The Raymond and Beverly Sackler
Faculty of Exact Sciences
School of Mathematical Sciences

The variance of sums of arithmetic functions

Thesis submitted in partial fulfillment of the requirements for the Ph.D. degree in the School of Mathematical Sciences, Tel Aviv University

by

Ofir Gorodetsky

Under the supervision of
Prof. Ze'ev Rudnick
Prof. Lior Bary-Soroker

January 2021

Abstract

Some of the fundamental problems in analytic number theory concern sums of arithmetic functions over short intervals. Such problems include the existence of primes in an interval, cancellation in Möbius sums and existence of squarefree integers in an interval.

It is conjectured for many naturally occurring functions that their mean in short intervals is asymptotic to their mean in a full interval. Moreover, sums over short intervals are expected to enjoy squareroot cancellation, in the sense that the error term is at most the squareroot of the number of terms. Even under the Riemann Hypothesis, these problems are open for most functions.

It is common and useful to introduce randomness into the problem, by picking a short interval at random. One can then study the sum of a arithmetic function over the random interval. Upper bounds on the variance of this random sum lead one to 'almost-all' versions of difficult conjectures.

In this thesis we study the variance of sums of arithmetic functions in function fields. We use combinatorial, analytic and geometric tools to prove stronger results than known over the integers. Our four main results are

1. An asymptotic formula for the variance of squarefree polynomials over short intervals, going well beyond R. R. Hall's 1982 work, and giving evidence to a recent conjecture of Keating and Rudnick.

2. A tight upper bound for the variance of factorization functions in short intervals.

3. An asymptotic formula for the variance of sums of two squares in short intervals, in the large-q limit. Our formula deviates from a naive probabilistic model and produces a prediction over the integers which agrees very well with numerical data.

4. A large-q asymptotic formula for the number of twin primes in a polynomial ring with an optimal error term.

Acknowledgments

I want to thank Ze'ev for his constant support and for always directing me to more interesting and fruitful areas of research. I have learned more from you about compelling pursuits in mathematics than I could ever put down in words.

Lior, you have my deep gratitude for always challenging me and encouraging me to aim higher in my work.

In mathematics, there are often porous boundaries between personal and professional lives. In that spirit, and with great appreciation, I must acknowledge my collaborators Brad and Zahi, who became my friends, and my friend Dor, who became my trusted collaborator.

This thesis could not have been written without the unwavering dedication of my remarkable family and beautiful wife Noga. You always give me what I need to do my best, and for this, I will always be thankful.

Contents

1 Introduction and results

1.1 Summary of results

Let $\mathbb{F}_q[T]$ be the polynomial ring over the finite field $\mathbb{F}_q$, $\mathcal{M}_{n,q} \subseteq \mathbb{F}_q[T]$ be the subset of monic polynomials of degree n and $\mathcal{M}_q = \cup_{n\geq 0}\mathcal{M}_{n,q}$ be the subset of all monic polynomials. Given a polynomial f, we define a short interval of size q^{h+1} around f to be

$$I(f,h) := \{f+g : \deg(g) \leq h\}.$$

Given a function $\alpha\colon \mathcal{M}_q \to \mathbb{C}$, its mean value over a finite subset $S \subseteq \mathcal{M}_q$ is

$$\langle \alpha \rangle_S := \frac{1}{\#S}\sum_{f\in S}\alpha(f).$$

Given $0 \leq h \leq n-1$, the variance of sums of α over intervals of size q^{h+1} around polynomials of degree n is defined as

$$\mathrm{Var}_\alpha(n,h) := \frac{1}{q^n}\sum_{f_0\in\mathcal{M}_{n,q}}\left|\sum_{f\in I(f_0,h)}\alpha(f) - q^{h+1}\langle\alpha\rangle_{\mathcal{M}_{n,q}}\right|^2.$$

For $\alpha = \mu_q^2$, the indicator of squareefree polynomials, we prove

Theorem 1 (Thm. 1.2)**.** *Fix q and $\varepsilon > 0$. As h, n tend to ∞ with $h/n \in (\varepsilon, 1/2-\varepsilon)$, we have*

$$\mathrm{Var}_{\mu_q^2}(n,h) \sim C_{q,h}q^{\frac{h}{2}}$$

where $C_{q,h} > 0$ depends only on q and the parity of h.

Previously this was only known to hold for $h/n < 2/9 - \varepsilon$ by work of R. R. Hall [Hal82], and it supports a recent conjecture of Keating and Rudnick [KR16]. See §1.9.1 for more background and results.

For the next theorem, we recall the notion of a factorization function. To $f \in \mathcal{M}_q$ that factorizes as $\prod_i P_i^{e_i}$ we associate the multiset $\omega_f = \{(\deg(P_i), e_i)\}_i$. We say that $\alpha\colon \mathcal{M}_q \to \mathbb{C}$ is a factorization function if $\alpha(f)$ is a function of ω_f.

Theorem 2 (Thm. 1.5)**.** *Fix q. Let $\alpha\colon \mathcal{M}_q \to \mathbb{C}$ be a factorization function. For $0 \leq h \leq n-1$ we have*

$$\mathrm{Var}_\alpha(n,h) \leq \max_{f\in\mathcal{M}_{n,q}}|\alpha(f)|^2 q^{h+1}e^{o_q(n)}$$

as $n \to \infty$.

Previously this was only known for specific functions, such as the von Mangoldt function. See §1.9.2 for further background.

Let $b_q\colon \mathcal{M}_q \to \{0,1\}$ be the indicator of polynomials of the form $A^2 + TB^2$. For real s define

$$G(s) := \mathbb{P}\Big(1 - \frac{s}{\alpha_1} \leq Y \leq \frac{s}{\alpha_1'}\Big),$$

for Y, α_1, α_1' independent random variables, with Y distributed as $\mathrm{Beta}(1/4, 1/4)$ and α_1, α_1' identically distributed copies of the largest part of the Thoma simplex distributed according to the spectral z-measure with parameters $1/2, 1/2$. (The spectral z-measure is defined in §6.6.)

Theorem 3 (Thm. 1.7, Prop. 1.8)**.** *Suppose $q = p^k$ for a fixed odd prime p. The limit*

$$T(n; n-h-1) := \lim_{k\to\infty} \frac{\mathrm{Var}_{b_q}(n,h)}{q^{h+1}}$$

exists when $0 \le h \le n - \max\{\sqrt{n}, 7\}$. Furthermore,

$$\lim_{N/n\to s} \sqrt{\pi n} T(n;N) = G(s)$$

for $s \in [0,1]$.

This theorem allows us to make a precise prediction, over the integers, for the variance of sums of two squares in short intervals. The prediction agrees very well with numerics. See §1.9.3 for more details.

Let $\Lambda_q \colon \mathcal{M}_q \to \mathbb{C}$ be the von Mangoldt function.

Theorem 4 (Thm. 1.10)**.** *Fix $n \ge 4$. We have*

$$\frac{\sum_{f\in\mathcal{M}_{n,q}} \Lambda_q(f)\Lambda_q(f+1)}{\#\mathcal{M}_{n,q}} = 1 + O_n\left(\frac{1}{q}\right) \tag{1.1}$$

as $q \to \infty$.

This is a large-q version of the twin prime conjecture, in quantitative form. Previously, (1.1) was known with error term $O_n(q^{-1/2})$ due to works of Pollack, Bender and Pollack, Bary-Soroker and Carmon [Pol08, BP09, BS14, Car15]. In fact, the error term $O_n(1/q)$ cannot be improved upon under the Hardy-Littlewood Prime Tuple Conjecture in $\mathbb{F}_q[T]$. For further discussion, see §1.9.4.

We now survey the background to these results and some of the previous literature.

1.2 Arithmetic functions

Arithmetic functions are functions from the positive integers $\mathbb{N}$ to $\mathbb{C}$. Some of the most fundamental questions in analytic number theory may be expressed as problems about particular arithmetic functions and their associated sums (or mean values). Let us introduce some well-known examples of arithmetic functions.

1. The von Mangoldt function Λ: it is defined as

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^k,\ p \text{ a prime},\ k \ge 1, \\ 0. & \text{otherwise.} \end{cases}$$

   This function is closely related to the indicator function of prime numbers. The celebrated Prime Number Theorem (PNT), proved independently by Hadamard and de la Vallée Poussin in 1896, says that

$$\#\{1 \le n \le x : n \text{ a prime}\} \sim \int_2^x \frac{dt}{\log t}$$

   as $x \to \infty$. An integration-by-parts argument [SS03, p. 189] shows that PNT is equivalent to

$$\sum_{n\le x} \Lambda(n) \sim x$$

as $x \to \infty$. The Riemann Hypothesis (henceforth known as RH), which states that the non-trivial zeros of the Riemann zeta function $\zeta(s) = \sum_{n\geq 1} n^{-s}$ lie on $\Re s = 1/2$, is equivalent to the estimate [IK04, Prop. 5.14]

$$\sum_{n\leq x} \Lambda(n) = x + O_\varepsilon(x^{1/2+\varepsilon}).$$

Unconditionally, it is not even known that the above error term is $O(x^{1-\delta})$ for some positive δ.

2. The Möbius function μ: it is defined as

$$\mu(n) = \begin{cases} (-1)^k & \text{if } n = p_1 p_2 \cdots p_k, \ p_i \text{ are distinct primes,} \\ 0 & \text{otherwise.} \end{cases}$$

PNT is equivalent to $\sum_{n\leq x} \mu(n) = o(x)$ [IK04, Ch. 2], while RH is equivalent to $\sum_{n\leq x} \mu(n) = O_\varepsilon(x^{1/2+\varepsilon})$ [IK04, Prop. 5.14]. Again, it is not even known that the error term is $O(x^{1-\delta})$.

3. The divisor function d: it is defined as

$$d(n) = \#\{m \in \mathbb{N} : m \mid n\}.$$

Dirichlet observed that the number of lattice points $(a, b) \in \mathbb{N}^2$ satisfying $ab \leq x$, that is, lying under an hyperbola, is given by $\sum_{n\leq x} d(n)$. He proved that $\sum_{n\leq x} d(n) = x \log x + (2\gamma - 1)x + O(x^{1/2})$ where γ is the Euler-Mascheroni constant. It is conjectured that the error term is actually $O_\varepsilon(x^{1/4+\varepsilon})$ ("Dirichlet's divisor problem"). There have been many works in that direction, starting with Voronoi [Vor03], who proved that the error is $O(x^{1/3} \log x)$. The current record is $O(x^{517/1648+\varepsilon})$, due to Bourgain and Watt [BW17]. That the error term is $\Omega(x^{1/4}(\log x)^{1/4} \log\log x)$ is due to Hardy [Har16]. Dirichlet divisor problem would follow from the "exponent pair hypothesis"; see [Mon94, Ch. 3, Conj. 2] for a statement of this hypothesis.

4. The indicator of squarefrees, μ^2: it is defined as

$$\mu^2(n) = \begin{cases} 1 & \text{if } p^2 \mid n \text{ for some prime } p, \\ 0 & \text{otherwise,} \end{cases},$$

and is indeed given by the square of the Möbius function. It is elementary to show that [MV07, Thm. 2.2]

$$\sum_{n\leq x} \mu^2(n) = \frac{x}{\zeta(2)} + O(x^{1/2}).$$

It is conjectured that the optimal exponent in the error term is $1/4 + \varepsilon$, but there is no unconditional improvement on it. Conditionally on RH this is also open, despite many works, see Axer [Axe11] who proved that RH implies an error of $O_\varepsilon(x^{2/5+\varepsilon})$. The strongest result is due to Liu, who proved an error term of $O_\varepsilon(x^{11/35+\varepsilon})$ under RH [Liu16]. It is known that the exponent cannot be $1/4$ or smaller [MV07, p. 471].

5. Hooley's Δ-function: it is defined by

$$\Delta(n) = \max_{u \geq 0} \#\{m \in (u, eu] : m \mid n\}.$$

It was introduced by Hooley, who used it in the study of "apparently unrelated topics in the fields of Diophantine approximation, Waring's problem, and divisor sums" [Hoo79, p. 115]. The pointwise bound $1 \leq \Delta \leq d$ yields bounds on $\sum_{n \leq x} \Delta(n)$. Despite various works [Erd74, Hoo79, HT82], the correct lower and upper bounds on $\sum_{n \leq x} \Delta(n)$ are not known.

1.3 Statistics of arithmetic functions

As we have seen, the mean value of arithmetic functions, that is, the asymptotics of

$$\frac{1}{x} \sum_{n \leq x} \alpha(n)$$

as $x \to \infty$, can hold important number-theoretic information. Moreover, once an asymptotic expression $M_\alpha(x)$ is found for $\sum_{n \leq x} \alpha(n)/x$, the error term

$$E_\alpha(x) = xM_\alpha(x) - \sum_{n \leq x} \alpha(n)$$

holds information which pertains to deep conjectures such as RH. For functions, such as μ, that exhibit cancellation in the sense that $\sum_{n \leq x} \alpha(n) = o(\sum_{n \leq x} |\alpha(n)|)$, we often take $M_\alpha \equiv 0$.

For applications, one wants to understand the behavior of arithmetic functions on a finer scale, for instance in *short intervals* or other sparse sets. By that we mean understanding the behavior of α when restricted to a set $[x, x+h] \cap \mathbb{Z}$ with $h = o(x)$. For many functions, it is conjectured that

$$\sum_{n \in [x, x+h]} \alpha(n) \sim hM_\alpha(x) \tag{1.2}$$

as long as h grows at least like a small power of x, that is $h \gg x^\varepsilon$ for some $\varepsilon > 0$. For functions with $M_\alpha \equiv 0$ one should replace (1.2) with

$$\sum_{n \in [x, x+h]} \alpha(n) = o(h). \tag{1.3}$$

However, neither (1.2) nor (1.3) is known for any interesting function in the 'full range' $h \gg x^\varepsilon$.

Good bounds on E_α yield a result of the form (1.2) or (1.3) in some range of h. For instance, in the case of $\alpha = \Lambda$, RH gives $E_\Lambda(x) = O_\varepsilon(x^{1/2+\varepsilon})$ with $M_\Lambda \equiv 1$, from which one deduces (1.2) for $h \geq x^{1/2+\delta}$. This still falls short from the expected truth. Unconditionally, the best known range where (1.2) holds for $\alpha = \Lambda$ is due to Heath-Brown [HB88] (building on Huxley [Hux72]), who obtains a result for $h \geq x^{7/12-o(1)}$.

For most functions, it is conjectured that not only (1.2) or (1.3) hold in a very wide range, but also that the error term is small. As an example, Montgomery and Vaughan [MV07, Conj. 13.4] conjecture that for $\alpha = \Lambda$, (1.2) holds with an error term of $O_\varepsilon(x^\varepsilon \sqrt{h})$:

$$\sum_{n \in [x, x+h]} \Lambda(n) = h + O_\varepsilon(x^\varepsilon \sqrt{h}). \tag{1.4}$$

That is, there is essentially square-root cancellation even when $\alpha = \Lambda - 1$ is summed over short intervals. This is obviously stronger than RH (as for $h = x$ it implies RH) and is completely open.

Another example is μ^2, for which one conjectures that (1.2) holds with an error term of $O_\varepsilon(x^\varepsilon h^{1/4})$[1]:

$$\sum_{n\in[x,x+h]} \mu^2(n) = \frac{h}{\zeta(2)} + O_\varepsilon(x^\varepsilon h^{1/4}). \tag{1.5}$$

Again, this is completely open.

1.4 The variance of arithmetic functions

One way to obtain useful results about short intervals is to introduce averaging, as we now explain. Let x be a uniformly drawn number from $[0, X]$. Let H be a function of X. Consider the random variable

$$S_\alpha(X, H) = \sum_{n\in[x,x+H]} \alpha(n),$$

that is, a sum of α over a random interval of length H. Instead of proving bounds for *every* short interval, one can try and study the distribution of $S_\alpha(X, H)$, from which one can obtain information on *most* intervals. Here is an example. Consider

$$V_\alpha(X, H) := \mathbb{E}(S_\alpha(X, H) - HM_\alpha(X))^2 = \frac{1}{X}\int_0^X \left(\sum_{n\in[x,x+H]} \alpha(n) - HM_\alpha(X)\right)^2 dx,$$

which is the second moment of $S_\alpha - HM_\alpha$ (itself an approximation to the variance of S_α). An application of Chebyshev's inequality shows that

$$\mathbb{P}(|S_\alpha(X, H) - HM_\alpha(X)| \geq \varepsilon HM_\alpha(X)) \leq \frac{V_\alpha(X, H)}{\varepsilon^2 H^2 M_\alpha(X)^2}.$$

If $V_\alpha(X, H) = o(H^2 M_\alpha^2(X))$, it follows that for most $x \in [0, X]$ we have $\sum_{n\in[x,x+H]} \alpha(n) = HM_\alpha(X)(1 + o(1))$ as $X \to \infty$, in the sense that the exceptional set is $o(X)$.

The quantity $V_\alpha(X, H)$ and its variants – which we informally call the variance of α in short intervals – are the main topic of this thesis. We survey a few of the results on $V_\alpha(X, H)$.

1. For the von Mangoldt function, Selberg [Sel43] studied the following variant of $V_\Lambda(X, H)$:

$$V'_\Lambda(X, \delta) = \frac{1}{X}\int_X^{2X} \left(\sum_{n\in[x,(1+\delta)x]} \Lambda(n) - \delta x\right)^2 x^{-2}\, dx.$$

   Unconditionally, he proved $V'_\Lambda(X, \delta) = O(\delta^2/\log^4 x)$ for $\delta \in [x^{\varepsilon-c}, 1]$ for some absolute $c \in (0, 1)$, from which he deduced that $\sum_{n\in[x,x+\Phi(x)]} \Lambda(n) \sim \Phi(x)$ for almost all x, where Φ is a positive, increasing function satisfying $\Phi(x) > x^{1-c+\varepsilon}$ and $\Phi(x) = o(x)$. Under RH, he proved

$$V'_\Lambda(X, \delta) = O\left(\frac{\delta \log^2 X}{X}\right)$$

[1] A q-analogue of this is stated in a paper of Croft [Cro75] and is attributed to Montgomery.

for $\delta \in [X^{-1}, X^{-1/4}]$, which allowed him to take any $\Phi(x)$ with $\Phi(x)/\log^2 x \to \infty$ instead of $\Phi(x) > x^{1-c+\varepsilon}$ in the short interval result. Saffari and Vaughan [SV77], building on Selberg, proved that

$$V_\Lambda(X, H) = O(H \log^2(2X/H)) \tag{1.6}$$

on RH. Observe that (1.6) implies (1.4) for almost all $x \in [X, 2X]$. We touch on some of the ideas behind these works in §4.4.2.

2. For μ, Ramachandra [Ram76] proved that

$$V_\mu(X, H) = O_{A,\varepsilon}(H^2(\log X)^{-A} + X^{c+\varepsilon})$$

unconditionally for $c = 1/6$, and under RH for $c = 0$. A relatively recent breakthrough result of Matomäki and Radziwiłł[MRl16] gives

$$V_\mu(X, H) = o(H^2),$$

as $X, H \to \infty$, which is optimal in the sense that it proves cancellation in $\sum_{n\in[x,x+H]} \mu(n)$ for almost all $x \in [X, 2X]$, once H grows to infinity with X.

3. R. R. Hall [Hal82, Thm. 2] proved that

$$V_{\mu^2}(X, H) \sim CH^{1/2}, \qquad C = \frac{\zeta(3/2)}{\pi} \prod_p \left(1 - \frac{3}{p^2} + \frac{2}{p^3}\right), \tag{1.7}$$

as X and H tend to infinity with $H < X^{2/9-\varepsilon}$. This shows that $\sum_{n\in[x,x+H]} \mu^2(n) = 6H/\pi^2 + O_\varepsilon(H^{1/4}X^\varepsilon)$ for almost all $x \in [0, X]$, as long as $H < X^{2/9-\varepsilon}$ and $H \to \infty$, and lends support to the conjectured estimate (1.5). Hall's method is explained in §1.9.1.

4. Ivić [Ivi09], building on work of Jutila [Jut84], proved that

$$V_d(X, H) \sim HP_3\left(\log \frac{\sqrt{X}}{H}\right)$$

uniformly for $X^\varepsilon \le H \le X^{1/2-\varepsilon}$, for some cubic polynomial P_3. For the complementary range $X^{1/2+\varepsilon} \le H \le X^{1-\varepsilon}$, Lester [Les16] proved that $V_d(X, H) \sim D\sqrt{X}$ for an absolute constant D. We give new proofs for these results, in the function setting, in §1.8.

An upper bound for V_α leads to a result about the mean value of α in almost all short intervals (for some range of H and X). We now illustrate why sometimes this is the best we can hope for, which serves as an additional reason for the study of the variance. Maier [Mai85] proved that for any $A > 1$, $\sum_{n\in[x,x+\log^A x]} \Lambda(n)$ is not asymptotic to $\log^A x$ for infinitely many x's. However, at least under RH, Selberg's work shows that $\sum_{n\in[x,x+\log^A x]} \Lambda(n) \sim \log^A x$ for almost all x once $A > 2$. We note that the irregularities in the distribution of primes exhibited by Maier are a part of a much more general phenomena, see the work of Granville and Soundararajan [GS07].

1.5 Interpretations of the variance

One reason for being interested in the asymptotics of the variance V_α, and not only in its bounds, is its relations to probabilistic models for Möbius values, primes, and other objects, and well as its connections with the distribution of zeros of L-functions.

A naive probabilistic model for the values of the Möbius function comes from considering a random arithmetic function R_n, such that $R_n = 0$ with probability 1 if $\mu^2(n) \neq 0$, while for n's with $\mu^2(n) = 1$, R_n are i.i.d random variables taking the values $-1, +1$ with equal probabilities. By studying $\sum_{n\in[x,x+H]} R_n$, Good and Churchhouse [GC68] conjectured that $\sum_{n\in[x,x+H]} \mu(n)/\sqrt{H/\zeta(2)}$, for random $x \in [0, X]$, tends in distribution to standard Gaussian (at least when $H \to \infty$ is not too close to X), and investigated this conjecture numerically. Their conjecture implies in particular that

$$\mathrm{Var}_\mu(X, H) \sim \frac{H}{\zeta(2)}. \tag{1.8}$$

So proving or giving evidence to (1.8) lends support to the randomness of $\mu(n)$. See [Ng08] for conditional evidence towards (1.8).

Cramér's model is the name for a similar probabilistic model for primes, where $\Lambda(n)$ is modeled by $R_n \cdot \log n$, where now $R_n = 1$ with probability $1/\log n$ and 0 otherwise, independently of the other R_m's. This model suggests that

$$\mathrm{Var}_\Lambda(X, H) \sim H \log X$$

as H tends to infinity with $X^\varepsilon < H < X^{1-\varepsilon}$. However, there is substantial evidence that the truth is in fact

$$\mathrm{Var}_\Lambda(X, H) \sim H \log \frac{X}{H}, \tag{1.9}$$

which deviates from $H \log X$ in the range $X^\varepsilon < H < X^{1-\varepsilon}$. However, there is currently no probabilistic model for the primes explaining why the variance decreases in that way!

1.6 Some Random Matrix Theory

The most satisfactory motivation for (1.9) comes from Random Matrix Theory, as we now (briefly) explain, following Goldston's historical account [Gol05]. Let us denote the non-trivial zeros of $\zeta(s)$ by $\rho = 1/2 + i\gamma$. We shall assume RH, so that the γ's are real. Up to height T, we have roughly $(T \log T)/2\pi$ zeros with $\gamma > 0$, and so their mean spacing is $2\pi/\log T$. Montgomery [Mon73] defined

$$F(\alpha) = F(\alpha, T) = \left(\frac{T}{2\pi}\log T\right)^{-1} \sum_{0<\gamma,\gamma'\leq T} T^{i\alpha(\gamma-\gamma')}\omega(\gamma - \gamma'),$$

where $\omega(u) = 4/(4 + u^2)$. Here the sum is over an ordered pair of non-trivial zeros of ζ in the upper half-plane. Montgomery proved (still under RH) that F is real, even and non-negative. For $0 \leq \alpha \leq 1 - \varepsilon$ he obtained

$$F(\alpha) = \alpha + o(1) + (1 + o(1))T^{-2\alpha} \log T;$$

this was extended by Goldston to $0 \leq \alpha \leq 1$ [Gol81]. Conditionally on a certain uniform version of the Hardy-Littlewood 2-Tuple Conjecture, Montgomery showed that $F(\alpha) =$

$1 + o(1)$ for $1 \le \alpha \le 2 - \varepsilon$. He speculated that $F(\alpha) = 1 + o(1)$ for $\alpha \in [1, M]$ for any fixed M - which is known as the Strong Pair Correlation Conjecture (SPC). To see its usefulness, observe that by integrating $F(\alpha)$ against a nice test function r, we have

$$\left(\frac{T}{2\pi}\log T\right)^{-1} \sum_{0<\gamma,\gamma'\le T} r\left((\gamma-\gamma')\frac{\log T}{2\pi}\right)\omega(\gamma-\gamma') = \int_{\mathbb{R}} \widehat{r}(\alpha)F(\alpha)\,d\alpha.$$

The right-hand side may be computed under SPC, at least for functions with $\widehat{r}$ having compact support. With some work, a particular choice of r leads to

$$\left(\frac{T}{2\pi}\log T\right)^{-1} \sum_{\substack{0<\gamma,\gamma'\le T \\ 0<\gamma'-\gamma\le \frac{2\pi\beta}{\log T}}} 1 \sim \int_0^\beta \left(1 - \frac{\sin^2(\pi u)}{(\pi u)^2}\right)\,du$$

for any fixed $\beta > 0$; this asymptotic is known as the Pair Correlation Conjecture (PC). The left-hand side counts pair of zeros that are close to each other, hence the name. As observed by Dyson, the right-hand side also arises as the pair correlation of eigenvalues of random unitary matrices, that is,

$$\frac{1}{N}\int_{U(N)} \left(\sum_{0<\theta_j-\theta_k<\frac{2\pi\beta}{N}} 1\right)\,dX \sim \int_0^\beta \left(1 - \frac{\sin^2(\pi u)}{(\pi u)^2}\right)\,du.$$

Here $U(N)$ is the N by N unitary group, endowed with Haar measure, and $\theta_i \in [0, 2\pi]$ are the angles of the eigenvalues of $X \in U(N)$ in some order. Quoting Conrey, "this important fact was fortuitously discovered at tea at the Institute for Advanced Study one afternoon in 1971 when Chowla introduced Hugh Montgomery and Freeman Dyson to each other". It took 20 years until the work of Montgomery on pair correlation was extended to n-level correlations by Rudnick and Sarnak [RS96]. See also the work of Keating and Snaith [KS00], predicting moments of L-functions through Random Matrix Theory.

The connection with the variance of primes is a beautiful theorem of Goldston and Montgomery [GM87], saying that SPC is *equivalent* to (1.9). So the surprising asymptotics for the variance of primes reflects deep facts about the distribution of the zeros of ζ.

Although SPC and PC are open, in the function field setting there are analogous of them that are known unconditionally, see for instance the book of Katz and Sarnak [KS99].

1.7 The function field setting

In this thesis, we study the variance of certain functions in the setting of function fields. Some of our results go beyond what is known in the integer setting, while others give predictions for the behavior of certain variances in the number field setting, which were not understood before. Below we define the basic objects we shall need.

We let q be a prime power and let $\mathbb{F}_q[T]$ be the polynomial ring over the finite field $\mathbb{F}_q$ with q elements. Let $\mathcal{M}_{n,q}$ denote the set of monic polynomials of degree n in $\mathbb{F}_q[T]$, and let $\mathcal{M}_q = \cup_{n\ge 0}\mathcal{M}_{n,q}$ denote the set of all monic polynomials in $\mathbb{F}_q[T]$. By a well-known analogy $\mathcal{M}_q$ serves as a substitute for the set of positive integers.

Let h, n be integers such that $-1 \leq h \leq n-1$. Given $f_0 \in \mathcal{M}_{n,q}$, a short interval around f_0 of size q^{h+1} is the subset

$$I(f_0, h) = \{f_0 + g : \deg(g) \leq h\} \subseteq \mathcal{M}_{n,q}.$$

The degree of the zero polynomial is defined to be $-\infty$.

Let $\mathcal{P}_{n,q} \subseteq \mathcal{M}_{n,q}$ be the subset of monic irreducible polynomials of degree n and $\mathcal{P}_q = \cup_{n \geq 0} \mathcal{P}_{n,q}$ be the set of all monic irreducible polynomials. The norm $\|f\|$ of $f \in \mathbb{F}_q[T]$ is $\#\mathbb{F}_q[T]/(f) = q^{\deg(f)}$ if $f \neq 0$, and $\|0\| = 0$. The zeta function of ζ_q is defined as

$$\zeta_q(s) = \prod_{P \in \mathcal{P}_q} (1 - \|P\|^{-s})^{-1} = \sum_{f \in \mathcal{M}_q} \|f\|^{-s} = \frac{1}{1 - q^{1-s}},$$

where both the product and the sum converge for $\Re s > 1$.

An arithmetic function in this setting is a function $\alpha \colon \mathcal{M}_q \to \mathbb{C}$. Its mean value over $\mathcal{M}_{n,q}$ is defined as

$$\langle \alpha \rangle_{\mathcal{M}_{n,q}} = \frac{\sum_{f \in \mathcal{M}_{n,q}} \alpha(f)}{\#\mathcal{M}_{n,q}}.$$

Given integers $-1 \leq h \leq n-1$, the variance of sums of α over intervals of length q^{h+1} around a polynomial of degree n is defined as

$$V_\alpha(n, h) = \frac{1}{q^n} \sum_{f_0 \in \mathcal{M}_{n,q}} \left| \sum_{f \in I(f_0, h)} \alpha(f) - q^{h+1} \langle \alpha \rangle_{\mathcal{M}_{n,q}} \right|^2.$$

Studying this quantity while taking n and h to infinity is similar to taking X and H to infinity in the integer setting. However, in this setting we have a new feature: the variance also depends on the parameter q, the size of the underlying finite field. In recent years, there has been a lot of fruitful work on understanding the variance of arithmetic functions, and many other important quantities, in the large-q limit, where one fixes all parameters and takes q to infinity. Once an asymptotic is found for $V_\alpha(n, q)$ in the large-q limit, one often recovers a combinatorial quantity that can be studied as n, h go to infinity and still retains important information which can be extracted.

