Received: 28 April 2022 | Revised: 9 September 2022 | Accepted: 21 October 2022

DOI: 10.1112/mtk.12181

RESEARCH ARTICLE

Mathematika

Correlations of multiplicative functions in function fields

Oleksiy Klurman[1] | **Alexander P. Mangerel**[2] | **Joni Teräväinen**[3]

[1]School of Mathematics, University of Bristol, Woodland Road, Bristol, UK

[2]Department of Mathematical Sciences, Durham University, Upper Mountjoy Campus, Stockton Road, Durham, UK

[3]Department of Mathematics and Statistics, University of Turku, Turku, Finland

Correspondence
Joni Teräväinen, Department of Mathematics and Statistics, University of Turku, 20014 Turku, Finland.
Email: joni.p.teravainen@gmail.com

Funding information
Titchmarsh Fellowship; Academy of Finland, Grant/Award Number: 340098

Abstract

We develop an approach to study character sums, weighted by a multiplicative function $f : \mathbb{F}_q[t] \to S^1$, of the form

$$\sum_{\substack{\deg(G)=N \\ G \text{ monic}}} f(G)\chi(G)\xi(G),$$

where χ is a Dirichlet character and ξ is a short interval character over $\mathbb{F}_q[t]$. We then deduce versions of the Matomäki–Radziwiłł theorem and Tao's two-point logarithmic Elliott conjecture over function fields $\mathbb{F}_q[t]$, where q is fixed. The former of these improves on work of Gorodetsky, and the latter extends the work of Sawin–Shusterman on correlations of the Möbius function for various values of q. Compared with the integer setting, we encounter a different phenomenon, specifically a low characteristic issue in the case that q is a power of 2. As an application of our results, we give a short proof of the function field version of a conjecture of Kátai on classifying multiplicative functions with small increments, with the classification obtained and the proof being different from the existing one in the integer case. In a companion paper, we use these results to characterize the limiting behavior of partial sums of multiplicative functions in

function fields and in particular to solve a "corrected" form of the Erdős discrepancy problem over $\mathbb{F}_q[t]$.

MSC 2020
11T55, 11N37 (primary)

1 | INTRODUCTION AND RESULTS

In the integer setting, there has been a lot of progress in understanding short sums

$$\sum_{x<n\leqslant x+H} f(n), \quad \text{with } 1\leqslant H\leqslant x, \tag{1}$$

of multiplicative functions $f:\mathbb{N}\to\mathbb{C}$, as well as their correlations

$$\frac{1}{x}\sum_{n\leqslant x} f_1(n)f_2(n+h), \quad \text{for } h\geqslant 1. \tag{2}$$

See [27], [29, Theorem A.1], [28] for some papers dealing with (1) and [20, 29, 33, 36] for some papers dealing with (2). These results have also led to a number of applications, including a solution by Tao [32] to the famous Erdős discrepancy problem.

Let q be a fixed prime power and denote by $\mathbb{F}_q[t]$ the ring of polynomials in t over $\mathbb{F}_q$. Our focus in this paper is on analogs of (1) and (2) over $\mathbb{F}_q[t]$. These results have applications, in particular, to the Erdős discrepancy problem over $\mathbb{F}_q[t]$, which we study in our follow-up paper [24]. In the course of the proofs of our main results, we develop a substantial amount of pretentious number theory over $\mathbb{F}_q[t]$.

Let $\mathcal{M}$ denote the set of monic polynomials in $\mathbb{F}_q[t]$. Also, denote by $\mathcal{M}_{\leqslant N}$ and $\mathcal{M}_N$ the sets of monic polynomials of degree $\leqslant N$ or $=N$, respectively. Let $\mathcal{P}$ be the set of irreducible monic polynomials in $\mathbb{F}_q[t]$. Again, define $\mathcal{P}_{\leqslant N}$ and $\mathcal{P}_N$ similarly. Finally, let $\mathbb{U}$ stand for the unit disc of the complex plane.

By a Dirichlet character $\chi:\mathbb{F}_q[t]\to\mathbb{C}$ modulo $M\in\mathcal{M}$, we mean a multiplicative homomorphism $\chi:(\mathbb{F}_q[t]/M\mathbb{F}_q[t])^\times\to\mathbb{C}\setminus\{0\}$, extended to all of $\mathbb{F}_q[t]$ by setting $\chi(G)=0$ whenever G and M are not coprime.

We first describe our result on short sums of multiplicative functions. This provides an analog of the celebrated Matomäki–Radziwiłł theorem [27] in function fields.

Matomäki and Radziwiłł showed that, for any bounded, real-valued multiplicative function $f:\mathbb{N}\to[-1,1]$, one has

$$\frac{1}{X}\int_X^{2X}\left|\frac{1}{H}\sum_{x<n\leqslant x+H} f(n)-\frac{1}{X}\sum_{X<n\leqslant 2X} f(n)\right|^2 dx = o(1),$$

as soon as $H=H(X)\to\infty$ with X. Thus, the short sums of f over $[x,x+H]$ are almost always asymptotic to the corresponding long sum of f over $[X,2X]$, which can either be understood asymptotically or upper bounded non-trivially by Halász's theorem (see [35, Section III.4.3] for further details).

In function fields, the role of a short interval is played by[†]

$$I_H(G_0) := \{G \in \mathcal{M} : \ \deg(G - G_0) < H\}, \quad G_0 \in \mathcal{M}.$$

We prove a function field version of the aforementioned result for sums over such short intervals, following a line of approach which differs somewhat from the result over the integers. We state this as follows.

Theorem 1.1 (Matomäki–Radziwiłł Theorem for Function Fields, Real Case). *Let $f : \mathcal{M} \to [-1,1]$ be a multiplicative function. Let N be large and let $1 \leqslant H \leqslant N - N^{3/4}$ with $H = H(N) \to \infty$ as $N \to \infty$.*

(i) *If q is odd, we have*

$$q^{-N} \sum_{G_0 \in \mathcal{M}_N} \left| q^{-H} \sum_{G \in I_H(G_0)} f(G) - q^{-N} \sum_{G \in \mathcal{M}_N} f(G) \right|^2 \ll \frac{\log H}{H} + N^{-1/18+o(1)}.$$

(ii) *If q is even, we have*

$$q^{-N} \sum_{G_0 \in \mathcal{M}_N} \left| q^{-H} \sum_{G \in I_H(G_0)} f(G) - q^{-N} \sum_{G \in \mathcal{M}_N} f(G)\overline{\chi_1^*}(G) \right|^2 \ll \frac{\log H}{H} + N^{-1/18+o(1)},$$

where $\chi_1 \bmod t^{N-H+1}$ is a real character that minimizes the map

$$\chi \mapsto \min_{\theta \in [0,1]} \sum_{P \in \mathcal{P}_{\leqslant N}} \frac{1 - Re(f(P)\overline{\chi}(P))}{q^{\deg(P)}}, \quad \chi \bmod t^{N-H+1},$$

and where χ_1^* is the completely multiplicative function satisfying $\chi_1^*(t) = 1$ and $\chi_1^*(G) := \chi_1(t^{\deg(G)}G(1/t))$ for all G coprime to t.

Remarks.

- The long sum $\sum_{G \in \mathcal{M}_N} f(G)$ appearing in Theorem 1.1 is very well understood, as in the integer setting. This is thanks to a version of Halász's theorem over function fields, established by Granville, Harper, and Soundararajan [9].
- The savings $(\log H)/H$ obtained is of the same quality as the h-dependence found in [27, Theorems 1 and 2] (replacing H by $\log h$ there). This term arises from a sieve-theoretic bound that allows us to restrict the support of the intervening sums in our analysis to polynomials G having prime factors with degrees in specified intervals, depending on H (see Lemma 4.10 below). This term is not expected to be optimal in general, and standard "square-root cancellation" heuristics for sufficiently pseudo-random multiplicative functions (such as the Liouville function $\lambda : \mathcal{M} \to \{-1,1\}$, the completely multiplicative function defined at all prime polynomials P by $\lambda(P) = -1$) suggests a bound of the shape $q^{-H(1/2-\varepsilon)}$. In this connection, it is worth not-

[†] For ease of comparison with prior function field literature, specifically the work of Keating and Rudnick [18, 19], we note that our short interval $I_H(G_0)$ corresponds to $I(G_0; H-1)$ in the notation of [18].

ing that in the integer setting, Chinis [3, Theorem 1.2] has shown that, assuming the Riemann Hypothesis, the short sums $h^{-1}\sum_{x<n\leqslant x+h}\lambda(n)$ of the Liouville function λ exhibit a corresponding error term of the quality $h^{-1/2+\varepsilon}$ in mean square whenever $h \geqslant (\log X)^A$ with $A = A(\varepsilon) > 0$. It might be interesting to pursue a similar result in the $\mathbb{F}_q[t]$ setting.

- Note that, interestingly, a low-characteristic issue emerges in the Matomäki–Radziwiłł theorem: In $\mathbb{F}_2[t]$, for instance, a real-valued multiplicative function can indeed have different mean values on short and long intervals. This is the reason why we have stated the cases of q odd and even separately in Theorem 1.1. Functions of the form χ_1^*, where χ_1 is a character modulo a power of t, are examples of *short interval characters*; see Definition 1.3 below, as well as Sections 3.3 and 3.4 for further details relating to the transformation $\chi_1 \mapsto \chi_1^*$.
- Theorem 1.1 can be viewed as generalizing and strengthening the work of Gorodetsky [7, Theorem 1.3], who proved that for any *factorization function*[†] f and for $H = H(N)$ satisfying $H \log\log N/\log N \to \infty$, the sum of f over a short interval $I_H(G_0)$ is almost always asymptotic to the corresponding long sum. Neither the class of factorization functions nor the class of multiplicative functions contains the other, but their intersection contains several interesting number theoretic functions; for example, one of the most important functions in both classes is the Möbius function $\mu\colon \mathcal{M} \to \{-1, 0, +1\}$ (defined as $\mu(G) := (-1)^s$ if G is squarefree and has s irreducible factors, and $\mu(G) := 0$ otherwise). In Theorem 1.1, we do not have any lower bound on how quickly the length H of the interval must grow, which is vital when we use this result to deduce Theorem 1.4.

We in fact establish a slightly more general version of Theorem 1.1 (namely, Theorem 6.1) that applies to bounded complex-valued multiplicative functions as well, but omit the more complicated statement here for the sake of simplicity.

It is also natural to study the variance of multiplicative functions in arithmetic progressions; see [11, 14] for some works on this topic. In the integer setting, an estimate for the variance of a multiplicative function in arithmetic progressions that is of comparable strength to the Matomäki–Radziwiłł theorem was established in [23]. Here, we generalize this result to function fields, obtaining in fact a stronger version that does not involve exceptional[‡] moduli. For multiplicative factorization functions, this also improves on a corresponding result of Gorodetsky [7, Theorem 1.3].

Theorem 1.2 (Variance of Multiplicative Functions in Arithmetic Progressions). *Let $1 \leqslant H \leqslant N - N^{3/4}$, such that $H = H(N) \to \infty$ as $N \to \infty$. Let $f\colon \mathcal{M} \to \mathbb{U}$ be a multiplicative function. For every $Q \in \mathcal{M}_{N-H}$, there is a character χ_1 modulo Q such that*

$$\sum_{A \bmod Q}^{*}\left|\sum_{\substack{G\in\mathcal{M}_N\\ G\equiv A \bmod Q}} f(G) - \frac{\chi_1(A)}{\phi(Q)}\sum_{G\in\mathcal{M}_N} f(G)\overline{\chi_1}(G)\right|^2 \ll \left(\frac{\log H}{H} + N^{-1/18+o(1)}\right)q^{2N-\deg(Q)}.$$

[†] A function $f(G)$ is called a factorization function if it only depends on the values of $\deg(P)$ and $v_P(G)$, where P runs through the irreducible divisors of G, and $v_P(G)$ denotes the largest integer k with $P^k \mid G$.

[‡] It should be noted that if one assumes GRH (generalised Riemann hypothesis) in the integer setting then [23, Theorem 1.4] also gives non-trivial estimates for the corresponding variance for *all* moduli q without exception, at least as long as q does not have too many "small" prime factors. Note crucially that such a constraint on Q in our setting is not needed for our result as it is stated, simply because our savings are given relative to the (possibly worse-than-trivial) bound $q^{2N-\deg(Q)}$, rather than the sharper $\phi(Q)q^{2(N-\deg(Q))}$. It is the factor $\phi(Q)q^{-\deg(Q)}$, ignored here, that is affected by the primes of small degree.

Precisely, χ_1 is any character modulo Q that minimizes the map

$$\chi \mapsto \min_{\theta\in[0,1]} \sum_{P\in\mathcal{P}_{\leqslant N}} q^{-\deg(P)}\Big(1-\mathrm{Re}(f(P)\overline{\chi}(P)e^{-2\pi i\theta\deg(P)})\Big).$$

Next, we turn to our result on two-point correlations of multiplicative functions in function fields, with the objective of analogizing Tao's groundbreaking work in [33]. Tao's result states that if $f_1, f_2 : \mathbb{N}\to\mathbb{U}$ are multiplicative functions such that at least one of f_1 and f_2, say f_1, satisfies the non-pretentiousness assumption

$$\inf_{|t|\leqslant x}\sum_{p\leqslant x}\frac{1-\mathrm{Re}(f_1(p)\overline{\chi}(p)p^{-it})}{p}\to\infty \text{ as } x\to\infty$$

for any fixed Dirichlet character χ, then we have

$$\frac{1}{\log x}\sum_{n\leqslant x}\frac{f_1(n)f_2(n+h)}{n}=o(1)$$

for any fixed $h\neq 0$. The analog of the logarithmic weight $n\mapsto 1/n$ in function fields is $G\mapsto q^{-\deg(G)}$.

Tao's result implies that if f_1 does not pretend to be a twisted Dirichlet character $n\mapsto\chi(n)n^{it}$, then the autocorrelations of f_1 are small. It turns out that in the function field setting there are two collections of Archimedean characters that play a role similar to $n\mapsto n^{it}$, namely, the characters $G\mapsto e^{2\pi i\theta\deg(G)}$ as well as the *short interval characters*, to be defined presently (the group they generate will be discussed in further detail in Section 3). Dirichlet characters twisted by at least one of these functions provide obstructions to $f_1:\mathcal{M}\to\mathbb{U}$ having small autocorrelations in the setting of $\mathbb{F}_q[t]$. While in terms of phenomenology this is consistent with the integer setting, some of the arguments in the function field setting require some additional care to address both types of twists.

Definition 1.3. A multiplicative function $\xi:\mathcal{M}\to\mathbb{C}$ which is not identically zero is called a *short interval character* if there exists ν such that $\xi(A)=\xi(B)$ whenever the $\nu+1$ highest degree coefficients of A and B agree (i.e., $A/t^{\deg(A)}-B/t^{\deg(B)}$ is a rational function of degree $<-\nu$). If ν is the smallest positive integer with this property, then we refer to ν as the *length* of ξ, and write $\mathrm{len}(\xi)=\nu$.

Theorem 1.4 (Two-point logarithmic Elliott conjecture in function fields). *Let $A,B\in\mathbb{F}_q[t]\backslash\{0\}$ be fixed, with A monic. Let $f_1,f_2:\mathcal{M}\to\mathbb{U}$ be multiplicative functions. Assume that f_1 satisfies the non-pretentiousness assumption*

$$\min_{M\in\mathcal{M}_{\leqslant W}}\min_{\psi \bmod M}\min_{\substack{\xi \text{ short}\\ \mathrm{len}(\xi)\leqslant N}}\min_{\theta\in[0,1]}\sum_{P\in\mathcal{P}_{\leqslant N}}\frac{1-\mathrm{Re}(f_1(P)\overline{\psi}(P)\overline{\xi}(P)e^{-2\pi i\theta\deg(P)})}{q^{\deg(P)}}\to\infty, \tag{3}$$

as $N\to\infty$ for every fixed $W\geqslant 1$. Then,

$$\frac{1}{N}\sum_{G\in\mathcal{M}_{\leqslant N}}q^{-\deg(G)}f_1(G)f_2(AG+B)=o(1) \tag{4}$$

as $N\to\infty$.

Moreover, if f_1 is real-valued and q is odd, then the same conclusion follows provided only that

$$\min_{M\in\mathcal{M}_{\leqslant W}} \min_{\psi \bmod M} \min_{\theta\in\{0,1/2\}} \sum_{P\in\mathcal{P}_{\leqslant N}} \frac{1-\mathrm{Re}(f_1(P)\overline{\psi}(P)e^{-2\pi i\theta\deg(P)})}{q^{\deg(P)}} \to \infty \tag{5}$$

as $N\to\infty$.

Remark 1.5. Observe that if ξ is a short interval character of length ν, $m\geqslant 2\nu$ and $\deg(B)<\nu$, then $\xi(AG+B)=\xi(A)\xi(G)$ for any $G\in\mathcal{M}_m$. As $\xi(A)\in S^1$, it follows that as $N\to\infty$,

$$\left|\frac{1}{N}\sum_{G\in\mathcal{M}_{\leqslant N}} q^{-\deg(G)}\xi(G)\overline{\xi}(AG+B)\right| \geqslant 1-\frac{2\nu}{N},$$

so short interval characters clearly present a class of functions with large two-point correlations. This explains why our non-pretentiousness assumption must rule out significant correlations of f_1 with such characters.

Since the Möbius function $\mu\colon \mathcal{M}\to\{-1,0,+1\}$ is non-pretentious in the sense of (3) (by an application of Lemma 3.1 below), this result has the following corollary regarding Chowla's conjecture in function fields.

Corollary 1.6 (Two-point logarithmic Chowla conjecture in function fields). *Let $B\in\mathbb{F}_q[t]\backslash\{0\}$ be fixed. Let $\mu\colon \mathbb{F}_q[t]\to\{-1,0,+1\}$ be the Möbius function. Then, as $N\to\infty$,*

$$\frac{1}{N}\sum_{G\in\mathcal{M}_{\leqslant N}} q^{-\deg(G)}\mu(G)\mu(G+B)=o(1).$$

Remarks.

- Theorem 1.4 indicates that functions f that pretend to be twisted products of Dirichlet and short interval characters $\chi\xi e_\theta(G)$ (where $e_\theta(G):=e^{2\pi i\theta\deg(G)}$) are obstructions to the autocorrelations of f being small. This shows a different phenomenon compared to *mean values* of multiplicative functions in function fields, wherein the only obstructions to the mean value being small are functions pretending to be e_θ (see, for instance, Lemma 3.5 below); this is not necessarily unexpected since the problem of estimating mean values is not one that relates to short interval averages.
- Theorem 1.4 and Corollary 1.6 compare to previous results as follows. A recent groundbreaking result of Sawin and Shusterman [31] established the Chowla conjecture in function fields in the form

$$\frac{1}{q^N}\sum_{G\in\mathcal{M}_{\leqslant N}} \mu(G+B_1)\cdots\mu(G+B_k)=o(1),$$

as $N\to\infty$ for any $k\geqslant 1$ and any distinct $B_1,\dots,B_k\in\mathbb{F}_q[t]$ in the large field case $q>p^2k^2e^2$, where $p=\mathrm{char}(\mathbb{F}_q)$. In particular, if $q=p^a$, then we must have $a\geqslant 3$ for this condition to hold. Theorem 1.4 is somewhat orthogonal to this result in the sense that, despite being limited to two-point correlations, it works for *any* non-pretentious multiplicative functions, unlike the theorem in [31] which is specific to the Möbius function, and Theorem 1.4 works in *any* finite

field $\mathbb{F}_q$, which will be important for us. We also point out that the one-point case ($f_2 \equiv 1$) of Theorem 1.4 is (a logarithmic version of) Halász's theorem in function fields, proved by Granville, Harper and Soundararajan in [9].

In a different direction, when N is fixed and $q \to \infty$, Gorodetsky and Sawin [8, Theorem 3] obtained cancellation in two-point correlations $q^{-N}\sum_{G\in\mathcal{M}_N}\alpha(G)\beta(G+B)$, where α and β are factorization functions; in the q-limit this yields cancellation for the unweighted sums $q^{-N}\sum_{G\in\mathcal{M}_{\leqslant N}}\mu(G)\mu(G+B)$, for example (see [8, Theorem 2] for the precise statement).

Finally, we describe a short application of our results on Elliott's conjecture to the function field analog of a question of Kátai. Kátai [16] conjectured in 1983 that if $f:\mathbb{N}\to S^1$ is completely multiplicative and the consecutive values of f are close to each other on average, in the sense that

$$\sum_{n\leqslant x}|f(n+1)-f(n)| = o(x), \tag{6}$$

then $f(n)=n^{it}$ for some real number t. This was proved in [20] by the first author. Later, the result was generalized by Kátai and Phong [17] who proved that if $f,g:\mathbb{N}\to S^1$ are completely multiplicative and

$$\sum_{n\leqslant x}|g(2n+1)-zf(n)| = o(x) \tag{7}$$

for some complex number z, then $f(n)=g(n)=n^{it}$. Since in the function field setting there are two varieties of Archimedean characters, namely e_θ and short interval characters ξ, our classification of completely multiplicative functions satisfying (6) (and in fact more generally (7)) in function fields takes a slightly different form.

Theorem 1.7 (Kátai's conjecture in function fields). *Let $f:\mathcal{M}\to S^1$ be completely multiplicative, and let $Q\in\mathcal{M}$. Let $z\in S^1$. Suppose that*

$$\sum_{G\in\mathcal{M}_{\leqslant N}}|f(QG+1)+zf(G)| = o(q^N) \tag{8}$$

as $N\to\infty$. Then there exist $\theta\in[0,1)$ and a short interval character $\xi:\mathcal{M}\to\mathbb{U}$ such that $f(G)=\xi(G)e^{2\pi i\theta\deg(G)}$. Conversely, any function of this form satisfies (8) for some z.

The proof we give for this result is different[†] in various aspects from the proof in the integer setting in [20], and could be translated to produce a new proof of Kátai's result over the integers.

1.1 | Proof ideas

The proof of Matomäki and Radziwiłł [27] in the integer setting uses harmonic analysis methods that do not translate directly to function fields. In particular, the characters that control the short

[†] While the method of proof in [20, Section 5] could in principle be adapted to the function field setting, it would require a function field derivation of binary correlation formulas for multiplicative functions, as in [20, Corollary 3.4], which would likely lengthen this paper even further.

sum behavior in function fields are not the Archimedean characters n^{it} as in the integer setting, but rather the short interval characters from Definition 1.3. For our result on the variance in arithmetic progressions, in turn, the set of characters that control it are the Dirichlet characters. Thus, in order to deal with both theorems simultaneously, we study character sums weighted by f of the form

$$\sum_{G\in\mathcal{M}_N} f(G)\chi(G)\xi(G), \tag{9}$$

where χ is a Dirichlet character and ξ is a short interval character. Products of Dirichlet characters and short interval characters are called *Hayes characters* (the same terminology is used in [7] and stems from the fact that Hayes introduced these characters in [12]). Roughly speaking, we are able to follow the proof strategy of [23] with this set of characters rather than Dirichlet characters alone. In [23], however, our results only applied to characters whose modulus lies outside a small set of exceptional moduli, because of our incomplete understanding of zero-free regions for Dirichlet L-functions. In the function field setting, however, we can make use of a consequence of Weil's Riemann hypothesis due to Rhin [30] that shows that the L-functions corresponding to Hayes characters satisfy GRH, which implies that there are no exceptional moduli in this setting.

We gain some information in passing from the physical to the Fourier space versions of the problem by applying an involution (which we learned from the work of Keating and Rudnick [18, 19] and which appears earlier in the work of Hayes [12, e.g., pp. 115–116]) that relates short interval sums to sums over arithmetic progressions, that is,

$$\sum_{\substack{\deg(G)=N\\ G\in I_H(G_0)}} f(G) \leftrightarrow \sum_{\substack{\deg(G)=N\\ G\equiv A(G_0) \bmod t^{N-H+1}}} f^*(G),$$

where $A(G_0)$ is a residue class modulo t^{N-H+1} determined by G_0, and f^* is a kind of dual to f under the correspondence,[†] see Section 3.3 for further details (as well as [18, Section 5] for a nice exposition of this idea). For example, this allows us to gain some insight, in the case that q is even in Theorem 1.1, about the nature of the main term in the variance.

For proving our two-point Elliott result, we in fact need a generalized version of our Matomäki–Radziwiłł theorem in function fields, where we twist the multiplicative function by an additive character, thus looking at the short exponential sum

$$\sum_{G\in\mathcal{M}_N\cap I_H(G_0)} f(G)e_{\mathbb{F}}(\alpha G) \tag{10}$$

for almost all G_0 (see Section 2 for the relevant notation). This exponential sum is analyzed by adapting the approach of Matomäki–Radziwiłł–Tao from [29] to function fields (see Theorem 7.1). In particular, this involves performing the circle method in function fields, which is perhaps of independent interest.

To complete the proof, we develop a version of Tao's entropy decrement argument from [33, section 3] that allows us to express the two-point correlation as a two-variable correlation. By a bit of Fourier analysis, we can reduce the necessary estimate for this two-variable correlation sum to the estimate for (10) that we proved.

[†] Strictly speaking, one needs to restrict to G with $G(0) = 1$ for this to work.

We use our correlation results in the proof of Theorem 1.7 in order to reduce our classification of functions f to functions that are pretentious to a Hayes character $\chi\xi e_\theta$. We eschew the need for correlation formulas, as found in [20], by instead appealing to a concentration inequality that forces $f\overline{\chi\xi}e_{-\theta}$ to be close to 1 along structured sequences of polynomials. A judicious construction of such sequences leads to Theorem 1.7.

1.2 | Structure of the paper

The paper is organized as follows. In Section 3, we present some preliminary lemmas on the pretentious distance, the involution mentioned above, and Hayes characters. In Section 4, we introduce the remaining relevant preliminaries relating especially to mean square and pointwise estimates for character sums that will be needed in the proofs of Theorems 1.2, 6.1, and 1.4. In Section 5, we prove Theorem 1.2 using these lemmas. The proof of the Matomäki–Radziwiłł theorem (Theorem 1.1) proceeds completely analogously and is described in Section 6. In Section 7, we establish cancellation in exponential sums over short intervals weighted by any non-pretentious multiplicative function. Finally, in Section 8, we adapt the entropy decrement argument of [33] to the function field setting and apply the short exponential sum estimate for multiplicative functions from Section 7 to establish Theorem 1.4. Section 9 is then devoted to the proof of our application, Theorem 1.7, on Kátai's conjecture.

2 | NOTATION

Throughout the paper, p is the characteristic of $\mathbb{F}_q$, and $q = p^k$ for some $k \geqslant 1$.

We denote by $\mathcal{M}$ the set of monic polynomials in $\mathbb{F}_q[t]$ (we do not denote q dependence in $\mathcal{M}$, since it will always be clear from the context), and $\mathcal{P}$ the set of monic irreducible (prime) polynomials in $\mathbb{F}_q[t]$. For $N \in \mathbb{N}$, we write $\mathcal{M}_N$, $\mathcal{M}_{\leqslant N}$, and $\mathcal{M}_{<N}$ to denote, respectively, the set of monic polynomials of degree exactly N, less than or equal N, and strictly less than N. Analogously, we define $\mathcal{P}_N$, $\mathcal{P}_{\leqslant N}$, and $\mathcal{P}_{<N}$ to be the corresponding sets of monic irreducible polynomials. We denote the degree of $M \in \mathbb{F}_q[t]$ by $\deg(M)$.

Given two polynomials $F, G \in \mathcal{M}$, not both zero, we define their greatest common divisor (F, G) as the unique polynomial $D \in \mathcal{M}$ such that $D \mid F, D \mid G$ and such that for any $D' \in \mathcal{M}$ satisfying $D' \mid F, D' \mid G$ we have $D' \mid D$. The least common multiple $[F, G]$ of F and G is in turn defined by $[F, G] := FG/(F, G)$.

Typically, G will be used to denote an element of $\mathcal{M}$, whereas R or P denotes an element of $\mathcal{P}$ and M denotes an element of $\mathbb{F}_q[t]$, monic or otherwise.

Given two polynomials $G_0, G \in \mathcal{M}$ and a parameter $H \geqslant 1$, we write

$$I_H(G_0) := \{G \in \mathcal{M} : \deg(G - G_0) < H\}$$

to denote the short interval centered at G_0 of size H.

As usual, given $t \in \mathbb{R}$, we write $e(t) := e^{2\pi i t}$. Given a parameter $\theta \in [0, 1]$ and a polynomial $G \in \mathbb{F}_q[t]$, we also write $e_\theta(G) := e(\theta\deg(G))$. Finally, given an element $\alpha \in \mathbb{K}_\infty(t)$ (see Section 7) with formal Laurent series $\alpha = \sum_{k=N}^{\infty} a_{-k}(\alpha)t^{-k}$, we define $e_{\mathbb{F}}(\alpha) := e(\mathrm{tr}_{\mathbb{F}_q/\mathbb{F}_p} a_{-1}(\alpha)/p)$, where $\mathrm{tr}_{\mathbb{F}_q/\mathbb{F}_p}$ denotes the usual field trace. We also define $\langle\alpha\rangle := q^{-N}$.

Throughout the paper, we write $\mathbb{U} := \{z \in \mathbb{C} : |z| \leqslant 1\}$ and $S^1 := \{z \in \mathbb{U} : |z| = 1\}$. We say that $f : \mathcal{M} \to \mathbb{C}$ is multiplicative if $f(G_1G_2) = f(G_1)f(G_2)$ whenever G_1, G_2 are coprime. Given multiplicative functions $f, g : \mathcal{M} \to \mathbb{U}$, we define the pretentious distance between them by

$$\mathbb{D}(f, g; N) := \left(\sum_{P \in \mathcal{P}_{\leqslant N}} q^{-\deg(P)}(1 - \mathrm{Re}(f(P)\overline{g}(P))) \right)^{1/2}, \tag{11}$$

and define $\mathbb{D}(f, g; M, N)$ similarly, but with the summation being over $P \in \mathcal{P}_{\leqslant N} \setminus \mathcal{P}_{\leqslant M}$. We also set

$$\mathcal{D}_f(N) := \min_{\theta \in [0,1]} \mathbb{D}(f, e_\theta; N)^2.$$

Given a monic polynomial $G \in \mathcal{M}$ with $G(0) = 1$, we put $G^*(t) := t^{\deg(G)}G(1/t)$ (see Section 3.3 for further discussion). For a multiplicative function $f : \mathcal{M} \to \mathbb{U}$, we define the associated multiplicative function $f^* : \mathcal{M} \to \mathbb{U}$ as $f^*(G) := f(G^*)$, whenever $G(0) = 1$, and set $f^*(G) = 0$ otherwise.

Given a Dirichlet character χ modulo Q (defined above), we define its conductor[†] as $\mathrm{cond}(\chi) := \deg(Q')$ if $Q' \mid Q$ is such that $\chi(M)$ agrees with a primitive Dirichlet character $\chi' \bmod Q'$ for all M coprime to Q. In this case, we say that χ' induces χ. We write $\mathcal{X}_Q$ to denote the set of Dirichlet characters modulo Q.

A *Hayes character* is a character of the form $\tilde{\chi} = \psi_Q\xi_\nu$, where ψ_Q is a Dirichlet character to modulus Q induced by a primitive character to some modulus Q' and ξ_ν is a length ν short interval character for some $\nu \geqslant 0$ (in Section 3 we first give a different definition and then note that it is equivalent to this one). We define the conductor of $\tilde{\chi}$ by $\mathrm{cond}_H(\tilde{\chi}) := \deg(Q') + \nu$. We say that $\tilde{\chi}$ is *non-principal* if $\mathrm{cond}_H(\tilde{\chi}) > 0$. We further say that $\tilde{\chi}'$ *induces* $\tilde{\chi}$ if $\tilde{\chi}' = \chi'\xi'$ and $\tilde{\chi} = \chi\xi$, with the Dirichlet character χ' inducing χ and $\xi' = \xi$. We also write $\mathcal{X}_{Q,\nu}$ to denote the collection of Hayes characters of the form $\psi_Q\xi_\nu$, where ψ has modulus Q and ξ has length ν. See Section 3.4 for further discussion.

We will sometimes write μ_k to denote the set of roots of unity of order k, where $k \in \mathbb{N}$.

The functions Λ, ω, λ, μ, ϕ, rad, and ν_P, defined on $\mathcal{M}$, are the analogs of the corresponding arithmetic functions in the number field setting. Thus,

- $\Lambda(G) = \deg(P)$ if $G = P^k$ for some $k \geqslant 1$ and $P \in \mathcal{P}$, and $\Lambda(G) = 0$ otherwise.
- $\omega(G)$ is the number of distinct irreducible divisors of G.
- $\lambda : \mathcal{M} \to \{-1, +1\}$ is the completely multiplicative function with $\lambda(P) = -1$ for all $P \in \mathcal{P}$.
- $\mu : \mathcal{M} \to \{-1, 0, +1\}$ is given by $\mu(G) = (-1)^{\omega(G)}$ for G not divisible by P^2 for any $P \in \mathcal{P}$, and $\mu(G) = 0$ otherwise.
- $\phi(G)$ is the size of the finite multiplicative group $(\mathbb{F}_q[t]/G\mathbb{F}_q[t])^\times$.
- $\mathrm{rad}(G) = 1$ if $G = 1$ and $\mathrm{rad}(G) = P_1 \cdots P_k$ if $P_1, \dots, P_k$ are the distinct irreducible factors of G.
- $\nu_P(G)$, for $P \in \mathcal{P}$, is the largest integer k such that $P^k \mid G$.

[†] This is strictly speaking an abuse of notation/terminology, as the conductor of a function field Dirichlet character χ ought to be a polynomial of least degree that is a period for χ. Here, we use it as an integer-valued measure of complexity of the character, which will be convenient for us in various estimates in the sequel.

Throughout this paper, the cardinality q of the underlying finite field $\mathbb{F}_q$ is fixed. For the sake of convenience, we have chosen to omit mention of dependencies of implicit constants in our estimates on q. In particular, the implicit constants in any estimate may depend on q throughout this paper.

3 | PRELIMINARIES I: MULTIPLICATIVE FUNCTIONS AND HAYES CHARACTERS

In this section, we establish some auxiliary lemmas, specifically related to multiplicative functions, that will be necessary in the proofs of Theorems 1.1 and 1.4. Recall the definition of Hayes characters from Section 2.

3.1 | Lemmas on character sums

When working over $\mathbb{F}_q[t]$, we have the generalized Riemann hypothesis at our disposal, arising from an application of Weil's Riemann hypothesis for curves over finite fields (see [39, p. 134]).[†]

Lemma 3.1 (Rhin). *Let $N \geqslant 1$. Let $\tilde{\chi}$ be a non-principal Hayes character. Then,*

$$\sum_{G \in \mathcal{M}_N} \tilde{\chi}(G)\Lambda(G) \ll \mathrm{cond}_H(\tilde{\chi}) q^{N/2}. \tag{12}$$

Proof. This is [30, Theorem 3]. □

A useful corollary of Lemma 3.1 is the following.

