ON THE CHOWLA AND TWIN PRIMES CONJECTURES OVER $\mathbb{F}_q[T]$

WILL SAWIN AND MARK SHUSTERMAN

Abstract. Using geometric methods, we improve on the function field version of the Burgess bound, and show that, when restricted to certain special subspaces, the Möbius function over $\mathbb{F}_q[T]$ can be mimicked by Dirichlet characters. Combining these, we obtain a level of distribution close to 1 for the Möbius function in arithmetic progressions, and resolve Chowla's k-point correlation conjecture with large uniformity in the shifts. Using a function field variant of a result by Fouvry-Michel on exponential sums involving the Möbius function, we obtain a level of distribution beyond 1/2 for irreducible polynomials, and establish the twin prime conjecture in a quantitative form. All these results hold for finite fields satisfying a simple condition.

1. Introduction

Our main results are the resolutions of two open problems in number theory, except with the ring of integers $\mathbb{Z}$ replaced by the ring of polynomials $\mathbb{F}_q[T]$ for a fixed prime power q, under suitable assumptions on q.

We first fix some notation. Define the norm of a nonzero $f \in \mathbb{F}_q[T]$ to be

$$|f| = q^{\deg(f)} = |\mathbb{F}_q[T]/(f)|. \tag{1.1}$$

The degree of the zero polynomial is negative ∞, so we set its norm to be 0.

1.1. **The main result - twin primes.** Our main result covers the twin prime conjecture in its quantitative form. The latter is the 2-point prime tuple conjecture of Hardy-Littlewood, predicting for a nonzero integer h that

$$\#\{X \le n \le 2X : n \text{ and } n+h \text{ are prime}\} \sim \mathfrak{S}(h)\frac{X}{\log^2(X)}, \quad X \to \infty, \tag{1.2}$$

where

$$\mathfrak{S}(h) = \prod_p (1-p^{-1})^{-2}(1-p^{-1}-p^{-1}\mathbf{1}_{p\nmid h}), \tag{1.3}$$

with $\mathbf{1}_{p\nmid h}$ equals 1 if h is not divisible by p, and 0 otherwise.

For the function field analogue, we set

$$\mathfrak{S}_q(h) = \prod_P \left(1-|P|^{-1}\right)^{-2}\left(1-|P|^{-1}-|P|^{-1}\mathbf{1}_{P\nmid h}\right) \tag{1.4}$$

W.S. served as a Clay Research Fellow, and, later, was supported by NSF grant DMS-2101491 while working on this paper.

where q is a prime power, P ranges over all primes (monic irreducibles) of $\mathbb{F}_q[T]$, and $h \in \mathbb{F}_q[T]$ is nonzero.

Theorem 1.1. *For an odd prime number p, and a power q of p satisfying $q > 685090p^2$, the following holds. For any nonzero $h \in \mathbb{F}_q[T]$ we have*

$$\#\{f \in \mathbb{F}_q[T] : |f| = X, f \text{ and } f+h \text{ are prime}\} \sim \mathfrak{S}_q(h)\frac{X}{\log_q^2(X)} \tag{1.5}$$

as $X \to \infty$ through powers of q. Moreover, we have a power saving (depending on q) in the asymptotic above.

For example, the 2-point Hardy-Littlewood conjecture holds over

$$\mathbb{F}_{3^{15}}, \mathbb{F}_{5^{11}}, \mathbb{F}_{7^9}, \mathbb{F}_{11^8}, \mathbb{F}_{685093^3}. \tag{1.6}$$

For h a constant, the fact that the count above tends to ∞ was proven in [Hal06, Corollary 14] for $q > 3$ and in [Pol08, Theorem 1] for $q > 2$. This has been extended to monomial h (assuming $q > 105$) in [CHLPT15, Theorem 1.4 (1)] using an idea of Entin. The latter work builds on the recent dramatic progress on this problem over the integers, particularly [Ma15]. The strongest result known over the integers is [PM14, Theorem 16(i)], which says that for any 'admissible tuple' of 50 integers, there exists at least one difference h between two elements in the tuple such that there are infinitely many pairs of primes separated by h.

Remark 1.2. Our proof of Theorem 1.1 establishes also the analog of the Goldbach problem over function fields, and can be modified to treat more general linear forms in the primes.

The proof of Theorem 1.1 passes through some intermediate results which may be of independent interest. We will discuss these results in the remainder of the introduction. Once (a uniform variant of) Theorem 1.3 (for $k = 2$) and Theorem 1.7 below are established, Theorem 1.1 will follow from arguments similar to those in [MV17]. These involve a convolution identity relating the von Mangoldt function, which can be used to count primes, to the Möbius function.

1.2. **The key ingredient - Chowla's conjecture.** The main ingredient in the proof of Theorem 1.1 is the removal of the 'parity barrier'. More precisely, we confirm Chowla's k-point correlation conjecture over $\mathbb{F}_q[T]$ for some prime powers q. Over the integers, this conjecture predicts that for any fixed distinct integers $h_1, \dots, h_k$, one has

$$\sum_{n \le X} \mu(n+h_1)\mu(n+h_2)\cdots\mu(n+h_k) = o(X), \quad X \to \infty. \tag{1.7}$$

The only completely resolved case is $k = 1$ which is essentially equivalent to the prime number theorem.

For the function field analogue, we recall that the Möbius function of a monic polynomial f is 0 if f is not squarefree, and is otherwise given by

$$\mu(f) = \begin{cases} 1, & \#\{P : P \mid f\} \equiv 0 \bmod 2 \\ -1, & \#\{P : P \mid f\} \equiv 1 \bmod 2. \end{cases} \tag{1.8}$$

We denote by $\mathbb{F}_q[T]^+$ the set of monic polynomials over $\mathbb{F}_q$, and Euler's number by $e = \exp(1) = 2.71828\ldots$.

Theorem 1.3. *For an odd prime number p, an integer $k \geq 1$, and a power q of p satisfying $q > p^2k^2e^2$, the following holds. For every fixed choice of k distinct polynomials $h_1, \ldots, h_k \in \mathbb{F}_q[T]$ we have*

$$\sum_{\substack{f \in \mathbb{F}_q[T]^+ \\ |f| \leq X}} \mu(f+h_1)\mu(f+h_2)\cdots\mu(f+h_k) = o(X), \quad X \to \infty. \tag{1.9}$$

For instance, the 2-point Chowla conjecture holds over

$$\mathbb{F}_{3^6}, \mathbb{F}_{5^5}, \mathbb{F}_{7^4}, \mathbb{F}_{31^3}. \tag{1.10}$$

In fact, in Theorem 1.3 we obtain a power saving exponent inversely proportional to p, and the sizes of the shifts $|h_1|, \ldots, |h_k|$ can be as large as any fixed power of X (the corresponding assumption on q becomes stronger as this power grows larger). In Corollary 6.1 we also get cancellation in case the sum is restricted to prime polynomials f.

Over the integers, the $k = 2$ case of the Chowla conjecture, with logarithmic averaging, was proven in [Tao16, Theorem 3], building on earlier breakthrough work of Matomäki and Radziwiłł [MR16]. The k odd case, again with logarithmic averaging, was handled by Tao and Teräväinen [TT19]. Generalizations of some of these arguments to the function field setting form a part of the work [KMT20].

In contrast to these works, which deal with any sufficiently general (i.e. non-pretentious) multiplicative function, our result relies on special properties of the Möbius function (in positive characteristic). Specifically, we observe that for any fixed polynomial r, the function $\mu(r + s^p)$ essentially equals $\chi_r(s+c_r)$ where χ_r is a quadratic Dirichlet character and c_r is a shift, both depending only on r (and not on s).

This observation is very closely related to the properties of the Möbius function described in [CCG08], specifically [CCG08, Theorem 4.8]. Conrad, Conrad, and Gross prove a certain quasiperiodicity property in s for a general class of expressions of the form $r + s^p$, while we give a more precise description via Dirichlet characters in a special case. Similar features of μ, or rather of μ^2, facilitate the counting of squarefrees in different contexts, as can be seen from [Car21, Lan15, Poo03].

Our observation on $\mu(r + s^p)$ arises from the connection between the parity of the number of prime factors of a squarefree $f \in \mathbb{F}_q[T]$, and the sign (inside the symmetric group) of the Frobenius automorphism acting on the roots of f. In odd characteristic, this sign is determined by the value of the

quadratic character of $\mathbb{F}_q^\times$ on the discriminant of f, i.e. the resultant of f and its derivative f'. In characteristic p, the derivative of $f = r + s^p$ is equal to the derivative of r, so the aforementioned sign of Frobenius is determined by the quadratic character of the resultant of f with the fixed polynomial r'. The latter is a quadratic Dirichlet character of f, and thus equals an additively shifted Dirichlet character of s.

To use our observation, we restrict the sum in Theorem 1.3 to f of the form $r + s^p$ for any fixed r, and obtain a short sum in s of a product of additively shifted Dirichlet characters. As the conductors of these characters are typically essentially coprime, using the Chinese remainder theorem we arrive at a short sum of a single Dirichlet character.

Typically in analytic number theory, short character sums are handled by the method of Burgess, who showed in [Bur63] that for a real number $\eta > 1/4$, a squarefree integer M, a real number $X \geq |M|^\eta$, and a nonprincipal Dirichlet character χ mod M, one has

$$\sup_{s\in\mathbb{Z}} \left| \sum_{|a|\leq X} \chi(s+a) \right| = o(X), \qquad |M| \to \infty. \tag{1.11}$$

Refining the method of Burgess is the focus of several works, but the exponent $1/4$ has not yet been improved (even conditionally). However, in the function field setting, we can do better by a geometric method, as long as q is sufficiently large.

Theorem 1.4. *Fix $\eta > 0$. Then for a prime power $q > e^2/\eta^2$ the following holds. For a squarefree $M \in \mathbb{F}_q[T]$, a real number $X \geq |M|^\eta$, and a nonprincipal Dirichlet character χ mod M, we have*

$$\sup_{s\in\mathbb{F}_q[T]} \left| \sum_{|a|\leq X} \chi(s+a) \right| = o(X), \qquad |M| \to \infty. \tag{1.12}$$

By further enlarging q, we get arbitrarily close to square root cancellation. This is stated more precisely in Corollary 2.7.

To prove Theorem 1.4, we express the problem geometrically, viewing the short interval $\{s + a : |a| \leq X\}$ as an affine space over $\mathbb{F}_q$, and the character χ as arising from a sheaf on that space. Following a strategy from [Hoo91, appendix by Katz], we use vanishing cycles theory to compare the cohomology of this sheaf for the $s = 0$ short interval and its cohomology for a general short interval. Vanishing cycles can only occur when the vanishing locus of χ is not (geometrically) a simple normal crossings divisor. Arguing as in [Kat89], we split the modulus M of χ into a product of distinct linear terms over $\overline{\mathbb{F}_q}$, which makes our vanishing locus a union of the hyperplanes where the linear terms vanish, so we can check that this is a simple normal crossings divisor away from some isolated points. This implies that the cohomology groups vanish until almost the middle degree. Since we can

precisely calculate the vanishing cycles at the isolated points, we get a very good control of the dimensions of cohomology groups as well.

Remark 1.5. The relation between the Möbius function and multiplicative characters is less powerful the larger p is, as then fewer polynomials share a given derivative. On the other hand, our geometric character sum bounds become stronger as q grows. Thus, to make this method of proving Chowla work, we need q to be sufficiently large with respect to p.

Remark 1.6. The study of the statistics of polynomial factorizations by examining Frobenius as an element of the symmetric group has been very fruitful in the 'large finite field limit', where (in the notation of Theorem 1.3) X is kept fixed and q is allowed to grow. We refer to [CaRu14], [Ca15], [GS20] (and references therein) for the large finite field analogs of Theorem 1.3 which save a fixed power of q. Our methods likely give an improved savings in the large finite field limit when the characteristic is fixed, as long as the degrees of the polynomials are sufficiently large with respect to the characteristic, but we have not carefully calculated the resulting bounds in this range.

1.3. **Further ingredients - level of distribution estimates.** Another ingredient in the proof of Theorem 1.1 is an improvement of the level of distribution of the Möbius and von Mangoldt functions in arithmetic progressions. Over the integers, assuming the Generalized Riemann Hypothesis (GRH), this level of distribution is (at least) 1/2, which means that

$$\sum_{\substack{n\leq X\\ n\equiv a \bmod M}} \mu(n) = o\left(\frac{X}{|M|}\right), \qquad \sum_{\substack{n\leq X\\ n\equiv a \bmod M}} \Lambda(n) = \frac{X}{\varphi(M)} + o\left(\frac{X}{|M|}\right) \tag{1.13}$$

where M, a are coprime integers, and $|M| \leq X^{\frac{1}{2}-\epsilon}$ (for any fixed $\epsilon > 0$ and real $X \geq 0$ going to infinity).

The importance of level of distribution for the twin prime conjecture was known for a long time, and is highlighted for instance in the groundbreaking work [Zha14] of Zhang, where attention is concentrated on (results on average over) smooth moduli. Improvements of Zhang's results on level of distribution have been obtained in [PM14B]. For results that hold on average over all moduli, and have inspired Zhang's work, we refer to [BFI89].

For the Möbius function over $\mathbb{F}_q[T]$, we obtain a level of distribution close to 1.

Theorem 1.7. *Fix $\eta > 0$. For an odd prime number p, and a power q of p with $q > p^2 e^2 \left(\frac{2}{\eta} - 1\right)^2$, the following holds. For coprime $M, a \in \mathbb{F}_q[T]$, and a real number X with $X^{1-\eta} \geq |M|$ we have*

$$\sum_{\substack{f\in\mathbb{F}_q[T]^+\\ |f|\leq X\\ f\equiv a \bmod M}} \mu(f) = o\left(\frac{X}{|M|}\right), \quad |M| \to \infty. \tag{1.14}$$

As in the previous theorems, we obtain a power savings estimate. Here however (as opposed to Theorem 1.3), every f in our sum may have a different derivative, so a somewhat more elaborate implementation of our observation on the Möbius function is required. We put $f = Mg + a$, and wishfully write

$$\mu(Mg+a) \approx \mu(M)\mu\left(g + \frac{a}{M}\right) \tag{1.15}$$

in order to create coincidences among the derivatives of the inputs to the Möbius function. This is carried out more formally in Lemma 3.2, where we show that for a power q of an odd prime p, and coprime $a, M \in \mathbb{F}_q[T]$, the function $s \mapsto \mu(a + s^p M)$ is essentially proportional to an additive shift of a (quadratic) Dirichlet characters in s, with the modulus of the character depending on a and M in an explicit way. To visualize the power of this claim, we view $\mathbb{F}_q[T]$ as a rank p lattice over its subring $\mathbb{F}_q[T^p]$. Restricting the Möbius function to any line in this lattice gives a Dirichlet character whose modulus varies with the line.

In order to deduce from Theorem 1.7 an improved level of distributions for primes, we establish in the appendix a function field variant of [FM98, Theorem 1.1] giving quasi-orthogonality of the Möbius function and 'inverse additive characters'. While Fouvry-Michel work with characters to prime moduli, in order to establish Theorem 1.1 we need arbitrary squarefree moduli.

Theorem 1.8. *Let q be a prime power, and let $\epsilon > 0$. Then for a squarefree $M \in \mathbb{F}_q[T]$, and an additive character ψ mod M, we have*

$$\sum_{\substack{f \in \mathbb{F}_q[T]^+ \\ |f| \leq X \\ (f,M)=1}} \mu(f)\psi\left(\overline{f}\right) \ll |M|^{\frac{3}{16}+\epsilon} X^{\frac{25}{32}}, \quad X, |M| \to \infty \tag{1.16}$$

where $\overline{f}$ denotes the inverse of f mod M, and the implied constant depends only on q and ϵ.

For nonzero $M \in \mathbb{F}_q[T]$ we recall that Euler's totient function is given by

$$\varphi(M) = \left|(\mathbb{F}_q[T]/(M))^\times\right|, \tag{1.17}$$

and for $f \in \mathbb{F}_q[T]^+$, we recall that the von Mangoldt function is

$$\Lambda(f) = \begin{cases} \deg(P), & f = P^n \\ 0, & \text{otherwise.} \end{cases} \tag{1.18}$$

Theorem 1.9. *Fix $\delta < \frac{1}{126}$. For an odd prime p and a power q of p with*

$$q > p^2 e^2 \left(\frac{51 - 26\delta}{1 - 126\delta}\right)^2, \tag{1.19}$$

the following holds. For a squarefree $M \in \mathbb{F}_q[T]$, a polynomial $a \in \mathbb{F}_q[T]$ coprime to M, and X a power of q with $X^{\frac{1}{2}+\delta} \geq |M|$ we have

$$\sum_{\substack{f \in \mathbb{F}_q[T]^+ \\ |f|=X \\ f \equiv a \bmod M}} \Lambda(f) = \frac{X}{\varphi(M)} + o\left(\frac{X}{|M|}\right), \quad |M| \to \infty. \tag{1.20}$$

We have not ventured too much into improving the constant $\frac{1}{126}$, as our method cannot give anything above $\frac{1}{6}$, even if Theorem 1.8 would give square root cancellation. Since our proof of Theorem 1.9 is based on the 'convolutional' connection of the von Mangoldt and Möbius functions, it is not surprising that a level of distribution of $\frac{2}{3} = \frac{1}{2} + \frac{1}{6}$, which is a longstanding barrier for the divisor function over $\mathbb{Z}$ (perhaps the most basic convolution), is a natural limit of our techniques. A large finite field variant of Theorem 1.7 and Theorem 1.9 was earlier proved in [BBSR15, Theorem 2.5].

Remark 1.10. It would be interesting to see whether our results can be extended to characteristic 2, perhaps in a manner similar to which [Ca15] extends the results of [CaRu14].

1.4. **Additional results in small characteristic.** Throughout this work, we have not made every possible effort to reduce the least values of the prime powers q to which our theorems apply. Instead, we present some results that hold for q as small as 3.

The first concerns sign changes of the Möbius function in short intervals. Improving on many previous works, Matömaki and Radziwiłł have shown in [MR16] that for any $\eta > 1/2$, and any large enough positive integer N, there exist integers a, b with $|a|, |b| \leq N^\eta$ such that $\mu(N+a) = 1$, $\mu(N+b) = -1$. In characteristic 3, we show that the exponent $1/2$ can be improved to $3/7$.

Theorem 1.11. *Let q be a power of 3, and fix $3/7 < \eta < 1$. Then for any $f \in \mathbb{F}_q[T]^+$ of large enough norm, there exist $g, h \in \mathbb{F}_q[T]$ with $|g|, |h| \leq |f|^\eta$ such that $\mu(f+g) = 1$ and $\mu(f+h) = -1$.*

We follow the same proof strategy relating the Möbius function to characters, but since we allow small values of q, we cannot apply Theorem 1.4 anymore. Now however, once we are interested in sign change only (and not cancellation), we can focus on just one of the derivatives appearing. It turns out that if this derivative has a relatively large order of vanishing at 0, the conductor of the associated character is relatively small, and we can apply a function field version (see [Hsu99]) of the aforementioned result of Burgess. For $p > 3$, the arising character sums are too short for the Burgess bound to apply.