We can often say something about the problem when q grows thanks to algebraic geometry being able to deal well with varying q. In particular, in $\mathbb{F}_q[T]$ we have versions of the Pair Correlation Conjecture in the large-q limit, but not so in the large-n limit.

As an example, Keating and Rudnick [KR14, KR16], using deep results of Katz on zeros of Dirichlet L-functions in the function field setting [Kat13, Kat15], established analogues of (1.8) and (1.9). To state their result, we define the polynomial von Mangoldt function $\Lambda_q \colon \mathcal{M}_q \to \mathbb{C}$ as

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

and the polynomial Möbius function $\mu_q \colon \mathcal{M}_q \to \mathbb{C}$ as

$$\mu_q(f) = \begin{cases} (-1)^k & \text{if } f = P_1 P_2 \cdots P_k \text{ for distinct } P_i \in \mathcal{P}_q, \\ 0 & \text{otherwise.} \end{cases}$$

It is well-known that

$$\langle \Lambda_q \rangle_{\mathcal{M}_{n,q}} = 1, \qquad \langle \mu_q \rangle_{\mathcal{M}_{n,q}} = 0 \tag{1.10}$$

for $n \geq 2$ [Ros02, Ch. 2]. Their results are

$$\lim_{q\to\infty} \frac{\mathrm{Var}_{\Lambda_q}(n,h)}{q^{h+1}} = n-h-2, \qquad \lim_{q\to\infty} \frac{\mathrm{Var}_{\mu_q}(n,h)}{q^{h+1}} = 1,$$

giving strong evidence for the corresponding statements over the integers.

An example of a different flavor relates to higher divisor functions,

$$d_k(n) := \#\{(n_1, \cdots, n_k) \in \mathbb{N}^k : n = n_1 n_2 n_3 \dots n_k\}, \qquad k \geq 2.$$

The work of Ivić and Lester helps us understand $\mathrm{Var}_{d_2}(X,H)$. Lester also obtains asymptotics for $\mathrm{Var}_{d_k}(X,H)$ in the range $H > X^{1-1/(k-1)}$ under the Lindelöf Hypothesis, but the techniques cannot handle smaller H. Keating, Rodgers, Roditty-Gershon and Rudnick [KRRGR18] studied a function-field analogue of this problem, first computing $\mathrm{Var}_{d_k}(n,h)/q^{h+1}$ as q tends to ∞, and then analyzing the remaining quantity as h/n tends to $\delta \in (0,1)$. Based upon their rigorous results, they gave a surprising prediction for $\mathrm{Var}_{d_k}(X,H)$ in the integers, involving phase changes. Precisely, they expect $\mathrm{Var}_{d_k}(X,H) \sim a_k H P_{k^2-1}(\log H/\log X)(\log X)^{k^2-1}$ for a particular constant a_k and for a piecewise polynomial P_{k^2-1} of degree k^2-1, behaving differently on each interval $[1-1/i, 1-1/(i+1)]$ $(1 \leq i \leq k-1)$ and on $[1-1/k, 1]$. This conjecture predicts a transition near $H = X^{1-1/i}$ for $2 \leq i \leq k$. This work has already spawned several subsequent works over the integers, which give results (both conditional on GRH and unconditional) in agreement with the curious prediction [HS17, RS18, dlBF20, BC20, Mas20].

In the function field setting, the connection between basic arithmetic questions and the distribution of zeros of L-functions is much more clearer. As we shall show, this is even true for problems such at twin primes in function fields, where we uncover such a connection for the first time.

In this thesis, in addition to large-q results (which usually follow from information about zeros of L-functions), we also prove some large-n results, where other benefits of function fields come into play.

1.8 Approaches to the computation of the variance

One can study $V_\alpha(n,h)$ in physical space and in Fourier space. We explain this through a worked out example. Recall that Ivić and Lester [Ivi09, Les16] showed

$$V_d(X,H) \sim \begin{cases} HP_3(\log \frac{\sqrt{X}}{H}) & \text{if } X^\varepsilon \leq H \leq X^{1/2-\varepsilon}, \\ D\sqrt{X} & \text{if } X^{1/2+\varepsilon} \leq H \leq X^{1-\varepsilon}. \end{cases}$$

The polynomial divisor function is

$$d_q(f) = \#\{m \in \mathcal{M}_q : m \mid f\}.$$

Over $\mathbb{F}_q[T]$, we can obtain the following closed-form expression for $V_{d_q}(n,h)$.

Theorem 1.1. *We have*

$$V_{d_q}(n,h) = \begin{cases} q^{h+1}\binom{n-2h-1}{3} & \text{if } -1 \leq h \leq \lfloor n/2 \rfloor - 1, \\ 0 & \text{if } \lfloor n/2 \rfloor - 1 \leq h \leq n-1. \end{cases}$$

We are going to present two completely different proofs for Theorem 1.1. The first approach is through *correlation sums*. Expanding $V_{d_q}(n,h)$, we find that

$$\frac{\mathrm{Var}_{d_q}(n,h)}{q^{h+1}} = \sum_{\Delta\in\mathbb{F}_q[T],\deg(\Delta)\leq h} \left(\langle d_q(f)d_q(f+\Delta)\rangle_{f\in\mathcal{M}_{n,q}} - (\langle d_q\rangle_{\mathcal{M}_{n,q}})^2\right), \tag{1.11}$$

see Lemma 3.4 for a general statement and a proof. We shall prove

$$\frac{1}{\#\mathcal{M}_{n,q}}\sum_{f\in\mathcal{M}_{n,q}} d_q(f)d_q(f+\Delta) = (n+1)^2 + \sum_{i=1}^{\lfloor\frac{n}{2}\rfloor}\frac{(n-2i+1)^2}{q^i}(a_{i,\Delta}-a_{i-1,\Delta}), \tag{1.12}$$

where $a_{i,\Delta}$ denotes the number of monic divisors of Δ of degree i, see Lemma 3.3 for a proof. For $\Delta = 1$, (1.12) is due to Andrade, Bary-Soroker and Rudnick [ABSR15, Thm. 7.1], and the statement for general Δ follows by generalizing their arguments. See the introduction of [ABSR15] for background on the integer sum $\sum_{n\leq X} d(n)d(n+\Delta)$, where weaker results are known.

Plugging (1.12) in (1.11), an elementary computation yields our theorem. Full details are given in §3.2.

Our second proof involves the functional equation for Dirichlet L-functions. In §2 We introduce a group $G(R_\ell)$ of 'short interval characters', which are able to detect whether two polynomial of degree n are in the same interval of size $q^{n-\ell}$. A Plancherel-type theorem yields

$$\mathrm{Var}_{d_q}(n,h) = \frac{\sum_{\chi_0\neq\chi\in G(R_{n-h-1})}|\sum_{f\in\mathcal{M}_{n,q}} d_q(f)\chi(f)|^2}{q^{2(n-h-1)}}. \tag{1.13}$$

Since we express the problem as an 'integral' over characters, we consider this a Fourier space approach. The sum $\sum_{f\in\mathcal{M}_{n,q}} d_q(f)\chi(f)$ is a function of $L(u,\chi)$. Applying the functional equation of $L(u,\chi)$ to the right-hand side of (1.13) and expanding the average over $G(R_{n-h-1})$, we reduce to a new problem in physical space, which is easier to deal with than the original one. Informally, the functional equation allows us to replace sums over high degree polynomials with sums over lower degree polynomials, which are easier to study. Full details are given in §3.3.

The approaches of Ivić and Lester over the integers are in the spirit of our second approach. However, they use 'ready-made' transformation formulas for the divisor functions. Our approach is more flexible, and led us eventually to make new progress on $\mathrm{Var}_{\mu_q^2}(n,h)$ and $\mathrm{Var}_{\mu^2}(X,H)$.

We note that the correlation sums approach was not used before in the study of the variance of divisor functions.

1.9 Results

We now motivate and state the four main results of the thesis. They are results in the function field setting. Our first two results are in the large-n limit, and the later two are in the large-q limit.

Since the proofs underlying the large-n results use only analytic tools, the results and proofs should have analogues over the integers (possibly conditional).

The large-q results, however, use techniques based on algebraic geometry which are unique to the function field setting, and so cannot be transported to $\mathbb{Z}$ without new ideas. However, they help predict new number-theoretic conjectures as well as give support to existing conjectures.

1.9.1 The variance of squarefree polynomials in short intervals

Before discussing our results, we review Hall's method in proving (1.7). Expanding the definition of $V_{\mu^2}(X,H)$, one can express the variance as a linear combination of the correlation sums

$$\sum_{n\leq X}\mu^2(n)\mu^2(n+c)$$

for c-s between 0 and H. Hall estimated such sums, obtaining [Hal82, Thm. 1]

$$\sum_{n\leq X}\mu^2(n)\mu^2(n+c)=\prod_{p\nmid c}\left(1-\frac{2}{p}\right)\prod_{p\mid c}\left(1-\frac{1}{p}\right)X+O\left(X^{\frac{2}{3}}(\log X)^{\frac{2}{3}}\right) \tag{1.14}$$

uniformly in $0<c<X$. The exponent of X in the error term in (1.14) leads to the range $H=O_\varepsilon(X^{2/9-\varepsilon})$. Inspecting Hall's argument, if the error term in (1.14) is $O(X^A)$ then one has $\mathrm{Var}_{\mu^2}(X,H)\sim CH^{1/2}$ for $H=O_\varepsilon(X^{2(1-A)/3-\varepsilon})$. Since the error term in (1.14) is expected to be $O_\varepsilon(X^{1/4+\varepsilon})$, the limit of Hall's argument is $H=O_\varepsilon(X^{1/2-\varepsilon})$. Although we are very much far from establishing (1.14) with an optimal error term, we shall extend Hall's result (in the function field setting) to the range $H=O_\varepsilon(X^{1/2-\varepsilon})$, unconditionally.

We now introduce $\mu_q^2\colon\mathcal{M}_q\to\{0,1\}$, the indicator function of squarefree polynomials. It is known that the mean value of μ_q^2 is *exactly* $1/\zeta_q(2)=1-1/q$ [Ros02, Ch. 2], that is,

$$\langle\mu_q^2\rangle_{\mathcal{M}_{n,q}}=\frac{1}{\zeta_q(2)}$$

for all $n\geq 2$. We are interested in the asymptotics of the variance $V_{\mu_q^2}(n,h)$. Keating and Rudnick [KR16, Thm. 1.4], using an equidistribution result of Katz [Kat15, Thm. 1.2], have shown that as q tends to infinity we have

$$V_{\mu_q^2}(n,h)\sim q^{\lfloor\frac{h}{2}\rfloor}$$

for any fixed $0\leq h\leq n-6$, as long as the characteristic of $\mathbb{F}_q$ is not 2 or 3. Their large-q result led them to conjecture that Hall's result (1.7) – both in $\mathbb{Z}$ and $\mathbb{F}_q[T]$ – holds for the complete range, that is, $X^\varepsilon\leq H\leq X^{1-\varepsilon}$ and $\varepsilon\leq h/n\leq 1-\varepsilon$, respectively. Let us see what this means in $\mathbb{F}_q[T]$. By adapting Hall's method of proof, Keating and Rudnick proved the following result in the large-n limit:

$$V_{\mu_q^2}(n,h)\sim q^{\lfloor\frac{h}{2}\rfloor}\lambda_{q,h}C_q \tag{1.15}$$

uniformly in h which satisfies $\varepsilon\leq h/n\leq 2/9-\varepsilon$ for some $\varepsilon>0$, where

$$C_q=\prod_{P\in\mathcal{P}_q}\left(1-\frac{3}{\|P\|^2}+\frac{2}{\|P\|^3}\right),\quad \lambda_{q,h}=\frac{1+q^{-1-1_{2|h}}}{1-q^{-3}}.$$

The constants C_q and $\lambda_{q,h}$ do not appear in the large-q result since C_q, $\lambda_{q,h}$ are asymptotic to 1 in the large-q limit. While C_q is analogous to C appearing in Hall's result, $\lambda_{q,h}$ is unique to the function field setting.

Keating and Rudnick conjectured that (1.15) holds for $\varepsilon\leq h/n\leq 1-\varepsilon$. We make progress towards their conjecture, replacing the exponent 2/9 with 1/2.

Theorem 1.2. *Fix a prime power q. For any $\varepsilon \in (0, 1/4)$, the asymptotic estimate* (1.15) *holds uniformly for h, n tending to infinity with $\varepsilon \le h/n < 1/2 - \varepsilon$.*

Additionally, in the full range we have an upper bound of essentially the predicted magnitude: $V_{\mu_q^2}(n, h) = O_{\varepsilon,q}(q^{\lfloor h/2 \rfloor} q^{n\varepsilon})$ as $n \to \infty$ and $-1 \le h \le n - 1$.

The starting point for our proof is the classical identity

$$\mu_q^2(f) = \sum_{d^2 | f} \mu_q(d) \tag{1.16}$$

which follows from an inclusion-exclusion argument. Alternatively, since both sides of (1.16) are multiplicative, it suffices to verify the identity on prime powers. We introduce a parameter $m \in \mathbb{N}$, and write μ_q^2 as a sum of two functions:

$$\mu_q^2(f) = \alpha_m(f) + \beta_m(f),$$

where

$$\alpha_m(f) = \sum_{d^2 | f,\, \deg(d) \le m} \mu_q(d), \qquad \beta_m(f) = \sum_{d^2 | f,\, \deg(d) > m} \mu_q(d).$$

We shall show that β_m contributes very little to the variance of μ_q^2, at least for large m. By passing to Fourier space (that is, an average over characters) and using estimates on characters sums and Möbius sums, we obtain

Proposition 1.3. *Fix a prime power q. For $1 \le h \le n-1$ and $-1 \le m \le n/2$, we have*

$$\mathrm{Var}_{\beta_m}(n, h) \le q^{\min\{\frac{h}{2}, h-m\} + o(n)}$$

as $n \to \infty$. (The $o(n)$ term may depend on q.)

Taking $m = -1$, β_m coincides with μ_q^2 and so we obtain the second part of Theorem 1.2.

To evaluate $\mathrm{Var}_{\alpha_m}(n, h)$ we again pass to Fourier space, and apply the functional equation there. We interpret the obtained expression back in physical space. It turns out that for small enough m, the problem we obtain in physical space is easy enough in the sense that only certain diagonal terms contribute to it. We obtain

Proposition 1.4. *Fix a prime power q. For $-1 \le h \le n-1$ and $-1 \le m \le n/4$, we have*

$$\mathrm{Var}_{\alpha_m}(n, h) = F(h, m)$$

for a function F depending only on h and on m (as well as q).

We have an explicit formula for $F(h, m)$, see (4.10). However, as we now show, we can easily conclude Theorem 1.2 from Propositions 1.3 and 1.4, without needing to work out the behavior of $F(h, m)$ directly. Indeed, fix small $\varepsilon > 0$ and suppose that $h/n \in (\varepsilon, 1/2 - \varepsilon)$. From Cauchy-Schwarz and the two propositions,

$$\mathrm{Var}_{\mu_q^2}(n, h) = F(h, m) + O_\varepsilon\left(q^{\frac{h}{2}(1-\frac{\varepsilon}{2})} + \sqrt{F(h, m) q^{\frac{h}{2}(1-\frac{\varepsilon}{2})}} \right) \tag{1.17}$$

if $m = h(1+\varepsilon)/2$. To conclude, it suffices to show that $F(h, h(1+\varepsilon)/2) \sim q^{\lfloor \frac{h}{2} \rfloor} \lambda_{q,h} C_q$. To do so, we compute $\mathrm{Var}_{\mu_q^2}(\lfloor h/(2\varepsilon) \rfloor, h)$ in two different ways. By (1.15),

$$\mathrm{Var}_{\mu_q^2}(\lfloor h/(2\varepsilon) \rfloor, h) \sim q^{\lfloor \frac{h}{2} \rfloor} \lambda_{q,h} C_q. \tag{1.18}$$

By (1.17),

$$\mathrm{Var}_{\mu_q^2}(\lfloor h/(2\varepsilon)\rfloor, h) = F(h,m) + O_\varepsilon\left(q^{\frac{h}{2}(1-\frac{\varepsilon}{2})} + \sqrt{F(h,m)q^{\frac{h}{2}(1-\frac{\varepsilon}{2})}}\right). \tag{1.19}$$

Equating (1.18) and (1.19) we find that $F(h, h(1+\varepsilon)/2) \sim q^{\lfloor h/2\rfloor}\lambda_{q,h}C_q$, as needed. What was crucial in this bootstrapping-type argument is the independence of $F(h,m)$ from the parameter n, allowing as to increase n however we desire.

The proofs of Propositions 1.3 and 1.4 are detailed in §4. Theorem 1.2 has led to a joint paper titled "On the variance of squarefree integers in short intervals and arithmetic progressions" [GMRR20] (to appear in GAFA), authored by myself, Kaisa Matomäki, Maksym Radziwiłł and Brad Rodgers, where we transport the method of proof to the integer setting. In §4.4 we explain some of the ideas required in proving an unconditional version of Theorem 1.2 over the integers.

We expect the methods of proof to apply to many more multiplicative functions, especially those that, in a sense, are close to the constant function **1** (e.g. are a convolution of **1** with a function of zero mean value), and plan to explore this in the future.

1.9.2 The variance of factorization functions in short intervals

To any $f \in \mathcal{M}_q$ with prime factorization $f = \prod_{i=1}^k P_i^{e_i}$ ($P_i \in \mathcal{P}_q$ distinct, $e_i \geq 1$), we can associate the following multiset, named the *extended factorization type of* f:

$$\omega_f := \{(\deg(P_i), e_i) : 1 \leq i \leq k\}.$$

(We often omit the word 'extended'.) Following Rodgers [Rod18], an arithmetic function $\alpha\colon \mathcal{M}_q \to \mathbb{C}$ is called a *factorization function* if $\alpha(f)$ depends only on ω_f. Some of the most commonly studied arithmetic functions in number theory, when considered in the function field setting, are instances of factorization functions: the von Mangoldt function Λ_q, the Möbius function μ_q, the divisor function d_q, the indicator of squarefrees μ_q^2, and many more. We prove

Theorem 1.5. *Let $\alpha\colon \mathcal{M}_q \to \mathbb{C}$ be a factorization function. Let $0 \leq h \leq n-1$ and $f_0 \in \mathcal{M}_{n,q}$. Then*

$$\left|\sum_{f\in I(f_0,h)} \alpha(f) - q^{h+1}\langle\alpha\rangle_{\mathcal{M}_{n,q}}\right| \leq \max_{f\in\mathcal{M}_{n,q}} |\alpha(f)|\, q^{\frac{n}{2}} e^{O_q(\frac{n\log\log(n+2)}{\log(n+2)})} \tag{1.20}$$

and

$$\mathrm{Var}_\alpha(n,h) \leq \max_{f\in\mathcal{M}_{n,q}} |\alpha(f)|^2\, q^{h+1} e^{O_q(\frac{n\log\log(n+2)}{\log(n+2)})}. \tag{1.21}$$

As long as $\max_{f\in\mathcal{M}_{n,q}}|\alpha(f)|$ grows subexponentially in n (as is the case for most functions), (1.20) is a non-trivial result in the range $\limsup_{n\to\infty} h/n > 1/2$.

For the variance $\mathrm{Var}_\alpha(n,h)$, we beat the trivial upper bound $q^{2(h+1)}\max_{f\in\mathcal{M}_{n,q}}|\alpha(f)|^2$ as long as $\limsup_{n\to\infty} h/n > 0$, which corresponds to $H \gg X^\varepsilon$ in the number field setting.

We now give a concrete application for Theorem 1.5. As far as the author is aware, Hooley's Δ-function was not studied in short intervals. The function field analogue of Δ is

$$\Delta_q\colon \mathcal{M}_q \to \mathbb{C}, \qquad \Delta_q(f) := \max_{0\leq i\leq \deg(f)} \sum_{\substack{d\mid f,\\ d\in\mathcal{M}_{i,q}}} 1.$$

As $\langle d_q\rangle_{\mathcal{M}_{n,q}} = n+1$ [Ros02, Prop. 2.5], it follows that $\max_{f\in\mathcal{M}_{n,q}} \Delta_q(f) \le \max_{f\in\mathcal{M}_{n,q}} d_q(f)$, which is known to grow slower than any power of q^n as n tends to infinity. Also, $\langle \Delta_q\rangle_{\mathcal{M}_{n,q}} \ge \langle d_q\rangle_{\mathcal{M}_{n,q}}/(n+1) = 1$. Applying Theorem 1.5 with $\alpha = \Delta_q$, we obtain the following

Corollary 1.6. *As $n\to\infty$, we have*

$$\sum_{f\in I(f_0,h)} \Delta_q(f) \sim q^{h+1}\langle \Delta_q\rangle_{\mathcal{M}_{n,q}}$$

as long as $\limsup_{n\to\infty} h/n > 1/2$, uniformly for $f_0\in\mathcal{M}_{n,q}$.

Over the integers, and conditionally on RH, one can prove results similar to Theorem 1.5 for certain arithmetic functions – see Ramachandra [Ram76] for a method that works both for μ and Λ. However, there is no general result similar to Theorem 1.5 in $\mathbb{Z}$, and in particular the work of Ramachandra requires the Dirichlet series of $\alpha\colon \mathbb{N}\to\mathbb{C}$ to have a very particular form in order to work.

Theorem 1.5 is a large-n result. It complements a beautiful theorem of Rodgers [Rod18] in the large-q limit, from which one obtains as a corollary that

$$\mathrm{Var}_\alpha(n,h) \le q^{h+1}(\langle |\alpha|^2\rangle_{\mathcal{M}_{n,q}} + o_{n,\alpha}(q^{-1/2})). \tag{1.22}$$

The quantity $o_{n,\alpha}(q^{-1/2})$ goes to zero with q, but the implied constant depends both on n and on $\max_{f\in\mathcal{M}_{n,q}}|\alpha(f)|$, and one cannot infer anything in the large-n from (1.22). It would be interesting if one could improve the ℓ_∞-dependence on α in (1.21) to an ℓ_2-dependence, as in (1.22).

The proof of Theorem 1.5 is detailed in §5. The material of §5 has appeared in the paper "Mean values of arithmetic functions in short intervals and in arithmetic progressions in the large-degree limit" published in Mathematika [Gor20]. In that paper we also prove corresponding results over arithmetic progressions, in addition to short intervals.

1.9.3 The variance of sums of two squares in short intervals

Consider the set $S = \{n^2+m^2 : n,m\in\mathbb{Z}\}$ of integers representable as sums of two squares, and let b be its indicator function. Landau [Lan08] proved that

$$\sum_{n\le x} b(n) = K\frac{x}{\sqrt{\log x}} + O\Big(\frac{x}{(\log x)^{3/2}}\Big), \tag{1.23}$$

where

$$K = \frac{1}{\sqrt{2}}\prod_{p\equiv 3 \bmod 4}(1-p^{-2})^{-1/2} \approx 0.764 \tag{1.24}$$

is the Landau-Ramanujan constant. Thus, roughly stated, the likelihood that a random integer near X will be the sum of two squares is around $K/\sqrt{\log X}$.

A naive probabilistic model would predict that $\mathrm{Var}_b(X,H) \sim KH/\sqrt{\log X}$. We shall give evidence, in terms of a function field theorem, that this natural prediction fails, and instead produce a prediction that matches very well numerically. As in the case of primes, we predict that the variance is asymptotically smaller than one would expect once H grows like a power of X.

In the study of the function field analogue of b, we shall let q be an odd prime power. In [BSSW16], Bary-Soroker, Smilansky and Wolf studied an analogue of Landau's problem in

$\mathbb{F}_q[T]$ by introducing the following set and indicator function, which by abuse of notation we will also denote by S and b:

$$S = \{A^2 + TB^2 : \ A, B \in \mathcal{M}_q\},$$

$$b\colon \mathcal{M}_q \to \mathbb{C},\ b_q(f) = \mathbf{1}_{f\in S}.$$

The analogy with integers can be seen in the following way: a positive integer lies in S if an only if it is the norm of some element of $\mathbb{Z}[i]$, and an element of $\mathcal{M}_q$ lies in S if and only if it is the norm of some element of $\mathbb{F}_q[\sqrt{-T}]$. In $\mathbb{F}_q[T]$, the mean value of b can be estimated as follows [Gor17, Thm. 1.1]:

$$\langle b\rangle_{\mathcal{M}_{n,q}} = K_q \binom{n-\frac{1}{2}}{n}\left(1 + O\left(\frac{1}{qn}\right)\right), \tag{1.25}$$

where the implied constant is absolute, and the constant K_q is positive and is an analytic function of q^{-1}. The constant K_q is given by

$$K_q = (1-q^{-1})^{-1/2} \prod_{\chi_2(P)=-1} (1-q^{-2\deg P})^{-1/2} = 1 + O\left(\frac{1}{q}\right),$$

where χ_2 is the unique non-trivial quadratic Dirichlet character modulo T. By Stirling's formula, $\binom{n-1/2}{n} = 1/\sqrt{\pi n} + O(1/n^{3/2})$, and so (1.25) has a resemblance to (1.23).

We evaluate $\mathrm{Var}_b(n,h)$ in the large-q limit. The evaluation involves the z-measure on partitions introduced in [KOV93], with $z = 1/2$. The z-measures arise in an evaluation of certain integrals over the unitary group (Theorem 6.6).

We give a brief introduction to these measures in §6.6, but for the moment we discuss only the notation; recall that we write $\lambda \vdash n$ to indicate that λ is a partition of n and λ_1 to indicate the largest part of a partition λ. For parameters $z \in \mathbb{C}$ and $n \in \mathbb{N}$, the z-measure is a probability measure $M_z^{(n)}(\lambda)$ on the set of partitions $\lambda \vdash n$. In fact these z-measures are a generalization of the well-known Plancherel measure on partitions. The notation $\mathbb{P}_z^{(n)}(\lambda_1 \le N)$ denotes the obvious thing, namely

$$\mathbb{P}_z^{(n)}(\lambda_1 \le N) := \sum_{\substack{\lambda\vdash n\\ \lambda_1\le N}} M_z^{(n)}(\lambda).$$

The actual definition of these probability measures $M_z^{(n)}$ will be given in §6.6. (By convention we set $M_z^{(0)}(\lambda_1 \le N) = 1$ for any N.) We show

Theorem 1.7. *For a fixed odd prime p, and fixed $n \ge 6$, take $0 \le h \le n-7$ and let $N := n-h-1$ and $q = p^k$. Define*

$$T(n;N) := \sum_{j=0}^{n} \frac{(1/4)_j (1/4)_{n-j}}{j!(n-j)!} \mathbb{P}_{1/2}^{(j)}(\lambda_1 \le N-1)\mathbb{P}_{1/2}^{(n-j)}(\lambda_1 \le N).$$

For $N(N-1) \ge n$,

$$\mathrm{Var}_b(n,h) = q^{h+1}T(n;N) + o_{n,p}(q^{h+1}),$$

as $q \to \infty$ (that is $k \to \infty$).

Here $(x)_j := x(x+1)\cdots(x+j-1)$ is the Pochhammer symbol.

We use this theorem to inform an analogous conjecture in the setting of the integers. We require for this purpose an understanding of the limiting behavior of $T(n;N)$ as $h,n\to\infty$ with $h/n\to\delta\in(0,1)$. Note that if h and n are both sufficiently large and $h\sim\delta n$, then $N(N-1)\geq n$ and $N\geq 6$ will both be satisfied.

Proposition 1.8. *For $n,N\to\infty$ with $N/n\to s\in[0,1]$, we have*

$$T(n;N)=\frac{1}{\sqrt{\pi n}}G(s)+o\Big(\frac{1}{\sqrt{n}}\Big),$$

where for real s we define

$$G(s):=\mathbb{P}\Big(1-\frac{s}{\alpha_1}\leq Y\leq\frac{s}{\alpha_1'}\Big),\tag{1.26}$$

for Y, α_1, α_1' independent random variables, with Y distributed as Beta$(1/4,1/4)$ *and α_1,α_1' identically distributed copies of the largest part of the Thoma simplex distributed according to the spectral z-measure with parameters $1/2,1/2$. (The spectral z-measure is defined in §6.6.)*

Note that the random variable $Y\sim\text{Beta}(1/4,1/4)$ is defined by $\mathbb{P}(a\leq Y\leq b):=\sqrt{\pi}\Gamma(1/4)^{-2}\int_a^b t^{-3/4}(1-t)^{-3/4}\,dt$ for $a,b\in[0,1]$, with $Y\in[0,1]$ almost surely. Interestingly, an explicit computations shows that $G''(s)$ has a jump discontinuity as $s=1/2$.

The random variables α_1 and α_1' also lie in $[0,1]$ almost surely, but an explicit characterization of their distribution takes more space to give. Historically they arose in formulas for the characters of certain important representations in the infinite symmetric group (see [KOV93]), but more concretely they are the limiting distribution of the random variable λ_1/n for $\lambda\vdash n$ drawn according to the z-measure of Theorem 1.7. That such a limiting distribution even exists is not obvious, but was shown in [Ols98]. We discuss z-measures on the Thoma simplex in more detail in §6.6.

Plainly for all $s\in[0,1]$ we have $0\leq G(s)\leq 1$. It also is easy to see (i) that $G(s)$ is non-decreasing (from the definition) and (ii) that $G(1)=1$ (from the fact that $Y,\alpha_1,\alpha_1'\in[0,1]$ almost surely). Very recent work on z-measures of Korotkikh [Kor18] and Olshanski [Ols18] tells us that $\mathbb{P}(\alpha_1\leq\varepsilon)$ is non-zero for any small ε, which forces G to be strictly positive.

Using Theorem 1.7 and Proposition 1.8 together, we can write somewhat more succinctly,

Corollary 1.9. *For a fixed odd prime p let $q=p^k$. If $h,n\to\infty$ in such a way that $h/n\to\delta\in(0,1)$, then*

$$\lim_{q\to\infty}\frac{\text{Var}_b(n,h)}{q^{h+1}}=\frac{G(1-\delta)+o(1)}{\sqrt{\pi n}},$$

where the function $G(s)$ is defined in Proposition 1.8.

