Partial sums of the Liouville function and further topics in analytic number theory

Jake Chinis

A Thesis

in

The Department

of

Mathematics and Statistics

Presented in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy (Mathematics) at

McGill University

Montreal, Quebec, Canada

August 2022

McGill University

Graduate and Postdoctoral Studies

This is to certify that the thesis prepared

By: Mr. Jake Chinis

Entitled: Partial sums of the Liouville function and further topics in analytic number theory

and submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Mathematics)

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the Final Examining Committee:

_____ Supervisor
Dr. Maksym Radziwiłł

_____ Co-Supervisor
Dr. Henri Darmon

_____ Internal Examiner
Dr. Michael Lipnowski

_____ External Examiner
Dr. Nathan Ng

_____ External Member
Dr. Chantal David

Approved by _____
Jacques Hurtubise, Chair
Department of Mathematics and Statistics

_____ 2022

_____
Bruce Lennox, Dean
Faculty of Science

ABSTRACT

A central problem in analytic number theory, which has fascinated mathematicians for over 2000 years, is that of understanding the distribution of the primes. From Riemann's famous 1859 memoir, we know that counting primes is directly related to understanding zero-free regions of the so-called Riemann Zeta-function. In turn, this latter problem is deeply connected to understanding partial sums of the Liouville function, a completely multiplicative function that counts the parity of the total number of prime factors of a given integer. In this manuscript based thesis, which is comprised of two original publications, we further study the relationship between the Liouville function and the distribution of the primes.

In the first paper, we assume Riemann's Hypothesis and study partial sums of the Liouville function. Adapting the methods developed by Matomäki and Radziwiłł in their seminal work on multiplicative functions in short intervals, we show that the Liouville function exhibits squareroot cancellation in almost all short intervals, provided that the length of the interval falls within some suitable range. This establishes the Liouville function analogue to Selberg's celebrated theorem on primes in short intervals.

In the second paper, we assume the existence of Siegel zeros and study multiple correlations of the Liouville function. More precisely, we build on earlier work of Germán and Kátai and show that the existence of Siegel zeros implies Chowla's Conjecture along a subsequence. This establishes the Liouville function analogue of the Hardy-Littlewood prime k-tuples conjecture, under the existence of Siegel zeros, and is closely related to Heath-Brown's work on Siegel zeros and twin primes. As an immediate corollary, which follows from a well-known argument due to Sarnak, we get that Sarnak's Conjecture on Möbius disjointness holds at infinitely-many scales, conditionally on the existence of Siegel zeros.

ABRÉGÉ

Un problème central de la théorie analytique des nombres, qui fascine les mathématiciens depuis plus de 2000 ans, est celui de la compréhension de la distribution des nombres premiers. D'après les célèbres mémoires de Riemann de 1859, nous savons que le comptage des nombres premiers est directement lié à la compréhension des régions sans zéro de la fonction dite Riemann Zeta. À son tour, ce dernier problème est profondément lié à la compréhension des sommes partielles de la fonction de Liouville, une fonction complètement multiplicative qui compte la parité du nombre total de facteurs premiers d'un entier donné. Dans cette thèse manuscrite, composée de deux publications originales, nous étudions plus avant la relation entre la fonction de Liouville et la distribution des nombres premiers.

Dans le premier article, nous supposons l'hypothèse de Riemann et étudions les sommes partielles de la fonction de Liouville. En adaptant les méthodes développées par Matomäki et Radziwiłł dans leurs travaux fondateurs sur les fonctions multiplicatives à intervalles courts, nous montrons que la fonction de Liouville présente une annulation racine carrée dans presque tous les intervalles courts, à condition que la longueur de l'intervalle se situe dans une plage appropriée. Cela établit l'analogue de la fonction de Liouville au célèbre théorème de Selberg sur les nombres premiers à intervalles courts.

Dans le second article, nous supposons l'existence de zéros de Siegel et étudions les corrélations multiples de la fonction de Liouville. Plus précisément, nous nous appuyons sur les travaux antérieurs de Germán et Kátai et montrons que l'existence de zéros de Siegel implique la conjecture de Chowla le long d'une sous-suite. Cela établit l'analogue de la fonction de Liouville de la conjecture k-tuples premiers de Hardy-Littlewood, sous l'existence de zéros de Siegel, et est étroitement liée aux travaux de Heath-Brown sur les zéros de Siegel et les nombres premiers jumeaux. Comme corollaire immédiat, qui découle d'un argument bien connu dû à Sarnak, nous obtenons que la conjecture de Sarnak sur la disjonction de Möbius est vraie à une infinité d'échelles, conditionnellement à l'existence de zéros de Siegel.

In memory of Dr. Elie Cohen and Πέγκυ Σταθόπουλος

Acknowledgements

Από πού ΄σαι κλωναράκι, από κείνο τό δεντράκι

First and foremost, I thank my supervisor, Maksym Radziwiłł. His continuous support, mathematical expertise, and kindness have proven invaluable on every occasion. It has been an absolute pleasure being his student and I could not have completed this thesis without him.

In general, the number theory group in Montreal has been of tremendous support throughout my studies and I thank Chantal David, Andrew Granville, Hershy Kisilevsky, and Dimitris Koukoulopoulos for everything they have taught me over the years and for their many helpful thoughts and comments on all my work. I am particularly grateful towards Henri Darmon for serving as my official supervisor while at McGill, as well as to Michael Lipnowski and Nathan Ng for serving as the examiners during my defense.

Much of this work was conducted while I was visiting CalTech and with the financial support of NSERC. I thank the math department at CalTech and, in particular, Michelle J. Vine, for their hospitality. I thank NSERC for allowing me the freedom to pursue these research projects.

I thank the teachers who were the first to light my passion for mathematics; in particular, Mr. Calame, Mrs. Stefatos, Mr. Mason, Prof. Shahabi, and Dr. Elie Cohen. They taught me the importance of a good teacher in a student's life and they are a big part of why I chose to study analytic number theory.

From the non-mathematical side of things (i.e., the most important side of things), I cannot thank my friends enough. I especially thank Peter Zenz for many fruitful mathematical discussions, as well as for all the intense squash, tennis, and chess matches; I could not have completed this thesis without his friendship. Then, in no particular order, I thank: Isabella Negrini and James Rickards for the hiking trips; Bart, Carlos, Damien, James^2, and Jordan for all the chess/poker/pool nights; Al, Brahim, Chris, Emilia, JB, Joe, Josh, Kenzy, and Patrick for their years of friendship; and Georgiana and Louise for always believing in me and for their continuous encouragement in all aspects of my life.

Finally, I thank my family and, in particular, my parents, for everything. Some of the fondest memories I had growing up were sitting down with my mother as she taught me the multiplication tables, or waking up in the middle of the night when my father would come home from the restaurant and I would help him count the register. These were my first forays into mathematics and I remember vividly the first time I learned that five 20 dollar bills made 100. If this thesis is worth anything, then my parents deserve all the praise.

List of Original Publications

This thesis contains the following original publications:

[Chi21a] Jake Chinis. On the Liouville Function in Short Intervals. *Int. Math. Res. Notices*, 04 2021.

[Chi21b] Jake Chinis. Siegel Zeros and Sarnak's Conjecture. Preprint, arXiv: math/2105.14653. Submitted to: *Adv. in Math.*, 2021

Contents

Summary of Research

1 Introduction

This manuscript based thesis contains two original publications [Chi21a, Chi21b] in the field of analytic number theory. In Section 1, we give a brief historical introduction on the distribution of primes, surveying recent work on partial sums/multiple correlations of the Liouville and von Mangoldt functions. The ultimate goal of Section 1 is to frame how the results contained in [Chi21a, Chi21b] fit into the grand scheme of things, before discussing future avenues of research. In Sections 2 and 3, we briefly discuss the work contained in [Chi21a] and [Chi21b], respectively. For the complete details concerning the new results in this thesis, we invite the reader to jump directly to Article I and Article II.

1.1 Naive Heuristics

Throughout this thesis, our main object of focus will be the Liouville function[1], λ, which is the completely multiplicative function defined by $\lambda(p) := -1$ for all primes p; that is, $\lambda : \mathbb{N} \to \{-1, 1\}$ with $\lambda(n) = 1$ if n has an even number of prime factors (counted with multiplicity) and -1 otherwise. To begin our study of the Liouville function, we first ask ourselves the following question: how many integers $n \leq x$ are such that $\lambda(n) = 1$? Since we do not expect any sort of bias in the number of prime factors that a given integer might have, it is reasonable to guess that about "half" of all integers have $\lambda(n) = 1$ and the other "half" have $\lambda(n) = -1$, so that $\#\{n \leq x : \lambda(n) = \pm 1\} \approx x/2$. If this were case, then we might also expect some cancellation in the partial sums $\sum_{n \leq x} \lambda(n)$, as many of the plus and minus ones would cancel. This seemingly innocuous heuristic already has deep implications. Indeed, showing that

$$\#\{n \leq x : \lambda(n) = \pm 1\} = \frac{x}{2}(1 + o(1))$$

is equivalent to showing that

$$\sum_{n \leq x} \lambda(n) = o(x),$$

as

$$\#\{n \leq x : \lambda(n) = \pm 1\} = \sum_{n \leq x} \frac{1 \pm \lambda(n)}{2} = \left\lfloor \frac{x}{2} \right\rfloor \pm \frac{1}{2} \sum_{n \leq x} \lambda(n).$$

Furthermore, showing that $\sum_{n \leq x} \lambda(n) = o(x)$ is equivalent to the Prime Number Theorem, which asserts that the density of the primes up to x is $1/\log x$:

[1] In all that follows, one can easily replace the Liouville function with the Möbius function, μ, which is simply the Liouville function supported on squarefree integers. Indeed, one can go between statements of λ and μ via the identity $\mu(n) = \sum_{d^2 | n} \mu(d) \lambda\left(\frac{n}{d^2}\right)$.

Theorem 1.1 (Prime Number Theorem)**.** *Let $\pi(x)$ denote the prime counting function, $\pi(x) := \#\{p \leq x\}$. Then,*

$$\pi(x) = (1 + o(1))\frac{x}{\log x},$$

as $x \to \infty$.