Lemma 3.2 (A pretentious distance bound). *Let $N \geqslant 3$, $A \geqslant 1$. Let $\tilde{\chi}$ be a non-principal Hayes character of conductor $\mathrm{cond}_H(\tilde{\chi}) \leqslant N^A$. Then,*

$$\max_{\theta \in [0,1]} \left| \sum_{P \in \mathcal{P}_{\leqslant N}} \tilde{\chi}(P) e_\theta(P) q^{-\deg(P)} \right| \ll_A \log\log N.$$

Proof. Splitting the sum according to degree, then separating the contribution of $\deg(P) \leqslant 10A \log N$ from its complement, we get

$$\begin{aligned} &\sum_{P \in \mathcal{P}_{\leqslant N}} \tilde{\chi}(P) e_\theta(P) q^{-\deg(P)} \\ &= \sum_{d \leqslant 10A \log N} e(\theta d) q^{-d} \sum_{P \in \mathcal{P}_d} \tilde{\chi}(P) + \sum_{10A \log N < d \leqslant N} e(\theta d) q^{-d} \sum_{P \in \mathcal{P}_d} \tilde{\chi}(P) =: T_1 + T_2. \end{aligned}$$

[†] Even though GRH is useful for us in certain parts of our arguments, specifically Lemmas 3.2 and 3.3, we emphasize that it is not the main driving force behind the proofs of our results. As noted implicitly in [27] and explicitly in [23], obtaining non-trivial bounds on the variance in short intervals and arithmetic progressions, respectively, in the integer setting only requires the existence of sufficiently wide zero-free regions, for example, of Korobov–Vinogradov type, for Dirichlet L-functions to the left of $\mathrm{Re}(s) = 1$ (and for L-functions of Hayes characters in our setting).

We bound the first sum trivially using the prime polynomial theorem, yielding

$$T_1 \ll \sum_{d \leqslant 10A \log N} q^{-d} |\mathcal{P}_d| \ll \sum_{d \leqslant 10A \log N} \frac{1}{d} = \log\log N + O_A(1).$$

We now consider T_2. Replacing $\sum_{P \in \mathcal{P}_d} \tilde{\chi}(P)$ by $\sum_{G \in \mathcal{M}_d} \tilde{\chi}(G)\Lambda(G)/d$ in the inner sum over primes in T_2 incurs an error of size $O(\sum_{d \leqslant N} q^{-d/2}) = O(1)$ from terms P^k with $k \geqslant 2$. This sum can thus be expressed as

$$T_2 = \sum_{10A \log N < d \leqslant N} \frac{e(\theta d)}{dq^d} \sum_{G \in \mathcal{M}_d} \Lambda(G)\tilde{\chi}(G) + O(1).$$

By Lemma 3.1, we can bound this as

$$\begin{aligned} |T_2| &\leqslant \sum_{10A \log N < d \leqslant N} \frac{1}{dq^d} \left| \sum_{G \in \mathcal{M}_d} \Lambda(G)\tilde{\chi}(G) \right| + O(1) \ll \sum_{10A \log N < d \leqslant N} \frac{N^A}{dq^{d/2}} + O(1) \\ &\ll N^A \cdot 2^{-5A \log N} + 1 \ll 1. \end{aligned}$$

Combining the contributions from T_1 and T_2, we obtain the claim. □

We will also need a bound on sums of Hayes characters over $\mathcal{M}$ (as opposed to $\mathcal{P}$).

Lemma 3.3 (Pointwise bound for character sums over monics). *Let $M > N \geqslant 1$. Let $\tilde{\chi}$ be either a non-principal Dirichlet character or a non-principal short interval character of conductor M. Then, we have*

$$\sum_{G \in \mathcal{M}_N} \tilde{\chi}(G) \ll q^{N/2} \binom{M-1}{N}. \tag{13}$$

Remark 3.4. This lemma will be applied in particular when $M \leqslant (1 + o(1))N$. For Dirichlet characters, we could instead have appealed to the Pólya–Vinogradov inequality (see [15, Proposition 2.1]) to produce a sharper bound in this range, rather than applying Weil's RH; however, a corresponding result for general Hayes characters does not exist in the literature. For this reason, we have resorted to appealing to RH instead.

Note that the same quality bound, with an essentially identical proof, appears as [6, Lemma 2.1] in the context of Dirichlet characters; for the sake of completeness, we include the short proof of the general case.

Proof. By the GRH for L-functions corresponding to Hayes characters [30], we can write the L-function

$$\mathcal{L}(z, \tilde{\chi}) = \sum_{G \in \mathcal{M}} \tilde{\chi}(G) z^{\deg(G)}$$

as

$$\mathcal{L}(z, \tilde{\chi}) = \prod_{j=1}^{M-1} (1 - \alpha_i z) \tag{14}$$

for some $\alpha_i = \alpha_i(\tilde{\chi})$ that all have modulus either 1 or $q^{1/2}$. Now, the sum in question is the coefficient of z^N on the right of (14), which by Vieta's formulas is equal to

$$\sum_{\substack{S\subseteq[1,M-1]\cap\mathbb{N}\\|S|=N}}\prod_{j\in S}(-\alpha_j).$$

This is trivially bounded in absolute value by $q^{N/2}\binom{M-1}{N}$, which yields the claim. □

3.2 | Multiplicative functions in function fields

Let $f:\mathcal{M}\to\mathbb{U}$ be a one-bounded multiplicative function. Define the Dirichlet series corresponding to f by

$$L(s,f) := \sum_{N\geqslant 0}\sum_{G\in\mathcal{M}_N} f(G)q^{-\deg(G)s} = \prod_{P\in\mathcal{P}}\sum_{k\geqslant 0} f(P^k)q^{-k\deg(P)s}, \tag{15}$$

for $\mathrm{Re}(s)>1$; in this region, both expressions converge absolutely.

Recall the pretentious distance

$$\mathbb{D}(f,g;N) := \left(\sum_{P\in\mathcal{P}_{\leqslant N}} q^{-\deg(P)}\left(1-\mathrm{Re}(f(P)\overline{g(P)})\right)\right)^{\frac{1}{2}}.$$

One can show [21] that $\mathbb{D}$ satisfies a triangle inequality of the shape

$$\mathbb{D}(f,h;N)\leqslant \mathbb{D}(f,g;N)+\mathbb{D}(g,h;N),$$

for any $f,g,h:\mathcal{M}\to\mathbb{U}$ multiplicative. Define also

$$\mathcal{D}_f(N) := \min_{\theta\in[0,1]}\mathbb{D}(f,e_\theta;N)^2.$$

The following variant of Halász's theorem then holds:

Theorem 3.5 (Halász's Theorem in Function Fields). *Let $N\geqslant 1$. Let $f:\mathcal{M}\to\mathbb{U}$ be multiplicative. Then,*

$$\frac{1}{q^N}\sum_{G\in\mathcal{M}_N} f(G)\ll(1+\mathcal{D}_f(N))e^{-\mathcal{D}_f(N)}.$$

Proof. We will reduce this to the Granville–Harper–Soundararajan formulation of Halász's inequality in [9]. Define the multiplicative function $\tilde{f}_N$ on prime powers by[†]

[†] This is technically different from the definition of $f^{\perp}$ used in [9]. However, it is always true that, in the notation there, $\Lambda_{\tilde{f}_N}(P)=\Lambda_{f^\perp}(P)$, and the difference lies only in values at powers P^k with $k\geqslant 2$. It is easy to check, then, that $|L(s,f^\perp)|$ and $|L(s,\tilde{f}_N)|$ differ in at most a factor of an absolute constant whenever f is one-bounded and $\mathrm{Re}(s)=1$.

$$\tilde{f}_N(P^k) := \begin{cases} f(P^k) & \text{if } \deg(P^k) \leqslant N \\ 0 & \text{otherwise.} \end{cases}$$

Then, [9, Corollary 1.2] (in the case $\varkappa = 1$) shows that

$$\frac{1}{q^N} \sum_{G \in \mathcal{M}_N} f(G) \ll (1+M)e^{-M},$$

where $M := \min_{\mathrm{Re}(s)=1} \log(2N/|L(s, \tilde{f}_N)|)$. Now, the prime polynomial theorem gives

$$\sum_{P \in \mathcal{P}_{\leqslant N}} q^{-\deg(P)} = \sum_{d \leqslant N} q^{-d} |\mathcal{P}_d| = \sum_{d \leqslant N} q^{-d} \left(\frac{q^d}{d} + O(q^{d/2}) \right)$$

$$= \sum_{d \leqslant N} \frac{1}{d} + O\left(\sum_{d \leqslant N} q^{-d/2} \right) = \log N + O(1).$$

Moreover, if s_0 maximizes $|L(s, \tilde{f}_N)|$ on $\mathrm{Re}(s) = 1$ and $q^{-s_0} = e(\theta)/q$ for some $\theta \in [0,1]$, then

$$\log |L(s_0, \tilde{f}_N)| = \log \left| \prod_{P \in \mathcal{P}_{\leqslant N}} \left(1 + f(P)e(\theta \deg(P)) q^{-\deg(P)} + O\left(\sum_{k \geqslant 2} q^{-k\deg(P)} \right) \right) \right|$$

$$= \sum_{P \in \mathcal{P}_{\leqslant N}} \mathrm{Re}(f(P)e_\theta(P)) q^{-\deg(P)} + O(1).$$

It follows that

$$M = \min_{\mathrm{Re}(s)=1} \log(2N/|L(s, \tilde{f}_N)|) = \min_{\theta \in [0,1]} \sum_{P \in \mathcal{P}_{\leqslant N}} q^{-\deg(P)} (1 - \mathrm{Re}(f(P)e_\theta(P))) + O(1)$$

$$= \mathcal{D}_f(N) + O(1).$$

The claim follows immediately. □

Using Lemma 3.1, we can also show that for any $N \geqslant 3$, there is at most one Hayes character $\tilde{\chi}$ with $\mathrm{cond}_H(\tilde{\chi}) \leqslant N$ for which $\mathcal{D}_{f\overline{\tilde{\chi}}}(N)$ can be "small" in some sense. In what follows, we denote $\tilde{\chi}_1 \sim \tilde{\chi}_2$ if $\tilde{\chi}_1$ and $\tilde{\chi}_2$ are induced by the same Hayes character, and otherwise write $\tilde{\chi}_1 \nsim \tilde{\chi}_2$.

Lemma 3.6 (Repulsion of pretentious distance). *Let $N \geqslant 3$. Let $f : \mathcal{M} \to \mathbb{U}$ be multiplicative. Let $\tilde{\chi}_1 \nsim \tilde{\chi}_2$ be two Hayes characters of conductors $\leqslant N$. Then,*

$$\max\{\mathcal{D}_{f\overline{\tilde{\chi}}_1}(N), \mathcal{D}_{f\overline{\tilde{\chi}}_2}(N)\} \geqslant \left(\frac{1}{4} - o(1) \right) \log N.$$

Proof. For each $j = 1, 2$, let θ_j be an angle for which $\mathcal{D}_{f\overline{\tilde{\chi}}_j}(N) = \mathbb{D}(f, \tilde{\chi}_j e_{\theta_j}; N)^2$. Suppose first that f is unimodular. Then, by the triangle inequality, we have

$$2\max\{\mathcal{D}_{f\overline{\tilde{\chi}}_1}(N)^{1/2}, \mathcal{D}_{f\overline{\tilde{\chi}}_2}(N)^{1/2}\} \geqslant \mathbb{D}(f\overline{\tilde{\chi}}_1, e_{\theta_1}; N) + \mathbb{D}(f\overline{\tilde{\chi}}_2, e_{\theta_2}; N) \geqslant \mathbb{D}(\tilde{\chi}_1, \tilde{\chi}_2 e_{\theta_2 - \theta_1}; N), \quad (16)$$

where the unimodularity of f was used in the second inequality to write $(f\overline{\tilde{\chi}_1})\overline{(f\overline{\tilde{\chi}_2})} = \overline{\tilde{\chi}_1}\tilde{\chi}_2$. Now, by definition we have

$$\mathbb{D}(\tilde{\chi}_1, \tilde{\chi}_2 e_{\theta_2-\theta_1}; N)^2 = \log N - \mathrm{Re}\left(\sum_{P\in\mathcal{P}_{\leqslant N}} \tilde{\chi}_1\overline{\tilde{\chi}_2}(P)e((\theta_1-\theta_2)\deg(P))q^{-\deg(P)}\right) + O(1).$$

Since $\tilde{\chi}_1\overline{\tilde{\chi}_2}$ has conductor $\leqslant N^2$ and it is non-principal, Lemma 3.2 (with $\theta := \theta_1 - \theta_2$ and $\tilde{\chi} := \tilde{\chi}_1\overline{\tilde{\chi}_2}$) yields

$$\mathrm{Re}\left(\sum_{P\in\mathcal{P}_{\leqslant N}} \tilde{\chi}_1\overline{\tilde{\chi}_2}(P)e((\theta_1-\theta_2)\deg(P))q^{-\deg(P)}\right) \ll \log\log N,$$

and so it follows that

$$\mathbb{D}(\tilde{\chi}_1, \tilde{\chi}_2 e_{\theta_2-\theta_1}; N)^2 \geqslant \log N - O(\log\log N). \tag{17}$$

Squaring both sides of (16), then inserting this last estimate into the result yields

$$\max\{\mathcal{D}_{f\overline{\tilde{\chi}}_1}(N), \mathcal{D}_{f\overline{\tilde{\chi}}_2}(N)\} \geqslant \left(\frac{1}{4} - o(1)\right)\log N.$$

Suppose then that f is not unimodular. Define a random completely multiplicative function $\boldsymbol{f}: \mathcal{M} \to S^1$ (on some associated probability space) at irreducibles P in such a way that $f(P) = \mathbb{E}\boldsymbol{f}(P)$ for every irreducible P. By linearity of expectation, it follows that for any multiplicative function g, we have

$$\mathbb{D}(f, g; N)^2 = \mathbb{E}\mathbb{D}(\boldsymbol{f}, g; N)^2. \tag{18}$$

It follows from this and (17) that for any $\theta \in [0,1]$, we have

$$\begin{aligned}
2\max\{\mathcal{D}_{f\overline{\tilde{\chi}}_1}(N), \mathcal{D}_{f\overline{\tilde{\chi}}_2}(N)\} &\geqslant \mathbb{D}(f\overline{\tilde{\chi}}_1, e_{\theta_1}; N)^2 + \mathbb{D}(f\overline{\tilde{\chi}}_2, e_{\theta_2}; N)^2 \\
&\geqslant \frac{1}{2}(\mathbb{D}(f\overline{\tilde{\chi}}_1, e_{\theta_1}; N) + \mathbb{D}(f\overline{\tilde{\chi}}_2, e_{\theta_2}; N))^2 \\
&= \frac{1}{2}(\mathbb{E}(\mathbb{D}(\boldsymbol{f}\overline{\tilde{\chi}}_1, e_{\theta_1}; N) + \mathbb{D}(\boldsymbol{f}\overline{\tilde{\chi}}_2, e_{\theta_2}; N)))^2 \\
&\geqslant \left(\frac{1}{2} - o(1)\right)\log N,
\end{aligned}$$

and the claim follows. □

Combining Theorem 3.5 with Lemma 3.6 immediately produces the following.

Corollary 3.7 (Sup norm estimate for weighted character sums). *Let $N \geqslant 3$. Let $f: \mathcal{M} \to \mathbb{U}$ be multiplicative. Let $\tilde{\chi}_1$ be the Hayes character of conductor $\leqslant N$ that minimizes*[†] *the map*

[†] If there are several minimizers, we choose one of them arbitrarily.

$\tilde{\chi} \mapsto \mathcal{D}_{f\overline{\tilde{\chi}}}(N)$. *Then,*

$$\max_{\substack{\mathrm{cond}_H(\tilde{\chi})\leqslant N \\ \tilde{\chi}\sim\tilde{\chi}_1}} \left| \frac{1}{q^N} \sum_{G\in\mathcal{M}_N} f(G)\overline{\tilde{\chi}}(G) \right| \ll N^{-1/4+o(1)}. \tag{19}$$

Finally, we will need the following simple upper bound estimate for non-negative multiplicative functions later in this paper (for a corresponding result about functions over the integers, see [10]).

Lemma 3.8 (Halberstam–Richert bound in function fields). *Let $g : \mathcal{M} \to [0,\infty)$ be multiplicative, and let $N \geqslant 1$. Let $\kappa > 0$, and assume that for all $P \in \mathcal{P}$ and $k \geqslant 1$ we have $g(P) \leqslant \kappa$ and $g(P^k) \ll_\varepsilon q^{k\varepsilon\deg(P)}$ for any $\varepsilon > 0$. Then,*

$$\frac{1}{|\mathcal{M}_N|} \sum_{G\in\mathcal{M}_N} g(G) \ll \frac{(\kappa+1)}{N} \exp\left(\sum_{P\in\mathcal{P}_{\leqslant N}} g(P) q^{-\deg(P)} \right).$$

Proof. Observe that for any $G \in \mathcal{M}_N$ we have $N = \sum_{\substack{P^k \,||\, G \\ P\in\mathcal{P}}} k\deg(P)$ (where $P^k \,||\, B$ means $P^k \mid B$ and $P^{k+1} \nmid B$), and thus

$$\begin{aligned}
\sum_{G\in\mathcal{M}_N} g(G) &= \frac{1}{N} \sum_{\substack{P^k B\in\mathcal{M}_N \\ (P,B)=1 \\ P\in\mathcal{P}}} g(P^k) g(B) k\deg(P) \\
&\leqslant \frac{1}{N} \sum_{B\in\mathcal{M}_{\leqslant N}} g(B) \sum_{P\in\mathcal{P}_{N-\deg(B)}} g(P)\deg(P) + \frac{1}{N} \sum_{\substack{P^k B\in\mathcal{M}_N \\ P\in\mathcal{P} \\ k\geqslant 2}} g(P^k) g(B) k\deg(P) \\
&=: \mathfrak{S}_1 + \mathfrak{S}_2.
\end{aligned}$$

Consider $\mathfrak{S}_1$ first. Bounding $g(P) \leqslant \kappa$ for each $P \in \mathcal{P}_{N-\deg(B)}$ and then using the prime polynomial theorem, we have

$$\sum_{P\in\mathcal{P}_{N-\deg(B)}} g(P)\deg(P) \leqslant \kappa \sum_{G\in\mathcal{M}_{N-\deg(B)}} \Lambda(G) \ll \kappa q^{N-\deg(B)},$$

for every $B \in \mathcal{M}_{\leqslant N}$. Summing over such B now gives

$$\mathfrak{S}_1 \ll \kappa \frac{q^N}{N} \sum_{B\in\mathcal{M}_{\leqslant N}} g(B) q^{-\deg(B)} \leqslant \kappa \frac{q^N}{N} \prod_{P\in\mathcal{P}_{\leqslant N}} \left(\sum_{k\geqslant 0} g(P^k) q^{-k\deg(P)} \right).$$

Using the condition $g(P^k) \ll q^{\frac{1}{4}k\deg(P)}$ for $k \geqslant 2$, we get

$$\sum_{P\in\mathcal{P}_{\leqslant N}} \sum_{k\geqslant 2} g(P^k) q^{-k\deg(P)} \ll \sum_{P\in\mathcal{P}_{\leqslant N}} q^{-\frac{3}{2}\deg(P)} \ll \sum_{d\leqslant N} q^{-d/2} \ll 1.$$

Thus, rewriting the product over $P \in \mathcal{P}_{\leqslant N}$ as an exponential, we get

$$\prod_{P\in\mathcal{P}_{\leqslant N}}\left(1+g(P)q^{-\deg(P)}+\sum_{k\geqslant 2}g(P^k)q^{-k\deg(P)}\right)$$

$$\leqslant \prod_{P\in\mathcal{P}_{\leqslant N}}\left(1+g(P)q^{-\deg(P)}\right)\left(1+\sum_{k\geqslant 2}g(P^k)q^{-k\deg(P)}\right)\ll \exp\left(\sum_{P\in\mathcal{P}_{\leqslant N}}g(P)q^{-\deg(P)}\right). \tag{20}$$

Inserting this into our bound for $\mathfrak{S}_1$ yields

$$\mathfrak{S}_1 \ll \kappa\frac{q^N}{N}\exp\left(\sum_{P\in\mathcal{P}_{\leqslant N}}g(P)q^{-\deg(P)}\right).$$

To bound $\mathfrak{S}_2$, we use the identity $1 = q^N/q^{\deg(B)+k\deg(P)}$ and the upper bound $k\deg(P)g(P^k) \ll q^{k\deg(P)/3}$ to get

$$\mathfrak{S}_2 = \frac{q^N}{N}\sum_{B\in\mathcal{M}_{\leqslant N}}g(B)q^{-\deg(B)}\sum_{\substack{P^k\in\mathcal{P}_{N-\deg(B)}\\ k\geqslant 2}}k\deg(P)g(P^k)q^{-k\deg(P)}$$

$$\ll \frac{q^N}{N}\left(\prod_{P_1\in\mathcal{P}_{\leqslant N}}\sum_{\ell\geqslant 0}g(P_1^\ell)q^{-\ell\deg(P_1)}\right)\sum_{k\geqslant 2}\sum_{P_2\in\mathcal{P}}q^{-2k\deg(P_2)/3}.$$

The sum over P_2 can be bounded by

$$\sum_{d\geqslant 1}|\mathcal{P}_d|\sum_{k\geqslant 2}q^{-2kd/3} \ll \sum_{d\geqslant 1}q^{-d/3} \ll 1.$$

Bounding the product in P_1 as in (20), we obtain

$$\mathfrak{S}_2 \ll \frac{q^N}{N}\exp\left(\sum_{P\in\mathcal{P}_{\leqslant N}}g(P)q^{-\deg(P)}\right).$$

Combining this with the bound for $\mathfrak{S}_1$ proves the claim. □

3.3 | An involution for monic polynomials

Let $G \in \mathcal{M}$, and assume that $(G, t) = 1$. Following Keating and Rudnick (see [18, section 5]), we define[†]

$$G^*(t) := t^{\deg(G)}G(1/t).$$

[†] We could extend this definition to other polynomials by writing $G^*(t) = t^{\nu(G)}(G/t^{\nu(G)})^*$, where $\nu(G)$ denotes the order of vanishing of G at $t = 0$. We could also modify the definition here when $G(0) \neq 0$ to give $G^* = G(0)^{-1}t^{\deg(G)}G(1/t)$, thus ensuring that G^* is monic whenever G is; however, we will not need this variant of the involution here.

The coefficients of G^* are the same as those of G, but in reverse order. One can easily check that when $(G,t)=1$ and $G(0)=1$, G^* is monic and $(G^*)^*=G$. Since $\deg(G^*)=\deg(G)$, the $*$-map is an involution on the set of monic degree N polynomials with $G(0)=1$, for each $N\geqslant 1$.

We observe, furthermore, that this involution is a multiplicative homomorphism on $\mathcal{M}$. Indeed, if $(FG,t)=1$, then

$$(FG)^*(t)=t^{\deg(FG)}FG(1/t)=t^{\deg(F)}F(1/t)\cdot t^{\deg(G)}G(1/t)=F^*(t)G^*(t).$$

In light of this, we can define a corresponding involution on the space of multiplicative functions. That is, suppose that $f:\mathcal{M}\to\mathbb{U}$ is multiplicative. We define a map $f\mapsto f^*$ via $f^*(G):=f(G^*)$ for all $(G,t)=1$ with $G(0)=1$, and $f^*(t^k)=0$ for $k\geqslant 1$. Under a suitable extension of f to $\mathbb{F}_q[t]$ (which we are free to choose, given that f is only defined on $\mathcal{M}$ by assumption), we may define $f^*(G)$ at all monic G irrespective of the condition $G(0)=1$. Then, f^* acts as a multiplicative function on $\mathcal{M}$, and if $g:\mathcal{M}\to\mathbb{U}$ is a second such multiplicative function, then $(fg)^*=f^*g^*$.

The next result, which is essentially contained in [18], shows that the $*$-operation maps short intervals to arithmetic progressions modulo a power of t.

Lemma 3.9. *Let $1\leqslant H\leqslant N$ and $G_0\in\mathcal{M}_N$. There is a reduced residue class A modulo t^{N-H+1} for which we have a bijection*

$$\{G\in\mathcal{M}_N:\ G\in I_H(G_0),(G,t)=1\}\leftrightarrow\{\deg(F)=N:\ F\equiv A \bmod t^{N-H+1},F(0)=1\};$$

the bijection is furnished by the map $G\mapsto G^*$. Moreover, the class $A=A(G_0)$ depends at most on the first $N-H$ coefficients of G_0 after the leading coefficient.

Proof. This is implied by [18, Lemma 5.1], using the fact that $I_H(G_0)=I_H(t^HG_0')$ whenever $\deg(G_0-t^HG_0')<H$. □

The following lemma shows how the pretentious distance is affected by replacing a multiplicative function f (whose behavior on $\mathbb{F}_q^\times$ is fixed) by its involution f^*. In the following, we fix a generator ρ for $\mathbb{F}_q^\times$ and write ν_c to be the minimal non-negative integer such that $\rho^{\nu_c}=c$ whenever $c\in\mathbb{F}_q^\times$.

Lemma 3.10. *Let $\zeta\in\mu_{q-1}$ and let $f:\mathbb{F}_q[t]\to\mathbb{U}$ be a multiplicative function. Extend f to $\mathbb{F}_q[t]$ so that $f(cF)=\zeta^{\nu_c}f(F)$ for all $c\in\mathbb{F}_q^\times$. Let χ be a Dirichlet character modulo t^M, for $M\geqslant 1$. Then, there is a character $\xi=\xi(\zeta,\chi)$ modulo t such that for any $N\geqslant 1$ we have*

$$\mathcal{D}_{f\xi\overline{\chi^*}}(N)=\mathcal{D}_{f^*\overline{\chi}}(N)+O(1).$$

Moreover, if $\zeta=\chi^*(\rho)$ then $\xi\equiv 1$.

Remark 3.11. Note that even though we are only concerned with the values of f on $\mathcal{M}$, in order to define f^* we need to choose an extension of f to $\mathbb{F}_q^\times$.

Proof. We claim first that there is a unique character ξ modulo t such that $\overline{\chi}^*(c)\zeta^{\nu_c}\xi(c)=1$. To see this, note first that $\chi^*(1)=\chi(1)=1$, so that $\chi^*(\rho)\in\mu_{q-1}$. The group of characters mod t may be

identified with that of $\mathbb{F}_q^\times$ via the isomorphism $(\mathbb{F}_q[t]/(t\mathbb{F}_q[t]))^\times \cong \mathbb{F}_q^\times$, so there is a ξ modulo t such that $\xi(\rho) = \overline{\zeta}\chi^*(\rho)$. Extending by complete multiplicativity, we obtain $\xi(c) = \chi^*(c)\zeta^{-\nu_c}$ for all $c \in \mathbb{F}_q^\times$. Moreover, if there is a second such character ξ' modulo t, then we must have $\xi'(\rho) = \xi(\rho)$, and thus $\xi' = \xi$, as required.

We select ξ to be the character modulo t determined above. Let $\theta_0 \in [0,1]$ be fixed. We will show that $\mathbb{D}(f^*, \chi e_{\theta_0}; N) = \mathbb{D}(f\xi, \chi^* e_{\theta_0}; N) + O(1)$. By minimizing over θ_0, we deduce the claimed estimate.

First, note that if $R \in \mathcal{P}, R \neq t$, then $R^*/R(0) \in \mathcal{P}$. For if $R^* = AB$ with $\deg(A)\deg(B) > 0$, then as $(R^*, t) = 1$ we have $R = A^*B^*$, with $\deg(A^*)\deg(B^*) = \deg(A)\deg(B) > 0$, a contradiction to irreduciblity. In particular, for each $c \in \mathbb{F}_q^\times$ and $d \geqslant 2$, we have a bijection

$$\{R \in \mathcal{P}_d : R(0) = c\} \leftrightarrow \{R' \in \mathcal{P}_d : R'(0) = c^{-1}\},$$

implied by the map $R \mapsto R' := R^*/R(0)$. Thus, we have

$$\begin{aligned}\mathbb{D}(f^*, \chi e_{\theta_0}; N)^2 &= \log N - \mathrm{Re}\left(\sum_{2\leqslant d\leqslant N} q^{-d} e(-\theta_0 d) \sum_{c\in\mathbb{F}_q^\times} \sum_{\substack{R\in\mathcal{P}_d\\ R(0)=c}} f^*(R)\overline{\chi}(R)\right) + O(1)\\ &= \log N - \mathrm{Re}\left(\sum_{2\leqslant d\leqslant N} q^{-d} e(-\theta_0 d) \sum_{c\in\mathbb{F}_q^\times} \sum_{\substack{R'\in\mathcal{P}_d\\ R'(0)=c^{-1}}} f(cR')\overline{\chi}^*(cR')\right) + O(1).\end{aligned}$$

Since $R'(0) = c^{-1}$ iff $R' \equiv c^{-1} \pmod t$, we get

$$1_{R'(0)=c^{-1}} = \frac{1}{\phi(t)} \sum_{\xi' \pmod t} \xi'(c)\xi'(R'),$$

and thus for each $2 \leqslant d \leqslant N$ we obtain

$$\begin{aligned}\sum_{c\in\mathbb{F}_q^\times} \sum_{\substack{R'\in\mathcal{P}_d\\ R'(0)=c^{-1}}} f(cR')\overline{\chi}^*(cR') &= \frac{1}{\phi(t)} \sum_{\xi' \pmod t} \left(\sum_{c\in\mathbb{F}_q^\times} \overline{\chi}^*(c)\zeta^{\nu_c}\xi'(c)\right) \sum_{R'\in\mathcal{P}_d} f(R')\xi'(R')\overline{\chi}^*(R')\\ &= \sum_{R'\in\mathcal{P}_d} f(R')\xi(R')\overline{\chi}^*(R'),\end{aligned}$$

where ξ is the character modulo t constructed earlier. It follows then that

$$\begin{aligned}\mathbb{D}(f^*, \chi e_{\theta_0}; N)^2 &= \log N - \mathrm{Re}\left(\sum_{2\leqslant d\leqslant N} q^{-d} e(-\theta_0 d) \sum_{R'\in\mathcal{P}_d} f(R')\xi(R')\overline{\chi}^*(R')\right) + O(1)\\ &= \mathbb{D}(f\xi, \chi^* e_{\theta_0}; N)^2 + O(1),\end{aligned}$$

proving the first claim.

For the second, note that if $\chi^*(\rho) = \zeta$, then by construction we have $\xi(\rho) = 1$, and thus ξ is trivial modulo t, as required. □

3.4 | Hayes characters

We introduce here the following notation. Let $F, G \in \mathbb{F}_q[t]$ with $G \neq 0$, and consider $F/G \in \mathbb{F}_q(t)$. When G is a power of t this rational function admits a finite Laurent polynomial representation (in $1/t$)

$$(F/G)(t) = \sum_{j=m_1}^{m_2} a_j t^{-j},$$

where $m_1 \leqslant m_2$ are integers and $a_{m_1} \neq 0$. We then set $\langle F/G\rangle := q^{-m_1}$. We note that the map $\langle\cdot\rangle$ satisfies the ultrametric inequality $\langle f_1 - f_2\rangle \leqslant \max\{\langle f_1\rangle, \langle f_2\rangle\}$, with equality if $\langle f_1\rangle \neq \langle f_2\rangle$, whenever $f_1, f_2 \in \mathbb{F}_q(t)$ have finite Laurent polynomial representations (in Section 7 we will extend this notation to all of $\mathbb{F}_q(t)$).

Let $\nu \geqslant 1$ and $M \in \mathcal{M}$. We define a relation $\mathcal{R}_{M,\nu}$ on $\mathcal{M}$ as follows: If $A, B \in \mathcal{M}$, then we say that

$$A \equiv B \bmod \mathcal{R}_{M,\nu} \text{ if, and only if, } A \equiv B \bmod M \text{ and } \left\langle At^{-\deg(A)} - Bt^{-\deg(B)}\right\rangle < q^{-\nu}.$$

This latter condition says that the leading $\nu + 1$ coefficients of A and B are the same; In the particular case where $A, B \in \mathcal{M}_N$ for some N, it is equivalent to $\deg(A - B) < N - \nu$.

It turns out that this defines an equivalence relation, and quotienting $\mathcal{M}$ by this relation yields a monoid whose multiplicative group of invertible elements is abelian. It thus admits a set of characters, which we call *Hayes characters*. We will denote by $\mathcal{X}_{M,\nu}$ the collection of all Hayes characters associated with the pair (M, ν). A Hayes character $\tilde{\chi}$ is characterized by the property that it is constant on sets of the form

$$\{G \in \mathcal{M} : G \equiv C \bmod M\} \cap \{G \in \mathcal{M} : \left\langle Gt^{-\deg(G)} - Dt^{-\deg(D)}\right\rangle < q^{-\nu}\},$$

where C is a reduced residue class modulo M, and $D \in \mathcal{M}_{\leqslant\nu}$. Any Hayes character in $\mathcal{X}_{M,\nu}$ can be uniquely decomposed as a product $\psi_M\xi_\nu$, where ψ_M is a Dirichlet character modulo M, and ξ_ν is a *short interval* character of *length* $\mathrm{len}(\xi_\nu) := \nu$, that is, for $\ell = \nu$ the multiplicative function ξ_ν fixes the set $\{G \in \mathcal{M}_N : \langle Gt^{-\deg(G)} - Dt^{-\deg(D)}\rangle < q^{-\ell}\}$ for all D, and the same does not hold for any $\ell < \nu$ (see, e.g., [12, Theorem 8.6]). Thus, this definition agrees with Definition 1.3. We say that $\tilde{\chi} \in \mathcal{X}_{M,\nu}$ is *primitive* if ψ_M is primitive and $\nu > 0$, and imprimitive otherwise. Likewise, a Hayes character is *non-principal* if it is either non-principal in the Dirichlet character aspect or if the length of its short interval character is non-zero. We define the *Hayes conductor* of $\chi = \psi\xi \in \mathcal{X}_{M,\nu}$ by $\mathrm{cond}_H(\chi) := \mathrm{cond}(\psi) + \mathrm{len}(\xi) := \deg(M) + \nu$.

The group $\mathcal{X}_{M,\nu}$ has size $\phi(M)q^\nu$, and the orthogonality relations are given by

$$\frac{1}{\phi(M)q^\nu} \sum_{A \bmod R_{M,\nu}} \widetilde{\chi_1}(A)\overline{\widetilde{\chi_2}(A)} = 1_{\widetilde{\chi_1}=\widetilde{\chi_2}} \tag{21}$$

and

$$\frac{1}{\phi(M)q^{\nu}}\sum_{\widetilde{\chi}\in\mathcal{X}_{M,\nu}}\widetilde{\chi}(A)\overline{\widetilde{\chi}(B)}=1_{A\equiv B \bmod R_{M,\nu}};\tag{22}$$

these are proved in [12].

An important fact about the relationship between Hayes characters and the $*$-involution from the previous subsection is the following.