Corollary 1.9 suggests a conjecture for the integers regarding the number of elements of S that lie in a short interval. Naively one might think it will suggest a conjecture regarding the quantity

$$\frac{1}{X}\int_X^{2X}\Big(\sum_{x\leq n\leq x+H}b(n)-M_{X,H}\Big)^2\,dx,\tag{1.27}$$

where $H = X^\delta$ with $\delta \in (0,1)$ and

$$M_{X,H} = \frac{1}{X}\int_X^{2X} \sum_{x<n\le x+H} b(n)\,dx \sim K\frac{H}{\sqrt{\log X}}. \tag{1.28}$$

Here (1.27) is the *probabilistic variance* of $\sum_{x<n\le x+H} b(n) = B(x+H) - B(x)$ where

$$B(x) = \sum_{n\le x} b(n) \tag{1.29}$$

and (1.28) is the *probabilistic mean*. This is not exactly the right quantity to look at, owing to the fact that $b(n)$ on average behaves like $1/\sqrt{\log n}$, and the slow change of this function means that the variance in (1.27) will be much larger than we would like. Indeed, even the probabilistic variance of $\sum_{x\le n\le x+H} 1/\sqrt{\log n}$ is quite large owing to this change; the probabilistic variance of this sum is

$$\frac{1}{X}\int_X^{2X} \Big(\sum_{x<n\le x+H} \frac{1}{\sqrt{\log n}} - \frac{1}{X}\int_X^{2X} \sum_{t<n\le t+H} \frac{1}{\sqrt{\log n}}\,dt\Big)^2 dx,$$

and with a little work one may see that this is at least of order $H^2/(\log X)^3$.

Thus instead of (1.27), we consider a variant in which $M_{X,H}$ has been replaced by a better approximation to $\sum_{x<n\le x+H} b(n)$ which changes with x; this approximation is given in terms of an integral of L-functions. Define the function $F(s)$ for $\Re s > 1$ by

$$F(s) = \sum_{n=1}^\infty \frac{b(n)}{n^s}.$$

Using the fact that n is an element of S if and only if n can be written in the form $2^\alpha \mu\nu^2$, for μ a product of primes congruent to 1 modulo 4 and ν a product of primes congruent to 3 modulo 4, it may be seen that for $\Re s > 1$,

$$\begin{aligned} F(s) =& \frac{1}{1-2^{-s}} \prod_{q\equiv 1 \bmod 4} \frac{1}{1-q^{-s}} \prod_{r\equiv 3 \bmod 4} \frac{1}{1-r^{-2s}} \\ =& \Big(\frac{\zeta(s)L(s,\chi_4)}{1-2^{-s}}\Big)^{1/2} \prod_{k=1}^\infty \Big(\frac{\zeta(2^k s)}{L(2^k s,\chi_4)}(1-2^{-2^k s})\Big)^{1/2^{k+1}}, \end{aligned}$$

where χ_4 is the non-trivial character modulo 4. The first Euler product here dates at least back to Landau [Lan08], while the second factorization has in effect been derived many times (see e.g. [Sha64, FV96]).

The second representation allows one to analytically continue $F(s)$ to the cut disc $\mathcal{E} = \{s : |s-1| < 1/2\} \setminus \{s : \Im s = 0, \Re s \le 1\}$: note that in this region, because neither $\zeta(s)$ nor $L(s,\chi)$ have low-lying zeros inside of it (see [LMF18] for a list of zeros), we can write

$$F(s) = (s-1)^{-1/2} f(s), \tag{1.30}$$

where $f(s)$ is an analytic function and where the principal branch of the function $(s-1)^{-1/2}$ is taken. Under RH for $\zeta(s)$ and $L(s,\chi_4)$, it may be shown that (see §6.10 for a full proof)

$$B(x) = \overline{B}(x) + O_\varepsilon(x^{1/2+\varepsilon}), \quad \text{where} \quad \overline{B}(x) = \frac{1}{\pi}\int_{1/2}^1 \frac{x^s}{(1-s)^{1/2}s} f(s)\,ds$$

Thus we approximate $B(x+H)-B(x)$ (the number of elements of S in a short interval $(x, x+H]$) by

$$I(x,H) := \overline{B}(x+H) - \overline{B}(x).$$

We will consider variance defined in the following sense:

$$V_b(X,H) := \frac{1}{X}\int_X^{2X} (B(x+H)-B(x)-I(x,H))^2\,dx,$$

Ramachandra [Ram76] investigated a quantity equivalent to this one and showed that there is some cancellation over the trivial bound of $H^2/\log X$; namely

$$V_b(X,H) = O(H^2\exp(-(\log X)^{1/6})),$$

for $H > X^{1/6+\varepsilon}$. Under density hypotheses for the zeros of $\zeta(s)$ and $L(s,\chi_4)$ (see [Ram76, Eq. (6)]) this is improved to the more complete range $H > X^{\varepsilon}$. Motivated by Corollary 1.9, we conjecture the following

Conjecture 1. Fix $\delta \in (0,1)$. As $X \to \infty$ with $H = X^{\delta}$, we have

$$V_b(X,H) = \Big(K\,G(1-\delta) + o(1)\Big)\frac{H}{\sqrt{\log X}},$$

for K as in (1.24) and $G(s)$ as in (1.26).

The proofs of Theorem 1.7 and Proposition 1.8 are detailed in §6. Their are taken from a joint paper titled "The variance of the number of sums of two squares in $\mathbb{F}_q[T]$ in short intervals" [GR18] (to appear in AJM), authored by myself and Brad Rodgers. In the paper we also deal with the variance of the generalized divisor functions d_z (z not necessarily an integer).

In Figure 1 we plot numerical data supporting the conjecture.

1.9.4 Twin primes in the large-q limit

One of the oldest open problems in number theory is the existence of infinitely many *twin primes*, that is, primes with distance 2 from one another. This can be expressed as asking whether

$$\sum_{n\le x}\Lambda(n)\Lambda(n+2) \to \infty$$

as $x\to\infty$. A heuristic computation based on the circle method led Hardy and Littlewood to conjecture a quantitative version of the above [HL23]. The Hardy-Littlewood 2-Tuple Conjecture says that, as $x\to\infty$,

$$\frac{1}{x}\sum_{n\le x}\Lambda(n)\Lambda(n+\Delta) \sim \prod_p \frac{1-\frac{\#\{0 \bmod p,\,\Delta \bmod p\}}{p}}{(1-\frac{1}{p})^2}$$

for even $\Delta\in\mathbb{N}$. Over $\mathbb{F}_q[T]$, the analogous conjecture is

$$\frac{1}{\#\mathcal{M}_{n,q}}\sum_{f\in\mathcal{M}_{n,q}}\Lambda_q(n)\Lambda_q(n+\Delta) \sim \prod_{P\in\mathcal{P}_q}\frac{1-\frac{\#\{0 \bmod P,\,\Delta \bmod P\}}{\|P\|}}{(1-\frac{1}{\|P\|})^2} \tag{1.31}$$

for all non-zero polynomials Δ, as $q^n\to\infty$. When q is fixed while n tends to ∞, two results are known:

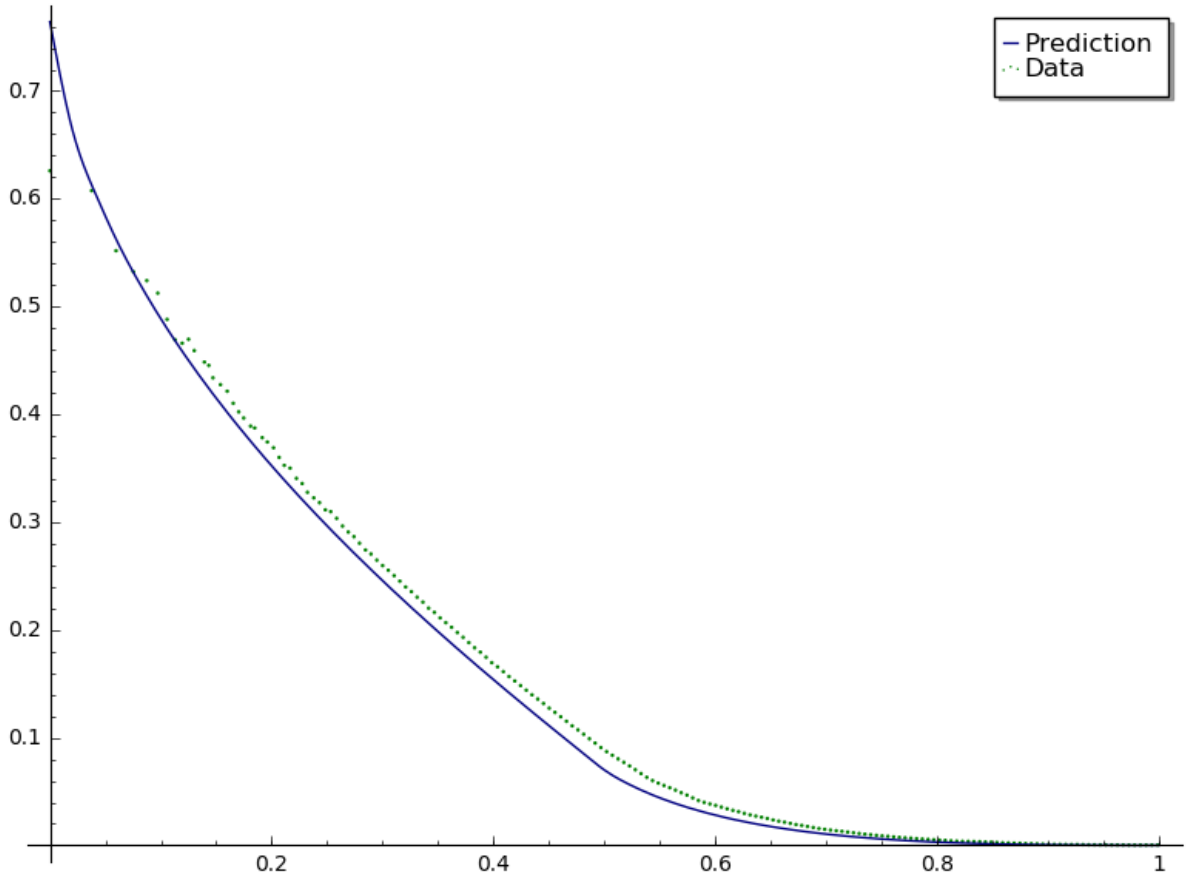

Figure 1: Numerically produced data compared to the z-measure induced prediction given in Conjecture 1 for variance in short intervals. Let $V_b(X, H)$ be the variance of counts of S in random short intervals $[x, x+H]$ for $n \leq X$. For $X = 10^8$ and $H \leq X$, set $\delta = \delta_H = \log(H)/\log(X)$. For a selection of H, we plot the points $(\delta, V_b(X,H)/(H/\sqrt{\log X}))$ under the label **data**, and the curve $(\delta, K\,G(1-\delta))$ under **prediction**.

- C. Hall [Hal03, Prop. 19], in his PhD Thesis, proved the existence of infinitely many twin primes with distance $c \in \mathbb{F}_q^\times$. His proof is constructive - he provided an explicit infinite family of $\ell \in \mathbb{N}$ for which there are $\beta \in \mathbb{F}_q$ with $T^\ell - \beta$, $T^\ell - \beta + c$ both irreducible. However, the family of ℓ's is quite sparse, and contains only perfect powers.
- Very recently, Sawin and Shusterman proved, for certain q's, the Hardy Littlewood 2-Tuple Conjecture in $\mathbb{F}_q[T]$ in the large-n limit [SS19]. Specifically, they require $q > 685090p^2$ where p is the characteristic of $\mathbb{F}_q$.

Here we concentrate on fixed n and growing q. The constant in the right-hand side of (1.31) is $1 + (a_\Delta - 1)/q + O_{\deg(\Delta)}(1/q^2)$, where a_Δ is the number of zeros of Δ in $\mathbb{F}_q$ (without multiplicities), see [GS20, Eq. (5.4)]. We may consider (1.31) for fixed n, and ask whether

$$\frac{1}{\#\mathcal{M}_{n,q}} \sum_{f \in \mathcal{M}_{n,q}} \Lambda_q(n)\Lambda_q(n+\Delta) = 1 + \frac{a_\Delta - 1}{q} + O_n\left(\frac{1}{q^2}\right)$$

as $q \to \infty$ (uniformly for Δ of degree between 0 and $n-1$). This question has received significant attention. Pollack [Pol08, Thm. 2], Bender and Pollack [BP09, Thm. 1.3],

Bary-Soroker [BS14, Thm. 1.1] and Carmon [Car15, §6] have shown that

$$\frac{1}{\#\mathcal{M}_{n,q}}\sum_{f\in\mathcal{M}_{n,q}}\Lambda_q(n)\Lambda_q(n+\Delta)=1+O_n\left(\frac{1}{\sqrt{q}}\right).$$

Their proof is based on Galois-theoretic methods, which give an implied constant of order $n!^2$ (see the statement of [BP09, Thm. 1.3]). We concentrate on scalar Δ and prove the following.

Theorem 1.10. *Fix $n\geq 4$. Uniformly for $c\in\mathbb{F}_q^\times$, we have*

$$\frac{\sum_{f\in\mathcal{M}_{n,q}}\Lambda_q(f)\Lambda_q(f+c)}{\#\mathcal{M}_{n,q}}=1+O\left(\frac{n^3}{\sqrt{q}}\right) \tag{1.32}$$

and

$$\frac{\sum_{f\in\mathcal{M}_{n,q}}\Lambda_q(f)\Lambda_q(f+c)}{\#\mathcal{M}_{n,q}}=1+O_n\left(\frac{1}{q}\right) \tag{1.33}$$

as $q\to\infty$.

The first estimate has the same dependence on q as the previous works, but polynomial dependence on n. Its proof is short and self-contained, taking barely 3 pages. The second estimate improves the exponent with which q appears, and is optimal, in the sense that the error term cannot be replaced with $O_n(1/q^\alpha)$, $\alpha>1$ (at least if the Hardy-Littlewood Conjecture is correct).

The main ideas of the proof of Theorem 1.10 are new. We introduce an L-function formula for the correlation of general arithmetic functions, which relates an average over polynomials to an average over short interval characters (Proposition 7.1). This falls into the general framework in analytic number theory where we replace an identity, in this case $f_2=f_1+c$, with an average over characters. This case may be surprising because we are detecting an additive identity using multiplicative Dirichlet characters. However, using Dirichlet characters ramified at ∞, it is possible to do this. The contribution of a given character is closely related to the Dirichlet L-function of that character.

We can compose a short interval character with a ring automorphism of $\mathbb{F}_q[T]$ to get a new character, which will have the same Dirichlet L-function. This gives an additional symmetry of the average over characters (Proposition 7.2) which we are able to use in order to derive (1.32), by first summing over compositions of a given character and then using pointwise estimates for Gauss sums that give a saving of $\sqrt{q}$. To prove (1.33), we use the fact that we sum over all characters in the family $G(R_n)$, and use L-function equidistribution results (Theorem 7.4) for a saving of an additional $\sqrt{q}$.

The proof of Theorem 1.10 is detailed in §7. The material of §7 is based on the joint paper "Correlation of arithmetic functions over $\mathbb{F}_q[T]$" published in Mathematische Annalen [GS20], authored by myself and Will Sawin. In the paper, we extend Theorem 1.10 to general factorization functions. We also study the easier problem where one averages over the larger space of non-monic polynomials. There one obtains stronger results than over monics, uncovering an (expected) lower term by relating the problem to a variance problem in intervals of size q.

2 Short interval characters

Here we review the notion of short interval characters in $\mathbb{F}_q[T]$. These are characters that can be used to detect whether two polynomials are 'close', that is, have a difference of small degree. These characters are ramified only at the prime at infinity, and are analogous to the functions $n \mapsto n^{it}$ ($t \in \mathbb{R}$) in the integers.

More general characters are studied, in an elementary fashion, by Hayes [Hay65], and in the context of class field theory by Weil [Wei74].

2.1 Equivalence relation

Let ℓ be a non-negative integer. We define an equivalence relation R_ℓ on $\mathcal{M}_q$ by saying that $A \equiv B \bmod R_\ell$ if and only if A and B have the same first ℓ next-to-leading coefficients. We adopt throughout the following convention: the j-th next-to-leading coefficient of a polynomial $f(T) \in \mathcal{M}_q$ with $j > \deg(f)$ is considered to be 0. It may be shown that there is a well-defined quotient monoid $\mathcal{M}_q/R_\ell$, where multiplication is the usual polynomial multiplication. Any element of $\mathcal{M}_q$ is invertible modulo R_ℓ, and $\mathcal{M}_q/R_\ell$ forms an abelian group, having as identity element the equivalence class of the polynomial 1. It may be shown that

$$|\mathcal{M}_q/R_\ell| = q^\ell.$$

2.2 Characters

For every character χ of the finite abelian group $\mathcal{M}_q/R_\ell$, we define $\chi^\dagger$ with domain $\mathcal{M}_q$ as follows. If $\mathfrak{c}$ is the equivalence class of A, then $\chi^\dagger(A) = \chi(\mathfrak{c})$. We shall abuse language somewhat and write χ instead of $\chi^\dagger$ to indicate a character of the relation R_ℓ derived from the character χ of the group $\mathcal{M}_q/R_\ell$. Thus we write χ_0 for the character of R_ℓ which is identically 1. We denote by $G(R_\ell)$ the set $\{\chi^\dagger : \chi \in \widehat{\mathcal{M}_q/R_\ell}\}$.

Elements of $G(R_\ell)$ are called "characters of the relation R_ℓ" or "characters modulo R_ℓ". We also call them "short interval characters of ℓ coefficients" (*short interval characters* for short), because for any $A \in \mathcal{M}_q$ of degree $\geq \ell$, $\chi \in G(R_\ell)$ is constant on the set

$$\{f \in \mathcal{M}_{\deg(A),q} : f \equiv A \bmod R_\ell\}$$

which is nothing but the short interval $I(A, \deg(A) - \ell - 1)$.

A character modulo R_ℓ is said to be primitive if it does not coincide with a character modulo $R_{\ell-1}$.

A set of polynomials in $\mathcal{M}_q$ is called a representative set modulo R_ℓ if the set contains one and only one polynomial from each equivalence class of R_ℓ. If $\chi_1, \chi_2 \in G(R_\ell)$, then

$$\frac{1}{q^\ell} \sum_F \chi_1(F)\overline{\chi_2}(F) = \begin{cases} 0 & \text{if } \chi_1 \neq \chi_2, \\ 1 & \text{if } \chi_1 = \chi_2, \end{cases} \tag{2.1}$$

F running through a representative set modulo R_ℓ. If $n \geq \ell$, then $\mathcal{M}_{n,q}$ is a disjoint union of $q^{n-\ell}$ representative sets. Thus, applying (2.1) with $\chi_2 = \chi_0$, we obtain that for all $n \geq \ell$,

$$\frac{1}{q^n} \sum_{F \in \mathcal{M}_{n,q}} \chi(F) = \begin{cases} 0 & \text{if } \chi \neq \chi_0, \\ 1 & \text{if } \chi = \chi_0. \end{cases} \tag{2.2}$$

We also have, for all $A, B \in \mathcal{M}_q$,

$$\frac{1}{q^\ell} \sum_{\chi \in G(R_\ell)} \chi(A)\overline{\chi}(B) = \begin{cases} 1 & \text{if } A \equiv B \bmod R_\ell, \\ 0 & \text{otherwise.} \end{cases} \tag{2.3}$$

2.3 L-functions

Let $\chi \in G(R_\ell)$. The L-function of χ is the following series in u:

$$L(u, \chi) = \sum_{f \in \mathcal{M}_q} \chi(f) u^{\deg(f)},$$

which also admits the Euler product

$$L(u, \chi) = \prod_{P \in \mathcal{P}} (1 - \chi(P) u^{\deg(P)})^{-1}. \tag{2.4}$$

The series converges in $|u| < 1/q$. If χ is the principal character χ_0 of $G(R_\ell)$, then

$$L(u, \chi) = \frac{1}{1 - qu}.$$

Otherwise, the orthogonality relation (2.2) implies that $L(u, \chi)$ is a polynomial in u of degree at most

$$\deg L(u, \chi) \leq \ell - 1. \tag{2.5}$$

The L-function $L(u, \chi)$ satisfies a functional equation (originally proved by Witt in the thirties; see Roquette [Roq18] for an historical account): if χ is a primitive character modulo R_ℓ then

$$L(u, \chi) = (\sqrt{q}u)^{\ell - 1} \varepsilon(\chi) L(\frac{1}{qu}, \overline{\chi}) \tag{2.6}$$

where $\varepsilon(\chi)$ is a quantity of modulus 1 (sometimes known as 'root number'). Comparing coefficients, we obtain that

$$\deg L(u, \chi) = \ell - 1$$

for primitive $\chi \in G(R_\ell)$, as well as the following equality:

$$\sum_{f \in \mathcal{M}_{i,q}} \chi(f) = \varepsilon(\chi) q^{i - \frac{\ell - 1}{2}} \sum_{f \in \mathcal{M}_{\ell - 1 - i, q}} \overline{\chi}(f), \tag{2.7}$$

for all $0 \leq i \leq \ell - 1$.

The first one to realize that Weil's proof of the Riemann Hypothesis for Function Fields [Wei74, Thm. 6, p. 134] implies the Riemann Hypothesis for the L-functions of $\chi \in G(R_\ell)$ was Rhin [Rhi72, Chapitre 2] in his thesis (cf. [EH91, Thm. 5.6] and the discussion following it). Hence we know that if we factor $L(u, \chi)$ as

$$L(u, \chi) = \prod_{i=1}^{\deg L(u, \chi)} (1 - \gamma_i(\chi) u), \tag{2.8}$$

then for any i,

$$|\gamma_i(\chi)| = \sqrt{q}. \tag{2.9}$$

We note the following standard consequences of (2.9), and include a proof for completeness.

Lemma 2.1. *Let $n, \ell \in \mathbb{N}$ and $\chi \in G(R_\ell)$. Then*

$$\left| \sum_{f \in \mathcal{M}_{n,q}} \Lambda_q(f)\chi(f) \right| \leq \begin{cases} (\ell-1)q^{\frac{n}{2}} & \text{if } \chi \neq \chi_0, \\ q^n & \text{if } \chi = \chi_0, \end{cases} \tag{2.10}$$

and

$$\left| \sum_{P \in \mathcal{P}_{n,q}} \chi(P) \right| \leq \begin{cases} \min\{\frac{q^{\frac{n}{2}}}{n}(\ell+1), \frac{q^n}{n}\} & \text{if } \chi \neq \chi_0, \\ \frac{q^n}{n} & \text{otherwise.} \end{cases}$$

Proof. The case $\chi = \chi_0$ of (2.10) follows from (1.10). For $\chi \neq \chi_0$, we equate (2.4) with (2.8) and take logarithmetic derivatives to obtain

$$\sum_{f \in \mathcal{M}_{n,q}} \Lambda_q(f)\chi(f) = - \sum_{i=1}^{\deg L(u,\chi)} \gamma_i(\chi)^n. \tag{2.11}$$

By (2.5), (2.9), (2.11) and the triangle inequality,

$$|\sum_{f \in \mathcal{M}_{n,q}} \Lambda_q(f)\chi(f)| \leq (\ell-1)q^{\frac{n}{2}}. \tag{2.12}$$

The bound $\left|\sum_{P \in \mathcal{P}_{n,q}} \chi(P)\right| \leq q^n/n$ follows from the bound $|\mathcal{P}_{n,q}| \leq q^n/n$ [Ros02, Prop. 2.1]. For $\chi \neq \chi_0$, we obtain the additional bound as follows. We can split (2.12) into the contribution of primes of degree n and proper prime powers:

$$\left| n \sum_{f \in \mathcal{P}_{n,q}} \chi(f) + \sum_{d|n,\, d\neq 1} \frac{n}{d} \sum_{f \in \mathcal{P}_{\frac{n}{d},q}} \chi^d(f) \right| \leq (\ell-1)q^{\frac{n}{2}}. \tag{2.13}$$

As

$$\left| \sum_{d|n,\, d\neq 1} \frac{n}{d} \sum_{f \in \mathcal{P}_{\frac{n}{d},q}} \chi^d(f) \right| \leq \sum_{d|n,\, d\neq 1} \frac{n}{d} |\mathcal{P}_{n/d,q}| \leq \sum_{d|n,\, d\neq 1} q^{n/d} \leq 2q^{\frac{n}{2}},$$

we obtain from (2.13) and the triangle inequality that

$$|n \sum_{f \in \mathcal{P}_{n,q}} \chi(f)| \leq (\ell+1)q^{\frac{n}{2}}.$$

After dividing by n, the lemma is established. □

2.4 Sums over short intervals and their variance

For an arithmetic function $\alpha\colon \mathcal{M}_q \to \mathbb{C}$ and $n \geq 1$, define

$$S(n, \alpha) = \sum_{f \in \mathcal{M}_{n,q}} \alpha(f). \tag{2.14}$$

The following lemma expresses sums over short intervals, and the variance of such sums, as sums over characters in $G(R_\ell)$. A variant of this lemma appeared in a paper of Keating and Rudnick [KR16].

Lemma 2.2. *Let $-1 \le h \le n-1$ and set $\ell = n-h-1$. Then the following hold.*

1. *For any $f, g \in \mathcal{M}_{n,q}$,*
$$\mathbf{1}_{g \in I(f,h)} = \frac{\sum_{\chi \in G(R_\ell)} \overline{\chi}(A)\chi(g)}{q^\ell}.$$
2. *For any arithmetic function $\alpha\colon \mathcal{M}_q \to \mathbb{C}$ and $f \in \mathcal{M}_{n,q}$ we have*
$$\begin{aligned}\sum_{g \in I(f,h)} \alpha(g) &= \frac{\sum_{\chi \in G(R_\ell)} \overline{\chi}(A) S(n, \alpha \cdot \chi)}{q^\ell} \\ &= q^{n-\ell}\langle \alpha \rangle_{\mathcal{M}_{n,q}} + \frac{\sum_{\chi_0 \neq \chi \in G(R_\ell)} \overline{\chi}(A) S(n, \alpha \cdot \chi)}{q^\ell}.\end{aligned} \tag{2.15}$$
3. *For any arithmetic function $\alpha\colon \mathcal{M}_q \to \mathbb{C}$, the variance $\mathrm{Var}_\alpha(n,h)$ of $\{\sum_{g \in I(f,h)} \alpha(g)\}_{f \in \mathcal{M}_{n,q}}$ may be expressed as*
$$\mathrm{Var}_\alpha(n,h) = \frac{\sum_{\chi_0 \neq \chi \in G(R_\ell)} |S(n, \alpha \cdot \chi)|^2}{q^{2\ell}}. \tag{2.16}$$

Proof. The first part of the lemma is a restatement of the orthogonality relation (2.3). For the second part of the lemma, we observe that $\sum_{g \in I(f,h)} \alpha(g) = \sum_{g \in \mathcal{M}_{n,q}} \alpha(g) \cdot \mathbf{1}_{g \in I(f,h)}$, and now we apply the first part of the lemma and interchange the order of summation. Note that
$$\frac{S(n, \alpha \cdot \chi_0)}{q^\ell} = q^{n-\ell}\langle \alpha \rangle_{\mathcal{M}_{n,q}}. \tag{2.17}$$
We now prove the last part of the lemma. Given $A \in \mathcal{M}_q/R_\ell$, write f_A for a polynomial in $\mathcal{M}_{n,q}$ in the equivalence class of A. We use (2.17) and (2.15) as follows:
$$\begin{aligned}\mathrm{Var}_\alpha(n,h) &= \frac{1}{q^n} \sum_{f \in \mathcal{M}_{n,q}} \left| \sum_{g \in I(f,h)} \alpha(g) - q^{h+1}\langle \alpha \rangle_{\mathcal{M}_{n,q}} \right|^2 \\ &= \frac{1}{q^\ell} \sum_{A \in \mathcal{M}/R_\ell} \left| \sum_{g \in I(f_A,h)} \alpha(g) - \frac{S(n, \alpha \cdot \chi_0)}{q^\ell} \right|^2 \\ &= \frac{1}{q^\ell} \sum_{A \in \mathcal{M}/R_\ell} \left| \frac{\sum_{\chi_0 \neq \chi \in G(R_\ell)} \overline{\chi}(A) S(n, \alpha \cdot \chi)}{q^\ell} \right|^2 \\ &= \frac{1}{q^{3\ell}} \sum_{A \in \mathcal{M}/R_\ell} \sum_{\chi_1, \chi_2 \in G(R_\ell) \setminus \{\chi_0\}} \overline{\chi_1}(A) S(n, \alpha \cdot \chi_1) \chi_2(A) \overline{S(n, \alpha \cdot \chi_2)}.\end{aligned}$$
Interchanging the order of summation and applying (2.1), we conclude the proof. □

3 The variance of divisor function in short intervals

3.1 Moments of d_2

Lemma 3.1. *' We have*
$$\langle d_q \rangle_{\mathcal{M}_{n,q}} = n+1$$

and

$$\langle d_q^2 \rangle_{\mathcal{M}_{n,q}} = \binom{n+3}{3} - \frac{1}{q}\binom{n+1}{3}. \tag{3.1}$$

Proof. The first part follows by interchanging the order of summation:

$$\begin{aligned}\frac{1}{q^n}\sum_{f\in\mathcal{M}_{n,q}} d_q(f) = \frac{1}{q^n}\sum_{f\in\mathcal{M}_{n,q}}\sum_{d|f} 1 = \frac{1}{q^n}\sum_{d\in\mathcal{M}_q}\sum_{f\in\mathcal{M}_{n,q}:\, d|f} 1 = \frac{1}{q^n}\sum_{d\in\mathcal{M}_q,\,\deg(d)\le n} q^{n-\deg(d)} \\ = \sum_{d\in\mathcal{M}_q,\,\deg(d)\le n} q^{-\deg(d)} = n+1.\end{aligned}$$

For the second part, let $Z(u) = \sum_{f\in\mathcal{M}_q} u^{\deg(f)} = \prod_{P\in\mathcal{P}_q}(1-u^{\deg(P)})^{-1}$. Since $Z(q^{-s}) = \zeta_q(s)$, we have $Z(u) = 1/1(1-qu)$. We use the multiplicativity of d_q to write

$$\begin{aligned}\sum_{f\in\mathcal{M}_q} d_q^2(f)u^{\deg(f)} &= \prod_{P\in\mathcal{P}_q}\left(\sum_{k\ge 0}(k+1)^2 u^{k\deg(P)}\right) \\ &= \prod_{P\in\mathcal{P}_q}\frac{1+u^{\deg(P)}}{(1-u^{\deg(P)})^3} \\ &= \prod_{P\in\mathcal{P}_q}\frac{(1-u^{2\deg(P)})}{(1-u^{\deg(P)})^4} \\ &= \frac{Z(u)^4}{Z(u^2)} = \frac{1-qu^2}{(1-qu)^4}.\end{aligned}$$

By the binomial theorem, $(1-qu)^{-4} = \sum_{n\ge 0} u^n(-q)^n\binom{-4}{n} = \sum_{n\ge 0} u^n q^n\binom{n+3}{3}$, and so $\sum_{f\in\mathcal{M}_{n,q}} d_q^2(f) = q^n\binom{n+3}{3} - q\cdot q^{n-2}\binom{(n-2)+3}{3}$, and dividing through by q^n concludes the proof. □

3.2 Correlation sums proof

Lemma 3.2. *Let $A, B \in \mathcal{M}_q$ be coprime polynomials. Let n be a positive integer such that*

$$n \ge \deg(A) + \deg(B).$$

Let Δ be a non-zero polynomial of degree $< n$. Then the number of solutions to the polynomial equation

$$Au - Bv = \Delta, \deg(Au) = \deg(Bv) = n, \quad u, v \text{ monic}$$

is $q^{n-\deg(A)-\deg(B)}$.