Remark 1.1. *The PNT was first conjectured by Gauss around $1792/3$, while looking at prime tables of Legendre. It was only in 1896 that the theorem was proved (independently) by Hadamard [Had96] and de la Vallée Poussin [dlVP96]. Their proofs followed ideas from Riemann's famous 1859 memoir [Rie59], which first established the link between the distribution of the primes and zero-free regions of the so-called Riemann Zeta-function.*

For $\Re(s) > 1$, the Riemann Zeta-function, denoted by $\zeta(s)$, is defined by the absolutely convergent infinite sum

$$\zeta(s) := \sum_{n \geq 1} \frac{1}{n^s}.$$

*A simple consequence of the Fundamental Theorem of Arithmetic allows us to see the connection between $\zeta(s)$ and the primes by writing the above infinite sum as an **infinite product**, indexed by the primes:*

$$\zeta(s) = \prod_{p} (1 - p^{-s})^{-1} \text{ for } \Re(s) > 1.$$

*We call such an infinite product an **Euler product**, after Leonhard Euler, who first observed the relationship between the primes and the infinite sum $\sum_{n \geq 1} \frac{1}{n^s}$.*

The connection between the primes and $\zeta(s)$ goes deeper: Riemann showed that $\zeta(s)$ can be analytically continued to the entire complex plane, save for a simple pole at $s = 1$, and gave an exact formula for the prime counting function (with appropriate weights) in terms of the zeros of $\zeta(s)$, which we call THE explicit formula for the Riemann Zeta-function. The idea is that counting primes is equivalent, via partial summation, to obtaining an asymptotic formula for the sum

$$\sum_{p^m \leq x} \log p$$

and one can express the above as a contour integral via Perron's Formula

$$\sum_{p^m \leq x} \log p = \frac{1}{2\pi i} \int_{(c)} \left(-\frac{\zeta'(s)}{\zeta(s)} \right) \frac{x^s}{s} \mathrm{d}s,$$

for any x which is not a power of a prime. Shifting the contour to the left and picking up the poles from the trivial and non-trivial zeros yields the explicit formula

$$\sum_{p^m \leq x} \log p = x - \sum_{\substack{\rho : \zeta(\rho) = 0 \\ 0 < \Re(\rho) < 1}} \frac{x^\rho}{\rho} - \frac{\zeta'(0)}{\zeta(0)} - \frac{1}{2} \log \left(1 - \frac{1}{x^2} \right),$$

provided x is not a prime power. For further details, see Section 17 of [Dav00].

The key thing to take away from the explicit formula is that zero-free regions of the Riemann Zeta-function are directly related to understanding the distribution of primes. For more on the history of the Prime Number Theorem and the relationship between primes and zeros, see [Gol73] and [Gra95].

From the above discussions, it is clear that understanding the Liouville function has deep implications concerning the distribution of prime numbers. The rest of this introduction looks to further establish the connection the Liouville function and the primes.

1.2 Probabilistic Heuristics

We now ask ourselves what happens if we make stronger (probabilistic) assumptions on the distribution of $\{\lambda(n)\}_n$; i.e., suppose we have a perfectly random, independent distribution, so that $\{\lambda(n)\}_n$ is a sequence of i.i.d. random variables, each taking the value ± 1 with probability $1/2$. In this case, the partial sums of the Liouville function mimic a *random walk* with mean zero and standard deviation equal to the squareroot of the number of steps. In particular, we might expect that the partial sums of the Liouville function exhibit *squareroot cancellation*; that is,

$$\sum_{n \leq x} \lambda(n) \ll x^{1/2+\epsilon}, \tag{1}$$

for any $\epsilon > 0$, as $x \to \infty$. It is a nice exercise, via a standard contour integral, to show that (1) is equivalent to the unsolved Riemann Hypothesis, which asserts that the non-trivial zeros[2] of $\zeta(s)$ lie on the line $\Re(s) = 1/2$; see [Tit32, Theorem 14.25 (C)], for example.

Continuing along these probabilistic lines, we may view the partial sums of λ as an *expectation*, where n is chosen uniformly randomly in the interval $[1, x]$:

$$\mathbb{E}_{n \leq x}[\lambda(n)] = \frac{1}{x} \sum_{n \leq x} \lambda(n).$$

Thus, we might also expect cancellation in partial sums of *multiple correlations* of the Liouville function. The analogy here comes from the fact that the expectation of a product of independent random variables is equal to product of the expectations; in other words,

$$\mathbb{E}_{n \leq x}[\lambda(n+1) \cdots \lambda(n+k)] \approx \prod_{i=1}^{k} \mathbb{E}_{n \leq x}[\lambda(n+i)] = o_k(1).$$

[2]Still in [Rie59], Riemann showed that $\zeta(s)$ satisfies the functional equation $\xi(1-s) = \xi(s)$ for all $s \in \mathbb{C}$, where $\xi(s)$ is the *completed Zeta-function*, defined by $\xi(s) := \frac{1}{2} s(s-1)\pi^{-s/2}\Gamma(s/2)\zeta(s)$, and where $\Gamma(s)$ is equal to the Gamma function. Since $\Gamma(s)$ has simple poles at $s = -n$ for all integers $n \in \mathbb{N}$, it is easy to see, via the functional equation, that $\zeta(s)$ has simple zeros at the negative integers and that these are the only zeros for $\Re(s) < 0$; we call these the *trivial zeros* of $\zeta(s)$. Furthermore, one can easily see that $\zeta(s) \neq 0$ for $\Re(s) > 1$, so that the remaining zeros of the Riemann Zeta-function must lie inside the so-called *critical strip* $0 \leq \Re(s) \leq 1$.

A famous conjecture due to Chowla [Cho65] states the above notion of independence as follows:

Conjecture 1.1 (Chowla's Conjecture[3])**.** *For any distinct $h_1, \dots, h_k \in \mathbb{N}$,*

$$\sum_{n \leq x} \lambda(n+h_1)\lambda(n+h_2)\cdots\lambda(n+h_k) = o_k(x),$$

as $x \to \infty$.

Remark 1.2. *The case $k = 1$ is equivalent to the Prime Number Theorem while the conjecture remains open for all $k \geq 2$. Up until recently, the case $k = 2$ was believed to be as difficult as the celebrated Twin Prime Conjecture, which asserts that there are infinitely-many* ***twin primes*** *(i.e., pairs of primes of the form $(p, p+2)$). Since the breakthrough work of Matomäki and Radziwiłł on multiplicative functions in short intervals [MR16], which helped resolve, among other things a logarithmically averaged form of Chowla's Conjecture [Tao16b, TT19b, TT18], as well as an averaged form of Chowla's Conjecture [MRT15], it is now believed that further progress on Chowla's Conjecture is within reach.*

1.3 From λ to Λ

As one might guess, in connection with Riemann's work, the aforementioned theorems/conjectures have analogues over the primes. More precisely, if we let Λ denote the von Mangoldt function (defined by $\Lambda(n) = \log p$ if $n = p^k$ with $\Lambda(n) = 0$ otherwise), then the Prime Number Theorem is equivalent to the estimate

$$\sum_{n \leq x} \Lambda(n) = (1 + o(1))x,$$

while Riemann's Hypothesis is equivalent to

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

Concerning multiple correlations of the von Mangoldt function, we have the famous Hardy-Littlewood Prime k-tuples Conjecture [HL23]:

Conjecture 1.2 (Hardy-Littlewood Conjecture[4])**.** *Let $\mathbf{h} := (h_1, \dots, h_k)$ be a k-tuple of distinct (positive) integers.*

[3] Chowla's Conjecture more generally states that $\sum_{n \leq x} \lambda(a_1 n + b_1)\cdots\lambda(a_k n + b_k) = o(x)$, for any integers a_i, b_i such $a_i b_j \neq a_j b_i$ and where the implied constant may depend on a_i, b_i.

[4] In Hardy and Littlewood's original paper, they predict an asymptotic formula for sums of the form $\sum_{n \leq x} \Lambda(a_1 n + b_1)\cdots\Lambda(a_k n + b_n)$ with $a_i b_j \neq a_j b_i$, much like the most general case of Chowla's Conjecture. The most general case of the Hardy-Littlewood Conjecture is thus a quantitative version of a conjecture due to Dickson [Dic04], which asserts that there are infinitely many prime k-tuples of the form $(a_1 n + b_1, \dots, a_k n + b_k)$, for any admissible a_i, b_i. See also Schinzel's Hypothesis [SS58] and the Bateman-Horn Conjecture [BH62] for analogues related to more general polynomials.

For every prime p, let $\nu_p(\mathbf{h})$ denote the number of distinct residue classes among $h_1, \dots, h_k$ modulo p. Then,

$$\sum_{n \leq x} \Lambda(n+h_1)\Lambda(n+h_2)\cdots\Lambda(n+h_k) = (\mathfrak{S}(\mathbf{h}) + o(1))x,$$

where

$$\mathfrak{S}(\mathbf{h}) := \prod_p \left(1 - \frac{\nu_p(\mathbf{h})}{p}\right)\left(1 - \frac{1}{p}\right)^{-k}.$$

Remark 1.3. *The case $k = 1$ is equivalent to the Prime Number Theorem while the conjecture remains open for all $k \geq 2$. In fact, the case $k = 2$ generalizes a quantitative version of the Twin Prime Conjecture. More precisely, the case $k = 2$ corresponds to a quantitative version of a conjecture due to de Polignac[5] [dP49], which asserts that there are infinitely-many consecutive primes p_n, p_{n+1} spaced h apart (i.e., $p_{n+1} = p_n + h$), for any even h.*

Remark 1.4. *Similarly to the heuristic which leads to Chowla's Conjecture, one obtains the Hardy-Littlewood Conjecture by thinking of an integer being prime as a "random" event: since there are roughly $x/\log x$ primes in the interval $(x, 2x]$, one can assume that the probability a given integer of size x is prime is roughly $1/\log x$. Cramér's random model for the primes assigns to each positive integer n a Bernoulli random variable such that $\mathbb{P}(X_n = 1) = 1/\log n$. If we let $\Pi(x) := \sum_{n \leq x} X_n$ (which counts the number of "random primes"), then one can show (using the law of iterated logarithm) that*

$$\Pi(x) = \int_2^x \frac{1}{\log t}\mathrm{d}t + O\left(\frac{(x\log\log x)^{1/2}}{(\log x)^{1/2}}\right),$$

almost surely as $x \to \infty$, which is in line with Riemann's Hypothesis for the prime counting function.

Concerning prime k-tuples, Cramér's random model predicts that

$$\#\{n \leq: n + h_1, \dots, n + h_k \text{ are all prime}\} \sim \frac{x}{(\log x)^k},$$

*for **any** choice of $h_1, \dots, h_k$. Of course, this is blatantly wrong, which one can see by taking $h_1 = 0$ and $h_2 = 1$, for example; the idea here is that the random model does not take into account the behaviour of the shifts h_i relative to small primes. One can rectify this using the Cramér-Granville random model, which is a version of Cramér's model which pre-sieves the small primes.*

In the Carmér-Granville model, one considers a sequence of Bernoulli random variables which are supported on integers free from small prime factors. We then note that for the k-tuple $(n + h_1, \dots, n + h_k)$ to contain only primes,

[5] Although Polignac's conjecture is far from resolved, there has been tremendous progress in recent years, notably by Zhang, Maynard, and Tao. Zhang [Zha14] was the first to show that the conjecture is true for some $h < 70000000$; he did this by overcoming some technical obstacles present in the GPY-sieve of Goldston, Pintz, and Yildirim [GPY09]. Around the same time, Maynard [May15] and Tao (independently) were able to reduce this bound quite a bit, by revisiting the so-called *Selberg Sieve*. Assuming the Generalized Elliott-Halberstam Conjecture, which is a result concerning the distribution of primes in arithmetic progressions, and pushing the methods of Maynard and Tao to their limits, the Polymath8b project has reduced the gap between primes to $h \leq 6$ [Pol14]. For more on bounded gaps between primes, see the expository articles of Soundararajan [Sou06] and Granville [Gra15] on the GPY-Sieve and Zhang-Maynard-Tao theorem, respectively.

it is necessary that n is not congruent to $-h_i$ modulo p for all small primes and for all $i = 1, \dots, k$. But we must also consider that the probability that any k random integers are all divisible by a prime p is $(1-1/p)^k$, which then yields the desired constant in Conjecture 1.2. See Chapter 17 and Exercise 17.4 of [Kou20] for more details on how to derive the Hardy-Littlewood constant from the Cramér-Granville model, as well as [Gra95].

