phi_n(a) = \frac{a^n - 1}{\prod_{d|n, d \neq n} \phi_d(a)} \mid \frac{a^n - 1}{\prod_{d|n/p'} \phi_d(a)} = \frac{a^n - 1}{a^{n/p'} - 1} = \sum_{i=0}^{p'-1} a^{in/p'} \equiv \sum_{i=0}^{p'-1} 1 = p' \pmod{p},$$

where we used that by definition of  $\text{ord}_p(a)$

$$a^{n/p'} = (a^{\text{ord}_p(a)})^{\frac{n}{p' \text{ord}_p(a)}} \equiv 1^{\frac{n}{p' \text{ord}_p(a)}} \equiv 1 \pmod{p}.$$

Hence,  $p'$  must be  $p$ , and  $n$  is indeed  $\text{ord}_p(a)$  times a power of  $p$ . This shows  $A_p \subseteq \{p^i \text{ord}_p(a) : i \geq 0\}$ . Finally, the relation  $\phi_{p^i m}(a) = \phi_m(a^{p^i})/\phi_m(a^{p^{i-1}})$  for  $\gcd(m, p) = 1$  and  $i \geq 1$  implies  $p^i \text{ord}_p(a) \in A_p$  for every  $i \geq 1$  since  $\phi_m(a^{p^i})/\phi_m(a^{p^{i-1}}) \equiv \phi_m(a)^{p^i - p^{i-1}} \pmod{p}$ .  $\square$

*Proof of Lemma 3.6.* Recall that  $q^d - 1 = \prod_{e|d} \phi_e(q)$ . Because  $\gcd(A, ab) \leq \gcd(A, a) \gcd(A, b)$ , it then follows that

$$\log \gcd(k, q^d - 1) \leq \sum_{e|d} \log \gcd(k, \phi_e(q)) = \sum_{e|d} \sum_{p|k} \log \gcd(p^{\nu_p(k)}, \phi_e(q)),$$

where  $p$  stands for prime and  $\nu_p(k)$  is the multiplicity of  $p$  in  $k$  (i.e., the  $p$ -adic valuation of  $k$ ). Hence,

$$B_{L,k} \leq \sum_{L \leq d < 2L} \sum_{\substack{p|k, \phi_e(q) \\ e|d}} \log \gcd(p^{\nu_p(k)}, \phi_e(q)).$$

We introduce a parameter  $T \geq 1$ . We shall show

$$B_{L,k} \ll T \log k + \frac{L^2 \log q}{T}, \quad (3.13)$$

and then take  $T = L/\sqrt{\log_q k}$ . We consider separately the contribution of  $e > d/T$  and  $e \leq d/T$ , obtaining

$$B_{L,k} \leq B_{L,k,1} + B_{L,k,2},$$

where

$$\begin{aligned} B_{L,k,1} &= \sum_{L \leq d < 2L} \sum_{\substack{p|k, \phi_e(q) \\ e|d \\ e > d/T}} \log \gcd(p^{\nu_p(k)}, \phi_e(q)), \\ B_{L,k,2} &= \sum_{L \leq d < 2L} \sum_{\substack{p|k, \phi_e(q) \\ e|d \\ e \leq d/T}} \log \gcd(p^{\nu_p(k)}, \phi_e(q)). \end{aligned}$$

Let us bound  $B_{L,k,1}$ . By interchanging the order of summation,

$$B_{L,k,1} \leq \sum_{\substack{e, p: p|k, \phi_e(q) \\ L/T < e < 2L}} \nu_p(k) \log p \sum_{\substack{L \leq d < 2L \\ e|d}} 1 \ll L \sum_{\substack{e, p: \\ p|k, \phi_e(q) \\ L/T < e < 2L}} \frac{\nu_p(k) \log p}{e}.$$