Proof. The case $\Delta = 1$ is proven in [ABSR15, Lem. 7.2]. The general case is proved in the same way. □

Lemma 3.3. *Let Δ be a non-zero polynomial of degree $< n$. Let $a_{i,\Delta}$ denote the number of monic divisors of Δ of degree i. Then*

$$\frac{\sum_{f\in\mathcal{M}_{n,q}} d_q(f)d_q(f+\Delta)}{q^n} = (n+1)^2 + \sum_{i=1}^{\lfloor \frac{n}{2}\rfloor}\frac{(n-2i+1)^2}{q^i}(a_{i,\Delta} - a_{i-1,\Delta}).$$

Proof. We follow and generalize the proof of [ABSR15, Thm. 7.1]. Fix a positive integer n. Let $\alpha, \beta, \gamma, \delta$ be non-negative integers such that $\alpha + \beta = \gamma + \delta = n$. Set

$$S(\alpha, \beta; \gamma, \delta; \Delta) = \#\{x \in \mathcal{M}_{\alpha,q}, y \in \mathcal{M}_{\beta,q}, z \in \mathcal{M}_{\gamma,q}, u \in \mathcal{M}_{\delta,q} : xy - zu = \Delta\}.$$

We have some obvious symmetries from the definition:

$$S(\alpha, \beta; \gamma, \delta; \Delta) = S(\beta, \alpha; \gamma, \delta; \Delta) = S(\alpha, \beta; \delta, \gamma; \Delta).$$

Hence, to evaluate $S(\alpha, \beta; \gamma, \delta; \Delta)$ it suffices to assume that

$$\alpha \leq \beta, \quad \gamma \leq \delta. \tag{3.2}$$

Assuming (3.2) we write

$$S(\alpha, \beta; \gamma, \delta; \Delta) = \sum_{\substack{g|\Delta \\ \text{monic}}} \sum_{\substack{x \in \mathcal{M}_{\alpha,q} \\ z \in \mathcal{M}_{\gamma,q} \\ \gcd(x,z)=g}} \#\{y \in \mathcal{M}_{\beta,q}, u \in \mathcal{M}_{\delta,q} : xy - zu = \Delta\}. \tag{3.3}$$

If $x \in \mathcal{M}_{\alpha,q}$, $z \in \mathcal{M}_{\gamma,q}$ and $\gcd(x, z) = g$, then x/g, y/g are coprime and $\deg(g) \leq \alpha \leq n/2$. Hence

$$\#\{y \in \mathcal{M}_{\beta,q}, u \in \mathcal{M}_{\delta,q} : xy - zu = \Delta\} = \#\{y \in \mathcal{M}_{\beta,q}, u \in \mathcal{M}_{\delta,q} : \frac{x}{g}y - \frac{z}{g}u = \frac{\Delta}{g}\}, \tag{3.4}$$

and by Lemma 3.2 we have

$$\#\{y \in \mathcal{M}_{\beta,q}, u \in \mathcal{M}_{\delta,q} : \frac{x}{g}y - \frac{z}{g}u = \frac{\Delta}{g}\} = q^{n-\alpha-\gamma+\deg(g)}. \tag{3.5}$$

Plugging (3.4) and (3.5) in (3.3), we get that

$$S(\alpha, \beta; \gamma, \delta; \Delta) = \sum_{\substack{g|\Delta, \text{ monic} \\ \deg(g) \leq \min\{\alpha,\beta,\gamma,\delta\}}} q^{n-\alpha-\gamma+\deg g} \sum_{\substack{x \in \mathcal{M}_{\alpha,q} \\ z \in \mathcal{M}_{\gamma,q} \\ \gcd(x,z)=g}} 1. \tag{3.6}$$

Note that by [ABSR15, Eqs. (7.14-7.17)],

$$\sum_{\substack{x \in \mathcal{M}_{\alpha,q} \\ z \in \mathcal{M}_{\gamma,q} \\ \gcd(x,z)=g}} 1 = \sum_{\substack{x' \in \mathcal{M}_{\alpha-\deg(g),q} \\ z' \in \mathcal{M}_{\gamma-\deg(g),q} \\ \gcd(x',z')=1}} 1 = q^{\alpha-\deg(g)+\gamma-\deg(g)} \cdot \begin{cases} 1 & \alpha = \deg(g) \text{ or } \gamma = \deg(g), \\ 1 - \frac{1}{q} & \alpha > \deg(g) \text{ or } \gamma > \deg(g). \end{cases} \tag{3.7}$$

Plugging (3.7) in (3.6) we get that

$$S(\alpha, \beta; \gamma, \delta; \Delta) = q^n \sum_{\substack{g|\Delta, \text{ monic} \\ \deg(g) \leq \min\{\alpha,\beta,\gamma,\delta\}}} q^{-\deg(g)} \left(1 - \frac{1}{q} \cdot \mathbf{1}_{\min\{\alpha,\beta,\gamma,\delta\} > \deg(g)}\right). \tag{3.8}$$

Consider the following sum:

$$\sum_{f \in \mathcal{M}_{n,q}} d_q(f) d_q(f+\Delta) = \#\{x, y, z, u \in \mathcal{M}_q : xy - zu = \Delta, \deg(xy) = \deg(zu) = n\}. \tag{3.9}$$

We partition the right-hand side of (3.9) into a sum over variables with fixed degree, that it

$$\sum_{f\in\mathcal{M}_{n,q}} d_q(f)d_q(f+\Delta) = \sum_{\substack{\alpha+\beta=\gamma+\delta=n,\\ \alpha,\beta,\gamma,\delta\geq 0}} S(\alpha,\beta;\gamma,\delta;\Delta). \tag{3.10}$$

Plugging (3.8) in (3.10) we get that

$$\frac{\sum_{f\in\mathcal{M}_{n,q}} d_q(f)d_q(f+\Delta)}{q^n} = \sum_{\substack{\alpha+\beta=\gamma+\delta=n,\\ \alpha,\beta,\gamma,\delta\geq 0}} \sum_{\substack{g|\Delta,\text{ monic}\\ \deg(g)\leq\min\{\alpha,\beta,\gamma,\delta\}}} q^{-\deg(g)}\left(1-\frac{1}{q}\cdot\mathbf{1}_{\min\{\alpha,\beta,\gamma,\delta\}>\deg(g)}\right).$$

Interchanging the order of summation we obtain

$$\begin{aligned}\frac{\sum_{f\in\mathcal{M}_{n,q}} d_q(f)d_q(f+\Delta)}{q^n} &= \sum_{\substack{g|\Delta,\text{ monic}\\ \deg(g)\leq\frac{n}{2}}} q^{-\deg(g)} \sum_{\substack{\alpha+\beta=\gamma+\delta=n,\\ \alpha,\beta,\gamma,\delta\geq 0\\ \deg(g)\leq\min\{\alpha,\beta,\gamma,\delta\}}} \left(1-\frac{1}{q}\cdot\mathbf{1}_{\min\{\alpha,\beta,\gamma,\delta\}>\deg(g)}\right)\\ &= \sum_{\substack{g|\Delta,\text{ monic}\\ \deg(g)\leq\frac{n}{2}}} q^{-\deg(g)}\Big((n-2\deg(g)+1)^2\\ &\qquad -\frac{(n-2\deg(g)-1)^2}{q}\cdot\mathbf{1}_{\deg(g)\leq\frac{n}{2}-1}\Big)\\ &= (n+1)^2+\sum_{i=1}^{\lfloor\frac{n}{2}\rfloor}\frac{(n-2i+1)^2}{q^i}(a_{i,\Delta}-a_{i-1,\Delta}),\end{aligned}$$

as claimed. □

Lemma 3.4. *Let $\alpha\colon \mathcal{M}_q\to\mathbb{C}$. For $-1\leq h\leq n-1$,*

$$\frac{\mathrm{Var}_\alpha(n,h)}{q^{h+1}} = \sum_{\Delta\in\mathbb{F}_q[T],\deg(\Delta)\leq h}\left(\langle\alpha(f)\overline{\alpha}(f+\Delta)\rangle_{f\in\mathcal{M}_{n,q}}-\left|\langle\alpha\rangle_{\mathcal{M}_{n,q}}\right|^2\right),$$

where the summation is over (not necessarily monic) polynomials Δ.

Proof. Expanding the square in $\mathrm{Var}_\alpha(n,h)$, we have

$$\mathrm{Var}_\alpha(n,h) = \frac{1}{q^n}\sum_{f_0\in\mathcal{M}_{n,q}}\sum_{f,g\in I(f_0,h)}\alpha(f)\overline{\alpha}(g)-q^{2h+2}|\langle\alpha\rangle_{\mathcal{M}_{n,q}}|^2.$$

Letting $\Delta:=f-g$, we see that $\deg(\Delta)\leq h$ and that $\alpha(f)\overline{\alpha}(f+\Delta)$ appears with weight q^{h+1-n} (since we may take f_0 to be any polynomial in $I(f,h)$), which concludes the proof once we divide by q^{h+1}. □

3.2.1 Conclusion

Applying Lemma 3.4 with $\alpha = d_q$, separating the contribution of $\Delta = 0$ from the rest of the terms and then using Lemmas 3.3 and 3.1 to evaluate the terms, we obtain

$$\begin{aligned}\frac{\mathrm{Var}_{d_q}(n,h)}{q^{h+1}} &= \langle |d_q|^2 \rangle_{\mathcal{M}_{n,q}} - |\langle d_q \rangle_{\mathcal{M}_{n,q}}|^2 + \sum_{0 \le \deg(\Delta) \le h} \sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} \frac{(n-2i+1)^2}{q^i} (a_{i,\Delta} - a_{i-1,\Delta}) \\ &= \binom{n+3}{3} - \frac{1}{q}\binom{n+1}{3} - (n+1)^2 + \sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} \frac{(n-2i+1)^2}{q^i} \sum_{0 \le \deg(\Delta) \le h} (a_{i,\Delta} - a_{i-1,\Delta}).\end{aligned} \tag{3.11}$$

To evaluate the inner sum, observe that

$$\begin{aligned}\sum_{\deg(\Delta)=k} a_{i,\Delta} &= \sum_{d \in \mathcal{M}_{d,q}} \sum_{\deg(\Delta)=k} \mathbf{1}_{d|\Delta} \\ &= \sum_{d \in \mathcal{M}_{i,q}} (q-1) q^{k-\deg(d)} \cdot \mathbf{1}_{k \ge \deg(d)} = q^k (q-1) \cdot \mathbf{1}_{k \ge i}.\end{aligned} \tag{3.12}$$

From (3.12),

$$\sum_{\deg(\Delta)=k} (a_{i,\Delta} - a_{i-1,\Delta}) = -q^k (q-1) \cdot \mathbf{1}_{k=i-1}. \tag{3.13}$$

Plugging (3.13) in (3.11), we obtain

$$\frac{\mathrm{Var}_{d_q}(n,h)}{q^{h+1}} = \binom{n+3}{3} - \frac{1}{q}\binom{n+1}{3} - (n+1)^2 - \frac{q-1}{q} \sum_{1 \le i \le \lfloor \frac{n}{2} \rfloor, h+1} (n-2i+1)^2.$$

Now it is a matter of school algebra to show that the above expression simplifies to $(1-1/q)\binom{n-2h-1}{3}$ for $h \le \lfloor n/2 \rfloor - 1$, and vanishes for larger h. □

3.3 Functional equation proof

We write $[u^n]f$ for the coefficient of u^n in a power series f. Since d_q is the Dirichlet convolution of the constant function $\mathbf{1}$ with itself, we have

$$S(n, d_q \cdot \chi) = [u^n] L(u,\chi)^2. \tag{3.14}$$

From the third part of Lemma 2.2 and from (3.14) we obtain

$$\mathrm{Var}_{d_q}(n,h) = \sum_{k=1}^{n-h-1} \frac{\sum_{\chi \in G(R_k) \setminus G(R_{k-1})} |[u^n] L(u,\chi)^2|^2}{q^{2(n-h-1)}}. \tag{3.15}$$

If $\chi \in G(R_k) \setminus G(R_{k-1})$ then $\deg L(u,\chi)^2 = 2(k-1)$. Hence, the k's for which $n > 2(k-1)$ contribute 0 to the sum in (3.15), so that

$$\mathrm{Var}_{d_q}(n,h) = \sum_{k=\lceil \frac{n}{2} \rceil + 1}^{n-h-1} \frac{\sum_{\chi \in G(R_k) \setminus G(R_{k-1})} |[u^n] L(u,\chi)^2|^2}{q^{2(n-h-1)}}.$$

From the functional equation (2.6) we obtain

$$\begin{aligned}\mathrm{Var}_{d_q}(n,h) &= \sum_{k=\lceil\frac{n}{2}\rceil+1}^{n-h-1} \frac{\sum_{\chi\in G(R_k)\setminus G(R_{k-1})}\left|[u^n]\left(L(\frac{1}{qu},\overline{\chi})(\sqrt{q}u)^{k-1}\varepsilon(\chi)\right)^2\right|^2}{q^{2(n-h-1)}} \\ &= \sum_{k=\lceil\frac{n}{2}\rceil+1}^{n-h-1} q^{2n-2k+2}\frac{\sum_{\chi\in G(R_k)\setminus G(R_{k-1})}\left|[u^{2k-2-n}]L(u,\overline{\chi})^2\right|^2}{q^{2(n-h-1)}}.\end{aligned} \tag{3.16}$$

By (3.14), we may replace $[u^{2k-2-n}]L(u,\overline{\chi})^2$ with $S(2k-2-n,d_q\cdot\overline{\chi})$. Doing so, and then interchanging the order of summation in (3.16), we have

$$\begin{aligned}\mathrm{Var}_{d_q}(n,h) &= \sum_{k=\lceil\frac{n}{2}\rceil+1}^{n-h-1} q^{2n-2k+2}\frac{\sum_{\chi\in G(R_k)\setminus G(R_{k-1})}\sum_{f,g\in\mathcal{M}_{2k-2-n,q}} d_q(f)\chi(f)\overline{d_q(g)\chi(g)}}{q^{2(n-h-1)}} \\ &= \sum_{k=\lceil\frac{n}{2}\rceil+1}^{n-h-1}\frac{\sum_{f,g\in\mathcal{M}_{2k-2-n,q}} d_q(f)d_q(g)}{q^{2k-2h-4}}\sum_{\chi\in G(R_k)\setminus G(R_{k-1})}\chi(f)\overline{\chi}(g).\end{aligned} \tag{3.17}$$

Since $2k-2-n\le k-1\le k$, we have

$$f\equiv g \bmod R_{k-1}\leftrightarrow f\equiv g \bmod R_k\leftrightarrow f=g$$

for all $f,g\in\mathcal{M}_{2k-2-n,q}$. Hence, the orthogonality relation (2.3) implies that

$$\sum_{\chi\in G(R_k)\setminus G(R_{k-1})}\chi(f)\overline{\chi}(g)=\left(q^k-q^{k-1}\right)\cdot\mathbf{1}_{f=g} \tag{3.18}$$

for all $f,g\in\mathcal{M}_{2k-2-n,q}$. Plugging (3.18) in (3.17), we find that

$$\mathrm{Var}_{d_q}(n,h)=\left(1-\frac{1}{q}\right)\sum_{k=\lceil\frac{n}{2}\rceil+1}^{n-h-1} q^{k+2h-n+2}\langle d_q^2\rangle_{\mathcal{M}_{2k-2-n,q}}. \tag{3.19}$$

From (3.1) and (3.19), we have that

$$\mathrm{Var}_{d_q}(n,h)=\left(1-\frac{1}{q}\right)\sum_{k=\lceil\frac{n}{2}\rceil+1}^{n-h-1}\left(q^{k+2h-n+2}\binom{2k+1-n}{3}-q^{k+2h-n+1}\binom{2k-1-n}{3}\right).$$

If $h\ge\lfloor n/2\rfloor-1$, the sum is 0 as it is empty and is 0. Otherwise, it telescopes to

$$\left(1-\frac{1}{q}\right)q^{h+1}\binom{n-2h-1}{3},$$

as needed. □

4 The variance of squarefree polynomials in short intervals

4.1 Bounds on character sums and Möbius sums

The following bounds are originally due to Bhowmick, Lê and Liu [BLL17, Thms. 1–2]. They also follow from Theorem 5.1 below.

Lemma 4.1. *[BLL17, Thms. 1–2] Let $k \geq 1$ and let $\chi_0 \neq \chi \in G(R_k)$. We have*

$$\sum_{f \in \mathcal{M}_{i,q}} \chi(f), \sum_{f \in \mathcal{M}_{i,q}} \mu_q(f)\chi(f) = O\left(q^{\frac{i}{2}} e^{O_q\left(\frac{i \log\log(k+2)}{\log(k+2)} + \frac{k+2}{\log^2(k+2)}\right)}\right).$$

4.2 Proof of Proposition 1.3

We begin by assuming that q is odd. By Lemma 2.2,

$$\mathrm{Var}_{\beta_m}(n,h) = \frac{\sum_{\chi_0 \neq \chi \in G(R_{n-h-1})} \left|\sum_{f \in \mathcal{M}_{n,q}} \beta_m(f)\chi(f)\right|^2}{q^{2(n-h-1)}}. \tag{4.1}$$

We have

$$\sum_{f \in \mathcal{M}_{n,q}} \beta_m(f)\chi(f) = \sum_{f \in \mathcal{M}_{n,q}} \sum_{\substack{f = d^2 g \\ \deg(d) > m}} \mu_q(d)\chi(d^2 g) = \sum_{i > m}^{\lfloor n/2 \rfloor} \sum_{d \in \mathcal{M}_{i,q}} \mu_q(d)\chi^2(d) \sum_{g \in \mathcal{M}_{n-2i,q}} \chi(g). \tag{4.2}$$

Observe that we may start the summation over i from $i \geq (h+1)/2$, since $\sum_{g \in \mathcal{M}_{n-2i,q}} \chi(g)$ vanishes for smaller i. In odd characteristic, if χ is non-trivial then so is χ^2 (since $\mathcal{M}_q/R_\ell$ is of odd order q^ℓ). Hence, we may apply Lemma 4.1 with $k = n \geq n - h - 1$ and with both χ and χ^2, obtaining

$$\left|\sum_{f \in \mathcal{M}_{n,q}} \beta_m(f)\chi(f)\right| \leq \sum_{i > m,\, i \geq (h+1)/2}^{\lfloor n/2 \rfloor} \left|q^{\frac{i}{2}} q^{\frac{n-2i}{2}} e^{o_q(n)}\right| \leq q^{\frac{n - \max\{m, h/2\}}{2}} e^{o_q(n)}. \tag{4.3}$$

by the triangle inequality. Plugging (4.3) in (4.1), we obtain

$$\mathrm{Var}_{\beta_m}(n,h) \leq q^{h - \max\{m, h/2\}} e^{o_q(n)}, \tag{4.4}$$

as needed. If q is even, there are non-trivial characters $\chi \in G(R_{n-h-1})$ such that χ^2 is trivial. For such characters,

$$\sum_{d \in \mathcal{M}_{i,q}} \mu_q(d)\chi^2(d) = \sum_{d \in \mathcal{M}_{i,q}} \mu_q(d) = 0$$

if $i > 1$, according to (1.10). For $i = 0, 1$ we have $\sum_{g \in \mathcal{M}_{n-2i,q}} \chi(g) = 0$. This implies $\sum_{f \in \mathcal{M}_{n,q}} \beta_m(f)\chi(f) = 0$. Hence (4.4) holds for all characters and the proof goes through for even q as well. □

4.3 Proof of Proposition 1.4

By Lemma 2.2,

$$\mathrm{Var}_{\alpha_m}(n,h) = \frac{\sum_{\chi_0 \neq \chi \in G(R_{n-h-1})} \left| \sum_{f \in \mathcal{M}_{n,q}} \alpha_m(f)\chi(f) \right|^2}{q^{2(n-h-1)}}. \tag{4.5}$$

As in (4.2),

$$\sum_{f \in \mathcal{M}_{n,q}} \alpha_m(f)\chi(f) = \sum_{i \leq m} \sum_{d \in \mathcal{M}_{i,q}} \mu_q(d)\chi^2(d) \sum_{g \in \mathcal{M}_{n-2i,q}} \chi(g).$$

If $\chi \in G(R_k) \setminus G(R_{k-1})$, then $\sum_{g \in \mathcal{M}_{n-2i,q}} \chi(g) = 0$ if $n - 2i \geq k$. Applying the functional equation in the form (2.7), we may write

$$\sum_{f \in \mathcal{M}_{n,q}} \alpha_m(f)\chi(f) = \varepsilon(\chi) \sum_{i \leq m,\, (n-k)/2} q^{n-2i-\frac{k-1}{2}} \sum_{d \in \mathcal{M}_{i,q}} \mu_q(d)\chi^2(d) \sum_{g \in \mathcal{M}_{k-1-(n-2i),q}} \overline{\chi}(g) \tag{4.6}$$

where $|\varepsilon(\chi)| = 1$. Plugging (4.6) in (4.5), we obtain

$$\begin{aligned}\mathrm{Var}_{\alpha_m}(n,h) = q^{2h+3} \sum_{k=1}^{n-h-1} & \sum_{\chi \in G(R_k) \setminus G(R_{k-1})} q^{-k} \\ & \times \left| \sum_{i \leq m,\, (n-k)/2} q^{-2i} \sum_{d \in \mathcal{M}_{i,q}} \mu_q(d)\chi^2(d) \sum_{g \in \mathcal{M}_{k-1-(n-2i),q}} \overline{\chi}(g) \right|^2. \end{aligned} \tag{4.7}$$

The orthogonality relation (2.3) tells us that

$$\sum_{\chi \in G(R_k) \setminus G(R_{k-1})} \chi(f_1)\overline{\chi}(f_2) = q^{k-1} \left(q \cdot \mathbf{1}_{f_1 \equiv f_2 \bmod R_k} - \mathbf{1}_{f_1 \equiv f_2 \bmod R_{k-1}} \right). \tag{4.8}$$

Expanding the square in (4.7), interchanging the order of summation and plugging (4.8) yields

$$\begin{aligned}\mathrm{Var}_{\alpha_m}(n,h) = q^{2h+2} & \sum_{k=1}^{n-h-1} \sum_{i_1, i_2 \leq m,\, (n-k)/2} \\ & \sum_{\substack{d_1 \in \mathcal{M}_{i_1,q} \\ d_2 \in \mathcal{M}_{i_2,q} \\ g_1 \in \mathcal{M}_{k-1-(n-2i_1),q} \\ g_2 \in \mathcal{M}_{k-1-(n-2i_2),q}}} \mu_q(d_1)\mu_q(d_2) q^{-2(i_1+i_2)} \left(q \cdot \mathbf{1}_{d_1^2 g_2 \equiv d_2^2 g_1 \bmod R_k} - \mathbf{1}_{d_1^2 g_2 \equiv d_2^2 g_1 \bmod R_{k-1}} \right). \end{aligned} \tag{4.9}$$

We have $\deg(d_1^2 g_2) = \deg(d_2^2 g_1) = 2i_1 + 2i_2 + k - 1 - n \leq k - 1$ since $m \leq n/4$. If $f_1 \equiv f_2 \bmod R_{k-1}$ or $f_1 \equiv f_2 \bmod R_k$ for polynomials f_1, f_2 of the same degree which is at most $k-1$, then f_1 must equal f_2. Hence, the only terms that contribute to (4.9) are those with $d_1^2 g_2 = d_2^2 g_1$ ('diagonal terms'), and so

$$\mathrm{Var}_{\alpha_m}(n,h) = q^{2h+2}(q-1) \sum_{k=1}^{n-h-1} \sum_{i_1, i_2 \leq m,\, (n-k)/2} \sum_{\substack{d_1 \in \mathcal{M}_{i_1,q} \\ d_2 \in \mathcal{M}_{i_2,q} \\ g_1 \in \mathcal{M}_{k-1-(n-2i_1),q} \\ g_2 \in \mathcal{M}_{k-1-(n-2i_2),q} \\ d_1^2 g_2 = d_2^2 g_1}} \mu_q(d_1)\mu_q(d_2) q^{-2(i_1+i_2)}.$$

The condition $d_1^2 g_2 = d_2^2 g_1$ can be written as

$$\left(\frac{d_1}{\gcd(d_1,d_2)}\right)^2 g_2 = \left(\frac{d_2}{\gcd(d_1,d_2)}\right)^2 g_1,$$

which is equivalent to

$$g_i = \left(\frac{d_i}{\gcd(d_1,d_2)}\right)^2 s$$

for a monic polynomial s of degree $2\deg\gcd(d_1,d_2) - (n+1-k)$ (of which there exist $q^{2\deg\gcd(d_1,d_2)-(n+1-k)}$ if $\deg\gcd(d_1,d_2) \geq (n+1-k)/2$, and otherwise there are no such s). Writing D for $\deg\gcd(d_1,d_2)$, we can rewrite the last sum as

$$\mathrm{Var}_{\alpha_m}(n,h) = q^{2h+2}(q-1)\sum_{k=1}^{n-h-1}\sum_{i_1,i_2\leq m,\,(n-k)/2}\sum_{\substack{d_1\in\mathcal{M}_{i_1,q}\\ d_2\in\mathcal{M}_{i_2,q}}}\frac{\mu_q(d_1)\mu_q(d_2)}{\|d_1\|^2\|d_2\|^2}q^{2D-(n+1-k)}\cdot\mathbf{1}_{2D\geq n+1-k}.$$

Summing first over i_1 and i_2, we may write

$$\mathrm{Var}_{\alpha_m}(n,h) = q^{2h+2}(q-1)\sum_{i_1,i_2\leq m}\sum_{\substack{d_1\in\mathcal{M}_{i_1,q}\\ d_2\in\mathcal{M}_{i_2,q}}}\frac{\mu_q(d_1)\mu_q(d_2)}{\|d_1\|^2\|d_2\|^2}\sum_{k=n+1-2D}^{n-h-1}q^{2D-(n+1-k)}.$$

The inner sum is $q^{2D-h-2}\sum_{r=0}^{2D-h-2}q^{-r} = q^{2D-h-2}(1-q^{-(2D-h-1)})/(1-q^{-1})$ if $D \geq (h+2)/2$, and is 0 otherwise. All in all, $\mathrm{Var}_{\alpha_m}(n,h) = F(h,m)$ where

$$F(h,m) = q^{h+1}\sum_{\substack{i=1,2:\deg(d_i)\leq m,\,d_i\text{ monic}\\ \deg\gcd(d_1,d_2)\geq(h+2)/2}}\frac{\mu_q(d_1)\mu_q(d_2)}{\|d_1\|^2\|d_2\|^2}\|\gcd(d_1,d_2)\|^2(1-q^{-(2\deg\gcd(d_1,d_2)-h-1)}). \tag{4.10}$$

This concludes the proof. □

4.4 Over the integers and far away

In [GMRR20] we prove the following integer analogue of Theorem 1.2.

Theorem 4.2. *We have* $\mathrm{Var}_{\mu^2}(X,H) \sim CH^{1/2}$ *as* $X\to\infty$ *in the range* $H = O_\varepsilon(X^{11/6-\varepsilon})$, $H\to\infty$.