Alternatively, one can justify the choice of constant $\mathfrak{S}(\mathbf{h})$ via the so-called Hardy-Littlewood circle method; see [HL23], as well as [Vau97].

1.4 Still on the Primes

In some sense, we know more about the primes (i.e. Λ), than we do about the Liouville function. For example, we have Selberg's celebrated theorem on the variance of primes in short intervals, as well as Heath-Brown's surprising work on Siegel zeros[6] and twin primes:

Theorem 1.2 ([Sel43] - Primes in Short Intervals)**.** *Assume RH and let $h = h(X) \leq X$. Then,*

$$\int_X^{2X} \Big| \sum_{x \leq n \leq x+h} \Lambda(n) - h \Big|^2 \mathrm{d}x \ll Xh(\log X)^2,$$

so that almost all short intervals contain the correct number of primes.

Theorem 1.3 ([HB83] - Siegel Zeros and Twin Primes)**.** *Let h_1, h_2 be distinct (positive) integers and let χ denote a real, primitive Dirichlet character of conductor q_χ. Suppose that the Dirichlet L-function $L(s,\chi)$ has a Siegel zero β and set $\eta := ((\beta-1)\log q_\chi)^{-1}$. Then,*

$$\frac{1}{x}\sum_{n \leq x} \Lambda(n+h_1)\Lambda(n+h_2) = \mathfrak{S}(\mathbf{h}) + O\left(\frac{1}{\log\log \eta}\right),$$

uniformly for $q_\chi^{250} \leq x \leq q_\chi^{500}$.

Remark 1.5. *Notice that Theorem 1.3 states that the Hardy-Littlewood Conjecture holds for $k = 2$, along some increasing subsequence, conditionally on the existence of Siegel zeros.*

One of the reasons why we know more about the primes than we do about the Liouville function has to do with the nature of their respective generating/Dirichlet series:

$$\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} = -\frac{\zeta'(s)}{\zeta(s)} \qquad \text{vs.} \qquad \sum_{n=1}^{\infty} \frac{\lambda(n)}{n^s} = \frac{\zeta(2s)}{\zeta(s)}.$$

[6] For those unfamiliar with Siegel zeros, the necessary background needed to understand Heath-Brown's work and the hypotheses in Theorem 1.3 can be found in Section 3. We borrow the notation and some notions from the recent work of Tao and Teräväinen [TT21], which we will return to shortly.

For example, Selberg's Theorem 1.2 is established via a standard contour integral, where one shifts the contour towards the half-line; assuming RH, the only pole of $\zeta'(s)/\zeta(s)$ occurs at $s = 1$. Of course, one still needs to take into account the contribution from s near the non-trivial zeros $\rho = 1/2 + i\gamma$, but these are well understood. In contrast, $\zeta(2s)/\zeta(s)$ has poles at $s = \rho$ and the residues are equal to $\zeta(2\rho)/\zeta'(\rho)$, but not much is known about $1/\zeta'(\rho)$. That said, Gao (unpublished) obtains the Liouville function analogue to Theorem 1.2 by using similar techniques as in work of Montgomery and Maier on partial sums of the Möbius function [MM09]. Indeed, Gao proceeds in a indirect manner by choosing a contour integral which avoids "clusters" of zeros of ζ near the half-line, much like in [MM09].

1.5 New Results

The above discussion thus establishes the bigger picture of the (new) results contained in this thesis. Guided by the fact that RH is equivalent to obtaining squareroot cancellation in the partial sums of the Liouville function, together with the breakthrough work of Matomäki and Radziwiłł on multiplicative functions in short intervals [MR16], we aim to prove that the partial sums of λ exhibit squareroot cancellation in almost all short intervals (conditionally on RH). This would then establish the Liouville function analogue to Selberg's Theorem 1.2. Similarly, we prove the Liouville function analogue to Heath-Brown's surprising work on Siegel zeros and twin primes, generalizing previous work of Germán and Kátai on 2-point correlations of the Liouville function [GK10]. More precisely, we prove that Chowla's Conjecture holds along a subsequence, conditionally on the existence of Siegel zeros.

Theorem 1.4 ([Chi21a] - Squareroot Cancellation in Short Intervals)**.** *Assume RH. Then,*

$$\int_X^{2X} \Big| \sum_{x \le n \le x+h} \lambda(n) \Big|^2 \mathrm{d}x \ll Xh(\log X)^6,$$

provided $h = h(X) \le \exp\left(\sqrt{\left(\frac{1}{2} - o(1)\right) \log X \log\log X}\right)$.

Remark 1.6. *Gao obtains the true analogue to Selberg's Theorem 1.2, in that his range of h includes larger values; namely, $(\log X)^A \le h \le X$, for some large constant $A > 0$. Although we cannot take h as large as X, our proof method is substantially simpler and allows us to take a much smaller power of* $\log X$.

Theorem 1.5 ([Chi21b] - Siegel Zeros and Chowla's Conjecture)**.** *Let $h_1, \dots, h_k$ be distinct (positive) integers and let χ denote a real, primitive Dirichlet character of conductor q_χ. Suppose that the Dirichlet L-function $L(s, \chi)$ has a Siegel zero β and set $\eta := ((\beta - 1)\log q_\chi)^{-1}$. Then,*

$$\frac{1}{x} \sum_{n \le x} \lambda(n + h_1) \cdots \lambda(n + h_k) \ll \frac{1}{(\log\log \eta)^{1/2} (\log \eta)^{1/12}},$$

for $q_\chi^{10} \le x \le q_\chi^{(\log\log \eta)/3}$.

1.6 Hybrid Results

The new results in the previous section further establish the rich analogy between the Liouville function (which counts the parity of the number of prime factors) and the primes themselves. Indeed, one might be inclined to combine statements of correlations of the Liouville function with correlations of the von Mangoldt function in a *hybrid* Hardy-Littlewood-Chowla Conjecture:

Conjecture 1.3 (Hybrid Hardy-Littlewood-Chowla)**.** *Let $k, \ell \geq 0$ and let $h_1, \dots, h_k$ and $h'_1, \dots, h'_\ell$ be fixed, distinct (positive) numbers. Then,*

$$\frac{1}{x}\sum_{n\leq x}\Lambda(n+h_1)\cdots\Lambda(n+h_k)\lambda(n+h'_1)\cdots\lambda(n+h'_\ell) = \mathfrak{S}(\mathbf{h})\cdot\mathbb{1}_{\ell=0}+o(1),$$

as $x\to\infty$, where $\mathbf{h} := (h_1, \dots, h_k)$.

Unconditionally, Conjecture 1.3 has only been proved when $k+\ell = 1$. Assuming the existence of Siegel zeros, we have the special cases where the conjecture holds along a subsequence for $k = 2$ and $\ell = 0$ by Heath-Brown, as well as $k = 0$ and any $\ell \geq 1$ by myself. Recently, Theorems 1.3 and 1.5 have been combined by Tao and Teräväinen [TT21], where they show that the existence of Siegel zeros implies that this hybrid version of the Hardy-Littlewood-Chowla Conjecture holds along a subsequence:

Theorem 1.6 ([TT21] - Siegel Zeros and Hardy-Littlewood-Chowla)**.** *Fix $0 \leq k \leq 2$ and $\ell \geq 0$. Let $h_1, \dots, h_k$ and $h'_1, \dots, h'_\ell$ be fixed, distinct natural numbers and let χ denote a real, primitive Dirichlet character of conductor q_χ. Suppose that the Dirichlet L-function $L(s,\chi)$ has a Siegel zero β and set $\eta := ((\beta-1)\log q_\chi)^{-1}$. Let $0 < \epsilon < 1$ be fixed and let x lie in the range*

$$q_\chi^{10k+1/2+\epsilon} \leq x \leq q_\chi^{\eta^{1/2}}.$$

Then,

$$\frac{1}{x}\sum_{n\leq x}\Lambda(n+h_1)\cdots\Lambda(n+h_k)\lambda(n+h'_1)\cdots\lambda(n+h'_\ell) = \mathfrak{S}(\mathbf{h})\cdot\mathbb{1}_{\ell=0}+O\left(\frac{1}{\log^{1/(10\max\{1,k\})}\eta}\right),$$

where $\mathbf{h} := (h_1, \dots, h_k)$.

Remark 1.7. *Theorem 1.6 builds and improves on earlier work of Heath-Brown, Germán-Kátai, and myself, by increasing the range in which our results hold, while simultaneously sharpening the bounds that one gets. Furthermore, their methods allow them to treat combinations of the Liouville function with the von Mangoldt function, proving results on a hybrid version of the Hardy-Littlewood-Chowla Conjecture, so long as the von Mangoldt function appears at most 2 times. As Tao and Teräväinen remark in their paper, this limitation on the number of copies of the von*

Mangoldt one can take is due to the fact that our knowledge of k-point correlations of the divisor function is limited to $k \leq 2$; see [TT21, Proposition 7.2].

1.7 Partial Progress in Recent Years

In the last few years, we have seen tremendous progress towards partial results of both Chowla's Conjecture and the Hardy-Littlewood Conjecture. Most notably, we have the work of Matomäki and Radziwiłł [MR16], which provided the necessary tools to prove, among other things, an averaged form of Chowla's Conjecture [MRT15], logarithmically averaged forms of Chowla's Conjecture [Tao16b, TT19b, TT18, HR21], and an averaged form of the Hardy-Littlewood Conjecture (as well as averaged forms concerning multiple correlations of divisor functions) [MRT18]. Most recently, we have the work of Lichtman and Teräväinen [LT21], which shows that Conjecture 1.3 holds on average (thus combining [MRT15] and [MRT18] and extending the latter by increasing the number of von Mangoldt correlations one can take from $k = 2$ to any $k \geq 0$):

Theorem 1.7 ([LT21] - HLC on Average)**.** *Let $\epsilon > 0$ and fix $k, \ell \geq 1$. Let $h_1, \dots, h_k$ and $h'_2, \dots, h'_\ell$ be fixed, distinct (positive) integers.*

- *Let $(\log X)^{k+\epsilon} \leq H \leq \exp((\log X)^a)$ with $a = a(\epsilon, k) > 0$ small enough. Then:*

$$\sum_{h'_1 \leq H} \left| \sum_{n \leq x} \Lambda(n+h_1) \dots \Lambda(n+h_k) \lambda(n+h'_1) \cdots \lambda(n+h'_\ell) \right| \ll HX \frac{\log\log H}{\log H}$$

- *Let $C \geq 1$ be fixed and let $(\log X)^{k+\epsilon} \leq H \leq (\log X)^C$. Then there exists an absolute constant $c > 0$ such that for any $10^5 k \epsilon^{-1} (\log\log X)/\log X \leq \delta \leq c/C$,*

$$\left| \sum_{n \leq x} \Lambda(n+h_1) \cdots \Lambda(n+h_k) \lambda(n+h'_1) \cdots \lambda(n+h'_\ell) \right| \leq \delta X,$$

for all but at most $\ll H^{1-c\delta}$ values of $h'_1 \leq H$.