Lemma 3.12. *Let $n\geqslant 2$ and $k\geqslant 2$. Let χ be a Dirichlet character modulo t^k. Then, there is a short interval character ψ of length $k-1$ such that $\chi^*(G)=\psi(G)$ for all G coprime to t. Moreover, if χ is non-principal, then ψ is also non-principal.*

Proof. It is enough to show that if $G_1,G_2\in\mathbb{F}_q[t]$ satisfy $(G_1G_2,t)=1$ and are close to each other in the sense that $\langle G_1t^{-\deg(G_1)}-G_2t^{-\deg(G_2)}\rangle\leqslant q^{-k}$, then $\chi^*(G_1)=\chi^*(G_2)$.

Without loss of generality suppose that $m_1:=\deg(G_1)\geqslant\deg(G_2)=:m_2$. Then, we can write $G_1=t^{m_1-m_2}G_2+M$, where $r:=\deg(M)\leqslant m_1-k$. Writing $G_2(t)=\sum_{0\leqslant j\leqslant m_2}b_jt^j$ and $M(t)=\sum_{0\leqslant j\leqslant r}a_jt^j$ (with $a_0b_{m_2}\neq 0$ by assumption) we find

$$\begin{aligned}G_1^*&=\left(\sum_{m_1-m_2\leqslant j\leqslant m_1}b_{m_2-(m_1-j)}t^j+\sum_{0\leqslant j\leqslant r}a_jt^j\right)^*\\&=t^{m_1}\left(\sum_{m_1-m_2\leqslant j\leqslant m_1}b_{m_2-(m_1-j)}t^{-j}+\sum_{0\leqslant j\leqslant r}a_jt^{-j}\right)\\&=t^{m_1-r}\sum_{0\leqslant j\leqslant r}a_{r-j}t^j+\sum_{0\leqslant l\leqslant m_2}b_{m_2-l}t^l\equiv\sum_{0\leqslant l\leqslant m_2}b_{m_2-l}t^l \bmod t^{m_1-r}\\&\equiv G_2^* \bmod t^{m_1-r}\equiv G_2^* \bmod t^k,\end{aligned}$$

since $k\leqslant m_1-r$. Thus, $\chi^*(G_1)-\chi^*(G_2)=\chi(G_1^*)-\chi(G_2^*)=0$, as claimed.

For the second claim, if ψ were principal, then $\chi(G^*)=1$ for all $(G^*,t)=1$. The set $\{G\in\mathbb{F}_q[t]:\ G(0)\neq 0\}$ is invariant under the involution, so this would imply that $\chi(G)=1$ whenever $G(0)\neq 0$; but since $\chi(G)=0$ whenever $G(0)=0$, this implies that $\chi(G)=1_{(G,t^k)=1}$, which implies that χ is principal, and the claim follows. □

Remark 3.13. Note that if χ is a character modulo t^k, then the previous lemma does not prescribe a value for $\chi^*(t)$. However, in keeping with our convention $f^*(t)=0$ for multiplicative functions f, we shall set $\chi^*(t)=0$. In any case, this particular definition will play no significant role in the sequel.

We shall distinguish between the following notions of non-pretentiousness.

Definition 3.14. Let $N\geqslant 1$. Let $f:\mathcal{M}\to\mathbb{U}$ be multiplicative. We say that f is *Hayes non-pretentious to level $W=W(N)$* if, as $N\to\infty$,

$$\min_{w\leqslant W}\min_{\substack{\psi \bmod M\\ M\in\mathcal{M}_w}}\min_{\substack{\xi \text{ short}\\ \mathrm{len}(\xi)\leqslant N}}\mathcal{D}_{f\overline{\chi\xi}}(N)\to\infty.$$

We say that f is *Dirichlet non-pretentious* to level $W=W(N)$ if, as $N\to\infty$,

$$\min_{w\leqslant W}\min_{\substack{\psi \bmod M\\ M\in\mathcal{M}_w}}\mathcal{D}_{f\overline{\chi}}(N)\to\infty.$$

An immediate corollary of Lemma 3.12 relating to Hayes non-pretentiousness (and utilized in Section 7) is the following.

Corollary 3.15 (Hayes non-pretentiousness implies Dirichlet non-pretentiousness of dual). *Let $N\geqslant 1$, and let $W=W(N)\leqslant N$. Let $f:\mathbb{F}_q[t]\to\mathbb{U}$ be multiplicative and even, that is, $f(cG)=f(G)$ for all $c\in\mathbb{F}_q^\times$. Then,*

$$\min_{\substack{\psi \bmod M\\ M\in\mathcal{M}_{\leqslant W(N)+1}}}\min_{\substack{\xi \text{ short}\\ \mathrm{len}(\xi)\leqslant N}}\mathcal{D}_{f\psi\overline{\xi}}(N)\leqslant\min_{M\in\mathcal{M}_{\leqslant W(N)}}\min_{\psi \bmod M}\min_{\substack{\chi \bmod t^\nu\\ 1\leqslant\nu\leqslant N}}\mathcal{D}_{(f\psi)^*\overline{\chi}}(N)+O(1).$$

In particular, if f is Hayes non-pretentious to level $W':=W+1$, then

$$\lim_{N\to\infty}\min_{\substack{\psi \bmod M\\ M\in\mathcal{M}_{\leqslant W(N)}}}\min_{\substack{\chi \bmod t^\nu\\ 1\leqslant\nu\leqslant N}}\mathcal{D}_{(f\psi)^*\overline{\chi}}(N)=\infty.$$

Proof. Let N be large and let $\psi \bmod M$ with $\deg(M)\leqslant W(N)$ and $\chi \bmod t^\nu$ with $1\leqslant\nu\leqslant N$ be chosen such that

$$\mathcal{D}_{(f\psi)^*\overline{\chi}}(N)=\min_{M'\in\mathcal{M}_{\leqslant W(N)}}\min_{\psi' \bmod M'}\min_{\substack{\chi' \bmod t^{\nu'}\\ 1\leqslant\nu'\leqslant N}}\mathcal{D}_{(f\psi')^*\overline{\chi'}}(N).$$

Since f is even and $\psi(c)\in\mu_{q-1}$ for all $c\in\mathbb{F}_q^\times$, we may apply Lemma 3.10 to conclude that there is a character $\xi \bmod t$, depending on ψ and χ, such that

$$\mathcal{D}_{f\psi\xi\overline{\chi^*}}(N)=\mathcal{D}_{(f\psi)^*\overline{\chi}}(N)+O(1).$$

By Lemma 3.12, χ^* coincides with a short interval character of length $\nu-1$ at all primes $P\neq t$, so that $\psi\xi\overline{\chi^*}$ coincides at all $P\in\mathcal{P}\backslash\{t\}$ with a Hayes character whose Dirichlet part has conductor $\leqslant\deg(Mt)\leqslant W(N)+1$ and whose short interval character part has conductor at most N. It follows then that

$$\min_{M\in\mathcal{M}_{\leqslant W(N)+1}}\min_{\psi \bmod M}\min_{\substack{\xi \text{ short}\\ \mathrm{len}(\xi)\leqslant N}}\mathcal{D}_{f\overline{\chi\xi}}(N)\leqslant\mathcal{D}_{f\psi\xi\overline{\chi^*}}(N)+O(1)\leqslant\mathcal{D}_{(f\psi)^*\overline{\chi}}(N)+O(1).$$

This implies the first claim. The second claim follows upon taking $N\to\infty$ and using the definition of Hayes non-pretentiousness. □

4 | PRELIMINARIES II: CHARACTER SUMS AND SIEVE ESTIMATES

Beginning in this section, we set out to prove (a generalization of) Theorem 1.1, as well as Theorem 1.4. We collect together the main general results we shall use for this purpose. Most of these are simple translations of the corresponding result in the number field setting, but we have not managed to locate such translations in the literature.

Remark 4.1. For brevity and to simplify notation, all of the lemmas below are stated for sums of Dirichlet characters, but as we will note in Section 6, all of them work equally well if $\chi \bmod Q$ is replaced with $\chi \in \mathcal{X}_{1,\nu}$ (i.e., we are summing over short interval characters of length ν), and $\deg(Q)$ is replaced with ν and $\phi(Q)$ is replaced with q^{ν}.

4.1 | Large sieve estimates in function fields

Lemma 4.2 (L^2 Mean Value Theorem). *Let $N \geqslant 1$. Let $\{a_G\}_{G\in\mathcal{M}_N} \subset \mathbb{C}$, and let $Q \in \mathcal{M}$. Then,*

$$\sum_{\chi \bmod Q} \left| \sum_{G\in\mathcal{M}_N} a_G \chi(G) \right|^2 \leqslant 2\Big(\phi(Q)q^{N-\deg(Q)} + \phi(Q)\Big) \sum_{\substack{G\in\mathcal{M}_N \\ (G,Q)=1}} |a_G|^2.$$

Remark 4.3. The short interval analog of this lemma reads as

$$\sum_{\xi\in\mathcal{X}_{1,\nu}} \left| \sum_{G\in\mathcal{M}_N} a_G \xi(G) \right|^2 \leqslant 2\big(q^{\nu}q^{N-\nu} + q^{\nu}\big) \sum_{G\in\mathcal{M}_N} |a_G|^2.$$

All the lemmas that follow in this section have short interval formulations in a completely analogous fashion.

Proof. Denote the left-hand side by Σ. Expanding the square and swapping orders of summation yields

$$\Sigma = \sum_{G,G'\in\mathcal{M}_N} a_G\overline{a_{G'}} \sum_{\chi \bmod Q} \chi(G)\overline{\chi}(G') = \phi(Q)\left(\sum_{\substack{G\in\mathcal{M}_N \\ (G,Q)=1}} |a_G|^2 + \sum_{\substack{G,G'\in\mathcal{M}_N \\ G\equiv G' \bmod Q \\ G\neq G',(GG',Q)=1}} a_G\overline{a_{G'}} \right).$$

Bounding the second sum trivially, using the AM–GM inequality in the form $|a_G a_{G'}| \leqslant \frac{1}{2}(|a_G|^2 + |a_{G'}|^2)$ and invoking symmetry in G and G', we get

$$\Sigma \leqslant \phi(Q) \sum_{\substack{G\in\mathcal{M}_N \\ (G,Q)=1}} |a_G|^2 \left(1 + \sum_{\substack{G'\in\mathcal{M}_N \\ Q\mid(G'-G)}} 1 \right).$$

Since $\deg(G'-G) \leqslant N$ for each $G \in \mathcal{M}_N$, and the number of polynomials in $\mathcal{M}_{\leqslant N}$ divisible by Q is precisely $|\mathcal{M}_{\leqslant N-\deg(Q)}| \leqslant 2q^{N-\deg(Q)}$, it follows that

$$\Sigma \leqslant \left(\phi(Q)+\phi(Q)|\mathcal{M}_{\leqslant N-\deg(Q)}|\right) \sum_{\substack{G\in\mathcal{M}_N\\(G,Q)=1}} |a_G|^2 \leqslant 2\Big(\phi(Q)+\phi(Q)q^{N-\deg(Q)}\Big) \sum_{\substack{G\in\mathcal{M}_N\\(G,Q)=1}} |a_G|^2,$$

as claimed. □

Lemma 4.4 (Halász–Montgomery Lemma). *Let $N \geqslant 1$. Let $\{a_G\}_{G\in\mathcal{M}_N} \subset \mathbb{C}$, and let $Q \in \mathcal{M}$, $\deg(Q) \leqslant (1+o(1))N$. Let $\Xi \subseteq \mathcal{X}_Q$. Then,*

$$\sum_{\chi\in\Xi}\left|\sum_{G\in\mathcal{M}_N} a_G\chi(G)\right|^2 \ll \left(\phi(Q)q^{N-\deg(Q)}+|\Xi|q^{(1/2+o(1))N}\right) \sum_{\substack{G\in\mathcal{M}_N\\(G,Q)=1}} |a_G|^2.$$

Proof. We may obviously assume that $\Xi \neq \emptyset$, since otherwise the claim is trivial. Moreover, by duality (see, e.g., [27, Lemma 10]), it suffices to show that for any set of coefficients $\{c_\chi\}_{\chi\in\Xi} \subset \mathbb{C}$, we have

$$\sum_{G\in\mathcal{M}_N}\left|\sum_{\chi\in\Xi} c_\chi\chi(G)\right|^2 \ll \left(\phi(Q)q^{N-\deg(Q)}+|\Xi|q^{\deg(Q)/2}\right) \sum_{\chi\in\Xi} |c_\chi|^2.$$

Expanding the square in the left-hand side and swapping the order of summations, we get

$$\sum_{\chi_1,\chi_2\in\Xi} c_{\chi_1}\overline{c_{\chi_2}} \sum_{G\in\mathcal{M}_N} \chi_1\overline{\chi}_2(G).$$

The diagonal contribution with $\chi_1=\chi_2$ yields

$$|\{G\in\mathcal{M}_N : (G,Q)=1\}| \sum_{\chi\in\Xi}|c_\chi|^2 \ll \phi(Q)q^{N-\deg(Q)} \sum_{\chi\in\Xi}|c_\chi|^2.$$

When $\chi_1\neq\chi_2$, $\chi_1\overline{\chi}_2$ is non-principal, so by Lemma 3.3 we have

$$\sum_{\substack{\chi_1,\chi_2\in\Xi\\\chi_1\neq\chi_2}} |c_{\chi_1}c_{\chi_2}|\left|\sum_{G\in\mathcal{M}_N}\chi_1\overline{\chi_2}(G)\right| \ll q^{(1/2+o(1))N} \sum_{\substack{\chi_1,\chi_2\in\Xi\\\chi_1\neq\chi_2}} |c_{\chi_1}||c_{\chi_2}|.$$

Applying AM–GM as in the proof of the previous lemma, the sum above is bounded by $|\Xi|\sum_{\chi\in\Xi}|c_\chi|^2$. Putting everything together, this proves the claim. □

Lemma 4.5 (Halász–Montgomery Lemma for Primes). *Let $N \geqslant 1$. Let $\{a_P\}_{P\in\mathcal{P}_N} \subset \mathbb{C}$, and let $Q \in \mathcal{M}$. For any $\Xi \subseteq \mathcal{X}_Q$, we have*

$$\sum_{\chi\in\Xi}\left|\sum_{P\in\mathcal{P}_N} a_P\chi(P)\right|^2 \ll \left(\frac{q^N}{N}+\deg(Q)\frac{q^{N/2}}{N}|\Xi|\right) \sum_{P\in\mathcal{P}_N}|a_P|^2.$$

Proof. We apply duality, as in the proof of Lemma 4.4. Given a sequence $\{c_\chi\}_{\chi\in\Xi}\subset\mathbb{C}$, we bound $1_{G\in\mathcal{P}_N}\leqslant N^{-1}\Lambda(G)$ to obtain

$$\sum_{P\in\mathcal{P}_N}\left|\sum_{\chi\in\Xi}c_\chi\chi(P)\right|^2\leqslant\sum_{G\in\mathcal{M}_N}\frac{\Lambda(G)}{N}\left|\sum_{\chi\in\Xi}c_\chi\chi(G)\right|^2=\frac{1}{N}\sum_{\chi_1,\chi_2\in\Xi}c_{\chi_1}\overline{c_{\chi_2}}\sum_{G\in\mathcal{M}_N}\Lambda(G)\chi_1\overline{\chi_2}(G).$$

When $\chi_1=\chi_2$, the prime polynomial theorem gives $\sum_{G\in\mathcal{M}_N}\Lambda(G)\ll q^N$, whence the diagonal contribution to the sum becomes $(q^N/N)\sum_{\chi\in\Xi}|c_\chi|^2$.

When $\chi_1\neq\chi_2$, we may apply Lemma 3.1 to give

$$\sum_{G\in\mathcal{M}_N}\Lambda(G)\chi_1\overline{\chi_2}(G)\ll\deg(Q)q^{N/2}.$$

It follows that

$$\sum_{\substack{\chi_1,\chi_2\in\Xi\\ \chi_1\neq\chi_2}}|c_{\chi_1}||c_{\chi_2}|\left|\sum_{G\in\mathcal{M}_N}\Lambda(G)\chi_1\overline{\chi_2}(G)\right|\ll\deg(Q)q^{N/2}|\Xi|\sum_{\chi\in\Xi}|c_\chi|^2,$$

upon applying AM–GM and using symmetry, as before.

Combined with the diagonal contribution, we get

$$\sum_{P\in\mathcal{P}_N}\left|\sum_{\chi\in\Xi}c_\chi\chi(P)\right|^2\ll\left(\frac{q^N}{N}+\deg(Q)\frac{q^{N/2}}{N}|\Xi|\right)\sum_{\chi\in\Xi}|c_\chi|^2.$$

Invoking duality as discussed above, the claim follows. □

Lemma 4.6 (A large values estimate). *Let $N,Z\geqslant 1$. Let $\{a_P\}_{P\in\mathcal{P}_N}\subset\mathbb{U}$, and let $Q\in\mathcal{M}$, with $\phi(Q)\geqslant q^N$. Then,*

$$\left|\left\{\chi \bmod Q:\ \frac{1}{q^N}\left|\sum_{P\in\mathcal{P}_N}a_P\chi(P)\right|\geqslant\frac{1}{Z}\right\}\right|$$
$$\ll\exp\left(\frac{\log(q^N\phi(Q))}{N\log q}\left(2\log\left(\frac{2\log\phi(Q)}{N\log q}\right)+\log\left(\frac{2Z^2}{N}\right)\right)\right).$$

Proof. The proof is essentially the same as in the number fields case [27, Lemma 8]. Let $k:=\lfloor\frac{\log\phi(Q)}{N\log q}\rfloor+1$. Let $\mathcal{N}$ denote the cardinality of the set of characters on the left-hand side. By Chebyshev's inequality, we have

$$\mathcal{N}\leqslant\left(\frac{Z}{q^N}\right)^{2k}\sum_{\chi \bmod Q}\left|\sum_{P\in\mathcal{P}_N}a_P\chi(P)\right|^{2k}=\left(\frac{Z}{q^N}\right)^{2k}\sum_{\chi \bmod Q}\left|\left(\sum_{P\in\mathcal{P}_N}a_P\chi(P)\right)^k\right|^2$$
$$=\left(\frac{Z}{q^N}\right)^{2k}\sum_{\chi \bmod Q}\left|\sum_{G\in\mathcal{M}_{kN}}b_G\chi(G)\right|^2,\tag{23}$$

where we have defined

$$b_G := \sum_{\substack{P_1\cdots P_k=G\\ P_j\in\mathcal{P}_N\ \forall j}} a_{P_1}\cdots a_{P_k}.$$

Applying Lemma 4.2, we get

$$\sum_{\chi \bmod Q}\left|\sum_{G\in\mathcal{M}_{kN}} b_G\chi(G)\right|^2 \ll \left(\phi(Q)+q^{kN}\right)\sum_{G\in\mathcal{M}_{kN}}|b_G|^2$$
$$\ll q^{kN}\sum_{\substack{P_1\cdots P_k=Q_1\cdots Q_k\\ P_i,Q_j\in\mathcal{P}_N}} a_{P_1}\cdots a_{P_k}\overline{a_{Q_1}\cdots a_{Q_k}},$$

according to our choice of k. Since the P_i and Q_j are irreducible, up to permutation we have $P_i = Q_i$ for all $1\leqslant i\leqslant k$, and thus by the prime polynomial theorem

$$\sum_{\substack{P_1\cdots P_k=Q_1\cdots Q_k\\ P_i,Q_j\in\mathcal{P}_N}} a_{P_1}\cdots a_{P_k}\overline{a_{Q_1}\cdots a_{Q_k}} \leqslant (k!)^2\left(\sum_{P\in\mathcal{P}_N}|a_P|^2\right)^k \ll (k!)^2(1.1q^N/N)^k.$$

Inserting this into our mean value estimate, we get that

$$\sum_{\chi \bmod Q}\left|\sum_{P\in\mathcal{P}_N} a_P\chi(P)\right|^{2k} \ll \left(\frac{q^{2N}}{N}\right)^k 1.1^k(k!)^2.$$

Combining this with (23) and using $\log k!\leqslant k\log k$ for $k\geqslant 2$, we find that

$$\mathcal{N} \ll 1.1^k(k!)^2(Z^2/N)^k \ll \exp\left(\left(1+\frac{\log\phi(Q)}{N\log q}\right)\left(2\log\left(\frac{2\log\phi(Q)}{N\log q}\right)+\log(2Z^2/N)\right)\right).$$

This implies the claim. □

Lemma 4.7 (A moment computation). *Let $1\leqslant d\leqslant m\leqslant N$. Let $\{a_P\}_{P\in\mathcal{P}_d}, \{b_G\}_{G\in\mathcal{M}_{N-m}}\subset\mathbb{U}$. Set*

$$U(\chi) := \frac{1}{d|\mathcal{P}_d|}\sum_{P\in\mathcal{P}_d} a_P\chi(P),$$

$$V(\chi) := \frac{1}{|\mathcal{M}_{N-m}|}\sum_{G\in\mathcal{M}_{N-m}} b_G\chi(G).$$

Set $\ell := \lceil m/d\rceil$. Then, for any $Q\in\mathcal{M}$, we have

$$\sum_{\chi \bmod Q}|U(\chi)^\ell V(\chi)|^2 \ll \left(\phi(Q)q^{-N}+\phi(Q)q^{-\deg(Q)}\right)\ell^{2\ell}.$$

Proof. This is similar to [27, Lemma 13]. Expanding out the product for each χ, we have

$$U(\chi)^\ell V(\chi) = \frac{1}{d^\ell |\mathcal{P}_d|^\ell |\mathcal{M}_{N-m}|} \sum_{M \in \mathcal{M}_{N-m+\ell d}} \chi(M) \left(\sum_{\substack{GP_1 \cdots P_\ell = M \\ P_j \in \mathcal{P}_d \ \forall j}} a_{P_1} \cdots a_{P_\ell} b_G \right).$$

We denote by $g(M)$ the bracketed sum on the right-hand side. Taking squares, summing over $\chi \bmod Q$, and then applying Lemma 4.2 (and the prime polynomial theorem) yields

$$\begin{aligned} \sum_{\chi \bmod Q} |U(\chi)^\ell V(\chi)|^2 &\ll \phi(Q)\left(1 + q^{N-\deg(Q)-m+\ell d}\right) \frac{1}{d^{2\ell} |\mathcal{P}_d|^{2\ell} |\mathcal{M}_{N-m}|^2} \sum_{M \in \mathcal{M}_{N-m+\ell d}} |g(M)|^2 \\ &\ll \left(\phi(Q) q^{-N} + \phi(Q) q^{-\deg(Q)}\right) \frac{1.1^\ell}{|\mathcal{M}_{N-m+\ell d}|} \sum_{M \in \mathcal{M}_{N-m+\ell d}} |g(M)|^2. \end{aligned}$$

Now, by the triangle inequality we can bound g as

$$|g(M)| \leqslant \sum_{\substack{GP_1 \cdots P_\ell = M \\ P_j \in \mathcal{P}_d \forall j}} 1 \leqslant (\ell!) 1 * \gamma(M) =: (\ell!) \widetilde{g}(M),$$

where γ is the indicator function of monic polynomials all of whose prime factors belong to $\mathcal{P}_d$; note that on prime powers, $\widetilde{g}(P^k) = 1 + k 1_{\mathcal{P}_d}(P)$, which is $\ll_\varepsilon q^{\varepsilon k \deg(R)}$ for any $\varepsilon > 0$ and $k \geqslant 1$, and $\widetilde{g}(P) \leqslant 2$ for all irreducibles P. We may thus apply Lemma 3.8 to get that

$$\begin{aligned} \frac{1}{|\mathcal{M}_{N-m+d\ell}|} \sum_{M \in \mathcal{M}_{N-m+\ell d}} \widetilde{g}(M)^2 &\ll \frac{1}{N-m+\ell d} \exp\left(\sum_{P \in \mathcal{P}_{\leqslant N-m+\ell d}} \widetilde{g}(P)^2 q^{-\deg(P)} \right) \\ &\ll \exp\left(\sum_{P \in \mathcal{P}_d} (2^2 - 1) q^{-\deg(P)} \right) \ll 1. \end{aligned}$$

Inserting this into the above estimate, we get

$$\sum_{\chi \bmod Q} |U(\chi)^\ell V(\chi)|^2 \ll \left(\phi(Q) q^{-N} + \phi(Q) q^{-\deg(Q)}\right) \ell^{2\ell},$$

as claimed. □

4.2 | Sieve bounds in function fields

Our next result shows that most monics have irreducible factors whose degrees belong to prescribed ranges, provided these ranges are large enough.

Lemma 4.8. *Let $P < Q$. Then,*

$$|\{G \in \mathcal{M}_N : R \in \mathcal{P} \text{ such that } R|G \Rightarrow \deg(R) \notin [P,Q]\}| \ll \frac{P}{Q}q^N.$$

Proof. Let g denote the indicator function for the set on the left-hand side. Then, $0 \leqslant g \leqslant 1$ and g is multiplicative. By Lemma 3.8, the left-hand side is

$$\sum_{G\in\mathcal{M}_N} g(G) \ll \frac{q^N}{N}\exp\left(\sum_{\substack{R\in\mathcal{P}_{\leqslant N}\\ \deg(R)\notin[P,Q]}} q^{-d}\right) \ll q^N \exp\left(-\sum_{P\leqslant d\leqslant Q} q^{-d}|\mathcal{P}_d|\right) \ll \frac{P}{Q}q^N,$$

as claimed. □

Definition 4.9. Let $J \geqslant 1$, and let $\boldsymbol{P} := \{P_j\}_{1\leqslant j\leqslant J}$ and $\boldsymbol{Q} := \{Q_j\}_{1\leqslant j\leqslant J}$ be collections of parameters satisfying $P_j < P_{j+1}$, $Q_j < Q_{j+1}$, and $P_j < Q_j$ for all j. We define the set $S_{\boldsymbol{P},\boldsymbol{Q}}(N)$ by

$$S_{\boldsymbol{P},\boldsymbol{Q}} := \{G \in \mathcal{M} : \forall\, 1 \leqslant j \leqslant J\ \exists\, d \in [P_j, Q_j], R \in \mathcal{P}_d \text{ such that } R|G\}.$$

If $J = 1$ then, for convenience, we write $S_{P_1,Q_1} = S_{\boldsymbol{P},\boldsymbol{Q}}$.

We will be able to restrict character-twisted sums over monic polynomials to monics belonging to sets of the form $S_{\boldsymbol{P},\boldsymbol{Q}}(N)$, on average.

Lemma 4.10. *Let $N \geqslant 1$, and let $Q \in \mathcal{M}$ with $\deg(Q) \leqslant N$. Let $\Xi \subseteq \mathcal{X}_Q$ be a set of characters modulo Q, and let $f : \mathcal{M} \to \mathbb{U}$ be multiplicative. Then,*

$$\sum_{\chi\in\Xi}\left|\frac{1}{|\mathcal{M}_N|}\sum_{G\in\mathcal{M}_N} f(G)\overline{\chi}(G)\right|^2 \ll \sum_{\chi\in\Xi}\left|\frac{1}{|\mathcal{M}_N|}\sum_{\substack{G\in\mathcal{M}_N\\ G\in S_{\boldsymbol{P},\boldsymbol{Q}}}} f(G)\overline{\chi}(G)\right|^2 + \phi(Q)q^{-\deg(Q)}\sum_{1\leqslant j\leqslant J}\frac{P_j}{Q_j}.$$

Proof. Given a map $g : \mathbb{F}_q[t] \to \mathbb{C}$, set $M_g(N) := \frac{1}{|\mathcal{M}_N|}\sum_{G\in\mathcal{M}_N} g(G)$. For each $\chi \bmod M$, we have

$$|M_{f\overline{\chi}}(N)|^2 \leqslant 2|M_{f\overline{\chi}1_{S_{\boldsymbol{P},\boldsymbol{Q}}}}(N)|^2 + 2|M_{f\overline{\chi}1_{S^c_{\boldsymbol{P},\boldsymbol{Q}}}}(N)|^2.$$

Summing the first of these terms over $\chi \in \Xi$ gives the first term in the estimate. Summing the second term over χ and applying Lemma 4.2 gives

$$\sum_{\chi\in\Xi}|M_{f\chi1_{S^c_{\boldsymbol{P},\boldsymbol{Q}}}}(N)|^2 \leqslant \sum_{\chi \bmod Q}|M_{f\chi1_{S^c_{\boldsymbol{P},\boldsymbol{Q}}}}(N)|^2 \ll \phi(Q)\left(q^{N-\deg(Q)}+1\right)\frac{1}{|\mathcal{M}_N|^2}\sum_{\substack{G\in\mathcal{M}_N\\ G\notin S_{\boldsymbol{P},\boldsymbol{Q}}}} 1.$$

By the union bound and Lemma 4.8, we have

$$\frac{1}{|\mathcal{M}_N|}\sum_{\substack{G\in\mathcal{M}_N\\ G\notin S_{\mathbf{P},\mathbf{Q}}}} 1 \leqslant \sum_{1\leqslant j\leqslant J}\frac{1}{|\mathcal{M}_N|}|\{G\in\mathcal{M}_N:\ R\in\mathcal{P}, R|G\Rightarrow \deg(R)\notin[P_j,Q_j]\}|$$

$$\ll \sum_{1\leqslant j\leqslant J}\frac{P_j}{Q_j}.$$

This implies the claim. □

We will also need the following estimate for smooth (otherwise known as friable) polynomials, that is, polynomials with no irreducible factors of large degree. For $1\leqslant M\leqslant N$, we write

$$S(N,M) := \{G\in\mathcal{M}_N:\ R\in\mathcal{P} \text{ and } R|G\Rightarrow \deg(R)\leqslant M\}.$$

Lemma 4.11. *Let $1\leqslant M\leqslant N$. Then, for some absolute constant $c>0$ we have*

$$|S(N,M)|\ll q^N\exp(-cN/M).$$

Proof. This follows from [37]. □

Lemma 4.12 (Selberg upper bound sieve in function fields). *Let $1\leqslant y\leqslant z, H\leqslant N$ and let $\mathcal{A}\subseteq\mathcal{M}_N$. Put*

$$\mathfrak{P}_{y,z}=\prod_{Q\in\mathcal{P}_{\leqslant z}\setminus\mathcal{P}_{\leqslant y}} Q.$$

Suppose g is a multiplicative function supported on squarefree monic polynomials such that for each $D\in\mathcal{M}$ squarefree with $D\in\mathcal{M}_{\leqslant H}$,

$$\sum_{\substack{G\in\mathcal{A}\\ D|G}} 1 = g(D)|\mathcal{A}| + r_D(\mathcal{A}). \tag{24}$$

Put $J=J(H)=\sum_{\substack{D|\mathfrak{P}_{y,z}\\ \deg(D)\leqslant H}}\prod_{\substack{R\in\mathcal{P}\\ R|D}} g(R)/(1-g(R))$. Then,

$$\sum_{\substack{G\in\mathcal{A}\\ (G,\mathfrak{P}_{y,z})=1}} 1 \leqslant |\mathcal{A}|J^{-1}+\sum_{\deg(D)\leqslant H}\tau_3(D)|r_D(\mathcal{A})|,$$

where $\tau_3(D)=\sum_{\substack{A,B,C\in\mathcal{M}\\ ABC=D}} 1$.

Proof. This follows from [38, Theorem 1] (take $\mathcal{P}:=\mathcal{P}_{\leqslant z}\setminus\mathcal{P}_{\leqslant y}$ and $\mathcal{D}:=\{D\in\mathcal{D}:\ \deg(D)\leqslant H\}$, which is divisor closed, as needed according to the hypotheses there). □

We have the following useful corollary.

Corollary 4.13 (Additive energy of irreducible polynomials). *Let $H \geqslant 1$. If $M \in \mathbb{F}_q[t]$ has $\deg(M) < H$, then*

$$|\{(P_1, P_2, P_3, P_4) \in \mathcal{P}_H^4 : P_1 + P_2 - P_3 - P_4 = M\}| \ll q^{3H}/H^4.$$

Proof. We begin by considering the case $M = 0$. Given $G \in \mathbb{F}_q[t]$ of degree H, let $r(G)$ denote the number of representations of G as a sum of two irreducible polynomials of degree $\leqslant H$. Note that if $q = 2$, then $r(G) > 0$ only when $\deg(G) < H$; otherwise, if $q > 2$, then $r(G) > 0$ only when $\deg(G) = H$. We thus have

$$\sum_{\substack{P_1,P_2,P_3,P_4 \in \mathcal{P}_H \\ P_1+P_2=P_3+P_4}} 1 = \sum_{\deg(G)\leqslant H} r(G)^2.$$

Let $z := H/2$, $y = 1$, and $\mathfrak{P}_z := \mathfrak{P}_{y,z} = \prod_{P \in \mathcal{P}_{\leqslant z} \setminus \mathcal{P}_{\leqslant y}} P$ as in the previous lemma. We then have

$$r(G) \leqslant \sum_{\substack{M \in \mathcal{M}_H \\ (M(G-M), \mathfrak{P}_z) = 1 \\ G-M \in \mathcal{M}_H}} 1 = \sum_{\substack{F \in \mathcal{A} \\ (F, \mathfrak{P}_z) = 1}} 1,$$

where $\mathcal{A} = \{B(G-B) : B \in \mathcal{M}_H\} \cap \mathcal{M}_{2H}$; as $\mathcal{A}$ and $\mathcal{M}_H$ are in bijection with one another, we have $|\mathcal{A}| = |\mathcal{M}_H| \asymp q^H$.

Note that for $D | \mathfrak{P}_z$ with $\deg(D) \leqslant H$

$$\sum_{\substack{F \in \mathcal{A} \\ D | F}} 1 = g(D)|\mathcal{M}_H|,$$

where g is the multiplicative function supported on squarefree polynomials and defined at irreducibles via $g(P) = 2q^{-\deg(P)}$ if $P \nmid G$ and $g(P) = q^{-\deg(P)}$ otherwise; note that $g(P) \leqslant 1/2$ for all $P | \mathfrak{P}_z$ and all $q \geqslant 2$, since such P must have $\deg(P) \geqslant 2$. By Lemma 4.12, we deduce

$$\sum_{\substack{F \in \mathcal{A} \\ (F, \mathfrak{P}_z) = 1}} 1 \ll q^H \left(\sum_{\substack{D | \mathfrak{P}_z \\ \deg(D) \leqslant H}} \prod_{P | D} \frac{g(P)}{1 - g(P)} \right)^{-1} \ll q^H \left(\sum_{D | \mathfrak{P}_z} g(D) - \sum_{\substack{D | \mathfrak{P}_z \\ \deg(D) > H}} g(D) \right)^{-1}.$$

Note that the full bracketed sum over $D | \mathfrak{P}_z$ has order of magnitude

$$= \prod_{R \in \mathcal{P}_{\leqslant z}} (1 + g(R)) \asymp \exp\left(2 \sum_{R \in \mathcal{P}_{\leqslant z}} q^{-\deg(R)} - \sum_{\substack{R \in \mathcal{P} \\ R | G}} q^{-\deg(R)} \right) \asymp \frac{\phi(G)}{q^{\deg(G)}} z^2.$$

The remaining sum over $D|\mathfrak{P}_z$ with $\deg(D) > H$ can be bounded above as

$$\leqslant \sum_{k>H} q^{-k} \sum_{\substack{D|P_z \\ \deg(D)=k}} 2^{\omega(D)} = \sum_{k>H} q^{-k} \sum_{\substack{a_1+2a_2+\cdots+za_z=k \\ 0\leqslant a_j\leqslant|\mathcal{P}_j|}} \prod_{1\leqslant j\leqslant z} 2^{a_j}$$

$$\leqslant \sum_{k>H} q^{-k} \sum_{\substack{a_1+2a_2+\cdots+za_z=k \\ 0\leqslant a_j\leqslant|\mathcal{P}_j|}} 2^{q+\frac{1}{2}(2a_2+\cdots+za_z)}$$

$$\ll \sum_{k>H} (\sqrt{2}/q)^k |\{\boldsymbol{a} \in \mathbb{N}\cup\{0\} : a_1 + 2a_2 + \cdots + za_z = k\}|.$$

Using standard results on partitions (see, e.g., [5]), the cardinality above is $\ll e^{c\sqrt{k}}$, for some $c > 0$ absolute. Thus, as $q \geqslant 2$ the series over k converges, and in fact

$$\sum_{\substack{D|\mathfrak{P}_z \\ \deg(D)>H}} g(D) \ll e^{-c'H},$$

for a suitable absolute $c' > 0$.