For q a large enough power of 3, Theorem 1.11 follows from Theorem 1.7 (since $T \mapsto 1/T$ allows one to think of short intervals as arithmetic progressions), and also from the $k = 1$ case of Theorem 1.3 (as we have sufficient uniformity in the shift).

Our last result, in the spirit of [Gal72], shows that the Möbius function enjoys cancellation in the arithmetic progression 1 mod a growing power of a fixed prime P, no matter how slowly does the length of the progression increase.

Theorem 1.12. *Let q be a power of 3. Fix an irreducible $P \in \mathbb{F}_q[T]$. Then for a positive integer n we have*

$$\sum_{\substack{f \in \mathbb{F}_q[T]^+ \\ |f| \leq X \\ f \equiv 1 \bmod P^n}} \mu(f) = o\left(\frac{X}{|P|^n}\right), \quad \frac{X}{|P|^n} \to \infty. \tag{1.21}$$

An analog of [Gal72, Theorem 2] in our setting would be cancellation for progressions whose length $X/|P|^n$ is at least $|X|^{3/5+\epsilon}$.

As before, we can obtain a power saving in Eq. (1.21). Since the progressions are so short, this result does not follow from the previous ones, even if q is large. On probabilistic grounds, one should not expect to obtain the theorem for all residue classes.

1.5. **Further directions.** In future work, we hope to use some of the methods introduced here to address the following problems:

- Obtaining cancellation in 'polynomial Möbius sums' such as

$$\sum_{\substack{f \in \mathbb{F}_q[T]^+ \\ |f| \leq X}} \mu\left(f^2 + T\right)$$

  which is relevant for counting primes of the form $f^2 + T$. We refer to our subsequent work [SS20] where this is carried out, culminating in a nonsplit analog of Theorem 1.1 and Theorem 1.3. In that work we also obtain analogs of Theorem 1.4 and Theorem 1.8 with certain more general 'trace functions' replacing the multiplicative/additive characters.

- Obtaining an asymptotic for the variance (and higher moments) of the Möbius function in short intervals (and arithmetic progressions) of length H for polynomials in $\mathbb{F}_q[T]$ of norm X, with H as close to X as possible. For sufficiently short intervals, namely with H a sufficiently small power of X (depending on q), an asymptotic for the variance follows from (a version with power savings of) Theorem 1.3 using the arguments of [MS04].

1.6. **Notation.** From this point on, it will be more convenient to work with degrees of polynomials instead of absolute values. For $g \in \mathbb{F}_q[T]$ we denote its degree by $d(g)$. By convention, the latter is $-\infty$ if $g = 0$. The letter q denotes a prime power, and is often suppressed from notation such as

$$\mathcal{M}_d = \left\{g \in \mathbb{F}_q[T]^+ : d(g) = d\right\}. \tag{1.22}$$

2. Character sums

We will make use of ℓ-adic sheaves, see for instance [KR14]. Often we apply the proper base change theorem, see [Mil, Theorem 17.10]. Another tool we use is the theory of vanishing cycles, see [Saw20].

The main result of this section is the following.

Theorem 2.1. *Let $t \leq m$ be natural numbers, let $f \in \mathbb{F}_q[T]$, let $g \in \mathcal{M}_m$ be squarefree, and let $\chi : (\mathbb{F}_q[T]/g)^\times \to \mathbb{C}^\times$ be a nontrivial character. Then*

$$\Big| \sum_{\substack{h \in \mathbb{F}_q[T] \\ d(h)<t \\ \gcd(f+h,g)=1}} \chi(f+h) \Big| \leq (q^{1/2}+1)\binom{m-1}{t} q^{\frac{t}{2}}. \tag{2.1}$$

To prove this theorem, we use the following geometric setup. View $\mathbb{A}^t$ as a space parameterizing polynomials h of degree less than t via their coefficients, namely a point $(a_0, \dots, a_{t-1})$ in $\mathbb{A}^t$ corresponds to the polynomial $h = a_0 + \cdots + a_{t-1}T^{t-1}$. Let (c_1, c_2) be coordinates on $\mathbb{A}^2$. Let $U \subseteq \mathbb{A}^t \times \mathbb{A}^2$ be the open set consisting of points $(h, (c_1, c_2))$ where $c_1 f + h + c_2 T^t$ is coprime to g. Let $j : U \to \mathbb{P}^t \times \mathbb{A}^2$ be the open immersion, embedding $\mathbb{A}^t$ into $\mathbb{P}^t$ in the usual way. Let $\pi : \mathbb{P}^t \times \mathbb{A}^2 \to \mathbb{A}^2$ be the projection.

Let $\mathcal{T}$ parametrize polynomials of degree less than m that are coprime to g, namely $\mathcal{T}$ is the open affine subscheme

$$\mathcal{T} = \{(b_0, \dots, b_{m-1}) : \mathrm{Res}(b_0 + \cdots + b_{m-1}T^{m-1}, g) \neq 0\}$$

of m-dimensional affine space. Multiplication of polynomials followed by reduction mod g endows $\mathcal{T}$ with the structure of a commutative algebraic group over $\mathbb{F}_q$, since inversion mod g can be written as a rational function of $b_0, \dots, b_{m-1}$ over $\mathbb{F}_q$ whose denominator is the above resultant. For every $\mathbb{F}_q$-algebra A, we therefore have $\mathcal{T}(A) = (A[T]/(g))^\times$.

Let $\alpha_1, \dots, \alpha_m \in \overline{\mathbb{F}_q}$ be the (distinct) roots of g. The evaluation map

$$b_0 + \cdots + b_{m-1}T^{m-1} \mapsto (b_0 + \cdots + b_{m-1}\alpha_1^{m-1}, \dots, b_0 + \cdots + b_{m-1}\alpha_m^{m-1})$$

is an isomorphism between $\mathcal{T}$ and $\mathbb{G}_m^m$, so $\mathcal{T}$ is a torus.

Fix a prime number ℓ different from the characteristic p of $\mathbb{F}_q$. Using the Lang isogeny, see [SGA4$\frac{1}{2}$, p. 171], one associates to our character χ: $\mathcal{T}(\mathbb{F}_q) \to \mathbb{C}^\times$ a $\overline{\mathbb{Q}_\ell}$-'character sheaf' $\mathcal{L}_\chi$ on $\mathcal{T}$ whose trace function is χ, namely

$$\iota\left(\mathrm{tr}\left(\mathrm{Frob}_q, (\mathcal{L}_\chi)_{\bar{f}}\right)\right) = \chi(f) \tag{2.2}$$

where $f \in \mathcal{T}(\mathbb{F}_q)$ is a polynomial, the map ι is an isomorphism between $\overline{\mathbb{Q}_\ell}$ and $\mathbb{C}$ (which will be tacitly used to identify the two fields in the sequel), and $(\mathcal{L}_\chi)_{\bar{f}}$ is the stalk of $\mathcal{L}_\chi$ at a geometric point $\bar{f}$ over f. In what follows, the passage to geometric points will at times not be explicit in our notation.

Let $\mathcal{L}_\chi(c_1 f + h + c_2 T^t)$ be the pullback of $\mathcal{L}_\chi$ to U along the natural map ν from U to $\mathcal{T}$ that sends $(h, (c_1, c_2))$ to $c_1 f + h + c_2 T^t$. The extension by

zero $j_!\mathcal{L}_\chi(c_1 f + h + c_2 T^t)$ of $\mathcal{L}_\chi(c_1 f + h + c_2 T^t)$ is lisse on U and tamely ramified along the complement of U in $\mathbb{P}^t \times \mathbb{A}^2$ since the monodromy of $\mathcal{L}_\chi$ is a finite group whose order is the order of χ, a number not divisible by p.

We will prove (2.1) using geometric properties of the (derived pushforward) complex $R\pi_* j_! \mathcal{L}_\chi(c_1 f + h + c_2 T^t)$ of (constructible) sheaves on $\mathbb{A}^2$. The relevance of this complex to the sum in Eq. (2.1) is explained in the proof of Theorem 2.1. (We write $R\pi_* j_!$ rather than the equivalent $R\pi_!$ because the key vanishing cycles steps of the argument use that factorization.)

Lemma 2.2. *The complex $R\pi_* j_! \mathcal{L}_\chi(c_1 f + h + c_2 T^t)$ is geometrically isomorphic to its pullback under the map $M_\lambda \colon \mathbb{A}^2 \to \mathbb{A}^2$ given by*

$$M_\lambda(c_1, c_2) = (\lambda c_1, \lambda c_2)$$

for any $\lambda \in \overline{\mathbb{F}_q}^\times$.

Proof. Let $M'_\lambda \colon U \to U$ be the map given by $M'_\lambda(h, (c_1, c_2)) = (\lambda h, (\lambda c_1, \lambda c_2))$. Applying the smooth base change theorem [Mil, Theorem 20.1] to the Cartesian squares

$$\begin{array}{ccc} \mathbb{P}^t \times \mathbb{A}^2 & \xrightarrow{\mathrm{id}_{\mathbb{P}^t} \times M_\lambda} & \mathbb{P}^t \times \mathbb{A}^2 \\ \downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \pi} \\ \mathbb{A}^2 & \xrightarrow{M_\lambda} & \mathbb{A}^2 \end{array} \qquad \begin{array}{ccc} U & \xrightarrow{M'_\lambda} & U \\ \downarrow{\scriptstyle j} & & \downarrow{\scriptstyle j} \\ \mathbb{P}^t \times \mathbb{A}^2 & \xrightarrow{\mathrm{id}_{\mathbb{P}^t} \times M_\lambda} & \mathbb{P}^t \times \mathbb{A}^2 \end{array}$$

we see that it suffices to check that $\mathcal{L}_\chi(c_1 f + h + c_2 T^t)$ is geometrically isomorphic to its pullback under M'_λ. For the map $T_\lambda \colon \mathcal{T} \to \mathcal{T}$ given by $T_\lambda(b_0 + \cdots + b_{m-1} T^{m-1}) = \lambda b_0 + \cdots + \lambda b_{m-1} T^{m-1}$ we have the commutative diagram

$$\begin{array}{ccc} U & \xrightarrow{\nu} & \mathcal{T} \\ \downarrow{\scriptstyle M'_\lambda} & & \downarrow{\scriptstyle T_\lambda} \\ U & \xrightarrow{\nu} & \mathcal{T} \end{array}$$

so it is sufficient that $\mathcal{L}_\chi$ on $\mathcal{T}$ is geometrically isomorphic to its pullback under T_λ, which follows from its construction as a character sheaf. □

Lemma 2.3. *The stalk of $R\pi_* j_! \mathcal{L}_\chi(c_1 f + h + c_2 T^t)$ at the point $(0, 1)$ is supported in degree t, where it has rank $\binom{m-1}{t}$.*

Proof. When $c_1 = 0, c_2 = 1$, the polynomial $c_1 f + h + c_2 T^t = T^t + h$ is monic and has degree t. By the proper base change theorem, our stalk is thus isomorphic to the compactly supported cohomology of the space of degree t monic polynomials that are prime to g, with coefficients in $\mathcal{L}_\chi(T^t + h)$. We may view this space as the quotient C^t / S_t where $C = \operatorname{Spec} \mathbb{F}_q[x, g(x)^{-1}]$, and denote by

$$\rho \colon C^t \to C^t / S_t \tag{2.3}$$

the quotient map. For any sheaf $\mathcal{F}$ on C^t/S_t, there is a natural action of S_t on $\rho_*\rho^*\mathcal{F}$. We define the S_t-invariants $(\rho_*\rho^*\mathcal{F})^{S_t}$ by taking the S_t-invariants of the sections on each open set. Then $(\rho_*\rho^*\overline{\mathbb{Q}_\ell})^{S_t} = \overline{\mathbb{Q}_\ell}$, so

$$\left(\rho_*\rho^*\mathcal{L}_\chi(T^t+h)\right)^{S_t} = \mathcal{L}_\chi(T^t+h) \tag{2.4}$$

and thus

$$\begin{aligned} &H_c^*\left(C^t/S_t, \mathcal{L}_\chi(T^t+h)\right) = \\ &H_c^*\left(C^t/S_t, \left(\rho_*\rho^*\mathcal{L}_\chi(T^t+h)\right)^{S_t}\right) = \\ &H_c^*\left(C^t/S_t, \rho_*\rho^*\mathcal{L}_\chi(T^t+h)\right)^{S_t} = \\ &H_c^*\left(C^t, \rho^*\mathcal{L}_\chi(T^t+h)\right)^{S_t}. \end{aligned} \tag{2.5}$$

For the second equality, note that taking S_t-invariants commutes with pushforward by definition. Thus their derived functors commute, but since we work with $\mathbb{Q}_\ell$-sheaves, taking derived S_t-invariants is the same as taking S_t-invariants. Similarly S_t-invariants commute with extension by zero and thus with compactly-supported cohomology.

Now ρ is the map defined by factorizing a polynomial into linear terms, so $\rho^*\mathcal{L}_\chi(T^t+h) = (\mathcal{L}_\chi(T-x))^{\boxtimes t}$. By the Künneth formula [SGA4-3, XVII, Theorem 5.4.3], it follows that the stalk of $R\pi_* j_! \mathcal{L}_\chi(c_1 f + h + c_2 T^t)$ at $(0,1)$ is

$$\left(\left(H_c^*\left(C, \mathcal{L}_\chi(T-x)\right)\right)^{\otimes t}\right)^{S_t}. \tag{2.6}$$

Because χ is nontrivial, $\mathcal{L}_\chi(T-x)$ has nontrivial monodromy, so

$$H_c^*\left(C, \mathcal{L}_\chi(T-x)\right) \tag{2.7}$$

vanishes in degrees other than 1. Because it arises from a character sheaf on a torus, $\mathcal{L}_\chi$ has tame local monodromy, so the rank of this cohomology group in degree 1 is the Euler characteristic of $\operatorname{Spec}\overline{\mathbb{F}}_q[x, g(x)^{-1}]$, which is $m-1$ [Ray66, Theorem 1]. Thus

$$\left(H_c^*\left(C, \mathcal{L}_\chi(T-x)\right)\right)^{\otimes t} \tag{2.8}$$

is supported in degree t, where it equals the t-th tensor power of an $(m-1)$-dimensional vector space, with S_t acting the usual way, twisted by the sign character because of the Koszul sign in the tensor product of a derived category. (In other words, because the isomorphism with the tensor product arises, after extension by zero, by pulling back classes along the different projections to C and taking the cup product. The pullback map is compatible with permuting the factors of C, but the cup product is multiplied by -1 each time we swap the order of two factors in degree 1 because cup produt is graded-commutative. This makes the isomorphism S_t-equivariant up to twisting by the sign representation.) Thus taking S_t-invariants is equivalent to taking the t-th wedge power of this $(m-1)$-dimensional vector space, which has dimension $\binom{m-1}{t}$. □

Lemma 2.4. *The complement of U in $\mathbb{P}^t \times \mathbb{A}^2$ is a divisor with simple normal crossings relative to $\mathbb{A}^2$ away from*

$$\left\{(h,(c_1,c_2)) \in \mathbb{A}^t \times \mathbb{A}^2 : d\left(\gcd(c_1 f + h + c_2 T^t, g)\right) > t\right\}. \tag{2.9}$$

Proof. Because this is a purely geometric question, we may assume that g splits completely, and let $\alpha_1, \dots, \alpha_m$ be its roots. Then the complement of U is the union of the hyperplane H_∞ at ∞ in $\mathbb{P}^t$ with the hyperplanes

$$H_i = \{(h,(c_1,c_2)) \in \mathbb{A}^t \times \mathbb{A}^2 : c_1 f(\alpha_i) + h(\alpha_i) + c_2 \alpha_i^t = 0\},\ 1 \le i \le m. \tag{2.10}$$

This union has simple normal crossings if the intersection of each k distinct hyperplanes from $\{H_1, \dots, H_m, H_\infty\}$ has codimension k.

We first consider the case of a subset not including the hyperplane H_∞. For any $S \subseteq \{1, \dots, m\}$ with $|S| = k$, we have

$$\bigcap_{i \in S} H_i = \left\{ (h,(c_1,c_2)) \in \mathbb{A}^t \times \mathbb{A}^2 : \prod_{i \in S} (T - \alpha_i) \;\middle|\; c_1 f + h + c_2 T^t \right\}. \tag{2.11}$$

Since $\prod_{i \in S}(T - \alpha_i)$ is a polynomial of degree k, the above has codimension k for any c_1, c_2 as long as $k \le t$, because in this case

$$\dim \left\{ h \in \mathbb{A}^t : h \equiv -c_1 f - c_2 T^t \mod \prod_{i \in S} (T - \alpha_i) \right\} = t - k.$$

On the other hand, when $k > t$ we get that

$$d\left(\gcd(c_1 f + h + c_2 T^t, g)\right) \ge k > t, \tag{2.12}$$

so removing these points, we obtain simple normal crossings.