Thus, Theorem 1.7 establishes Conjecture 1.3 for all but at most $o(H)$ values of $h'_1 \leq H$. Furthermore, Theorem 1.7 establishes an averaged form of a folklore conjecture dating back to at least Hildebrand [Hil89]:

Conjecture 1.4 (Folklore Conjecture)**.**

$$\sum_{p \leq x} \lambda(p+1) = o(\pi(x)).$$

Remark 1.8. *Notice that Conjecture 1.4 is equivalent to the simplest case of the Hardy-Littlewood-Chowla Conjecture, corresponding to $k = \ell = 1$ with $h_1 = 0$ and $h_1' = 1$, as*

$$\sum_{n \leq x} \Lambda(n)\lambda(n+1) = \sum_{p \leq x} \lambda(p+1)\log p + O(x^{1/2}(\log x)^2)$$

and where one removes the $\log p$ factor via partial summation.

Corollary 1.1 ([LT21] - Average Folklore)**.** *Let $\epsilon > 0$. Then, for $(\log X)^{1+\epsilon} \leq H \leq X$,*

$$\sum_{p \leq x} \lambda(p+h) = o(\pi(x))$$

for all but at most $o(H)$ values of $h \leq H$.

Remark 1.9. *Conjecture 1.4 is intimately related to the so-called parity barrier of sieve theory and is thus believed to be as difficult as the Twin Prime Conjecture. Roughly speaking, the parity problem states that sieve methods cannot distinguish between integers which have an even or an odd number or prime factors. This phenomenon was first discussed by Selberg in 1947, in his seminal work on sieves [Sel47]. For a complete introduction to sieve methods, including the parity problem and Selberg's Sieve, see [Kou20, Part 4].*

Recently, it was shown that a sufficiently strong averaged version of Conjecture 1.4 in arithmetic progressions (in the same spirit of the Bombieri-Vonagradov Theorem/Elliott-Halberstam Conjecture for primes) allows one to overcome the parity barrier and conclude the existence of infinitely-many twin primes; see [MV17].

1.8 The Möbius Randomness Law

One can interpret both new and old results concerning partial sums of the Liouville function as instances of the so-called *Möbius Randomness Law* [IK04, p. 338], which states that the values of the Liouville function (rather, the Möbius function) are random enough so that the partial sums

$$\sum_{n \leq x} a_n \lambda(n)$$

should exhibit cancellation for any "reasonable" sequence of complex numbers $\{a_n\}_n$. For example, $a_n = \lambda(n+1)$ is considered reasonable as the multiplicative structure in the shift $n+1$ should have no relation to the multiplicative structure of n (i.e., there is no obvious correlation between $\lambda(n+1)$ and $\lambda(n)$). A famous conjecture due to Sarnak characterizes one such family of reasonable sequences as those which are *deterministic*:

Definition 1.1. *Given a bounded sequence $f : \mathbb{N} \to \mathbb{C}$, its **topological entropy** is equal to the least exponent σ for*

which the set

$$\{(f(n+1), f(n+2), \dots, f(n+m))\}_{n=1}^{\infty} \subset \mathbb{C}^m$$

can be covered by $\mathcal{O}(\exp(\sigma m + o(m)))$ *balls of radius* ϵ *(in the* ℓ^∞ *metric), for any fixed* $\epsilon > 0$*, as* $m \to \infty$*. In the case where* $\sigma = 0$*, we say that* f *is* ***deterministic****.*

Conjecture 1.5 (Sarnak's Conjecture)**.** *Let* $f : \mathbb{N} \to \mathbb{C}$ *be a deterministic sequence. Then,*

$$\sum_{n \leq x} \lambda(n) f(n) = o_f(x),$$

as $x \to \infty$*.*

Remark 1.10. *In the case where* f *is constant, Sarnak's Conjecture is equivalent to the Prime Number Theorem; in the case where* f *is periodic, the conjecture is equivalent to the Prime Number Theorem in arithmetic progressions. For a thorough survey on instances for which Sarnak's Conjecture holds, see [FKPL18, KPL20], for example.*

What is interesting is that Liouville function is not believed to be deterministic (as one expects all 2^m possible sign patterns to occur infinitely-often in the m-tuples $(\lambda(n+1), \dots, \lambda(n+m))$, as n varies). This is also a conjecture due to Sarnak, which remains open, although there is partial progress for $m \leq 3$ due to Hildebrand [Hil86] and recent improvements by Matomäki-Radziwiłł-Tao [MRT16], following the seminal work of the first two authors on multiplicative functions in short intervals, as well as [TT19a], among others. Furthermore, Sarnak has shown that Chowla's Conjecture implies Sarnak's Conjecture [Sar11]; in particular, as immediate corollaries to Theorems 1.5 and 1.6, we have that Sarnak's Conjecture holds along a subsequence, conditionally on the existence of Siegel zeros:

Corollary 1.2 ([Chi21b, TT21] - Siegel Zeros Imply Sarnak's Conjecture)**.** *Let* χ *denote a real, primitive Dirichlet character of conductor* q_χ *and suppose that the Dirichlet* L*-function* $L(s, \chi)$ *has a Siegel zero. For any deterministic sequence* f *and any fixed* $\epsilon > 0$*,*

$$\sum_{n \leq x} \lambda(n) f(n) = o_f(x),$$

in the range $q_\chi^{1/2+\epsilon} \leq x \leq q_\chi^{\eta^{1/2}}$ *[TT21]* $(q_\chi^{10} \leq x \leq q_\chi^{\log\log \eta/3}$ *[Chi21b]).*

Remark 1.11. *Sarnak's proof that Chowla's Conjecture implies Sarnak's Conjecture is purely combinatorial and holds equally well for any other uncorrelated sequence. For another proof of this fact , which is based on ideas in ergodic theory, see [AKPLdlR17].*

1.9 Future Avenues of Research

The main takeaway from this brief historical introduction on the distribution of the primes is that there is a rich analogy between results concerning the Liouville function and the corresponding results over the primes. Furthermore, one can use this analogy in order to explore further avenues of research, say by looking at what is known in one setting and attempting to prove it in the other; we list some open problems below:

- In 1972, Huxley [Hux72] showed that every interval $[x, x+h]$ of length $h \geq x^{7/12+\epsilon}$ contains the correct number of primes; i.e.,

$$\sum_{x \leq n \leq x+h} \Lambda(n) = (1+o(1))h, \tag{2}$$

  for all $h \geq x^{7/12+\epsilon}$. Shortly thereafter, Ramachandra [Ram76] and Motohashi [Mot76] (independently) proved the analogous result for the Liouville function; namely,

$$\sum_{x \leq n \leq x+h} \lambda(n) = o(h), \tag{3}$$

  for all $h \geq x^{7/12+\epsilon}$. Recently, Matomäki and Teräväinen [MT22] were able to improve on (3) to include the range $h \geq x^{\theta}$ for any $\theta > 0.55$, which moves closer to the natural barrier of $\theta = 0.5$, which one obtains on the Riemann Hypothesis ([IK04, Section 10.5]). Still in [MT22], the authors mention that there is some limitations to their work which do not allow them to extend Huxley's (2) to include the larger range $h \geq x^{\theta}$ for $\theta > 0.55$. It is thus natural to ask whether or not one might be able to improve on (2) to establish the von Mangoldt function analogue in this wider range.

- Similarly, we have the recent improvements to Heath-Brown's Theorem 1.3 and Tao and Teräväinen's Theorem 1.6 by Matomäki and Merikoski [MM22, Theorem 1.4], where they prove a quantitative version of the Twin Prime Conjecture under the existence of Siegel zeros which is **uniform** in the shift; namely, they obtain an asymptotic formula for $\sum_{n \leq x} \Lambda(n)\Lambda(n+h)$, under the existence of Siegel zeros, which holds uniformly for $h \ll x$. This uniformity for large values of h then allows them to study other results on primes, conditionally on Siegel zeros, such as Goldbach's Conjecture [MM22, Theorem 1.4] and primes in short intervals [MM22, Corollary 1.5]. This latter result on Siegel zeros and primes in short intervals [MM22, Corollary 1.5] establishes a stronger result than Selberg's Theorem 1.2, by proving that almost all intervals as short as $h = h(X)$ such that $h/\log X \to \infty$ contain the correct number of primes, which is in line with a probabilistic model for the prime number theorem and which is known to hold under the Pair Correlation Conjecture (by work of Heath-Brown [HB82]). Thus, it would be interesting to see if the methods in [Chi21b, TT21] can be adapted to include shifts

in Theorems 1.5, 1.6 which grow with x and also if one could prove a stronger version of Theorem 1.4 under the existence of Siegel zeros, much like in [MM22].

- Concerning future work directly related to [Chi21a], it is of interest to further strengthen my results by increasing the range of h in Theorem 1.4, as well as proving Omega-type results by showing that squareroot cancellation is indeed optimal; this is an ongoing work.

This concludes our brief historical introduction to the distribution of primes. In the upcoming sections, we briefly discuss the work in [Chi21a, Chi21b]. For more information on the primes and the Riemann Zeta-funciton, see the excellent books of Davenport [Dav00], Iwaniec-Kowalski [IK04], Koukoulopoulos [Kou20], Montgomery-Vaughan [MV07], and Titchmarsh [THB86].

2 Multiplicative Function in Short Intervals

In the seminal work of Matomäki and Radziwiłł [MR16] on multiplicative functions in short intervals, the authors show that the "local" averages of 1-bounded multiplicative functions are equal to the corresponding "global" averages almost everywhere. More precisely, they prove the following:

Theorem 2.1 ([MR16])**.** *Let $f : \mathbb{N} \to [-1, 1]$ be a multiplicative function and let $h = h(X) \to \infty$ arbitrarily slowly as $X \to \infty$. Then, for almost all $x \in [X, 2X]$,*

$$\frac{1}{h} \sum_{x \leq n \leq x+h} f(n) = \frac{1}{X} \sum_{X \leq n \leq 2X} f(n) + o(1),$$

with $o(1)$ not depending on f.

As stated in their abstract, the so-called "MR-Theorem" has the following immediate corollaries:

> This result has several consequences. First, for the Möbius function we show that there are cancellations in the sum of $\mu(n)$ in almost all intervals of the form $[x, x + \psi(x)]$ with $\psi(x) \to \infty$ arbitrarily slowly. This goes beyond what was previously known conditionally on the Density Hypothesis or the stronger Riemann Hypothesis. Second, we settle the long-standing conjecture on the existence of x^{ϵ}-smooth numbers in intervals of the form $[x, x + c(\epsilon)\sqrt{x}]$, recovering unconditionally a conditional (on the Riemann Hypothesis) result of Soundararajan. Third, we show that the mean-value of $\lambda(n)\lambda(n + 1)$, with $\lambda(n)$ Liouville's function, is non-trivially bounded in absolute value by $1 - \delta$ for some $\delta > 0$. This settles an old folklore conjecture and constitutes progress towards Chowla's conjecture. Fourth, we show that a (general) real-valued multiplicative function f has a positive proportion of sign changes if and only if f is negative on at least one integer and non-zero on a positive proportion of the integers. This improves on many previous works, and is new already in the case of the Möbius function. We also obtain some additional results on smooth numbers in almost all intervals, and sign changes of multiplicative functions in all intervals of square-root length.