Let  $A_p := \{n \geq 1 : p \text{ divides } \phi_n(q)\}$ . By Lemma 3.7,  $A_p$  is either empty or is a geometric progression with step size  $p$ , so that

$$\sum_{\substack{e: p|\phi_e(q) \\ L/T < e < 2L}} \frac{1}{e} \ll \frac{T}{L} \sum_{i \geq 0} p^{-i} \ll \frac{T}{L}$$

and it follows that

$$B_{L,k,1} \ll T \sum_{p|k} \nu_p(k) \log p = T \log k.$$

Next, omitting the restriction  $p \mid k$  and using  $\phi_e(q) \mid q^e - 1 \leq q^e$ , we have

$$B_{L,k,2} \leq \sum_{L \leq d < 2L} \sum_{\substack{e|d \\ e \leq d/T}} \log \phi_e(q) \leq \log q \sum_{L \leq d < 2L} \sum_{\substack{e|d \\ e \leq d/T}} e \ll L \log q \sum_{1 \leq e < 2L/T} 1 \ll \frac{L^2 \log q}{T}.$$

This implies (3.13), and thus the statement of the lemma.  $\square$

### 3.5 Symmetry of character sums

**Lemma 3.8.** *Let  $\psi: \mathbb{F}_q \rightarrow \mathbb{C}^\times$  be an additive character. Let  $k, n \geq 1$ , and let  $k' = \gcd(k, q^n - 1)$ . Then,*

$$\sum_{f \in \mathcal{M}_{n,q}} \Lambda(f) \chi_{k,\psi}(f) = \sum_{f \in \mathcal{M}_{n,q}} \Lambda(f) \chi_{k',\psi}(f).$$

*Proof.* Consider a surjective map

$$\Phi: \mathbb{F}_{q^n}^\times \rightarrow \{f \in \mathcal{M}_{n,q} \setminus \{T\} : \Lambda(f) \neq 0\}$$

given by

$$\Phi(x) = \prod_{\sigma \in \text{Gal}(\mathbb{F}_{q^n}/\mathbb{F}_q)} (T - \sigma(x)) = \prod_{i=0}^{n-1} (T - x^{q^i}).$$

Alternatively, one can also write  $\Phi(x) = m_\alpha(x)^{n/\deg m_\alpha}$ , where  $m_\alpha$  is the minimal polynomial of  $x$  over  $\mathbb{F}_q$ . Every element  $f$  in the image of  $\Phi$  has  $\Lambda(f)$  preimages given by the Galois conjugates of a given preimage, namely,  $\{x^{q^i} : 0 \leq i < d\}$  if  $\deg m_x = d$ . Because  $\chi_{k,\psi}(f)$  is defined in terms of the zeros of  $f$ , the map  $\Phi$  allows us to conveniently write

$$\sum_{f \in \mathcal{M}_{n,q}} \Lambda(f) \chi_{k,\psi}(f) = \sum_{x \in \mathbb{F}_{q^n}^\times} \psi\left(\sum_{i=0}^{n-1} (x^{q^i})^{-k}\right).$$

By Euclid's algorithm, we can write  $k' = \gcd(k, q^n - 1) = ak + b(q^n - 1)$  for some coprime integers  $a$  and  $b$ . In particular,  $x^{k'} = (x^k)^a$ . Conversely,  $x^k = (x^{k'})^{k/k'}$ . It follows that the group endomorphisms  $x \mapsto x^k$  and  $x \mapsto x^{\gcd(k, q^n - 1)}$  defined on  $\mathbb{F}_{q^n}^\times$  have the same image and kernel. This implies that every element in the image is attained the same number of times, and so

$$\sum_{x \in \mathbb{F}_{q^n}^\times} \psi\left(\sum_{i=0}^{n-1} (x^{q^i})^{-k}\right) = \sum_{x \in \mathbb{F}_{q^n}^\times} \psi\left(\sum_{i=0}^{n-1} (x^{q^i})^{-k'}\right) = \sum_{f \in \mathcal{M}_{n,q}} \Lambda(f) \chi_{k',\psi}(f)$$

as required.  $\square$

From Lemmas 2.1, 3.3, and 3.8, we conclude the following.