Now we consider the case where we have a k-element set of (distinct) hyperplanes indexed by $S \subseteq \{1, \dots, m\}$ and also H_∞. We can take coordinates on H_∞ to be $(h,(c_1,c_2))$, with h a nonzero polynomial of degree less than t, well-defined up to scaling. In these coordinates, the equation for the intersection of any hyperplane with H_∞ is its original equation with all terms having degree zero in h removed. Thus the equation for $H_i \cap H_\infty$ is simply $h(\alpha_i) = 0$, so the intersection of H_∞ with the hyperplanes indexed by S consists of those $(h,(c_1,c_2))$ where h is a multiple of $\prod_{i \in S}(T - \alpha_i)$. Such an h exists only if $k < t$, as k is the degree of this polynomial and $d(h) < t$, but then our intersection always has codimension k in H_∞ hence codimension $k + 1$ overall, so we have simple normal crossings. □

Lemma 2.5. *Away from a finite union of lines through the origin, the complex $R\pi_* j_! \mathcal{L}_\chi(c_1 f + h + c_2 T^t)$ is supported in degree t with rank $\binom{m-1}{t}$.*

Proof. Let $l_1 \colon \mathbb{P}^t \times \mathbb{A}^1 \to \mathbb{P}^t \times \mathbb{A}^2$ be the direct product of the identity map on $\mathbb{P}^t$ and the map sending $c \in \mathbb{A}^1$ to $(c,1) \in \mathbb{A}^2$. We consider the vanishing cycles at zero $R\Phi_c l_1^* j_! \mathcal{L}_\chi(c_1 f + h + c_2 T^t)$ of the pullback of $j_! \mathcal{L}_\chi(c_1 f + h + c_2 T^t)$ under l_1. Let us first check that the complement of U is a simple normal crossings divisor everywhere in the fiber over zero in $\mathbb{A}^1$. To do this, we apply

Lemma 2.4 and observe that we cannot have $d(\gcd(c_1 f + h + c_2 T^t, g)) > t$ in this fiber as $c_1 = 0, c_2 = 1$ and

$$d\left(\gcd(h + T^t, g)\right) \leq d(h + T^t) \leq t. \tag{2.13}$$

Since the complement of U is a simple normal crossings divisor in the fiber over zero, and $j_!\mathcal{L}_\chi(c_1 f + h + c_2 T^t)$ is lisse on U and tamely ramified along the complement of U, it follows from [SGA7-II, XIII Lemma 2.1.11] that the vanishing cycles

$$R\Phi_c l_1^* j_! \mathcal{L}_\chi(c_1 f + h + c_2 T^t) \tag{2.14}$$

vanish everywhere. Hence the cohomology of the nearby fiber is isomorphic to the cohomology of the general fiber. Using Lemma 2.3 to compute the cohomology of the special fiber, it follows that the cohomology of the general fiber is supported in degree t with rank $\binom{m-1}{t}$. By constructibility, the stalk cohomology must have the same description at every point in some nonempty open set. By Lemma 2.2, the same description holds for the stalks in the $\mathbb{G}_m$-orbit of this open set, which is the complement of finitely many lines. □

Let $l_2 \colon \mathbb{P}^t \times \mathbb{A}^1 \to \mathbb{P}^t \times \mathbb{A}^2$ be the direct product of the identity map on $\mathbb{P}^t$ and the map sending $c \in \mathbb{A}^1$ to $(1, c) \in \mathbb{A}^2$.

Lemma 2.6. *Taking vanishing cycles at zero, the complex*

$$\mathcal{K} = R\Phi_c l_2^* j_! \mathcal{L}_\chi(c_1 f + h + c_2 T^t) \tag{2.15}$$

has the following properties:

- *it is supported on* $C = \left\{(h, 0) \in \mathbb{A}^t \times \mathbb{A}^1 : d\left(\gcd(f + h, g)\right) > t\right\}$;
- *it is supported in degree* t;
- *the rank of its stalk at* $(h, 0) \in C$ *is* $\binom{d(\gcd(f+h,g))-1}{t}$.

Proof. The first property follows immediately from [SGA7-II, XIII Lemma 2.1.11] and Lemma 2.4.

We note that $l_2^* j_! \mathcal{L}_\chi(c_1 f + h + c_2 T^t)[t + 1]$ is the extension by zero of a lisse sheaf in degree $-(t + 1)$ on a variety of dimension $t + 1$ and is thus semiperverse. The dual complex is the pushforward of a lisse sheaf in degree $-(t+1)$ on a variety of dimension $t+1$ along an affine open immersion and thus is semiperverse by Artin's affine theorem [BBD82, Theorem 4.1.1]. We conclude that

$$l_2^* j_! \mathcal{L}_\chi(c_1 f + h + c_2 T^t)[t + 1] \tag{2.16}$$

is a perverse sheaf. By [Ill94, Corollary 4.6], the vanishing cycles of a perverse sheaf are perverse up to a shift by one, so

$$\mathcal{K}[t] = R\Phi_c l_2^* j_! \mathcal{L}_\chi(c_1 f + h + c_2 T^t)[t] \tag{2.17}$$

is perverse.

The support of $\mathcal{K}[t]$ is the closed set C. Because C does not intersect the divisor at ∞, C is finite. A perverse sheaf supported on a finite set is necessarily a sum of skyscraper sheaves supported in degree zero, so we obtain the second property in the statement of this lemma.

It remains to calculate the rank of the stalk at a particular point $(h_0, 0) \in C$. We do that by working locally in an étale neighborhood.

We set

$$g' = \gcd(f + h_0, g), \quad g^* = g/g', \tag{2.18}$$

and factor $\mathcal{T}$ as the product of the torus $\mathcal{T}'$ of residue classes mod g' and the torus $\mathcal{T}^*$ of residue classes mod g^*. This lets us factor $\mathcal{L}_\chi$ as the tensor product of $\mathcal{L}_{\chi'}$ (the pullback of a character sheaf from $\mathcal{T}'$) and $\mathcal{L}_{\chi^*}$ (the pullback of a character sheaf from $\mathcal{T}^*$). Because $f + h_0$ is relatively prime to g^*, the map $U \to \mathcal{T} \to \mathcal{T}^*$ extends to a well-defined map in a neighborhood of the point $(h_0, 0)$, so $\mathcal{L}_{\chi^*}(c_1 f + h + c_2 T^t)$ extends to a lisse sheaf in a neighborhood of $(h_0, 0)$. Because tensoring with a lisse rank one sheaf does not affect vanishing cycles, it suffices to calculate

$$R\Phi_c l_2^* j'_! \mathcal{L}_{\chi'}(c_1 f + h + c_2 T^t), \tag{2.19}$$

where j' is the inclusion of the open set where $\gcd(c_1 f + h + c_2 T^t, g') = 1$.

By changing variables, we may replace (f, h) with f', h' where $f' = f + h_0$ and $h' = h - h_0$. We are then tasked with calculating the vanishing cycles $R\Phi_c l_2^* j'_! \mathcal{L}_{\chi'}(c_1 f' + h' + c_2 T^t)$ at zero. Having done this, we observe that f' is a multiple of g', so translation by f' does not affect $\mathcal{L}_{\chi'}(c_1 f' + h' + c_2 T^t)$, thus these vanishing cycles are the same as $R\Phi_c l_2^* j'_! \mathcal{L}_{\chi'}(h' + c_2 T^t)$.

As $d(h' + c_2 T^t) \leq t$, we can only have

$$d\left(\gcd(h' + c_2 T^t, g')\right) > t \tag{2.20}$$

if $h' + c_2 T^t = 0$. Hence, by Lemma 2.4, the complement of the image of j' is a simple normal crossings divisor way from the point $(0, 0)$. Therefore, by [SGA7-II, XIII Lemma 2.1.11], the vanishing cycles are supported at this point.

For a geometric generic point η of $\mathbb{A}^1$ we have a vanishing cycles long exact sequence

$$\begin{aligned}(R\pi_* j'_! \mathcal{L}_{\chi'}(h' + c_2 T^t))_{(0,0)} &\to (R\pi_* j'_! \mathcal{L}_{\chi'}(h' + c_2 T^t))_{(0,\eta)} \\ &\to H^*(\mathbb{P}^t, R\Phi_c l_2^* j'_! \mathcal{L}_{\chi'}(h' + c_2 T^t)).\end{aligned}$$

Therefore the Euler characteristic of the vanishing cycles complex is the difference between the Euler characteristic of the generic fiber and the Euler characteristic of the special fiber. Because the vanishing cycles complex is supported at a single point in a single degree, its Euler characteristic is $(-1)^t$ times its rank at that point. We will calculate these Euler characteristics and thereby calculate the rank.

By Lemma 2.5, the Euler characteristic of the generic fiber is $(-1)^t \binom{d(g')-1}{t}$. So it remains to check the Euler characteristic at the special point is zero. Because $\mathcal{L}_\chi$ is lisse of rank one and tame, the Euler characteristic of the special fiber is the Euler characteristic of the space of polynomials of degree less than t and prime to g'. Because this admits a free action of $\mathbb{G}_m$ by scaling, its Euler characteristic is zero in view of [BB73, Corollary 2]. $\square$

We can now prove Theorem 2.1.

Proof. In view of Eq. (2.2) we have

$$\sum_{\substack{h\in\mathbb{F}_q[T]\\ d(h)<t\\ \gcd(f+h,g)=1}} \chi(f+h) = \sum_{\substack{h\in\mathbb{F}_q[T]\\ d(h)<t\\ \gcd(f+h,g)=1}} \operatorname{tr}\left(\operatorname{Frob}_q, (\mathcal{L}_\chi)_{f+h}\right).$$

From the definition of the maps j and π we get

$$\begin{aligned}\sum_{\substack{h\in\mathbb{F}_q[T]\\ d(h)<t\\ \gcd(f+h,g)=1}} \operatorname{tr}\left(\operatorname{Frob}_q, (\mathcal{L}_\chi)_{f+h}\right) &= \sum_{\substack{u\in U(\mathbb{F}_q)\\ \pi(u)=(1,0)}} \operatorname{tr}\left(\operatorname{Frob}_q, (\mathcal{L}_\chi(c_1 f+h+c_2T^t))_u\right)\\ &= \sum_{\substack{x\in(\mathbb{P}^t\times\mathbb{A}^2)(\mathbb{F}_q)\\ \pi(x)=(1,0)}} \operatorname{tr}\left(\operatorname{Frob}_q, (j_!\mathcal{L}_\chi(c_1 f+h+c_2T^t))_x\right).\end{aligned}$$

Abusing notation, we denote the fiber of π over $(1,0)$ also by $\mathbb{P}^t$. The Grothendieck–Lefschetz fixed point formula [SGA4$\frac{1}{2}$, Rapport, Theorem 3.2] then gives

$$\begin{aligned}&\sum_{x\in\mathbb{P}^t(\mathbb{F}_q)} \operatorname{tr}\left(\operatorname{Frob}_q, (j_!\mathcal{L}_\chi(c_1 f+h+c_2T^t))_x\right) =\\ &\sum_{i=0}^{2t}(-1)^i \operatorname{tr}\left(\operatorname{Frob}_q, H^i(\mathbb{P}^t, j_!\mathcal{L}_\chi(c_1 f+h+c_2T^t))\right).\end{aligned}$$

By Deligne's Riemann hypothesis [Del80, Corollary 3.3.4], the absolute values of the eigenvalues of Frob_q on the i-th cohomology group are bounded from above by $q^{i/2}$ so

$$\begin{aligned}&\left|\sum_{i=0}^{2t}(-1)^i \operatorname{tr}\left(\operatorname{Frob}_q, H^i(\mathbb{P}^t, j_!\mathcal{L}_\chi(c_1 f+h+c_2T^t))\right)\right| \le\\ &\sum_{i=0}^{2t} q^{i/2}\dim H^i(\mathbb{P}^t, j_!\mathcal{L}_\chi(c_1 f+h+c_2T^t)).\end{aligned}$$

Applying the proper base change theorem to the Cartesian square

$$\begin{array}{ccc}\mathbb{P}^t & \hookrightarrow & \mathbb{P}^t\times\mathbb{A}^2\\ \downarrow & & \downarrow{\scriptstyle\pi}\\ \operatorname{Spec}\overline{\mathbb{F}_q} & \longrightarrow & \mathbb{A}^2\end{array}$$

where the upper horizontal map is the immersion of the fiber of π over $(1,0)$, and the lower horizontal map corresponds to the point $(1,0)\in\mathbb{A}^2$, we get that $H^i(\mathbb{P}^t, j_!\mathcal{L}_\chi(c_1 f+h+c_2T^t)) \cong (R\pi_* j_!\mathcal{L}_\chi(c_1 f+h+c_2T^t))_{(1,0)}$.

For a geometric generic point η of $\mathbb{A}^1$ we have a vanishing cycles long exact sequence

$$(R\pi_* j_! \mathcal{L}_\chi(c_1 f + h + c_2 T^t))_{(1,0)} \to (R\pi_* j_! \mathcal{L}_\chi(c_1 f + h + c_2 T^t))_{(1,\eta)} \\ \to H^*(\mathbb{P}^t, R\Phi_c l_2^* j_! \mathcal{L}_\chi(c_1 f + h + c_2 T^t)).$$

By Lemma 2.5, $(R\pi_* j_! \mathcal{L}_\chi(c_1 f + h + c_2 T^t))_{(1,\eta)}$ is supported in degree t with rank $\binom{m-1}{t}$. By Lemma 2.6, the complex $R\Phi_c l_2^* j_! \mathcal{L}_\chi(c_1 f + h + c_2 T^t)$ is supported in degree t and at finitely many points, so the third term above is also supported in degree t and is simply the sum of the stalks at those points, and thus has rank

$$r(f,g,t) = \sum_{\substack{h \in \overline{\mathbb{F}_q}[T] \\ d(h)<t \\ d(\gcd(f+h,g))>t}} \binom{d\left(\gcd(f+h,g)\right)-1}{t}, \tag{2.21}$$

again using Lemma 2.6.

We conclude that, upon suppressing $c_1 f + h + c_2 T^t$ for brevity, the vanishing cycles long exact sequence becomes

$$0 \to (R^t \pi_* j_! \mathcal{L}_\chi)_{(1,0)} \to (R^t \pi_* j_! \mathcal{L}_\chi)_{(1,\eta)} \\ \to H^t(\mathbb{P}^t, R\Phi_c l_2^* j_! \mathcal{L}_\chi) \to (R^{t+1} \pi_* j_! \mathcal{L}_\chi)_{(1,0)} \to 0. \tag{2.22}$$

Thus $(R\pi_* j_! \mathcal{L}_\chi(c_1 f + h + c_2 T^t))_{(1,0)}$ is supported in degrees t and $t+1$, with rank at most $\binom{m-1}{t}$ in degree t and rank bounded by $r(f,g,t)$ in degree $t+1$.

We thus have

$$\sum_{i=0}^{2t} q^{i/2} \dim H^i(\mathbb{P}^t, j_! \mathcal{L}_\chi(c_1 f + h + c_2 T^t)) \leq \binom{m-1}{t} q^{\frac{t}{2}} + r(f,g,t) q^{\frac{t+1}{2}}. \tag{2.23}$$

Finally, we check that $r(f,g,t) \leq \binom{m-1}{t}$. To do this fix a root α of g, and note that $\binom{d(\gcd(f+h,g))-1}{t}$ does not exceed the number of degree t divisors of g, prime to $T-\alpha$, that divide $f+h$. Each such divisor of g prime to $T-\alpha$ contributes at most once to the sum in Eq. (2.21), so this sum is bounded by the number of such divisors, which is $\binom{m-1}{t}$. □

Corollary 2.7. *Fix $\eta > 0$ and $0 < \beta < 1/2$. Then for a prime power $q \geq (e\eta^{-1})^{\frac{2}{1-2\beta}}$ the following holds. For a nonprincipal character χ to a squarefree modulus $g \in \mathbb{F}_q[T]$, $f \in \mathbb{F}_q[T]$, and $t \geq \eta \cdot d(g)$, we have*

$$\sum_{d(h)<t} \chi(f+h) \ll q^{(1-\beta)t} \tag{2.24}$$

with the implied constant depending only on q.

Furthermore, if we have $t \leq \eta \cdot d(g) \leq t'$, then we still have

$$\sum_{d(h)<t} \chi(f+h) \ll q^{(1-\beta)t'}. \tag{2.25}$$

Proof. If $\eta \geq 1$, the left hand side vanishes and the bound is trivial. Otherwise, we apply Theorem 2.1 to the left side, taking $m = d(g)$, to obtain

$$\begin{aligned} \sum_{d(h)<t} \chi(f+h) &\leq (q^{1/2+1})\binom{m-1}{t} q^{t/2} \\ &\ll \binom{m}{t} q^{t/2} \leq \left(\frac{t}{m}\right)^{-t} \left(\frac{m-t}{m}\right)^{t-m} q^{t/2} \end{aligned} \tag{2.26}$$

where the last inequality follows from

$$\begin{aligned} 1 = \left(\frac{t}{m} + \frac{m-t}{m}\right)^m &= \sum_{k=0}^{m} \binom{m}{k} \left(\frac{t}{m}\right)^k \left(\frac{m-t}{m}\right)^{m-k} \\ &\geq \binom{m}{t} \left(\frac{t}{m}\right)^t \left(\frac{m-t}{m}\right)^{m-t}. \end{aligned} \tag{2.27}$$

From the Taylor series we can see that $-\log(1-x) \leq x/(1-x)$ if $x > 0$ so $(1-x)^{-(1-x)/x} \leq e$. Applying this to $x = t/m$, we get

$$\left(\frac{m-t}{m}\right)^{t-m} \leq e^t \tag{2.28}$$

so we obtain

$$\sum_{d(h)<t} \chi(f+h) \ll \left(\frac{t}{m}\right)^{-t} e^t q^{t/2}. \tag{2.29}$$

Because $q \geq (e\eta^{-1})^{\frac{2}{1-2\beta}}$ and $t/m \geq \eta$, we have

$$e \cdot \left(\frac{t}{m}\right)^{-1} \leq e\eta^{-1} \leq q^{\frac{1}{2}-\beta} \tag{2.30}$$

hence

$$\left(\frac{t}{m}\right)^{-t} e^t q^{t/2} \leq q^{(1-\beta)t}, \tag{2.31}$$

as desired.

To handle the case where $t \leq \eta \cdot d(g) \leq t'$, first note that we may assume $t' \leq m$. We observe that the left hand side of Eq. (2.31) is an increasing function of t because its logarithm

$$t \log m - t \log t + t + t(\log q)/2 \tag{2.32}$$

has derivative

$$\log m - \log t + (\log q)/2 \tag{2.33}$$

which is positive in the range $t \leq m$. Thus we can get a bound for the shorter sum which is at least as good as our bound for the longer sum. □

3. The Möbius Function

From now on, we will assume that the characteristic p of $\mathbb{F}_q$ is odd. Because of this, $\mathbb{F}_q^\times$ admits a unique quadratic character, which we denote ψ. We use freely the basic properties of resultants (see [Jan07]) and the Jacobi symbol (see [Ros02, Chapter 3]). For example, the quadratic reciprocity law

$$\left(\frac{a}{b}\right)\left(\frac{b}{a}\right) = (-1)^{\frac{q-1}{2}mn}\psi(a_m)^{d(b)}\psi(b_n)^{d(a)}$$

where a (respectively, b) is a nonzero polynomial in $\mathbb{F}_q[T]$ of degree m (respectively, n) and leading coefficient a_m (respectively, b_n).

The following lemma recalls the relation between the (real valued) Jacobi symbol and the quadratic character of a resultant.

Lemma 3.1. *Let $f \in \mathbb{F}_q[T]$, $g = a_nT^n + \cdots + a_0$ of degree $n \geq 1$. Then*

$$\left(\frac{f}{g}\right) = \psi(a_n)^{\max\{d(f),0\}}\psi\left(\mathrm{Res}(g,f)\right). \tag{3.1}$$

Proof. Fix $f \neq 0$, and note that both sides above are completely multiplicative in g, so we may assume that g is irreducible. For a root θ of g we have

$$\mathrm{Res}(g,f) = a_n^{d(f)}\prod_{i=0}^{d(g)-1} f\left(\theta^{q^i}\right) = a_n^{d(f)}\prod_{i=0}^{d(g)-1} f(\theta)^{q^i} = a_n^{d(f)}f(\theta)^{\frac{q^{d(g)}-1}{q-1}} \tag{3.2}$$

so we get the mod p congruence

$$\begin{aligned}\psi\left(\mathrm{Res}(g,f)\right) &= \psi(a_n)^{d(f)}\psi\left(f(\theta)^{\frac{q^{d(g)}-1}{q-1}}\right)\\ &\equiv \psi(a_n)^{d(f)}f(\theta)^{\frac{q^{d(g)}-1}{2}} \equiv \psi(a_n)^{d(f)}\left(\frac{f}{g}\right)\end{aligned} \tag{3.3}$$

which implies the lemma. □

Given a $D \in \mathbb{F}_q[T]$ we write $\mathrm{rad}(D)$ for the product of the primes that appear in the factorization of D, and $\mathrm{rad}_1(D)$ for the product of the primes that appear with odd multiplicity in the factorization of D. The derivative of D (with respect to T) is denoted by D'.