It follows that for large H,

$$r(G) \ll \frac{q^{\deg(G)}}{\phi(G)}\frac{q^H}{z^2} \ll \frac{q^{\deg(G)}}{\phi(G)}\frac{q^H}{H^2}.$$

Squaring this bound and summing over $G \in \mathbb{F}_q[t]$ of degree $\leqslant H$ for which $r(G) \neq 0$, we get that

$$|\{(P_1,P_2,P_3,P_4) \in \mathcal{P}_H : P_1 + P_2 = P_3 + P_4\}| \leqslant \sum_{\deg(G)\leqslant H} r(G)^2 \ll \frac{q^{2H}}{H^4} \sum_{\deg(G)\leqslant H} \left(\frac{q^{\deg(G)}}{\phi(G)}\right)^2.$$

We claim that the sum over G is $\ll q^H$, which will then imply the claim for $M = 0$. To see this, write $\psi(G) := (q^{\deg(G)}/\phi(G))^2$; note that ψ is independent of the leading coefficient of G, and so we may replace G by $G/G(0)$ and assume G is monic (this changes the sum by at most a factor depending only on q). Note that for any $k \geqslant 1$, $\psi(R^k) = \psi(R) = (1 - q^{-\deg(R)})^{-2} \leqslant 4$ uniformly over $R \in \mathcal{P}$. Hence, we may apply Lemma 3.8 to get

$$\sum_{G\in\mathcal{M}_{\leqslant H}} \psi(G) \ll \sum_{0\leqslant h\leqslant H} \frac{q^h}{h} \exp\left(\sum_{R\in\mathcal{P}_{\leqslant h}} \psi(R) q^{-\deg(R)}\right).$$

We may directly evaluate the sum over R here for each $h \leqslant H$ by the prime polynomial theorem, getting

$$\sum_{R\in\mathcal{P}_{\leqslant h}} \psi(R) q^{-\deg(R)} = \sum_{k\leqslant h} \frac{q^{-k}}{(1-q^{-k})^2} |\mathcal{P}_k| = \sum_{k\leqslant h} \left(\frac{1}{k}(1-q^{-k})^{-2} + O(q^{-k/2})\right) = \log h + O(1),$$

which leads to

$$\sum_{G\in\mathcal{M}_{\leqslant H}}\psi(G)\ll\sum_{h\leqslant H}q^h\ll q^H,$$

as required. Next, let $M\in\mathcal{M}_{<H}$. Then,

$$|\{(P_1,P_2,P_3,P_4)\in\mathcal{P}_H^4:\ P_1+P_2-P_3-P_4=M\}|=\sum_{\deg(G)\leqslant H}r(G)r(M+G)\leqslant\sum_{\deg(G)\leqslant H}r(G)^2,$$

by the AM–GM inequality and the fact that $\deg(G+M)\leqslant H$ whenever $\max\{\deg(G),\deg(M)\}\leqslant H$. The second claim now follows from the first. □

4.3 | Dirichlet polynomial decompositions

Let $Q>P\geqslant 1$. Recall that $S_{P,Q}$ denotes the set of monic G that have an irreducible factor R satisfying $\deg(R)\in[P,Q]$.

Lemma 4.14 (Ramaré's identity). *Let $P<Q$. Let $f:\ \mathcal{M}\to\mathbb{U}$ be multiplicative. Then, for any $G\in S_{P,Q}$,*

$$f(G)=\sum_{\substack{RM=G\\R\in\mathcal{P}\\\deg(R)\in[P,Q]}}\frac{f(RM)}{1_{(R,M)=1}+\omega_{[P,Q]}(M)},$$

where $\omega_{[P,Q]}(M):=|\{R\in\mathcal{P}:\ \deg(R)\in[P,Q],R|M\}|$.

Proof. Since $\omega_{[P,Q]}(G)\geqslant 1$ by assumption we have

$$1=\sum_{\substack{R|G\\R\in\mathcal{P}\\\deg(R)\in[P,Q]}}\frac{1}{\omega_{[P,Q]}(G)}=\sum_{\substack{RM=G\\R\in\mathcal{P}\\\deg(R)\in[P,Q]}}\frac{1}{\omega_{[P,Q]}(RM)}=\sum_{\substack{RM=G\\R\in\mathcal{P}\\\deg(R)\in[P,Q]}}\frac{1}{1_{(R,M)=1}+\omega_{[P,Q]}(M)}.$$

This implies the claim. □

We will use Ramaré's identity to decompose Dirichlet polynomials supported on $S_{P,Q}$, as in the following lemma.

Lemma 4.15. *Let $N\geqslant 1$. Let $L\in\mathcal{M}_{\leqslant N}$ and suppose $\Xi\subseteq\mathcal{X}_L$. Finally, let $f:\ \mathcal{M}\to\mathbb{U}$ be multiplicative. Then, for any $1\leqslant P<Q\leqslant N$,*

$$\sum_{\chi\in\Xi}\left|q^{-N}\sum_{G\in\mathcal{M}_N}f(G)\overline{\chi}(G)1_{S_{P,Q}}(G)\right|^2\ll(Q-P+1)\sum_{P\leqslant d\leqslant Q}\sum_{\chi\in\Xi}|A_d(\chi)|^2|B_{N-d}(\chi)|^2+\phi(L)(q^{-N}+q^{-\deg(L)})q^{-P},$$

where for $d \geqslant 1$ and a character χ modulo L, we set

$$A_d(\chi) := q^{-d} \sum_{R \in \mathcal{P}_d} f(R)\overline{\chi}(R).$$

$$B_{N-d}(\chi) := q^{-N+d} \sum_{D \in \mathcal{M}_{N-d}} \frac{f(D)\overline{\chi}(D)}{1+\omega_{[P,Q]}(D)}.$$

Proof. This is analogous to [27, Lemma 12]. By Lemma 4.14, for any $\chi \in \Xi$ we have

$$\begin{aligned}
&q^{-N} \sum_{G \in \mathcal{M}_N} f(G)\overline{\chi}(G) 1_{S_{P,Q}}(G) \\
&= q^{-N} \sum_{RM \in \mathcal{M}_N} \frac{f\overline{\chi}(R) f\overline{\chi}(M)}{1+\omega_{[P,Q]}(M)} + q^{-N} \sum_{RM \in \mathcal{M}_N} \frac{(f(RM)-f(R)f(M))\overline{\chi}(RM)}{1+\omega_{[P,Q]}(M)} \\
&\quad + q^{-N} \sum_{RM \in \mathcal{M}_N} f(RM)\overline{\chi}(RM)\left(\frac{1}{1_{(R,M)=1}+\omega_{[P,Q]}(M)} - \frac{1}{1+\omega_{[P,Q]}(M)}\right) \\
&= \sum_{P \leqslant d \leqslant Q} \left(q^{-d} \sum_{R \in \mathcal{P}_d} f(R)\overline{\chi}(R)\right)\left(q^{-N+d} \sum_{M \in \mathcal{M}_{N-d}} \frac{f(M)\overline{\chi}(M)}{1+\omega_{[P,Q]}(M)}\right) + \mathcal{R}_{1,\chi} + \mathcal{R}_{2,\chi} \\
&= \sum_{P \leqslant d \leqslant Q} A_d(\chi) B_{N-d}(\chi) + \mathcal{R}_{1,\chi} + \mathcal{R}_{2,\chi}. \qquad (25)
\end{aligned}$$

Note that for each $\chi \in \Xi$, both of $\mathcal{R}_{1,\chi}$ and $\mathcal{R}_{2,\chi}$ are supported on polynomials M such that $R|M$ for some $R \in \mathcal{P}$, $\deg(R) \in [P,Q]$. We now take squares and sum the whole expression over all $\chi \in \Xi$ to see that the mean square of (25) is

$$\ll \sum_{\chi \in \Xi} \left| \sum_{P \leqslant d \leqslant Q} A_d(\chi) B_{N-d}(\chi) \right|^2 + \sum_{\chi \in \Xi} |\mathcal{R}_{1,\chi}|^2 + \sum_{\chi \in \Xi} |\mathcal{R}_{2,\chi}|^2.$$

To treat the first term, we use the Cauchy–Schwarz inequality in the inner sum to get

$$\sum_{\chi \in \Xi} \left| \sum_{P \leqslant d \leqslant Q} A_d(\chi) B_{N-d}(\chi) \right|^2 \leqslant (Q-P+1) \sum_{P \leqslant d \leqslant Q} \sum_{\chi \in \Xi} |A_d(\chi)|^2 |B_{N-d}(\chi)|^2.$$

To treat $\sum_{\chi \in \Xi} |\mathcal{R}_{j,\chi}|^2$ for $j = 1, 2$ we use Lemma 4.2; since the arguments are similar, we shall restrict ourselves to proving the bound for $\mathcal{R}_{1,\chi}$. By Lemma 4.2,

$$\sum_{\chi \in \Xi} |R_{1,\chi}|^2 \ll \phi(L)\left(q^{-N} + q^{-\deg(L)}\right) q^{-N} \sum_{G \in \mathcal{M}_N} \left| \sum_{\substack{RM=G \\ R \in \mathcal{P}, R|M \\ \deg(R) \in [P,Q]}} \frac{(f(RM)-f(R)f(M))}{1+\omega_{[P,Q]}(M)} \right|^2.$$

Expanding the square and bounding the summands trivially, we bound the sum on the right-hand side as

$$\sum_{\substack{R_1,R_2\in\mathcal{P}\\ \deg(R_j)\in[P,Q]}}\sum_{\substack{G\in\mathcal{M}_N\\ [R_1,R_2]^2\mid G}} 1 \ll q^N\left(\sum_{P\leqslant d\leqslant Q}|\mathcal{P}_d|q^{-2d}+\sum_{P\leqslant d_1,d_2\leqslant Q}|\mathcal{P}_{d_1}||\mathcal{P}_{d_2}|q^{-2(d_1+d_2)}\right)$$
$$\ll q^N\sum_{P\leqslant d\leqslant Q}q^{-d}\ll q^{N-P},$$

which implies the claim. □

Lemma 4.16 (Pointwise bound with Ramaré weight). *Let $1\leqslant P<Q<N^{0.9}$. Let $f:\mathcal{M}\to\mathbb{U}$ be multiplicative. There is a Hayes character $\widetilde{\chi_1}$ of conductor $\leqslant N$ such that*

$$\max_{\substack{\mathrm{cond}_H(\widetilde{\chi})\leqslant N\\ \widetilde{\chi}\nsim\widetilde{\chi_1}}}\frac{1}{|\mathcal{M}_N|}\left|\sum_{G\in\mathcal{M}_N}\frac{f(G)\overline{\widetilde{\chi}}(G)}{1+\omega_{[P,Q]}(G)}\right|\ll(Q/P)^3N^{-1/4+o(1)}.$$

Moreover, we can take $\widetilde{\chi_1}$ to be the Hayes character of conductor $\leqslant N$ that minimizes $\chi\mapsto\mathcal{D}_{f\overline{\widetilde{\chi}}}(N)$.

Proof. Let $\widetilde{\chi_1}$ be the character that minimizes $\mathcal{D}_{f\overline{\widetilde{\chi}}}(N)$ among all $\widetilde{\chi}$ of conductor $\leqslant N$, and let $\widetilde{\chi}\nsim\widetilde{\chi_1}$. Write $\mathcal{I}:=[P,Q]$ and $\mathcal{I}^c:=\mathbb{N}\backslash\mathcal{I}$. We can express $f=f_{\mathcal{I}}*f_{\mathcal{I}^c}$, where for $\mathcal{J}\in\{\mathcal{I},\mathcal{I}^c\}$, we define the multiplicative function $f_{\mathcal{J}}$ at powers of irreducibles via

$$f_{\mathcal{J}}(P^k):=\begin{cases}f(P^k) & \text{if } \deg(P)\in\mathcal{J}\\ 0 & \text{otherwise.}\end{cases}$$

Let $N':=\lfloor N/2\rfloor$. By the hyperbola method,

$$\sum_{G\in\mathcal{M}_N}\frac{f(G)\overline{\widetilde{\chi}}(G)}{1+\omega_{\mathcal{I}}(G)}=\sum_{AB\in\mathcal{M}_N}\frac{f_{\mathcal{I}}(A)f_{\mathcal{I}^c}(B)\overline{\widetilde{\chi}}(AB)}{1+\omega_{\mathcal{I}}(A)}$$
$$=\sum_{A\in\mathcal{M}_{\leqslant N'}}\frac{f_{\mathcal{I}}(A)\overline{\widetilde{\chi}}(A)}{1+\omega_{\mathcal{I}}(A)}\sum_{B\in\mathcal{M}_{N-\deg(A)}}f_{\mathcal{I}^c}(B)\overline{\widetilde{\chi}}(B)+\sum_{B\in\mathcal{M}_{\leqslant N-N'}}f_{\mathcal{I}^c}(B)\overline{\widetilde{\chi}}(B)\sum_{A\in\mathcal{M}_{N-\deg(B)}}\frac{f_{\mathcal{I}}(A)\overline{\widetilde{\chi}}(A)}{1+\omega_{\mathcal{I}}(A)}$$
$$-\left(\sum_{A\in\mathcal{M}_{N'}}\frac{f_{\mathcal{I}}(A)\overline{\widetilde{\chi}}(A)}{1+\omega_{\mathcal{I}}(A)}\right)\left(\sum_{B\in\mathcal{M}_{N-N'}}f_{\mathcal{I}^c}(B)\overline{\widetilde{\chi}}(B)\right)=:T_1+T_2-T_3.$$

We first treat T_1. Let $0\leqslant K\leqslant N'$. Since $\widetilde{\chi}\nsim\widetilde{\chi_1}$, Lemma 3.6 implies that

$$\mathcal{D}_{f_{\mathcal{I}^c}\overline{\widetilde{\chi}}}(N-K)=\mathcal{D}_{f_{\mathcal{I}^c}\overline{\widetilde{\chi}}}(N)-O(1)\geqslant\mathcal{D}_{f\overline{\widetilde{\chi}}}(N)-2\sum_{P\leqslant d\leqslant Q}q^{-d}|\mathcal{P}_d|$$
$$\geqslant\left(\frac{1}{4}-o(1)\right)\log N-2\log(Q/P).$$

Combining this with Theorem 3.5, we obtain

$$\sum_{G\in\mathcal{M}_{N-K}} f_{I^c}(G)\overline{\widetilde{\chi}}(G) \ll q^{N-K}\mathcal{D}_{f_{I^c}\overline{\widetilde{\chi}}}(N-K)\exp\left(-\mathcal{D}_{f_{I^c}\overline{\widetilde{\chi}}}(N-K)\right) \ll q^{N-K}(Q/P)^2N^{-1/4+o(1)}. \tag{26}$$

Applying this with $K=\deg(A)$ in T_1 and summing over $A\in\mathcal{M}_{\leqslant N'}$ yields

$$\begin{aligned} T_1 &\ll q^N(Q/P)^2N^{-1/4+o(1)} \sum_{\substack{A\in\mathcal{M}_{\leqslant N'}\\ R|A\Rightarrow \deg(R)\in[P,Q]}} q^{-\deg(A)} \\ &\ll q^N(Q/P)^2N^{-1/4+o(1)}\exp\left(\sum_{\substack{R\in\mathcal{P}\\ P\leqslant\deg(R)\leqslant Q}} q^{-\deg(R)}\right) \\ &\ll q^N(Q/P)^3N^{-1/4+o(1)}. \end{aligned}$$

We next consider T_2. Using Lemma 4.11, for every $0\leqslant K\leqslant N-N'\leqslant N/2+1$, we have

$$\left|\sum_{A\in\mathcal{M}_{N-K}}\frac{f_I(A)}{1+\omega_I(A)}\right| \leqslant |\{A\in\mathcal{M}_{N-K}:\ R|A, R\in\mathcal{P}\Rightarrow\deg(R)\leqslant Q\}| \ll q^{N-K}\exp(-c(N-K)/Q) \tag{27}$$

$$\ll q^{N-K}\exp\left(-c\frac{N}{3Q}\right) \tag{28}$$

for some $c>0$. Applying this with $K=\deg(B)$, then summing over B in T_2 yields

$$\begin{aligned} T_2 &\ll \sum_{B\in\mathcal{M}_{\leqslant N-N'}}\left|\sum_{A\in\mathcal{M}_{N-\deg(B)}}\frac{f_I(A)}{1+\omega_I(A)}\right| \ll q^N\exp\left(-c\frac{N}{3Q}\right)\sum_{B\in\mathcal{M}_{\leqslant N-N'}} q^{-\deg(B)} \\ &\ll q^NN\exp\left(-c\frac{N}{3Q}\right). \end{aligned}$$

Finally, consider T_3. Using the estimates (26) and (27) (with $N-K$ replaced by N' and $N-N'$, respectively) yields

$$\begin{aligned} T_3 &\ll q^NN\exp\left(-c\frac{N}{6Q}\right)(Q/P)^2N^{-1/4+o(1)} \\ &\ll q^NN^3\exp\left(-c\frac{N}{6Q}\right). \end{aligned}$$

Combining the estimates for T_1, T_2, and T_3 and noting that $N^3\exp(-c\frac{N}{6Q})\ll N^{-100}$ establishes the claim. □

5 | VARIANCE OF MULTIPLICATIVE FUNCTIONS IN PROGRESSIONS TO LARGE DEGREE MODULI

In this section, we will prove Theorem 1.2. In the next section, we will apply a very similar argument to deduce the Matomäki–Radziwiłł type theorem that we shall need.

Let $1 \leqslant H \leqslant N - N^{3/4}$ with $H = H(N) \to \infty$ as $N \to \infty$. Let $f : \mathcal{M} \to \mathbb{U}$ be multiplicative, and let $Q \in \mathcal{M}_{N-H}$. Let $\chi_1 \in \mathcal{X}_Q$ be the Dirichlet character mod Q that minimizes $\chi \mapsto \mathcal{D}_{f\overline{\chi}}(N)$. By orthogonality,

$$\sum_{A \bmod Q}^{*} \left| \sum_{\substack{G \in \mathcal{M}_N \\ G \equiv A \bmod Q}} f(G) - \frac{\chi_1(A)}{\phi(Q)} \sum_{G \in \mathcal{M}_N} f(G)\overline{\chi_1}(G) \right|^2 = \frac{1}{\phi(Q)} \sum_{\chi \neq \chi_1} \left| \sum_{G \in \mathcal{M}_N} f(G)\overline{\chi}(G) \right|^2. \tag{29}$$

Let $\eta \in (0, 1/6)$ be fixed, and set

$$Q_1 := \min\{H, N^{1/5}\}, \quad \text{and} \quad P_1 := \frac{400}{\eta \log q} \log Q_1.$$

Fix $J \geqslant 1$ to be the least integer such that $J^{4J+2} Q_1^J \geqslant N^{1/2}$, and if $J \geqslant 2$ set

$$P_j := j^{4j} Q_1^{j-1} P_1 \quad \text{and} \quad Q_j := j^{4j+2} Q_1^j$$

for each $2 \leqslant j \leqslant J$. We define $S_{\boldsymbol{P},\boldsymbol{Q}}$ as in Definition 4.9 with these collections of parameters P_j and Q_j, and for $1 \leqslant j \leqslant J$, we let $S^{(j)}_{\boldsymbol{P},\boldsymbol{Q}}$ denote the set of $G \in \mathcal{M}$ with an irreducible factor R with $\deg(R) \in [P_i, Q_i]$ for all $i \neq j$.

For each j, $1 \leqslant d \leqslant N$ and a character χ modulo Q, set

$$A_{j,d}(\chi) := \frac{1_{[P_j,Q_j]}(d)}{d|\mathcal{P}_d|} \sum_{R \in \mathcal{P}_d} f(R)\overline{\chi}(R).$$

$$B_{j,d}(\chi) := \frac{1}{|\mathcal{M}_d|} \sum_{\substack{D \in \mathcal{M}_d \\ D \in S^{(j)}_{\boldsymbol{P},\boldsymbol{Q}}}} \frac{f(D)\overline{\chi}(D)}{1 + \omega_{[P_j,Q_j]}(D)}.$$

Thus, $A_{j,d}(\chi) = 0$ except when $d \in [P_j, Q_j]$. Following [27] (see also [23, Section 5.2]), we split the set $\Xi := \mathcal{X}_Q \backslash \{\chi_1\}$ into the following sets.

Definition 5.1. For $j \geqslant 1$ put $\beta_j := \frac{1}{4} - \frac{\eta}{2}(1 + 1/j)$. Define

$$\mathcal{X}_1 := \{\chi \in \Xi : |A_{1,d}(\chi)| \leqslant q^{-\beta_1 d} \ \forall P_1 \leqslant d \leqslant Q_1\}$$

$$\mathcal{X}_j := \{\chi \in \Xi : |A_{j,d}(\chi)| \leqslant q^{-\beta_j d} \ \forall P_j \leqslant d \leqslant Q_j\} \backslash \bigcup_{1 \leqslant i \leqslant j-1} \mathcal{X}_i \ \ (2 \leqslant j \leqslant J)$$

$$\mathcal{U} := \Xi \backslash \bigcup_{1 \leqslant j \leqslant J} \mathcal{X}_j.$$

We shall bound the contribution of the characters from each of $\mathcal{X}_j$ and $\mathcal{U}$ using the lemmata from the previous sections.

We begin by estimating the contribution from $\chi \notin \mathcal{U}$. By Lemma 4.10, we have

$$\frac{1}{\phi(Q)} \sum_{\substack{\chi \neq \chi_1 \\ \chi \notin \mathcal{U}}} \left| \sum_{G \in \mathcal{M}_N} f(G)\overline{\chi}(G) \right|^2 \ll \frac{1}{\phi(Q)} \sum_{\substack{\chi \neq \chi_1 \\ \chi \notin \mathcal{U}}} \left| \sum_{\substack{G \in \mathcal{M}_N \\ G \in S_{P,Q}}} f(G)\overline{\chi}(G) \right|^2 + q^{2N-\deg(Q)} \sum_{1 \leqslant j \leqslant J} \frac{P_j}{Q_j}$$

$$\ll \frac{1}{\phi(Q)} \sum_{\substack{\chi \neq \chi_1 \\ \chi \notin \mathcal{U}}} \left| \sum_{\substack{G \in \mathcal{M}_N \\ G \in S_{P,Q}}} f(G)\overline{\chi}(G) \right|^2 + q^{2N-\deg(Q)} \frac{P_1}{Q_1}. \tag{30}$$

For each $\chi \in \Xi \backslash \mathcal{U}$, write

$$F(\chi) := \frac{1}{|\mathcal{M}_N|} \sum_{\substack{G \in \mathcal{M}_N \\ G \in S_{P,Q}}} f(G)\overline{\chi}(G).$$

We apply Lemma 4.15 for each $1 \leqslant j \leqslant J$ (with $P = P_j$ and $Q = Q_j$ in the notation there) to get

$$\sum_{\chi \in \mathcal{X}_j} |F(\chi)|^2 \ll (Q_j - P_j + 1) \sum_{P_j \leqslant d \leqslant Q_j} \sum_{\chi \in \mathcal{X}_j} |A_{j,d}(\chi)|^2 |B_{j,N-d}(\chi)|^2 + \phi(Q) q^{-\deg(Q)} q^{-P_j}$$

$$=: \mathfrak{M}_j + \mathfrak{R}_j.$$

Summing the error terms arising from the terms $1 \leqslant j \leqslant J$ yields

$$\sum_{1 \leqslant j \leqslant J} \mathfrak{R}_j \ll \phi(Q) q^{-\deg(Q)} \cdot q^{-P_1} \ll Q_1^{-100} \phi(Q)/q^{\deg(Q)}, \tag{31}$$

using the definitions of P_j and Q_j above. We thus focus on the main terms arising in the above estimate.

Case 1: $j = 1$

In this case, we bound $|A_{1,d}(\chi)| \leqslant q^{-\beta_1 d}$ for each $\chi \in \mathcal{X}_1$ and then apply Lemma 4.2 to get

$$\mathfrak{M}_1 \ll (Q_1 - P_1 + 1) \sum_{P_1 \leqslant d \leqslant Q_1} q^{-2\beta_1 d} \sum_{\chi \bmod Q} |B_{1,N-d}(\chi)|^2$$

$$\leqslant Q_1 \left(\phi(Q) q^{Q_1 - N} + \phi(Q) q^{-\deg(Q)} \right) \sum_{P_1 \leqslant d \leqslant Q_1} q^{-2\beta_1 d}$$

$$\ll \phi(Q) q^{-\deg(Q)} \cdot Q_1 q^{-2\beta_1 P_1} \ll \phi(Q) q^{-\deg(Q)} \cdot Q_1 q^{-P_1/6},$$

since $\beta_1 = 1/4 - \eta \geqslant 1/12$. Thus,

$$\mathfrak{M}_1 \ll Q_1^{-100} \phi(Q)/q^{\deg(Q)}. \tag{32}$$

Case 2: $2 \leqslant j \leqslant J$

We know that for each $\chi \in \mathcal{X}_j$, we can find a $d_\chi \in [P_{j-1}, Q_{j-1}]$ for which $|A_{j-1,d_\chi}(\chi)| \geqslant q^{-\beta_{j-1}d_\chi}$. Thus, similarly as in [27], we can estimate

$$\begin{aligned}\mathfrak{M}_j &\ll (Q_j - P_j + 1) \sum_{P_{j-1} \leqslant r \leqslant Q_{j-1}} \sum_{\substack{\chi \in \mathcal{X}_j \\ d_\chi = r}} \sum_{P_j \leqslant d \leqslant Q_j} |A_{j,d}(\chi)|^2 |B_{j,N-d}(\chi)|^2 \\ &\leqslant (Q_j - P_j + 1)(Q_{j-1} - P_{j-1} + 1) \max_{P_{j-1} \leqslant r \leqslant Q_{j-1}} \sum_{\substack{\chi \in \mathcal{X}_j \\ d_\chi = r}} \sum_{P_j \leqslant d \leqslant Q_j} q^{-2\beta_j d} |B_{j,N-d}(\chi)|^2 \\ &\leqslant Q_j^2 \sum_{P_j \leqslant d \leqslant Q_j} q^{-2\beta_j d} q^{2\ell_d r_0 \beta_{j-1}} \sum_{\chi \bmod Q} |A_{j-1,r_0}(\chi)^{\ell_d} B_{j,N-d}(\chi)|^2,\end{aligned}$$

for some $r_0 \in [P_{j-1}, Q_{j-1}]$, with $\ell_d := \lceil d/r_0 \rceil$. Applying Lemma 4.7, we have

$$\sum_{\chi \bmod Q} |A_{j-1,r_0}(\chi)^{\ell_d} B_{j,N-d}(\chi)|^2 \ll \phi(Q) q^{-\deg(Q)} \ell_d^{2\ell_d}.$$

Combining this with the estimates from the previous line, we get

$$\mathfrak{M}_j \ll \phi(Q) q^{-\deg(Q)} \cdot Q_j^2 \sum_{P_j \leqslant d \leqslant Q_j} q^{2(\ell_d r_0 \beta_{j-1} - d\beta_j)} \ell_d^{2\ell_d}.$$

By definition, $\ell_d \leqslant d/r_0 + 1$, so that since $r_0 \leqslant Q_{j-1}$,

$$\ell_d r_0 \beta_{j-1} - d\beta_j \leqslant d(\beta_{j-1} - \beta_j) + r_0 \beta_{j-1} \leqslant -\frac{\eta d}{2j^2} + Q_{j-1}\beta_{j-1}.$$

Furthermore, we have

$$\ell_d \log \ell_d \leqslant \frac{d \log d}{r_0} + \log d \leqslant (\log Q_j)(d/P_{j-1} + 1).$$

We thus may bound $\mathfrak{M}_j$ as

$$\mathfrak{M}_j \ll \phi(Q) q^{-\deg(Q)} \cdot Q_j^4 q^{2Q_{j-1}\beta_{j-1}} \sum_{P_j \leqslant d \leqslant Q_j} q^{-2d(\eta/(2j^2) - (\log Q_j)/(P_{j-1} \log q))}.$$

We record the following easy-to-check bounds, contingent on Q_1 being sufficiently large and $j \geqslant 2$:

(i) $\frac{\log Q_j}{P_{j-1}} \leqslant \frac{j(\log Q_1)(1+5\log j)}{(j-1)^{4j-4} Q_1^{j-2} P_1} \leqslant \frac{50 \log Q_1}{P_1} \cdot \frac{1}{j^2} \leqslant \frac{\eta \log q}{8j^2}$.

(ii) $Q_j^4 q^{2\beta_{j-1}Q_{j-1}} \leqslant q^{Q_{j-1}/2}$.

(iii) $Q_{j-1} \leqslant j^{4j-2} Q_1^{j-1} \leqslant P_j/(j^2 P_1)$.

Using these bounds, we get

$$\mathfrak{M}_j \ll \phi(Q)q^{-\deg(Q)} \cdot q^{Q_{j-1}/2} \sum_{P_j \leqslant d \leqslant Q_j} q^{-\eta d/(2j^2)} \ll \phi(Q)q^{-\deg(Q)} \cdot j^2\eta^{-1}q^{P_j/(2j^2P_1)}q^{-\eta P_j/(2j^2)}$$

$$\ll_\eta \phi(Q)q^{-\deg(Q)} \cdot j^2 q^{-\eta P_j/(4j^2)} \leqslant \phi(Q)q^{-\deg(Q)} \cdot j^{-2}q^{-\eta Q_1 P_1/4}.$$

Summing over $2 \leqslant j \leqslant J$, we get

$$\sum_{2\leqslant j\leqslant J} \mathfrak{M}_j \ll_\eta \phi(Q)q^{-\deg(Q)} \cdot q^{-\eta Q_1 P_1/4} \ll Q_1^{-100}\phi(Q)q^{-\deg(Q)}. \tag{33}$$

Case 3: $\mathcal{U}$

We now treat the remaining characters $\chi \in \mathcal{U}$. We make an additional choice of parameters $\tilde{P} := N^{2/3}$, $\tilde{Q} := N^{13/18}$. Combining Lemma 4.10 with Lemma 4.15, we find $\tilde{P} \leqslant d_0 \leqslant \tilde{Q}$ such that

$$\sum_{\chi\in\mathcal{U}} \left| \sum_{G\in\mathcal{M}_N} f(G)\overline{\chi}(G) \right|^2$$

$$\ll \sum_{\chi\in\mathcal{U}} \left| \sum_{\substack{G\in\mathcal{M}_N \\ G\in S_{\tilde{P},\tilde{Q}}}} f(G)\overline{\chi}(G) \right|^2 + \phi(Q)q^{-\deg(Q)}\frac{\tilde{P}}{\tilde{Q}}$$

$$\ll (\tilde{Q}-\tilde{P}+1) \sum_{\tilde{P}\leqslant d\leqslant\tilde{Q}} \sum_{\chi\in\mathcal{U}} |A_d(\chi)|^2|B_{N-d}(\chi)|^2 + \phi(Q)q^{-\deg(Q)}\left(q^{-\tilde{P}} + \frac{\tilde{P}}{\tilde{Q}}\right)$$

$$\ll \tilde{Q}^2 \sum_{\chi\in\mathcal{U}} |A_{d_0}(\chi)|^2|B_{N-d_0}(\chi)|^2 + \phi(Q)q^{-\deg(Q)}\left(q^{-\tilde{P}} + \frac{\tilde{P}}{\tilde{Q}}\right),$$

where A_d and B_{N-d}, $\tilde{P} \leqslant d \leqslant \tilde{Q}$, are defined as in Lemma 4.15 with respect to the parameters $\tilde{P}$ and $\tilde{Q}$. We now split the set $\mathcal{U}$ further. Following [27], we define the subsets

$$\mathcal{U}_S := \{\chi \in \mathcal{U} : |A_{d_0}(\chi)| \leqslant N^{-10}\},$$

$$\mathcal{U}_L := \{\chi \in \mathcal{U} : |A_{d_0}(\chi)| > N^{-10}\}.$$

We begin by treating the contribution from $\mathcal{U}_S$. By Lemma 4.4 and the fact that $d_0 \leqslant \tilde{Q} = o(N)$, we have

$$\sum_{\chi\in\mathcal{U}_S} |A_{d_0}(\chi)|^2|B_{N-d_0}(\chi)|^2 \leqslant N^{-20} \sum_{\chi\in\mathcal{U}_S} |B_{N-d_0}(\chi)|^2$$

$$\ll N^{-20}\phi(Q)q^{-\deg(Q)}\left(1 + |\mathcal{U}_S|q^{(1/2+o(1))N-N+d_0}\right)$$

$$\ll N^{-20}\phi(Q)q^{-\deg(Q)}\left(1 + |\mathcal{U}|q^{(-1/2+o(1))N}\right). \tag{34}$$

To estimate the size of $\mathcal{U}$, we note that whenever $\chi \in \mathcal{U}$, there is $d_J \in [P_J, Q_J]$ such that $|A_{J,d_J}(\chi)| \geqslant q^{-\beta_J d_J}$, and thus

$$|\mathcal{U}| \leqslant \left| \bigcup_{P_J \leqslant d_J \leqslant Q_J} \{\chi \bmod Q : |A_{J,d_J}(\chi)| \geqslant q^{-\beta_J d_J}\} \right| \leqslant Q_J \max_{P_J \leqslant d_J \leqslant Q_J} |\{\chi \bmod Q : |A_{J,d_J}(\chi)| \geqslant q^{-\beta_J d_J}\}|.$$

By choice, we have that $N^{1/2} \leqslant Q_J \ll N^{1/2} J^4 Q_1 \ll N^{7/10+o(1)}$, so that from (i) above we have $P_J \geqslant \frac{8J^2}{\eta \log q} \log Q_{J+1} \geqslant \frac{4}{\eta \log q} \log N$ and $\phi(Q) \geqslant q^{Q_J} \geqslant q^{d_J}$ for all $d_J \in [P_J, Q_J]$. Hence, $(\log\log \phi(Q))/P_J < \frac{1}{2}\eta \log q$ for N large enough, and Lemma 4.6 may be applied to give

$$|\mathcal{U}| \ll Q_J \exp\left(\frac{\log(q^{d_J}\phi(Q))}{d_J \log q} \left(2\log\left(\frac{\log(2\phi(Q))}{d_J \log q} \right) + \log(2q^{2\beta_J d_J}/d_J) \right) \right) \leqslant \phi(Q)^{\frac{1}{2}-\eta/2}.$$

Inserting this into (34), the off-diagonal term becomes $O(q^{(-\eta/2+o(1))N}) = o(1)$. Thus, we find that

$$\sum_{\chi \in \mathcal{U}_S} |A_{d_0}(\chi)|^2 |B_{N-d_0}(\chi)|^2 \ll \phi(Q) q^{-\deg(Q)} \cdot N^{-20}.$$

We now consider the contribution from $\mathcal{U}_L$. Since $\chi_1 \notin \mathcal{U}$, and since $2(N-d_0) > N$, we may apply Lemma 4.16 to obtain

$$B_{N-d_0}(\chi) \ll (\tilde{Q}/\tilde{P})^3 (N-d_0)^{-1/4+o(1)} \ll \phi(Q) q^{-\deg(Q)} \cdot (\tilde{Q}/\tilde{P})^3 N^{-1/4+o(1)},$$

since $d_0 \leqslant \tilde{Q} = o(N)$ and $\phi(Q) q^{-\deg(Q)} \gg (\log N)^{-1}$. It follows that

$$\sum_{\chi \in \mathcal{U}_L} |A_{d_0}(\chi)|^2 |B_{N-d_0}(\chi)|^2 \ll (\phi(Q) q^{-\deg(Q)})^2 (\tilde{Q}/\tilde{P})^6 N^{-1/2+o(1)} \sum_{\chi \in \mathcal{U}_L} |A_{d_0}(\chi)|^2.$$

Applying Lemma 4.5, we deduce that

$$\sum_{\chi \in \mathcal{U}_L} |A_{d_0}(\chi)|^2 \ll \frac{1}{d_0^2} \left(1 + \deg(Q) q^{-d_0/2} |\mathcal{U}_L| \right) \ll \tilde{P}^{-2} \left(1 + \deg(Q) |\mathcal{U}_L| q^{-\tilde{P}/2} \right).$$

Since $\phi(Q) \geqslant q^{(1-o(1))N^{3/4}} \geqslant q^{\tilde{Q}} \geqslant q^{d_0}$, we may appeal once again to Lemma 4.6, this time with $Z = N^{10}$, getting

$$|\mathcal{U}_L| \ll \exp\left(\frac{\log(q^{d_0}\phi(Q))}{d_0 \log q} \left(2\log\left(\frac{\log(2\phi(Q))}{d_0 \log q} \right) + \log(2N^{19}) \right) \right) \leqslant e^{N^{1+o(1)}/\tilde{P}}.$$

Inserting this into the previous bound and using the fact that $\tilde{P} = N^{2/3} \geqslant 2N^{1.01}/\tilde{P}$ yields

$$\sum_{\chi \in \mathcal{U}_L} |A_{d_0}(\chi)|^2 \ll \tilde{P}^{-2} + \deg(Q) q^{-\tilde{P}/2} e^{N^{1+o(1)}/\tilde{P}} \ll \tilde{P}^{-2}.$$

It follows that

$$\sum_{\chi\in\mathcal{U}_L}|A_{d_0}(\chi)|^2|B_{N-d_0}(\chi)|^2 \ll \left(\phi(Q)q^{-\deg(Q)}\right)^2\cdot N^{-1/2+o(1)}\tilde{Q}^6\tilde{P}^{-8}.$$

Combined with the bounds for $\mathcal{U}_S$, we get

$$\begin{aligned}\sum_{\chi\in\mathcal{U}}|F(\chi)|^2 &\ll (\phi(Q)q^{-\deg(Q)})^2\cdot\left(\tilde{Q}^2\left(N^{-18}+\tilde{Q}^6\tilde{P}^{-8}N^{-1/2+o(1)}\right)+\tilde{P}/\tilde{Q}+q^{-\tilde{P}}\right)\\ &\ll \phi(Q)q^{-\deg(Q)}\cdot\left(N^{-1/2+o(1)}\left(\tilde{Q}/\tilde{P}\right)^8+\tilde{P}/\tilde{Q}\right)\\ &\ll N^{-1/18+o(1)}\phi(Q)q^{-\deg(Q)}. \end{aligned}\tag{35}$$

Finally, putting (31), (32), (33), and (35) together with (29) and (30), Theorem 1.2 follows.