As can be seen, we go slightly beyond the $X^{1/2-\varepsilon}$ range! We explain some of the ideas that go into the proof of Theorem 4.2. The starting point is again a decomposition of μ^2:

$$\mu^2 = \alpha_M + \beta_M, \qquad \alpha_M(n) = \sum_{d^2|n,\,d\leq M}\mu(d), \qquad \beta_M(n) = \sum_{d^2|n,\,d>M}\mu(d).$$

4.4.1 On Proposition 1.4 in $\mathbb{Z}$

We want to evaluate

$$I = \mathrm{Var}_{\alpha_M}(X,H) = \frac{1}{X}\int_0^X\left(\sum_{\substack{d^2m\in[x,x+H]\\ d\leq M}}\mu(d) - H\sum_{d\leq M}\frac{\mu(d)}{d^2}\right)^2 dx.$$

Here $\sum_{d\le M}\mu(d)/d^2$ is the mean value of α_M. In place of the functional equation of $L(u,\chi)$ used originally, we use a Poisson summation argument (which is philosophically the same). To apply it, we first introduce smoothing. For smooth $\sigma,\ \rho\colon \mathbb{R}\to\mathbb{R}$ concentrated on $[0,1]$, consider the smoothed variance

$$\widetilde{I}=\frac{1}{X}\int_0^X \sigma\left(\frac{x}{X}\right)\left(\sum_{m\in\mathbb{Z},\, d\le M}\mu(d)\rho\left(\frac{d^2m-x}{H}\right)-\hat{\rho}(0)H\sum_{d\le M}\frac{\mu(d)}{d^2}\right)^2 dx. \tag{4.11}$$

We can now apply Poisson summation to

$$f(t)=\sum_{d\le M}\mu(d)\rho\left(\frac{d^2t-x}{H}\right)$$

obtaining

$$\sum_{m\in\mathbb{Z},\, d\le M}\mu(d)\rho\left(\frac{d^2m-x}{H}\right)=H\sum_{\nu\in\mathbb{Z},\, d\le M}\frac{\mu(d)}{d^2}\hat{\rho}(\frac{H\nu}{d^2})e^{-2\pi i\nu x/d^2}.$$

The term $\hat{\rho}(0)\sum_{d\le M}\mu(d)/d^2$ cancels with the contribution of $\nu=0$. Plugging in (4.11), we have

$$\widetilde{I}=H^2\sum_{d_1,d_2\le M}\sum_{\nu_1,\nu_2\in\mathbb{Z}\setminus\{0\}}\frac{\mu(d_1)\mu(d_2)}{d_1^2d_2^2}\hat{\rho}(\frac{H\nu_1}{d_1^2})\overline{\hat{\rho}(\frac{H\nu_2}{d_2^2})}\hat{\sigma}(X(\frac{\nu_1}{d_1^2}-\frac{\nu_2}{d_2^2})).$$

We choose σ so that the support of $\hat{\sigma}$ is in $(-1,1)$. If M is small enough, namely $M\le X^{1/4}$, only the 'diagonal terms' (i.e. those with $\nu_1/d_1^2=\nu_2/d_2^2$) survive, since otherwise we have a lower bound on $|X(\nu_1/d_1^2\nu_2/d_2^2)|$ which is outside the support: $|X(\nu_1/d_1^2-\nu_2/d_2^2)|\ge X/(d_1^2d_2^2)\ge X/M^4\ge 1$. Hence,

$$\widetilde{I}=\hat{\sigma}(0)H^2\sum_{\substack{d_1,d_2\le M\\ \nu_1,\nu_2\in\mathbb{Z}\setminus\{0\}\\ \nu_1d_2^2=\nu_2d_1^2}}\frac{\mu(d_1)\mu(d_2)}{d_1^2d_2^2}\left|\hat{\rho}(\frac{H\nu_1}{d_1^2})\right|^2.$$

This expression can either be evaluated directly, or as in $\mathbb{F}_q[T]$, we can evaluate it indirectly by a bootstrapping-type argument. A useful feature is that $\widetilde{I}$ does not depend directly on X, but rather on H and M only. Removing the weights σ and ρ is routine, by taking a limit of weight functions.

In the paper, we are in fact able estimate $\widetilde{I}$ for M larger than $X^{1/4}$, by studying some of the non-diagonal terms that arise. That is, terms with $\nu_1/d_1^2-\nu_2/d_2^2$ small but non-zero. We bound the contribution of such terms using the theory of quadratic forms.

4.4.2 On Proposition 1.3 in $\mathbb{Z}$

The starting point for Proposition 1.3 is a Plancherel-type formula, namely (2.16). Over the integers, we do not have such a nice, simple formula available. Fortunately, since we only want to upper bound

$$J=\mathrm{Var}_{\beta_M}(X,H)=\frac{1}{X}\int_0^X\left(\sum_{\substack{d^2m\in[x,x+H]\\ d>M}}\mu(d)-H\sum_{d>M}\frac{\mu(d)}{d^2}\right)^2 dx,$$

we can make some simplifications which allow us to pass to Fourier space. An inequality of Saffari and Vaughan [SV77, p. 25] tells us (roughly) that J is bounded by

$$\widetilde{J} = \frac{1}{X}\int_0^{2X}\left(\sum_{\substack{d^2m\in[x,x(1+\theta)]\\ d>M}}\mu(d) - \theta x\sum_{d>M}\frac{\mu(d)}{d^2}\right)^2 dx$$

for some θ with $\theta \asymp H/X$ (so that the length of the interval $[x, x(1+\theta)]$ is of order $O(H)$). In this new form, there is Plancherel-type formula, first found by Selberg in his study of the variance of Λ [Sel43]. We explain his formula. We have, by Perron's formula,

$$\sum_{n\le x}\Lambda(n) = \frac{1}{2\pi i}\int_{(c)}\frac{\zeta'(s)}{\zeta(s)}\frac{x^s}{s}\,ds$$

for any $c > 1$ and non-integer x. At least under RH, we can shift the contour to $c \in (1/2, 1)$, picking a pole at $s = 1$:

$$\sum_{n\le x}\Lambda(n) = x + \frac{1}{2\pi i}\int_{(c)}\frac{\zeta'(s)}{\zeta(s)}\frac{x^s}{s}\,ds.$$

Using this with x and $x(1+\theta)$, we find

$$\sum_{n\in[x,x(1+\theta)]}(\Lambda(n)-1) \approx \frac{1}{2\pi i}\int_{(c)}\frac{\zeta'(s)}{\zeta(s)}\frac{x^s((1+\theta)^s-1)}{s}\,ds = \frac{1}{2\pi}\int_{\mathbb{R}}\frac{\zeta'(c+it)}{\zeta(c+it)}x^{it}\frac{x^c((1+\theta)^s-1)}{c+it}\,dt.$$

Writing e^x in place of x,

$$\frac{\sum_{n\in[e^x,e^x(1+\theta)]}(\Lambda(n)-1)}{e^{xc}} \approx \frac{1}{2\pi}\int_{\mathbb{R}}e^{ixt}\frac{\zeta'(c+it)}{\zeta(c+it)}\frac{(1+\theta)^{c+it}-1}{c+it}\,dt. \tag{4.12}$$

The function of x in the left-hand side of (4.12) is the Fourier transform of

$$\frac{\zeta'(c+it)}{\zeta(c+it)}\frac{(1+\theta)^{c+it}-1}{c+it},$$

and upon applying Plancherel theorem we obtain

$$\int_{\mathbb{R}}|\sum_{n\in[e^u,e^u(1+\theta)]}(\Lambda(n)-1)|^2\frac{du}{e^{2cu}} \approx 2\pi\int_{\mathbb{R}}\left|\frac{\zeta'(c+it)}{\zeta(c+it)}\right|^2\left|\frac{(1+\theta)^{c+it}-1}{c+it}\right|^2 dt.$$

(As written, the integrals do not necessarily converge. However, introducing certain weights solves the issue.) At this point, we can input information on the zeros and growth of ζ to study the variance of primes.

A similar 'trick' can be used to relate $\widetilde{J}$ to an integral of a Dirichlet series. In practice, we dissect the series $\sum_{n\ge1}\beta_m(n)/n^s = \zeta(s)\sum_{d>M}\mu(d)/d^{2s}$ into dyadic pieces: $\zeta(s)\sum_{d\in I_k}\mu(d)/d^{2s}$ for $I_k = (2^kM, 2^{k+1}m]$. This is beneficial, since Dirichlet polynomials are easier to work with than with series (for instance, we may take $c = 1/2$ without worrying about convergence issues), and more importantly, the contributions of the different Dirichlet polynomials are different so they should be estimated separately. To summarize,

the study of J reduces to that of $\widetilde{J}$, and by a dydadic decomposition and an application of Plancherel, we want to bound

$$\int_{\mathbb{R}} |\zeta(1/2+it)|^2 \left|\sum_{d\in I_k} \frac{\mu(d)}{d^{2it}}\right|^2 \left|\frac{(1+\theta)^{1/2+it}-1}{1/2+it}\right|^2 dt.$$

RH gives strong pointwise bounds on $\sum_{d\in I_k} \mu(d)/d^{2it}$, which allow us to bound the integral quite easily. However, we can do nicely without RH, by combining standard tools for handling Dirichlet polynomials. These include Montgomery's mean value theorem, Huxley's large value theorem and subconvexity bounds for ζ. For more details, see [GMRR20].

5 The variance of factorization functions in short intervals

5.1 Strategy of proof of Theorem 1.5

By Lemma 2.2 and the triangle inequality, we have

$$\left|\sum_{f\in I(f_0,h)} \alpha(f) - q^{h+1}\langle\alpha\rangle_{\mathcal{M}_{n,q}}\right| \leq \max_{\chi_0\neq\chi\in G(R_{n-h-1})} |\sum_{f\in\mathcal{M}_{n,q}} \alpha(f)\chi(f)|$$

and

$$\mathrm{Var}_\alpha(n,h) \leq \frac{\max_{\chi_0\neq\chi\in G(R_{n-h-1})} \left|\sum_{f\in\mathcal{M}_{n,q}} \alpha(f)\chi(f)\right|^2}{q^{n-h-1}}.$$

Thus, the result follows at once from the following theorem.

Theorem 5.1. *Let $\alpha\colon \mathcal{M}_q \to \mathbb{C}$ be a factorization function, and let $\chi_0 \neq \chi \in G(R_\ell)$.*

$$\left|\sum_{f\in\mathcal{M}_{n,q}} \alpha(f)\chi(f)\right| \leq \max_{f\in\mathcal{M}_{n,q}} |\alpha(f)|\, q^{\frac{n}{2}} e^{O_q(\frac{n\log\log(\ell+2)}{\log(\ell+2)} + \frac{\ell}{\log^2(\ell+2)})}. \tag{5.1}$$

Indeed, if $\ell = n-h-1$ then the exponent of e in (5.1) is $O(n\log\log(n+2)/\log(n+2))$. The rest of this section is dedicated to the proof of Theorem 5.1. We manage to prove such a theorem, which works for any α, by reducing it to estimating a single character sum, which we now describe.

Let Ω be the set of finite multisets of elements from $\mathbb{N}\times\mathbb{N}$, so that ω_f, the factorization type of a polynomial f, is an element of Ω. For an element $\omega = \{(d_i,e_i) : 1\leq i\leq k\} \in \Omega$, we define its size to be $|\omega| := \sum_{i=1}^k d_i e_i$ and its length to be $\ell(\omega) := k$. For a factorization function α and a factorization type $\omega \in \Omega$, we denote by $\alpha(\omega)$ the value of α on a polynomial f with $\omega_f = \omega$ if such a polynomial exists, and otherwise set $\alpha(\omega) = 0$. We have, by the triangle inequality,

$$\left|\sum_{f\in\mathcal{M}_{n,q}} \alpha(f)\chi(f)\right| = \left|\sum_{\substack{\omega\in\Omega\\|\omega|=n}} \alpha(\omega) \sum_{\substack{f\in\mathcal{M}_{n,q}\\ \omega_f=\omega}} \chi(f)\right| \leq \max_{f\in\mathcal{M}_{n,q}} |\alpha(f)| \sum_{\substack{\omega\in\Omega\\|\omega|=n}} \left|\sum_{\substack{f\in\mathcal{M}_{n,q}\\ \omega_f=\omega}} \chi(f)\right|. \tag{5.2}$$

Thus, it suffices to bound the sum on the right-hand side of (5.2). In order to bound the character sums

$$\sum_{f \in \mathcal{M}_{n,q}:\, \omega_f = \omega} \chi(f),$$

we relate them to products of *monomial symmetric polynomials* evaluated at $\chi(P)$ where P runs over irreducible polynomials of certain degrees, see Lemmas 5.2 and 5.3 below. We use tools from symmetric function theory in order to be able to use RH efficiently in bounding these evaluations of symmetric polynomials.

We shall use the notation $[u^n]f(u)$ for the coefficient of u^n in a power series f. We also write $\exp(\bullet)$ for $e^{\bullet}$.

5.2 Preparation for proof of Theorem 5.1

5.2.1 Symmetric function theory

A partition of size n is a finite (possibly empty) non-increasing sequence of positive integers that sum to n. The length of a partition $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ is the number of its elements and is denoted by $\ell(\lambda) := k$. We write $\lambda \vdash n$ to indicate that λ sums to n. The empty partition is of size and length 0. We denote by $\mathbb{Y}$ the set of all partitions.

An important class of symmetric polynomials is the *monomial symmetric polynomials*. Given a partition $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ and variables $X_1, \ldots, X_k$, then the monomial symmetric polynomial $m_\lambda(X_1, \ldots, X_k)$ is the symmetric polynomial

$$m_\lambda(X_1, \ldots, X_k) := \sum_{\exists \pi \in S_k : (\lambda'_1, \ldots, \lambda'_k) = (\lambda_{\pi(1)}, \ldots, \lambda_{\pi(k)})} \prod_{i=1}^{k} X_i^{\lambda'_i} \in \mathbb{Z}[X_1, \ldots, X_k],$$

where the sum is over the *distinct* permutations of λ. It is useful to extend m_λ to the case of a general number of variables $X_1, \ldots, X_n$. If $n < k$ we define $m_\lambda(X_i : 1 \le i \le n)$ to be zero. If $n > k$ we set $\lambda_j = 0$ for $j = k+1, \ldots, n$ and define

$$m_\lambda(X_i : 1 \le i \le n) := \sum_{\exists \pi \in S_n : (\lambda'_1, \ldots, \lambda'_n) = (\lambda_{\pi(1)}, \ldots, \lambda_{\pi(n)})} \prod_{i=1}^{n} X_i^{\lambda'_i} \in \mathbb{Z}[X_1, \ldots, X_n],$$

where the sum is over the *distinct* permutations of λ followed by $n-k$ zeros. In particular, m_λ is the elementary symmetric polynomial e_k if $\lambda = (1, 1, \ldots, 1)$ with k ones.

Another class of symmetric polynomials is the *power sum symmetric polynomials*. Given a positive integer r, the power sum symmetric polynomial $p_r(X_i : 1 \le i \le n)$ is the symmetric polynomial

$$p_r(X_i : 1 \le i \le n) := \sum_{i=1}^{n} X_i^r \in \mathbb{Z}[X_1, \ldots, X_n].$$

More generally, given a partition $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$, then the power sum symmetric polynomial $p_\lambda(X_i : 1 \le i \le n)$ is the symmetric polynomial

$$p_\lambda(X_i : 1 \le i \le n) := \prod_{i=1}^{k} p_{\lambda_i}(X_i : 1 \le i \le n).$$

A basic result in symmetric function theory says that whenever $m \geq n$, $\{m_\lambda(X_i : 1 \leq i \leq m)\}_{\lambda \vdash n}$ and $\{p_\lambda(X_i : 1 \leq i \leq m)\}_{\lambda \vdash n}$ are both bases for homogeneous symmetric polynomials of degree n with rational coefficients. In particular, $m_\lambda(X_i : 1 \leq i \leq m)$ can be expressed uniquely as a linear combination of the symmetric polynomials p_μ for partitions μ of size n, that is, there are unique coefficients $c_{\lambda,\mu} \in \mathbb{Q}$ such that

$$m_\lambda(X_i : 1 \leq i \leq m) = \sum_{\mu \vdash n} c_{\lambda,\mu} p_\mu(X_i : 1 \leq i \leq m) \tag{5.3}$$

for all $m \geq n$ ($c_{\lambda,\mu}$ are independent of m).

5.2.2 Multiplicativity of character sums

Given $d \geq 1$ and a factorization type

$$\omega = \{(d_i, e_i) : 1 \leq i \leq k\} \in \Omega,$$

we denote by $\omega(d) \subseteq \omega$ the factorization type $\{(d_i, e_i) : 1 \leq i \leq k,\ d_i = d\}$. By definition, ω is the disjoint of union of the $\omega(d)$-s. Let $\mathbf{1}_\omega$ be the indicator function of polynomials f with $\omega_f = \omega$. The following lemma shows that the character sums $S(n, \chi \cdot \mathbf{1}_\omega)$ (recall (2.14)) enjoy a multiplicative property.

Lemma 5.2. *Let $\omega \in \Omega$ with $|\omega| = n$. Let χ be a Hayes character. Then*

$$S(n, \chi \cdot \mathbf{1}_\omega) = \prod_{d=1}^{n} S(|\omega(d)|, \chi \cdot \mathbf{1}_{\omega(d)}). \tag{5.4}$$

Proof. Each f with $\omega_f = \omega$ can be written uniquely as $f = \prod_{d=1}^n f_d$ where f_d is divisible only by primes of degree d. We then have $\omega_{f_d} = \omega(d)$, and the lemma follows by expanding the right-hand side of (5.4). □

5.2.3 Symmetric function theory

The following lemma expresses character sums, of the form appearing in the right-hand side of (5.4), as an evaluation of a monomial symmetric polynomial.

Lemma 5.3. *Let $\omega = \{(d, e_i) : 1 \leq i \leq k\} \in \Omega$. Let $\lambda \in \mathbb{Y}$ be the partition whose parts are $\{e_i\}_{i=1}^k$ in non-increasing order. Let χ be a Hayes character. We have*

$$S(|\omega|, \chi \cdot \mathbf{1}_\omega) = m_\lambda(\chi(P) : P \in \mathcal{P}_d).$$

Proof. A polynomial f with $\omega_f = \omega$ is necessarily given by a product $\prod_{i=1}^k P_i^{e_i}$ where P_i are distinct elements from $\mathcal{P}_{d,q}$. Equivalently, f may be expressed as $\prod_{P \in \mathcal{P}_{d,q}} P^{e(P)}$ where the multiset $\{e(P) : P \in \mathcal{P}_{d,q}\}$ is equal to

$$E := \{e_i : 1 \leq i \leq k\} \cup \{0 : 1 \leq i \leq |\mathcal{P}_{d,q}| - k\}.$$

Moreover, by unique factorization, this form is unique. Thus,

$$\begin{aligned} S(|\omega|, \chi \cdot \mathbf{1}_\omega) &= \sum_{f \in \mathcal{M}_{|\omega|} : \omega_f = \omega} \chi(f) = \sum_{\substack{e \colon \mathcal{P}_{d,q} \to \mathbb{N}_{\geq 0} \\ \{e(P) : P \in \mathcal{P}_{d,q}\} = E}} \chi\left(\prod_{P \in \mathcal{P}_{d,q}} P^{e(P)} \right) \\ &= \sum_{\substack{e \colon \mathcal{P}_{d,q} \to \mathbb{N}_{\geq 0} \\ \{e(P) : P \in \mathcal{P}_{d,q}\} = E}} \prod_{P \in \mathcal{P}_{d,q}} \big(\chi(P)\big)^{e(P)}, \end{aligned}$$

which is just m_λ evaluated at $\{\chi(P) : P \in \mathcal{P}_{d,q}\}$, as needed. □

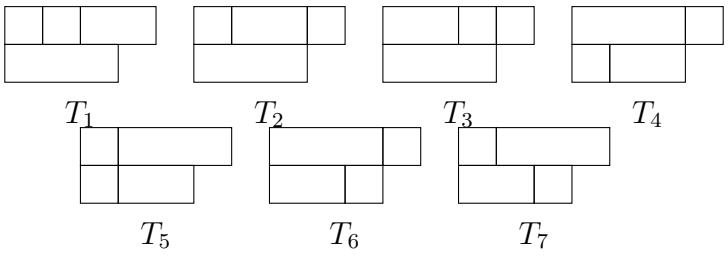

Figure 2: $(3,2,1,1)$-brick tabloids of shape $(4,3)$

Eğecioğlu and Remmel [ER91, pp. 107–111] gave a combinatorial interpretation of $c_{\lambda,\mu}$ in (5.3) which we now describe. We begin with their definition of λ-brick tabloids.

Let $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$, $\mu = (\mu_1, \mu_2, \ldots, \mu_r)$ be two partitions. Recall that the *Young diagram* Y_μ is the diagram which consists of left justified rows of squares of lengths $\mu_1, \mu_2, \ldots, \mu_k$ reading from top to bottom. For instance, if $\mu = (4,3)$ then Y_μ is given by

$$Y_\mu = \begin{array}{|c|c|c|c|}\hline \multicolumn{1}{|c}{} & \multicolumn{1}{c|}{} & \multicolumn{1}{c|}{} & \multicolumn{1}{c|}{} \\ \hline \multicolumn{1}{|c}{} & \multicolumn{1}{c|}{} & \multicolumn{1}{c|}{} \\ \cline{1-3} \end{array}.$$

A *λ-brick tabloid T of shape μ* is a filling of Y_μ with bricks $b_1, \ldots, b_k$ of lengths $\lambda_1, \ldots, \lambda_k$, respectively, such that

1. each brick b_i covers exactly λ_i squares of Y_μ all of which lie in a single row of Y_μ,
2. no two brick overlap.

For example, if $\lambda = (3,2,1,1)$ and $\mu = (4,3)$, then we must cover Y_μ with the bricks

$$\underbrace{\begin{array}{|c|c|c|}\hline \multicolumn{1}{|c}{} & \multicolumn{1}{c|}{} & \multicolumn{1}{c|}{} \\ \hline \end{array}}_{b_1}, \quad \underbrace{\begin{array}{|c|c|}\hline \multicolumn{1}{|c}{} & \multicolumn{1}{c|}{} \\ \hline \end{array}}_{b_2}, \quad \underbrace{\begin{array}{|c|}\hline \multicolumn{1}{|c}{} \\ \hline \end{array}}_{b_3}, \quad \underbrace{\begin{array}{|c|}\hline \multicolumn{1}{|c}{} \\ \hline \end{array}}_{b_4}.$$

Here, bricks of the same size are indistinguishable. There are in total seven λ-brick tabloids of shape μ, given in Figure 2.

We let $B_{\lambda,\mu}$ denote the set of λ-brick tabloids of shape μ. We define a weight $w(T)$ for each λ-brick tabloid $T \in B_{\lambda,\mu}$ by

$$w(T) = \prod_{b \in T} w_T(b),$$

where for each brick b in T, $|b|$ denotes the length of b and

$$w_T(b) = \begin{cases} 1 & \text{if } b \text{ is not at the end of a row in } T, \\ |b| & \text{if } b \text{ is at the end of a row in } T. \end{cases}$$

Thus $w(T)$ is the product of the lengths of the rightmost bricks in T. For example, for the seven $(3,2,1,1)$-brick tabloids of shape $(4,3)$ given in Figure 2, the weights are computed to be $w(T_1) = 6$, $w(T_2) = 3$, $w(T_3) = 3$, $w(T_4) = 2$, $w(T_5) = 6$, $w(T_6) = 1$ and $w(T_7) = 3$. We let

$$w(B_{\lambda,\mu}) := \sum_{T \in B_{\lambda,\mu}} w(T).$$

Eğecioğlu and Remmel [ER91, p. 111, Rel. (11)] proved that for partitions $\lambda, \mu \vdash n$ we have

$$c_{\lambda,\mu} = (-1)^{\ell(\lambda)-\ell(\mu)} w(B_{\lambda,\mu}) \mathbb{P}_\mu, \tag{5.5}$$

where

$$\mathbb{P}_\mu := \mathbb{P}_{\pi \in S_n}(\pi \text{ has cycle type } \mu).$$

Here $\mathbb{P}_{\pi \in S_n}$ is the uniform probability measure on the symmetric group S_n, and we say that π has a cycle type $(\mu_1, \mu_2, \ldots, \mu_r)$ if the cycle sizes of π are given by $\mu_1, \ldots, \mu_r$.

Lemma 5.4. *Let n and k be positive integers. Let $\mu \vdash n$. We have*

$$\sum_{\substack{\lambda \vdash n \\ \ell(\lambda) \leq k}} w(B_{\lambda,\mu}) \leq \sum_{i=0}^{k} \binom{n}{i}. \tag{5.6}$$

Proof. Write μ as $(\mu_1, \ldots, \mu_r)$. A λ-brick tabloid of shape μ determines the partition λ uniquely. Indeed, λ can be recovered by reading the lengths of the bricks in each row of the tabloid. Thus, the set $\cup_{\lambda \vdash n,\, \ell(\lambda) \leq k} B_{\lambda,\mu}$ may be identified with a sequence $\{b_i\}_{i=1}^r$ of positive integers with $\sum_{i=1}^r b_i \leq k$, and a double sequence $\{a_{i,j}\}_{1 \leq i \leq r,\, 1 \leq j \leq b_i}$ of positive integers with $\sum_{j=1}^{b_i} a_{i,j} = \mu_i$ for each i as follows. The number b_i is set to be the number of blocks in the i-th topmost row of the tabloid, and the number $a_{i,j}$ is set to be the length of th j-th leftmost brick in the i-topmost row. Under this identification, $w(B_{\lambda,\mu})$ is given by the product $\prod_{i=1}^r a_{i,b_i}$, and it follows that for any $t \geq 0$ we have

$$\sum_{\substack{\lambda \vdash n \\ \ell(\lambda) = t}} w(B_{\lambda,\mu}) = \sum_{\substack{b_1, \ldots, b_r \geq 1 \\ b_1 + \ldots + b_r = t}} \sum_{\substack{\forall 1 \leq i \leq r: \\ a_{i,1}, \ldots, a_{i,b_i} \geq 1 \\ \sum_{j=1}^{b_i} a_{i,j} = \mu_i}} \prod_{i=1}^{r} a_{i,b_i}. \tag{5.7}$$

Consider the generating function

$$B(u) := \sum_{\lambda \vdash n} \omega(B_{\lambda,\mu}) u^{\ell(\lambda)}.$$

Letting $c(n_1, n_2, n_3)$ be the number of solutions to $x_1 + x_2 + \ldots + x_{n_1} = n_3$ with $x_{n_1} = n_2$ and $x_i \geq 1$, it follows from (5.7) that

$$B(u) = \sum_{\substack{\forall 1 \leq i \leq r: \\ b_i, y_i \geq 1}} \prod_{i=1}^{r} c(b_i, y_i, \mu_i) y_i u^{b_i} = \prod_{i=1}^{r} \Big(\sum_{b_i, y_i \geq 1} c(b_i, y_i, \mu_i) y_i u^{b_i} \Big).$$

As $c(n_1, n_2, n_3)$ is also the number of solutions to $x_1 + \ldots + x_{n_1 - 1} = n_3 - n_2$ in positive integers, a standard combinatorial result says that

$$c(n_1, n_2, n_3) = \begin{cases} \binom{n_3 - n_2 - 1}{n_1 - 2} & \text{if } n_3 \geq n_2 + n_1 - 1,\ n_1 \geq 2, \\ \mathbf{1}_{n_2 = n_3} & \text{if } n_1 = 1, \\ 0 & \text{otherwise,} \end{cases}$$

so that

$$\begin{aligned} B(u) &= \prod_{i=1}^{r}(\mu_i u + \sum_{y=1}^{\mu_i-1}\sum_{b=2}^{\mu_i-y+1}\binom{\mu_i-y-1}{b-2}yu^b) \\ &= \prod_{i=1}^{r}(\mu_i u + \sum_{y=1}^{\mu_i-1} yu^2(1+u)^{\mu_i-y-1}) \\ &= \prod_{i=1}^{r}((1+u)^{\mu_i}-1), \end{aligned}$$

where in the last passage we made use of the identity $\sum_{i=1}^{d} ix^i = x(dx^{d+1} - (d+1)x^d + 1)/(x-1)^2$ with $d = \mu_i - 1$ and $x = 1/(1+u)$. As the left-hand side of (5.6) is the sum of the first $k+1$ coefficients of $B(u)$, and they are bounded from above by the corresponding coefficients of $\prod_{i=1}^{r}(1+u)^{\mu_i} = (1+u)^n = \sum_{i=0}^{n}\binom{n}{i}u^i$, the proof is concluded. □

5.2.4 Permutation statistics

We denote the expectation of a function $f\colon S_n \to \mathbb{R}$ with respect to the uniform probability measure on S_n by $\mathbb{E}_{\pi\in S_n} f(\pi)$. We denote by $\ell(\pi)$ the number of cycles in a permutation π.

Lemma 5.5. *Let $n \geq 1$, $m \geq 2$ be positive integers. Let $z_1, z_2 \in \mathbb{C}$. Define the following function on S_n:*

$$f(\pi) = \prod_{C\in\pi,\, m\nmid|C|} z_1 \prod_{C\in\pi,\, m||C|} z_2 \tag{5.8}$$

where the product is over the disjoint cycles of π. Then

$$\mathbb{E}_{\pi\in S_n} f(\pi) = [u^n](1-u)^{-z_1}(1-u^m)^{(-z_2+z_1)/m}.$$

Proof. The exponential formula for permutations [Sta99, Cor. 5.1.9] states the following. Given a function $g\colon \mathbb{N} \to \mathbb{C}$, we construct a corresponding function on permutations (on arbitrary number of elements) as follows:

$$G(\pi) = \prod_{C\in\pi} g(|C|),$$

where the product is over the disjoint cycles of π. We then have the following identity of formal power series:

$$1 + \sum_{i\geq 1}\left(\mathbb{E}_{\pi\in S_i} G(\pi)\right)u^i = \exp(\sum_{j\geq 1}\frac{g(j)}{j}u^j).$$

Applying the ideneity with

$$g(j) = \begin{cases} z_1 & \text{if } m \nmid j, \\ z_2 & \text{otherwise,} \end{cases}$$

we find that $G(\pi) = f(\pi)$ for every $\pi \in S_n$, and the lemma follows by a short computation. □

5.2.5 Bounds on certain finite sums

Lemma 5.6. *Let $x \geq 2$ be a real number and n be a positive integer. Then*

1. $\sum_{d_1 d_2 = n} 2^{d_1} x^{d_2} \leq 8x^n$.
2. *If furthermore $x \geq 4$,* $\sum_{d_1 d_2 = n,\, d_2 < n} 2^{d_1} x^{d_2} \leq 10 x^{n/2}$.

Proof. We begin with the first part. We may assume $n > 1$. Consider the function $f(t) = 2^{\frac{n}{t}} x^t$ on $[1, n]$. Its derivative is

$$f'(t) = f(t)(\log x - \frac{n \log 2}{t^2}),$$

so that f is either increasing from 1 to n (if $n \log 2 / \log x \leq 1$), or decreasing from 1 to $\sqrt{n \log 2 / \log x}$ and increasing from $\sqrt{n \log 2 / \log x}$ to n (otherwise). As $f(1) = 2^n x \leq 2x^n = f(n)$, it follows that

$$\sum_{d_1 d_2 = n} 2^{d_1} x^{d_2} = \sum_{d_2 | n} f(d_2) \leq 4x^n + (n-2) \max\{f(2), f(n/2)\} \cdot \mathbf{1}_{n \text{ not a prime}}.$$

If n is a prime, we are done. Otherwise, it suffices to show that $(n-2)2^2 x^{n/2} \leq 4x^n$ and that $(n-2)2^{n/2} x^2 \leq 4x^n$. For $n = 4, 5$ these are easy to verify, and for larger n they follow by induction. This establishes the first part of the lemma. The proof of the second part follows similar lines and is therefore omitted. □

The following variant of Lemma 5.6 is also needed. We omit the similar proof.