The next lemma interprets the Möbius function (on an arithmetic progression) as a Dirichlet character that 'depends only on the derivative'.

Lemma 3.2. *Let $m, k \geq 0$, $d \geq 1$ be integers with $k \neq d+m$. For $M \in \mathcal{M}_m$, $g \in \mathcal{M}_d$, and $a \in \mathcal{M}_k$ coprime to M, define the polynomials*

$$D = M^2\left(g + \frac{a}{M}\right)', \quad E = \frac{\mathrm{rad}(D)}{\gcd\left(M, \mathrm{rad}(D)\right)}, \quad E_1 = \frac{\mathrm{rad}_1(D)}{\gcd\left(M, \mathrm{rad}_1(D)\right)}. \tag{3.4}$$

Then

$$\mu(a + gM) = S \cdot \chi(w + g) \tag{3.5}$$

where $w = w_{a,M,g'} \in \mathbb{F}_q[T]$, $\chi = \chi_{a,M,g'}$ *is a (real) multiplicative character mod* E *with conductor* E_1, *and*

$$S = S_{d,a,M,g'} \in \{0, 1, -1\} \tag{3.6}$$

with $S = 0$ *if and only if* $D = 0$.

Proof. Pellet's formula (see [Con05, Lemma 4.1]) gives

$$\mu(a + gM) = (-1)^{d(a+gM)} \psi\left(\mathrm{Disc}(a + gM)\right) \tag{3.7}$$

and since $d(a + gM) = \max\{k, d + m\}$, we see that $(-1)^{d(a+gM)}$ can be absorbed into S. Our assumption that $k \neq d + m$ implies $a + gM$ is monic, so $\psi\left(\mathrm{Disc}(a + gM)\right)$ equals, up to a sign that S absorbs,

$$\psi\left(\mathrm{Res}(a + gM, a' + g'M + gM')\right). \tag{3.8}$$

By Lemma 3.1, and the fact that $a + gM$ is monic, the above equals

$$\left(\frac{a' + g'M + gM'}{a + gM}\right) \tag{3.9}$$

and since $\gcd(a, M) = 1$ by multiplicativity, this equals

$$\left(\frac{a'M + g'M^2 + gMM'}{a + gM}\right) \left(\frac{M}{a + gM}\right)^{-1}. \tag{3.10}$$

Using quadratic reciprocity, we can absorb $\left(\frac{M}{a+gM}\right)$ into S. Subtracting $M'(a + gM)$ from the numerator of the first Jacobi symbol above, we get (in the notation of equation (3.4))

$$\left(\frac{D}{a + gM}\right) \tag{3.11}$$

and set $w = 0$, $S = 0$ in case $D = 0$. Otherwise (if $D \neq 0$) we apply quadratic reciprocity once again to obtain

$$\left(\frac{a + gM}{D}\right) \tag{3.12}$$

up to a sign that goes into S.

We write $\xi(a + gM)$ for the Jacobi symbol above, so that ξ is a multiplicative character mod $\mathrm{rad}(D)$ with conductor $\mathrm{rad}_1(D)$. Since $\mathrm{rad}(D)$ is squarefree, we see that M is coprime to E (from equation (3.4)). Therefore, by the Chinese remainder theorem, there exists a unique factorization $\xi = \xi_E \xi_N$ to characters mod E and $N = \gcd\left(M, \mathrm{rad}(D)\right)$ respectively. In this notation, our Jacobi symbol equals $\xi_E(a + gM)\xi_N(a + gM)$ and the second factor is simply $\xi_N(a)$, so we immerse it in S. Taking $\overline{M} \in \mathbb{F}_q[T]$ with $\overline{M}M \equiv 1\ (E)$ we can write

$$\xi_E(a + gM) = \xi_E(a\overline{M} + g)\xi_E(M) \tag{3.13}$$

and conclude by setting $w = a\overline{M}$, $\chi = \xi_E$, and dumping $\xi_E(M)$ into S. $\square$

Remark 3.3. With notation as in Lemma 3.2, suppose that χ is principal. We can then write $D = AB^2$ for some $A, B \in \mathbb{F}_q[T]$ with $A \mid M$, simply by taking $A = \mathrm{rad}_1(D)$ since principality gives $E_1 = 1$.

4. Linear forms in the Möbius function

Proposition 4.1. *Let p be a prime, let q be a power of p, let m, d be non-negative integers, let $M \in \mathcal{M}_m$ be squarefree, and let $a \in \mathbb{F}_q[T]$. Then*

$$\frac{\#\left\{h : h = g' \text{ for some } g \in \mathcal{M}_d,\ h \equiv a \bmod M\right\}}{\#\left\{g' : g \in \mathcal{M}_d\right\}} \leq q^{-\min\left\{m, \left\lfloor \frac{d-1}{p} \right\rfloor\right\}}.$$

Proof. Let $0 \leq j \leq p-1$ be the unique integer congruent to d mod p. Any $g \in \mathcal{M}_d$ can then be uniquely expressed as

$$g = \sum_{i=0}^{p-1} T^i g_i^p, \quad g_j \in \mathcal{M}_{\frac{d-j}{p}}, \quad d(g_i) \leq \left\lfloor \frac{d-i}{p} \right\rfloor \text{ for } i \neq j. \tag{4.1}$$

For the derivative we then have

$$g' = \sum_{i=1}^{p-1} i T^{i-1} g_i^p \tag{4.2}$$

so we set $a_i = iT^{i-1}$ for $1 \leq i \leq p-1$, and consider the congruence

$$\sum_{i=1}^{p-1} a_i g_i^p \equiv a \bmod M. \tag{4.3}$$

Since $a_1 = 1$, we have $g_1^p \equiv a - \sum_{i=2}^{p-1} a_i g_i^p \equiv a \bmod M$, and thus the value of $g_1^p \bmod M$ is uniquely determined by $g_2, \ldots, g_{p-1}$. Since M is squarefree, the pth power map mod M is injective, and thus $g_1 \bmod M$ is uniquely determined by $g_2, \ldots, g_{p-1}$.

This implies the bound: If $m \geq \left\lfloor \frac{d-1}{p} \right\rfloor$ then a fraction of $\frac{1}{q^m}$ of values of g_1 satisfy this congruence, and if $m < \left\lfloor \frac{d-1}{p} \right\rfloor$ then at most one, meaning a fraction $\frac{1}{q^{\left\lfloor \frac{d-1}{p} \right\rfloor}}$ of all values, satisfies it.

□

The following technical proposition follows from the arguments of [BGP92, Page 371] or [CG07, Section 9], which obtain stronger statements over $\mathbb{Z}$ in place of $\mathbb{F}_q[T]$.

Proposition 4.2. *Fix $\alpha, \epsilon > 0$, and a prime power q. Then for integers $d, m, k \geq 0$ with $d \geq \epsilon(m+k)$, and $M \in \mathcal{M}_m$, $A \in \mathcal{M}_k$, $a \in \mathbb{F}_q[T]$, we have*

$$\#\left\{g \in \mathbb{F}_q[T] : d(g) < d,\ a + gM = \lambda AB^2, \quad \lambda \in \mathbb{F}_q, B \in \mathbb{F}_q[T]\right\} \ll q^{(\frac{1}{2}+\alpha)d}$$

as $d \to \infty$, with the implied constant depending only on ϵ, α and q.

This proposition is also a special case of the remark following [BSE21, Theorem 1.1].

We can now prove Theorem 1.3. We first prove the "generic" special case where the derivatives of certain parameters are distinct, and then prove the general case.

Denote by d_2 the binary divisor function on $\mathbb{F}_q[T]$. Namely for a nonzero polynomial $M \in \mathbb{F}_q[T]$ we denote by $d_2(M)$ the number of distinct monic polynomials dividing M, that is

$$d_2(M) = \#\{h \in \mathbb{F}_q[T]^+ : h \mid M\}.$$

Proposition 4.3. *Fix $\epsilon, \delta > 0$, $0 < \beta < 1/2$, and a positive integer n. Let q be a power of an odd prime p such that*

$$q > \left(\frac{pne}{\min\left\{ \frac{\epsilon}{\epsilon+2}, \frac{\epsilon\delta}{\epsilon+\delta} \right\}} \right)^{\frac{2}{1-2\beta}}. \tag{4.4}$$

Then for nonnegative integers $d, m_1, \dots, m_n, k_1, \dots, k_n$ with

$$d \geq \max\{\epsilon m_1, \dots, \epsilon m_n, \delta k_1, \dots, \delta k_n\}, \quad k_i \neq d + m_i,\ 1 \leq i \leq n, \tag{4.5}$$

and pairs $(a_i, M_i) \in \mathcal{M}_{k_i} \times \mathcal{M}_{m_i}$ for $1 \leq i \leq n$ such that the derivatives $\left(\frac{a_i}{M_i}\right)'$ are all distinct, we have

$$\sum_{g \in \mathcal{M}_d} \prod_{i=1}^{n} \mu\left(a_i + g M_i\right) \ll |\mathcal{M}_d|^{1-\frac{\beta}{p}} \tag{4.6}$$

as $d \to \infty$, with the implied constant depending only on $\beta, \epsilon, \delta, n$ and q.

Remark 4.4. Note that the statement of this proposition remains meaningful even if ϵ and δ are very large, though it is at its strongest when ϵ and δ are small.

Proof. Let us first assume that for every $1 \leq i \leq n$ we have $\gcd(a_i, M_i) = 1$.

We say that $g_1, g_2 \in \mathcal{M}_d$ are equivalent if $g_1' = g_2'$, and let $\mathcal{R}$ be a complete set of representatives of equivalence classes. So for each $g \in \mathcal{M}_d$ there exists a unique $r \in \mathcal{R}$ such that $(g - r)' = 0$, and therefore also a unique $s \in \mathbb{F}_q[T]$ such that $g - r = s^p$. We can thus write our sum as

$$\sum_{r \in \mathcal{R}} \sum_{d(s) < t} \prod_{i=1}^{n} \mu\left(a_i + (r + s^p) M_i\right), \quad t = \frac{d}{p}. \tag{4.7}$$

By Lemma 3.2 (the notation of which is used throughout), our sum equals

$$\sum_{r \in \mathcal{R}} \prod_{i=1}^{n} S_r^{(i)} \sum_{d(s) < t} \prod_{i=1}^{n} \chi_r^{(i)}\left(w_r^{(i)} + s^p\right) \tag{4.8}$$

with $\chi_r^{(i)}$ a character to a squarefree modulus $E_r^{(i)}$ (defined in Lemma 3.2 using $D_r^{(i)}$). Hence, there exist $f_r^{(i)} \in \mathbb{F}_q[T]$ with

$$f_r^{(i)p} \equiv w_r^{(i)} \bmod E_r^{(i)}, \quad 1 \leq i \leq n, \tag{4.9}$$

so using the fact that $\chi_r^{(i)}$ is real, we see that our inner sum equals

$$\sum_{d(s)<t} \prod_{i=1}^{n} \chi_r^{(i)}\left(f_r^{(i)} + s\right). \tag{4.10}$$

For $1 \leq i < j \leq n$ we set

$$G_r^{(i,j)} = \gcd\left(E_r^{(i)}, E_r^{(j)}\right), \quad U_r = \mathrm{lcm}_{i,j}\left(G_r^{(i,j)}\right), \quad \ell_r = d(U_r), \tag{4.11}$$

and claim that, for every integer $\ell \geq 0$ and $\gamma > 0$, we have

$$\#\left\{r \in \mathcal{R} : \ell_r \geq \ell\right\} \ll q^{d\left(1-\frac{1}{p}+\gamma\right)-\min\left(\frac{d-1}{p},\ell\right)}, \quad d \to \infty. \tag{4.12}$$

To prove this, first note that, for any $1 \leq i < j \leq n$, since $G_r^{(i,j)}$ divides $E_r^{(i)}$, and divides $E_r^{(j)}$, it also divides $D_r^{(i)}$, and divides $D_r^{(j)}$ so therefore it divides the polynomial

$$M_j^2 D_r^{(i)} - M_i^2 D_r^{(j)} = M_i^2 M_j^2 \left(\frac{a_i}{M_i}\right)' - M_i^2 M_j^2 \left(\frac{a_j}{M_j}\right)' \tag{4.13}$$

which is nonzero by our initial assumption. The degree of the above polynomial is at most $d/\delta + 3d/\epsilon$, so by the divisor bound (see [IK04, Equation 1.81]), the polynomial $G_r^{(i,j)}$ attains $\ll q^{2\gamma d/(n(n-1))}$ values, for any $\gamma > 0$. Hence, the tuple $\left(G_r^{(i,j)}\right)_{1\leq i<j\leq n}$ attains $\ll q^{\gamma d}$ values. For each possible tuple $G_*^{(i,j)}$, we can recover the residue class of r' mod $G_*^{(i,j)}$ from the congruence

$$M_j^2 r' \equiv M_j' a_j - a_j' M_j \bmod G_r^{(i,j)}. \tag{4.14}$$

and the fact that $G_r^{(i,j)}$ is prime to M_j (because $E_r^{(j)}$ is prime to M_j, by definition). Combining these for all i, j we can also recover the residue class of r' mod the least common multiple U_* of $G_*^{(i,j)}$.

Proposition 4.1 tells us that for any $\alpha \in \mathbb{F}_q[T]$ we have

$$\#\{r \in \mathcal{R} : r' \equiv \alpha \bmod U_*\} \leq q^{-\min\left\{d(U_*),\left\lfloor\frac{d-1}{p}\right\rfloor\right\}} \#\mathcal{R} \ll q^{d\left(1-\frac{1}{p}\right)-\min\left(\frac{d-1}{p},\ell\right)}$$

since $d(U_*) \geq \ell$, so our claim is established.

Next, we observe from Eq. (4.12) that the contribution of those $r \in \mathcal{R}$ with $\ell_r \geq \frac{d-1}{p}$ to Eq. (4.7) is $\ll q^{d\left(1-\frac{1}{p}+\gamma\right)}$ and thus can be ignored as we can choose γ small enough that $1 - \frac{1}{p} + \gamma \leq 1 - \frac{\beta}{p}$. So we may assume that $\ell_r < \frac{d-1}{p}$.

We further set

$$\widetilde{E}_r^{(i)} = \gcd\left(E_r^{(i)}, U_r\right), \quad \widehat{E}_r^{(i)} = \frac{E_r^{(i)}}{\widetilde{E}_r^{(i)}}, \quad 1 \le i \le n, \tag{4.15}$$

and note that $\gcd(\widetilde{E}_r^{(i)}, \widehat{E}_r^{(i)}) = 1$ as $E_r^{(i)}$ is squarefree. The Chinese remainder theorem then gives a unique decomposition

$$\chi_r^{(i)} = \widetilde{\chi}_r^{(i)} \widehat{\chi}_r^{(i)}, \quad 1 \le i \le n, \tag{4.16}$$

to characters mod $\widetilde{E}_r^{(i)}$ and $\widehat{E}_r^{(i)}$ respectively. In this notation, our sum reads

$$\sum_{d(s)<t} \prod_{i=1}^{n} \widetilde{\chi}_r^{(i)}\left(f_r^{(i)} + s\right) \prod_{i=1}^{n} \widehat{\chi}_r^{(i)}\left(f_r^{(i)} + s\right) \tag{4.17}$$

so splitting according to the residue class u of s mod U_r we get

$$\sum_{d(u)<\ell_r} \prod_{i=1}^{n} \widetilde{\chi}_r^{(i)}\left(f_r^{(i)} + u\right) \sum_{d(h)<t-\ell_r} \prod_{i=1}^{n} \widehat{\chi}_r^{(i)}\left(f_r^{(i)} + u + hU_r\right). \tag{4.18}$$

Since $\gcd(\widetilde{E}_r^{(i)}, \widehat{E}_r^{(i)}) = 1$, we get that $\gcd(U_r, \widehat{E}_r^{(i)}) = 1$ for $1 \le i \le n$. Hence, there exist $V_r^{(i)} \in \mathbb{F}_q[T]$ with $U_r V_r^{(i)} \equiv 1 \bmod \widehat{E}_r^{(i)}$. Summing trivially over u, we may thus consider

$$\prod_{i=1}^{n} \widehat{\chi}_r^{(i)}(U_r) \sum_{d(h)<t-\ell_r} \prod_{i=1}^{n} \widehat{\chi}_r^{(i)}\left(f_r^{(i)} V_r^{(i)} + u V_r^{(i)} + h\right). \tag{4.19}$$

From $\gcd(\widetilde{E}_r^{(i)}, \widehat{E}_r^{(i)}) = 1$ we moreover conclude that $\{\widehat{E}_r^{(i)}\}_{i=1}^n$ are pairwise coprime. The Chinese remainder theorem then gives an $f_{r,u} \in \mathbb{F}_q[T]$ with

$$f_{r,u} \equiv f_r^{(i)} V_r^{(i)} + u V_r^{(i)} \bmod \widehat{E}_r^{(i)}, \quad 1 \le i \le n, \tag{4.20}$$

so defining the character $\chi_r = \widehat{\chi}_r^{(1)} \cdots \widehat{\chi}_r^{(n)}$, mod $E_r = \widehat{E}_r^{(1)} \cdots \widehat{E}_r^{(n)}$, the sum above becomes

$$\sum_{d(h)<t-\ell_r} \chi_r(f_{r,u} + h). \tag{4.21}$$

We have

$$\frac{t}{d(E_r)} \ge \frac{t}{n \max\{d + 2d/\epsilon, d/\delta + d/\epsilon\}} = \frac{\min\left\{\frac{\epsilon}{\epsilon+2}, \frac{\epsilon\delta}{\epsilon+\delta}\right\}}{pn} =: \eta \tag{4.22}$$

so if χ_r is nonprincipal, Corollary 2.7 bounds the sum above by $\ll q^{(1-\beta')t}$, for some $\beta' > \beta + p\gamma$, with $\gamma > 0$ arbitrarily small. Because we have

$$q > \left(\frac{pne}{\min\left\{\frac{\epsilon}{\epsilon+2}, \frac{\epsilon\delta}{\epsilon+\delta}\right\}}\right)^{\frac{2}{1-2\beta}} \tag{4.23}$$

by assumption, using our definition of η this can be writeen as

$$q > \left(e\eta^{-1}\right)^{\frac{2}{1-2\beta}} \tag{4.24}$$

and therefore

$$q > \left(e\eta^{-1}\right)^{\frac{2}{1-2\beta'}} \tag{4.25}$$

for any sufficiently small choice of γ.