**Corollary 3.9.** *Let  $\psi: \mathbb{F}_q \rightarrow \mathbb{C}^\times$  be a nontrivial additive character. Let  $k, n \geq 1$ . We have*

$$\left| \sum_{f \in \mathcal{M}_{n,q}} \Lambda(f) \chi_{k,\psi}(f) \right| \leq q^{\frac{n}{2}} \gcd(k, q^n - 1).$$

### 3.6 Proof of Proposition 1.6

Let  $m := \lceil 12 \log_q n \rceil$ , and

$$S = \{m \leq d \leq n : \gcd(k, q^d - 1) < q^{d/3}\}. \quad (3.14)$$

Let  $S^c = \{1, \dots, n\} \setminus S$ . As before, let

$$q^{-n} \sum_{f \in \mathcal{M}_{n,q}} \chi_{k,\psi}(f) = q^{-n} \sum_{\substack{f \in \mathcal{M}_{n,q} \\ \exists P|f \text{ such that } \deg P \in S}} \chi_{k,\psi}(f) + q^{-n} \sum_{\substack{f \in \mathcal{M}_{n,q} \\ P|f \implies \deg P \in S^c}} \chi_{k,\psi}(f).$$

Applying Lemma 3.4 for  $\alpha = \chi_{k,\psi}$  and our chosen  $S$  together with Corollary 3.9, we get

$$\begin{aligned} & q^{-n} \sum_{\substack{f \in \mathcal{M}_{n,q} \\ \exists P|f \text{ such that } \deg P \in S}} \chi_{k,\psi}(f) \\ & \ll \exp\left(\sum_{d \in S^c} d^{-1}\right) \left(\exp\left(\sum_{d \in S} \frac{q^{-d/6}}{d} + O(q^{-m/2}/m)\right) - 1\right) \ll \frac{nq^{-m/6}}{m} \ll n^{-1}. \end{aligned}$$

By Lemma 3.5,

$$q^{-n} \sum_{\substack{f \in \mathcal{M}_{n,q} \\ P|f \Rightarrow \deg P \in S^c}} \chi_{k,\psi}(f) \ll q^{-n} \sum_{\substack{f \in \mathcal{M}_{n,q} \\ P|f \Rightarrow \deg P \in S^c}} 1 \ll \exp\left(-\sum_{d \in S} d^{-1}\right).$$

*Remark 3.3.* The same proof works as is if we replace  $\chi_{k,\psi}$  by  $\mu \cdot \chi_{k,\psi}$ , because

$$\sum_{f \in \mathcal{M}_{d,q}} \chi_{k,\psi}(f) \Lambda(f) = - \sum_{f \in \mathcal{M}_{d,q}} \mu(f) \chi_{k,\psi}(f) \Lambda(f) + O(q^{d/2}).$$

### 3.7 Conclusion of proof of Theorem 1.5

Due to the trivial estimate  $|S(n, \chi)| \leq q^n$ , we may assume  $n \geq 12(1 + \sqrt{\log_q k})$ . In view of Proposition 1.6 it suffices to show

$$-\sum_{d \in S} d^{-1} \leq \log(1 + \sqrt{\log_q k} + \log_q n) - \log n + O(1),$$

where  $S$  is as in (3.14). To show this, first observe that  $-\sum_{d \in S} d^{-1} = \sum_{d \in S^c} d^{-1} - \log n + O(1)$ . Let  $m' = \max\{\lceil 12 \log_q n \rceil, \sqrt{\log_q k}\}$ . Since  $\mathbf{1}_{A \geq B} \leq \log A / \log B$ ,

$$\sum_{d \in S^c} d^{-1} \leq \sum_{d \leq m'} d^{-1} + \sum_{\substack{m' \leq d \leq n \\ \gcd(k, q^d - 1) \geq q^{d/3}}} d^{-1} \leq \log m' + O(1) + \sum_{m' \leq d \leq n} d^{-1} \frac{\log \gcd(k, q^d - 1)}{\log(q^{d/3})}.$$

By Lemma 3.6, the last  $d$ -sum is  $O(\sqrt{\log_q k}/m')$ , finishing the proof.