Remark 5.2. Note that if we began by assuming that the sums in the variable G in (29) are supported on $S_{\boldsymbol{P},Q}$, then the same proof would give the sharper estimate

$$\begin{aligned}&\sum_{A \bmod Q}^{*}\left|\sum_{\substack{G\in\mathcal{M}_N\\ G\equiv A \bmod Q}} f1_{S_{\boldsymbol{P},Q}}(G)-\frac{\chi_1(A)}{\phi(Q)}\sum_{G\in\mathcal{M}_N} f1_{S_{\boldsymbol{P},Q}}\overline{\chi_1}(G)\right|^2\\ &\ll q^{2N-\deg(Q)}\left(Q_1q^{-P_1/6}+N^{-1/18+o(1)}\right).\end{aligned}\tag{36}$$

We will use this sharper version of the theorem in Section 7.

6 | MATOMÄKI–RADZIWIŁŁ THEOREM IN FUNCTION FIELDS

In this section, we prove the following analog of the main result in [27] (see Section 3.3 for the definition of g^*, for g multiplicative).

Theorem 6.1 (Matomäki–Radziwiłł Theorem in Function Fields). *Let $f:\mathcal{M}\to\mathbb{U}$ be a multiplicative function and let $1\leqslant H\leqslant N-N^{3/4}$, with $H=H(N)$ tending to infinity with N. Then,*

$$\begin{aligned}&\frac{1}{|\mathcal{M}_N|}\sum_{G_0\in\mathcal{M}_N}\left|\frac{1}{|\mathcal{M}_{<H}|}\sum_{\substack{G\in\mathcal{M}_N\\ G\in I_H(G_0)}} f(G)-\frac{1}{|\mathcal{M}_N|}\sum_{G\in\mathcal{M}_N} f(G)\overline{\chi_1^*}(G)\right|^2\\ &\ll (\log H)/H+N^{-1/18+o(1)},\end{aligned}\tag{37}$$

where χ_1 is the Dirichlet character modulo t^{N-H+1} that minimizes the map $\chi\mapsto\mathcal{D}_{f\overline{\chi^*}}(N)$.

Theorem 1.1 will follow as a special case, as we will see later in this section.

Remark 6.2. In light of Remark 5.2, if we replace f by $f1_{S_{P,Q}}$, with the choice of parameters $P_1, Q_1, P_j = j^{4j}P_1Q_1^{j-1}$, and $Q_j = j^{4j+2}Q_1^j$ for $j \geqslant 2$, then the bound in Theorem 6.1 improves to $\ll Q_1^{1/2}q^{-P_1/12} + N^{-1/18+o(1)}$. The additional flexibility in choosing P_1 and Q_1 will be used in the next section.

Remark 6.3. We can obtain the same estimate as in Theorem 6.1 if (37) is replaced with

$$\frac{1}{|\mathcal{M}_N|}\sum_{G_0\in\mathcal{M}_N}\left|\frac{1}{|\mathcal{M}_{<H}|}\sum_{\substack{G\in\mathcal{M}_N\\ G\in I_H(G_0)}} f(G) - \frac{1}{|\mathcal{M}_N|}\sum_{G\in\mathcal{M}_N} f(G) - \frac{1_{\chi_1\neq\chi_0}}{|\mathcal{M}_N|}\sum_{G\in\mathcal{M}_N} f(G)\overline{\chi_1^*}(G)\right|^2.$$

The proof is the same, and in fact it will be clear from the application of the orthogonality relations in the proof that this quantity is never larger than (37).

Proof of Theorem 6.1. The proof is very similar to the proof of Theorem 1.2, just with a different set of characters.

By the orthogonality relation (22), we can write (37) as

$$\frac{1}{q^{2N}}\sum_{\substack{\xi\in\mathcal{X}_{1,N-H}\\ \xi\neq\chi_1^*}}\left|\sum_{G\in\mathcal{M}_N} f(G)\overline{\xi}(G)\right|^2.$$

This is analogous to (29), just with a different group of characters (see also [7, (2.12)]). Now the rest of the proof follows precisely as the proof of Theorem 1.2 up to notation. Indeed, the only properties of the Dirichlet characters used in the proof of Theorem 1.2 were the lemmas from Sections 3 and 4. In Section 3, all the lemmas are readily stated for Hayes characters, which includes both short interval characters and Dirichlet characters as special cases. Also in Section 4, all the mean value estimates have perfect analogs for short interval characters, and the proofs are identical, as noted in Remark 4.1. Moreover, the pointwise bound offered by Lemma 4.16 is written for more general Hayes characters, and we can take $\widetilde{\chi_1}$ there to be the short interval character of length $N-H$ that minimizes $\tilde{\chi}\mapsto \mathcal{D}_{f\overline{\tilde{\chi}}}(N)$. Hence, all the steps in the proof of Theorem 1.2 work in the same way. □

To deduce the real-valued case of the Matomäki–Radziwiłł theorem from this, we will need the following variant of Corollary 3.7, applicable to real-valued multiplicative functions twisted by Dirichlet characters modulo powers t^m, $m \geqslant 2$.

Lemma 6.4 (Sup norm estimate for weighted Dirichlet character sums). *Assume* $\mathrm{char}(\mathbb{F}_q) \neq 2$. *Let* $N \geqslant 1$ *and* $2 \leqslant k \leqslant N$. *Let* $f : \mathcal{M} \to [-1,1]$ *be multiplicative. Let* $2 \leqslant k \leqslant N$. *Then,*

$$\max_{\substack{\chi \bmod t^k\\ \chi\neq\chi_0}}\frac{1}{q^N}\left|\sum_{G\in\mathcal{M}_N} f(G)\chi(G)\right| \ll N^{-1/4+o(1)}.$$

Proof. In light of Theorem 3.5, it suffices to show that

$$\min_{\substack{\chi \bmod t^k \\ \chi \neq \chi_0}} \mathcal{D}_{f\overline{\chi}}(N) \geqslant (1/4 - o(1)) \log N.$$

Note that if χ is a character modulo t^k and χ' is the primitive character inducing χ, then

$$\mathcal{D}_{f\overline{\chi}}(N) = \mathcal{D}_{f\overline{\chi'}}(N) + O(1),$$

so it suffices to consider primitive characters modulo t^k. Now, suppose χ is a primitive character that is *not* real. Since χ^2 is not principal, arguing precisely as in the proof of Lemma 3.6, we obtain

$$\mathbb{D}(f, \chi e_\theta; N)^2 \geqslant (1/4 - o(1)) \log N$$

(for instance, when f takes values in S^1 the triangle inequality immediately yields $\mathbb{D}(f, \chi e_\theta; N) \geqslant \frac{1}{2}\mathbb{D}(1, \chi^2 e_{2\theta}; N)$, and the general case follows from this as in the proof of Lemma 3.6). Thus, we may conclude that for any *primitive non-quadratic* Dirichlet character modulo t^k, we have

$$\mathcal{D}_{f\overline{\chi}}(N) \geqslant \left(\frac{1}{4} - o(1)\right) \log N.$$

Furthermore, it is easy to see that there are *no* primitive quadratic (non-principal) characters modulo t^k for any $k \geqslant 2$. Indeed, suppose χ is real and primitive modulo t^k. Then, χ cannot be periodic modulo t^j, for any $j < k$. To deduce a contradiction from this, set now $m := \lceil k/2 \rceil < k$. Since $q \geqslant 3$ is odd and χ is real, we have $\chi^q = \chi$, and also $mq > k$. Thus, for any $A, B \in \mathbb{F}_q[t]$, we have

$$\chi(B + At^m) = \chi^q(B + At^m) = \chi((B + At^m)^q) = \chi(B^q + qAB^{q-1}t^m) = \chi(B)^q = \chi(B).$$

Thus, in fact, χ is periodic modulo t^m, contradicting the fact that χ is primitive modulo t^k.

Therefore, we obtain

$$\min_{\substack{\chi \bmod t^k \\ \chi \neq \chi_0}} \mathcal{D}_{f\overline{\chi}}(N) = \min_{\substack{\chi \bmod t^k \\ \chi^2 \neq \chi_0 \\ \chi \text{ primitive}}} \mathcal{D}_{f\overline{\chi}}(N) + O(1) \geqslant (1/4 - o(1)) \log N,$$

as claimed. □

Proof of Theorem 1.1. Assume that $f : \mathcal{M} \to [-1, 1]$. We extend f to a map on $\mathbb{F}_q[t]$ by requiring that $f(c) = \chi_1^*(c)$ for all $c \in \mathbb{F}_q^\times$, where χ_1 is given by Theorem 6.1.

Theorem 1.1 follows immediately from Theorem 6.1 when q is even, aside from the claim that χ_1 may be assumed to be real. To see this, note that by Lemma 6.4, if χ_1 were not real since then the mean value of $f\overline{\chi_1}$ would contribute $\ll N^{-1/4+o(1)}$, which is anyway dwarfed by the error term in the statement of the theorem.

When q is odd, it suffices to show that if $\chi_1 \neq \chi_0$, then $\mathcal{D}_{f\overline{\chi_1^*}}(N) \geqslant (1/4 - o(1)) \log N$, so that once again the sum in $f\overline{\chi_1^*}$ contributes negligibly.

Since $f(c) = \chi_1^*(c)$ on $\mathbb{F}_q^\times$ and f^* is necessarily real, Lemma 3.10 combines with Lemma 6.4 to show that

$$\mathcal{D}_{f\overline{\chi_1}^*}(N) = \mathcal{D}_{f^*\overline{\chi_1}}(N) + O(1) \geqslant \left(\frac{1}{4} - o(1)\right)\log N.$$

This completes the proof. □

7 | SHORT EXPONENTIAL SUMS OF NON-PRETENTIOUS FUNCTIONS

In this section, we apply the results of the previous section to derive two function field analogs of estimates for short exponential sums weighted by a multiplicative function, due to Matomäki, Radziwiłł, and Tao [29]. To explain the formulation of our results, we begin by recording some of the relevant definitions. See [26] for an excellent reference to the definitions given here.

We write $\mathbb{F}_q(t)$ to denote the field of rational functions of t over $\mathbb{F}_q$. This comes equipped with the non-Archimedean valuation $\langle\cdot\rangle$ such that if $G = \sum_{j=N}^{\infty} a_{-j}t^{-j}$ for an integer N with $a_N \neq 0$, then $\langle G\rangle = q^{-N}$. The completion of $\mathbb{F}_q(t) = \mathbb{F}_q(1/t)$ with respect to this valuation is the set $\mathbb{K}_\infty(t) := \mathbb{F}_q((1/t))$ of formal Laurent series in $1/t$ with a finite number of non-negative power terms. We define $\mathbb{T}$ to be the unit ball of $\mathbb{K}_\infty(t)$ with respect to $\langle\cdot\rangle$, that is,

$$\mathbb{T} := \{\alpha \in \mathbb{K}_\infty(t) : \langle\alpha\rangle < 1\} \cong \mathbb{K}_\infty(t)/\mathbb{F}_q[t].$$

That is, $\mathbb{T}$ is the set of formal power series in $1/t$. This set forms a compact abelian group under addition, and thus comes equipped with a normalized Haar measure, which we shall denote by $d\alpha$. The Pontryagin dual group consists of the characters $\{\alpha \mapsto e_{\mathbb{F}}(G\alpha)\}_{G\in\mathbb{F}_q[t]}$, where, given $\alpha \in \mathbb{K}_\infty(t)$, we have written

$$e_{\mathbb{F}}(\alpha) := e\left(\frac{\mathrm{tr}_{\mathbb{F}_q/\mathbb{F}_p}(a_{-1}(\alpha))}{p}\right),$$

writing $a_{-1}(\alpha)$ to denote the coefficient of the term t^{-1} in the expansion of α. An important feature of these characters is that

$$\int_{\mathbb{T}} e_{\mathbb{F}}(G\alpha)d\alpha = \begin{cases} 1 & \text{if } G = 0 \\ 0 & \text{otherwise,} \end{cases}$$

in analogy to the orthogonality of additive characters on $\mathbb{R}/\mathbb{Z}$.

Our goal in this section will be to prove the following two results. The first is an estimate for exponential sums with multiplicative coefficients over short intervals that applies to complex-valued f, provided that f is *Hayes* non-pretentious (see Definition 3.14). The second result concerns such exponential sums with *real-valued* functions f, for which only the usual notion of (Dirichlet) non-pretentiousness needs to be assumed. The first of these theorems will be of relevance in proving the logarithmically averaged binary Chowla conjecture in this context.

In the theorems below, given $1 \leqslant H \leqslant N - N^{3/4}$ and $f : \mathcal{M} \to \mathbb{U}$ a multiplicative function, set

$$M_{\text{Hayes}}(f; N, H) := \min_{M \in \mathcal{M}_{\leqslant H}} \min_{\psi \bmod M} \min_{\substack{\xi \text{ short} \\ \text{len}(\xi) \leqslant N}} \mathcal{D}_{f\overline{\psi}\overline{\xi}}(N),$$

$$M_{\text{Dir}}(f; N, H) := \min_{M \in \mathcal{M}_{\leqslant H}} \min_{\psi \bmod M} \mathcal{D}_{f\overline{\psi}}(N).$$

It is clear from the definitions that $M_{\text{Hayes}}(f; N, H) \leqslant M_{\text{Dir}}(f; N, H)$.

Theorem 7.1. *Let $1 \leqslant H \leqslant N - N^{3/4}$. Let $f : \mathcal{M} \to \mathbb{U}$ be multiplicative. Then,*

$$\sup_{\alpha \in \mathbb{T}} \frac{1}{|\mathcal{M}_N|} \sum_{G_0 \in \mathcal{M}_N} \frac{1}{|\mathcal{M}_{<H}|} \left| \sum_{\substack{G \in \mathcal{M}_N \\ G \in I_H(G_0)}} f(G) e_{\mathbb{F}}(G\alpha) \right| \ll \frac{\log H}{H} + N^{-1/(2000 \log q)} + M e^{-M/100},$$

where $M := 1 + M_{\text{Hayes}}(f; N, H)$.

Theorem 7.2. *Assume q is odd. Let $1 \leqslant H \leqslant N - N^{3/4}$. Let $f : \mathcal{M} \to [-1, 1]$ be a multiplicative function. Then,*

$$\sup_{\alpha \in \mathbb{T}} \frac{1}{|\mathcal{M}_N|} \sum_{G_0 \in \mathcal{M}_N} \frac{1}{|\mathcal{M}_{<H}|} \left| \sum_{\substack{G \in \mathcal{M}_N \\ G \in I_H(G_0)}} f(G) e_{\mathbb{F}}(G\alpha) \right| \ll \frac{\log H}{H} + N^{-1/(2000 \log q)} + M e^{-M/100}, \quad (38)$$

where $M := 1 + M_{\text{Dir}}(f; N, H)$.

We will deduce both of these results from the following result about completely multiplicative functions.

Theorem 7.3. *Let $1 \leqslant H \leqslant N - N^{3/4}$. Let $f : \mathcal{M} \to \mathbb{U}$ be completely multiplicative. Then, Theorem 7.1 holds for f. Moreover, if f is real-valued and q is odd, then Theorem 7.2 holds for f.*

We will begin by proving Theorem 7.3; we will prove the deduction of Theorems 7.1 and 7.2 for general one-bounded multiplicative functions at the end of this section. The proofs of the complex and real cases begin the same way. We shall thus begin both simultaneously, then highlight where the differences arise below.

We proceed using the circle method, as in [29], splitting into cases according to whether α lies in a major or minor arc (to be defined momentarily). In the function field setting, arcs can be determined via the following form of Dirichlet's theorem.

Lemma 7.4 (Dirichlet's Theorem in Function Fields). *Suppose $\alpha \in \mathbb{T}$. Given $M \geqslant 1$, we can find $g \in \mathcal{M}_{\leqslant M}$ and $a \in \mathbb{F}_q[t]$ coprime to g with $\deg(a) < \deg(g)$ such that $\langle g\alpha - a \rangle \leqslant q^{-M}$.*

Proof. This follows from the pigeonhole principle, just as in the integer setting. □

We are now ready to embark on the proof of Theorem 7.3. We will first prove the following closely related statement.

Proposition 7.5. *Let $1 \leqslant H \leqslant N - N^{3/4}$. Let $f : \mathcal{M} \to \mathbb{U}$ be completely multiplicative. Suppose that*

$$10 \log H \leqslant \min\{(\log N)/(100 \log q), H/10, M_*(f; N, H)/100\}, \tag{39}$$

where $M_*(f; N, H) := 1 + M_{Dir}(f; N, H)$ if f is real-valued and q is odd, and $M_*(f; N, H) := 1 + M_{\text{Hayes}}(f; N, H)$ otherwise. Then,

$$\sup_{\alpha \in \mathbb{T}} \frac{1}{|\mathcal{M}_N|} \sum_{G_0 \in \mathcal{M}_N} \frac{1}{|\mathcal{M}_{<H}|} \left| \sum_{\substack{G \in \mathcal{M}_N \\ G \in I_H(G_0)}} f(G) e_{\mathbb{F}}(G\alpha) \right| \ll (\log H)/H + N^{-1/40} + e^{-M_{Hayes}(f;N,H)/20}.$$

Moreover, if f is real-valued and q is odd, we can replace M_{Hayes} with M_{Dir}.

By (39), we can choose $1 \leqslant W \leqslant X \leqslant H$ such that

$$10 \log H \leqslant W \leqslant \min\{(\log N)/(100 \log q), H/10, M_*(f; N, H)/10\},$$

and set $X = H - W$ (so that $W \leqslant X/2$). In general, we define arcs of the form

$$\mathfrak{M}_{a,g}(X) := \{\alpha \in \mathbb{T} : \langle g\alpha - a \rangle \leqslant q^{-X}\}.$$

The major arcs of length X and degree W are defined by

$$\mathfrak{M}(X, W) := \bigcup_{\deg(g) \leqslant W} \bigcup_{\substack{a \bmod g \\ (a,g)=1}} \mathfrak{M}_{a,g}(X),$$

and the minor arcs are then defined by

$$\mathfrak{m} = \mathfrak{m}(X, W) := \mathbb{T} \backslash \mathcal{M}(X, W).$$

Let $P_1 := 100W$ and $Q_1 := H/3$, and let $S = S_{\boldsymbol{P},\boldsymbol{Q}}$, with P_j, Q_j defined in terms of P_1 and Q_1 as in Section 5. For the same reason as in Section 5, it will be advantageous to replace the expression on the left-hand side in (38) by

$$\sup_{\alpha \in \mathbb{T}} \frac{1}{|\mathcal{M}_N|} \sum_{G_0 \in \mathcal{M}_N} \frac{1}{|\mathcal{M}_{<H}|} \left| \sum_{\substack{G \in \mathcal{M}_N \\ G \in I_H(G_0)}} f 1_S(G) e_{\mathbb{F}}(G\alpha) \right|. \tag{40}$$

By Lemma 4.8 and the triangle inequality, the difference between this latter expression and the one in (38) is $O(P_1/Q_1)$. We will thus focus our attention mostly on the estimation of (40).

As mentioned, the expression in (40) will be treated differently according to whether α lies in a major arc or a minor arc. We start with the minor arc case, where the argument has some resemblance to the derivation of the orthogonality criterion for multiplicative functions [2], and which can be derived independently of the results of the last two sections.

7.1 | The minor arcs

We fix $\alpha \in \mathfrak{m}$. In order to proceed in estimating (40), we shall need the following basic result.

Lemma 7.6. *Let $\alpha \in \mathbb{K}_\infty(t)$ and $H \geqslant 1$. Then,*

$$\sum_{deg(F)<H} e_{\mathbb{F}}(F\alpha) = q^H 1_{\langle \alpha \bmod 1\rangle \leqslant q^{-H-1}}.$$

Proof. This is standard, see, for example, [25, Lemma 7]. □

We will also need the following estimate, connected with Lemma 7.6.

Lemma 7.7. *Let $\alpha \in \mathfrak{M}_{a,g}(X)$, where $W < \deg(g) \leqslant X$ and $(a,g)=1$. Let $100W \leqslant k \leqslant H/3$. Then,*

$$|\{\deg(F) < k : \langle F\alpha \bmod 1\rangle < q^{-H+k-1}\}| \ll q^{k-W}.$$

Proof. Write $\beta := \alpha - a/g$. By assumption, we have $\langle \beta\rangle \leqslant q^{-X-\deg(g)}$. Since $(A,B) \mapsto \langle A-B\rangle$ is an ultrametric, for any $F \in \mathbb{F}_q[t]$, we have

$$\langle F\alpha \bmod 1\rangle \leqslant \max\{\langle Fa/g \bmod 1\rangle, \langle F\beta \bmod 1\rangle\},$$

with equality whenever the two valuations on the right-hand side differ.

Note that if $g \nmid F$, then as $(a,g)=1$, we can write $Fa = Mg + L$ with $L \not\equiv 0 \bmod g$. Hence, if $\deg(F) \leqslant k$, then

$$\langle Fa/g \bmod 1\rangle = q^{\deg(L)-\deg(g)} \geqslant q^{-\deg(g)} > q^{\deg(F)-X-\deg(g)} \geqslant \langle F\beta \bmod 1\rangle,$$

using $X \geqslant H/2 > k$. On the other hand, if $g|F$, then $\langle Fa/g \bmod 1\rangle = 0 \leqslant \langle F\beta \bmod 1\rangle$. In particular, we have

$$\langle F\alpha \bmod 1\rangle \begin{cases} = \langle Fa/g \bmod 1\rangle & \text{if } g \nmid F \\ \leqslant \langle F\beta \bmod 1\rangle & \text{if } g|F. \end{cases}$$

Let $\mathcal{E} := \{\deg(F) < k : \langle F\alpha \bmod 1\rangle < q^{-H+k-1}\}$. Consider separately the number of $F \in \mathcal{E}$ with $g \mid F$ and $g \nmid F$. Note that

$$|\{F \in \mathcal{E} : g|F\}| \leqslant |\{\deg(F) < k : g|F\}| \leqslant 1 + q^{k-\deg(g)} \ll q^{k-W}.$$

Next, consider the contribution to $\mathcal{E}$ from F that are not divisible by g. We first observe that there are no $F \in \mathcal{E}$ with $\deg(F) \geqslant \deg(g)$. Indeed, if such an F belonged to $\mathcal{E}$, then $\langle F\alpha \bmod 1\rangle = \langle Fa/g \bmod 1\rangle \geqslant q^{-\deg(g)}$. This implies the chain of inequalities

$$q^{-k} < q^{-\deg(F)} \leqslant q^{-\deg(g)} \leqslant \langle F\alpha \bmod 1\rangle < q^{-H+k-1},$$

which are conflicting since $k \leqslant H/3$.

We may therefore assume that $\deg(F) < \deg(g)$. Suppose next that $\deg(Fa) < \deg(g)$ as well. Then,

$$\langle F\alpha \bmod 1\rangle = \langle Fa/g \bmod 1\rangle = q^{\deg(Fa)-\deg(g)} \geqslant q^{\deg(F)-\deg(g)}.$$

Thus, if $F \in \mathcal{E}$, then we must have

$$\deg(F) \leqslant \deg(g) + k - H - 1 \leqslant k - H + X - 1 < k - W,$$

since $W = H - X$. Hence,

$$|\{F \in \mathcal{E} : \ g \nmid F, \text{ and } \deg(F) \geqslant \deg(g) \text{ or } \deg(Fa) < \deg(g)\}| \leqslant |\{\deg(F) < k - W\}| \ll q^{k-W}.$$

It remains to consider those F with $\deg(F) < \deg(g) \leqslant \deg(Fa)$. Observe that

$$\begin{aligned} &|\{F \in \mathcal{E} : \ g \nmid F, \deg(F) < \deg(g) \leqslant \deg(Fa)\}| \\ &= \sum_{0\leqslant m<\deg(g)} \sum_{\deg(B)=m} |\{F : \ \deg(g) - \deg(a) \leqslant \deg(F) < \min\{k, \deg(g)\} : \ Fa \equiv B \bmod g \\ &\qquad\qquad \text{and} \quad \langle Fa/g \bmod 1\rangle < q^{-H+k-1}\}|. \end{aligned}$$

Note that $\langle Fa/g \bmod 1\rangle = \langle B/g\rangle = q^{m-\deg(g)}$ whenever $\deg(B) = m$ and $Fa \equiv B \bmod g$, and so $F \in \mathcal{E}$ under these conditions only if $0 \leqslant m < \deg(g) - H + k - 1$. This condition is empty if $\deg(g) \leqslant H - k + 1$, so we assume otherwise (and hence $\deg(g) > k$). We may thus bound the above by

$$\leqslant \sum_{0\leqslant m<\deg(g)-H+k-1} \sum_{\deg(B)=m} |\{\deg(g) - \deg(a) \leqslant \deg(F) < k : \ F \equiv \overline{a}B \bmod g\}|,$$

where $\overline{a}$ is the inverse of $a \bmod g$. Since $\deg(F) < \deg(g)$, the cardinality above is $\leqslant 1$, and thus

$$\begin{aligned} |\{F \in \mathcal{E} : \ g \nmid F, \deg(F) < \deg(g) \leqslant \deg(Fa)\}| &\leqslant 1 + \sum_{0\leqslant m<\deg(g)-H+k-1} \sum_{\deg(B)=m} 1 \\ &\leqslant 1 + q \sum_{0\leqslant m<\deg(g)-H+k-1} q^m \ll 1 + q^{\deg(g)-H+k} \\ &\ll q^{X-H+k} = q^{k-W}. \end{aligned}$$

It follows that

$$|\mathcal{E}| \leqslant |\{F \in \mathcal{E} :\ g|F\}| + |\{F \in \mathcal{E} :\ g \nmid F, \deg(F) \geqslant \deg(g) \text{ or } \deg(Fa) < \deg(g)\}|$$
$$+ |\{F \in \mathcal{E} :\ \deg(F) < \deg(g) \leqslant \deg(Fa)\}| \ll q^{k-W},$$

as claimed. □

Let $\alpha \in \mathfrak{m}$. For each $G_0 \in \mathcal{M}_N$ let $\theta(G_0) \in S^1$ be chosen so as to write (40) as

$$\Sigma_S(\alpha) := \frac{1}{|\mathcal{M}_N|} \sum_{G_0 \in \mathcal{M}_N} \frac{\theta(G_0)}{|\mathcal{M}_{<H}|} \sum_{\substack{G \in \mathcal{M}_N \\ G \in I_H(G_0)}} f 1_S(G) e_{\mathbb{F}}(\alpha G). \tag{41}$$

Since f is completely multiplicative, upon applying Lemma 4.14, we obtain

$$\Sigma_S(\alpha)$$
$$= \frac{1}{|\mathcal{M}_N|} \sum_{G_0 \in \mathcal{M}_N} \frac{\theta(G_0)}{|\mathcal{M}_{<H}|} \sum_{G' \in \mathcal{M}_{\leqslant N}} \frac{f(G')}{1 + \omega_{[P_1,Q_1]}(G')} \sum_{\substack{R \in \mathcal{P} \\ P_1 \leqslant \deg(R) \leqslant \min\{Q_1, N - \deg(G')\} \\ RG' \in I_H(G_0) \\ \deg(RG') = N}} f(R) e_{\mathbb{F}}(RG'\alpha)$$
$$+ O\left(q^{-N-H} \sum_{G_0 \in \mathcal{M}_N} \sum_{P_1 \leqslant d \leqslant Q_1} \sum_{R \in \mathcal{P}_d} |\{G \in I_H(G_0) :\ R^2 | G\}| \right)$$
$$= \frac{1}{|\mathcal{M}_N|} \sum_{G_0 \in \mathcal{M}_N} \frac{\theta(G_0)}{|\mathcal{M}_{<H}|} \sum_{G' \in \mathcal{M}_{\leqslant N}} \frac{f(G')}{1 + \omega_{[P_1,Q_1]}(G')} \sum_{\substack{R \in \mathcal{P} \\ P_1 \leqslant \deg(R) \leqslant \min\{Q_1, N - \deg(G')\} \\ RG' \in I_H(G_0) \\ \deg(RG') = N}} f(R) e_{\mathbb{F}}(RG'\alpha)$$
$$+ O(q^{-P_1}).$$

We pull the summation over G' out, split the sum over R according to degree and apply the triangle inequality to get

$$|\Sigma_S(\alpha)| \leqslant \sum_{P_1 \leqslant k \leqslant Q_1} \frac{1}{|\mathcal{M}_N|} \sum_{G' \in \mathcal{M}_{N-k}} \left| \sum_{R \in \mathcal{P}_k} f(R) e(G'R\alpha) \frac{1}{|\mathcal{M}_{<H}|} \sum_{\substack{G_0 \in \mathcal{M}_N \\ G'R \in I_H(G_0)}} \theta(G_0) \right| + O(H^{-50}),$$

since $P_1 \geqslant 1000 \log H$. We apply Hölder's inequality to the sum over G', getting

$$|\Sigma_S(\alpha)| \ll q^{-N/4-H} \sum_{P_1 \leqslant k \leqslant Q_1} q^{-3k/4} \mathcal{T}_k^{1/4} + H^{-50}, \tag{42}$$

where for each $P_1 \leqslant k \leqslant Q_1$ we define

$$\mathcal{T}_k := \sum_{R_1,R_2,R_3,R_4\in\mathcal{P}_k} f(R_1)f(R_2)\overline{f(R_3)f(R_4)} \sum_{G_1,G_2,G_3,G_4\in\mathcal{M}_N} \theta(G_1)\theta(G_2)\overline{\theta(G_3)\theta(G_4)}$$
$$\sum_{\substack{G'\in\mathcal{M}_{N-k}\\ G'R_j\in I_H(G_j)\ \forall j}} e_{\mathbb{F}}(G'\alpha(R_1+R_2-R_3-R_4)).$$

Fix $P_1 \leqslant k \leqslant Q_1$ for the time being. Split the sums over G_j according to their residue classes $A_j \bmod R_j$. Writing $G_j = D_jR_j + A_j$, we know that $I_H(G_j) = I_H(D_jR_j)$ since $\deg(A_j) < \deg(R_j) \leqslant Q_1 < H$. Thus, we can rewrite $\mathcal{T}_k$ as

$$\mathcal{T}_k = \sum_{\substack{R_1,R_2,R_3,R_4\in\mathcal{P}_k\\ 1\leqslant j\leqslant 4}} f(R_1)f(R_2)\overline{f(R_3)f(R_4)} \sum_{\substack{A_1,A_2,A_3,A_4\\ A_j \bmod R_j\ \forall 1\leqslant j\leqslant 4}} \sum_{D_1\in\mathcal{M}_{N-k}} \theta(D_1R_1+A_1)$$
$$\cdot \sum_{D_2,D_3,D_4\in\mathcal{M}_{N-k}} \theta(D_2R_2+A_2)\overline{\theta(D_3R_3+A_3)\theta(D_4R_4+A_4)}$$
$$\cdot \sum_{\substack{G'\in\mathcal{M}_{N-k}\\ G'\in I_{H-k}(D_j)\ \forall 1\leqslant j\leqslant 4}} e_{\mathbb{F}}(G'\alpha(R_1+R_2-R_3-R_4)).$$