Hence, those $r \in \mathcal{R}$ for which $\ell_r = \ell$ and χ_r is nonprincipal, contribute $\ll q^{(1-\beta')t}q^{\ell}$ individually, so the total contribution to our initial sum is $\ll$

$$q^{(1-\beta')t+\ell}q^{d\left(1-\frac{1}{p}+\gamma\right)-\ell} = q^{(1-\beta)t}q^{d\left(1-\frac{1}{p}\right)}q^{\gamma d-(\beta'-\beta)t} = q^{d\left(1-\frac{\beta}{p}\right)}q^{d(\gamma-(\beta'-\beta)/p)}.$$

The increase from summing over the possible values of ℓ is linear in d and thus can be bounded by the exponential $q^{d((\beta'-\beta)/p-\gamma)}$, so the contribution of all the terms where χ is nonprincipal is $\ll q^{d\left(1-\frac{\beta}{p}\right)}$.

Let $r \in \mathcal{R}$ for which χ_r is principal. Pairwise coprimality implies that $\widehat{\chi}_r^{(1)}$ is principal as well, so the conductor of $\chi_r^{(1)}$ divides that of $\widetilde{\chi}_r^{(1)}$, and the latter divides U_r. Thus, for some $\lambda \in \mathbb{F}_q$ and monic $B \in \mathbb{F}_q[T]$ we have

$$M_1^2 r' + a_1' M_1 - a_1 M_1' = D_r^{(1)} = \mathrm{rad}_1\left(D_r^{(1)}\right)\lambda B^2 = \lambda A\overline{A}B^2,\ A \mid M_1, \overline{A} \mid U_r$$

where A is the greatest common divisor of M_1 and $\mathrm{rad}_1(D_r^{(1)})$, and $\overline{A}$ is the conductor of $\chi_r^{(1)}$. Since U_r divides the nonzero polynomial Q defined to be

$$\prod_{1\le i<j\le n} M_i^2 M_j^2 \left[\left(\frac{a_i}{M_i}\right)' - \left(\frac{a_j}{M_j}\right)'\right], \tag{4.26}$$

it follows that $\overline{A} \mid Q$ as well, so Proposition 4.2 ensures that the number of r for which the equation above holds is

$$\ll d_2(M_1 Q)q^{(1/2+\zeta)d}. \tag{4.27}$$

The divisor bound then allows us to neglect those r for which χ_r is principal, as long as $\frac{1}{2}+\zeta+\frac{1}{p} < 1-\frac{\beta}{p}$, which is alright as $\beta < \frac{1}{2} \le p\left(\frac{1}{2}-\frac{1}{p}\right)$ so we can choose ζ small enough that this inequality holds.

Let us now handle the case when a_i and M_i are not coprime. Set

$$H_i = \gcd(a_i, M_i), \quad 1 \le i \le n. \tag{4.28}$$

If some H_i is not squarefree, the sum vanishes and the bound is trivial. Otherwise, we have the identity

$$\mu(a_i + gM_i) = \begin{cases} \mu\left(\frac{a_i}{H_i} + g\frac{M_i}{H_i}\right)\mu(H_i) & \text{if } \gcd\left(\frac{a_i}{H_i} + g\frac{M_i}{H_i}, H_i\right) = 1 \\ 0 & \text{if } \gcd\left(\frac{a_i}{H_i} + g\frac{M_i}{H_i}, H_i\right) \ne 1. \end{cases} \tag{4.29}$$

Thus we can write the sum as the constant factor $\prod_{i=1}^n \mu(H_i)$ times a similar sum, except that the degrees of a_i and M_i are reduced and the terms where $\gcd\left(\frac{a_i}{H_i} + g\frac{M_i}{H_i}, H_i\right) \ne 1$ are removed. The degree reduction preserves

the system of inequalities (4.5) and so is no trouble. Removing the terms with $\gcd\left(\frac{a_i}{H_i}+g\frac{M_i}{H_i}, H_i\right) \neq 1$ amounts to removing those g which lie in a particular residue class modulo each prime factor of H_i, for a total of at most $\Gamma = \sum_{i=1}^{n} \omega(H_i)$ residue classes.

We perform the same argument to this restricted sum. The only change that occurs is when we write $g = r + s^p$, we must assume that s avoids a corresponding set of residue classes modulo these primes. By inclusion-exclusion, we can write a Dirichlet character sum avoiding Γ residue classes as an alternating sum of Dirichlet character sums in at most 2^Γ residue classes, and hence as an alternating sum of at most 2^Γ shorter Dirichlet character sums. Because the sums over each residue class are shorter, we can get the same bound for them by Corollary 2.7. Thus our final bound for this case is worse by a factor of

$$2^\Gamma = q^{\sum_{i=1}^{n} o(d(H_i))} = q^{o(d)}. \tag{4.30}$$

We can absorb this into our bound by slightly increasing β so that it still satisfies the strict inequality (4.4). □

Theorem 4.5. *Fix $\epsilon, \delta > 0$, $0 < \beta < 1/2$, and a positive integer n. Let q be a power of an odd prime p such that*

$$q > \left(\frac{pne}{\min\left\{\frac{\epsilon}{\epsilon+2}, \frac{\epsilon\delta}{\epsilon+\delta}\right\}}\right)^{\frac{2}{1-2\beta}}. \tag{4.31}$$

Then for nonnegative integers $d, m_1, \dots, m_n, k_1, \dots, k_n$ with

$$d \geq \max\{\epsilon m_1, \dots, \epsilon m_n, \delta k_1, \dots, \delta k_n\}, \quad k_i \neq d + m_i,\ 1 \leq i \leq n, \tag{4.32}$$

and pairs $(a_i, M_i) \in \mathcal{M}_{k_i} \times \mathcal{M}_{m_i}$ for $1 \leq i \leq n$ with a_i/M_i distinct, we have

$$\sum_{g \in \mathcal{M}_d} \prod_{i=1}^{n} \mu\left(a_i + gM_i\right) \ll |\mathcal{M}_d|^{1-\frac{\beta}{p}} \tag{4.33}$$

as $d \to \infty$, with the implied constant depending only on $\epsilon, \delta, \beta, n$ and q.

Proof. Our initial assumption is that there are no coincidences among (a_i/M_i), for $1 \leq i \leq n$, so we can find a prime P not dividing

$$\prod_{i=1}^{n} M_i \prod_{1 \leq i < j \leq n} (a_i M_j - a_j M_i), \tag{4.34}$$

with $d(P) = o(d)$. Let $t' = d(P)$. Splitting our initial sum according to the residue class z of g mod P we get

$$\sum_{d(z) < t'} \sum_{f \in \mathcal{M}_{d-t'}} \prod_{i=1}^{n} \mu\left(a_i + zM_i + fM_iP\right). \tag{4.35}$$

We can bound the sums over residue classes by applying Proposition 4.3. Indeed, suppose toward a contradiction that for some $1 \leq i < j \leq n$ we have

$$\left(\frac{a_i + zM_i}{M_i P}\right)' = \left(\frac{a_j + zM_j}{M_j P}\right)'. \tag{4.36}$$

The left hand side above equals

$$\frac{a_i' M_i P + z' M_i^2 P + z M_i' M_i P - a_i M_i' P - a_i M_i P' - z M_i M_i' P - z M_i^2 P'}{M_i^2 P^2}$$

so equating it with the right hand side and clearing denominators, we get that

$$M_j^2(a_i' M_i P + z' M_i^2 P + z M_i' M_i P - a_i M_i' P - a_i M_i P' - z M_i M_i' P - z M_i^2 P')$$

equals

$$M_i^2(a_j' M_j P + z' M_j^2 P + z M_j' M_j P - a_j M_j' P - a_j M_j P' - z M_j M_j' P - z M_j^2 P').$$

Reducing mod P and multiplying by -1 we get the congruence

$$M_j^2(a_i M_i P' + z M_i^2 P') \equiv M_i^2(a_j M_j P' + z M_j^2 P') \bmod P$$

from which we deduce that

$$M_j^2 a_i M_i P' \equiv M_i^2 a_j M_j P' \bmod P.$$

Since P does not divide $M_i M_j$ by assumption, we get that

$$M_j a_i P' \equiv M_i a_j P' \bmod P.$$

The primality of P implies that P' is coprime to P so we conclude that P divides the polynomial $a_i M_j - a_j M_i$. This contradicts our choice of P.

Because the length of the sum in this case is $q^{d-t'}$, we obtain a savings in each term of $q^{(d-t')\beta/2p}$ from Proposition 4.3. To obtain our desired savings of $q^{d\beta/2p}$, we must choose $\beta + o(1)$ instead of β in the statement of Proposition 4.3 . Similarly to ensure that $d - t' \geq \max\{\epsilon m_1, \ldots, \epsilon m_n, \delta k_1, \ldots, \delta k_n\}$ we must choose $\epsilon - o(1)$ and $\delta - o(1)$. However because the inequality from Eq. (4.31) is strict, we may increase β by $o(1)$ and reduce ϵ and δ by $o(1)$ in such a way that this inequality is still satisfied. □

We prove two corollaries that give weaker results under conditions that are simpler to state.

Corollary 4.6. *Fix $\epsilon, \delta > 0$ and a positive integer n. Let q be a power of an odd prime p such that*

$$q > p^2 n^2 e^2 \max\left(1 + \frac{2}{\epsilon}, \frac{1}{\epsilon} + \frac{1}{\delta}\right)^2. \tag{4.37}$$

Then for nonnegative integers $d, m_1, \ldots, m_n, k_1, \ldots, k_n$ with

$$d \geq \max\{\epsilon m_1, \ldots, \epsilon m_n, \delta k_1, \ldots, \delta k_n\}, \quad k_i \neq d + m_i, \ 1 \leq i \leq n, \tag{4.38}$$

and distinct coprime pairs $(a_i, M_i) \in \mathcal{M}_{k_i} \times \mathcal{M}_{m_i}$ *for* $1 \leq i \leq n$*, we have*

$$\sum_{g \in \mathcal{M}_d} \prod_{i=1}^{n} \mu\left(a_i + gM_i\right) = o\left(|\mathcal{M}_d|\right) \tag{4.39}$$

as $d \to \infty$ *for fixed* ϵ, δ, n, q*.*

Proof. The assumed lower bound on q is equivalent to

$$q > \left(\frac{pne}{\min\left\{ \frac{\epsilon}{\epsilon+2}, \frac{\epsilon\delta}{\epsilon+\delta} \right\}} \right)^2. \tag{4.40}$$

If q satisfies this inequality, we can take β small enough that the inequality

$$q > \left(\frac{pne}{\min\left\{ \frac{\epsilon}{\epsilon+2}, \frac{\epsilon\delta}{\epsilon+\delta} \right\}} \right)^{\frac{2}{1-2\beta}} \tag{4.41}$$

holds, and then apply Theorem 4.5. □

Corollary 4.7. *Let* q *be a power of an odd prime* p *such that*

$$q > p^2 n^2 e^2, \tag{4.42}$$

and let $(a_i, M_i) \in \mathcal{M}_{k_i} \times \mathcal{M}_{m_i}$ *be distinct coprime pairs for* $1 \leq i \leq n$*. Then*

$$\sum_{g \in \mathcal{M}_d} \prod_{i=1}^{n} \mu\left(a_i + gM_i\right) = o\left(|\mathcal{M}_d|\right) \tag{4.43}$$

as $d \to \infty$ *for fixed* q, n, a_i, M_i*.*

Proof. We can take ϵ, δ large enough that

$$q > p^2 n^2 e^2 \max\left(1 + \frac{2}{\epsilon}, \frac{1}{\epsilon} + \frac{1}{\delta}\right)^2. \tag{4.44}$$

For this ϵ and δ, the inequalities

$$d \geq \max\{\epsilon m_1, \ldots, \epsilon m_n, \delta k_1, \ldots, \delta k_n\}, \quad k_i \neq d + m_i, \ 1 \leq i \leq n, \tag{4.45}$$

will be satisfied for all d sufficiently large. We can then apply Corollary 4.6 to deduce the claim. □

5. Level of distribution

The following will be obtained using the techniques of [FM98] and [FKM14] in the appendix. For coprime polynomials g, M we denote by $\overline{g}$ the residue class of the (multiplicative) inverse of g mod M.

Theorem 5.1. *Fix an odd prime power q. Then for any $\theta > 0$, for nonnegative integers d, m with $d \leq m$, squarefree $M \in \mathcal{M}_m$, and an additive character ψ mod M, we have*

$$\sum_{\substack{g \in \mathcal{M}_d \\ (g,M)=1}} \mu(g)\psi(\overline{g}) \ll q^{\left(\frac{3}{16}+\theta\right)m+\frac{25}{32}d} \tag{5.1}$$

as $d \to \infty$, with the implied constant depending only on q and θ.

The following proposition allows us to identify the main term in sums of von Mangoldt in arithmetic progressions.

Proposition 5.2. *Fix a prime power q. For nonnegative integers d, m, and a squarefree $M \in \mathcal{M}_m$ we have*

$$\sum_{k=1}^{d} kq^{-k} \sum_{\substack{A \in \mathcal{M}_k \\ (A,M)=1}} \mu(A) = -\frac{q^m}{\varphi(M)} + q^{o(m+d)-d}. \tag{5.2}$$

Proof. The left hand side above is the sum of the first d coefficients of the power series

$$\begin{aligned} u\frac{d}{du}\sum_{k=1}^{\infty} q^{-k}u^k \sum_{\substack{A \in \mathcal{M}_k \\ (A,M)=1}} \mu(A) &= u\frac{d}{du}\prod_{P \nmid M}\left(1-u^{d(P)}|P|^{-1}\right) \\ &= u\frac{d}{du}\left((1-u)\prod_{P|M}\left(1-u^{d(P)}|P|^{-1}\right)^{-1}\right). \end{aligned}$$

Summing all the coefficients of a power series, evaluates it at $u = 1$. Hence, the main term comes from the equality

$$\left(u\frac{d}{du}\left((1-u)F(u)\right)\right)(1) = -F(1), \quad F(u) = \prod_{P|M}\left(1-u^{d(P)}|P|^{-1}\right)^{-1}.$$

The coefficients of degree greater than d contribute to the error term. To bound the sum of these coefficients, we can write the degree k coefficient as

$$k \oint_{|u|=r} \frac{(1-u)F(u)}{u^{k+1}}du \tag{5.3}$$

for $r < q$, getting a bound of

$$k(1+r)r^{-k-1}\max_{|u|=r} F(u). \tag{5.4}$$

As long as the above maximum is subexponential in m for all $r < q$, the expression above will be $q^{o(m)}k(1+r)r^{-k-1}$, so the sum of the coefficients of degree greater than d is

$$q^{o(m)}\sum_{k>d} k(1+r)r^{-k-1} = q^{o(m)}(r-o(1))^{-d}. \tag{5.5}$$

Taking r arbitrarily close to q, the above is $q^{o(m)}(q-o(1))^{-d} = q^{o(m+d)-d}$.

The value of $F(u)$ is indeed subexponential in m, because M has $o(m)$ prime divisors and each contributes at most

$$\left(1-\left(\frac{r}{q}\right)^{d(P)}\right)^{-1} \leq \left(1-\frac{r}{q}\right)^{-1}. \tag{5.6}$$

□

The proof above is written using asymptotic notation for m (and d), namely it gives uniformity in M. Clearly the same argument also applies to the case of a fixed polynomial M (or a fixed value of m), and $d \to \infty$.

In the following we deduce, from our results on the Möbius function, a level of distribution beyond $1/2$ for primes in every individual arithmetic progression to a squarefree modulus. We shall use the convolution identity $\Lambda = -1 * (\mu \cdot \deg)$ which for $f \in \mathbb{F}_q[T]^+$ of degree $d \geq 0$ says that

$$\Lambda(f) = -\sum_{k=1}^{d} k \sum_{A\in\mathcal{M}_k} \sum_{\substack{B\in\mathcal{M}_{d-k}\\ AB=f}} \mu(A). \tag{5.7}$$

Corollary 5.3. *For any $0<\omega<1/32$, for any odd prime p and power q of p such that $q > p^2e^2\left(1+\frac{50}{1-32\omega}\right)^2$, the following holds. For nonnegative integers d,m with $d \geq (1-\omega)m$, squarefree $M \in \mathcal{M}_m$, and $a \in \mathbb{F}_q[T]$ with $d(a) < d+m$ and coprime to M, we have*

$$\sum_{g\in\mathcal{M}_d} \Lambda(a+gM) = \frac{q^{d+m}}{\varphi(M)} + O\left(q^{(1-\delta)d}\right) \tag{5.8}$$

as $d \to \infty$, for some $\delta > 0$ depending only on q,ω (a power savings error term).

Proof. We can assume $d(a) = m$, and use Eq. (5.7) to write our sum as

$$\sum_{\substack{f\in\mathcal{M}_{m+d}\\ f\equiv a \bmod M}} \Lambda(f) = -\sum_{k=1}^{d+m} k \sum_{\substack{A\in\mathcal{M}_k\\ (A,M)=1}} \mu(A) \sum_{\substack{B\in\mathcal{M}_{m+d-k}\\ AB\equiv a \bmod M}} 1 \tag{5.9}$$

so by Proposition 5.2 the range $k \leq d$ contributes the main term.