## Acknowledgments

We thank Brad Rodgers and Zeev Rudnick for comments on an earlier version of the paper. We appreciate the comments and suggestions given by the anonymous referee, which improved the presentation of the paper. O.G. was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement number 851318). V.K. was supported by ERC LogCorRM (grant number 740900) and by CRM-ISM postdoctoral fellowship.

## A Squareroot cancellation in character sums

Under the generalized Riemann hypothesis for  $L(s, \chi)$  one has  $\sum_{n \leq x} \chi(n) \ll \sqrt{x} \exp(C \log m / \log \log m)$  for a nonprincipal Dirichlet character  $\chi$  modulo  $m$  [3]. Bhowmick, Lê and Liu [3] proved the function field analogue

$$\sum_{f \in \mathcal{M}_{n,q}} \chi(f) \ll q^{n/2} \exp\left(C_q \left(\frac{n \log \log \deg Q}{\log \deg Q} + \frac{\deg Q}{\log^2 \deg Q}\right)\right) \quad (\text{A.1})$$

unconditionally, for a constant  $C_q$  that depends only on  $q$ , where  $\chi$  is a character modulo  $Q \in \mathbb{F}_q[T]$ . The estimate (A.1) implies  $\sum_{f \in \mathcal{M}_{n,q}} \chi(f) \ll q^{\frac{n}{2}(1+o(1))}$  for  $\deg Q = o(n \log^2 n)$ . Here we prove that squareroot cancellation holds in the wider range  $\deg Q \leq n^{1+o(1)}$ .

**Proposition A.1.** *Let  $\chi$  be a nonprincipal Dirichlet character modulo  $Q \in \mathbb{F}_q[T]$ . There is a constant  $C_q$  that depends only on  $q$  such that*

$$\sum_{f \in \mathcal{M}_{n,q}} \chi(f) \ll_q q^{\frac{n}{2}} \exp\left(\max\left\{C_q n \frac{\log(\deg Q \log \deg Q/n)}{\log(1 + \deg Q)}, 0\right\}\right).$$



Recall  $d(\chi) := \deg L(u, \chi) < \deg Q$ . If  $d(\chi) < n$  we have  $S(n, \chi) = 0$ , so from now on we assume that  $d(\chi) \geq n$ . In view of the trivial bound  $|S(n, \chi)| \leq q^n$  we may also assume that  $n$  is larger than a fixed constant, as well as that  $d \leq n^{1+\delta}$  for a fixed  $\delta > 0$ . From (3.2),

$$L(u, \chi) = \exp \left( \sum_{k \geq 1} \frac{u^k}{k} \sum_{f \in \mathcal{M}_{k,q}} \chi(f) \Lambda(f) \right).$$

Set  $L := \lfloor 2 \log_q d \rfloor$ . For  $k \leq L$  the trivial bound

$$\left| \sum_{f \in \mathcal{M}_{k,q}} \chi(f) \Lambda(f) \right| \leq \sum_{f \in \mathcal{M}_{k,q}} \Lambda(f) = q^k \quad (\text{A.2})$$

coming from (3.1) is superior to (3.4). We see from (3.4) and (A.2) that the coefficients of  $L(u, \chi)$  are bounded (in absolute value) from above by those of

$$\exp \left( \sum_{k \geq 1} \frac{u^k}{k} \min\{q^k, d(\chi) q^{k/2}\} \right),$$

hence, writing  $[u^n]F$  for the  $n$ th coefficient of a power series  $F$ ,

$$\begin{aligned} |S(n, \chi)| &= q^{n/2} |[u^n]L(u/\sqrt{q}, \chi)| \leq q^{n/2} [u^n] \exp \left( \sum_{k \geq 1} \frac{u^k}{k} \min\{q^{k/2}, d(\chi)\} \right) \\ &\leq q^{n/2} \exp \left( \sum_{k \geq 1} \frac{R^k}{k} \min\{q^{k/2}, d(\chi)\} \right) R^{-n} \end{aligned}$$

for any  $R \in (0, 1)$  (this is the same observation used in the proof of Lemma 3.4). If  $6/(5\sqrt{q}) < R < 1$  then