We observe now that

$$G' \in \bigcap_{1\leqslant j\leqslant 4} I_{H-k}(D_j) \iff G' \in I_{H-k}(D_1) \text{ and } \deg(D_j - D_1) < H-k \text{ for all } 1 \leqslant j \leqslant 4.$$

Hence, making the change of variables $L := G' - D_1$, we can recast the above expression for $\mathcal{T}_k$ as

$$\sum_{\substack{R_j\in\mathcal{P}_k\\ 1\leqslant j\leqslant 4}} f(R_1)f(R_2)\overline{f(R_3)f(R_4)} \sum_{D_1\in\mathcal{M}_{N-k}} \sum_{\substack{A_j \bmod R_j\\ 1\leqslant j\leqslant 4}} \theta(D_1R_1+A_1)e_{\mathbb{F}}(D_1\alpha(R_1+R_2-R_3-R_4))$$
$$\cdot \sum_{\substack{D_2,D_3,D_4\in\mathcal{M}_{N-k}\\ \deg(D_j-D_1)<H-k\ \forall j}} \theta(D_2R_2+A_2)\overline{\theta(D_3R_3+A_3)\theta(D_4R_4+A_4)}$$
$$\cdot \sum_{\deg(L)<H-k} e_{\mathbb{F}}(L\alpha(R_1+R_2-R_3-R_4)).$$

Note that now the inner sum over L is decoupled from the sums over A_j and D_j. Given $D_1 \in \mathcal{M}_{N-k}$ fixed, there are $\ll q^{H-k}$ choices of each of D_2, D_3, and D_4 to satisfy the condition $\deg(D_j - D_1) < H-k$. Furthermore, there are $\ll q^{4k}$ choices of four-tuples of residue classes A_1, A_2, A_3, A_4 to their, respective, moduli R_1, R_2, R_3, and R_4. Recalling that $\theta(\cdot)$ is unimodular and bounding trivially in $D_1 \in \mathcal{M}_{N-k}$, it follows that

$$\mathcal{T}_k \ll q^{4k}\cdot q^{3(H-k)}\cdot q^{N-k} \sum_{R_1,R_2,R_3,R_4\in\mathcal{P}_k} \left| \sum_{\deg(L)<H-k} e_{\mathbb{F}}(L\alpha(R_1+R_2-R_3-R_4)) \right|$$
$$\ll q^{3H+N} \sum_{R_1,R_2,R_3,R_4\in\mathcal{P}_k} \left| \sum_{\deg(L)<H-k} e_{\mathbb{F}}(L\alpha(R_1+R_2-R_3-R_4)) \right|.$$

We arrange the four-tuples $(R_1, R_2, R_3, R_4) \in \mathcal{P}_k^4$ according to the values of $F := R_1 + R_2 - R_3 - R_4 \in \mathbb{F}_q[t]$; note that since the R_j are all monic, $\deg(F) < k$. By Corollary 4.13, there are $\ll q^{3k}/k^4$ such representations of F in terms of irreducibles $R_j \in \mathcal{P}_k$. It follows that

$$\mathcal{T}_k \ll \frac{q^{3(H+k)+N}}{k^4} \sum_{\deg(F)<k} \left| \sum_{\deg(L)<H-k} e_{\mathbb{F}}(LF\alpha) \right|.$$

By Lemma 7.6, we can evaluate the exponential sum to yield

$$\begin{aligned}\mathcal{T}_k &\ll \frac{q^{3(H+k)+N}}{k^4} \sum_{\deg(F)<k} q^{H-k} 1_{\langle F\alpha \bmod 1\rangle < q^{-H+k-1}} \\ &= \frac{q^{4H+2k+N}}{k^4} |\{\deg(F) < k : \langle F\alpha \bmod 1\rangle < q^{-H+k-1}\}|.\end{aligned}$$

Since $\alpha \in \mathfrak{m}$, there must be a $g \in \mathcal{M}$ and a reduced residue class a mod g such that $W < \deg(g) \leqslant X$, $(a, g) = 1$ and $\alpha \in \mathfrak{M}_{a,g}(X)$. Since $P_1 \leqslant k \leqslant Q_1$ and given our choice of P_1, Q_1, Lemma 7.7 yields

$$|\{\deg(F) < k : \langle F\alpha \bmod 1\rangle < q^{-H+k-1}\}| \ll q^{k-W},$$

so that we finally obtain the estimate

$$\mathcal{T}_k \ll \frac{q^{4H+3k+N}}{k^4} q^{-W}.$$

Taking fourth roots of both sides and inserting this into (42), we get

$$\begin{aligned}|\Sigma_S(\alpha)| &\ll q^{-N/4-H} \sum_{P_1 \leqslant k \leqslant Q_1} q^{-3k/4} \left(q^{3k+4H+N-W}/k^4\right)^{1/4} + H^{-50} \ll q^{-W/4} \sum_{P_1 \leqslant k \leqslant Q_1} 1/k + H^{-50} \\ &\ll \log(Q_1/P_1) q^{-W/4} + H^{-50}.\end{aligned}$$

In light of the choices $W \geqslant 10 \log H$, $P_1 = 100W$, $Q_1 = H/3$, this leads, finally, to the bound

$$\begin{aligned}\max_{\alpha \in \mathfrak{m}(X,W)} \frac{1}{|\mathcal{M}_N|} \sum_{G_0 \in \mathcal{M}_N} \left| \frac{1}{|\mathcal{M}_{<H}|} \sum_{\substack{G \in \mathcal{M}_N \\ G \in I_H(G_0)}} f(G) e_{\mathbb{F}}(G\alpha) \right| &\ll \max_{\alpha \in \mathfrak{m}(X,W)} |\Sigma_S(\alpha)| + \frac{P_1}{Q_1} \\ &\ll \log(Q_1/P_1) q^{-W/4} + \frac{P_1}{Q_1} + H^{-50} \ll \frac{P_1}{Q_1}.\end{aligned}$$

7.2 | The major arcs

Next, we turn to the estimation of the major arcs, where the Matomäki–Radziwiłł theorem in function fields will be put into use. Fix $g \in \mathcal{M}_{\leqslant W}$ and a reduced residue class a modulo g coprime to g. Suppose that $\alpha \in \mathfrak{M}_{a,g}(X)$. We shall estimate $\Sigma_S(\alpha)$ (given by (40)) in this case as well.

Write $e_{\mathbb{F}}(G\alpha) = e_{\mathbb{F}}(Ga/g)e_{\mathbb{F}}(G\beta)$, and set $\gamma := \deg(g) - 1$. Since $X + \gamma < H$, for each $G_0 \in \mathcal{M}_N$ we may decompose

$$I_H(G_0) = \bigsqcup_{\deg(G')<H-X-\gamma} I_{X+\gamma}(G_0 + t^{X+\gamma}G').$$

As $\beta = \sum_{j \geqslant X+\gamma+1} b_j t^{-j}$, it follows that $e_{\mathbb{F}}(\beta G)$ is constant on $I_{X+\gamma}(G_0 + t^{X+\gamma}G')$, for each G' in the union. Splitting the inner sum over G in $\Sigma_S(\alpha)$ into pieces supported on each of these shorter intervals and applying the triangle inequality, we obtain

$$\begin{aligned}|\Sigma_S(\alpha)| &\leqslant \frac{1}{|\mathcal{M}_N|}\sum_{G_0\in\mathcal{M}_N}\frac{1}{|\mathcal{M}_{<H}|}\sum_{\deg(G')<H-X-\gamma}\left|\sum_{\substack{G\in\mathcal{M}_N\\ G\in I_{X+\gamma}(G_0+t^{X+\gamma}G')}} f1_S(G)e_{\mathbb{F}}(Ga/g)\right|.\\ &=: \frac{1}{|\mathcal{M}_N|}\sum_{G_0\in\mathcal{M}_N}\frac{1}{|\mathcal{M}_{<H}|}\sum_{\deg(G')<H-X-\gamma}|\Sigma_S(\alpha;G_0,G')|.\end{aligned}$$

For a Dirichlet character ψ mod Q, recall that the Gauss sum $\tau(\psi)$ of ψ is defined as

$$\tau(\psi) := \sum_{G \bmod Q}\psi(G)e_{\mathbb{F}}(G/M).$$

It is well known, as in the number field setting, that $|\tau(\psi)| \leqslant q^{\frac{1}{2}\deg(Q)}$. Expanding $e_{\mathbb{F}}(Ga/g)$ in terms of Dirichlet characters mod g, separating G according to the greatest common divisor $D = (G, g)$, we can rewrite

$$\begin{aligned}\Sigma_S(\alpha;G_0,G') &= \sum_{\substack{G\in\mathcal{M}_N\\ G\in I_{X+\gamma}(G_0+t^{X+\gamma}G')}} f1_S(G)e_{\mathbb{F}}(Ga/g)\\ &= \sum_{D|g}\frac{f(D)}{\phi(g/D)}\sum_{\psi \bmod g/D}\overline{\psi}(a)\tau(\psi)\sum_{\substack{G\in\mathcal{M}_{N-\deg(D)}\\ GD\in I_{X+\gamma}(G_0+t^{X+\gamma}G')}} f1_S(G)\overline{\psi}(G)\end{aligned}$$

for each G_0, G' in their respective ranges; here, we have used the fact that as $\deg(D) \leqslant W < P_1$ we have $1_S(DG) = 1_S(G)$.

As in the minor arc case, we separate G_0 and G' according to residue classes modulo D. Write $G_0 = DG_0' + A$, $G' = DG'' + B$, and $t^{X+\gamma} = DT_D + C$, so that

$$t^{X+\gamma}G' = D(DG''T + CG'' + BT_D) + BC.$$

Then, as $\deg(D) \leqslant W < X/2$, we have $\deg(A), \deg(BC) < X + \gamma$ and thus

$$I_{X+\gamma}(G_0 + t^{X+\gamma}G') = I_{X+\gamma}(D(G_0' + DG''T_D + BT_D + CG'')) = I_{X+\gamma}(D(G_0' + t^{X+\gamma}G'' + BT_D)).$$

Hence, we see that

$$GD \in I_{X+\gamma}(G_0 + t^{X+\gamma}G') \text{ if and only if } G \in I_{X+\gamma-\deg(D)}(G_0' + t^{X+\gamma}G'' + BT_D).$$

It follows, using the triangle inequality and the bound $|\tau(\psi)| \leqslant q^{\deg(g/D)/2}$ that

$$\begin{aligned}
&|\Sigma_S(\alpha)| \\
&\ll \sum_{D|g} \frac{q^{\frac{1}{2}\deg(g/D)}}{\phi(g/D)} \sum_{\psi \bmod g/D} \sum_{A,B \bmod D} \frac{1}{|\mathcal{M}_N|} \sum_{DG_0'+A \in \mathcal{M}_N} \frac{|\mathcal{M}_{<X+\gamma-\deg(D)}||\mathcal{M}_{<H-X-\gamma-\deg(D)}|}{|\mathcal{M}_{<H}|} \\
&\cdot \frac{1}{|\mathcal{M}_{<H-X-\gamma-\deg(D)}|} \sum_{\deg(DG''+B)<H-X-\gamma} \left| \frac{1}{|\mathcal{M}_{<X+\gamma-\deg(D)}|} \sum_{\substack{G \in \mathcal{M}_{N-\deg(D)} \\ G \in I_{X+\gamma-\deg(D)}(G_0'+t^{X+\gamma}G''+BT_D)}} f1_S(G)\overline{\psi}(G) \right| \\
&\ll q^{\frac{1}{2}\deg(g)} \sum_{D|g} q^{-\frac{3}{2}\deg(D)} \max_{\psi \bmod g/D} \max_{\deg(G'')<H-X-\gamma-\deg(D)} \max_{B \bmod D} \\
&\cdot \frac{1}{|\mathcal{M}_{N-\deg(D)}|} \sum_{G_0' \in \mathcal{M}_{N-\deg(D)}} \left| \frac{1}{|\mathcal{M}_{<X+\gamma-\deg(D)}|} \sum_{\substack{G \in \mathcal{M}_{N-\deg(D)} \\ G \in I_{X+\gamma-\deg(D)}(G_0'+t^{X+\gamma}G''+BT_D)}} f1_S(G)\overline{\psi}(G) \right|.
\end{aligned}$$

We observe that $\deg(BT_D) < X + \gamma \leqslant \deg(G''t^{X+\gamma})$, so that $\deg(G''t^{X+\gamma} + BT_D) \leqslant H - \deg(D) < N - \deg(D)$. We can thus make the change of variables $G_0'' := G_0' + G''t^{X+\gamma} + BT_D$ to finally obtain

$$\begin{aligned}
|\Sigma_S(\alpha)| &\ll q^{\frac{1}{2}\deg(g)} \max_{\substack{D|g \\ \psi \bmod g/D}} \frac{1}{|\mathcal{M}_{N-\deg(D)}|} \\
&\cdot \sum_{G_0'' \in \mathcal{M}_{N-\deg(D)}} \left| \frac{1}{|\mathcal{M}_{<X+\gamma-\deg(D)}|} \sum_{\substack{G \in \mathcal{M}_{N-\deg(D)} \\ G \in I_{X+\gamma-\deg(D)}(G_0'')}} f1_S(G)\overline{\psi}(G) \right|.
\end{aligned}$$

Let D_1 be a divisor of g such that some character $\psi_1 \bmod D_1$ yields the maximal contribution among the characters whose modulus divides g. Put $d := \deg(D_1)$, so that $d \leqslant W < N/2$. By Theorem 3.5 and Lemma 4.8, we get

$$\begin{aligned}
&\frac{1}{|\mathcal{M}_{N-d}|} \left| \sum_{G \in \mathcal{M}_{N-d}} f1_S(G)\overline{\psi}_1(G) \right| \\
&\leqslant \frac{1}{|\mathcal{M}_{N-d}|} \left| \sum_{G \in \mathcal{M}_{N-d}} f(G)\overline{\psi}_1(G) \right| + \frac{1}{|\mathcal{M}_{N-d}|} \left| \sum_{G \in \mathcal{M}_{N-d}} f(G)\overline{\psi}_1(G)1_{S^c}(G) \right|
\end{aligned}$$

$$\ll (1+\mathcal{D}_{f\overline{\psi}_1}(N))e^{-\mathcal{D}_{f\overline{\psi}_1}(N)} + \min\left\{\frac{P_1}{Q_1}, (1+\mathcal{D}_{f\overline{\psi}_1 1_{\mathcal{S}^c}}(N))e^{-\mathcal{D}_{f\overline{\psi}_1 1_{\mathcal{S}^c}}(N)}\right\}$$
$$\ll (1+\mathcal{D}_{f\overline{\psi}_1}(N))e^{-\mathcal{D}_{f\overline{\psi}_1}(N)} + \frac{P_1}{Q_1}\min\left\{1, (Q_1/P_1)^3(1+\mathcal{D}_{f\overline{\psi}_1}(N))e^{-\mathcal{D}_{f\overline{\psi}_1}(N)}\right\},$$

where in the last step we used the fact that

$$\sum_{R\in\mathcal{P}_{\leqslant N}} \frac{1-\mathrm{Re}(f(R)\overline{\psi}_1(R)1_{\mathcal{S}^c}(R)e_{-\theta}(R))}{q^{\deg(R)}} \geqslant \sum_{R\in\mathcal{P}_{\leqslant N}} \frac{1-\mathrm{Re}(f(R)\overline{\psi}_1(R)e_{-\theta}(R))}{q^{\deg(R)}} - 2\sum_{\substack{R\in\mathcal{P}\\ P_1\leqslant\deg(R)\leqslant Q_1}} q^{-\deg(R)}$$
$$= \sum_{R\in\mathcal{P}_{\leqslant N}} \frac{1-\mathrm{Re}(f(R)\overline{\psi}_1(R)e_{-\theta}(R))}{q^{\deg(R)}} - 2\log(Q_1/P_1) + O(1).$$

By the triangle inequality and the assumption $\deg(g)\leqslant W$, we thus have

$$|\Sigma_{\mathcal{S}}(\alpha)|$$
$$\ll \frac{q^{\frac{W}{2}}}{|\mathcal{M}_{N-d}|}\sum_{G_0''\in\mathcal{M}_{N-d}}\left|\frac{1}{|\mathcal{M}_{<X+\gamma-d}|}\sum_{\substack{G\in\mathcal{M}_{N-d}\\ G\in I_{X+\gamma-d}(G_0'')}} f1_{\mathcal{S}}\overline{\psi}_1(G) - \frac{1}{|\mathcal{M}_{N-d}|}\sum_{G\in\mathcal{M}_{N-d}} f1_{\mathcal{S}}(G)\overline{\psi}_1(G)\right|$$
$$+ q^{\frac{W}{2}}\left(\mathcal{D}_{f\overline{\psi}_1}(N)e^{-\mathcal{D}_{f\overline{\psi}_1}(N)} + \frac{P_1}{Q_1}\min\left\{1,(Q_1/P_1)^3(1+\mathcal{D}_{f\overline{\psi}_1}(N))e^{-\mathcal{D}_{f\overline{\psi}_1}(N)}\right\}\right).$$

Applying Theorem 6.1 (in the form given in Remarks 6.2 and 6.3, bounding the long sum using Theorem 3.5), the first expression above is

$$\ll q^{W/2}\left(Q_1^{\frac{1}{2}}q^{-\frac{1}{12}P_1} + N^{-1/36+o(1)} + 1_{\substack{f\overline{\psi}_1 \text{ not real}\\ \text{or } 2|q}}1_{\chi_1\neq\chi_0}(1+\mathcal{D}_{(f\overline{\psi}_1)^*\overline{\chi}_1}(N))e^{-\mathcal{D}_{(f\overline{\psi}_1)^*\overline{\chi}_1}(N)}\right),$$

where χ_1 denotes the character modulo $t^{N-X-\gamma+1}$ such that $\chi\mapsto\mathcal{D}_{(f\overline{\psi}_1)^*\overline{\chi}}(N)$ is minimal (and χ_0 is the principal character to the same modulus). Recalling that $P_1=100W$ and $Q_1=H/3$, it follows that

$$\max_{\alpha\in\mathcal{M}(X,W)}|\Sigma_{\mathcal{S}}(\alpha)| \ll q^{W/2}(H^{1/2}q^{-8W}+N^{-1/36+o(1)}$$
$$+\max_{\substack{M\in\mathcal{M}_{\leqslant W}\\ \psi \bmod M}}\left((1+\mathcal{D}_{f\overline{\psi}}(N))e^{-\mathcal{D}_{f\overline{\psi}}(N)} + (P_1/Q_1)\min\{1,(Q_1/P_1)^3(1+\mathcal{D}_{f\overline{\psi}}(N))e^{-\mathcal{D}_{f\overline{\psi}}(N)}\}\right) \quad (43)$$
$$+q^{W/2}\max_{M\in\mathcal{M}_{\leqslant W}}\max_{\substack{\psi \bmod M\\ f\overline{\psi}\text{ not real}\\ \text{if } 2\nmid q}}\max_{\substack{\chi\neq\chi_0 \bmod t^{N-j+1}\\ X<j\leqslant H}}(1+\mathcal{D}_{(f\overline{\psi})^*\overline{\chi}}(N))e^{-\mathcal{D}_{(f\overline{\psi})^*\overline{\chi}}(N)}. \quad (44)$$

In order to estimate this quantity further and to prove Proposition 7.5, we split the remainder of the analysis into two cases.

Case 1: f is not real-valued or q is even

By Corollary 3.15, we have

$$M_{\mathrm{Hayes}}(f;N,H) \leqslant \min_{M\in\mathcal{M}_{\leqslant X}} \min_{\substack{\psi \bmod M\\ f\overline{\psi} \text{ not real}}} \min_{X<j\leqslant H} \min_{\substack{\chi \bmod t^{N-j+1}\\ \chi\neq\chi_0}} \mathcal{D}_{(f\overline{\psi})^*\overline{\chi}}(N) + O(1).$$

Of course, we also have

$$M_{\mathrm{Hayes}}(f;N,H) \leqslant M_{\mathrm{Dir}}(f;N,H) \leqslant \min_{M\in\mathcal{M}_{\leqslant X}} \min_{\psi \bmod M} \mathcal{D}_{f\overline{\psi}}(N).$$

Inserting these bounds into (44) and using $10\log H \leqslant W \leqslant (\log N)/(100\log q)$, we get

$$\begin{aligned}\max_{\alpha\in\mathcal{M}(X,W)} |\Sigma_S(\alpha)| &\ll H^{-50} + N^{-1/40}\\ &\quad + q^{W/2}\Big(e^{-M_{\mathrm{Dir}}(f;N,H)/2} + (P_1/Q_1)\min\{1,(Q_1/P_1)^3 e^{-M_{\mathrm{Dir}}(f;N,H)/2}\} + e^{-M_{\mathrm{Hayes}}(f;N,H)/2}\Big).\end{aligned}$$

By assumption, we also have $W \leqslant M_{\mathrm{Hayes}}(f;N,H)/10 \leqslant M_{\mathrm{Dir}}(f;N,H)/10$. Furthermore, if $Q_1/P_1 \geqslant e^{M_{\mathrm{Dir}}(f;N,H)/5}$, then

$$q^{W/2}P_1/Q_1 \leqslant e^{M_{\mathrm{Dir}}(f;N,H)/20 - M_{\mathrm{Dir}}(f;N,H)/5} \leqslant e^{-M_{\mathrm{Dir}}(f;N,H)/10},$$

whereas if $Q_1/P_1 < e^{M_{\mathrm{Dir}}(f;N,H)/5}$, then

$$q^{W/2}(Q_1/P_1)^2 e^{-M_{\mathrm{Dir}}(f;N,H)/2} < e^{M_{\mathrm{Dir}}(f;N,H)(1/20+2/5-1/2)} = e^{-M_{\mathrm{Dir}}(f;N,H)/20}.$$

Thus, we deduce the bound

$$\begin{aligned}\max_{\alpha\in\mathcal{M}(X,W)} |\Sigma_S(\alpha)| &\ll (\log H)H^{-1} + N^{-1/40} + e^{-M_{\mathrm{Dir}}(f;N,H)/20} + e^{-M_{\mathrm{Hayes}}(f;N,H)/2}q^{W/2}\\ &\ll H^{-50} + N^{-1/40} + e^{-M_{\mathrm{Hayes}}(f;N,H)/20},\end{aligned}$$

since $10\log H \leqslant W \leqslant \min\{H/10, (\log N)/(100\log q), M_{\mathrm{Hayes}}(f;N,H)/10\}$.

Case 2: f is real-valued and q odd

We claim that

$$\max_{M\in\mathcal{M}_{\leqslant W}} \max_{\substack{\psi \bmod M\\ f\overline{\psi} \text{ not real}}} \max_{\substack{X<j\leqslant H\\ \chi\neq\chi_0 \bmod t^{N-j}}} \mathcal{D}_{(f\overline{\psi})^*\overline{\chi}}(N)e^{-\mathcal{D}_{(f\psi)^*\overline{\chi}}(N)} \ll N^{-1/4+o(1)}. \tag{45}$$

Inserting this into (44) and then repeating the arguments in Case 1 to simplify the terms in (43), we obtain

$$\max_{\alpha\in\mathcal{M}(X,W)}|\Sigma_S(\alpha)|\ll H^{-50}+N^{-1/40}+e^{-M_{\mathrm{Dir}}(f;N,H)/20}.$$

Let ψ be a character of modulus M with $\deg(M)\leqslant W\leqslant \log N$ for which $f\overline{\psi}$ is not real-valued. Since f is real-valued, it follows that ψ is not, nor is $(f\overline{\psi})^*$. Put $M=\tilde{M}t^r$, where $(\tilde{M},t)=1$, and write $\psi=\psi_{\tilde{M}}\psi_{t^r}$. We consider two subcases, depending on whether or not $\psi_{\tilde{M}}$ is real-valued.

Case 2.1. Suppose first that $\psi_{\tilde{M}}$ is real. Since ψ^2 is non-principal, it follows that $\psi_{t^r}^2$ is non-principal. Applying the triangle inequality as in the proof of Lemma 6.4, we can show that

$$\mathcal{D}_{(f\overline{\psi})^*\overline{\chi}}(N)\geqslant\frac{1}{4}\mathcal{D}_{(\psi_{t^r}^2)^*\chi^2}(N).$$

By Lemma 3.12, $(\psi_{t^r}^2)^*$ is a non-principal short interval character of length $\leqslant r$, so that $(\psi_{t^r}^2)^*\chi^2$ is a non-principal Hayes character of conductor $\leqslant r+\mathrm{cond}(\chi^2)\leqslant W+N-H+1<N$. Lemma 3.2 now implies that, for some $\theta_0\in[0,1]$,

$$\begin{aligned}\mathcal{D}_{(\psi_{t^r}^2)^*\chi^2}(N)&=\log N-\mathrm{Re}\left(\sum_{d\leqslant N}\frac{e(-\theta_0 d)}{dq^d}\sum_{G\in\mathcal{M}_d}(\psi_{t^r}^2)^*\chi^2(G)\Lambda(G)\right)+O(1)\\&=(1-o(1))\log N.\end{aligned}$$

In particular, we find that

$$\mathcal{D}_{(f\overline{\psi})^*\overline{\chi}}(N)\geqslant(1/4-o(1))\log N,$$

which implies (45) in this case.

Case 2.2. Next, suppose $\psi_{\tilde{M}}$ is not real, so that $\psi_{\tilde{M}}^2$ is non-principal. Without loss of generality, we may assume that f is extended to $\mathbb{F}_q[t]$ by $f(c)\overline{\psi}(c)=\chi(c)$. By Lemma 3.10, we see that

$$\mathcal{D}_{(f\overline{\psi})^*\overline{\chi}}(N)\geqslant\mathcal{D}_{f\overline{\psi}_{\tilde{M}}\overline{\psi}_{t^r}\overline{\chi}^*}(N)+O(1).$$

Applying a similar argument as in the previous subcase, we have then that

$$\mathcal{D}_{(f\overline{\psi})^*\overline{\chi}}(N)\geqslant\frac{1}{4}\mathcal{D}_{\psi_{\tilde{M}}^2(\psi_{t^r})^2(\chi^2)^*}(N)+O(1).$$

Since $(\tilde{M},t)=1$, $\psi_{\tilde{M}}^2(\psi_{t^r})^2(\chi^2)^*$ is a non-principal Hayes character. Similarly as in the previous subcase, we obtain

$$\mathcal{D}_{(f\overline{\psi})^*\chi}(N)\geqslant(1/4-o(1))\log N.$$

Thus, (45) is valid in this case as well, and thus in all cases in which f is real-valued. This completes the proof of Proposition 7.5.

Proof of Theorem 7.3. Let $1 \leqslant W, H' \leqslant N$, and put $P_1 := 100W$ and $Q_1 := H'/3$. If we assume the condition

$$10 \log H' \leqslant W \leqslant \min\{M_*(f;N,H')/10, (\log N)/(100 \log q), H'/10\}, \tag{46}$$

where we recall that $M_* = M_{\text{Hayes}}$ unless f is real and q is odd in which case $M_* = M_{\text{Dir}}$, then we have

$$\max_{\alpha \in \mathbb{T}} \frac{1}{|\mathcal{M}_N|} \sum_{G_0 \in \mathcal{M}_N} \left| \sum_{\substack{G \in \mathcal{M}_N \\ G \in I_{H'}(G_0)}} f(G) e_{\mathbb{F}}(G\alpha) \right|$$
$$\ll q^{H'} \Big((\log H')(H')^{-1} + P_1/Q_1 + N^{-1/40} + e^{-M_*(f;N,H')/20} \Big).$$

Suppose now that $1 \leqslant H \leqslant N - N^{3/4}$, and define $1 \leqslant H_0 \leqslant N$ by

$$\log H_0 := \min\{M_*(f;N,H)/100, (\log N)/(1000 \log q), H/100\}.$$

We will make a choice of W that suits our current choice of H.

If $H \leqslant H_0$, then $W := 10 \log H$ is admissible in (46) with $H' = H$, and Theorem 7.3 is verified in this case (here $P_1 \ll \log H$, so $P_1/Q_1 \ll (\log H)H^{-1}$).

Next, suppose $H > H_0$. For each $G_0 \in \mathcal{M}_N$ we can split $I_H(G_0)$ into $\ll q^{H-H_0}$ short intervals $I_{H_0}(G_0 + t^{H_0}M)$, where $\deg(M) < H - H_0$. We then have

$$\max_{\alpha \in \mathbb{T}} \frac{1}{|\mathcal{M}_N|} \sum_{G_0 \in \mathcal{M}_N} \left| \sum_{\substack{G \in \mathcal{M}_N \\ G \in I_H(G_0)}} f(G) e_{\mathbb{F}}(G\alpha) \right|$$
$$\ll q^{H-H_0} \max_{\deg(M) < H-H_0} \max_{\alpha \in \mathbb{T}} \frac{1}{|\mathcal{M}_N|} \sum_{G_0 \in \mathcal{M}_N} \left| \sum_{\substack{G \in \mathcal{M}_N \\ G \in I_{H_0}(G_0 + t^{H_0}M)}} f(G) e_{\mathbb{F}}(G\alpha) \right|$$
$$= q^{H-H_0} \max_{\alpha \in \mathbb{T}} \frac{1}{|\mathcal{M}_N|} \sum_{G_0' \in \mathcal{M}_N} \left| \sum_{\substack{G \in \mathcal{M}_N \\ G \in I_{H_0}(G_0')}} f(G) e_{\mathbb{F}}(G\alpha) \right|.$$

We have thus reduced matters to the case $H = H_0$, which was addressed previously. Since $H \mapsto M_*(f;N,H)$ is non-increasing, we see that

$$\log H_0 \leqslant \min\{M_*(f;N,H_0)/100, (\log N)/(1000 \log q), H_0/100\}$$

when N (and therefore H) is large enough. Selecting $W := 10 \log H_0$ gives an admissible choice in relation to (46) with $H = H_0$, picking P_1, Q_1 in terms of H_0 so that $P_1/Q_1 \ll (\log H_0)/H_0$. We

thus have

$$\begin{aligned}\max_{\alpha\in\mathbb{T}} \frac{1}{|\mathcal{M}_N|}\sum_{G_0\in\mathcal{M}_N}\left|\sum_{\substack{G\in\mathcal{M}_N\\ G\in I_H(G_0)}} f(G)e_{\mathbb{F}}(G\alpha)\right| &\ll q^{H-H_0}\cdot q^{H_0}\left((\log H_0)H_0^{-1}+N^{-1/40}+e^{-M_*(f;N,H_0)/20}\right)\\ &\ll q^{H}\left(N^{-1/(2000\log q)}+M_*(f;N,H)e^{-M_*(f;N,H)/100}\right),\end{aligned}$$

using the fact that $H'\mapsto M_*(f;N,H')$ is non-increasing and the definition of H_0 in the last line. Theorem 7.3 then follows in this case as well. □

Proof of Theorems 7.1 and 7.2. We will only prove Theorem 7.1 from the first statement in Theorem 7.3, as Theorem 7.2 follows in the same way from the second statement in Theorem 7.3.