The (absolute value of the) contribution of any $k > d$ is

$$\begin{aligned} &kq^{d-k}\Big|\sum_{\substack{A\in\mathcal{M}_k\\ (A,M)=1}} \mu(A) \sum_{\substack{\psi\colon \mathbb{F}_q[T]/M\to\mathbb{C}^\times\\ \psi(f)=0 \text{ if } d(f)<m+d-k}} \psi\left(a\overline{A}\right)\overline{\psi(T^{m+d-k})}\Big| \leq \\ &kq^{d-k} \sum_{\substack{\psi\colon \mathbb{F}_q[T]/M\to\mathbb{C}^\times\\ \psi(f)=0 \text{ if } d(f)<m+d-k}} \Big|\sum_{\substack{A\in\mathcal{M}_k\\ (A,M)=1}} \mu(A)\psi\left(a\overline{A}\right)\Big| \end{aligned} \tag{5.10}$$

so by Theorem 5.1 we get

$$\ll kq^{d-k}q^{k-d}q^{\left(\frac{3}{16}+\theta\right)m+\frac{25}{32}k} = kq^{\left(\frac{3}{16}+\theta\right)m+\frac{25}{32}k} \tag{5.11}$$

which gives a power saving as long as

$$\left(\frac{3}{16}+2\theta\right)m+\frac{25}{32}k < d. \tag{5.12}$$

The (absolute value of the) contribution of any other k is at most

$$k\sum_{\substack{B\in\mathcal{M}_{m+d-k}\\(B,M)=1}}\left|\sum_{\substack{A\in\mathcal{M}_k\\AB\equiv a \bmod M}}\mu(A)\right| = k\sum_{\substack{B\in\mathcal{M}_{m+d-k}\\(B,M)=1}}\left|\sum_{g\in\mathcal{M}_{k-m}}\mu(b+gM)\right| \tag{5.13}$$

where b is the unique monic polynomial of degree m congruent to aB^{-1} modulo m. We apply Theorem 4.5 with some fixed $\beta > 0$ and with

$$\epsilon = \delta = \frac{32}{25}\left(\frac{1}{32}-\omega-2\theta\right). \tag{5.14}$$

Then because (5.12) does not hold, we have

$$k-m \geq \frac{32}{25}\left(d-\left(\frac{31}{32}+2\theta\right)m\right) \geq \frac{32}{25}\left((1-\omega)m-\left(\frac{31}{32}+2\theta\right)m\right) = \epsilon m$$

so the conditions of Theorem 4.5 are satisfied as long as

$$q > \left(pe\left(1+\frac{25}{16}\frac{1}{\frac{1}{32}-\omega-2\theta}\right)\right)^{\frac{2}{1-2\beta}}. \tag{5.15}$$

Because we may take β and θ arbitrarily small, it suffices to have

$$q > p^2e^2\left(1+\frac{25}{16}\frac{1}{\frac{1}{32}-\omega}\right)^2 = p^2e^2\left(1+\frac{50}{1-32\omega}\right)^2 \tag{5.16}$$

Summation over k gives only an extra logarithmic factor, so it preserves our power savings. □

6. The twin primes conjecture

6.1. **Chowla sums over primes.** We establish cancellation in Möbius autocorrelation over primes.

Corollary 6.1. *Fix $\tilde{\epsilon},\tilde{\delta} > 0$, $0 < \alpha < 1$, $0 < \beta < 1/2$, and a positive integer n. Let q be a power of an odd prime p such that*

$$q > \left(p(n+1)e\max\left(1+\frac{2+2\alpha+4\tilde{\epsilon}^{-1}}{1-\alpha}, \frac{1+\alpha+2\tilde{\epsilon}^{-1}+2\tilde{\delta}^{-1}}{1-\alpha}\right)\right)^{\frac{2}{1-2\beta}}.$$

Take nonnegative integers $d, m_1, \dots, m_n, k_1, \dots, k_n$ with

$$d \geq \max\{\tilde{\epsilon}m_1, \dots, \tilde{\epsilon}m_n, \tilde{\delta}k_1, \dots, \tilde{\delta}k_n\}, \quad k_i \neq d+m_i,\ 1 \leq i \leq n, \tag{6.1}$$

and distinct coprime pairs $(a_i, M_i) \in \mathcal{M}_{k_i} \times \mathcal{M}_{m_i}$ for $1 \leq i \leq n$.

Furthermore let m, k be nonnegative integers with $m \leq \alpha d$, $k < m + d$, let $M \in \mathcal{M}_m$, and let $a \in \mathcal{M}_k$ with (a, M) distinct from (a_i, M_i) for every $1 \leq i \leq n$. Then

$$\sum_{g \in \mathcal{M}_d} \Lambda(a + gM) \prod_{i=1}^{n} \mu\left(a_i + gM_i\right) \ll |\mathcal{M}_d|^{1 - \frac{\beta(1-\alpha)}{2p}} \tag{6.2}$$

as $d \to \infty$, with the implied constant depending only on $\tilde{\epsilon}, \tilde{\delta}, \alpha, n, \beta$ and q.

Proof. We can assume $(a, M) = 1$, set $z = \left\lfloor \frac{d+m}{2} \right\rfloor$, and use Eq. (5.7) to write

$$\Lambda(a+gM) = -\sum_{b=1}^{z} b \sum_{\substack{B \in \mathcal{M}_b \\ B \mid a+gM}} \mu(B) - \sum_{c=0}^{d+m-z-1} (d+m-c) \sum_{\substack{C \in \mathcal{M}_c \\ C \mid a+gM}} \mu\left(\frac{a+gM}{C}\right).$$

For every B above, taking a monic $N = N(B) \in \mathbb{F}_q[T]$ with

$$MN \equiv -a \bmod B, \quad d(N) \neq k_i - m_i, \quad d(N) \leq b + n, \tag{6.3}$$

and writing $g = N + hB$ with $h \in \mathcal{M}_{d-b}$, we see that the sum over b contributes

$$\sum_{b=1}^{\left\lfloor \frac{d+m}{2} \right\rfloor} b \sum_{B \in \mathcal{M}_b} \mu(B) \sum_{h \in \mathcal{M}_{d-b}} \prod_{i=1}^{n} \mu\left(a_i + NM_i + hBM_i\right) \tag{6.4}$$

so we can apply Theorem 4.5 to the innermost sum, taking

$$\epsilon = \frac{1-\alpha}{1+\alpha+2\tilde{\epsilon}^{-1}} \leq \frac{d-m}{d+m+2m_i} = \frac{d - \frac{d+m}{2}}{\frac{d+m}{2} + m_i} \leq \frac{d-b}{b+m_i} \tag{6.5}$$

and

$$\begin{aligned} \delta = \min\left(\frac{1-\alpha}{2\tilde{\delta}^{-1}}, \frac{1-\alpha}{1+\alpha+2\tilde{\epsilon}^{-1}}\right) &\leq \min\left(\frac{d-b}{k_i}, \frac{d-b}{b+m_i}\right) \\ &\leq \min\left(\frac{d-b}{\deg a_i}, \frac{d-b}{\deg NM_i} + o(1)\right). \end{aligned} \tag{6.6}$$

To obtain the condition on q, note that, when calculating $\max(1+\frac{2}{\epsilon}, \frac{1}{\epsilon}+\frac{1}{\delta})$, we can treat δ as $\frac{1-\alpha}{2\tilde{\delta}^{-1}}$, because if the other term is smaller, then $\frac{1}{\epsilon} + \frac{1}{\delta}$ is dominated by $1 + \frac{2}{\epsilon}$ anyways.

Taking $L \in \mathbb{F}_q[T]^+$ with

$$ML \equiv -a \bmod C, \quad d(L) \neq k_i - m_i, \quad d(L) \neq d, \quad d(L) \leq c + n + 1, \tag{6.7}$$

and writing $g = L + hC$ with $h \in \mathcal{M}_{d-c}$, we see that the sum over c contributes

$$\sum_{c=0}^{d+m-z-1} (d+m-c) \sum_{C \in \mathcal{M}_c} \sum_{h \in \mathcal{M}_{d-c}} \mu\left(\frac{a+LM}{C} + hM\right) \prod_{i=1}^{n} \mu\left(a_i + LM_i + hCM_i\right)$$

and we again apply Theorem 4.5 to the innermost sum, with the same values of ϵ and δ. Everything is the same as before, with c replaced by b, except for two things.

(1) We can no longer use the inequality $b \leq \frac{m+d}{2}$, but rather the slightly weaker inequality $c \leq \frac{m+d+1}{2}$.
(2) The term $\mu\left(\frac{a+LM}{C} + hM\right)$ appears, which means we must check that

$$(d-c) \geq \epsilon m, (d-c) \geq \epsilon d \left(\frac{a+LM}{C}\right). \tag{6.8}$$

However (1) is no difficulty as we may assume d sufficiently large and perturb the parameters ϵ and δ slightly to insure the inequalities of Theorem 4.5 still hold. For this reason we will assume $c \leq \frac{m+d}{2}$ while handling (2) as well.

To check that $(d-c) \geq \epsilon m$ we observe that, because $c \leq \frac{m+d}{2}$ and $d < \alpha m$, we have $d - c \geq \frac{d-m}{2}$, and thus $\frac{d-c}{m} \geq \frac{1-\alpha}{2\alpha} \geq \frac{1-\alpha}{1+\alpha} \geq \epsilon$.

To check that

$$d\left(\frac{a+LM}{C}\right) \leq \max(m+d-c, m+n) \leq \delta^{-1}(d-c) \tag{6.9}$$

we observe

$$\delta \leq \frac{1-\alpha}{1+\alpha} = \frac{2(d-m)}{d+m} = \frac{1}{\frac{m}{(d-m)/2}+1} \leq \frac{1}{\frac{m}{d-c}+1} = \frac{d-c}{m+d-c}.$$

□

6.2. **Singular series.** For nonzero $a \in \mathbb{F}_q[T]$ define

$$\mathfrak{S}_q(a) = \prod_{P|a} \left(1 - |P|^{-1}\right)^{-1} \prod_{P\nmid a} \left(1 - (|P|-1)^{-2}\right). \tag{6.10}$$

The following propositon allows us to identify the main term in our twin prime number theorem. An analogous result over the integers is proved in [GY03, Lemma 2.1].

Proposition 6.2. *Fix a prime power q. Then for an integer $n \geq 1$ and a nonzero $a \in \mathbb{F}_q[T]$ we have*

$$\sum_{k=1}^{n} k \sum_{\substack{M \in \mathcal{M}_k \\ (M,a)=1}} \frac{\mu(M)}{\varphi(M)} = -\mathfrak{S}_q(a) + (q-1)^{o(n+d(a))-n}. \tag{6.11}$$

Proof. We are interested in the sum of the first n coefficients of

$$\begin{aligned} Z(u) = u\frac{d}{du}\sum_{k=1}^{\infty} u^k \sum_{\substack{M \in \mathcal{M}_k \\ (M,a)=1}} \frac{\mu(M)}{\varphi(M)} &= u\frac{d}{du}\prod_{P\nmid a}\left(1 - u^{d(P)}\left(|P|-1\right)^{-1}\right) \\ &= u\frac{d}{du}\left((1-u)G(u)\right) \end{aligned} \tag{6.12}$$

where $G(u)$ is

$$\prod_{P|a}\left(1-u^{d(P)}|P|^{-1}\right)^{-1}\prod_{P\nmid a}\left(1-u^{d(P)}|P|^{-1}\right)^{-1}\left(1-u^{d(P)}\left(|P|-1\right)^{-1}\right).$$

The sum of all coefficients equals

$$Z(1)=-G(1)=-\mathfrak{S}_q(a) \tag{6.13}$$

because

$$\left(1-|P|^{-1}\right)^{-1}\left(1-\left(|P|-1\right)^{-1}\right)=1-\left(|P|-1\right)^{-2}. \tag{6.14}$$

As in Proposition 5.2, to prove the bound for the error term it suffices to prove that $G(u)$ is bounded subexponentially in a for u on each circle of radius $<q-1$.

Note that

$$\begin{gathered}\left(1-u^{d(P)}|P|^{-1}\right)^{-1}\left(1-u^{d(P)}\left(|P|-1\right)^{-1}\right)=\\ 1-\frac{u^{d(P)}}{|P|\left(|P|-1\right)\left(1-u^{d(P)}|P|^{-1}\right)}\end{gathered} \tag{6.15}$$

so $G(u)$ can be rewritten as

$$\prod_{P}\left(1-\frac{u^{d(P)}}{|P|\left(|P|-1\right)\left(1-u^{d(P)}|P|^{-1}\right)}\right)\prod_{P|a}\left(1-u^{d(P)}\left(|P|-1\right)^{-1}\right)^{-1}.$$

The first product above is independent of a and converges on the disc where $|u|<q$. On a circle of radius r, the value of each term in the second product is at most

$$\left(1-\frac{r^{d(P)}}{q^{d(P)}-1}\right)^{-1}\leq\left(1-\frac{r}{q-1}\right)^{-1} \tag{6.16}$$

and thus is bounded, and the number of terms is $o(d(a))$, so the product is subexponential in a.

□

6.3. **Hardy-Littlewood conjecture for pairs.**

Theorem 6.3. *For every odd prime number p, and power q of p with*

$$q>685090p^2, \tag{6.17}$$

there exists $\lambda>0$ such that the following holds. For nonnegative integers $d>\ell$, and $a\in\mathcal{M}_\ell$ we have

$$\sum_{f\in\mathcal{M}_d}\Lambda(f)\Lambda(f+a)=\mathfrak{S}_q(a)q^d+O\left(q^{(1-\lambda)d}\right) \tag{6.18}$$

as $d\to\infty$, with the implied constant depending only on q.

Proof. Using Eq. (5.7) our sum becomes

$$-\sum_{k=0}^{d} k \sum_{M \in \mathcal{M}_k} \mu(M) \sum_{N \in \mathcal{M}_{d-k}} \Lambda(a+NM). \tag{6.19}$$

The appearance of the factor $\mu(M)$ allows us to consider only squarefree M, so by Corollary 5.3, the contribution of the range $0 \leq k \leq d/(2-\omega)$ is

$$-q^d \sum_{k=0}^{d/(2-\omega)} k \sum_{\substack{M \in \mathcal{M}_k \\ (a,M)=1}} \frac{\mu(M)}{\varphi(M)} + O\left(dq^{\frac{d}{2-\omega}} E_q\right) \tag{6.20}$$

which equals by Proposition 6.2 to

$$q^d \mathfrak{S}_q(a) + O\left(dq^{\frac{d}{2-\omega}} E_q + q^d E'_q\left(\frac{d}{2-\omega}\right)\right). \tag{6.21}$$

The error term is of power savings size.

In the other range, we need to prove a power savings bound for

$$\sum_{k > d/(2-\omega)}^{d} k \sum_{N \in \mathcal{M}_{d-k}} \sum_{M \in \mathcal{M}_k} \Lambda(a+MN)\mu(M) \tag{6.22}$$

and this is done by applying Corollary 6.1 to the innermost sum with

$$n = 1, \quad \tilde{\epsilon} = \infty, \quad \tilde{\delta} = \infty, \quad \alpha = 1 - \omega > \frac{d-k}{k} \tag{6.23}$$

and $\beta > 0$ but very small. This requires

$$q > \left(2pe\left(1 + \frac{2+2-2\omega}{\omega}\right)\right)^2 = \left(2pe\left(\frac{4}{\omega} - 1\right)\right)^2 \tag{6.24}$$

and Corollary 5.3 requires

$$q > p^2 e^2 \left(1 + \frac{50}{1-32\omega}\right)^2 \tag{6.25}$$

so the optimal bound is obtained by solving

$$2\left(\frac{4}{\omega} - 1\right) = 1 + \frac{50}{1-32\omega} \tag{6.26}$$

whose solution is

$$\omega = \frac{103 - \sqrt{\frac{30803}{3}}}{64} = .0261\ldots \tag{6.27}$$

which satisfies

$$\left(2e\left(\frac{4}{\omega} - 1\right)\right)^2 < 685090. \tag{6.28}$$

□

7. Results for Small q

We prove Theorem 1.11.

Proof. Set $d(f) = d$. We need to show that by suitably changing the coefficients of f in degree at most ηd, one can arrive at a polynomial with a given (nonzero) Möbius value.

Let $c < \eta d$ be the largest even integer not divisible by 3. Note that

$$c \geq \eta d - 4. \tag{7.1}$$

We take the coefficient of T^c in f to be 1, and the coefficient of T^k to be 0 for every $k < c$ that is not divisible by 3. Hence, it is enough to show that

$$1, -1 \in \left\{\mu(f + b^3) : b \in \mathcal{M}_{\lfloor c/3 \rfloor}\right\}. \tag{7.2}$$

By Lemma 3.2 with $M = 1$, $a = f$, and $g = b^3$, our set equals

$$\left\{S \cdot \chi(w + b^3) : b \in \mathcal{M}_{\lfloor c/3 \rfloor}\right\} \tag{7.3}$$

and since the highest power of T that divides $(f + b^3)' = f'$ is T^{c-1}, we conclude that $S = \pm 1$ and from Remark 3.3 that χ is a nonprincipal character (to a squarefree modulus E). Arguing as in Eq. (4.9) to 'extract third roots', we are thus led to consider

$$\left\{\chi(\widetilde{w} + b) : b \in \mathcal{M}_{\lfloor c/3 \rfloor}\right\} \tag{7.4}$$

where $\widetilde{w}$ is the unique residue class modulo E with $\widetilde{w}^3 \equiv w \mod E$.

From Lemma 3.2 we further conclude that

$$d(E) \leq d - c + 1 \leq (1 - \eta)d + 5, \tag{7.5}$$

and on the other hand

$$\lfloor c/3 \rfloor \geq \frac{c}{3} - 1 \geq \frac{\eta d - 4}{3} - 1 \geq \frac{\eta}{3} d - 3, \tag{7.6}$$

so combining the two we get

$$\frac{\lfloor c/3 \rfloor}{d(E)} \geq \frac{\frac{\eta}{3} d - 3}{(1 - \eta)d + 5}. \tag{7.7}$$

Since we have assumed that $3/7 < \eta < 1$, the right hand side of the above tends to a quantity greater than $1/4$ as $d \to \infty$. Consequently, we can use the (function field version of the) Burgess bound (as stated for instance in [Bur63, Theorem 2]) to show that 1 and -1 belong to the set above. Such a version is obtained in [Hsu99]. □

Now we prove Theorem 1.12.

Proof. Set $k = d(P)$. For positive integers d, n, we seek cancellation in

$$\sum_{g \in \mathcal{M}_d} \mu(a + gP^n) \tag{7.8}$$

where $a \in \mathbb{F}_q[T]^+$ satisfies $d(a) < nk$ and $a \equiv 1 \bmod P^{n-2}$. We assume first that $3 \mid n$, and follow the proof of Proposition 4.3 up to Eq. (4.21) getting

$$\sum_{d(h)<t} \chi_r(f+h) \tag{7.9}$$

with χ_r a character mod E_r. If χ_r is principal, then by Remark 3.3 we have

$$P^{2n}r' + a'P^n = D_r^{(1)} = AB^2, \quad A, B \in \mathbb{F}_q[T],\ A \mid P^n,\ P \nmid B \tag{7.10}$$

and since $P^{n-3} \mid a'$, we conclude that

$$P^3 r' + \frac{a'}{P^{n-3}} = \widetilde{A}\widetilde{B}^2, \quad \widetilde{A}, \widetilde{B} \in \mathbb{F}_q[T],\ \widetilde{A} \mid P. \tag{7.11}$$

There are $\ll q^{d/2}$ choices of $r \in \mathcal{R}$ satisfying the above, so those can be neglected.

For $r \in \mathcal{R}$ with χ_r nonprincipal, we note that

$$d(E_r) \le d\left(\operatorname{rad}\left(P^{2n}r' + a'P^n\right)\right) \le d\left(P^3 r' + \frac{a'}{P^{n-3}}\right) + k \tag{7.12}$$

so for large enough d we have $t/d(E_r) > 1/4$, hence cancellation in Eq. (7.9) is guaranteed by Burgess.