$$\sum_{1 \leq k \leq L} \frac{R^k}{k} \min\{q^{k/2}, d(\chi)\} = \sum_{1 \leq k \leq L} \frac{(R\sqrt{q})^k}{k} \leq \sum_{1 \leq k \leq L} (R\sqrt{q})^k \leq \frac{(R\sqrt{q})^L}{1 - (R\sqrt{q})^{-1}} \leq 6d(\chi)R^L$$

and

$$\sum_{k > L} \frac{R^k}{k} \min\{q^{k/2}, d(\chi)\} \leq d(\chi) \sum_{k > L} \frac{R^k}{k} \leq \frac{d(\chi)}{L+1} \sum_{k > L} R^k = \frac{d(\chi)}{L+1} \frac{R^{L+1}}{1-R},$$

so that

$$|S(n, \chi)/q^{n/2}| \leq \exp \left( 6d(\chi)R^L \left( 1 + \frac{R}{(L+1)(1-R)} \right) - n \log R \right).$$

Set  $a := \log(d(\chi) \log d(\chi)/n)$  and take  $R = q^{-a/(2 \log d(\chi))}$ . The assumption  $d(\chi) \geq n$  implies  $a > 0$  (at least if  $n \geq 3$ ) and so  $R < 1$ . (This choice of  $R$  differs from the choice in [3], where  $a$  is chosen to be of order  $\log \log d(\chi)$ .) Since we may assume  $n \leq d(\chi) \leq n^{1+\delta}$  and  $n \geq C$  for  $\delta$  and  $C$  we choose, we may also assume  $a/\log d(\chi)$  is at most  $\delta'$  for a  $\delta'$  we want. In particular,  $R > 6/(5\sqrt{q})$  may be assumed from now on by taking small enough  $\delta'$ . Moreover,

$$-n \log R \leq (\log q) n \frac{\log(d(\chi) \log d(\chi)/n)}{\log d(\chi)}, \quad R^L \leq q^{-\frac{\log(d(\chi) \log d(\chi)/n)}{2 \log d(\chi)} (2 \log_q d(\chi) - 1)} \leq \frac{2}{(d(\chi) \log d(\chi)/n)}$$

for large enough  $n$ . Additionally,  $a \log q / \log d(\chi) \in (0, 1)$  holds (by taking small enough  $\delta'$ ) which implies  $R \leq 1 - a \log q / (4 \log d(\chi))$ , and so

$$\frac{R}{(L+1)(1-R)} \leq \frac{1}{L+1} \frac{1}{1-R} \leq \frac{\log q}{2 \log d(\chi)} \frac{4 \log d(\chi)}{\log q \log(d(\chi) \log d(\chi)/n)} = \frac{2}{\log(d(\chi) \log d(\chi)/n)} \leq 1$$

for large enough  $n$ . All in all,

$$|S(n, \chi)/q^{n/2}| \leq \exp \left( \frac{24n}{\log d(\chi)} + (\log q) n \frac{\log(d(\chi) \log d(\chi)/n)}{\log d(\chi)} \right) \leq \exp \left( C_q n \frac{\log(d(\chi) \log d(\chi)/n)}{\log d(\chi)} \right).$$

Since  $\deg Q > d(\chi)$ , the claim follows.