Lemma 5.7. *Let $x \in \{\sqrt{2}, \sqrt{3}, 2, 3\}$. For any positive integer n we have*

$$\sum_{d_1 d_2 = n,\, d_1 \neq n} 2^{d_1} x^{d_2} \leq 7x^n.$$

For $x \in \{2, 3\}$ we also have $\sum_{d_1 d_2 = n,\, 1 < d_2 < n} 2^{d_1} x^{d_2} + 1.4^n x \leq 14 x^{n/2}$.

Lemma 5.8. *For any $n \geq 1$ we have $\sum_{i=0}^{3} \binom{n}{i} \leq 7 \cdot 1.4^n$.*

Proof. For $n \leq 9$ this is checked by a short computation. For $n \geq 10$ we have $\sum_{i=0}^{3} \binom{n}{i} = (n^3 + 5n + 6)/6 \leq 2n^3$, and an inductive argument shows that $2n^3 \leq 7 \cdot 1.4^n$. □

5.2.6 Bounds on coefficients of a generating function

The following lemma is based on recent proofs by Bhowmick, Lê and Liu [BLL17, Thms. 1–2].

Lemma 5.9. *Let $t, r \geq 2$. Set $L := \lfloor 2 \log_t r \rfloor$ and*

$$Z(u) := \exp\Big(\sum_{1 \leq k \leq L} \frac{20 t^k}{k} u^k + \sum_{k > L} 20(r+1) \frac{t^{\frac{k}{2}}}{k} u^k \Big).$$

Then

$$|[u^n] Z(u)| \leq t^{\frac{n}{2}} \binom{20(r+1) + n - 1}{n} \tag{5.9}$$

for all $n \geq 1$. If $r \geq \max\{200000, t^{\log^2 t}\}$, we have

$$|[u^n]Z(u)| \leq t^{\frac{n}{2}} t^{\frac{n \log\log r}{\log r}} \exp(140 \frac{(r+1)}{(\log r)^2} t) \tag{5.10}$$

for all $n \geq 1$. If $r = O(n)$, we have

$$|[u^n]Z(u)| \leq t^{\frac{n}{2} + O_t(\frac{n \log\log(n+2)}{\log(n+2)})}. \tag{5.11}$$

Proof. We work with the modified function

$$\widetilde{Z}(u) := Z(u/\sqrt{t}),$$

so that

$$[u^n]Z(u) = t^{n/2}[u^n]\widetilde{Z}(u).$$

As $t^{k/2} \leq r+1$ for $k \leq L$, we have

$$\left|[u^n]\widetilde{Z}(u)\right| \leq \left|[u^n]\exp(\sum_{k\geq 1} 20(r+1)\frac{u^k}{k})\right| = [u^n](1-u)^{-20(r+1)},$$

which establishes (5.9). As $\widetilde{Z}$ has non-negative coefficients and radius of convergence 1, we have

$$|[u^n]\widetilde{Z}(u)| \leq \sum_{i\geq 0} R^{i-n}[u^i]\widetilde{Z}(u) = \frac{\widetilde{Z}(R)}{R^n}$$

for every $R \in (0,1)$. If $R \in (1.2/\sqrt{t}, 1)$, we can bound $\sum_{1\leq k\leq L} 20t^{k/2}R^k/k$ from above by

$$\sum_{1\leq k\leq L} \frac{20t^{k/2}}{k} R^k \leq \frac{20(R\sqrt{t})^L}{1 - 1/(R\sqrt{t})} \leq 120(r+1)R^L,$$

and the sum $\sum_{k>L} 20(r+1)R^k/k$ by

$$\sum_{k>L} 20(r+1)\frac{R^k}{k} \leq \frac{20(r+1)}{L+1}\frac{R^L}{1-R}.$$

Thus, for every $R \in (1.2/\sqrt{t}, 1)$,

$$|[u^n]\widetilde{Z}(u)| \leq \exp\Big(10(r+1)R^L(6 + \frac{1}{(L+1)(1-R)}) - n\log R\Big). \tag{5.12}$$

Assume $r \geq 200000$ and choose $R = t^{-\frac{\log\log r}{\log r}}$ in (5.12). We then have

$$-n\log R = n\frac{\log\log r}{\log r}\log t, \qquad R^L \leq t^{-\frac{\log\log r}{\log r}(2\log_t r - 1)} = \frac{t^{\frac{\log\log r}{\log r}}}{(\log r)^2} \leq \frac{t}{(\log r)^2}. \tag{5.13}$$

Assuming further $r \geq t^{\log^2 t}$, we have $\frac{\log t \log\log r}{\log r} \in (0,1)$, which implies $R \leq 1 - \log t \log\log r/(2\log r)$, and so

$$\frac{1}{(L+1)(1-R)} \leq \frac{\log t}{2\log r}\frac{2\log r}{\log t \log\log r} \leq 1. \tag{5.14}$$

Plugging (5.13) and (5.14) in (5.12), we obtain (5.10). To prove (5.11), use (5.10) if $r \geq \max\{200000, t^{\log^2 t}, \sqrt{n}\}$, and otherwise use (5.9) together with the bound $\binom{n+k}{n} \leq (n+k)^{\min\{n,k\}}$. □

5.3 Proof of Theorem 5.1

By (5.2), it suffices to bound

$$\sum_{\substack{\omega\in\Omega\\ |\omega|=n}} |S(n,\chi\cdot\mathbf{1}_\omega)|,$$

where $\mathbf{1}_\omega$ is the indicator function of polynomials f with $\omega_f=\omega$ (see (2.14) for the definition of S). Let

$$\Omega_d\subseteq\Omega$$

be the subset of factorization types containing only pairs $(x,y)\in\mathbb{N}^2$ with $x=d$. The elements of Ω_d, for each d, may be parametrized by partitions – to each $\lambda=(\lambda_1,\lambda_2,\dots,\lambda_k)\in\mathbb{Y}$ we associate

$$\omega_{\lambda,d}:=\{(d,\lambda_1),\dots,(d,\lambda_k)\}\in\Omega_d.$$

Note that $|\omega_{\lambda,d}|=d|\lambda|$. By Lemma 5.2, $\sum_{\omega\in\Omega,\,|\omega|=n}|S(n,\chi\cdot\mathbf{1}_\omega)|$ is the coefficient of u^n in the following power series:

$$F_\chi(u):=\prod_{d\geq1}F_d(u^d)$$

where

$$F_d(u):=\sum_{\lambda\in\mathbb{Y}}\big|S(d|\lambda|,\chi\cdot\mathbf{1}_{\omega_{\lambda,d}})\big|\,u^{|\lambda|}.$$

The terms with $\ell(\lambda)>|\mathcal{P}_{d,q}|$ do not contribute to $F_d(u)$, as there is no factorization type with more than $|\mathcal{P}_{d,q}|$ distinct primes of degree d. By Lemma 5.3 and (5.3), for each $\lambda\in\mathbb{Y}$ and $d\geq1$ we have

$$S(d|\lambda|,\chi\cdot\mathbf{1}_{\omega_{\lambda,d}})=\sum_{\mu\vdash|\lambda|}c_{\lambda,\mu}p_\mu(\chi(P):P\in\mathcal{P}_{d,q}). \tag{5.15}$$

Let $\mathrm{ord}(\chi)\geq2$ be the order of χ. Writing μ as $(\mu_1,\mu_2,\dots,\mu_r)$, we may bound $p_\mu(\chi(P):P\in\mathcal{P}_{d,q})$ using Lemma 2.1 as follows:

$$\begin{aligned}|p_\mu(\chi(P):P\in\mathcal{P}_{d,q})|&=\prod_{i=1}^{r}|p_{\mu_i}(\chi(P):P\in\mathcal{P}_{d,q})|\\&\leq\prod_{\substack{1\leq i\leq r\\ \mathrm{ord}(\chi)\nmid\mu_i}}\left(\min\{\frac{q^{\frac{d}{2}}}{d}(\ell+1),\frac{q^d}{d}\}\right)\prod_{\substack{1\leq i\leq r\\ \mathrm{ord}(\chi)|\mu_i}}\left(\frac{q^d}{d}\right).\end{aligned} \tag{5.16}$$

From (5.15), (5.16), (5.8) and (5.5), we have for all $n\geq1$

$$\begin{aligned}\sum_{\substack{\lambda\vdash n\\ \ell(\lambda)\leq|\mathcal{P}_{d,q}|}}\big|S(d|\lambda|,\chi\cdot\mathbf{1}_{\omega_{\lambda,d}})\big|&\leq\sum_{\substack{\lambda,\mu\vdash n\\ \ell(\lambda)\leq|\mathcal{P}_{d,q}|}}|c_{\lambda,\mu}|\prod_{\substack{1\leq i\leq\ell(\mu)\\ \mathrm{ord}(\chi)\nmid\mu_i}}\left(\min\{\frac{q^{\frac{d}{2}}}{d}(\ell+1),\frac{q^d}{d}\}\right)\prod_{\substack{1\leq i\leq\ell(\mu)\\ \mathrm{ord}(\chi)|\mu_i}}\left(\frac{q^d}{d}\right)\\&=\mathbb{E}_{\pi\in S_n}f(\pi)\Big(\sum_{\substack{\lambda\vdash n\\ \ell(\lambda)\leq|\mathcal{P}_{d,q}|}}w(B_{\lambda,\mu_\pi})\Big),\end{aligned} \tag{5.17}$$

where f is defined as in (5.8) with $m:=\mathrm{ord}(\chi)$ and

$$z_1=z_{1,d}:=\min\{\frac{q^{\frac{d}{2}}}{d}(\ell+1),\frac{q^d}{d}\},\qquad z_2=z_{2,d}:=\frac{q^d}{d},$$

and μ_π is the partition of n whose parts are the cycle sizes of π. By Lemma 5.4 and (5.17), the coefficients of F_d are bounded from above by the coefficients of

$$G_d(u) := 1 + \sum_{n\geq 1} \Big(\sum_{i=0}^{|\mathcal{P}_{d,q}|} \binom{n}{i} \Big) \mathbb{E}_{\pi\in S_n} f(\pi) u^n.$$

If $q \geq 4$, we do the following. Replacing $\sum_{i=0}^{|\mathcal{P}_q|} \binom{n}{i}$ with 2^n, we use Lemma 5.5 to bound the coefficients of G_d from above by the coefficients of

$$H_d(u) := (1-2u)^{-z_{1,d}} (1-(2u)^{\mathrm{ord}(\chi)})^{(-z_{2,d}+z_{1,d})/\mathrm{ord}(\chi)},$$

and so the coefficients of F_χ are bounded from above by the coefficients of

$$H_\chi(u) := \prod_{d\geq 1} H_d(u^d).$$

Summarizing,

$$\begin{aligned} \sum_{\omega\in\Omega,\, |\omega|=n} |S(n, \chi\cdot \mathbf{1}_\omega)| &= [u^n] F_\chi(u) \leq [u^n] H_\chi(u) \\ &= [u^n] \prod_{d\geq 1} (1-2u^d)^{-z_{1,d}} (1-(2u)^{d\,\mathrm{ord}(\chi)})^{(-z_{2,d}+z_{1,d})/\mathrm{ord}(\chi)}. \end{aligned} \tag{5.18}$$

The logarithm of the power series H_χ is given by

$$\log H_\chi(u) = \sum_{i,d\geq 1} \frac{2^i u^{di}}{i} z_{1,d} + \sum_{i,d\geq 1} \frac{2^{i\mathrm{ord}(\chi)} u^{di\,\mathrm{ord}(\chi)}}{i} \frac{z_{2,d}-z_{1,d}}{\mathrm{ord}(\chi)},$$

so that

$$[u^k] \log H_\chi(u) = \frac{1}{k} \sum_{di=k} 2^i (z_{1,d} d) + \frac{\mathbf{1}_{\mathrm{ord}(\chi)|k}}{k} \cdot \sum_{di=\frac{k}{\mathrm{ord}(\chi)}} 2^{i\mathrm{ord}(\chi)} i (z_{2,d}-z_{1,d}) d.$$

Set

$$L := \lfloor 2\log_q(\ell+1) \rfloor.$$

For $d \leq L$, we have $z_{1,d} = z_{2,d} = q^d/d$, so that by the first part of Lemma 5.6 with $x = q$,

$$[u^k] \log H_\chi(u) \leq \frac{1}{k} \sum_{di=k} 2^i q^d \leq \frac{8q^k}{k}.$$

for all $1 \leq k \leq L$. As $z_{2,d} - z_{1,d} \leq q^d/d$ and $z_{1,d} \leq q^{\frac{d}{2}}(\ell+1)/d$, we also have, by the first part of Lemma 5.6 with $x = \sqrt{q}$ and the second part of the lemma with $x = q$, that

$$\begin{aligned} [u^k] \log H_\chi(u) &\leq \frac{(\ell+1)}{k} \sum_{di=k} 2^i q^{\frac{d}{2}} + \frac{\mathbf{1}_{\mathrm{ord}(\chi)|k}}{k} \cdot \sum_{di=\frac{k}{\mathrm{ord}(\chi)}} 2^{i\mathrm{ord}(\chi)} q^d \\ &\leq 8(\ell+1)\frac{q^{\frac{k}{2}}}{k} + 10\frac{q^{\frac{k}{2}}}{k} \leq 10(\ell+2)\frac{q^{\frac{k}{2}}}{k} \end{aligned} \tag{5.19}$$

for all $k \geq 1$. From (5.18) and (5.19), we have

$$\sum_{\omega\in\Omega,\,|\omega|=n} |S(n,\chi\cdot\mathbf{1}_\omega)| \leq [u^n]\exp\Big(\sum_{1\leq k\leq L}\frac{10q^k}{k}u^k + \sum_{k>L}10(\ell+2)\frac{q^{\frac{k}{2}}}{k}u^k\Big),$$

and by Lemma 5.9 with $t=q$, $r=\ell+1$, we establish the theorem for $q\geq 4$. We now suppose $q\in\{2,3\}$. We define $\widetilde{H_d} := H_d$ for $d\geq 2$, while for $d=1$

$$\widetilde{H_1}(u) := 1+\sum_{n\geq 1} 1.4^n \mathbb{E}_{\pi\in S_n} f(\pi)u^n = (1-1.4u)^{-z_{1,1}}(1-(1.4u)^{\mathrm{ord}(\chi)})^{(-z_{2,1}+z_{1,1})/\mathrm{ord}(\chi)},$$

where in the last passage we used Lemma 5.5. As $|\mathcal{P}_{1,q}| = q \leq 3$ for $q\in\{2,3\}$, it follows that $\sum_{i=0}^{|\mathcal{P}_{1,q}|}\binom{n}{i} \leq \sum_{i=0}^{3}\binom{n}{i}$, which is at most $7\cdot 1.4^n$ by Lemma 5.8. Thus the coefficients of G_1 are bounded from above by those of $\widetilde{H_1}$ times 7, and so the coefficients of F_χ are bounded from above by those of

$$\widetilde{H_\chi}(u) := \prod_{d\geq 1}\widetilde{H_d}(u^d),$$

times 7. As in the case $q\geq 4$, we proceed to upper bound the coefficients of $\log\widetilde{H_\chi}$. For $d\leq L$, we have $z_{1,d}=z_{2,d}=q^d/d$, so that by Lemma 5.7 with $x=q$,

$$[u^k]\log\widetilde{H_\chi}(u) \leq \frac{1}{k}\Big(1.4^k q + \sum_{di=k,\,d\neq 1} 2^i q^d\Big) \leq 10\frac{q^k}{k},$$

for all $1\leq k\leq L$, where we used $1.4<q\leq 3$. For any $k\geq 1$ we have by Lemma 5.7 that

$$\begin{aligned}[u^k]\log\widetilde{H_\chi}(u) &\leq \frac{(\ell+1)}{k}\Big(1.4^k q^{\frac12} + \sum_{di=k,\,d\neq 1} 2^i q^{\frac{d}{2}}\Big) + \frac{\mathbf{1}_{\mathrm{ord}(\chi)|k}}{k}\cdot\Big(\sum_{di=\frac{k}{\mathrm{ord}(\chi)},\,d\neq 1} 2^{i\mathrm{ord}(\chi)}q^d + 1.4^k q\Big)\\ &\leq \frac{(\ell+1)}{k}\Big(2\cdot 1.4^k + 7q^{\frac{k}{2}}\Big) + 14\frac{q^{\frac{k}{2}}}{k} \leq 20(\ell+2)\frac{q^{\frac{k}{2}}}{k},\end{aligned}$$

From this point we continue as in the case $q\geq 4$ and conclude the proof of the theorem. □

6 The variance of sums of two squares in short intervals

6.1 Outline of proof

In this section we write $[u^n]f$ for the coefficient of u^n in a power series f, and χ_2 for the unique non-principal quadratic Dirichlet character modulo T.

We have discussed short interval characters $\chi\in G(R_\ell)$ and their L-functions. For such characters, $\chi\cdot\chi_2$ is a Dirichlet character ramified at ∞ and at T. Its L-function is defined as one would expect: $L(u,\chi\cdot\chi_2) = \sum_{f\in\mathcal{M}_q}\chi(f)u^{\deg(f)}$. If $\chi\in G(R_\ell)$ then $\chi(f)$ depends only on the first ℓ coefficients of f, we find that, for $n\geq\ell+1$,

$$\sum_{f\in\mathcal{M}_{n,q}}\chi(f)\chi_2(f) = \sum_{c\in\mathbb{F}_q^\times}\sum_{f\in\mathcal{M}_{n,q},\,f(0)=c}\chi(f)\chi_2(T+c) = \sum_{c\in\mathbb{F}_q^\times}\chi_2(T+c)\sum_{f\in\mathcal{M}_{n-1,q}}\chi(f) = 0,$$

so that $L(u, \chi \cdot \chi_2)$ is a polynomial of degree at most ℓ. If χ is primitive then the degree is exactly ℓ. Weil's RH applied to $L(u, \chi)$, $L(u, \chi \cdot \chi_2)$ implies

$$[u^n]L(u,\chi),\ [u^n]L(u,\chi\cdot\chi_2) = O_{n,\ell}(q^{\frac{n}{2}}). \tag{6.1}$$

Given a short interval character χ, the following lemma relates $\sum_{f\in\mathcal{M}_{n,q}} b(f)\chi(f)$ to the L-functions of χ and $\chi\cdot\chi_2$. It is proved in §6.2.

Lemma 6.1. *Let $\chi \in G(R_\ell)$. Then*

$$\sum_{f\in\mathcal{M}_{n,q}} b(f)\chi(f) = [u^n]\Big(\sqrt{L(u,\chi)L(u,\chi\cdot\chi_2)}\prod_{i\geq 1}\left(\frac{L(u^{2^i},\chi^{2^i})}{L(u^{2^i},\chi^{2^i}\cdot\chi_2)}\right)^{2^{-i-1}} \times (1-\chi(T)u)^{-1/2}\prod_{i\geq 1}(1-\chi^{2^i}(T)u^{2^i})^{2^{-i-1}}\Big). \tag{6.2}$$

(The roots in the right-hand side of (6.2) *are chosen so that the constant terms remain 1.)*

In the large-q limit, this identity can be significantly simplified, as the following lemma shows, whose proof is given in §6.3.

Lemma 6.2. *Let $\chi_0 \neq \chi \in G(R_\ell)$. Then*

$$\sum_{f\in\mathcal{M}_{n,q}} b(f)\chi(f) = [u^n]\sqrt{L(u,\chi)L(u,\chi\cdot\chi_2)} + O_n(q^{\frac{n}{2}-\frac{1}{4}}), \tag{6.3}$$

$$\sum_{f\in\mathcal{M}_{n,q}} b(f)\chi(f) = O_n(q^{\frac{n}{2}}). \tag{6.4}$$

For $\chi \in G(R_N)\backslash G(R_{N-1})$, we factorize $L(u,\chi) = \prod_{i=1}^{N-1}(1-\gamma_i(\chi)u)$ and $L(u,\chi\cdot\chi_2) = \prod_{i=1}^{N}(1-\gamma_i(\chi\cdot\chi_2)u)$, and define $\Theta_\chi \in U(N-1)$ to be a unitary matrix with eigenvalues $\gamma_i(\chi)/\sqrt{q}$ and $\Theta_{\chi\cdot\chi_2} \in U(N)$ to be a unitary matrix with eigenvalues $\gamma_i(\chi\cdot\chi_2)/\sqrt{q}$. In the following proposition, proved in §6.4, we express $\mathrm{Var}_b(n,h)$ in terms of Θ_χ and $\Theta_{\chi\cdot\chi_2}$.

Proposition 6.3. *Let $-1 \leq h \leq n-1$. We have*

$$\frac{\mathrm{Var}_b(n,h)}{q^{h+1}} = q^{-(n-h-1)}\sum_{\chi\in G(R_{n-h-1})\backslash G(R_{n-h-2})}\left|[u^n]\sqrt{\det(I-u\Theta_\chi)\cdot\det(I-u\Theta_{\chi\cdot\chi_2})}\right|^2 + O_n\left(q^{-\frac{1}{4}}\right). \tag{6.5}$$

We are able to evaluate the integral in (6.5) by making use of a recent equidistribution theorem of Sawin [Saw18]. We adopt the following notation: for a continuous class function $f\colon U(N-1)\times U(N)\to\mathbb{C}$, where $U(n)$ is the $n\times n$ unitary group, we define the function $\langle f\rangle\colon U(1)\times U(1)\to\mathbb{C}$ as the unique continuous map such that

$$\int_{U(N-1)\times U(N)} f(g_1,g_2)\psi(\det g_1,\det g_2)\,dg_1\,dg_2 = \int_{U(N-1)\times U(N)}\langle f\rangle(\det g_1,\det g_2)\psi(\det g_1,\det g_2)\,dg_1\,dg_2,$$

for all continuous functions $\psi\colon U(1)\times U(1)\to\mathbb{C}$. That is, $\langle f\rangle(c_1,c_2)$ is the integral of f over the coset of $SU(N-1)\times SU(N)\le U(N-1)\times U(N)$ consisting of elements with determinants c_1,c_2, against the unique $SU(N-1)\times SU(N)$-invariant measure on that coset, of total mass 1. We can now state the special case of Sawin's result that we need.

Theorem 6.4. *[Saw18, Thm. 1.2] If $f\colon U(N-1)\times U(N)$ is a continuous class function and $N\ge 6$, then*

$$\lim_{q\to\infty}\left[\frac{1}{q^N}\sum_{\chi\in G(R_N)\backslash G(R_{N-1})} f(\Theta_\chi,\Theta_{\chi\cdot\chi_2})-\frac{1}{q^N}\sum_{\chi\in G(R_N)\backslash G(R_{N-1})}\langle f\rangle(\det\Theta_\chi,\det\Theta_{\chi\cdot\chi_2})\right]=0$$

where the limit is taken for q of fixed characteristic.

We introduce the notation, for a unitary matrix g,

$$A_{k,(z)}(g):=[u^k]\det(1-ug)^z. \tag{6.6}$$

Note that $A_{k,(z)}(g)$ is a symmetric homogeneous polynomial of degree k in the eigenvalues of g. Because we will make use especially of the case $z=1/2$, we introduce the abbreviation $A_k(g):=A_{k,(1/2)}(g)$. Theorem 6.4 allows us to deduce the following corollary, proved in §6.5.

Corollary 6.5. *Fix constants $-1\le h\le n-1$ and let $N=n-h-1$. For $n\le N(N-1)$ and $N\ge 6$,*

$$\begin{aligned}\lim_{q\to\infty} q^{-(n-h-1)}\sum_{\chi\in G(R_{n-h-1})\backslash G(R_{n-h-2})}&\left|[u^n]\sqrt{\det(1-u\Theta_\chi)\det(1-u\Theta_{\chi\cdot\chi_2})}\right|^2\\ &=\sum_{\substack{j+k=n\\ j,k\ge 0}}\int_{U(N-1)}|A_j(g_1)|^2\,dg_1\int_{U(N)}|A_k(g_2)|^2\,dg_2,\end{aligned} \tag{6.7}$$

with the limit taken along a sequence of q of fixed characteristic.

In order to give a succinct evaluation of the integrals on the right-hand side of Corollary 6.7, we make use of z-measures on partitions, first introduced by Kerov, Olshanski, and Vershik [KOV93]. We briefly survey them in §6.6. The following theorem evaluates the random matrix integrals appearing in Corollary 6.5. It is proved in §6.7.

Theorem 6.6. *For $g\in U(N)$, with $A_{n,(z)}(g)$ defined by (6.6), we have*

$$\int_{U(N)}A_{n,(z)}(g)A_{n,(z')}(g^{-1})\,dg=\frac{(zz')_n}{n!}\mathbb{P}^{(n)}_{z,z'}(\lambda_1\le N). \tag{6.8}$$

In §6.8 we quickly prove Theorem 1.7 from Proposition 6.3, Corollary 6.5 and Theorem 6.6. In §6.9 we prove Proposition 1.8 from basic properties of z-measures.

6.2 Proof of Lemma 6.1

The lemma is equivalent to the following formal identity:

$$\begin{aligned}\sum_{f\in\mathcal{M}_q}b(f)\chi(f)u^{\deg f}=\sqrt{L(u,\chi)L(u,\chi\cdot\chi_2)}\prod_{i\ge 1}\left(\frac{L(u^{2^i},\chi^{2^i})}{L(u^{2^i},\chi^{2^i}\cdot\chi_2)}\right)^{2^{-i-1}}\\ \times(1-\chi(T)u)^{-1/2}\prod_{i\ge 1}(1-\chi^{2^i}(T)u^{2^i})^{2^{-i-1}}.\end{aligned} \tag{6.9}$$

We verify (6.9) by comparing the Euler product of both sides. By [BSSW16, Prop. 2.4], the function b is multiplicative (that is, $b(fg) = b(f)b(g)$ for coprime $f, g \in \mathcal{M}_q$), and at prime powers we have

$$b(P^k) = \begin{cases} 1 & \text{if } 2 \mid k \text{ or } \chi_2(P) \in \{0, 1\}, \\ 0 & \text{otherwise.} \end{cases} \tag{6.10}$$

Since $b \cdot \chi$ is multiplicative, (6.10) implies that the left-hand side of (6.9) factors as

$$\prod_{P:\chi_2(P)=1} (1 - \chi(P)u^{\deg P})^{-1} \prod_{Q:\chi_2(Q)=-1} (1 - \chi(Q^2)u^{2\deg Q})^{-1}(1 - \chi(T)u^{\deg T})^{-1}, \tag{6.11}$$

where P, Q denote monic irreducible polynomials. We have

$$\begin{aligned} L(u, \chi) &= \prod_{P:\chi_2(P)=1} (1 - \chi(P)u^{\deg P})^{-1} \prod_{Q:\chi_2(Q)=-1} (1 - \chi(Q)u^{\deg Q})^{-1}(1 - \chi(T)u)^{-1}, \\ L(u, \chi \cdot \chi_2) &= \prod_{P:\chi_2(P)=1} (1 - \chi(P)u^{\deg P})^{-1} \prod_{Q:\chi_2(Q)=-1} (1 + \chi(Q)u^{\deg Q})^{-1}. \end{aligned} \tag{6.12}$$

In particular, (6.12) implies that

$$\frac{L(u, \chi)}{L(u, \chi \cdot \chi_2)} = \prod_{Q:\chi_2(Q)=-1} \frac{(1 - \chi(Q)u^{\deg Q})^{-1}}{(1 + \chi(Q)u^{\deg Q})^{-1}}(1 - \chi(T)u)^{-1} \tag{6.13}$$

and that

$$\begin{aligned} \sqrt{L(u, \chi)L(u, \chi \cdot \chi_2)} = \prod_{P:\chi_2(P)=1} (1 - \chi(P)u^{\deg P})^{-1} \prod_{Q:\chi_2(Q)=-1} (1 - \chi^2(Q)u^{2\deg Q})^{-1/2} \\ \times (1 - \chi(T)u)^{-1/2}. \end{aligned} \tag{6.14}$$

Using (6.12), (6.13) and (6.14), we find that the right-hand side of (6.9) factors as

$$\begin{aligned} &\prod_{P:\chi_2(P)=1} (1 - \chi(P)u^{\deg P})^{-1} \cdot \prod_{Q:\chi_2(Q)=-1} (1 - \chi^2(Q)u^{2\deg Q})^{-1/2} \\ &\quad \cdot \prod_{i\geq 1} \prod_{Q:\chi_2(Q)=-1} \left(\frac{(1 - \chi^{2^i}(Q)u^{2^i \deg Q})^{-1}}{(1 + \chi^{2^i}(Q)u^{2^i \deg Q})^{-1}} \right)^{2^{-i-1}} \\ &\quad \cdot (1 - \chi(T)u)^{-1/2} \prod_{i\geq 1} (1 - \chi^{2^i}(T)u^{2^i})^{-2^{-i-1}} \\ &\quad \cdot (1 - \chi(T)u)^{-1/2} \prod_{i\geq 1} (1 - \chi^{2^i}(T)u^{2^i})^{2^{-i-1}} \\ &\quad = \prod_{P:\chi_2(P)=1} (1 - \chi(P)u^{\deg P})^{-1} \cdot \prod_{Q:\chi_2(Q)=-1} (1 - \chi^2(Q)u^{2\deg Q})^{-1/2} \\ &\quad \cdot \prod_{i\geq 1} \prod_{Q:\chi_2(Q)=-1} \left(\frac{(1 - \chi^{2^i}(Q)u^{2^i \deg Q})^{-1}}{(1 + \chi^{2^i}(Q)u^{2^i \deg Q})^{-1}} \right)^{2^{-i-1}} (1 - \chi(T)u)^{-1}. \end{aligned} \tag{6.15}$$