Suppose now that $3 \mid n + \beta$ for some $\beta \in \{1, 2\}$, and write

$$\sum_{g \in \mathcal{M}_d} \mu(1 + gP^n) = \sum_{d(g_0)<\beta k} \sum_{g_1 \in \mathcal{M}_{d-\beta k}} \mu(1 + g_0 P^n + g_1 P^{n+\beta}) \tag{7.13}$$

for $d \ge \beta$. We have thus reduced to the previous case with $a = 1 + g_0 P^n$. □

Remark 7.1. We see from the proof that 1 is not the only residue class for which the argument works. Also, the modulus does not have to be a power of a fixed prime, but it has to be 'multiplicatively close' to a cube.

Appendix A. Orthogonality of the Möbius function and inverse additive characters

We explain how a variant of the results of [FM98] carries over to function fields and gives Theorem 5.1.

A standard strategy in the treatment of sums such as those from Eq. (5.1) is to use a combinatorial identity for the Möbius function. Following [FM98], we use Vaughan's identity, which for $f \in \mathbb{F}_q[T]^+$ gives

$$\mu(f) = -\sum_{\substack{d(g)\le\alpha \\ gh \mid f}} \sum_{d(h)\le\beta} \mu(g)\mu(h) + \sum_{\substack{d(g)>\alpha \\ gh \mid f}} \sum_{d(h)>\beta} \mu(g)\mu(h) \tag{A.1}$$

where summation is over monic polynomials, and α, β are nonnegative integers with $\max\{\alpha, \beta\} < d(f)$. For a proof (that is also valid for function fields) see [IK04, Proposition 13.5].

By applying Vaughan's identity as in [FM98, Section 6], we reduce our task to bounding sums of type I:

$$\Sigma_{k,r}^{(\mathrm{I})} = \sum_{\substack{f\in\mathcal{M}_k\\(fg,M)=1}}\sum_{g\in\mathcal{M}_r}\gamma_f\psi(\overline{fg}) \tag{A.2}$$

where $k \le \frac{1}{8}d(M)+\frac{7}{16}d,\ k+r\le d,\ |\gamma_f|\le 1$ and sums of type II:

$$\Sigma_{k,r}^{(\mathrm{II})} = \sum_{\substack{f\in\mathcal{M}_k\\(fg,M)=1}}\sum_{g\in\mathcal{M}_r}\gamma_f\delta_g\psi(\overline{fg}) \tag{A.3}$$

with $k \ge \frac{1}{8}d(M)+\frac{7}{16}d,\ r\ge\frac{7}{16}d-\frac{3}{8}d(M),\ k+r\le d,\ |\delta_g|\le 1$. For every $\epsilon>0$, we need the bounds

$$\Sigma_{k,r}^{(\mathrm{I})} \ll q^{\left(\frac{3}{16}+\epsilon\right)d(M)+\frac{25}{32}d},\quad \Sigma_{k,r}^{(\mathrm{II})} \ll q^{d+\epsilon d(M)-\frac{r}{2}}+q^{d+\left(\frac{1}{4}+\epsilon\right)d(M)-\frac{k}{2}} \tag{A.4}$$

that are analogous to [FM98, Equation 6.4]. The bounds (A.4) then imply (5.1) by the argument of [FM98, Section 6].

A.1. **Sums of type I.** Following first the bilinear shifting trick argument of [FM98, §4], we obtain the inequality

$$\Sigma_{k,r}^{(\mathrm{I})} \ll \frac{1}{V}\sum_{\substack{a\in\mathcal{M}_{d_A}\\(a,M)=1}}\sum_{f\in\mathcal{M}_k}\sum_{g\in\mathcal{M}_r}\Big|\sum_{\substack{b\in\mathcal{M}_{d_B}\\(a^{-1}g+b,M)=1}}\psi\left(\overline{af(a^{-1}g+b)}\right)\Big| \tag{A.5}$$

where, in analogy with the variables A, B from [FM98, (4.8)],

$$d_A=\frac{3r-k}{4},\quad d_B=\frac{k+r}{4}-1 \tag{A.6}$$

and

$$V=\#\{ab: a\in\mathcal{M}_{d_A}, b\in\mathcal{M}_{d_B}, (a,M)=1\}\gg q^{d_A+d_B-\epsilon d(M)}. \tag{A.7}$$

The parameter t and the term $e(-bt)$ from [FM98] do not appear here, as they arise from the failure of an archimedean interval to be perfectly invariant under a shift.

Put $\mathbb{M}=\mathbb{F}_q[T]/(M)$. Following the Hölder's inequality argument from [FM98, §4.a], we get as in [FM98, Equation 4.6], the bound

$$\begin{aligned}&\Big(\sum_{\substack{a\in\mathcal{M}_{d_A}\\(a,M)=1}}\sum_{f\in\mathcal{M}_k}\sum_{g\in\mathcal{M}_r}\Big|\sum_{\substack{b\in\mathcal{M}_{d_B}\\a^{-1}g+b\in\mathbb{M}^\times}}\psi(\overline{af(a^{-1}g+b)})\Big|\Big)^6\ll\\ &q^{5(k+r+d_A)+\epsilon d(m)}\sum_{b_1,\dots,b_3'\in\mathcal{M}_{d_B}}\Big|\sum_{\substack{h\in\mathbb{M}\\h+b_i\in\mathbb{M}^\times\\h+b_i'\in\mathbb{M}^\times}}\sum_{\substack{d(s)\le k+d_A\\(s,M)=1}}\psi\left(\Delta_{\mathbf{b}}(h,s)\right)\Big|\end{aligned} \tag{A.8}$$

where

$$\mathbf{b}=(b_1,b_2,b_3,b_1',b_2',b_3')\in\mathbb{F}_q[T]^6, \tag{A.9}$$

and the function $\Delta_{\mathbf{b}}(h,s)$ is defined by

$$\Delta_{\mathbf{b}}(h,s) = \sum_{i=1}^{3} \left[\frac{1}{(h+b_i)s} - \frac{1}{(h+b_i')s} \right]. \tag{A.10}$$

Because of our conditions $h+b_i \in \mathbb{M}^\times, h+b_i' \in \mathbb{M}^\times$, and $(s,M)=1$, the function $\Delta_{\mathbf{b}}$ is never evaluated where it is undefined. The Hölder's inequality argument is slightly more involved because the invertibility assumptions are more complex in case M is not prime, but the idea is essentially the same.

As in [FM98], we complete the right hand side of Eq. (A.8) and get

$$\begin{aligned} &\sum_{\substack{h \in \mathbb{M} \\ h+b_i \in \mathbb{M}^\times \\ h+b_i' \in \mathbb{M}^\times}} \sum_{\substack{d(s) \le k+d_A \\ (s,M)=1}} \psi\left(\Delta_{\mathbf{b}}(h,s)\right) = \\ &q^{k+d_A-d(M)} \sum_{\substack{z \in \mathbb{F}_q[T] \\ d(z) < d(M)-k-d_A}} \sum_{\substack{h \in \mathbb{M} \\ h+b_i \in \mathbb{M}^\times \\ h+b_i' \in \mathbb{M}^\times}} \sum_{s \in \mathbb{M}^\times} \psi\left(\Delta_{\mathbf{b}}(h,s) + zs\right). \end{aligned} \tag{A.11}$$

In order to treat the innermost sum on the right hand side, we note that

$$\Delta_{\mathbf{b}}(h,s) + zs = \frac{1}{s} \sum_{i=1}^{3} \left[\frac{1}{(h+b_i)} - \frac{1}{(h+b_{i'})} \right] + zs \tag{A.12}$$

so the aforementioned innermost sum over s is a Kloosterman sum. We put

$$\mathbb{M}_{\mathbf{b}} = \left\{ x \in \mathbb{M} : (x+b_1)\cdots(x+b_3') \in \mathbb{M}^\times \right\}, \tag{A.13}$$

define a function $R_{\mathbf{b}} \colon \mathbb{M}_{\mathbf{b}} \to \mathbb{M}$ by

$$R_{\mathbf{b}}(x) = \sum_{i=1}^{3} \left(\frac{1}{x+b_i} - \frac{1}{x+b_i'} \right), \tag{A.14}$$

and a Kloosterman sum

$$S(x,z) = \sum_{y \in \mathbb{M}^\times} \psi\left(xy^{-1} + zy\right), \quad x \in \mathbb{M},\ z \in \mathbb{F}_q[T]. \tag{A.15}$$

In this notation, for any $z \in \mathbb{F}_q[T]$ we have

$$\sum_{h \in \mathbb{M}_{\mathbf{b}}} \sum_{s \in \mathbb{M}^\times} \psi\left(\Delta_{\mathbf{b}}(h,s) + zs\right) = \sum_{x \in \mathbb{M}_{\mathbf{b}}} S(R_{\mathbf{b}}(x), z) \tag{A.16}$$

and the following claim.

Proposition A.1. *For $\sigma \in S_3$ and*

$$M_\sigma^{\mathbf{b}} = \gcd\left(M, b_1 - b_{\sigma(1)}', b_2 - b_{\sigma(2)}', b_3 - b_{\sigma(3)}'\right) \tag{A.17}$$

we have if $p > 3$ the bound

$$\left| \sum_{x \in \mathbb{M}_{\mathbf{b}}} S(R_{\mathbf{b}}(x), z) \right| \ll |M| d_2(M)^4 \left| \operatorname{lcm}_{\sigma \in S_3} \gcd(M_\sigma^{\mathbf{b}}, z) \right| \tag{A.18}$$

with the implied constant depending only on q. As before, we write $d_2(M)$ for the number of monic divisors of M.

If $p = 3$, with

$$M_{\triangle}^{\mathbf{b}} = \gcd\left(M, b_1 - b_2, b_2 - b_3, b_1' - b_2', b_2' - b_3'\right), \tag{A.19}$$

we have the bound

$$\left|\sum_{x \in \mathbb{M}_{\mathbf{b}}} S(R_{\mathbf{b}}(x), z)\right| \ll |M| d_2(M)^4 \left|\mathrm{lcm}_{\sigma \in S_3 \cup \{\triangle\}} \gcd(M_\sigma^{\mathbf{b}}, z)\right| \tag{A.20}$$

with the implied constant depending only on q.

Proof. Since both the bound and the sum are multiplicative in M, it suffices to handle the case when M is prime, where we show that

$$\left|\sum_{x \in \mathbb{M} \setminus \{b_1, \dots, b_3'\}} S(R_{\mathbf{b}}(x), m)\right| \leq 16|\mathbb{M}| \tag{A.21}$$

unless $z = 0$ and either

- for some $\sigma \in S_3$ we have $b_i = b'_{\sigma(i)}$ for all $1 \leq i \leq 3$;
- or $p = 3$, $b_1 = b_2 = b_3$, and $b_1' = b_2' = b_3'$;

in which case we have the trivial bound

$$\left|\sum_{x \in \mathbb{M} \setminus \{b_1, \dots, b_3'\}} S(R_{\mathbf{b}}(x), 0)\right| \leq |\mathbb{M}|^2. \tag{A.22}$$

The relevance of these conditions is that the residue of the pole of $R_{\mathbf{b}}$ at a point x equals

$$\#\{1 \leq i \leq 3 : b_i = -x\} - \#\{1 \leq i \leq 3 : b_i' = -x\} \tag{A.23}$$

so it is nonzero whenever these two numbers are not equal, except when $p = 3$, one of these numbers is 3, and the other is zero. Hence, $R_{\mathbf{b}}$ has a pole unless each b_i is equal to some $b_{\sigma(i)'}$, or $p = 3$, all the b_i are equal, and the b_i' are also all equal.

Excluding these 'trivial' values of $\mathbf{b}$ for $z = 0$, we get that the rational function $R_{\mathbf{b}}$ is nonconstant and at most 6 to 1. Hence, if $R_{\mathbf{b}}(x) \neq 0$ we get $S(R_{\mathbf{b}}(x), 0) = -1$, while for the values of x with $R_{\mathbf{b}}(x) = 0$, at most 6 in number, we have

$$S(0, 0) = |\mathbb{M}| - 1. \tag{A.24}$$

In total, we get

$$\left|\sum_{x \in \mathbb{M} \setminus \{b_1, \dots, b_3'\}} S(R_{\mathbf{b}}(x), 0)\right| \leq |\mathbb{M}| + 6|\mathbb{M}| = 7|\mathbb{M}|. \tag{A.25}$$

Suppose from now on that $z \neq 0$. Note that if the rational function $R_{\mathbf{b}}$ is constant then it necessarily vanishes identically, so we get

$$\left| \sum_{x \in \mathbb{M} \setminus \{b_1, \dots, b_3'\}} S(0,z) \right| \leq |\mathbb{M}|. \tag{A.26}$$

We can thus assume that $R_{\mathbf{b}}$ is nonconstant, and let $\mathcal{K}\ell_2$ be the Kloosterman sheaf, as in [Ka88]. With this notation, our sum can be written as

$$\begin{aligned} &- \sum_{x \in \mathbb{M} \setminus \{b_1, \dots, b_3'\}} \operatorname{tr}\left(\mathrm{Frob}_{|\mathbb{M}|}, (\mathcal{K}\ell_2)_{R_{\mathbf{b}}(x)z}\right) = \\ &- \sum_{x \in \mathbb{M} \setminus \{b_1, \dots, b_3'\}} \operatorname{tr}\left(\mathrm{Frob}_{|\mathbb{M}|}, ([zR_{\mathbf{b}}]^* \mathcal{K}\ell_2)_x\right) = \\ &- \sum_{i=0}^{2} (-1)^i \operatorname{tr}\left(\mathrm{Frob}_{|\mathbb{M}|}, H_c^i\left(\mathbb{A}^1_{\overline{\mathbb{F}}_q} \setminus \{-b_1, \dots, -b_3'\}, [zR_{\mathbf{b}}]^* \mathcal{K}\ell_2\right)\right) \end{aligned} \tag{A.27}$$

by the Grothendieck-Lefschetz fixed point formula. By [Ka88, Theorem 11.1] the geometric monodromy group of $\mathcal{K}\ell_2$ is SL_2, which is connected, so the geometric monodromy group of the pullback of $\mathcal{K}\ell_2$ by any finite covering map is SL_2, whose standard representation has no nontrivial monodromy coinvariants, so the cohomology groups in degree 0 and 2 vanish. As $\mathcal{K}\ell_2$ is pure of weight 1, its pullback by a finite covering map is mixed of weight at most 1, so by Deligne's Riemann Hypothesis, the eigenvalues of $\mathrm{Frob}_{|\mathbb{M}|}$ on $H_c^1(\mathbb{A}^1_{\overline{\mathbb{F}}_q} \setminus \{-b_1, \dots, -b_3'\}, [zR_{\mathbf{b}}]^* \mathcal{K}\ell_2)$ have absolute value at most $|\mathbb{M}|$. Hence, in order to bound our sum by $16|\mathbb{M}|$, it suffices to prove that the dimension of the above cohomology group is at most 16.

By the aforementioned vanishing of cohomology in degrees 0 and 2, the dimension of our cohomology group equals minus the Euler characteristic. Since $[zR_{\mathbf{b}}]^* \mathcal{K}\ell_2$ is lisse of rank 2 on $\mathbb{A}^1_{\overline{\mathbb{F}}_q} \setminus \{-b_1, \dots, -b_3'\}$, its Euler characteristic is twice the Euler characteristic of $\mathbb{A}^1_{\overline{\mathbb{F}}_q} \setminus \{-b_1, \dots, -b_3'\}$, which is

$$2\left(1 - \#\{-b_1, \dots, -b_3'\}\right), \tag{A.28}$$

minus the sum of the Swan conductors at each singular point, in view of [SGA5, X, Theorem 7.1]. Because the rational function $mR_{\mathbf{b}}$ has a zero at ∞ and a pole of order at most 1 at each b_i or b_i', the Swan conductor of $[mR_{\mathbf{b}}]^* \mathcal{K}\ell_2$ at ∞ vanishes and the Swan conductor of $[zR_{\mathbf{b}}]^* \mathcal{K}\ell_2$ at b_i or b_i' is at most 1, so the total Euler characteristic is at least

$$2 - 3\#\{-b_1, \dots, -b_3'\} \geq 2 - 3 \cdot 6 = -16. \tag{A.29}$$

□

Corollary A.2. *Keeping the same notation, we have if $p > 3$*

$$\sum_{h\in\mathbb{M}_{\mathbf{b}}}\sum_{\substack{s\in\mathbb{M}^{\times}\\ d(s)\leq k+d_A}}\psi\left(\Delta_{\mathbf{b}}(h,s)\right)\ll d_2(M)^5\left(|M|+q^{\frac{3(r+k)}{4}}\left|\mathrm{lcm}_{\sigma\in S_3}\,M_\sigma^{\mathbf{b}}\right|\right).$$

and, if $p=3$, the same bound but with $M_{\triangle}^{\mathbf{b}}$ also included in the lcm.