Regarding work which has emerged after the publication of [MR16], one can use the methods developed in the proof of Theorem 2.1 in order to prove, among other things, the following:

- An averaged form of Chowla's Conjecture [MRT15];
- Logarithmically averaged forms of Chowla's and Elliott's Conjectures [Tao16b, TT18, TT19b, HR21];
- The Erdős Discrepancy Problem [Tao16a] (see also [Sou18]);
- An averaged form of a hybrid version of the Hardy-Littlewood-Chowla/Elliott Conjectures [LT21]

In this section, we discuss the main ideas behind the proof of Theorem 2.1, before discussing related ideas which can be used in order to study the Liouville function more closely.

2.1 Sketch of the Proof of the MR-Theorem

We begin with a sketch of the proof of Theorem 2.1, hoping to highlight the key differences between the work contained in [MR16] and my own [Chi21a]; we closely follow Section 2 of [MR16]. We highly recommend that the reader consults the wonderful expository articles by Maynard [May19] and Soundararajan [Sou16].

There are two standard lemmas needed in order to make the problem of understanding the local averages of a multiplicative function more manageable. The first lemma allows us to write averages of multiplicative functions in terms of mean values of the corresponding Dirichlet polynomials. This, together with the fact that the mean values of Dirichlet polynomials are well understood, provide us with the necessary starting point in understanding the proof of Theorem 2.1.

Lemma 2.1 (Plancherel - [MR16, Lemma 14]). *Let* $|a_n| \leq 1$ *and suppose* $1 \leq h_1 \leq h_2 = X/(\log X)^{1/5}$. *For* $X \leq x \leq 2X$, *let*

$$S_j(x) := \sum_{x \leq n \leq x+h_j} a_n$$

and write

$$F(s) := \sum_{X \leq n \leq 4X} \frac{a_n}{n^s}.$$

Then,

$$\frac{1}{X}\int_X^{2X} \left|\frac{1}{h_1}S_1(x) - \frac{1}{h_2}S_2(x)\right|^2 \mathrm{d}x \ll \frac{1}{(\log X)^{2/15}} + \int_{1+i(\log X)^{1/15}}^{1+iX/h_1} |F(s)|^2 |\mathrm{d}s| + \max_{T \geq X/h_1} \frac{X/h_1}{T} \int_{1+iT}^{1+2iT} |F(s)|^2 |\mathrm{d}s|.$$

Sketch of the Proof. The proof follows from Perron's Formula, together with a few other tricks. The purpose of this sketch is to highlight some of these tricks.

From Perron's Formula, we have that

$$S_j(x) = \frac{1}{2\pi i}\int_{1-i\infty}^{1+i\infty} F(s)\frac{(x+h_j)^s - x^s}{s}\mathrm{d}s,$$

which we split into two integrals, according to the size of $|t|$. For small values of $|t|$ (i.e., $|t| \leq (\log X)^{1/5}$), trivial bounds suffice. For larger $|t|$, one hopes to remove a factor of the form x^s: once this is done, we can introduce a smoothing, expand the square, and change the order of integration, which allows us to exploit the rapid decay of the Mellin transform and thus get the desired cancellation. This is indeed the case, but the term $(x+h_j)^s$ prevents us from factoring such an x^s. We can overcome this obstacle by writing

$$\frac{(x+h_j)^s - x^s}{s} = \frac{1}{2h_j}\left(\int_{h_j}^{3h_j} \frac{(x+w)^s - x^s}{s}\mathrm{d}w - \int_{h_j}^{3h_j} \frac{(x+w)^s - (x+h_j)^s}{s}\mathrm{d}w\right).$$

Making a simple change of variables ($w = xu$ and $w = h_j + (x+h_j)u$, respectively) allows us to extract a factor of the form x^s from each of the integrals in the above.

The remaining part of the proof continues as described above, with the final step being an application of Fubini's Theorem; for further details, see [MR16, Lemma 14]. □

Having moved to the Fourier side via Plancherel, we have now changed the question of bounding sums of multiplicative functions in short intervals, to a question of bounding mean values of Dirichlet polynomials. The next lemma will be our main tool in understanding these mean values.

Lemma 2.2 (Mean Value Theorem)**.** *For any sequence of complex numbers $\{a_n\}_n$, we have that*

$$\int_0^T \Big|\sum_{1\leq n\leq N} a_n n^{it}\Big|^2 \mathrm{d}t = (T + \mathcal{O}(N)) \sum_{1\leq n\leq N} |a_n|^2.$$

Proof. See [IK04, Theorem 9.1]. □

Remark 2.1. *Notice that the MVT is best possible when the length of the sum and the length of integration are of the same size.*

Thus, after converting to Dirichlet polynomials via Plancherel, and disposing of the large values of T trivially through the MVT, we are left to bound integrals of the form

$$\int_{(\log X)^{1/15}}^T \Big|\sum_{X\leq n\leq 4X} \frac{a_n}{n^s}\Big|^2 \mathrm{d}t,$$

where we should think of $T \approx X/h$. The remaining part of the proof amounts to decomposing the Dirichlet polynomial according to the values of t for which some cancellation takes place. In [MR16], this decomposition is rather technical (we will discuss it shortly), but one can get a more intuitive understanding of the procedure from the proof of Theorem 1 in [GMRR21].

In [GMRR21], the authors use the same general set up to understand the variance of squarefree integers in short

intervals. At some point in their proof, obtaining an asymptotic for

$$\frac{1}{x}\int_X^{2X} \Big| \sum_{x\leq n\leq x+h} \mu^2(n) - \frac{h}{\zeta(2)}\Big|^2 \mathrm{d}x$$

is reduced to bounding an integral of the form

$$\int_0^T \Big|\zeta(1/2+it) \sum_{D\leq n\leq 2D} \frac{\mu(n)}{n^{1+2it}}\Big|^2 \mathrm{d}t, \tag{4}$$

for some parameters T and D. They proceed in a similar manner as in [MR16], by decomposing the domain of integration into regions where the corresponding Dirichlet polynomial attains some form of cancellation; in this case, they decompose the region of integration according to when the Dirichlet polynomial attains squareroot cancellation and whatever values of t remain. For values of t for which $\sum_{D\leq n\leq 2D} \mu(n)n^{-1+2it} \ll D^{-1/2+\epsilon}$, a pointwise bound on (4) suffices. For the remaining values of t, they go a step further by considering the sets

$$S(V) := \{t\in[0,T] : V \leq \Big| \sum_{D\leq n\leq D} \frac{\mu(n)}{n^{1+2it}}\Big| \leq 2V\},$$

for $V\in[D^{-1/2},1]$; in this range of t, one can bound the Dirichlet polynomial by $\ll V$, which is as sharp as one can hope for. Then, one simply needs to obtain a bound for $|S(V)|$, which is done via the standard Mean Value Theorem (Lemma 2.2) and a result of Huxley's on large values of Dirichlet polynomials ([IK04, Theorem 9.7 and Corollary 9.9]). Once we have this bound on $|S(V)|$, we can apply Cauchy-Schwartz to move to the 4-th moment of the Riemann Zeta-function; this approach is now winning as we have additional savings coming from the sharp bounds on $|S(V)|$.

Going back to the proof of Theorem 2.1, the authors decompose their Dirichlet polynomial into two sums, one of which contains integers with "typical" factorization, and the other which contains the remaining integers. They call this set of "typical" integers $\mathcal{S}$ and define it by setting $n\in\mathcal{S}\subset[X,2X]$ iff n has a prime factor in each interval $[P_j,Q_j]$, for some increasing sequence of disjoint intervals with $j=1,\dots,J$. Under some technical conditions, they show that the set of integers contained in $[X,2X]\setminus\mathcal{S}$ produce a negligible error, so that the bulk of their proof rests in bounding

$$\int_{(\log X)^{1/15}}^T \Big|\sum_{n\in\mathcal{S}} \frac{f(n)}{n^{1+it}}\Big|^2 \mathrm{d}t, \tag{5}$$

for a multiplicative function f. Once we have chosen appropriate intervals $[P_j,Q_j]$, the authors then decompose the interval $[(\log X)^{1/15},T]$ into $J+1$ disjoint sets $T_1,\dots,T_J,U$, which are defined by the following:

$$t\in T_j \text{ iff } j \text{ is the smallest index s.t. } \sum_{P\leq p\leq Q} \frac{f(p)}{p^{1+it}} \text{ is "small" for all } [P,Q]\subset[P_j,Q_j],$$

with $U := [(\log X)^{1/15}, T] \setminus \cup_{j=1}^{J} T_j$. This sort of decomposition is useful as, by the definition of $\mathcal{S}$, every integer $n \in \mathcal{S}$ has a prime factor $p \in [P_j, Q_j]$ for all j, so that one can extract a sum over primes from the Dirichlet polynomial in (5). Bounding the integral over T_j then amounts to bounding an integral of the form

$$\int_{T_j} \Big| \sum_{P_j \leq p \leq Q_j} \frac{f(p)}{p^{1+it}} \sum_{\substack{X/p \leq m \leq 2X/p \\ m \in S_j}} \frac{f(m)}{m^{1+it}} \cdot \frac{1}{\#\{q \text{ prime: } q|m\} + \mathbb{1}_{(m,p)=1}} \Big|^2 \mathrm{d}t, \tag{6}$$

where S_j is the set of integers $m \in [X, 2X]$ such that m has a prime factor in each interval $[P_i, Q_i]$ for $i = 1, \dots, J$, except possibly a prime factor in the interval $[P_j, Q_j]$. The upshot is that the Dirichlet polynomial over the primes is small for all $t \in T_j$, so that a pointwise bound on the sum over primes and the MVT on the sum over m should be enough to bound (6). This is indeed the case for $j = 1$, after dealing with the minor annoyance of separating the variables and moving from completely multiplicative f to general multiplicative functions. For $j > 1$, the MVT is not enough. After factoring out the sum over primes $p \in [P_j, Q_j]$, we are left to deal with the mean value of

$$\sum_{\substack{X/P \leq m \leq 2X/P \\ m \in S_j}} \frac{f(m)}{m^{1+it}} \cdot \frac{1}{\#\{P_j \leq p \leq Q_j : p|m\} + 1}, \tag{7}$$

for $P \in [P_j, Q_j]$. The problem here is that intervals $[P_j, Q_j]$ form an increasing sequence so that the length of the Dirichlet polynomial in (7) is too short relative to the length of of integration for the MVT to provide an optimal bound. To overcome this issue, the authors use the definition of T_j, which allows them to find an interval $[P', Q'] \subset [P_{j-1}, Q_{j-1}]$ such that

$$\sum_{P' \leq p \leq Q'} \frac{f(p)}{p^{1+it}}$$

is large, say $\asymp V$. Thus, one can replace computing the mean value of (7) by computing the mean value of

$$\left(\frac{1}{V} \sum_{P' \leq p \leq Q'} \frac{f(p)}{p^{1+it}} \right)^{\ell} \sum_{\substack{X/P \leq m \leq 2X/P \\ m \in S_j}} \frac{f(m)}{m^{1+it}} \cdot \frac{1}{\#\{P_j \leq p \leq Q_j : p|m\} + 1},$$

for some positive integer ℓ. The idea here is that one can take ℓ sufficiently large so that the length of the sum matches the length of integration, which ensures that applying the MVT is optimal. Furthermore, one can iterate this procedure by pulling out more and more prime factors; see [May19, Section 6.3], for example.