It remains to establish equality between the Euler products (6.11) and (6.15). The contribution of the prime T is the same in both, and so is the contribution of primes P satisfying $\chi_2(P) = 1$. Now let Q be a prime satisfying $\chi_2(Q) = -1$. It is sufficient to prove that the contribution of this prime in both products is the same, that is

$$(1-\chi(Q^2)u^{2\deg Q})^{-1} = (1-\chi^2(Q)u^{2\deg Q})^{-1/2}\prod_{i\geq 1}\left(\frac{(1-\chi^{2^i}(Q)u^{2^i\deg Q})^{-1}}{(1+\chi^{2^i}(Q)u^{2^i\deg Q})^{-1}}\right)^{2^{-i-1}}. \quad (6.16)$$

Letting $z = \chi^2(Q)u^{2\deg Q}$, the identity (6.16) becomes

$$(1-z)^{-1/2} = \prod_{i\geq 1}\left(\frac{1+z^{2^{i-1}}}{1-z^{2^{i-1}}}\right)^{2^{-i-1}}, \quad (6.17)$$

which follows by noting the telescoping nature of the right-hand side of (6.17):

$$\prod_{i\geq 1}\left(\frac{1+z^{2^{i-1}}}{1-z^{2^{i-1}}}\right)^{2^{-i-1}} = \prod_{i\geq 1}\frac{(1-z^{2^i})^{2^{-i-1}}}{(1-z^{2^{i-1}})^{2^{-i}}} = (1-z)^{-1/2}.$$

□

6.3 Proof of Lemma 6.2

By Lemma 6.1,

$$\sum_{f\in\mathcal{M}_{n,q}} b(f)\chi(f) = [u^n]\sqrt{L(u,\chi)L(u,\chi\cdot\chi_2)} \times\prod_{i\geq 1}\left(\frac{L(u^{2^i},\chi^{2^i})}{L(u^{2^i},\chi^{2^i}\cdot\chi_2)}\right)^{2^{-i-1}}(1-\chi(T)u)^{-1/2}\prod_{i\geq 1}(1-\chi^{2^i}(T)u^{2^i})^{2^{-i-1}}. \quad (6.18)$$

Although the products in (6.18) are infinite, we may truncate them because only the coefficient of u^n is of interest to us:

$$\sum_{f\in\mathcal{M}_{n,q}} b(f)\chi(f) = [u^n]\sqrt{L(u,\chi)L(u,\chi\cdot\chi_2)} \times\prod_{i=1}^{n}\left(\frac{L(u^{2^i},\chi^{2^i})}{L(u^{2^i},\chi^{2^i}\cdot\chi_2)}\right)^{2^{-i-1}}(1-\chi(T)u)^{-1/2}\prod_{i=1}^{n}(1-\chi^{2^i}(T)u^{2^i})^{2^{-i-1}}. \quad (6.19)$$

For any $i \geq 1$, the character χ^{2^i} is non-trivial, since the order of χ is odd (it divides $|G(R_\ell)| = q^\ell$). Hence, by making use of (6.1) with χ^{2^i} and $\chi^{2^i}\cdot\chi_2$, we see that the j-th coefficients of $L(u,\chi^{2^i})$ and of $L(u,\chi^{2^i}\cdot\chi_2)$ are both of size $O_{j,n}(q^{j/2})$. In particular, for any $i \geq 1$,

$$[u^j]L(u^{2^i},\chi^{2^i}), [u^j]L(u^{2^i},\chi^{2^i}\cdot\chi_2) = O_{j,n}(q^{j/2^{i+1}}) = O_{j,n}(q^{j/4}). \quad (6.20)$$

From (6.20) we deduce that

$$[u^j]\prod_{i=1}^{n}\left(\frac{L(u^{2^i},\chi^{2^i})}{L(u^{2^i},\chi^{2^i}\cdot\chi_2)}\right)^{2^{-i-1}}(1-\chi(T)u)^{-1/2}\prod_{i=1}^{n}(1-\chi^{2^i}(T)u^{2^i})^{2^{-i-1}} = O_{j,n}(q^{j/4}). \quad (6.21)$$

Additionally, from (6.1),

$$[u^j]\sqrt{L(u,\chi)L(u,\chi\cdot\chi_2)} = O_{j,n}(q^{j/2}). \tag{6.22}$$

Plugging the estimates (6.21) and (6.22) in (6.19), we establish (6.3). From (6.3) and (6.22) with $j = n$, we obtain (6.4). □

6.4 Proof of Proposition 6.3

By Lemma 6.2,

$$\sum_{f\in\mathcal{M}_{n,q}} b(f)\chi(f) = [u^n]\sqrt{L(u,\chi)L(u,\chi\cdot\chi_2)} + O_n(q^{\frac{n}{2}-\frac{1}{4}}) = O_n(q^{\frac{n}{2}}). \tag{6.23}$$

From (6.23) and (2.16) we obtain

$$\mathrm{Var}_b(n,h) = \frac{\sum_{\chi_0\neq\chi\in G(R_{n-h-1})}\left|[u^n]\sqrt{L(u,\chi)L(u,\chi\cdot\chi_2)} + O_n(q^{\frac{n}{2}-\frac{1}{4}})\right|^2}{q^{2(n-h-1)}}. \tag{6.24}$$

Since the number of characters in $G(R_{n-h-1})$ which are in $G(R_{n-h-2})$ as well is $O(q^{n-h-2})$, (6.23) and (6.24) imply that

$$\begin{aligned}\mathrm{Var}_b(n,h) &= \frac{\sum_{\chi_0\neq\chi\in G(R_{n-h-1})}\left|[u^n]\sqrt{L(u,\chi)L(u,\chi\cdot\chi_2)} + O_n(q^{\frac{n}{2}-\frac{1}{4}})\right|^2}{q^{2(n-h-1)}} + O_n(q^h)\\ &= \frac{\sum_{\chi_0\neq\chi\in G(R_{n-h-1})\setminus G(R_{n-h-2})}\left|[u^n]\sqrt{L(u,\chi)L(u,\chi\cdot\chi_2)}\right|^2}{q^{2(n-h-1)}} + O_n(q^{h-\frac{1}{4}}).\end{aligned}$$

We now write $L(u,\chi)$ as $\det(I-u\sqrt{q}\Theta_\chi)(1-u)$ and $L(u,\chi\cdot\chi_2)$ as $\det(I-u\sqrt{q}\Theta_{\chi\cdot\chi_2})$ to obtain

$$\mathrm{Var}_b(n,h) = \frac{\sum_{\chi\in G(R_{n-h-1})\setminus G(R_{n-h-2})}\left|[u^n]\sqrt{\det(I-u\sqrt{q}\Theta_\chi)\det(I-u\sqrt{q}\Theta_\chi)(1-u)}\right|^2}{q^{2(n-h-1)}} + O_n(q^{h-\frac{1}{4}}). \tag{6.25}$$

The proof is concluded by writing in (6.25) $[u^n]\sqrt{\det(I-u\sqrt{q}\Theta_\chi)\det(I-u\sqrt{q}\Theta_\chi)(1-u)} = q^{n/2}[u^n]\sqrt{\det(I-u\Theta_\chi)\det(I-u\Theta_{\chi\cdot\chi_2})} + O_n(q^{(n-1)/2})$ and dividing both sides by q^{h+1}. □

6.5 Proof of Corollary 6.5

We begin with the following lemma.

Lemma 6.7. *Let $f\colon U(N)\to\mathbb{C}$ be a function such that for $g\in U(N)$, $f(g)$ is a symmetric homogeneous Laurent polynomial of degree k in the eigenvalues of g. The following hold.*

1. *If $k\not\equiv 0 \bmod N$, then*
$$\langle f\rangle(z) = 0, \quad \textit{for all } |z| = 1,$$

2. If $k = 0$, then

$$\langle f \rangle(z) = \int_{U(N)} f(g)\, dg, \quad \textit{for all } |z| = 1.$$

Proof. Both i) and ii) make use of the following assertion: that if $F\colon U(N) \to \mathbb{C}$ is a symmetric homogeneous polynomial of degree $k \neq 0$ in the eigenvalues of a matrix from $U(N)$, then

$$\int_{U(N)} F(g)\, dg = 0. \tag{6.26}$$

For, the Haar measure is invariant under scalar multiplication, so for any $c \in U(1)$,

$$0 = \int_{U(N)} F(g)\, dg - \int_{U(N)} F(cg)\, dg = (1 - c^k) \int_{U(N)} F(g)\, dg.$$

If $k \neq 0$, there exists $c \in U(1)$ such that $(1 - c^k) \neq 0$ and (6.26) follows.

Turning to i), note that this will be proved if we show for $k \not\equiv 0 \bmod N$ that

$$\int_{U(N)} f(g)\psi(\det g)\, dg = 0, \tag{6.27}$$

for all continuous $\psi\colon U(1) \to \mathbb{C}$. In turn by Fourier analysis, since $\det g \in U(1)$ for all $g \in U(N)$, to establish (6.27) we need only establish it for $\psi(z) = z^\ell$ with $\ell \in \mathbb{Z}$. But if $f(g)$ is of degree k in the eigenvalues of g, then

$$f(g)(\det g)^\ell$$

is of degree $k + N\ell$. As $k \not\equiv 0 \bmod N$, we have $k + N\ell \neq 0$, and hence

$$\int_{U(N)} f(g)(\det g)^\ell\, dg = 0,$$

establishing the claim i). For ii), our proof is similar. We must show

$$\int_{U(N)} f(g)\psi(\det g)\, dg = \int_{U(N)} f(g)\, dg \int_{U(N)} \psi(\det g)\, dg. \tag{6.28}$$

As before it suffices to verify this claim when $\psi(z) = z^\ell$. For $\ell = 0$ this is clear, and when $\ell \neq 0$, note that

$$\int_{U(N)} (\det g)^\ell\, dg = 0,$$

so that we establish (6.28) by showing

$$\int_{U(N)} f(g)(\det g)^\ell\, dg = 0.$$

But as $f(g)(\det g)^\ell$ is of degree $\ell \neq 0$, this is indeed the case, establishing the claim. □

Note that

$$[u^n] \sqrt{\det(1 - u\Theta_\chi) \det(1 - u\Theta_{\chi\cdot\chi_2})} = \sum_{\substack{j+k=n \\ j,k \geq 0}} A_j(\Theta_\chi) A_k(\Theta_{\chi\cdot\chi_2}).$$

Hence the left-hand side of (6.7) is

$$\lim_{q\to\infty} \frac{1}{q^{n-h-1}} \sum_{\chi\in G(R_{n-h-1})\backslash G(R_{n-h-2})} \sum_{\substack{j+k=n\\ j,k\geq 0}} \sum_{\substack{j'+k'=n\\ j',k'\geq 0}} A_j(\Theta_\chi)A_k(\Theta_{\chi\cdot\chi_2})\overline{A_{j'}(\Theta_\chi)A_{k'}(\Theta_{\chi\cdot\chi_2})}. \tag{6.29}$$

We will need to evaluate the random matrix coset integral

$$\langle A_j A_k \overline{A_{j'} A_{k'}}\rangle = \langle A_j \overline{A_{j'}}\rangle \langle A_k \overline{A_{k'}}\rangle.$$

Note that if $j = j'$, then $k = k'$ also. Noting that $A_j(g)\overline{A_j(g)} = A_j(g)A_j(g^{-1})$ and likewise for A_k, one may see that $A_j\overline{A_j}$ and $A_k\overline{A_k}$ are homogeneous symmetric Laurent polynomials of degree 0. Thus by Lemma 6.7, we have for all $|z| = 1$,

$$\langle A_j\overline{A_j}\rangle(z) = \int_{U(N-1)} |A_j(g)|^2\, dg,$$

$$\langle A_k\overline{A_k}\rangle(z) = \int_{U(N)} |A_k(g)|^2\, dg.$$

Furthermore, in the sum (6.29), if $j \neq j'$ and $k \neq k'$, we may reason in the same way to see that $A_j\overline{A_{j'}}$ and $A_k\overline{A_{k'}}$ are homogeneous symmetric Laurent polynomials of non-zero degrees, say ℓ and $-\ell$ respectively, with $|\ell| \leq n < N(N-1)$. As no non-zero number smaller in magnitude than $N(N-1)$ is divisible by both N and $N-1$, Lemma 6.7 implies that one of

$$\int_{U(N-1)} A_j(g)\overline{A_{j'}(g)}\, dg = 0 \quad \text{or} \quad \int_{U(N)} A_k(g)\overline{A_{k'}(g)}\, dg = 0$$

holds, so in particular the product is always 0. From this analysis it follows that for all matrices Θ_χ and $\Theta_{\chi\cdot\chi_2}$

$$\langle A_j A_k \overline{A_{j'} A_{k'}}\rangle(\det\Theta_\chi, \det\Theta_{\chi\cdot\chi_2}) = \begin{cases} \int_{U(N-1)} |A_j(g_1)|^2 dg_1 \cdot \int_{U(N)} |A_k(g_2)|^2\, dg_2 & \text{if } j = j', \\ 0 & \text{otherwise.} \end{cases}$$

Thus using Theorem 6.4, (6.29) simplifies to

$$\sum_{\substack{j+k=n\\ j,k\geq 0}} \int_{U(N-1)} |A_j(g_1)|^2\, dg_1 \int_{U(N)} |A_k(g_2)|^2\, dg_2,$$

as claimed. □

6.6 z-measures on partitions

The z-measures are a two-parameter family of measures on partitions, though it is often natural to specialize to a one-parameter subfamily. In order to define the z-measures we make use of standard notation in enumerative combinatorics, along the lines of e.g. [Sta99, Ch. 7]. We view partitions $\lambda \vdash n$ as Young diagrams with n boxes. Recall (from e.g. [Sta99, Sec. 7.21]) that for a square $\square$ in λ with position (i, j) (where $1 \leq j \leq \lambda_i$), the *content* $c(\square)$ is defined by

$$c(\square) = j - i.$$

We let $\dim(\lambda)$ be the dimension of the irreducible representation of S_n associated to the partition λ; equivalently $\dim(\lambda)$ is equal to the number of standard Young tableaux of shape λ. The z-measure on partitions of n with parameters z and z', written $M^{(n)}_{z,z'}$ is the measure on the set of all partitions λ of n satisfying

$$M^{(n)}_{z,z'}(\lambda) := \frac{\dim(\lambda)^2}{n!(zz')_n} \prod_{\square \in \lambda} (z + c(\square))(z' + c(\square)). \tag{6.30}$$

Recall that $(x)_j := (x)(x+1)\cdots(x+j-1)$ is the Pochhammer symbol. The expression (6.30) is well-defined for all $z, z' \in \mathbb{C}$ with $zz' \notin \mathbb{Z}_{\leq 0}$. Furthermore we use the convention that $\varnothing$ is the sole partition of 0 and for any z, z',

$$M^{(0)}_{z,z'}(\varnothing) = 1.$$

For any n and $z, z' \in \mathbb{C}$ with $zz' \notin \mathbb{Z}_{\leq 0}$ one has

$$\sum_{\lambda \vdash n} M^{(n)}_{z,z'}(\lambda) = 1,$$

though this fact is not obvious (see e.g. [Oko01] for a proof). It is not always the case that $M^{(n)}_{z,z'}(\lambda) \geq 0$ for all λ (so that in some cases $M^{(n)}_{z,z'}$ must be viewed as a signed measure) but when, for instance, $z' = \overline{z}$, plainly (6.30) is always non-negative.

Note from the definition (6.30), for fixed n, this measure tends toward the Plancherel measure as $z, z' \to \infty$. In this sense, z-measures can be thought of as a generalization of Plancherel measures.

We denote $M^{(n)}_z(\lambda) := M^{(n)}_{z,\overline{z}}(\lambda)$, and moreover for a subset A of the set of all partitions of n, we use the notations

$$\mathbb{P}^{(n)}_{z,z'}(\lambda \in A) = \sum_{\lambda \in A} M^{(n)}_{z,z'}(\lambda), \quad \text{and} \quad \mathbb{P}^{(n)}_{z}(\lambda \in A) = \sum_{\lambda \in A} M^{(n)}_{z}(\lambda).$$

It is known that there exists a scaling limit of the z-measures as $n \to \infty$; these scaling limits were first investigated as a part of representation theory on the infinite symmetric group. We do not review the full theory here, instead referring the reader to [Ols03] for an introduction. The result from this theory that we will make use of is

Theorem 6.8. *For any $z \in \mathbb{C} \setminus \mathbb{Z}_{\leq 0}$, there exists a random variable $\alpha_1^{(z)}$ lying almost surely in the interval $[0,1]$ such that for $\lambda \vdash n$ chosen according to the z-measure with parameters $z, \overline{z}$ we have*

$$\lim_{n \to \infty} \mathbb{P}\Big(\frac{\lambda_1}{n} \leq x\Big) = \mathbb{P}(\alpha_1^{(z)} \leq x),$$

for all real x.

Moreover for $z \in \mathbb{C} \setminus \mathbb{Z}_{\leq 1}$ with $z' = \overline{z}$ as above, the function $F_z(x) = \mathbb{P}(\alpha_1^{(z)} \leq x)$ is continuous for all $x \in \mathbb{R}$.

We simply take this theorem as our definition of $\alpha_1^{(z)}$ – that is, $\alpha_1^{(z)}$ is the random variable with distribution function given by this limit – but we note that there exists a more sophisticated perspective in which the random variable $\alpha_1^{(z)}$ is *the largest part of the z-measure point process with parameters $z, \overline{z}$ on the Thoma simplex*; see [Bor98] for more about this latter object and its connection to the infinite symmetric group. We adopt the notational convention that $\alpha_1 = \alpha_1^{(1/2)}$.

6.7 Proof of Theorem 6.6

We use Schur functions to prove this theorem. For $x = (x_1, ..., x_N)$ and λ a partition, we use the notation $s_\lambda(x)$ to denote the Schur function of shape λ (see [Sta99, Ch. 7]).

We will use of the following well-known results:

First, we recall the *dual Cauchy identity* (see [Sta99, Thm 7.14.3]),

$$\prod_{i,j}(1 + x_i y_j) = \sum_\lambda s_\lambda(x) s_{\lambda'}(y), \tag{6.31}$$

where λ' is the dual partition to λ.

Second, we recall the following evaluation of Schur functions (proved by combining [Sta99, Cor 7.21.4] and [Sta99, Cor 7.21.6]),

$$s_\lambda(\underbrace{1, ..., 1}_{k}) = \frac{\dim(\lambda)}{n!} \prod_{\square \in \lambda}(k + c(\square)), \tag{6.32}$$

for $\lambda \vdash n$.

Third, we recall the orthogonality relations for Schur functions in eigenvalues of the unitary group (see e.g. [Bum04]). If $g \in U(N)$ has eigenvalues $y_1, ..., y_N$ and we use the Schur function notation $s_\lambda(g) := s_\lambda(y_1, ..., y_N)$, for any two partitions λ and ν,

$$\int_{U(N)} s_\lambda(g) s_\nu(g^{-1})\, dg = \delta_{\lambda=\nu, \ell(\lambda)\leq N}. \tag{6.33}$$

We start by specializing to the case where z is a positive integer; later on we will consider more general z. We make use of the dual Cauchy identity (6.31) in the variables $x_1, ..., x_z$ and $y_1, ..., y_N$ where for all i, $x_i = -u$, and $y_1, ..., y_N$ are the N eigenvalues of $g \in U(N)$. The dual Cauchy identity tells that

$$\det(1 - ug)^z = \sum_\lambda s_\lambda(-u, ..., -u) s_{\lambda'}(g) = \sum_\lambda (-u)^{|\lambda|} s_\lambda(\underbrace{1, ..., 1}_{z}) s_{\lambda'}(g).$$

Yet from (6.32), we see we can write this as

$$\det(1 - ug)^z = \sum_{n=0}^\infty u^n \Big((-1)^n \sum_{\lambda \vdash n} \frac{\dim(\lambda)}{n!} \prod_{\square \in \lambda}(z + c(\square)) s_{\lambda'}(g) \Big), \tag{6.34}$$

where we adopt the convention that the coefficient for $n = 0$ is 1. Note that we have so far only proved (6.34) for positive integer z.

For $|u| < 1$, the binomial series tells us that

$$(1 - uy_i)^z = \sum_{n=0}^\infty u^n \Big((-y_i)^n \frac{(z)_n}{n!} \Big),$$

for all complex z. In particular the coefficients of u^n in this series are polynomials in z. Multiplying N such identities, it follows that for $|u| < 1$ and all complex z,

$$\det(1 - ug)^z = \prod_{i=1}^N (1 - uy_i)^z = \sum_{n=0}^\infty u^n P_{n,g}(z),$$

where $P_{n,g}(z)$ are polynomials in z. From (6.34) we obtain the expression

$$P_{n,g}(z) = (-1)^n \sum_{\lambda \vdash n} \frac{\dim(\lambda)}{n!} \prod_{\square \in \lambda} (z + c(\square)) s_{\lambda'}(g), \tag{6.35}$$

valid for positive integer z. But as both the left and right-hand sides are polynomials in z equal at all positive integers, it follows that this identity holds for all $z \in \mathbb{C}$.

But of course, $P_{n,g}(z) = A_{n,(z)}(g)$, so that using (6.35) and orthogonality relations (6.33) for Schur functions,

$$\begin{aligned}\int_{U(N)} A_{n,(z)}(g) A_{n,(z')}(g^{-1})\, dg &= \sum_{\lambda \vdash n} \frac{(\dim \lambda)^2}{(n!)^2} \prod_{\square \in \lambda} (z + c(\square))(z' + c(\square)) \delta_{\ell(\lambda') \le N} \\ &= \frac{(zz')_n}{n!} \mathbb{P}^{(n)}_{z,z'}(\lambda_1 \le N).\end{aligned}$$

This verifies (6.8). □

6.8 Proof of Theorem 1.7

We note from Proposition 6.3 and Corollary 6.5 that for $N = n - h - 1$,

$$\mathrm{Var}_b(n, h) = q^{h+1} \sum_{\substack{j+k=n \\ j,k \ge 0}} \int_{U(N-1)} |A_j(g_1)|^2\, dg_1 \int_{U(N)} |A_k(g_2)|^2\, dg_2 + o(q^{h+1}),$$

for $n \le N(N-1)$ and $0 \le h \le n - 7$ (the upper bound restriction comes from requiring that $N \ge 6$ in Corollary 6.5). But then these integrals are evaluated using Theorem 6.6 with $z = z' = 1/2$, and the result is Theorem 1.7. □

6.9 Proof of Proposition 1.8

By Stirling's approximation, we have

$$\frac{(1/4)_j}{j!} = \frac{j^{-3/4}}{\Gamma(1/4)} + o(j^{-3/4})$$

as $j \to \infty$. In general, $\frac{(1/4)_j}{j!} = O(j^{-3/4})$ for all $j \ge 1$. By Theorem 6.8, we have

$$\mathbb{P}^{(j)}_{1/2}(\lambda_1 \le N) = \mathbb{P}(\alpha_1 \le N/j) + o(1)$$

for $j \to \infty$. This convergence is uniform as N varies (because $\mathbb{P}^{(j)}_{1/2}(\lambda_1 \le N) = 1 = \mathbb{P}(\alpha_1 \le N/j)$ for $N \ge j$ and otherwise N/j lies in a compact interval). Furthermore, we have $\mathbb{P}^{(j)}_{1/2}(\lambda_1 \le N - 1) = O(1)$ in general. Fix an arbitrary $\varepsilon \in (0, 1)$, and decompose

$$\begin{aligned}T(n; N) =& \Big(\sum_{\varepsilon n \le j \le (1-\varepsilon)n} + \sum_{\substack{j < \varepsilon n, \text{ or} \\ j > (1-\varepsilon)n}} \Big) \frac{(1/4)_j (1/4)_{n-j}}{j!(n-j)!} \mathbb{P}^{(j)}_{1/2}(\lambda_1 \le N-1) \mathbb{P}^{(n-j)}_{1/2}(\lambda_1 \le N) \\ =& \frac{1}{\Gamma(1/4)^2} \sum_{\varepsilon n \le j \le (1-\varepsilon)n} j^{-3/4} (n-j)^{-3/4} \mathbb{P}(\alpha_1 \le \frac{N-1}{j}) \mathbb{P}(\alpha_1 \le \frac{N}{n-j}) \\ &+ o\Big(\sum_{\varepsilon n \le j \le (1-\varepsilon)n} j^{-3/4} (n-j)^{-3/4} \Big) + O\Big(\sum_{\substack{0 < j < \varepsilon n, \text{ or} \\ n > j > (1-\varepsilon)n}} j^{-3/4} (n-j)^{-3/4} \Big) + O(n^{-3/4})\end{aligned} \tag{6.36}$$

as $n \to \infty$, where the rate at which the error term $o(\cdots)$ tends to zero depends upon ε, but the constants of other error terms are absolute, with the last error term $O(n^{-3/4})$ coming from the terms $j = 0$ and $j = n$ in the sum. If $N/n \to s$ as $n \to \infty$, then

$$\mathbb{P}(\alpha_1 \le \frac{N-1}{j}) = \mathbb{P}(\alpha_1 \le \frac{(N-1)/n}{j/n}) = \mathbb{P}(\alpha_1 \le \frac{s}{j/n}) + o(1),$$

uniformly for $1 \le j \le n$. (The reason for uniformity is again due to compactness.) Of course we have

$$\mathbb{P}(\alpha_1 \le \frac{N}{n-j}) = \mathbb{P}(\alpha_1 \le \frac{s}{1-j/n}) + o(1)$$

as $n \to \infty$. Moreover,

$$\sum_{\substack{0<j<\varepsilon n, \text{ or} \\ n>j>(1-\varepsilon)n}} j^{-3/4}(n-j)^{-3/4} = O(\varepsilon^{1/4} n^{-1/2}),$$

and

$$\sum_{0<j<n} j^{-3/4}(n-j)^{-3/4} = O(n^{-1/2})$$

Hence we can simplify (6.36) to

$$\begin{aligned}(6.36) =& \frac{1}{\Gamma(1/4)^2}\frac{1}{\sqrt{n}} \sum_{\varepsilon n \le j \le (1-\varepsilon)n} \frac{1}{n}(j/n)^{-3/4}(1-j/n)^{-3/4}\mathbb{P}(\alpha_1 \le \frac{s}{j/n})\mathbb{P}(\alpha_1 \le \frac{s}{1-j/n}) \\ &+ O(\varepsilon^{1/4}n^{-1/2}) + o(n^{-1/2}) \\ =& \frac{1}{\Gamma(1/4)^2}\frac{1}{\sqrt{n}} \int_{\varepsilon}^{1-\varepsilon} t^{-3/4}(1-t)^{-3/4}\mathbb{P}(\alpha_1 \le \frac{s}{t})\mathbb{P}(\alpha_1 \le \frac{s}{1-t})\, dt \\ &+ O(\varepsilon^{1/4}n^{-1/2}) + o(n^{-1/2}),\end{aligned}$$

with the second line following because the sum in the previous line is a Riemann sum. Completing the integral from the interval $[\varepsilon, 1-\varepsilon]$ to $[0,1]$ adds only an error of $O(\varepsilon^{1/4}n^{-1/2})$. Hence

$$\begin{aligned}T(n;N) =& \frac{1}{\Gamma(1/4)^2\sqrt{n}} \int_0^1 t^{-3/4}(1-t)^{-3/4}\mathbb{P}(\alpha_1 \le \frac{s}{t})\mathbb{P}(\alpha_1 \le \frac{s}{1-t})\, dt + O(\varepsilon^{1/4}n^{-1/2}) + o(n^{-1/2}) \\ =& \frac{1}{\sqrt{\pi n}}\Big(\int_0^1 \mathbb{E}\mathbf{1}(1-\frac{s}{\alpha_1} \le t \le \frac{s}{\alpha_1'})\sqrt{\pi}\Gamma(1/4)^{-2}t^{-3/4}(1-t)^{-3/4}\, dt + O(\varepsilon^{1/4}) + o(1)\Big) \\ =& \frac{1}{\sqrt{\pi n}}\Big(\mathbb{P}\Big(1-\frac{s}{\alpha_1} \le Y \le \frac{s}{\alpha_1'}\Big) + O(\varepsilon^{1/4}) + o(1)\Big),\end{aligned}$$

where α_1' is an independent copy of α_1 and $Y \sim \text{Beta}(1/4, 1/4)$. As ε is arbitrary this establishes the claim. □

6.10 Approximating $B(x)$

Recall the definitions of $B(x)$ and f, given in (1.29) and (1.30).

Theorem 6.9. *On the assumption of RH for $\zeta(s)$ and $L(s, \chi_4)$, for any $\varepsilon > 0$,*

$$B(x) = \frac{1}{\pi}\int_{1/2}^1 \frac{x^s}{(1-s)^{1/2}s} f(s)\, ds + O_\varepsilon(x^{1/2+\varepsilon}).$$

Proof. We have by Perron's formula (see [MV07, Cor. 5.3]), for $T = x^{100}$,

$$B(x) = \frac{1}{2\pi i}\int_{2-iT}^{2+iT} \frac{x^s}{s} F(s)\, ds + O(1).$$

For arbitrary $\varepsilon > 0$, let $\sigma = 1/2 + \varepsilon$, and let $\mathcal{K}_\delta$ be a contour from $\sigma - i\delta$ to $1 + \delta - i\delta$ to $1+\delta+i\delta$ to $\sigma+i\delta$ for $\delta > 0$. On the Riemann Hypothesis the contour from $2-iT$ to $2+iT$ may be shifted to a contour from $2-iT$ to $\sigma - iT$ to $\sigma - i\delta$, followed by $\mathcal{K}_\delta$, followed by a contour from $\sigma + i\delta$ to $\sigma + iT$ to $2 + iT$. The Lindelöf estimates $\zeta(s), L(s, \chi_4) = O_\varepsilon(|s|^\varepsilon)$ for $\Re s \geq 1/2$, $|s-1| \geq 1/10$ can be used to bound those contours other than $\mathcal{K}_\delta$, yielding

$$B(x) = \frac{1}{2\pi i}\int_{\mathcal{K}_\delta} \frac{x^s}{s}\frac{f(s)}{(s-1)^{1/2}}\, ds + O_\varepsilon(x^{1/2+10\varepsilon}).$$

Letting $\delta \to 0$ shows this is

$$= \frac{1}{\pi}\int_\sigma^1 \frac{x^s}{(1-s)^{1/2}s} f(s)\, ds + O_\varepsilon(x^{1/2+10\varepsilon}) = \frac{1}{\pi}\int_{1/2}^1 \frac{x^s}{(1-s)^{1/2}s} f(s)\, ds + O_\varepsilon(x^{1/2+10\varepsilon}),$$

which yields the claim. □

7 Twin primes in the large-q limit

7.1 Fundamental identity

We start by proving an identity, transporting the study of twin primes to Fourier space.