Proof. Using Eq. (A.11), Eq. (A.16), and Proposition A.1 we get a bound of

$$\ll q^{k+d_A}d_2(M)^4\sum_{\substack{z\in\mathbb{F}_q[T]\\ d(z)<d(M)-k-d_A}}\left|\mathrm{lcm}_{\sigma\in S_3}\gcd\left(M_\sigma^{\mathbf{b}},z\right)\right| \tag{A.30}$$

for the left hand side above. Summing over the possible values of the least common multiple, we get

$$\begin{aligned}&\ll q^{\frac{3(r+k)}{4}}d_2(M)^4\sum_{L|\mathrm{lcm}_{\sigma\in S_3}M_\sigma^{\mathbf{b}}}|L|\sum_{\substack{z\in\mathbb{F}_q[T]\\ d(z)<d(M)-\frac{3(r+k)}{4}\\ L|z}}1\\ &\ll q^{\frac{3(r+k)}{4}}d_2(M)^4\sum_{L|\mathrm{lcm}_{\sigma\in S_3}M_\sigma^{\mathbf{b}}}|L|\max\left\{q^{d(M)-\frac{3(r+k)}{4}-d(L)},1\right\}.\end{aligned} \tag{A.31}$$

The contribution of $q^{d(M)-\frac{3(r+k)}{4}-d(L)}$ (respectively, of 1) is the first (respectively, the second) summand of the right hand side in our corollary. □

Corollary A.3. *Notation unchanged, we have*

$$\sum_{b_1,\dots,b_3'\in\mathcal{M}_{d_B}}\left|\sum_{h\in\mathbb{M}_{\mathbf{b}}}\sum_{\substack{s\in\mathbb{M}^{\times}\\ d(s)\leq k+d_A}}\psi\left(\Delta_{\mathbf{b}}(h,s)\right)\right|\ll|M|d_2(M)^{12}q^{\frac{6}{4}(k+r)}. \tag{A.32}$$

Proof. By Corollary A.2 we have a bound of

$$\ll\sum_{b_1,\dots,b_3'\in\mathcal{M}_{d_B}}d_2(M)^5\left(|M|+q^{\frac{3(r+k)}{4}}\left|\mathrm{lcm}_{\sigma\in S_3}\,M_\sigma^{\mathbf{b}}\right|\right). \tag{A.33}$$

Summing first over tuples M_σ of divisors of M we get

$$\sum_{\sigma\in S_3}\sum_{M_\sigma|M}\sum_{\substack{b_1,\dots,b_3'\in\mathcal{M}_{d_B}\\ \forall\sigma\ M_\sigma=M_\sigma^{\mathbf{b}}}}d_2(M)^5\left(|M|+q^{\frac{3(r+k)}{4}}\left|\mathrm{lcm}_{\sigma\in S_3}\,M_\sigma\right|\right). \tag{A.34}$$

For each such tuple, the conditions

$$M_\sigma=M_\sigma^{\mathbf{b}},\quad \sigma\in S_3 \tag{A.35}$$

imply the congruences

$$b_i\equiv b'_{\sigma(i)}\bmod M_\sigma,\quad 1\leq i\leq 3,\ \sigma\in S_3 \tag{A.36}$$

which determine b_1', b_2', b_3' mod $\mathrm{lcm}_{\sigma \in S_3} M_\sigma$ once b_1, b_2, b_3 are chosen. Hence, for each tuple of divisors of M, the number of possible values of $\mathbf{b}$ is at most

$$\max\left\{q^{6d_B}\left|\mathrm{lcm}_{\sigma\in S_3} M_\sigma\right|^{-3}, q^{3d_B}\right\}. \tag{A.37}$$

If $p = 3$, we need a slightly more complicated argument. There are at most $M_\triangle^2$ ways to choose the congruence classes of $\mathbf{b}$ modulo $M_\triangle$, and then choosing b_1, b_2, b_3 arbitrarily now determines b_1', b_2', b_3' mod $\mathrm{lcm}_{\sigma \in S_3 \cup \{\triangle\}} M_\sigma$. Because the number of ways to choose $\mathbf{b}$ modulo M_σ and then choose b_1, b_2, b_3 is

$$\begin{cases} M_\triangle^2 (q^{3d_B}/M_\triangle^3) \le q^{3d_B} & \text{if } M_\triangle \le q^{d_B} \\ q^{2d_B} \le q^{3d_B} & \text{if } M_\triangle > q^{d_B} \end{cases} \tag{A.38}$$

in either case the number of possible values of $\mathbf{b}$ is at most

$$\max\left\{q^{6d_B}\left|\mathrm{lcm}_{\sigma\in S_3\cup\{\triangle\}} M_\sigma\right|^{-3}, q^{3d_B}\right\}. \tag{A.39}$$

Setting $\tau = d(\mathrm{lcm}_{\sigma\in S_3} M_\sigma)$ (or adding $\triangle$ if $p = 3$), and taking the maximal possible contribution for every tuple of divisors of M, we get the bound

$$\max_{0\le\tau\le d(M)} d_2(M)^{|S_3|+1}\left(q^{6d_B-3\tau} + q^{3d_B}\right) d_2(M)^5 \left(|M| + q^{\frac{3(r+k)}{4}+\tau}\right). \tag{A.40}$$

Expanding the brackets above, we see that each exponent is a linear function of τ, hence maximized either at $\tau = 0$ or at $\tau = d(M)$. Using Eq. (A.6), one observes that the maximal terms $q^{6d_B+d(M)}$ and $q^{3d_b+\frac{3(r+k)}{4}+d(M)}$ agree and arrives at the right hand side of Eq. (A.32). □

It then follows from Eq. (A.8), Eq. (A.32), and the divisor bound that

$$\sum_{\substack{a\in\mathcal{M}_{d_A}\\(a,M)=1}}\sum_{f\in\mathcal{M}_k}\sum_{g\in\mathcal{M}_r}\left|\sum_{b\in\mathcal{M}_{d_B}}\psi\left(af(a^{-1}g+b)\right)\right| \ll q^{\frac{41}{24}r+\frac{21}{24}k+(\frac{1}{6}+\epsilon)d(M)}$$

and thus (matching the $\ell = 3$ case of [FM98, (1.2.3)]) we get

$$\Sigma_{k,r}^{(\mathrm{I})} \ll q^{\frac{17}{24}r+\frac{7}{8}k+(\frac{1}{6}+\epsilon)d(M)} \tag{A.41}$$

using Eq. (A.5). Using the fact that $k \le \frac{1}{8}d(M) + \frac{7}{16}d$ and $k + r \le d$ we arrive at the first bound in Eq. (A.4).

A.2. **Sums of type II.** Keeping the same notation, we follow the proof of [FKM14, Theorem 1.17]. Applying Cauchy's inequality and Polya-Vinogradov

completion as in [FKM14, Section 3], we get

$$\begin{aligned}\left|\Sigma_{k,r}^{(\mathrm{II})}\right|^2 &\le q^k \sum_{g_1,g_2\in\mathcal{M}_r} \overline{\delta_{g_1}}\delta_{g_2} \sum_{\substack{f\in\mathcal{M}_k\\(fg_1,M)=(fg_2,M)=1}} \psi(\overline{fg_1}-\overline{fg_2}) \\ &\le q^k \sum_{g_1,g_2\in\mathcal{M}_r} q^{k-d(M)} \sum_{\substack{h\in\mathbb{F}_q[T]\\d(h)<d(M)-k}} \mathcal{C}(\overline{g_1}-\overline{g_2},h).\end{aligned} \tag{A.42}$$

where

$$\mathcal{C}(g,h)=\sum_{z\in\mathbb{M}^\times}\psi(g\overline{z})e_p\left(\mathrm{Tr}_{\mathbb{F}_p}^{\mathbb{M}}(hz)\right),\quad g,h\in\mathbb{F}_q[T]. \tag{A.43}$$

We have the following analog of [FKM14, Proposition 3.1].

Proposition A.4. *For $g,h\in\mathbb{F}_q[T]$ and $M_{g,h}=\gcd(M,g,h)$ we have*

$$|\mathcal{C}(g,h)|\le d_2(M)\sqrt{|M|}\sqrt{|M_{g,h}|}. \tag{A.44}$$

Proof. Since both $\mathcal{C}(g,h)$ and our putative bound are multiplicative in M, it suffices to show that for a prime P we have

$$|\mathcal{C}(g,h)|\le 2\sqrt{|P|} \tag{A.45}$$

unless $g\equiv h\equiv 0 \bmod P$. To demonstrate that, take $f\in\mathbb{F}_q[T]/(P)$ with

$$\psi(z)=e_p\left(\mathrm{Tr}_{\mathbb{F}_p}^{\mathbb{F}_q[T]/(P)}(fz)\right) \tag{A.46}$$

and note that

$$\mathcal{C}(g,h)=\sum_{z\in(\mathbb{F}_q[T]/(P))^\times} e_p\left(\mathrm{Tr}_{\mathbb{F}_p}^{\mathbb{F}_q[T]/(P)}(gfz^{-1}+hz)\right). \tag{A.47}$$

We have a Weil bound of $2\sqrt{|P|}$ for this exponential sum unless the rational function $gfz^{-1}+hz$ is an Artin-Schreier polynomial, which can only happen if it is constant, as all its poles have order at most 1. The latter happens only if $g=h=0$, as desired. □

By Eq. (A.42) we have

$$\left|\Sigma_{k,r}^{(\mathrm{II})}\right|^2 \le q^{2k-d(M)}\sum_{L|M}\sum_{\substack{h\in\mathbb{F}_q[T]\\d(h)<d(M)-k}}\sum_{\substack{g_1,g_2\in\mathcal{M}_r\\M_{\overline{g}_1-\overline{g}_2,h}=L}}\mathcal{C}(\overline{g_1}-\overline{g_2},h) \tag{A.48}$$

so from Proposition A.4 and the divisor bound, this is at most

$$q^{2k-\frac{d(M)}{2}+\epsilon d(M)}\sum_{L|M}q^{\frac{d(L)}{2}}\sum_{\substack{h\in\mathbb{F}_q[T]\\d(h)<d(M)-k}}\sum_{\substack{g_1,g_2\in\mathcal{M}_r\\M_{g_1-g_2,h}=L}}1. \tag{A.49}$$

Since $M_{g_1-g_2,h}=L$ implies that $g_2\equiv g_1,\ h\equiv 0 \bmod L$, the above is at most

$$q^{2k-\frac{d(M)}{2}+\epsilon d(M)}\sum_{L|M}q^{\frac{d(L)}{2}+r+\max(r-d(L),0)+\max(d(M)-k-d(L),0)} \tag{A.50}$$

so setting $\xi = d(L)$ and applying the divisor bound once again, we arrive at

$$\max_{0\leq\xi\leq d(M)} q^{2k+\frac{\xi}{2}+r+\max\{r-\xi,0\}+\max\{d(M)-k-\xi,0\}-\frac{d(M)}{2}+\epsilon d(M)}. \tag{A.51}$$

As a function of ξ, the exponent above is convex, so its values do not exceed those at $\xi = 0$ and $\xi = d(M)$, which are

$$q^{\frac{d(M)}{2}+\epsilon d(M)+2r+k}, \quad q^{r+2k+\epsilon d(M)} \tag{A.52}$$

in view of our assumption that $r \leq d \leq d(M)$. Taking a square root and using the fact that $k + r \leq d$ we get the second bound in (A.4).

Remark A.5. There are (at least) two other potential approaches for bounding our type II sums. The first is to follow the proof of [FM98, Proposition 1.3] that uses Bombieri's bound on complete exponential sums [FM98, Lemma 4.3]. One then argues as in [FM98, Section 5]. The second is to follow the proof of [FM98, Theorem 1.4] given in [FM98, Section 7].

References

[BBSR15] E. Bank, L. Bary-Soroker, L. Rosenzweig, *Prime polynomials in short intervals and in arithmetic progressions*, Duke Math. J. 164.2 (2015): 277-295.

[BSE21] L. Bary-Soroker, A. Entin, *Explicit Hilbert's irreducibility theorem in function fields*, Contemp. Math. 767 (2021), 125-134.

[BBD82] A. A. Beilinson, J. Bernstein, P. Deligne, *Faisceaux Pervers*, Asterisque 100 (1982).

[BB73] A. Bialynicki–Birula, *On fixed point schemes of actions of multiplicative and additive groups*, Topology 12 (1973), 99-103.

[BGP92] E. Bombieri, A. Granville, J. Pintz, *Squares in arithmetic progressions*, Duke Math. J. 66.3 (1992): 369-385.

[BFI89] E. Bombieri, J. Friedlander, H. Iwaniec, (1989). *Primes in arithmetic progressions to large moduli III*, JAMS, 2(2), 215-224.

[Bur63] D. A. Burgess (1963), *On Character Sums and L-Series II*, Proc. Lond. Math. Soc. 3(1), 524-536.

[Ca15] D. Carmon, *The autocorrelation of the Möbius function and Chowla's conjecture for the rational function field in characteristic* 2, Phil. Trans. R. Soc. A 2015.

[Car21] D. Carmon, *On square-free values of large polynomials over the rational function field*, with an appendix by A. Entin, Mathematical Proceedings of the Cambridge Philosophical Society. Vol. 170. No. 2. Cambridge University Press, 2021.

[CaRu14] D. Carmon, Z. Rudnick, *The autocorrelation of the Möbius function and Chowla's conjecture for the rational function field*, Q. J. Math. (2014) 65 (1):53–61.

[CHLPT15] A. Castillo, C. Hall, R. Lemke Oliver, P. Pollack, L. Thompson, (2015) *Bounded gaps between primes in number fields and function fields*, Proc. Amer. Math. Soc.143(7), 2841-2856.

[CG07] J. Cilleruelo, A. Granville, *Lattice points on circles, squares in arithmetic progressions and sumsets of squares*, Additive combinatorics, 43, 241-262, Amer. Math. Soc. Providence, RI, 2007.

[CCG08] B. Conrad, K. Conrad, and R. Gross (2008) *Prime specialization in genus 0*, Transactions of the AMS, 360(6), 2867-2908.

[Con05] K. Conrad (2005), *Irreducible values of polynomials: a non-analogy*, Number Fields and Function Fields: Two Parallel Worlds, 71-85, Progress in Mathematics, 239, Birkhauser, Basel.

[Del80] P. Deligne (1980), *La Conjecture De Weil II*, Publications Mathématiques de l'IHÉS, 52. 137-252.

[FM98] E. Fouvry, P. Michel, *Sur certaines sommes d'exponentielles sur les nombres premiers*, Ann. scient. Ec. Norm. Sup., 4e serie, t. 31, 1998, 93-130.

[FKM14] E. Fouvry, E. Kowalski, P. Michel (2014), *Algebraic trace functions over the primes*, Duke Math. J. 163(9), 1683-1736.

[Gal72] P. X. Gallagher (1972), *Primes in progressions to prime-power modulus*, Invent. math. 16(3), 191-201.

[GY03] D. Goldston, C. Yıldırım, *Higher correlations of divisor sums related to primes I: Triple correlations*, Integers 3 (2003) A5.

[GS20] O. Gorodetsky, W. Sawin, *Correlation of Arithmetic Functions over* $\mathbb{F}_q[T]$, Mathematische Annalen, 376, 1059-1106, 2020.

[Hal06] C. Hall (2006), *L-functions of twisted Legendre curves*, J. of Number Theory, 119, 128-147.

[Hoo91] C. Hooley, *On the number of points on a complete intersection over a finite field*, J. of Number Theory, 38.3 (1991), 338-358.

[Hsu99] C. N. Hsu (1999), *Estimates for Coefficients of L-Functions for Function Fields*, Fin. Fields App. 5(1), 76-88.

[Ill94] L. Illusie (1994) *Autour du théorême de monodromie locale*, Astérisque 223 (1994) 9-57..

[IK04] H. Iwaniec, E. Kowalski, *Analytic number theory*, Vol. 53. Amer. Math. Soc. 2004.

[Jan07] S. Janson (2007), *Resultant and discriminant of polynomials*, Note N5. http://www2.math.uu.se/~svante/papers/sjN5.pdf.

[Ka88] N. Katz, (1988). *Gauss Sums, Kloosterman Sums, and Monodromy Groups*, Princeton university press.

[Kat89] N. Katz (1989), *An Estimate for Character Sums*, JAMS, 2, 2, 197-200.

[KR14] L. Kindler, K. Rülling, (2014). *Introductory course on ℓ-adic sheaves and their ramification theory on curves*, arXiv preprint, 1409.6899.

[KMT20] O. Klurman, A. Mangerel, J. Teräväinen, (2020). *Correlations of multiplicative functions in function fields*, preprint, arXiv:2009.13497.

[Lan15] G. Lando (2015), *Square-free values of polynomials evaluated at primes over a function field*, The Quarterly Journal of Mathematics, 66, 3, 905-924.

[Ma15] J. Maynard (2015), *Small gaps between primes*, Ann. Math. 181, 1-31.

[Mil] J. Milne, *Lectures on Étale Cohomology*, https://www.jmilne.org/math/CourseNotes/LEC.pdf.

[MR16] K. Matomäki, M. Radziwiłł (2016), *Multiplicative functions in short intervals*, Ann. Math. 183, 1015-1056.

[MS04] H. Montgomery, K. Soundararajan, (2004). *Primes in Short Intervals*, Comm. Math. Phys. 252, 1-3, 589-617.

[MV17] R. Murty, A. Vatwani (2017), *Twin primes and the parity problem*, J. Number Theory, 180, 643-659.

[Pol08] P. Pollack, *An explicit approach to Hypothesis H for polynomials over a finite field*, in *Anatomy of integers*, CRM Proc. Lecture Notes 46, 259–273. Amer. Math. Soc., Providence, RI, 2008.

[PM14] Polymath, D. H. J. (2014), *Variants of the Selberg sieve, and bounded intervals containing many primes*, Res. Math. sci. 1(1), 12.

[PM14B] Polymath, D. H. J. (2014), *New equidistribution estimates of Zhang type*, Algebra & Number Theory 8.9, 2067-2199.

[Poo03] B. Poonen (2003), *Squarefree values of multivariable polynomials*, Duke Math. J. 118, 2, 353-373.

[Ray66] M. Raynaud (1966), *Caractŕistique d'Euler-Poincaré d'un faisceau et cohomologie des variétés abéliennes*, Séminaire N. Bourbaki, exp. no 286, p. 129-147.

[Ros02] M. Rosen, *Number theory in function fields*, Graduate Texts in Mathematics Vol. 210. Springer Science & Business Media, 2002.

[Saw20] W. Sawin, *Singularities and vanishing cycles in number theory over function fields*, arXiv preprint, 2020.

[SS20] W. Sawin, M. Shusterman, (2020). *Möbius cancellation on polynomial sequences and the quadratic Bateman–Horn conjecture over function fields*, preprint, arXiv:2008.09905.

[SGA4-3] M. Artin, A. Grothendieck. J.-L. Verdier, eds, *Séminaire de Géométrie Algébrique du Bois Marie - 1963-64 - Théorie des topos et cohomologie étale des schémas - (SGA 4) - vol. 3*, Lecture Notes in Mathematics **305** Springer-Verlag, 1972.

[SGA4$\frac{1}{2}$] P. Deligne, ed, *Séminaire de Géométrie Algébrique du Bois Marie - Cohomologie étale - (SGA 4$\frac{1}{2}$)*, Lecture Notes in Mathematics **569** Springer-Verlag, 1977.

[SGA5] A. Grothendieck, (1977), *Séminaire de Géométrie Algébrique du Bois Marie – 1965-66 – Cohomologie ℓ-adique et Fonctions L – (SGA 5)*, Lecture notes in mathematics, **589**, Berlin; New York: Springer-Verlag.

[SGA7-II] P. Deligne, N. Katz, eds. *Séminaire de Géométrie Algébrique du Bois Marie - 1967-69 - Groupes de monodromie en géométrie algébrique - (SGA 7) - vol. 2*, Lecture Notes in Mathematics (in French), Vol. 340. Springer-Verlag.

[Tao16] T. Tao, *The logarithmically averaged Chowla and Elliott conjectures for two-point correlations*, Forum of Mathematics, Pi, Vol 4, 2016, e8, Cambridge University Press.

[TT19] T. Tao, J. Teräväinen (2019), *The structure of logarithmically averaged correlations of multiplicative functions, with applications to the Chowla and Elliott conjectures*, Duke Math. J. 168(11), 1977-2077.

[Zha14] Y. Zhang (2014), *Bounded gaps between primes*, Ann. Math. 179, 1121-1174.

Department of Mathematics, Columbia University, New York, NY 10027, USA

Email address: sawin@math.columbia.edu

Department of Mathematics, University of Wisconsin-Madison, 480 Lincoln Drive, Madison, WI 53706, USA

Email address: mshusterman@wisc.edu