Let $f:\mathcal{M}\to\mathbb{U}$ be a multiplicative function. Define $\tilde{f}$ to be the completely multiplicative function such that $\tilde{f}(P)=f(P)$ for all $P\in\mathcal{P}$. We may thus find a multiplicative function $h:\mathcal{M}\to\mathbb{C}$, supported on squarefull monic polynomials (i.e., if $h(P^k)\neq 0$ for $P\in\mathcal{P}$ and $k\in\mathbb{N}$ then $k\geqslant 2$) such that $f=\tilde{f}*h$; in particular, h is bounded by the divisor function $d(G)=\sum_{D|G}1$, and hence $|h(G)|\ll_\varepsilon q^{\varepsilon\deg(G)}$ for any $G\in\mathcal{M}$, a fact we will use shortly. We thus have

$$\begin{aligned}&\sup_{\alpha\in\mathbb{T}}\frac{1}{|\mathcal{M}_N|}\sum_{G_0\in\mathcal{M}_N}\frac{1}{|\mathcal{M}_{<H}|}\left|\sum_{\substack{G\in\mathcal{M}_N\\ G\in I_H(G_0)}} f(G)e_{\mathbb{F}}(G\alpha)\right|\\ &\leqslant\sum_{D\in\mathcal{M}_{\leqslant N}}|h(D)|\sup_{\alpha\in\mathbb{T}}\frac{1}{|\mathcal{M}_N|}\sum_{G_0\in\mathcal{M}_N}\frac{1}{|\mathcal{M}_{<H}|}\left|\sum_{\substack{G'\in\mathcal{M}_{N-\deg(D)}\\ DG'\in I_H(G_0)}}\tilde{f}(G')e_{\mathbb{F}}(G'D\alpha)\right|\\ &=:\mathcal{T}_{<H}+\mathcal{T}_{\geqslant H}.\end{aligned}$$

We first estimate $\mathcal{T}_{\geqslant H}$, which corresponds to the terms with $\deg(D)\geqslant H$ above. If $\deg(D)\geqslant H$, then if $G_0\in\mathcal{M}_N$ is such that $I_H(G_0)\cap D\mathcal{M}_{N-\deg(D)}\neq\emptyset$, then in fact $|I_H(G_0)\cap D\mathcal{M}_{N-\deg(D)}|=1$ and G_0 lies in one of at most $\ll q^H$ residue classes modulo D. It follows that

$$\mathcal{T}_{\geqslant H}\ll\sum_{\substack{D\in\mathcal{M}\\ \deg(D)\geqslant H}}|h(D)|\cdot\frac{1}{|\mathcal{M}_N||\mathcal{M}_{<H}|}\cdot q^{N+H-\deg(D)}\ll\sum_{\substack{D\in\mathcal{M}\\ \deg(D)\geqslant H}}|h(D)|q^{-\deg(D)}.$$

Since h is supported on squarefull polynomials, all of which are of the form A^2B^3 for some $A,B\in\mathcal{M}$, and moreover $|h(D)|\ll q^{\deg(D)/5}$ for all D, we obtain

$$\mathcal{T}_{\geqslant H}\ll\sum_{\substack{D\in\mathcal{M}\\ \deg(D)\geqslant H}}|h(D)|q^{-\deg(D)}\ll q^{-H/5}\sum_{A^2B^3\in\mathcal{M}}q^{-0.6(2\deg(A)+3\deg(B))}\ll q^{-H/5}. \tag{47}$$

Next, we estimate $\mathcal{T}_{<H}$. Writing $G_0 = DG_0' + B$ for some $B \bmod D$ and $D \in \mathcal{M}_{<H}$, we have

$$\mathcal{T}_{<H} = \sum_{D\in\mathcal{M}_{<H}} |h(D)| \sup_{\alpha\in\mathbb{T}} \frac{1}{|\mathcal{M}_N|} \sum_{B \bmod D} \sum_{G_0'\in\mathcal{M}_{N-\deg(D)}} \frac{1}{|\mathcal{M}_{<H}|} \left| \sum_{\substack{G'\in\mathcal{M}_{N-\deg(D)} \\ DG'\in I_H(DG_0'+B)}} \tilde{f}(G') e_{\mathbb{F}}(G'D\alpha) \right|.$$

Since $\deg(B) < \deg(D) < H$, we see that $I_H(DG_0' + B) = I_H(DG_0')$ for all $B \bmod D$. Moreover, we also have that $DG' \in I_H(DG_0')$ if and only if $G' \in I_{H-\deg(D)}(G_0')$. Thus,

$$\begin{aligned}
\mathcal{T}_{<H} &= \sum_{D\in\mathcal{M}_{<H}} |h(D)| \sup_{\alpha\in\mathbb{T}} \frac{q^{\deg(D)}}{|\mathcal{M}_N|} \sum_{G_0'\in\mathcal{M}_{N-\deg(D)}} \frac{1}{|\mathcal{M}_{<H}|} \left| \sum_{\substack{G'\in\mathcal{M}_{N-\deg(D)} \\ G'\in I_{H-\deg(D)}(G_0')}} \tilde{f}(G') e_{\mathbb{F}}(G'D\alpha) \right| \\
&\ll \sum_{D\in\mathcal{M}_{<H}} \frac{|h(D)|}{q^{\deg(D)}} \sup_{\alpha\in\mathbb{T}} \frac{1}{|\mathcal{M}_{N-\deg(D)}|} \sum_{G_0'\in\mathcal{M}_{N-\deg(D)}} \frac{1}{|\mathcal{M}_{<H-\deg(D)}|} \left| \sum_{\substack{G'\in\mathcal{M}_{N-\deg(D)} \\ G'\in I_{H-\deg(D)}(G_0')}} \tilde{f}(G') e_{\mathbb{F}}(G'D\alpha) \right| \\
&\leqslant \sum_{D\in\mathcal{M}_{<H}} \frac{|h(D)|}{q^{\deg(D)}} \sup_{\beta\in\mathbb{T}} \frac{1}{|\mathcal{M}_{N-\deg(D)}|} \sum_{G_0'\in\mathcal{M}_{N-\deg(D)}} \frac{1}{|\mathcal{M}_{<H-\deg(D)}|} \left| \sum_{\substack{G'\in\mathcal{M}_{N-\deg(D)} \\ G'\in I_{H-\deg(D)}(G_0')}} \tilde{f}(G') e_{\mathbb{F}}(G'\beta) \right|.
\end{aligned}$$

Since the supremum over β is $\leqslant 1$ for all $D \in \mathcal{M}_{<H}$, we may further bound the contribution from $\deg(D) \geqslant H/2$ (as in (47), with H replaced by $H/2$) by $O(q^{-H/10})$. Applying Theorem 7.3 for each $D \in \mathcal{M}_{<H/2}$, we find

$$\begin{aligned}
\mathcal{T}_{<H} \ll q^{-H/10} + \sum_{D\in\mathcal{M}_{<H/2}} |h(D)| q^{-\deg(D)} \Bigg(& \frac{\log(H-\deg(D))}{H-\deg(D)} \\
& + (N-\deg(D))^{-1/(2000\log q)} + M_D e^{-M_D/100} \Bigg),
\end{aligned} \tag{48}$$

where we have set $M_D := M_{\mathrm{Hayes}}(\tilde{f}; N-\deg(D), H-\deg(D)) + 1$. We note from its definition that M_{Hayes} is non-increasing in H, and since $\tilde{f}$ takes the same values as f on primes we get

$$M_D \geqslant M_{\mathrm{Hayes}}(\tilde{f}; N-\deg(D), H) = M_{\mathrm{Hayes}}(f; N-\deg(D), H).$$

Finally, as $\mathcal{D}_g(N-\deg(D)) \geqslant \mathcal{D}_g(N/2) = \mathcal{D}_g(N) - O(1)$ for any one-bounded function $g : \mathcal{M} \to \mathbb{U}$, it follows that

$$M_D \geqslant M_{\mathrm{Hayes}}(f; N, H) - O(1) =: M - 1 - O(1)$$

for all $D \in \mathcal{M}_{<H/2}$. Invoking this in (48), we obtain the bound

$$\mathcal{T}_{<H} \ll q^{-H/10} + \left(\frac{\log H}{H} + N^{-1/(2000\log q)} + Me^{-M/100}\right) \sum_{D\in\mathcal{M}} |h(D)| q^{-\deg(D)}$$
$$\ll \frac{\log H}{H} + N^{-1/(2000\log q)} + Me^{-M/100}.$$

Combining this with our earlier estimate for $\mathcal{T}_{\geqslant H}$, the proof of Theorem 7.1 follows. □

8 | ELLIOTT'S CONJECTURE

In this section, we shall prove the two-point case of the logarithmically averaged Elliott's conjecture on correlations of non-pretentious multiplicative functions in function fields, Theorem 1.4. Here, we only treat the case $A = 1$ for simplicity; the proof of the general case of fixed monic A, which is essentially the same, is left to the interested reader.

In the sequel, we will adopt the following notational conventions: If $S \subset \mathbb{F}_q[t]$ and $g : \mathbb{F}_q[t] \to \mathbb{C}$, then

$$\mathbb{E}_{G\in S} g(G) := |S|^{-1} \sum_{G\in S} g(G),$$
$$\mathbb{E}^{\log}_{G\in S} g(G) := \left(\sum_{G\in S} q^{-\deg(G)}\right)^{-1} \sum_{G\in S} g(G) q^{-\deg(G)} \quad \text{if} \quad 0 \notin S.$$

To prove Theorem 1.4, we will combine the exponential sum estimate of Theorem 7.2 with a function field version of the entropy decrement argument that Tao developed in [33] for the corresponding problem in the integer setting. The key proposition arising from this is the following.

Proposition 8.1 (Introducing an extra averaging variable). *Let $N \geqslant 100$, and let $B \in \mathbb{F}_q[t]\backslash\{0\}$ be fixed. For any $1 \leqslant K \leqslant \log\log\log N$, there exists $H \in [K, \exp(\exp(10K))]$ such that the following is true. Suppose that f_1, f_2 satisfy the hypotheses of Theorem 1.4. For each $R \in \mathcal{P}_H$ set $c_R := \overline{f_1(R) f_2(R)}$. Then,*

$$\mathbb{E}^{\log}_{G\in\mathcal{M}_{\leqslant N}} q^{-\deg(G)} f_1(G) f_2(G+B) = \mathbb{E}_{P\in\mathcal{P}_H} c_P \mathbb{E}^{\log}_{G\in\mathcal{M}_{\leqslant N}} f_1(G) f_2(G+PB) + O(K^{-0.1}).$$

Proposition 8.1 will be deduced from the following proposition, which is based on our version of the entropy decrement argument.

Proposition 8.2 (Entropy decrement argument in function fields). *Let $k \geqslant 1$, and let $a_1, \dots, a_k : \mathbb{F}_q[t] \to \mathbb{U}$ be arbitrary one-bounded functions. Also let $B_1, \dots, B_k \in \mathbb{F}_q[t]$ be any fixed polynomials. Then, for any large enough N and for $1 \leqslant K \leqslant \log\log\log N$, there exists $H \in [K, \exp(\exp(10K))]$ such that*

$$\mathbb{E}^{\log}_{G\in\mathcal{M}_{\leqslant N}} |\mathbb{E}_{P\in\mathcal{P}_H} a_1(G+PB_1)\cdots a_k(G+PB_k)(q^{\deg(P)} 1_{P|G} - 1)| \ll K^{-0.1}.$$

Proof of Proposition 8.1 assuming Proposition 8.2. We may assume that K is larger than any fixed constant, otherwise the claim of Proposition 8.1 (with a suitably large implicit constant in the error term).

Thus, let $H \in [K, \exp(\exp(10K))]$, which may be assumed to be sufficiently large. By multiplicativity, for each $P \in \mathcal{P}_H$, we have

$$f_1(G)f_2(G+B) = c_P f_1(GP)f_2(PG+BP),$$

unless $P|G$ or $P|(G+B)$. Averaging over $P \in \mathcal{P}_H$, for a suitable choice of H we have

$$\mathbb{E}^{\log}_{G\in\mathcal{M}_{\leqslant N}} f_1(G)f_2(G+B) = \mathbb{E}_{P\in\mathcal{P}_H} c_P \mathbb{E}^{\log}_{G\in\mathcal{M}_{\leqslant N}} f_1(G)f_2(G+PB)q^{\deg(P)}1_{P|G} + O(q^{-H}+N^{-1}),$$

since (accounting for $G=-B$ in case this is monic)

$$\mathbb{E}_{P\in\mathcal{P}_H}\mathbb{E}^{\log}_{G\in\mathcal{M}_{\leqslant N}} 1_{G\equiv 0 \text{ or } -B \bmod P} \ll q^{-H}+N^{-1}.$$

By Proposition 8.2 with $a_i = f_i$ and the triangle inequality, we have

$$\begin{aligned}&\mathbb{E}_{P\in\mathcal{P}_H} c_P \mathbb{E}^{\log}_{G\in\mathcal{M}_{\leqslant N}} f_1(G)f_2(G+PB)q^{\deg(P)}1_{P|G}\\ &= \mathbb{E}_{P\in\mathcal{P}_H} c_P \mathbb{E}^{\log}_{G\in\mathcal{M}_{\leqslant N}} f_1(G)f_2(G+PB) + O(K^{-0.1}),\end{aligned}$$

and the claim follows. □

In the next subsection, we will establish Proposition 8.2.

8.1 | The entropy decrement argument in function fields

Definition 8.3. Let $\boldsymbol{X}, \boldsymbol{Y}$ be random variables on a probability space $(\Omega, \mathbb{P})$ with finite ranges $\mathcal{X}, \mathcal{Y}$, respectively. We define the entropy

$$\mathbb{H}(\boldsymbol{X}) := \sum_{x\in\mathcal{X}} \mathbb{P}(\boldsymbol{X}=x)\log\frac{1}{\mathbb{P}(\boldsymbol{X}=x)}$$

and the joint entropy

$$\mathbb{H}(\boldsymbol{X},\boldsymbol{Y}) := \sum_{x\in\mathcal{X},y\in\mathcal{Y}} \mathbb{P}(\boldsymbol{X}=x,\boldsymbol{Y}=y)\log\frac{1}{\mathbb{P}(\boldsymbol{X}=x,\boldsymbol{Y}=y)}.$$

Let $E \subset \Omega$. We define the conditional entropy of $\boldsymbol{X}$ with respect to the event E by

$$\mathbb{H}(\boldsymbol{X}|E) = \sum_{x\in\mathcal{X}} \mathbb{P}(\boldsymbol{X}=x \mid E)\log\frac{1}{\mathbb{P}(\boldsymbol{X}=x \mid E)}$$

and further define the conditional entropy of $\boldsymbol{X}$ given $\boldsymbol{Y}$ by

$$\mathbb{H}(\boldsymbol{X}|\boldsymbol{Y}) = \sum_{y\in\mathcal{Y}} \mathbb{H}(\boldsymbol{X}|\boldsymbol{Y}=y)\mathbb{P}(\boldsymbol{Y}=y).$$

Note that this satisfies the identity

$$\mathbb{H}(\boldsymbol{X},\boldsymbol{Y}) = \mathbb{H}(\boldsymbol{X}|\boldsymbol{Y}) + \mathbb{H}(\boldsymbol{Y}).$$

Finally, we define the mutual information between $\boldsymbol{X}$ and $\boldsymbol{Y}$ by

$$\mathbb{I}(\boldsymbol{X},\boldsymbol{Y}) := \mathbb{H}(\boldsymbol{X}) + \mathbb{H}(\boldsymbol{Y}) - \mathbb{H}(\boldsymbol{X},\boldsymbol{Y}).$$

The non-negativity of $\mathbb{I}(\boldsymbol{X},\boldsymbol{Y})$ follows from the following lemma.

Lemma 8.4 (Shannon inequalities). *Let $\boldsymbol{X},\boldsymbol{Y}$ be random variables on a probability space $(\Omega,\mathbb{P})$ with finite ranges $\mathcal{X},\mathcal{Y}$. Then, we have the bounds*

$$0 \leqslant \mathbb{H}(\boldsymbol{X}) \leqslant \log|\mathcal{X}|$$

and

$$\mathbb{H}(\boldsymbol{X}) \leqslant \mathbb{H}(\boldsymbol{X},\boldsymbol{Y}) \leqslant \mathbb{H}(\boldsymbol{X}) + \mathbb{H}(\boldsymbol{Y}).$$

Proof. These inequalities are proved by applying Jensen's inequality to the concave function $x \mapsto x\log\frac{1}{x}$; see [1] for the details. □

Proof of Proposition 8.2. We may assume that K (and thus N) are sufficiently large, since otherwise the claim of the proposition is trivial. We adapt Tao's proof in [33] to the function field setting. Let $\varepsilon = K^{-0.1}$. It suffices to show that there exists $H \in [K, \exp(\exp(10K))]$ for which

$$|\mathbb{E}^{\log}_{G\in\mathcal{M}_{\leqslant N}} \mathbb{E}_{P\in\mathcal{P}_H} c_P a_1(G+PB_1)\cdots a_k(G+PB_k)(q^{\deg(P)}1_{P|G}-1)| \ll \varepsilon$$

uniformly for all choices of $c_P \in \mathbb{U}$. We discretize the functions a_i by defining $\tilde{a}_i(F)$ for each $F \in \mathbb{F}_q[t]$ to be $a_i(F)$ rounded to the nearest element in the Gaussian lattice $\varepsilon\mathbb{Z}[i]$, breaking ties using the lexicographic ordering, say. Then, it suffices to prove

$$|\mathbb{E}_{P\in\mathcal{P}_H} \mathbb{E}^{\log}_{G\in\mathcal{M}_{\leqslant N}} c_P \tilde{a}_1(G+PB_1)\cdots \tilde{a}_k(G+PB_k)(q^{\deg(P)}1_{P|G}-1)| \ll \varepsilon \tag{49}$$

for some H as above and for any $c_P \in \mathbb{U}$. Since each polynomial $G \in \mathcal{M}_{\leqslant N}$ of degree $\geqslant H$ belongs to the same number of short intervals $I_H(G_0)$, where G_0 ranges through $\mathcal{M}_{\leqslant N}$, and $H/N \ll \varepsilon$, the left-hand side of (49) can be rewritten as

$$|\mathbb{E}^{\log}_{G_0\in\mathcal{M}_{\leqslant N}} \mathbb{E}_{\substack{G\in\mathcal{M}_{\leqslant N}\\ G\in I_H(G_0)}} \mathbb{E}_{P\in\mathcal{P}_H} c_P \tilde{a}_1(G+PB_1)\cdots \tilde{a}_k(G+PB_k)(q^{\deg(P)}1_{P|G}-1)| + O(\varepsilon). \tag{50}$$

Let $(\Omega, \mathbb{P})$ be the probability space where $\Omega = \mathcal{M}_{\leqslant N}$ and $\mathbb{P}$ is the probability measure

$$\mathbb{P}(A) := \mathbb{E}^{\log}_{G\in\mathcal{M}_{\leqslant N}} 1_A(G),$$

for any $A \subset \Omega$. Since each c_P and each map $\tilde{a}_i$ is uniformly bounded in absolute value, (50) can be bounded trivially by

$$\ll \varepsilon + \mathbb{P}(G_0 \in \mathcal{M}_{\leqslant N} : \ |\mathbb{E}_{\substack{G\in\mathcal{M}_{\leqslant N}\\ G\in I_H(G_0)}} \mathbb{E}_{P\in\mathcal{P}_H} c_P \tilde{a}_1(G+PB_1)\cdots\tilde{a}_k(G+PB_k)(q^{\deg(P)}1_{P|G}-1)| \geqslant \varepsilon). \tag{51}$$

Let $b := \max_{j\leqslant k}\deg(B_j)$. Introduce the random variables $\boldsymbol{X}_H$ and $\boldsymbol{Y}_H$ defined on Ω and given by

$$\boldsymbol{X}_H(G_0) := (\tilde{a}_1(F),\dots,\tilde{a}_k(F))_{F\in I_{H+b}(G_0)}, \quad \boldsymbol{Y}_H(G_0) := (G_0 \bmod P)_{P\in\mathcal{P}_H}, \quad G_0 \in \Omega.$$

Then, there is a deterministic function $\mathcal{F}$ such that we can write the probability in (51) as

$$\mathbb{P}(G_0 \in \mathcal{M}_{\leqslant N} : \ |\mathcal{F}(\boldsymbol{X}_H(G_0), \boldsymbol{Y}_H(G_0))| \geqslant \varepsilon);$$

more precisely, $\mathcal{F}$ is of the form

$$\mathcal{F}(x,y) = \mathbb{E}_{P\in\mathcal{P}_H} c_P \mathcal{Z}_P(x,y) := \mathbb{E}_{P\in\mathcal{P}_H} c_P \mathbb{E}_{\deg(J)<H+b} \phi_J(x,P)(q^{\deg(P)}1_{P|y+J}-1) \tag{52}$$

for some one-bounded functions ϕ_J and for $x \in \mathcal{X}_H, y \in \mathcal{Y}_H$, where $\mathcal{X}_H, \mathcal{Y}_H$ are the ranges of $\boldsymbol{X}_H, \boldsymbol{Y}_H$, respectively. Therefore, by the triangle inequality we have the bound $|\mathcal{Z}_P(x,y)| \leqslant 2$ for all $x \in \mathcal{X}_H, y \in \mathcal{Y}_H$.

It suffices to show that $\mathbb{P}(|\mathcal{F}(\boldsymbol{X}_H, \boldsymbol{Y}_H)| \geqslant \varepsilon) \ll \varepsilon$, for some $H \in [K, \exp(\exp(10K))]$. To do this, we start by bounding the probabilities $\mathbb{P}(|\mathcal{F}(x, \boldsymbol{Y}_H)| \geqslant \varepsilon)$ without conditioning and then we will deduce a bound on the corresponding conditional probabilities with $\boldsymbol{X}_H = x$.

By the Chinese remainder theorem, $\boldsymbol{Y}_H(F) = y$ for any $y \in \mathcal{Y}_H$ corresponds to a unique congruence for F modulo $\prod_{P\in\mathcal{P}_H} P$. Thus, this happens with probability exactly equal to $q^{-\sum_{P\in\mathcal{P}_H}\deg(P)}$ as long as

$$\sum_{P\in\mathcal{P}_H} \deg(P) < N,$$

which by the prime polynomial theorem holds whenever $H \leqslant \frac{\log N}{4\log q}$ for N large enough, say. Hence, $\boldsymbol{Y}_H$ is a uniform random variable on $\mathcal{Y}_H$ under the aforementioned condition. In particular, all the random variables $G_0 \mapsto G_0 \bmod P$ for $P \in \mathcal{P}_H$ are jointly independent of each other. By (52), we may write

$$\mathcal{F}(x, \boldsymbol{Y}_H) = \mathbb{E}_{P\in\mathcal{P}_H} c_P \mathcal{Z}_P(x, \boldsymbol{Y}_H),$$

and the random variables $\{\mathcal{Z}_P(x, \boldsymbol{Y}_H) : \ P \in \mathcal{P}_H\}$ are jointly independent, all having mean 0. Moreover, the number of different P here is $\geqslant \frac{1}{2}q^H/H$, say, again by the prime polynomial theorem. By

Hoeffding's inequality [13], there is an absolute constant $C > 0$ such that

$$\mathbb{P}(|\mathcal{F}(x, \boldsymbol{Y}_H)| \geqslant \varepsilon) = \mathbb{P}(|\mathbb{E}_{P \in \mathcal{P}_H} c_P \mathcal{Z}_P(x, \boldsymbol{Y}_H)| \geqslant \varepsilon) \leqslant \exp(-C\varepsilon^2 q^H/H) \tag{53}$$

for any $x \in \mathcal{X}_H$.

To bound the conditional probability $\mathbb{P}(|\mathcal{F}(x, \boldsymbol{Y}_H)| \geqslant \varepsilon | \boldsymbol{X}_H = x)$, we use a Pinsker-type inequality from [34]. This is applicable since $\boldsymbol{Y}_H$ is a uniform random variable. We get

$$\mathbb{P}(|\mathcal{F}(x, \boldsymbol{Y}_H)| \geqslant \varepsilon | \boldsymbol{X}_H = x) \leqslant \frac{\mathbb{H}(\boldsymbol{Y}_H) - \mathbb{H}(\boldsymbol{Y}_H | \boldsymbol{X}_H = x) + \log 2}{\log \frac{1}{\mathbb{P}(|\mathcal{F}(x, \boldsymbol{Y}_H)| \geqslant \varepsilon)}}.$$

Since $H \geqslant K$ and K is large, we may bound this from above using (53) and the prime polynomial theorem, obtaining

$$\leqslant \varepsilon + C^{-1}\varepsilon^{-2} \frac{\mathbb{H}(\boldsymbol{Y}_H) - \mathbb{H}(\boldsymbol{Y}_H | \boldsymbol{X}_H = x)}{q^H/H}. \tag{54}$$

Recalling that

$$\mathbb{P}(|\mathcal{F}(\boldsymbol{X}_H, \boldsymbol{Y}_H)| \geqslant \varepsilon) = \sum_{x \in \mathcal{X}_H} \mathbb{P}(|\mathcal{F}(x, \boldsymbol{Y}_H)| \geqslant \varepsilon | \boldsymbol{X}_H = x) \mathbb{P}(\boldsymbol{X}_H = x),$$

we multiply the bound in (54) by $\mathbb{P}(\boldsymbol{X}_H = x)$ and sum over $x \in \mathcal{X}_H$ to get

$$\mathbb{P}(|\mathcal{F}(\boldsymbol{X}_H, \boldsymbol{Y}_H)| \geqslant \varepsilon) \leqslant \varepsilon + C^{-1}\varepsilon^{-2} \frac{\mathbb{H}(\boldsymbol{Y}_H) - \mathbb{H}(\boldsymbol{Y}_H | \boldsymbol{X}_H)}{q^H/H} = \varepsilon + C^{-1}\varepsilon^{-2} \frac{\mathbb{I}(\boldsymbol{X}_H, \boldsymbol{Y}_H)}{q^H/H}$$

by the definition of mutual information $\mathbb{I}(\boldsymbol{X}_H, \boldsymbol{Y}_H)$ from Definition 8.3. Now what remains to be shown is that

$$\mathbb{I}(\boldsymbol{X}_H, \boldsymbol{Y}_H) \leqslant \varepsilon^3 \frac{q^H}{H} \tag{55}$$

holds for some H satisfying the conditions in Proposition 8.2. We will prove (55) by appealing to Shannon's inequality (Lemma 8.4) and pigeonholing in the parameter H.

Consider the conditional entropy

$$\mathbb{H}(\boldsymbol{X}_{H+j} | \boldsymbol{Y}_H)$$

for $H, j \leqslant \frac{\log N}{4 \log q}$, say. We may write

$$\boldsymbol{X}_{H+j} = \bigotimes_{\deg(M) \leqslant j} \boldsymbol{X}_H^{(M)},$$

where each $\boldsymbol{X}_H^{(M)}$ is a shifted copy of $\boldsymbol{X}_H$ given by

$$\boldsymbol{X}_H^{(M)}(G_0) := \boldsymbol{X}_H(G_0 + Mt^{H+b}).$$

Define also

$$\boldsymbol{Y}_H^{(M)}(G_0) := \boldsymbol{Y}_H(G_0 + Mt^{H+b}).$$

Then, by Shannon's inequality

$$\mathbb{H}(\boldsymbol{X}_{H+j}|\boldsymbol{Y}_H) \leqslant \mathbb{H}(\boldsymbol{X}_{H+j}, \boldsymbol{Y}_H) \leqslant \sum_{\deg(M)\leqslant j} \mathbb{H}(\boldsymbol{X}_H^{(M)}|\boldsymbol{Y}_H).$$

Since the sigma algebra given by $\boldsymbol{Y}_H$ is shift-invariant and $\mathbb{P}$ is *almost* shift-invariant in the sense that

$$\sup_{A\subset\mathcal{M}_{\leqslant N}} |\mathbb{P}(G\in A) - \mathbb{P}(G+J\in A)| \leqslant \deg(J)/N$$

we obtain

$$\begin{aligned}\mathbb{H}(\boldsymbol{X}_{H+j}|\boldsymbol{Y}_H) &\leqslant \sum_{\deg(M)\leqslant j} \mathbb{H}(\boldsymbol{X}_H^{(M)}|\boldsymbol{Y}_H^{(M)}) + \varepsilon q^j \\ &\ll q^j \mathbb{H}(\boldsymbol{X}_H \mid \boldsymbol{Y}_H) + \varepsilon q^j \\ &= q^j \mathbb{H}(\boldsymbol{X}_H, \boldsymbol{Y}_H) - q^j \mathbb{H}(\boldsymbol{Y}_H) + \varepsilon q^j \\ &= q^j \mathbb{H}(\boldsymbol{X}_H) - q^j \mathbb{I}(\boldsymbol{X}_H, \boldsymbol{Y}_H) + \varepsilon q^j.\end{aligned}$$

On the other hand, from Shannon's inequality, we also have the lower bound

$$\mathbb{H}(\boldsymbol{X}_{H+j}|\boldsymbol{Y}_H) = \mathbb{H}(\boldsymbol{X}_{H+j}, \boldsymbol{Y}_H) - \mathbb{H}(\boldsymbol{Y}_H) \geqslant \mathbb{H}(\boldsymbol{X}_{H+j}) - \mathbb{H}(\boldsymbol{Y}_H).$$

Comparing the upper and lower bounds for $\mathbb{H}(\boldsymbol{X}_{H+j}|\boldsymbol{Y}_H)$, we now have

$$\frac{\mathbb{I}(\boldsymbol{X}_H, \boldsymbol{Y}_H)}{q^H} \leqslant \frac{\mathbb{H}(\boldsymbol{X}_H)}{q^H} - \frac{\mathbb{H}(\boldsymbol{X}_{H+j})}{q^{H+j}} + \frac{\mathbb{H}(\boldsymbol{Y}_H) + \varepsilon q^j}{q^{H+j}}. \tag{56}$$

Since $\boldsymbol{Y}_H$ is a uniform random variable, we have $\mathbb{H}(\boldsymbol{Y}_H) = \log|\mathcal{Y}_H| \leqslant 2q^H$ by the prime polynomial theorem. Since $\boldsymbol{X}_H$ has kq^{H+b} components, each taking values in $\varepsilon\mathbb{Z}[i]\cap\mathbb{U}$, we have $\mathbb{H}(\boldsymbol{X}_H) \leqslant 10(\log\frac{1}{\varepsilon})kq^{H+b} \leqslant C_{k,b}\varepsilon^{-1}q^H$ for some $C_{k,b} > 0$. Now, if we denote $w_H := \mathbb{H}(\boldsymbol{X}_H)/q^H$, then from (56) we have the information bound

$$\frac{\mathbb{I}(\boldsymbol{X}_H, \boldsymbol{Y}_H)}{q^H} \leqslant w_H - w_{H+j} + \frac{\varepsilon}{q^H},$$

and $w_H \in [0, \varepsilon^{-1}C_{k,b}]$. Suppose that (55) failed for all $H \in [K, \exp(\exp(K/2))]$. Then, we would have

$$\frac{\varepsilon^3}{H} \leqslant w_H - w_{H+j} + \frac{\varepsilon}{q^H} + 2q^{-j} \tag{57}$$

for all $H \in [K, \exp(\exp(K/2))]$, $j \leqslant \frac{\log N}{4\log q}$. Define $H_1, H_2, \ldots$ recursively by $H_1 = \lceil K \rceil$ and $H_{r+1} := H_r + 2\log H_r + 1000\log\frac{1}{\varepsilon}$. Then, $H_r \leqslant \exp(\exp(K/2))$ for $r \leqslant \exp(\exp(K/3))$, say. Telescoping (57) with $H = H_r$ and $j = H_{r+1} - H_r$ then yields

$$\sum_{r\leqslant\exp(\exp(K/3))} \frac{\varepsilon^3}{2H_r} \leqslant \sum_{r\leqslant\exp(\exp(K/3))} (w_{H_r} - w_{H_{r+1}} + \frac{\varepsilon}{q^{H_r}} + 2q^{H_r - H_{r+1}}) \leqslant \varepsilon^{-1}C_{k,b} + 1.$$

Since, by telescoping, we have

$$H_r = H_1 + 2\sum_{1\leqslant j\leqslant r-1} \log H_j + 1000(r-1)\log(1/\varepsilon),$$

by induction on r we find that $H_r \leqslant C_0(r\log r + r\log\frac{1}{\varepsilon})$ for some absolute constant $C_0 > 0$ whenever $r \geqslant K \geqslant 1000\log(1/\varepsilon)$. Therefore,

$$\varepsilon^4 \sum_{K\leqslant r\leqslant\exp(\exp(K/3))} \frac{1}{10C_0 r\log r} \leqslant C_{k,b} + \varepsilon.$$

However, given our choice $\varepsilon = K^{-0.1}$, the left-hand side is

$$\gg K^{-0.4}(K - O(\log K)),$$

which is a contradiction for K large enough. This completes the proof. □

Now that we have established Proposition 8.2, which relates one-variable correlations to two-variable ones, we can apply the circle method to complete the proof of Theorem 1.4.

Proposition 8.5. *Assume the hypotheses of Proposition 8.1 and let H be chosen as in the conclusion of that proposition. Let $f_1, f_2 : \mathcal{M} \to \mathbb{U}$ be multiplicative functions. Set $H' := H + \deg(B)$. Then, for any $\varepsilon > 0$,*

$$\begin{aligned}&\mathbb{E}^{\log}_{G\in\mathcal{M}_{\leqslant N}} \left|\mathbb{E}_{P\in\mathcal{P}_H} c_P \mathbb{E}_{\deg(J)<H'} f_1(G+J) f_2(G+J+PB)\right| \\ &\ll \varepsilon^{-8}\left((\log H)H^{-1} + N^{-1/(100\log q)} + e^{-M_{Hayes}(f_1;N/H,H')/100}\right) + \varepsilon^2.\end{aligned} \tag{58}$$

Moreover, if f_1 is real-valued and q is odd, we may replace M_{Hayes} with M_{Dir} in (58).

Proof. Let $\mathcal{T}$ denote the expression on the left-hand side in (58). Fix $G \in \mathcal{M}_{\leqslant N}$ for the time being. For each $j = 1, 2$, define the sequence $x_{j,J} := f_j(G+J)$ for all $\deg(J) < H + \deg(B)$. For each $G \in \mathcal{M}_{\leqslant N}$, consider the double sum

$$\mathcal{T}_G := \mathbb{E}_{P\in\mathcal{P}_H} c_P \mathbb{E}_{\deg(J)<H'} x_{1,J}\overline{x}_{2,J+PB},$$

noting that $\mathcal{T} = \mathbb{E}^{\log}_{G\in\mathcal{M}_{\leqslant N}} |\mathcal{T}_G|$. We may view the set of polynomials J with $\deg(J) < H'$ as the representatives of residue classes modulo $t^{H'}$, and thus extending the sequences $\{x_{1,J}\}_J$ and $\{x_{2,J}\}_J$

periodically in modulo $t^{H'}$, we can consider them as maps on $\mathbb{F}_q[t]/(t^{H'}\mathbb{F}_q[t])$. We may thus expand these sequences in the corresponding Fourier basis, giving in the inner sum over J:

$$\begin{aligned}
&\mathbb{E}_{\deg(J)<H'} x_{1,J}\overline{x}_{2,J+PB} \\
&= \sum_{\xi_1,\xi_2 \bmod t^{H'}} \hat{x}_1(\xi_1)\overline{\hat{x}_2(\xi_2)} e_{\mathbb{F}}(-\xi_2 PB/t^{H'}) \mathbb{E}_{J \bmod t^{H'}} e_{\mathbb{F}}\left(\frac{J}{t^{H'}}(\xi_1-\xi_2)\right) \\
&= \sum_{\xi \bmod t^{H'}} \hat{x}_1(\xi)\overline{\hat{x}_2(\xi)} e_{\mathbb{F}}(-\xi PB/t^{H'}),
\end{aligned}$$

where we have defined

$$\hat{x}_j(\xi) := q^{-H'} \sum_{J \bmod t^{H'}} x_{j,J} e_{\mathbb{F}}(-J\xi/t^{H'}).$$

Inserting this into the definition of $\mathcal{T}_G$ thus gives

$$\mathcal{T}_G = \sum_{\xi \bmod t^{H'}} \hat{x}_1(\xi)\overline{\hat{x}_2(\xi)} \cdot \mathbb{E}_{P\in\mathcal{P}_H} c_P e_{\mathbb{F}}(-\xi PB/t^{H'}).$$

Now, define the large spectrum set

$$\Xi_H := \{\xi \bmod t^{H'} : \left|\mathbb{E}_{P\in\mathcal{P}_H} c_P e_{\mathbb{F}}(-\xi PB/t^{H'})\right| \geqslant \varepsilon^2\}.$$

We decompose $\mathcal{T}_G = \mathcal{T}_{G,s} + \mathcal{T}_{G,l}$, where

$$\begin{aligned}
\mathcal{T}_{G,s} &:= \sum_{\xi\notin\Xi_H} \hat{x}_1(\xi)\overline{\hat{x}_2(\xi)} \cdot \mathbb{E}_{P\in\mathcal{P}_H} c_P e_{\mathbb{F}}(-\xi PB/t^{H'}), \\
\mathcal{T}_{G,l} &:= \sum_{\xi\in\Xi_H} \hat{x}_1(\xi)\overline{\hat{x}_2(\xi)} \cdot \mathbb{E}_{P\in\mathcal{P}_H} c_P e_{\mathbb{F}}(-\xi PB/t^{H'}).
\end{aligned}$$

If $\xi \notin \Xi_H$, then we can bound the inner sum over P by ε^2. It follows from the Cauchy–Schwarz inequality and Plancherel's theorem that

$$\begin{aligned}
|\mathcal{T}_{G,s}| &\ll \varepsilon^2 \sum_{\xi\notin\Xi_H} |\hat{x}_1(\xi)||\hat{x}_2(\xi)| \leqslant \varepsilon^2 \prod_{j=1}^{2} \left(\sum_{\xi \bmod t^{H'}} |\hat{x}_j(\xi)|^2\right)^{\frac{1}{2}} \\
&= \varepsilon^2 \prod_{j=1}^{2} \left(q^{-H'} \sum_{J \bmod t^{H'}} |x_{j,J}|^2\right)^{\frac{1}{2}} \ll \varepsilon^2.
\end{aligned}$$

It remains to consider the case $\xi \in \Xi_H$. In this case, bounding the exponential sum in P trivially, this contribution is

$$|\mathcal{T}_{G,I}| \ll \sum_{\xi \in \Xi_H} |\hat{x}_1(\xi)||\hat{x}_2(\xi)|.$$

Note that $\|\hat{x}_j\|_\infty \leqslant 1$ for $j = 1, 2$. Averaging over $G \in \mathcal{M}_{\leqslant N}$ yields

$$\mathcal{T} \ll \varepsilon^2 + \sum_{\xi \in \Xi_H} \mathbb{E}^{\log}_{G \in \mathcal{M}_{\leqslant N}} q^{-H'} \left| \sum_{\deg(J) < H'} f_1(G+J) e_{\mathbb{F}}(-\xi J / t^{H'}) \right|$$

$$\leqslant \varepsilon^2 + |\Xi_H| \max_{\alpha \in \mathbb{T}} \mathbb{E}^{\log}_{G_0 \in \mathcal{M}_{\leqslant N}} \left| q^{-H'} \sum_{\substack{G \in \mathcal{M}_{\deg(G_0)} \\ G \in I_{H'}(G_0)}} f_1(G) e_{\mathbb{F}}(G\alpha) \right|.$$

To estimate $|\Xi_H|$, we use a fourth moment estimate. Indeed,

$$|\Xi_H| \leqslant \frac{\varepsilon^{-8}}{|\mathcal{P}_H|^4} \sum_{\xi \bmod t^{H'}} \left| \sum_{P \in \mathcal{P}_H} c_P e(-PB\xi / t^{H'}) \right|^4$$

$$\ll \varepsilon^{-8} H^4 q^{-4H} \left| \sum_{P_1,P_2,P_3,P_4 \in \mathcal{P}_H} c_{P_1} c_{P_2} \overline{c_{P_3} c_{P_4}} \sum_{\xi \bmod t^{H'}} e_{\mathbb{F}}(-B(P_1+P_2-P_3-P_4)\xi / t^{H'}) \right|$$

$$\ll \varepsilon^{-8} H^4 q^{-3H} \sum_{\substack{P_1,P_2,P_3,P_4 \in \mathcal{P}_H \\ P_1+P_2=P_3+P_4}} 1,$$

since $\deg(B(P_1+P_2-P_3-P_4)) < H'$. By Lemma 4.13, the sum over tuples (P_1, P_2, P_3, P_4) above is bounded by $O(q^{3H}/H^4)$, and hence $|\Xi_H| \ll \varepsilon^{-8}$. Splitting the average in $G_0 \in \mathcal{M}_{\leqslant N}$ according to degree, we get

$$\mathcal{T} \ll \varepsilon^2 + \varepsilon^{-8} \frac{1}{N} \sum_{k \leqslant N} \max_{\alpha \in \mathbb{T}} q^{-k} \sum_{G_0 \in \mathcal{M}_k} \left| q^{-H'} \sum_{\substack{G \in \mathcal{M}_k \\ G \in I_{H'}(G_0)}} f_{j_0}(G) e_{\mathbb{F}}(G\alpha) \right|.$$

The inner sum is trivially bounded as $\ll 1$ for $1 \leqslant k \leqslant N/H$, which contributes a term of size $\ll H^{-1}$. Since $H < N^{1/4}$, say, for each $N/H < k \leqslant N$ we may apply Theorem 7.1 to get

$$q^{-k} \sum_{G_0 \in \mathcal{M}_k} \left| q^{-H'} \sum_{\substack{G \in \mathcal{M}_k \\ G \in I_{H'}(G_0)}} f_{j_0}(G) e_{\mathbb{F}}(G\alpha) \right| \ll (\log H) H^{-1} + N^{-1/(2000 \log q)} + e^{-M_{\text{Hayes}}(f_1; N/H, H')/100}$$

in this range. Averaging this estimate over $N/H < k \leqslant N$ gives

$$\mathcal{T} \ll \varepsilon^2 + \varepsilon^{-8}\Big((\log H)H^{-1} + N^{-1/(2000\log q)} + e^{-M_{\mathrm{Hayes}}(f_1;N/H,H')/100}\Big).$$

This implies the first claim.