This establishes a **very** rough idea as to how the proof of Theorem 2.1 proceeds, but this is all that is necessary for our purposes when discussing the differences/similarities in [Chi21a]. For a more thorough outline of the proof of Theorem 2.1, see [MR16, Section 2], as well as the expository articles [Sou16, May19].

Remark 2.2. *In the case of the Liouville function and large integrals ($h = X^{\epsilon}$), one can avoid these messy details;*

see [MR15]. Furthermore, assuming RH and taking h not too large, we can not only avoid the technical difficulties of [MR16], but also obtain much stronger bounds (i.e., squareroot cancellation), as we shall presently see.

2.2 New Results on the Liouville Function in Short Intervals

In [Chi21a], I use methods inspired by the work of Matomäki and Radziwiłł in order to study partial sums of the Liouville function in short intervals. I prove that the Liouville function exhibits squareroot cancellation in almost all short intervals, conditionally on the Riemann Hypothesis and provided that the length of the short interval falls within a suitable range:

Theorem 2.2 ([Chi21a, Theorem 1.2])**.** *Assume the Riemann Hypothesis. Then,*

$$\int_X^{2X} \Big| \sum_{x \leq n \leq x+h} \lambda(n) \Big|^2 \mathrm{d}x \ll Xh(\log X)^6,$$

provided $h = h(X) \leq \exp\left(\sqrt{\left(\frac{1}{2} - o(1)\right) \log X \log\log X}\right)$.

The main tool in proving Theorem 2.2 is still the Mean Value Theorem (Lemma 2.2), with one minor modification: rather than decomposing the Dirichlet polynomials according to the values of $|t|$ for which the corresponding Dirichlet polynomial obtains some cancellation, I decompose the Dirichlet polynomial so that the length of the sum is small relative to the length of the integration. This can be thought of as the dual regime to the work in [MR16], where they apply the MVT in the range $T = o(N)$. In the case where $N = o(T)$, the MVT yields an optimal upper bound, which accounts for our sharper bounds. Of course, there is a cost to this: it is necessary to assume RH, as RH is equivalent to obtaining squareroot cancellation when h is large. Below, we give a quick sketch of the proof of Theorem 2.2.

After converting to Dirichlet polynomials via Plancherel (Lemma 2.1), our goal is to bound integrals of the form

$$\int_0^T \Big| \sum_{X \leq n \leq 4X} \frac{\lambda(n)}{n^{\frac{1}{2}+it}} \Big|^2 \mathrm{d}t,$$

where we should think of $T \approx X/h$. In order to make full use of the MVT, we would like to have a Dirichlet polynomial of length $\approx T$. Unfortunately, our polynomial has length $\approx X$. In order to circumvent this, we split the sum over $X \leq n \leq 4X$ into two sums

$$\Big| \sum_{X \leq n \leq 4X} \frac{\lambda(n)}{n^{\frac{1}{2}+it}} \Big|^2 \ll \Big| \sum_{\substack{X \leq n \leq 4X \\ \exists p | n : p > h}} \frac{\lambda(n)}{n^{\frac{1}{2}+it}} \Big|^2 + \Big| \sum_{\substack{X \leq n \leq 4X \\ p | n \Rightarrow p \leq h}} \frac{\lambda(n)}{n^{\frac{1}{2}+it}} \Big|^2,$$

where the first is over integers which have at least one prime factor $p > h$ and the second is over the integers all of whose prime factors are $\leq h$.

In the first sum, we can write such integers as $n = pm$ with $p > h$ and $m \approx X/p \ll X/h$. After factoring the sum over the primes $p > h$ (which is done via a Ramaré-type identity as in [MR16]), we are left with a sum over $X/p \leq m \leq 4X/p$, which is of the desired length to apply the MVT:

$$\sum_{\substack{X \leq n \leq 4X \\ \exists p | n : p > h}} \frac{\lambda(n)}{n^{\frac{1}{2}+it}} = \sum_{h < p \leq 4X} \frac{\lambda(p)}{p^{\frac{1}{2}+it}} \sum_{X/p \leq m \leq 4X/p} \frac{\lambda(m)}{m^{\frac{1}{2}+it}} \frac{1}{\#\{q \text{ prime: } q | m, q > h\} + \mathbb{1}_{p \nmid m}},$$

where the additional factor in the sum over m corresponds to the number of ways an integer $X \leq n \leq 4X$ can be written as $n = pm$ with $p > h$. Ignoring the minor inconvenience of separating the variables p and m (which can be done via Perron's Formula), we may apply a pointwise bound on the sum over p and the MVT on the sum over m, and this is enough to obtain squareroot cancellation for the integers which have at least one large "prime" factor.

For the remaining integers all of whose prime factors are $\leq h$, one can use the fact that there are few of these so-called h-smooth integers (at least when h is small), so that the Mean Value Theorem can be applied directly to give the desired bound. The idea here is that the length of the sum is equal to $\#\{n \in [X, 4X] : p | n \Rightarrow p \leq h\}$, which is $\ll X/h$ whenever $h = h(X) \leq \exp\left(\sqrt{\left(\frac{1}{2} - o(1)\right) \log X \log\log X}\right)$.

Remark 2.3. *As stated earlier, the above decomposition is convenient for many reasons: first, we avoid the messy decomposition of [MR16]; secondly, only standard tools are needed (such as "trivial" bounds on the sum over primes, regular dyadic decomposition to handle the extracted sum over primes, and basic facts about smooth numbers); third, we avoid the earlier approach through zeros of the Zeta-function (as in Selberg [Sel43] and Gao (unpublished)). Of course, these tricks require that h lies within a suitable range, so that it is an outstanding problem to determine what happens for larger h; this is an ongoing work.*

For the remaining details on my work in [Chi21a], see Article I.

3 Siegel Zeros and Multiple Correlations of the Liouville Function

In this section, we briefly discuss the classic work of Heath-Brown [HB83] on Siegel zeros and twin primes, before outlining the proofs of Germán-Kátai [GK10] and myself [Chi21b] on Siegel zeros and multiple correlations of the Liouville function. We begin with a standard introduction to Siegel zeros.

3.1 Background on Siegel Zeros and Consequences of Their Existence

To begin our study of Siegel zeros, we must first discuss zero-free regions of Dirichlet L-functions associated to Dirichlet characters $\chi \pmod q$; we borrow most of Section 2 from [Chi21b], following Chapter 12 of [Kou20] and

using the notation/terminology of [TT21].

Theorem 3.1. *Let $q \geq 3$ and set $Z_q(s) := \prod_{\chi \pmod q} L(s,\chi)$. Then, there is an absolute constant $c > 0$ such that the region $\Re(s) \geq 1 - \frac{c}{\log(q\tau)}$, where $\tau = \max\{1, |\Im(s)|\}$, contains at most one zero of Z_q. Furthermore, if this exceptional zero exists, then it is necessarily a real, simple zero of Z_q, say $\beta_1 \in [1 - c/\log q, 1]$, and there is a real, non-principal character $\chi_1 \pmod q$ such that $L(\beta_1, \chi_1) = 0$.*

Proof. See [Kou20, Theorem 12.3], for example. □

We call the character χ_1 in Theorem 3.1 an *exceptional character* and its zero, β_1, is the associated *exceptional zero*, or *Siegel/Landau–Siegel zero*. Note that this exceptional character depends on the choice of absolute constant and that this relationship implies some interesting facts:

- If we have one exceptional character, then we actually have infinitely many exceptional characters: if we had only finitely many exceptional characters $\chi_i \pmod{q_i}$, we could set $c' := \frac{1}{2}\min_i\{(1-\beta_i)\log q_i\}$ and we would then have that
$$1 - \frac{c}{\log q_i} \leq \beta_i < 1 - \frac{c'}{\log q_i}$$
for all i; in particular, replacing c with c' in Theorem 3.1, we no longer have any exceptional zeros.
- Similarly, we can take c to be arbitrarily small: if there are no exceptional zeros for c small enough, then we are done.

Thus, when we talk about Siegel zeros/exceptional characters, we are actually talking about an infinite sequence of real, primitive Dirichlet characters $\{\chi_\ell \pmod{q_\ell}\}_{\ell=1}^{\infty}$ for which $L(s,\chi_\ell)$ has a real zero
$$\beta_\ell = 1 - o_{\ell\to\infty}\left(\frac{1}{\log q_\ell}\right), \tag{8}$$
and such that no product $\chi_\ell\chi_{\ell'}$ is principal for any $\ell \neq \ell'$. Using Siegel's Theorem, we can quantify the rate of convergence in (8):

Theorem 3.2 (Siegel)**.** *Let $\epsilon > 0$. Then, there is a constant $c(\epsilon) > 0$, which cannot be computed effectively, such that $L(\sigma,\chi) \neq 0$ for $\sigma > 1 - c(\epsilon)q^{-\epsilon}$ and for all real, non-principal Dirichlet characters $\chi \pmod q$.*

Proof. See [Kou20, Theorem 12.10], for example. □

In particular,
$$\eta_\ell := ((\beta_\ell - 1)\log q_\ell)^{-1} \ll q_\ell, \tag{9}$$

as $\ell \to \infty$. In fact, one could show that $\eta_\ell \ll$ any fixed power of q_ℓ, but the above is all we need for our purposes.

Remark 3.1. *In most cases, such as in all that follows, this dependence on ℓ is implicit and often forgotten about; instead, we consider a "fixed" modulus q_χ with corresponding Siegel zero $\beta = 1 - \frac{1}{\eta \log q_\chi}$. Following [TT21], we call*

$$\eta := ((\beta - 1)\log q_\chi)^{-1}$$

*the **quality** of the zero.*

In an ideal world, these exceptional zeros do not exist: one expects that the non-trivial zeros of $L(s,\chi)$ all lie on the half-line, which is known as the Generalized Riemann Hypothesis (for Dirichlet L-functions). Unfortunately, all attempts at proving that Siegel zeros do not exist have so far failed. That said, it is fruitful to consider what happens in the "illusory" world where Siegel zeros do exist: one hopes that, by exploring this hypothesis further, we will eventually be led to a contradiction, ruling out their existence all together. Other reasons for studying problems under the assumption of Siegel zeros include the following:

- Siegel zeros allow us to overcome parity barriers of sieve theory and serve as a sanity check for certain conjectures concerning primes; for example, we have the work of Heath-Brown [HB83] on twin primes and, more recently, the work of Granville [Gra20] on lower bounds for the longest gaps between primes, among others.
- Similarly, one hopes that by studying problems under these assumptions, one can then also study the same problem by ruling out these assumptions, thereby proving the original claim unconditionally (much like in the proof of Linnik's Theorem, on the least prime in an arithmetic progression; see [Lin44a, Lin44b] for further details, as well as Tao's blog post on the subject. Note also that similar ideas have been used with GRH in place of Siegel zeros, such as in the proof of Gauss's conjecture on class numbers of imaginary quadratic fields [BM20] (whose proof is closely related to Siegel zeros).