# References

- [1] W. D. Banks and I. E. Shparlinski. Bounds on short character sums and  $L$ -functions with characters to a powerful modulus. *J. Anal. Math.*, 139(1):239–263, 2019.
- [2] A. Bhowmick and T. H. Lê. On primitive elements in finite fields of low characteristic. *Finite Fields Appl.*, 35:64–77, 2015.
- [3] A. Bhowmick, T. H. Lê, and Y.-R. Liu. A note on character sums in finite fields. *Finite Fields Appl.*, 46:247–254, 2017.
- [4] D. A. Burgess. On character sums and  $L$ -series. II. *Proc. Lond. Math. Soc. (3)*, 13:524–536, 1963.
- [5] D. A. Burgess. The character sum estimate with  $r = 3$ . *J. London Math. Soc. (2)*, 33(2):219–226, 1986.
- [6] K. Courteaut and K. Johansson. Multivariate normal approximation for traces of orthogonal and symplectic matrices. *Ann. Inst. Henri Poincaré, Probab. Stat.*, 60(1):312–342, 2024.
- [7] K. Courteaut, K. Johansson, and G. Lambert. From Berry-Esseen to super-exponential. *Electron. J. Probab.*, 29:48, 2024. Id/No 11.
- [8] N. G. de Bruijn. On the number of positive integers  $\leq x$  and free of prime factors  $> y$ . *Nederl. Akad. Wet., Proc., Ser. A*, 54:50–60, 1951.
- [9] P. Diaconis and M. Shahshahani. On the eigenvalues of random matrices. *J. Appl. Probab.*, 31A:49–62, 1994.
- [10] K. Ford. Cycle type of random permutations: a toolkit. *Discrete Anal.*, 2022. Paper No. 9, 36.
- [11] M. Gerstenhaber. On the number of nilpotent matrices with coefficients in a finite field. *Illinois J. Math.*, 5:330–333, 1961.
- [12] O. Gorodetsky and B. Rodgers. Traces of powers of matrices over finite fields. *Trans. Amer. Math. Soc.*, 374(7):4579–4638, 2021.
- [13] A. Granville and K. Soundararajan. Large character sums. *J. Amer. Math. Soc.*, 14(2):365–397, 2001.
- [14] H. Halberstam and H.-E. Richert. On a result of R. R. Hall. *J. Number Theory*, 11(1):76–89, 1979.
- [15] R. R. Hall. Halving an estimate obtained from Selberg’s upper bound method. *Acta Arith.*, 25:347–351, 1973/74.
- [16] D. R. Hayes. The distribution of irreducibles in  $GF[q, x]$ . *Trans. Amer. Math. Soc.*, 117:101–127, 1965.
- [17] H. Iwaniec and E. Kowalski. *Analytic number theory*, volume 53 of *Colloq. Publ., Am. Math. Soc.* Providence, RI: American Mathematical Society (AMS), 2004.
- [18] K. Johansson. On random matrices from the compact classical groups. *Ann. of Math. (2)*, 145(3):519–545, 1997.
- [19] K. Johansson and G. Lambert. Multivariate normal approximation for traces of random unitary matrices. *Ann. Probab.*, 49(6):2961–3010, 2021.
- [20] J. Keating and Z. Rudnick. Squarefree polynomials and Möbius values in short intervals and arithmetic progressions. *Algebra Number Theory*, 10(2):375–420, 2016.
- [21] J. P. Keating and Z. Rudnick. The variance of the number of prime polynomials in short intervals and in residue classes. *Int. Math. Res. Not. IMRN*, (1):259–288, 2014.
- [22] H. L. Montgomery and R. C. Vaughan. Exponential sums with multiplicative coefficients. *Invent. Math.*, 43(1):69–82, 1977.

- [23] P. Pollack. Irreducible polynomials with several prescribed coefficients. *Finite Fields Appl.*, 22:70–78, 2013.
- [24] E. M. Rains. High powers of random elements of compact Lie groups. *Probab. Theory Related Fields*, 107(2):219–241, 1997.
- [25] I. Reiner. On the number of matrices with given characteristic polynomial. *Illinois J. Math.*, 5:324–329, 1961.
- [26] M. Roitman. On Zsigmondy primes. *Proc. Amer. Math. Soc.*, 125(7):1913–1919, 1997.
- [27] M. Rosen. *Number theory in function fields*, volume 210 of *Graduate Texts in Mathematics*. Springer, New York, 2002.
- [28] A. Weil. *Basic number theory*, volume 144 of *Grundlehren Math. Wiss.* Springer, Cham, 1967.

MATHEMATICAL INSTITUTE, UNIVERSITY OF OXFORD, OXFORD, OX2 6GG, UK  
*E-mail address:* ofir.goro@gmail.com

DÉPARTEMENT DE MATHÉMATIQUES ET STATISTIQUE, UNIVERSITÉ DE MONTRÉAL, MONTRÉAL, QC H3C 3J7, CANADA  
*E-mail address:* valkovaleva42@gmail.com