The second claim is proved in an identical manner, except that at the end we appeal to Theorem 7.2. □

Proof of Theorem 1.4. Let W be fixed but large, and let $K = W/100$. Set

$$\varepsilon := \min\{e^{-M_{\mathrm{Hayes}}(f_{j_0};N/H,H+\deg(B))/1600}, K^{-0.1}\},$$

where H is chosen as in Proposition 8.1. Combining Propositions 8.1 and 8.5, we find

$$\left|\frac{1}{N}\sum_{G\in\mathcal{M}_{\leqslant N}} q^{-\deg(G)} f_1(G)f_2(G+B)\right| \ll K^{0.8}\Big(N^{-1/(2000\log q)} + (\log H)H^{-1}\Big) + e^{-M_{\mathrm{Hayes}}(f_1;N/H,H+\deg(B))/200} + K^{-0.1},$$

where $H \in [K, \exp(\exp(10K)))]$ is chosen as in Proposition 8.1. Since f_1 is Hayes non-pretentious to level W and $H + \deg(B) \leqslant 2H < (\log N)/(2\log q) \leqslant \log N$, it follows that $M_{\mathrm{Hayes}}(f_1; N/H, H + \deg(B)) \to \infty$ as $N \to \infty$. Since $H \geqslant K \geqslant W/100$ the above is $o_{W\to\infty}(1)$ as $N \to \infty$, and letting W tend to infinity very slowly in terms of N, the first part of Theorem 1.4 follows.

Consider then the second part of the theorem, where f_1 is real-valued and q is odd. Applying the same argument as before, save that M_{Hayes} is replaced in every instance by M_{Dir}, we see that (4) holds unless there exists an infinite sequence $N_j \to \infty$, Dirichlet characters ψ_j mod M_j with $\deg(M_j) = O(1)$, and $\theta_j \in [0,1]$ such that

$$\mathbb{D}(f_1, \psi_j e_{\theta_j}; N_j) = O(1). \tag{59}$$

If (59) holds, then by the pretentious triangle inequality also

$$\mathbb{D}(f_1^2, \psi_j^2 e_{2\theta_j}; N_j) = O(1).$$

By pigeonholing, we may assume that $\psi_j = \psi$ is independent of j. Moreover, by passing to a subsequence, we may assume that θ_j converges to some $\theta \in [0,1]$. Then,

$$\mathbb{D}(f_1^2, \psi^2 e_{2\theta}; N_j) = O(1), \tag{60}$$

since by (59) and the triangle inequality we have

$$\mathbb{D}(e_\theta, e_{\theta_j}; N_j) = \limsup_{k\to\infty} \mathbb{D}(e_{\theta_{j+k}}, e_{\theta_j}; N_j) \leqslant \mathbb{D}(f_1\overline{\psi}, e_{\theta_j}; N_j) + \limsup_{k\to\infty} \mathbb{D}(f_1\overline{\psi}, e_{\theta_{j+k}}; N_{j+k}) = O(1).$$

Assume first that $\mathbb{D}(1, f_1^2; \infty) < \infty$. Then by another application of the pretentious triangle inequality, we deduce that

$$\mathbb{D}(1, \psi^2 e_{2\theta}; N_j) = O(1). \tag{61}$$

By Lemma 3.2, this implies that ψ^2 is principal, so we may assume that $\psi^2 \equiv 1$ in (61). Then arguing as in [9, p. 15] we have

$$\mathbb{D}(1, e_{2\theta}; N_j)^2 = \log N_j - \sum_{n \leqslant N_j} \frac{\cos(4\pi\theta n)}{n} + O(1) = \log(\max\{N_j \|2\theta\|, 1\}) + O(1),$$

which in view of (61) implies that $2\theta \equiv 0 \bmod 1$. But this contradicts (5). Hence, we must have $\mathbb{D}(1, f_1^2; \infty) = \infty$. But as f_1^2 is non-negative, we have

$$\mathbb{D}(f_1^2, \psi^2 e_{2\theta}; N_j)^2 \geqslant \sum_{P \in \mathcal{P}_{\leqslant N_j}} \frac{1 - |f_1(P)\overline{\psi}(P)e_{-\theta}(P)|^2}{q^{\deg(P)}} \geqslant \sum_{P \in \mathcal{P}_{\leqslant N_j}} \frac{1 - f_1(P)^2}{q^{\deg(P)}} = \mathbb{D}(f_1^2, 1; N_j)^2,$$

so upon letting $j \to \infty$ this contradicts (60). The claim follows. □

9 | A CONJECTURE OF KÁTAI IN FUNCTION FIELDS

In this section, we establish Theorem 1.7 as an application of our two-point Elliott conjecture result (Theorem 1.4). Since short interval characters and Archimedean characters satisfy

$$\xi(QG + 1) = \xi(QG) = \xi(Q)\xi(G) \quad \text{and} \quad e_\theta(QG + 1) = e_\theta(Q)e_\theta(G) \tag{62}$$

whenever $\deg(QG)$ is sufficiently large relative to $\text{len}(\xi)$, the function $f = \xi e_\theta$ clearly obeys (8) for suitably chosen $z \in S^1$. Thus, the essence of Theorem 1.7 lies in showing that there are no other such functions.

Before beginning with the proof of Theorem 1.7, we state the following useful proposition.

Proposition 9.1 (Concentration inequality for multiplicative functions). *Let $f : \mathcal{M} \to \mathbb{U}$ be a multiplicative for which $\mathbb{D}(f, 1; N) \ll 1$ as $N \to \infty$, and let $\varepsilon > 0$. Then, there is an infinite increasing sequence $\{M_j\}_{j \geqslant 1} \subset [1, \infty)$, depending only on f and ε, for which the following holds:*

Let $W \in \mathcal{M}$ satisfy $P \mid W$ for all $P \in \mathcal{P}_{\leqslant M_j}$. Then, for any B coprime to W and of degree $< \deg(W)$, and for $N := M_{j+j'}$ with j' large enough as a function of M_j and ε, we have

$$\sum_{G \in \mathcal{M}_{\leqslant N}} |f(WG + B) - 1| \ll q^N \big(\varepsilon + \mathbb{D}(f, 1; M_j, \infty) + o_{j \to \infty}(1)\big).$$

To prove this proposition, we begin with the following general lemma. In the sequel, given scales $1 \leqslant A < B$, we define

$$\mathfrak{T}_f(A, B) := \sum_{\substack{P \in \mathcal{P} \\ A < \deg(P) \leqslant B}} \frac{\text{Im}(f(P))}{q^{\deg(P)}}.$$

Lemma 9.2. *Let $N \geqslant M \geqslant 1$ and let $f : \mathcal{M} \to \mathbb{U}$ be multiplicative. Let $W \in \mathcal{M}$ satisfy $P \mid W$ for all $P \in \mathcal{P}_{\leqslant M}$. Then, for any B coprime to W and of degree $< \deg(W)$, and for N large enough as a function of M, we have*

$$\sum_{G \in \mathcal{M}_{\leqslant N}} |f(WG+B) - e^{\Im_f(M,N)}| \ll q^N \max_{\beta \in \{1,2\}} \mathbb{D}(f,1;M,\infty)^\beta + o_{M\to\infty}(q^N).$$

Proof. Let $h : \mathcal{M} \to \mathbb{C}$ be the additive function given by $h(P^\alpha) = f(P^\alpha) - 1$. Note that $\mathrm{Re}(h(P^\alpha)) \leqslant 0$, so that $e^{h(P^\alpha)} \in \mathbb{U}$ for all P and $\alpha \geqslant 1$.

We apply the Taylor approximation

$$z = e^{z-1} + O(|z-1|^2), \text{ for } |z| \leqslant 1$$

with $z = f(P^\alpha) = 1 + h(P^\alpha)$ for $P \in \mathcal{P}$ and $\alpha \geqslant 1$, together with the simple inequality

$$|z_1 \cdots z_k - w_1 \cdots w_k| \leqslant \sum_{1 \leqslant j \leqslant k} |z_j - w_j|,$$

valid whenever $z_j, w_j \in \mathbb{U}$ for all $1 \leqslant j \leqslant k$ (with z_j and w_j, respectively, playing the roles of $e^{h(P^\alpha)}$ and $f(P^\alpha)$ here). Ultimately, this yields

$$f(WG+B) = e^{h(WG+B)} + O\bigg(\sum_{P^\alpha || WG+B} |h(P^\alpha)|^2\bigg).$$

Since $(B,W) = 1$, note that $P^\alpha \,||\, WG+B \Rightarrow \deg(P) > M$ and $P \nmid W$. Summing over $G \in \mathcal{M}_{\leqslant N}$ thus leads to

$$\sum_{G \in \mathcal{M}_{\leqslant N}} \Big| f(WG+B) - e^{h(WG+B)} \Big| \ll q^N \sum_{\substack{P \in \mathcal{P} \\ \deg(P) > M \\ \deg(P^\alpha) \leqslant N}} |h(P^\alpha)|^2 q^{-\deg(P^\alpha)}. \tag{63}$$

Next, set

$$A_h(Y,X) := \sum_{\substack{P \in \mathcal{P} \\ Y < \deg(P^\alpha) \leqslant X}} h(P^\alpha) q^{-\deg(P^\alpha)} (1 - q^{-\deg(P)}), \quad X > Y \geqslant 1.$$

Since $\mathrm{Re}(h(M)) \leqslant 0$ for all $M \in \mathcal{M}$, we have $\mathrm{Re}(A_h(M,N)) \leqslant 0$ as well, thus

$$|e^{h(WG+B)} - e^{A_h(M,N)}| \ll |h(WG+B) - A_h(M,N)|.$$

Summing this expression over $G \in \mathcal{M}_{\leqslant N}$, then applying the Cauchy–Schwarz inequality followed by the Turán–Kubilius inequality for h (see [4, Lemma 7] for the function field version of this[†]),

[†] In [4], the Turán–Kubilius inequality was stated for the linear forms $G \mapsto G + B$, but the same proof works for any linear forms $G \mapsto WG + B$.

we obtain

$$\sum_{G\in\mathcal{M}_{\leqslant N}} |e^{h(WG+B)} - e^{A_h(M,N)}| \ll q^{N/2}\left(\sum_{G\in\mathcal{M}_{\leqslant N}} |h(WG+B) - A_h(M,N)|^2\right)^{1/2}$$
$$\ll q^N\left(\sum_{\substack{P\in\mathcal{P}\\ \deg(P)>M\\ \deg(P^\alpha)\leqslant N}} |h(P^\alpha)|^2 q^{-\deg(P^\alpha)}\right)^{1/2}.$$

We note that

$$\sum_{\substack{P\in\mathcal{P}\\ \deg(P)>M\\ \deg(P^\alpha)\leqslant N}} |h(P^\alpha)|^2 q^{-\deg(P^\alpha)} = \sum_{\substack{P\in\mathcal{P}\\ M<\deg(P)\leqslant N}} |1-f(P^\alpha)|^2 q^{-\deg(P^\alpha)} + O(M^{-1/2}),$$

and this simplifies to $2\mathbb{D}(f,1;M,N)^2 + O(M^{-1/2})$. Combining this with (63), we thus find that

$$\sum_{G\in\mathcal{M}_{\leqslant N}} |f(WG+B) - e^{A_h(M,N)}| \ll q^N\Big(\mathbb{D}(f,1;M,\infty) + \mathbb{D}(f,1;M,\infty)^2 + M^{-1/4}\Big). \tag{64}$$

Now, observe that

$$A_h(M,N) = \sum_{\substack{P\in\mathcal{P}\\ M<\deg(P)\leqslant N}} \frac{\mathrm{Re}(f(P))-1}{q^{\deg(P)}} + i\sum_{\substack{P\in\mathcal{P}\\ M<\deg(P)\leqslant N}} \frac{\mathrm{Im}(f(P))}{q^{\deg(P)}} + O(M^{-1/2})$$
$$= -\mathbb{D}(f,1;M,N)^2 + i\mathfrak{T}_f(M,N) + O(M^{-1/2}).$$

When M is large enough, we thus have

$$e^{A_h(M,N)} = e^{i\mathfrak{T}_f(M,N)} + O(\mathbb{D}(f,1;M,\infty)^2 + M^{-1/2}),$$

which, when combined with (64) yields the claim. □

The following result allows us to pick suitable scales M and N in order to control the distribution of $\mathfrak{T}_f(M,N)$ mod 2π, and therefore the direction of $e^{i\mathfrak{T}_f(M,N)}$. Below, as usual we write $\|t\| := \min_{m\in\mathbb{Z}}|t-m|$ for $t\in\mathbb{R}$.

Lemma 9.3. *Let $\eta>0$. Then there is an infinite increasing sequence $\{M_j\}_{j\geqslant 1}\subset[1,\infty)$ such that*

$$\|\mathfrak{T}_f(M_j, M_{j+j'})/2\pi\| < \eta$$

for any choice of j, j' sufficiently large relative to η.

Proof. A proof of this claim appears implicitly in the proof of [22, Lemma 2.11] in the integer setting. We give here a shorter proof in the function field setting that would also be applicable (without change) over the integers.

Write $\mathfrak{T}_f(T) := \mathfrak{T}_f(1,T)$ for $T \geqslant 1$, and define $\mathfrak{T}_f(\infty) := \lim_{T\to\infty} \mathfrak{T}_f(T)$.

As $[0,1]$ is compact, the sequence $\{\mathfrak{T}_f(n)/2\pi \bmod 1\}_n$ has a limit point, say α. We select $\{M_j\}_{j\geqslant 1}$ to be a sequence for which $\frac{\mathfrak{T}_f(M_j)}{2\pi} \bmod 1 \to \alpha$. Let $j' \geqslant 1$. By the triangle inequality, we then have

$$\|\mathfrak{T}_f(M_j, M_{j+j'})/2\pi\| = \|\frac{\mathfrak{T}_f(M_{j+j'})}{2\pi} - \frac{\mathfrak{T}_f(M_j)}{2\pi}\| \leqslant \|\frac{\mathfrak{T}_f(M_j)}{2\pi} - \alpha\| + \|\alpha - \frac{\mathfrak{T}_f(M_{j+j'})}{2\pi}\| < \eta,$$

provided j is chosen large enough. □

Proof of Proposition 9.1. Let $\varepsilon > 0$. Applying Lemma 9.3 with $\eta = \varepsilon$, we may choose an infinite increasing sequence $\{M_j\}_{j\geqslant 1}$ such that, if j, j' are large then upon setting $M := M_j$ and $N := M_{j+j'}$ we find that

$$\left|e^{i\mathfrak{T}_f(M,N)} - 1\right| \ll \|\mathfrak{T}_f(M,N)\| \ll \varepsilon.$$

Combining this with Lemma 9.2 and the condition $\mathbb{D}(f,1;\infty) < \infty$, we deduce that

$$\sum_{G\in\mathcal{M}_{\leqslant N}} |f(WG+B) - 1| \ll q^N(\varepsilon + \mathbb{D}(f,1;M,\infty) + o_{M\to\infty}(1)),$$

which implies the claim. □

Proof of Theorem 1.7. By partial summation, if (S_n) is a non-negative sequence for which $q^{-N}\sum_{n\leqslant N} S_n = o(1)$, then $\frac{1}{N}\sum_{n\leqslant N} S_n/q^n = o(1)$. Thus, (8) implies

$$\sum_{G\in\mathcal{M}_{\leqslant N}} |f(QG+1) + zf(G)|/q^{\deg(G)} = o(N).$$

Since $|f(QG+1) + zf(G)| \leqslant 2$, this further gives

$$\sum_{G\in\mathcal{M}_{\leqslant N}} |f(QG+1) + zf(G)|^2/q^{\deg(G)} = o(N),$$

so that expanding the modulus squared and recalling that f is unimodular, we find

$$\sum_{G\in\mathcal{M}_{\leqslant N}} (1 + \mathrm{Re}(zf(G)\overline{f}(QG+1)))q^{-\deg(G)} = o(N), \tag{65}$$

We will use this in two ways as follows. First, since the summands are all $\in [0,2]$, for a logarithmic proportion $1 - o(1)$ of $G \in \mathcal{M}_{\leqslant N}$ we have

$$\mathrm{Re}(zf(G)\overline{f}(QG+1)) = -1 + o(1), \text{ that is, } f(G)\overline{f}(QG+1) = -\overline{z} + o(1), \tag{66}$$

by unimodularity. This will be applied shortly.

Second, from (65) and the triangle inequality we deduce that

$$1+o(1)\leqslant \frac{1}{N}\left|\mathrm{Re}\left(z\sum_{G\in\mathcal{M}_{\leqslant N}} f(G)\overline{f}(QG+1)q^{-\deg(G)}\right)\right|\leqslant \frac{1}{N}\left|\sum_{G\in\mathcal{M}_{\leqslant N}} f(G)\overline{f}(QG+1)q^{-\deg(G)}\right|. \quad (67)$$

By Theorem 1.4, (67) implies that for every $N\geqslant 1$ there exists a Dirichlet character χ_N of bounded conductor, a short interval character ξ_N of bounded length, and an angle $\theta_N\in[0,1]$ such that

$$\mathbb{D}(f,\chi_N\xi_N e_{\theta_N};N)\ll 1.$$

By pigeonholing, we may assume that $\chi_N=\chi$ and $\xi_N=\xi$ for some fixed Dirichlet character χ, short interval character ξ, and for an infinite sequence of integers N. Since the interval $[0,1]$ is compact, we may find an infinite strictly increasing subsequence (N_j) and a fixed $\theta\in[0,1)$ such that $\lim_{j\to\infty}\theta_{N_j}=\theta$ exists and

$$\mathbb{D}(f,\chi\xi e_{\theta_{N_j}};N_j)\ll 1. \quad (68)$$

By the triangle inequality and the fact that $N_j<N_{j+k}$, from (68) we see that

$$\mathbb{D}(e_{\theta_{N_j}},e_{\theta_{N_{j+k}}};N_j)\ll 1$$

uniformly for $k\geqslant 1$. Letting $k\to\infty$ yields

$$\mathbb{D}(e_{\theta_{N_j}},e_{\theta};N_j)\ll 1,$$

and hence

$$\mathbb{D}(f,\chi\xi e_{\theta};N_j)\leqslant \mathbb{D}(f,\chi\xi e_{\theta_{N_j}};N_j)+\mathbb{D}(e_{\theta},e_{\theta_{N_j}};N_j)\ll 1.$$

Since every N belongs to some interval $[N_j,N_{j+1})$, we finally see that

$$\mathbb{D}(f,\chi\xi e_{\theta};N)\ll 1$$

uniformly in N.

Let us now write

$$f(G)=\chi_1(G)\xi(G)f_1(G), \quad (69)$$

where χ_1 is the completely multiplicative function given at irreducibles P by $\chi_1(P)=\chi(P)$ if $P\nmid \mathrm{cond}(\chi)$ and $\chi_1(P)=1$ otherwise, and where f_1 satisfies $\mathbb{D}(f_1,1;N)\ll 1$.

Recalling (62), (66) gives

$$\chi_1 f_1(G)\overline{\chi_1 f_1}(QG+1)=z'+o(1) \quad (70)$$

for logarithmic proportion $1-o(1)$ of $G\in\mathcal{M}$ and some complex number $z'\in S^1$.

Suppose first that $\{P \in \mathcal{P}: \ f_1(P) \neq z'\chi_1 f_1(Q+1)\}$ is infinite. Let P_0 be an element of this set of degree $> \max\{\mathrm{cond}(\chi), \deg(Q)\}$, and let $\eta > 0$ be such that $|f_1(P_0) - z'\chi_1 f_1(Q+1)| > \eta$; since this condition becomes less stringent as η decreases, we may assume that η is smaller than any fixed constant. Let w be a large integer to be chosen shortly, subject in particular to the condition $w > \deg(Q)\mathrm{cond}(\chi)$. Consider the infinite sets

$$\mathcal{A} := \{G \in \mathcal{M}: \ G \equiv 1 \bmod \prod_{P \in \mathcal{P}_{\leqslant w} \setminus \{P_0\}} P^w, \ G \equiv P_0 \bmod P_0^2\},$$

$$\mathcal{B} := \{(QG+1)/(Q+1): \ G \in \mathcal{A}\}.$$

By the Chinese remainder theorem, the elements of $\mathcal{A}$ may be parameterized by an arithmetic progression $P_0(WG+B)$, where $W \in \mathcal{M}$ is divisible by all $P \in \mathcal{P}_{\leqslant w}$, and $B = B(P_0)$ is some residue class modulo W, necessarily coprime to W. Moreover, as $G-1$ is divisible by a $P^{\deg(Q)}$ for every $P|(Q+1)$, it follows that whenever $G \in \mathcal{A}$,

$$\frac{QG+1}{Q+1} = G - \frac{G-1}{Q+1} \in \mathcal{M}.$$

Hence, $\mathcal{B} \subset \mathcal{M}$, and as $\deg(P_0) > \deg(Q+1)$ the set $\mathcal{B}$ may similarly be parameterized as $P_0WG' + D$, for $D = D(P_0)$ coprime to P_0.

For $G \in \mathcal{A}$, we have $\chi_1(G) = 1$, $\chi_1(QG+1) = \chi_1(Q+1)$, and $G/P_0, (QG+1)/(Q+1)$ are both coprime to $\prod_{P \in \mathcal{P}_{\leqslant w}} P$. We apply Proposition 9.1 with $f = f_1$ and $\varepsilon = \eta^2$, say, along both $\mathcal{A}$ and $\mathcal{B}$: thus, we may find a common choice of parameters M, N (depending only on f_1 and η) such that, upon taking $w = M$, along both $\mathcal{A}$ and $\mathcal{B}$ a proportion $1 - o_{w\to\infty}(1)$ of $G \in \mathcal{M}_{\leqslant N}$ satisfy

$$|f_1(G) - f_1(P_0)| < \eta/10, \quad |f_1(QG+1) - f_1(Q+1)| < \eta/10.$$

Combined with (70) restricted to $\mathcal{A}$, we see that $|f_1(P_0) - z'\chi_1 f_1(Q+1)| < \eta$. However, this is a contradiction to our assumption, so $\{P \in \mathcal{P}: \ f_1(P) \neq z'\chi_1 f_1(Q+1)\}$ must be finite. Now, since f_1 pretends to be 1, we must have $z'\chi_1 f_1(Q+1) = 1$.

Now, let N_0 be such that $f_1(P) = 1$ whenever $P \in \mathcal{P}$, $\deg(P) \geqslant N_0$. Let M_0 be the modulus of χ. Let w' be large enough in terms of the aforementioned quantities, and set $W' = \prod_{P \in \mathcal{P}_{\leqslant w'}} P^{\max\{1, v_P(M_0)\}}$. Let C be arbitrary, subject to $(C, W') = 1$. By the Chinese remainder theorem, there exists a residue class $G_0 \bmod W'$ such that $G \equiv G_0 \bmod W'$ implies $G \equiv C \bmod M_0$ and $G \equiv 1 \bmod W'/M_0$. Thus, if $G = W'F$, for any $F \equiv G_0 \bmod W'$, then $f_1(F) = f_1(QG+1) = 1$, and thus

$$\chi_1(G)f_1(G) = \chi_1(W')f_1(W')\chi(C), \quad \chi_1(QG+1)f_1(QG+1) = f_1(QG+1) = 1.$$

By (70) restricted to such G, we conclude that

$$\chi_1(W')f_1(W')\chi(C) = z' + o(1).$$

But this implies that χ is constant on residue classes coprime to M_0, so χ is principal.

Now, (70) simplifies to

$$f_1(G) = z' f_1(QG+1) + o(1)$$

for logarithmic proportion $1-o(1)$ of G. Let us restrict to polynomials G of the form $G = W'F$, where W' is as above (in particular, $P \mid W'$ for $P \in \mathcal{P}_{\leqslant N_0}$). Since $(QW'F+1, W') = 1$, we deduce that $f_1(W')f_1(F) = z' + o(1)$ for logarithmic proportion $1-o(1)$ of $F \in \mathcal{M}$. Since $f_1(W') = \prod_{P\in\mathcal{P}_{\leqslant N_0}} f_1 s(P)^{\max\{1,\nu_P(M)\}}$, which is independent of w, there exists a constant c such that $f_1(F) = c + o(1)$ log-almost everywhere. But now if $P_0 \in \mathcal{P}$ is arbitrary, we can find an infinite sequence of polynomials G for which $f_1(P_0G) = f_1(G) + o(1)$, so $f_1(P_0) = 1 + o(1)$, which means that $f_1(P_0) = 1$. Thus, $f_1 \equiv 1$, and so $f = \xi e_\theta$. □

ACKNOWLEDGMENTS

This work began when the authors were in residence for the "Probability in Number Theory" Workshop at CRM in the spring of 2018, and continued in particular at the "Sarnak's Conjecture" workshop at AIM that winter. We would like to thank both institutions for their hospitality and for excellent working conditions.

We are grateful to the anonymous referee for a very careful reading of the paper and for numerous comments leading to improvements to the exposition. We would also like to thank Andrew Granville and Maksym Radziwiłł for their encouragement, and Ofir Gorodetsky for pointing out a correction to a smooth number estimate referenced in an earlier version of this paper. The third author was supported by a Titchmarsh Fellowship and Academy of Finland grant no. 340098.

JOURNAL INFORMATION

Mathematika is owned by University College London and published by the London Mathematical Society. All surplus income from the publication of *Mathematika* is returned to mathematicians and mathematics research via the Society's research grants, conference grants, prizes, initiatives for early career researchers and the promotion of mathematics.

REFERENCES

1. P. Billingsley, *Probability and measure*, 3rd ed., Wiley Series in Probability and Statistics, New York, 1995.
2. J. Bourgain, P. Sarnak, and T. Ziegler, *Disjointness of Moebius from horocycle flows*. In From Fourier analysis and number theory to Radon transforms and geometry, volume 28 of Dev. Math., Springer, New York, 2013, pp. 67–83.
3. J. Chinis, *Siegel zeros and Sarnak's conjecture*, *arXiv e-prints*, arXiv:2105.14653, May 2021.
4. P. Darbar and A. Mukhopadhyay, *Correlation of multiplicative functions over function fields*, *arXiv e-prints*, arXiv:1905.09303, May 2019.
5. P. Erdős, *On an elementary proof of some asymptotic formulae in the theory of partitions*, Ann. of Math. **43** (1942), 437–450.
6. D. Faifman and Z. Rudnick, *Statistics for zeros of hyperelliptic zeta functions*, Compositio Math. **146** (2010), 81–101.
7. O. Gorodetsky, *Mean values of arithmetic functions in short intervals and in arithmetic progressions in the large-degree limit*, Mathematika **66** (2020), no. 2, 373–394.
8. O. Gorodetsky and W. Sawin, *Correlation of arithmetic functions over* $\mathbb{F}_q[t]$, Math. Ann. **376** (2020), 1059–1106.
9. A. Granville, A. J. Harper, and K. Soundararajan, *Mean values of multiplicative functions over function fields*, Res. Number Theory (2015), Paper No. 25, 18 pp.
10. H. Halberstam and H.-E. Richert, *On a result of R. R. Hall*, J. Number Theory **11** (1979), no. 1, 76–89.
11. A. J. Harper and K. Soundararajan, *Lower bounds for the variance of sequences in arithmetic progressions: primes and divisor functions*, Q. J. Math. **68** (2017), no. 1, 97–123.

12. D. R. Hayes, *The distribution of irreducibles in GF*$[q, x]$, Trans. Amer. Math. Soc. **117** (1965), 101–127.
13. W. Hoeffding, *Probability inequalities for sums of bounded random variables*, J. Amer. Statist. Assoc. **58** (1963), 13–30.
14. C. Hooley, *On the Barban-Davenport-Halberstam theorem. IX*, Acta Arith. **83** (1998), no. 1, 17–30.
15. C.-N. Hsu, *Estimates for coefficients of L-functions fields*, Finite Fields Appl. **5** (1999), no. 1, 76–88.
16. I. Kátai, *Some problems in number theory*, Studia Sci. Math. Hungar. **16** (1983), no. 3-4, 289–295.
17. I. Kátai and B. M. Phong, *On the pairs of completely multiplicative functions satisfying some relation*, Acta Sci. Math. (Szeged) **85** (2019), no. 1-2, 139–145.
18. J. Keating and Z. Rudnick, *Squarefree polynomials and Möbius values in short intervals and arithmetic progressions*, Algebra Number Theory **10** (2016), no. 2, 375–420.
19. J. P. Keating and Z. Rudnick, *The variance of the number of prime polynomials in short intervals and in residue classes*, Int. Math. Res. Not. (2014), no. 1, 259–288.
20. O. Klurman, *Correlations of multiplicative functions and applications*, Compos. Math. **153** (2017), no. 8, 1622–1657.
21. O. Klurman, *Mean values and correlations of multiplicative functions: the "pretentious" approach*, Mathématiques Ph.D., Université de Montréal, 2017.
22. O. Klurman and A. Mangerel, *On the orbits of multiplicative pairs*, Algebra and Number Theory **14** (2020), no. 1, 155–189.
23. O. Klurman, A. P. Mangerel, and J. Teräväinen, *Multiplicative functions in short arithmetic progressions. arXiv e-prints*, arXiv:1909.12280, September 2019.
24. O. Klurman, A. P. Mangerel, and J. Teräväinen, *Beyond the Erdős discrepancy problem in function fields. arXiv e-prints*, arXiv:2202.10370, February 2022.
25. R. M. Kubota, *Waring's Problem in* $F_q[x]$, Ph.D. thesis, University of Michigan, 1971.
26. Y.-R. Liu and T. D. Wooley, *Waring's problem in function fields*, J. Reine Angew. Math. **638** (2010), 1–67.
27. K. Matomäki and M. Radziwiłł, *Multiplicative functions in short intervals*, Ann. of Math. (2) **183** (2016), no. 3, 1015–1056.
28. K. Matomäki and M. Radziwiłł, *Multiplicative functions in short intervals II*, *arXiv e-prints*, arXiv:2007.04290, July 2020.
29. K. Matomäki, M. Radziwiłł, and T. Tao, *An averaged form of Chowla's conjecture*, Algebra and Number Theory **9** (2015), 2167–2196.
30. G. Rhin, *Répartition modulo 1 dans un corps de séries formelles sur un corps fini*, Dissertationes Math. **95** (1972).
31. W. Sawin and M. Shusterman, *On the Chowla and twin primes conjectures over* $\mathbb{F}_q[T]$, Ann. of Math. (2) **196** (2022), no. 2, 457–506.
32. T. Tao, *The Erdős discrepancy problem*, Discrete Anal. **1** (2016), 29 p.
33. T. Tao, *The logarithmically averaged Chowla and Elliott conjectures for two-point correlations*, Forum Math. Pi **4e8** (2016), 36 p.
34. T. Tao and J. Teräväinen, *The structure of logarithmically averaged correlations of multiplicative functions, with applications to the Chowla and Elliott conjectures*, Duke Math. J. **168** (2019), no. 11, 1977–2027.
35. G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, Graduate Studies in Mathematics vol. 163, AMS Publications, Providence, RI, 2015.
36. J. Teräväinen, *On binary correlations of multiplicative functions*, Forum Math. Sigma **6** (2018), e10, 41.
37. R. Warlimont, *Arithmetical semigroups. II. Sieving by large and small prime elements. Sets of multiples*, Manuscripta Math. **71** (1991), no. 2, 197–221.
38. W. A. Webb, *Sieve methods for polynomial rings over finite fields*, J. Number Theory **16** (1983), no. 3, 343–355.
39. A. Weil, *Basic number theory*, 3rd ed., Springer, New York, Berlin, 1974.