As contemplating the existence of Siegel zeros is worthwhile per the above reasons, we now turn to one of the simplest implications; namely, if $L(\beta,\chi) = 0$, then $L(s,\chi)$ should be "small" for any s near $\beta \approx 1$, which in turn implies that $1/L(s,\chi) \approx \sum_{n\geq 1} \frac{\mu(n)\chi(n)}{n^s} = \prod_p \left(1 - \frac{\chi(p)}{p^s}\right)$ is large for s near 1. Thus, we might expect that $\chi(p)$ is biased towards -1; this is indeed the case for primes comparable to q_χ in the log scale, as first shown by Heath-Brown:

Lemma 3.1 ([HB83, Lemma 3])**.** *Suppose $L(s,\chi)$ has a Siegel zero β with corresponding conductor q_χ and quality η. Then,*

$$\sum_{\substack{p\leq x \\ \chi(p)=1}} \frac{\log p}{p} \ll \frac{\log q_\chi}{\sqrt{\log \eta}},$$

for $q_\chi^{250} \leq x \leq q_\chi^{500}$.

Sketch of the Proof. The idea behind Heath-Brown's proof is to notice that, for $s \approx 1 + 1/\log x$,

$$\sum_{\substack{p \le x \\ \chi(p)=1}} \frac{\log p}{p} \le \sum_{\substack{n \le x \\ \chi(n)=1}} \frac{\Lambda(n)}{n} \ll \sum_{\substack{n=1 \\ \chi(n)=1}}^{\infty} \frac{\Lambda(n)}{n^s} \ll -\frac{\zeta'(s)}{\zeta(s)} - \frac{L'(s,\chi)}{L(s,\chi)}.$$

Using Equations (8) and (17) from Section 12 of [Dav00], we can write the RHS in terms of the non-trivial zeros of $\zeta(s)$ and $L(s,\chi)$: the idea is to take the difference in (8) with $s = 1 + 1/\log q$ and $s = 2$ and similarly for (17) with $s = 1 + 1/\log q$ and $s = 1 + a/\log q$, for some $a \gg 1$ to be optimized later. Once this is done, we do the standard thing by treating the contribution from the Siegel zeros separately and using the Deuring-Heilbronn Phenomenon[7]/trivial bounds for the remaining zeros. □

Of course, this is not the strongest result of this form; using similar ideas, Germán and Kátai [GK10] prove the following:

Lemma 3.2 ([GK10])**.** *Suppose $L(s,\chi)$ has a Siegel zero β with corresponding conductor q_χ and quality $\eta > \exp(\exp(30))$. Then,*

$$\sum_{\substack{p \le x \\ \chi(p)=1}} \frac{\log p}{p} \ll \exp\left(\frac{\log x}{\log q_\chi}\right)(\log q_\chi)(\log \eta)^{-1/2},$$

uniformly for $x \in [q_\chi^{10}, q_\chi^{(\log\log \eta)/3}]$.

Remark 3.2. *Note that the upper bound in Lemma 3.2 is worse than that in Lemma 3 of [HB83], but the range of admissible x is larger: Lemma 3 of [HB83] yields the upper bound $\ll (\log q_\chi)(\log \eta_\ell)^{-1/2}$, uniformly for $x \in [q_\chi^{250}, q_\chi^{500}]$ (so that Lemma 3.2 recovers Heath-Brown's result when x is restricted to the interval $[q_\chi^{250}, q_\chi^{500}]$).*

One can do better than both Lemma 3.1 and Lemma 3.2, while avoiding this complex-analytic approach relying on the Deuring-Heilbronn phenomenon; we have the following variant due to Tao and Teräväinen:

Lemma 3.3 ([TT21, Proposition 3.5])**.** *Let $\epsilon > 0$. Then, for any $x \ge q_\chi^{(1+\epsilon)/2}$,*

$$\sum_{\substack{q_\chi^{(1+\epsilon)/2} < p \le x \\ \chi(p) \neq -1}} \frac{1}{p} \ll_\epsilon \frac{\log x}{\eta \log q_\chi}$$

and, for any natural number $m \ge 2$,

$$\sum_{\substack{q_\chi^{(1+\epsilon)/2m} < p \le q_\chi^{(1+\epsilon)/2(m-1)} \\ \chi(p) \neq -1}} \frac{1}{p} \ll_\epsilon \frac{m}{\eta^{1/m}}.$$

[7] The Deuring-Heilbronn Phenomenon states that the existence of exceptional zeros forces a sort of repulsion amongst the other zeros; namely, if β is our Siegel zero, then $r := \min\{|1-\rho| : \rho \neq \beta\} \gg \log \eta / \log q_\chi$ [Jut77, Theorem 2].

Remark 3.3. *Notice that the above sums are over primes for which $\chi(p) \neq -1$ (which includes all p such that $\chi(p) = 1$, as well as those for which $\chi(p) = 0$).*

Sketch of the Proof. The proof of Lemma 3.3 uses many of the same ingredients as Heath-Brown's Lemma 3.1, but rather than working directly with $\sum_{\substack{p \leq x \\ \chi(p)=1}} \frac{\log p}{p}$, Tao and Teräväinen consider the Dirichlet convolution $\sum_{n \leq x} \frac{(1 \star \chi)(n)}{n}$. The idea is to note that

$$\begin{aligned}\sum_{n \leq x} \frac{(1 \star \chi)(n)}{n} &= \sum_{d \leq x} \frac{\chi(d)}{d} \sum_{n \leq x/d} \frac{1}{n} \\ &= \sum_{d \leq x} \frac{\chi(d)}{d} \left(\log(x/d) + \gamma + O(\frac{d}{x}) \right) \\ &\approx (\log x + \gamma) L(1, \chi) + L'(1, \chi),\end{aligned}$$

for any Dirichlet character $\chi \pmod q$. Of course, one needs to be a little careful with the error terms; see [MV07, Exercise 11.2.3]. From there, the proof continues by using the explicit formula to show that $\frac{L'}{L}(1, \chi) \gg \eta \log q_\chi$ when χ is exceptional; see Theorem 11.4 of [MV07], for example. It then follows that

$$\sum_{n \leq x} \frac{(1 \star \chi)(n)}{n} \gg L(1, \chi) \eta \log q_\chi,$$

for x in the desired range. The trick now is to exploit the non-negative and multiplicative structure of $1 \star \chi$ to extract a sum over primes, ultimately proving the first half of the statement. The second half follows in a similar manner. □

The key point to take away from Lemmas 3.1, 3.2, and 3.3 is that there are fewer primes $p \leq x$ for which $\chi(p) = -1$ than there are for which $\chi(p) = 1$, at least when $\log x \approx \log q_\chi$. To see this, recall that the PNT yields the following:

$$\sum_{p \leq x} \frac{\log p}{p} = \log x + O(1)$$

in particular, Lemma 3.1 yields a saving of $\sqrt{\log \eta}$ over the trivial bound, so that there more primes for which $\chi(p) = -1$ than $\chi(p) = 1$. A similar analysis holds for the more recent improvements to Lemma 3.1. Point being that exceptional characters offer a good approximation to the Liouville function on "large" primes, as $\lambda(p) = -1$ for all p and $\chi(p) = -1$ often. This is the basic driving principle in [HB83, GK10, Chi21b, TT21]. This is more apparent for the work in [GK10] and [Chi21b], as we work with the Liouville function directly, the details of which are discussed in the next section. In the case of [HB83], one simply needs to note that $\Lambda(n) = (\mu \star \log)(n) \approx (\chi \star \log)(n)$. Tao and Teräväinen also use these general approximations for their hybrid results, but go a step further by decomposing what they call λ_{Siegel} and Λ_{Siegel} into "sharp" and "flat" approximants, which have structures of "Type I sums," which are well understood.

3.2 New Results on Chowla's Conjecture assuming the existence of Siegel Zeros

As mentioned in the previous section, the presence of a Siegel zero allows us to approximate the Liouville function by an exceptional character on large primes. Following Germán and Kátai, we define a completely multiplicative function λ_r by

$$\lambda_r(p) := \begin{cases} \lambda(p) & \text{if } p \leq r \\ \chi(p) & \text{if } p > r, \end{cases} \tag{10}$$

where r is some parameter to be optimised, but which one should think of as a small power of x (which eventually depends on η). By making a good choice for r, Lemma 3.2 allows us to trivially approximate

$$\sum_{n \leq x} \lambda(n+h_1) \cdots \lambda(n+h_k)$$

by

$$\sum_{n \leq x} \lambda_r(n+h_1) \cdots \lambda_r(n+h_k).$$

Then, one hopes to sieve the r-smooth integers, so that the above can be approximated by something like

$$\sum_{n \leq x} \chi(n+h_1) \cdots \chi(n+h_k).$$

The upshot here is that these types of character sums are well understood via Weil's Bound (e.g., [Bur63, Sch76]).

This is indeed the approach that is used in both [GK10] and [Chi21b], but the former is restricted to $k = 2$, while the latter handles all $k \geq 1$ and produces stronger error terms. One of the reasons for this is that [GK10] does not make use of the Weil Bound. Instead, they express χ in terms of a Gauss sum and evaluate the corresponding character sum

$$\sum_{n \leq x} \chi(n)\chi(n+1)$$

directly as -1. Another reason why Germán and Kátai are limited to $k = 2$ comes from the fact that their choice of sieve makes things complicated when dealing with general k-point correlations, even with the use of Weil's Bound. In [Chi21b], I use a different version of the Fundamental Lemma of Sieve Theory which is easier to work with and produces better error terms. Indeed, it was pointed out to me by Dimitris Koukoulopoulos and Maksym Radziwiłł that one can do even better by working with sieve weights directly, but this would not improve the error term in Theorem 1.5 without further inputs: to obtain any further improvement to Theorem 1.5, we would first need to improve Lemma 3.2, as we are limited by the error incurred from our initial approximation of λ by λ_r. For the complete details on my work in [Chi21b], see Article II.

Bibliography

[AKPLdlR17] Houcein El Abdalaoui, Joanna Kułaga-Przymus, Mariusz Lemańczyk, and Thierry de la Rue. The Chowla and the Sarnak conjectures from ergodic theory point of view. *Discrete and Continuous Dynamical Systems*, 37:2899–2944, 2017.

[BH62] Paul T. Bateman and Roger A. Horn. A heuristic asymptotic formula concerning the distribution of prime numbers. *Math. Comp*, 16:363–367, 1962.

[BM20] Ajit Bhand and M. Ram Murty. Class Numbers of Quadratic Fields. *Hardy-Ramanujan Journal*, Volume 42 - Special Commemorative volume in honour of Alan Baker - 2019, May 2020.

[Bur63] David A. Burgess. On character sums and L-series, II. *Proceedings of the London Mathematical Society*, s3-13(1):524–536, 1963.

[Chi21a] Jake Chinis. On the Liouville Function in Short Intervals. *Int. Math. Res. Notices*, 04 2021.

[Chi21b] Jake Chinis. Siegel Zeros and Sarnak's Conjecture. Preprint, arXiv: math/2105.14653. Submitted to: *Adv. in Math.*, 2021.