Proposition 7.1. *Let $\alpha, \beta \colon \mathcal{M}_q \to \mathbb{C}$ be two arithmetic functions. Let n be a positive integer and let $c \in \mathbb{F}_q^\times$. We have*

$$\frac{\sum_{f\in\mathcal{M}_{n,q}} \alpha(f)\overline{\beta}(f+c)}{\#\mathcal{M}_{n,q}} = \frac{1}{q^{2n}}\sum_{\chi\in G(R_n)} \chi(T^n + c) S(n, \alpha\cdot\chi)\overline{S(n, \beta\cdot\chi)}.$$

Proof. The orthogonality relation (2.3) implies that for any $f \in \mathcal{M}_{n,q}$ we have

$$\alpha(f) = \frac{1}{q^n}\sum_{\substack{g_1\in\mathcal{M}_{n,q}\\ \chi\in G(R_n)}} \alpha(g_1)\chi(g_1)\overline{\chi(f)}, \tag{7.1}$$

and similarly we have

$$\beta(f+c) = \frac{1}{q^n}\sum_{\substack{g_2\in\mathcal{M}_{n,q}\\ \chi\in G(R_n)}} \beta(g_2)\chi(g_2)\overline{\chi(f+c)}. \tag{7.2}$$

From (7.1) and (7.2) we obtain

$$\begin{aligned}&\frac{\sum_{f\in\mathcal{M}_{n,q}} \alpha(f)\overline{\beta}(f+c)}{\#\mathcal{M}_{n,q}}\\ &\quad= \frac{1}{q^{3n}}\sum_{\substack{g_1,g_2\in\mathcal{M}_{n,q}\\ \chi_1,\chi_2\in G(R_n)}} \Big(\alpha(g_1)\chi_1(g_1)\overline{\beta(g_2)\chi_2(g_2)}\sum_{f\in\mathcal{M}_{n,q}} \overline{\chi_1(f)}\chi_2(f+c)\Big).\end{aligned} \tag{7.3}$$

We have

$$f + c \equiv (T^n + c) \cdot f \bmod R_n \tag{7.4}$$

for all $f \in \mathcal{M}_{n,q}$. For any pair of characters $\chi_1, \chi_2 \in G(R_n)$ we have, by (7.4) and the orthogonality relation (2.1) with $F = \mathcal{M}_{n,q}$,

$$\begin{aligned} \sum_{f \in \mathcal{M}_{n,q}} \overline{\chi_1(f)}\chi_2(f+c) &= \chi_2(T^n + c) \sum_{f \in \mathcal{M}_{n,q}} \overline{\chi_1(f)}\chi_2(f) \\ &= \chi_2(T^n + c) \cdot q^n \cdot \mathbf{1}_{\chi_1 = \chi_2}. \end{aligned} \tag{7.5}$$

Plugging (7.5) in (7.3), we conclude the proof. □

7.2 Hidden symmetry

The following key proposition introduces an action of $\mathbb{F}_q^\times$ on $G(R_\ell)$, which preserves primitivity and L-functions.

Proposition 7.2. *Let ℓ be a positive integer. Let $\chi \in G(R_\ell)$ be a primitive character. For any $c \in \mathbb{F}_q^\times$, define a function $\chi_c \colon \mathcal{M}_q \to \mathbb{C}$ by*

$$\chi_c(f) = \chi(f(cT)/c^{\deg f}).$$

Then χ_c is well defined on $\mathcal{M}_q/R_\ell$ and in fact is a primitive character in $G(R_\ell)$. Moreover,

$$L(u, \chi) = L(u, \chi_c).$$

Proof. Fix $c \in \mathbb{F}_q^\times$. Let $f_1, f_2 \in \mathcal{M}_q$ be polynomials such that $f_1 \equiv f_2 \bmod R_{\ell,}$. Then f_1, f_2 have the same first ℓ next-to-leading coefficients. The i-th next to leading coefficient of $f_j(cT)/c^{\deg f_j}$ ($j \in \{1, 2\}$) is the i-th next-to-leading coefficient of $f_j(T)$, divided by c^i. Thus, $f_1(cT)/c^{\deg f_1} \equiv f_2(cT)/c^{\deg f_2} \bmod R_{\ell,1}$. This shows that χ_c can be regarded as a function of $\mathcal{M}_q/R_{\ell,1}$. By definition, χ_c is multiplicative, and it takes 1 to 1, so $\chi_c \in G(R_{\ell,1})$.

The coefficients of u^i in $L(u, \chi)$ and $L(u, \chi_c)$ are given by $\sum_{f \in \mathcal{M}_{i,q}} \chi(f(T))$ and $\sum_{f \in \mathcal{M}_{i,q}} \chi(f(cT)/c^i)$, respectively. The map $f \mapsto f(cT)/c^i$ is a permutation of $\mathcal{M}_{i,q}$, whose inverse is given by $f \mapsto f(T/c)c^i$. Thus, $L(u, \chi) = L(u, \chi_c)$. As $\deg L(u, \chi_c) = \deg L(u, \chi) = \ell - 1$, it follows that χ_c is a primitive character. □

Lemma 7.3. *Let ℓ be a positive integer. Let $\chi \in G(R_{\ell,1})$. Let $c \in \mathbb{F}_q^\times$. For any factorization function α, we have*

$$S(n, \alpha \cdot \chi) = S(n, \alpha \cdot \chi_c).$$

Proof. Since $f(T), f(cT)/c^{\deg f}$ have the same extended factorization type for any $c \in \mathbb{F}_q^\times$, and the inverse of $f \mapsto f(cT)/c^{\deg f}$ is $f \mapsto f(T/c)c^{\deg f}$, we have

$$\begin{aligned} S(n, \alpha \cdot \chi_c) &= \sum_{f \in \mathcal{M}_{n,q}} \alpha(f)\chi(f(cT)/c^n) = \sum_{f \in \mathcal{M}_{n,q}} \alpha(f(T/c)c^n)\chi(f) \\ &= \sum_{f \in \mathcal{M}_{n,q}} \alpha(f)\chi(f) = S(n, \alpha \cdot \chi), \end{aligned}$$

as needed. □

7.3 An equidistribution result

Let $\chi \in G(R_\ell) \setminus G(R_{\ell-1})$. By §2.3, $L(u,\chi) = \prod_{i=1}^{\ell-1}(1-\gamma_i(\chi)u)$ with $|\gamma_i(\chi)| = \sqrt{q}$. We denote by Θ_χ any unitary matrix in $U(\ell-1)$ whose eigenvalues are $\gamma_1(\chi)/\sqrt{q}, \dots, \gamma_{\ell-1}(\chi)/\sqrt{q}$. The following theorem is proved in Appendix A of [GS20].

Theorem 7.4. *[GS20, Thm. 8] Let $\ell \geq 4$. For any $\chi \in G(R_\ell)$ and $c \in \mathbb{F}_q^\times$, set*

$$A(\chi, c) = \frac{\sum_{\lambda \in \mathbb{F}_q^\times} \chi(T^\ell + c\lambda^\ell)}{\sqrt{q}}.$$

Let ρ be an irreducible representation of $\mathrm{PU}(\ell-1)$. Then

$$\frac{\sum_{\chi \in G(R_\ell)\setminus G(R_{\ell-1})} \mathrm{tr}(\rho(\Theta_\chi))A(\chi,c)}{|G(R_\ell) \setminus G(R_{\ell-1})|} = O_\rho\left(\frac{1}{\sqrt{q}}\right).$$

Informally, it says that the zeros of $L(u,\chi)$ are do not correlate, in the large-q limit, with the Gauss sum $A(\chi, c)$.

7.4 Conclusion of proof

Applying Proposition 7.1 with $\alpha = \beta = \Lambda_q$, we find that

$$\frac{\sum_{f \in \mathcal{M}_{n,q}} \Lambda_q(f)\Lambda_q(f+c)}{q^n} = \frac{1}{q^{2n}} \sum_{\chi \in G(R_n)} \chi(T^n + c)|S(n, \Lambda_q \cdot \chi)|^2. \tag{7.6}$$

The term corresponding to $\chi = \chi_0$ is 1, since $\langle \Lambda_q \rangle_{\mathcal{M}_{n,q}} = 1$ (1.10). Since $S(n, \Lambda_q \cdot \chi) = O(nq^{n/2})$ by Lemma 2.1, and since there are $O(q^{n-1})$ non-primitive characters modulo R_n, we have

$$\frac{\sum_{f \in \mathcal{M}_{n,q}} \Lambda_q(f)\Lambda_q(f+c)}{q^n} = 1 + \frac{1}{q^{2n}} \sum_{\chi \in G(R_n)/G(R_{n-1})} \chi(T^n + c)|S(n, \Lambda_q \cdot \chi)|^2 + O\left(\frac{n^2}{q}\right).$$

We claim that the multiset $A = \{\chi_\lambda : \lambda \in \mathbb{F}_q^\times, \chi \in G(R_n)\}$ consists of $q-1$ copies of $G(R_n)$. Indeed, the map $\chi \mapsto \chi_\lambda$ is a bijection for any $\lambda \in \mathbb{F}_q^\times$. Thus, in (7.6) we may sum over primitive characters in A and divide by $q-1$, instead of summing over primitive characters in $G(R_n)$, and obtain from Lemma 7.3 the following:

$$\frac{\sum_{f \in \mathcal{M}_{n,q}} \Lambda_q(f)\Lambda_q(f+c)}{q^n} = 1 + q^{-2n} \sum_{\chi \in G(R_n)\setminus G(R_{n-1})} \frac{\sum_{\lambda \in \mathbb{F}_q^\times} \chi_\lambda(T^n + c)}{q-1} |S(n, \Lambda_q \cdot \chi)|^2 + O\left(\frac{n^2}{q}\right). \tag{7.7}$$

When $\chi \in G(R_n)$, $\psi_\chi(x) := \chi(T^n + x)$ is an additive character of $\mathbb{F}_q$, since $(T^n + x_1)(T^n + x_2) \equiv T^n + x_1 + x_2 \bmod R_n$ and $(T^n + x)^p \equiv 1 \bmod R_n$. Moreover, we claim that if χ is primitive then ψ_χ is non-trivial. Otherwise, whenever $f \equiv g \bmod R_{n-1}$ we may write $f \equiv g \cdot (T^n + x) \bmod R_n$ for some $x \in \mathbb{F}_q$, and then $\chi(f) = \chi(g)\chi(T^n + x) = \chi(g)$, implying χ is not primitive, a contradiction. We set

$$A(\chi, c) = \frac{\sum_{\lambda \in \mathbb{F}_q^\times} \chi_\lambda(T^n + c)}{\sqrt{q}} = \frac{\sum_{\lambda \in \mathbb{F}_q^\times} \chi(T^n + \frac{c}{\lambda^n})}{\sqrt{q}} = \frac{\sum_{\lambda \in \mathbb{F}_q^\times} \chi(T^n + c\lambda^n)}{\sqrt{q}} = \frac{\sum_{\lambda \in \mathbb{F}_q^\times} \psi_\chi(c\lambda^n)}{\sqrt{q}}.$$

We express (7.7) as

$$\frac{\sum_{f\in\mathcal{M}_{n,q}}\Lambda_q(f)\Lambda_q(f+c)}{q^n}-1=\frac{\sqrt{q}}{q-1}\frac{\sum_{\chi\in G(R_n)\backslash G(R_{n-1})}A(\chi,c)|S(n,\Lambda_q\cdot\chi)|^2}{q^{2n}}+O\left(\frac{n^2}{q}\right).$$

To establish the bound (1.32), all we need is $S(n,\Lambda_q,\chi)=O(nq^{n/2})$ and Weil's bound on additive character sums [Sch76, Thm. 2E], implying $|A(\chi,c)|\leq n$.

To obtain a better saving in q we proceed as follows. By (2.11), if $\chi\in G(R_\ell)\backslash G(R_{\ell-1})$ then we may write $S(n,\Lambda_q\cdot\chi)$ as $-q^{n/2}\mathrm{Tr}(\Theta_\chi^n)$. Hence it suffices to prove that

$$\frac{\sum_{\chi\in G(R_n)\backslash G(R_{n-1})}|\mathrm{Tr}(\Theta_\chi^n)|^2A(\chi,c)}{|G(R_n)\setminus G(R_{n-1})|}=O\left(\frac{1}{\sqrt{q}}\right). \tag{7.8}$$

We decompose $|\mathrm{Tr}(U)^n|^2$ as a linear combination of irreducible characters of $\mathrm{PU}(n-1)$, and apply Theorem 7.4 with $\ell=n$ to each character and conclude that (7.8) holds, which concludes the proof of the theorem. □

References

[ABSR15] J. C. Andrade, L. Bary-Soroker, and Z. Rudnick. Shifted convolution and the Titchmarsh divisor problem over $\mathbb{F}_q[t]$. *Philos. Trans. Roy. Soc. A*, 373(2040):20140308, 18, 2015.

[Axe11] A. Axer. Über einige Grenzwertsätze. *Wien. Ber.*, 120:1253–1298, 1911.

[BC20] Sandro Bettin and J. Brian Conrey. Averages of long Dirichlet polynomials. *arXiv preprint arXiv:2002.09466*, 2020.

[BLL17] Abhishek Bhowmick, Thái Hoàng Lê, and Yu-Ru Liu. A note on character sums in finite fields. *Finite Fields Appl.*, 46:247–254, 2017.

[Bor98] Alexei Borodin. Point processes and the infinite symmetric group. Part II: Higher correlation functions. *arXiv preprint math/9804087*, 1998.

[BP09] Andreas O Bender and Paul Pollack. On quantitative analogues of the Goldbach and twin prime conjectures over $\mathbb{F}_q[t]$. *arXiv preprint arXiv:0912.1702*, 2009.

[BS14] Lior Bary-Soroker. Hardy-Littlewood tuple conjecture over large finite fields. *Int. Math. Res. Not. IMRN*, 2014(2):568–575, 2014.

[BSSW16] Lior Bary-Soroker, Yotam Smilansky, and Adva Wolf. On the function field analogue of Landau's theorem on sums of squares. *Finite Fields Appl.*, 39:195–215, 2016.

[Bum04] Daniel Bump. *Lie groups*. Springer, 2004.

[BW17] Jean Bourgain and Nigel Watt. Mean square of zeta function, circle problem and divisor problem revisited. *arXiv preprint arXiv:1709.04340*, 2017.

[Car15] Dan Carmon. The autocorrelation of the Möbius function and Chowla's conjecture for the rational function field in characteristic 2. *Philos. Trans. Roy. Soc. A*, 373(2040):20140311, 14, 2015.

[Cro75] M. J. Croft. Square-free numbers in arithmetic progressions. *Proc. London Math. Soc. (3)*, 30:143–159, 1975.

[dlBF20] Régis de la Bretèche and Daniel Fiorilli. Major arcs and moments of arithmetical sequences. *American Journal of Mathematics*, 142(1):45–77, 2020.

[EH91] Gove W. Effinger and David R. Hayes. *Additive number theory of polynomials over a finite field*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 1991. Oxford Science Publications.

[ER91] Ömer Eğecioğlu and Jeffrey B. Remmel. Brick tabloids and the connection matrices between bases of symmetric functions. *Discrete Appl. Math.*, 34(1-3):107–120, 1991. Combinatorics and theoretical computer science (Washington, DC, 1989).

[Erd74] Paul Erdős. On abundant-like numbers. *Canad. Math. Bull.*, 17(4):599–602, 1974.

[FV96] Philippe Flajolet and Ilan Vardi. Zeta function expansions of classical constants. 1996.

[GC68] I. J. Good and R. F. Churchhouse. The Riemann hypothesis and pseudo-random features of the Möbius sequence. *Math. Comp.*, 22:857–861, 1968.

[GM87] Daniel A. Goldston and Hugh L. Montgomery. Pair correlation of zeros and primes in short intervals. In *Analytic number theory and Diophantine problems (Stillwater, OK, 1984)*, volume 70 of *Progr. Math.*, pages 183–203. Birkhäuser Boston, Boston, MA, 1987.

[GMRR20] Ofir Gorodetsky, Kaisa Matomäki, Maksym Radziwiłł, and Brad Rodgers. On the variance of squarefree integers in short intervals and arithmetic progressions. *arXiv preprint arXiv:2006.04060 (to appear in GAFA)*, 2020.

[Gol81] Daniel Alan Goldston. *LARGE DIFFERENCES BETWEEN CONSECUTIVE PRIME NUMBERS*. ProQuest LLC, Ann Arbor, MI, 1981. Thesis (Ph.D.)–University of California, Berkeley.

[Gol05] D. A. Goldston. Notes on pair correlation of zeros and prime numbers. In *Recent perspectives in random matrix theory and number theory*, volume 322 of *London Math. Soc. Lecture Note Ser.*, pages 79–110. Cambridge Univ. Press, Cambridge, 2005.

[Gor17] Ofir Gorodetsky. A polynomial analogue of Landau's theorem and related problems. *Mathematika*, 63(2):622–665, 2017.

[Gor20] Ofir Gorodetsky. Mean values of arithmetic functions in short intervals and in arithmetic progressions in the large-degree limit. *Mathematika*, 66(2):373–394, 2020.

[GR18] Ofir Gorodetsky and Brad Rodgers. The variance of the number of sums of two squares in $\mathbb{F}_q[T]$ in short intervals. *arXiv preprint arXiv:1810.06002 (to appear in AJM)*, 2018.

[GS07] Andrew Granville and K. Soundararajan. An uncertainty principle for arithmetic sequences. *Ann. of Math. (2)*, 165(2):593–635, 2007.

[GS20] Ofir Gorodetsky and Will Sawin. Correlation of arithmetic functions over $\mathbb{F}_q[T]$. *Math. Ann.*, 376(3-4):1059–1106, 2020.

[Hal82] R. R. Hall. Squarefree numbers on short intervals. *Mathematika*, 29(1):7–17, 1982.

[Hal03] Christopher James Hall. *L-functions of twisted Legendre curves*. ProQuest LLC, Ann Arbor, MI, 2003. Thesis (Ph.D.)–Princeton University.

[Har16] G. H. Hardy. On Dirichlet's Divisor Problem. *Proc. London Math. Soc. (2)*, 15:1–25, 1916.

[Hay65] David R. Hayes. The distribution of irreducibles in GF[q, x]. *Trans. Amer. Math. Soc.*, 117:101–127, 1965.

[HB88] D. R. Heath-Brown. The number of primes in a short interval. *J. Reine Angew. Math.*, 389:22–63, 1988.

[HL23] G. H. Hardy and J. E. Littlewood. Some problems of 'Partitio numerorum'; III: On the expression of a number as a sum of primes. *Acta Math.*, 44(1):1–70, 1923.

[Hoo79] C. Hooley. On a new technique and its applications to the theory of numbers. *Proc. London Math. Soc. (3)*, 38(1):115–151, 1979.

[HS17] Adam J. Harper and Kannan Soundararajan. Lower bounds for the variance of sequences in arithmetic progressions: primes and divisor functions. *Q. J. Math.*, 68(1):97–123, 2017.

[HT82] R. R. Hall and G. Tenenbaum. On the average and normal orders of Hooley's Δ-function. *J. London Math. Soc. (2)*, 25(3):392–406, 1982.

[Hux72] M. N. Huxley. On the difference between consecutive primes. *Invent. Math.*, 15:164–170, 1972.

[IK04] Henryk Iwaniec and Emmanuel Kowalski. *Analytic number theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.

[Ivi09] Aleksandar Ivić. On the divisor function and the Riemann zeta-function in short intervals. *Ramanujan J.*, 19(2):207–224, 2009.

[Jut84] Matti Jutila. On the divisor problem for short intervals. *Ann. Univ. Turku. Ser. A I*, (186):23–30, 1984. Studies in honour of Arto Kustaa Salomaa on the occasion of his fiftieth birthday.

[Kat13] Nicholas M. Katz. Witt vectors and a question of Keating and Rudnick. *Int. Math. Res. Not. IMRN*, (16):3613–3638, 2013.

[Kat15] Nicholas M. Katz. Witt vectors and a question of Entin, Keating, and Rudnick. *Int. Math. Res. Not. IMRN*, (14):5959–5975, 2015.

[Kor18] Sergei Korotkikh. Transition functions of diffusion processes with the Jack parameter on the Thoma simplex. *arXiv preprint arXiv:1806.07454*, 2018.

[KOV93] Sergei Kerov, Grigori Olshanski, and Anatoly Vershik. Harmonic analysis on the infinite symmetric group: a deformation of the regular representations. *CR Acad. Sci. Paris Ser. I Math.*, 316:773–778, 1993.

[KR14] Jonathan P. Keating and Zeév Rudnick. The variance of the number of prime polynomials in short intervals and in residue classes. *Int. Math. Res. Not. IMRN*, (1):259–288, 2014.

[KR16] Jonathan Keating and Zeev Rudnick. Squarefree polynomials and Möbius values in short intervals and arithmetic progressions. *Algebra Number Theory*, 10(2):375–420, 2016.

[KRRGR18] J. P. Keating, B. Rodgers, E. Roditty-Gershon, and Z. Rudnick. Sums of divisor functions in $\mathbb{F}_q[t]$ and matrix integrals. *Math. Z.*, 288(1-2):167–198, 2018.

[KS99] Nicholas M. Katz and Peter Sarnak. *Random matrices, Frobenius eigenvalues, and monodromy*, volume 45 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 1999.

[KS00] J. P. Keating and N. C. Snaith. Random matrix theory and $\zeta(1/2 + it)$. *Comm. Math. Phys.*, 214(1):57–89, 2000.

[Lan08] Edmund Landau. Über die einteilung der positiven ganzen zahlen in vier klassen nach der mindestzahl der zu ihrer additiven zusammensetzung erforderlichen quadrate. *Arch. Math. Phys.*, 13:305–312, 1908.

[Les16] Stephen Lester. On the variance of sums of divisor functions in short intervals. *Proc. Amer. Math. Soc.*, 144(12):5015–5027, 2016.

[Liu16] H.-Q. Liu. On the distribution of squarefree numbers. *J. Number Theory*, 159:202–222, 2016.

[LMF18] The LMFDB Collaboration. The l-functions and modular forms database. `http://www.lmfdb.org`, 2018. [Online; accessed 20 August 2018].

[Mai85] Helmut Maier. Primes in short intervals. *Michigan Math. J.*, 32(2):221–225, 1985.

[Mas20] Daniele Mastrostefano. A lower bound for the variance of generalized divisor functions in arithmetic progressions. *arXiv preprint arXiv:2004.05602*, 2020.

[Mon73] H. L. Montgomery. The pair correlation of zeros of the zeta function. In *Analytic number theory (Proc. Sympos. Pure Math., Vol. XXIV, St. Louis Univ., St. Louis, Mo., 1972)*, pages 181–193, 1973.

[Mon94] Hugh L. Montgomery. *Ten lectures on the interface between analytic number theory and harmonic analysis*, volume 84 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1994.

[MRl16] Kaisa Matomäki and Maksym Radziwił ł. Multiplicative functions in short intervals. *Ann. of Math. (2)*, 183(3):1015–1056, 2016.

[MV07] Hugh L. Montgomery and Robert C. Vaughan. *Multiplicative number theory. I. Classical theory*, volume 97 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2007.

[Ng08] Nathan Ng. The Möbius function in short intervals. In *Anatomy of integers*, volume 46 of *CRM Proc. Lecture Notes*, pages 247–257. Amer. Math. Soc., Providence, RI, 2008.

[Oko01] Andrei Okounkov. SL(2) and z-measures. In *Random matrix models and their applications*, volume 40 of *Math. Sci. Res. Inst. Publ.*, pages 407–420. Cambridge Univ. Press, Cambridge, 2001.

[Ols98] Grigori Olshanski. Point processes and the infinite symmetric group. Part I: The general formalism and the density function. *arXiv preprint math/9804086*, 1998.

[Ols03] Grigori Olshanski. An introduction to harmonic analysis on the infinite symmetric group. In *Asymptotic Combinatorics with Applications to Mathematical Physics*, pages 127–160. Springer, 2003.

[Ols18] G. I. Olshanskiĭ. The topological support of z-measures on the Thoma simplex. *Funktsional. Anal. i Prilozhen.*, 52(4):86–88, 2018.

[Pol08] Paul Pollack. Simultaneous prime specializations of polynomials over finite fields. *Proc. Lond. Math. Soc. (3)*, 97(3):545–567, 2008.

[Ram76] K. Ramachandra. Some problems of analytic number theory. *Acta Arith.*, 31(4):313–324, 1976.

[Rhi72] Georges Rhin. Répartition modulo 1 dans un corps de séries formelles sur un corps fini. *Dissertationes Math. (Rozprawy Mat.)*, 95, 1972.

[Rod18] Brad Rodgers. Arithmetic functions in short intervals and the symmetric group. *Algebra Number Theory*, 12(5):1243–1279, 2018.

[Roq18] Peter Roquette. *The Riemann hypothesis in characteristic p in historical perspective*, volume 2222 of *Lecture Notes in Mathematics*. Springer, Cham, 2018. History of Mathematics Subseries.

[Ros02] Michael Rosen. *Number theory in function fields*, volume 210 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2002.

[RS96] Zeév Rudnick and Peter Sarnak. Zeros of principal L-functions and random matrix theory. volume 81, pages 269–322. 1996. A celebration of John F. Nash, Jr.

[RS18] Brad Rodgers and Kannan Soundararajan. The variance of divisor sums in arithmetic progressions. *Forum Math.*, 30(2):269–293, 2018.

[Saw18] Will Sawin. The equidistribution of L-functions of twists by Witt vector Dirichlet characters over function fields. *arXiv preprint arXiv:1805.04330*, 2018.

[Sch76] Wolfgang M. Schmidt. *Equations over finite fields. An elementary approach.* Lecture Notes in Mathematics, Vol. 536. Springer-Verlag, Berlin-New York, 1976.

[Sel43] Atle Selberg. On the normal density of primes in small intervals, and the difference between consecutive primes. *Arch. Math. Naturvid.*, 47(6):87–105, 1943.

[Sha64] Daniel Shanks. The second-order term in the asymptotic expansion of $B(x)$. *Math. Comp.*, 18:75–86, 1964.

[SS03] Elias M. Stein and Rami Shakarchi. *Complex analysis*, volume 2 of *Princeton Lectures in Analysis*. Princeton University Press, Princeton, NJ, 2003.

[SS19] Will Sawin and Mark Shusterman. On the Chowla and twin primes conjectures over $\mathbb{F}_q[t]$. *arXiv preprint arXiv:1808.04001*, 2019.

[Sta99] Richard P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin.

[SV77] B. Saffari and R. C. Vaughan. On the fractional parts of x/n and related sequences. II. *Ann. Inst. Fourier (Grenoble)*, 27(2):v, 1–30, 1977.

[Vor03] Georges Voronoi. Sur un problème du calcul des fonctions asymptotiques. *J. Reine Angew. Math.*, 126:241–282, 1903.

[Wei74] André Weil. *Basic number theory*. Springer-Verlag, New York-Berlin, third edition, 1974. Die Grundlehren der Mathematischen Wissenschaften, Band 144.

תקציר

חלק מהבעיות הבסיסיות בתורת המספרים האנליטית נוגעות לסכומים של פונקציות אריתמטיות על פני קטעים קצרים. בעיות אלו כוללות קיום של ראשוניים בקטע, ביטול בסכומי מביוס והקיום של מספרי חסרי־ריבועים בקטע.

משוער לגבי רבות מהפונקציות הצצות באופן טבעי שהממוצע שלהן על פני קטע קצר אסימפטוטי לממוצע שלהן בקטע מלא. בנוסף, צפוי כי סכומים על פני קטע קצר יהנו מביטול מסוג שורש, במובן שהשגיאה היא לכל היותר שורש מספר המחוברים. אפילו תחת השערת רימן, בעיות אלו פתוחות למרבית הפונקציות.

רעיון נפוץ ושימושי הוא להכניס אקראיות לתוך הבעיה, על ידי בחירת קטע קצר באקראי. אז ניתן לחקור סכום של פונקציה על פני הקטע המקרי. חסמים עליונים על השונות של סכום מקרי זה מובילים לגרסאות 'כמעט כל' של השערות קשות.

בתזה זו אנו חוקרים את השונות של סכומים של פונקציות אריתמטיות בשדות פונקציות. אנו משתמשים בכלים קומבינטורים, אנליטים וגיאומטריים כדי להוכיח תוצאות חזקות מאלו הידועות מעל השלמים. ארבעת התוצאות המרכזיות שלנו הן

1. נוסחה אסימפטוטית לשונות של פולינומים חסרי־ריבועים על פני קטעים קצרים, המגיעה הרבה מעבר לעבודה של ר. ר. הול (1982) ונותנת ראיות לטובת השערה שנוסחה לאחרונה על ידי קיטינג ורודניק.

2. חסם עליון הדוק על השונות של פונקציות התפרקות על פני קטעים קצרים.

3. נוסחה אסימפטוטית לשונות של סכומים של שני ריבועים בקטעים קצרים, בגבול של q־גדול. הנוסחה שלנו סוטה מהמודל ההסתברותי הנאיבי ומייצרת תחזית מעל השלמים שמסכימה היטב עם הדאטא הנומרי.

4. נוסחה אסימפטוטית בגבול של q־גדול למספר התאומים הראשונים בחוג פולינומים עם שגיאה אופטימלית.

הפקולטה למדעים מדויקים
ע"ש ריימונד ובברלי סאקלר
בית הספר למדעי המתמטיקה

השונות של סכומים של פונקציות אריתמטיות

תזה זו הוגשה כמילוי חלקי של הדרישות לתואר דוקטור לפילוסופיה בבית הספר למדעי המתמטיקה, אוניברסיטת תל־אביב

על ידי

אופיר גורודצקי

תחת הנחייתם של
פרופ' זאב רודניק
פרופ' ליאור ברי־סורוקר

ינואר 2021