[Cho65] Sarvadaman Chowla. *The Riemann Hypothesis and Hilbert's Tenth Problem*, volume 4 of *Mathematics and its Applications*. Gordon and Breach Science Publishers, New York-London-Paris, 1965.

[Dav00] Harold Davenport. *Multiplicative Number Theory*. Graduate Texts in Mathematics 74. Springer, 3rd edition, 2000.

[Dic04] Leonard E. Dickson. A new extension of Dirichlet's theorem on prime numbers. *Messenger of Mathematics*, 33:155–161, 1904.

[dlVP96] Charles de la Vallée Poussin. Recherches analytiques sur la théorie des nombres premiers. *Ann. Soc. Sci. Bruxelles I-III*, 20:183–256, 281–362, 363–397, 1896.

[dP49] Alphonse de Polignac. Recherches nouvelles sur les nombres premiers. *Comptes rendus*, 29:397–401, 1849.

[FKPL18] Sébastien Ferenczi, Joanna Kułaga-Przymus, and Mariusz Lemańczyk. *Sarnak's Conjecture – what's new. In: Ferenczi, S., Kułaga-Przymus, J., Lemańczyk, M. (eds) Ergodic Theory and Dynamical Systems in their Interactions with Arithmetics and Combinatorics - Lecture Notes in Mathematics, vol 2213*. Springer International Publishing, 2018.

[GK10] László Germán and Imre Kátai. Multiplicative functions at consecutive integers. *Lithuanian Mathematical Journal*, 50(1):43–53, 2010.

[GMRR21] Ofir Gorodetsky, Kaisa Matomäki, Maksym Radziwiłł, and Brad Rodgers. On the variance of square-free integers in short intervals and arithmetic progressions. *Geometric and Functional Analysis*, 31(1):111–149, February 2021.

[Gol73] Larry J. Goldstein. A History of the Prime Number Theorem. *Am. Math. Mon.*, 80(6):599–615, 1973.

[GPY09] Daniel Goldston, János Pintz, and Cem Yıldırım. Primes in tuples I. *Annals of Mathematics*, 170(2):819–862, September 2009.

[Gra95] Andrew Granville. Harald Cramér and the Distribution of the Prime Numbers. *Scand. Actuarial J.*, 1:12–28, 1995.

[Gra15] Andrew Granville. Primes in intervals of bounded length. *Bulletin of the American Mathematical Society*, 52(2):171–222, February 2015.

[Gra20] Andrew Granville. Sieving intervals and Siegel zeros. Preprint, arXiv: math/2010.01211, 2020.

[Had96] Jacques Hadamard. Sur la distribution des zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques. *Bulletin de la Société Mathématique de France*, 24:199–220, 1896.

[HB82] David R. Heath-Brown. Prime Numbers in Short Intervals and a Generalized Vaughan Identity. *Canadian Journal of Mathematics*, 34(6):1365–1377, 1982.

[HB83] David R. Heath-Brown. Prime twins and Siegel zeros. *Proc. Lond. Math.*, 47(3):193–224, 1983.

[Hil86] Adolf Hildebrand. On consecutive values of the Liouville function. *L'Enseignement Mathématique*, pages 219–226, 1986.

[Hil89] Adolf Hildebrand. Additive and multiplicative functions on shifted primes. *Proceedings of the London Mathematical Society*, s3-59(2):209–232, September 1989.

[HL23] George H. Hardy and John E. Littlewood. Some problems of 'Partitio numerorum'; III: On the expression of a number as a sum of primes. *Acta Mathematica*, 44:1 – 70, 1923.

[HR21] Harald A. Helfgott and Maksym Radziwiłł. Expansion, divisibility, and parity. Preprint, arXiv: math/2103.06853, 2021.

[Hux72] Martin N. Huxley. On the Difference Between Consecutive Primes. *Inventiones Mathematicae*, 15:164–170, 1971/72.

[IK04] Henryk Iwaniec and Emmanuel Kowalski. *Analytic Number Theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.

[Jut77] Matti Jutila. On Linnik's Constant. *Mathematica Scandinavica*, 41(1):45–62, 1977.

[Kou20] Dimitris Koukoulopoulos. *The Distribution of Prime Numbers*. Graduate Studies in Mathematics 203. American Mathematical Society, 2020.

[KPL20] Joanna Kułaga-Przymus and Mariusz Lemańczyk. Sarnak's conjectures from the ergodic theory point of view. To appear in Encyclopedia of Complexity and Systems Science, arXiv: math/2009.04757, 2020.

[Lin44a] Yuri V. Linnik. On the least prime in an arithmetic progression I. The basic theorem. *Rec. Math. (Mat. Sbornik)*, 15(57):139–178, 1944.

[Lin44b] Yuri V. Linnik. On the least prime in an arithmetic progression II. The Deuring-Heilbronn phenomenon. *Rec. Math. (Mat. Sbornik)*, 15(57):347–368, 1944.

[LT21] Jared Duker Lichtman and Joni Teräväinen. On the Hardy-Littlewood-Chowla conjecture on average. Preprint, arXiv: math/2111.08912, 2021.

[May15] James Maynard. Small gaps between primes. *Annals of Mathematics*, 181(1):383–413, January 2015.

[May19] James Maynard. On the twin prime conjecture. Preprint, arXiv: math/1910.14674, 2019.

[MM09] Helmut Maier and Hugh L. Montgomery. The sum of the Möbius function. *Bulletin of the London Mathematical Society*, 41(2):213–226, 2009.

[MM22] Kaisa Matomäki and Jori Merikoski. Siegel zeros, twin primes, Goldbach's conjecture, and primes in short intervals, 2022. Preprint, arXiv: math/2112.11412.

[Mot76] Yoichi Motohashi. On the sum of the Möbius function in a short segment. *Proc. Japan Acad.*, 52(9):477–479, 1976.

[MR15] Kaisa Matomäki and Maksym Radziwiłł. A note on the Liouville function in short intervals. Preprint, arXiv: math/1502.02374, 2015.

[MR16] Kaisa Matomäki and Maksym Radziwiłł. Multiplicative functions in short intervals. *Annals of Mathematics*, 183(3):1015–1056, 2016.

[MRT15] Kaisa Matomäki, Maksym Radziwiłł, and Terence Tao. An averaged form of Chowla's conjecture. *Algebra and Number Theory*, 9(9):2167–2196, 2015.

[MRT16] Kaisa Matomäki, Maksym Radziwiłł, and Terence Tao. Sign Patterns of the Liouville and Möbius Functions. *Forum of Mathematics, Sigma*, 4, 2016.

[MRT18] Kaisa Matomäki, Maksym Radziwiłł, and Terence Tao. Correlations of the von Mangoldt and higher divisor functions I. Long shift ranges. *Proceedings of the London Mathematical Society*, 118(2):284–350, July 2018.

[MT22] Kaisa Matomäki and Joni Teräväinen. On the Möbius function in all short intervals. *Journal of the European Mathematical Society*, January 2022.

[MV07] Hugh L. Montgomery and Robert C. Vaughan. *Multiplicative Number Theory I: Classical Theory*. Cambridge studies in advanced mathematics. Cambrigde University Press, 1st edition, 2007.

[MV17] M. Ram Murty and Akshaa Vatwani. Twin primes and the parity problem. *Journal of Number Theory*, 180:643–659, 2017.

[Pol14] D. H. J. Polymath. The "bounded gaps between primes" polymath project - a retrospective, 2014. Preprint, arXiv: math/1409.8361.

[Ram76] Kanakanahalli Ramachandra. Some problems of analytic number theory. *Acta Arithmetica*, 31(4):313–324, 1976.

[Rie59] Bernhard Riemann. Ueber die Anzahl der Primzahlen unter einer gegebenen Grösse. *Monatsberichte der Berliner Akademie*, 1859.

[Sar11] Peter Sarnak. Three lectures on the Möbius function, randomness and dynamics, 2011. Lecture notes, http://publications.ias.edu/sarnak/.

[Sch76] Wolfgang M. Schmidt. *Equations over Finite Fields An Elementary Approach*. Lecture Notes in Mathematics. Springer-Berlin-Heidelberg, 1st edition, 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.

[Sel47] Atle Selberg. On an elementary method in the theory of primes. *Norske Vid. Selsk. Forhdl.*, 19:64–67, 1947.

[Sou06] Kannan Soundararajan. Small gaps between prime numbers: The work of Goldston-Pintz-Yildirim, 2006. Preprint, arXiv: math/0605696.

[Sou16] Kannan Soundararajan. The Liouville function in short intervals [after Matomäki and Radziwiłł], 2016. Preprint, arXiv: math/1606.08021.

[Sou18] Kannan Soundararajan. Tao's resolution of the Erdős discrepancy problem. *Bulletin (New Series) of the American Mathematical Society*, 55(1):81–92, 2018.

[SS58] Andrzej Schinzel and Wacław Sierpiński. Sur certaines hypothèses concernant les nombres premiers. *Acta Arithmetica*, 4(3):185–208, 1958.

[Tao16a] Terence Tao. The Erdős discrepancy problem. *Discrete Analysis*, 1:1–27, 2016.

[Tao16b] Terence Tao. The logarithmically averaged Chowla and Elliott conjectures for two-point correlations. *Forum of Mathematics, Pi*, 4(8):1–36, 2016. doi:10.1017/fmp.2016.6.

[THB86] Edward C. Titchmarsh and David R. Heath-Brown. *The theory of the Riemann zeta-function*. Clarendon Press, Oxford, 1986.

[Tit32] Edward C. Titchmarsh. *The Theory of Functions*. Oxford University Press, Oxford, 1932.

[TT18] Terence Tao and Joni Teräväinen. Odd order cases of the logarithmically averaged Chowla conjecture. *Journal de Théorie de Nombres de Bordeaux*, 30(3):997–1015, 2018.

[TT19a] Terence Tao and Joni Teräväinen. Value patterns of multiplicative functions and related sequences. *Forum of Mathematics, Sigma*, 7:e33, 2019.

[TT19b] Terence Tao and Joni Teräväinen. The structure of logarithmically averaged correlations of multiplicative functions, with applications to the Chowla and Elliott conjectures. *Duke Math J.*, 168(11):1977–2027, 2019. doi:10.1215/00127094-2019-0002.

[TT21] Terence Tao and Joni Teräväinen. The Hardy-Littlewood-Chowla conjecture in the presence of a Siegel zero, 2021. Preprint, arXiv: math/2109.06291.

[Vau97] Robert C. Vaughan. *The Hardy-Littlewood Method*. Cambridge University Press, Cambridge, 1997.

[Zha14] Yitang Zhang. Bounded gaps between primes. *Annals of Mathematics*, 179(3):1121–1174, May 2014.

Original Publications

Article I

[Chi21a] Jake Chinis. On the Liouville Function in Short Intervals. *Int. Math. Res. Notices*, 04 2021

J. Chinis (2021) "On the Liouville Function in Short Intervals,"
International Mathematics Research Notices, Vol. 00, No. 0, pp. 1–17
doi:10.1093/imrn/rnab063

On the Liouville Function in Short Intervals

Jake Chinis*

Department of Mathematics and Statistics, McGill University, Montreal, QC H3A 2K6, Canada

**Correspondence to be sent to: e-mail: iakov.chinis@gmail.com*

Let λ denote the Liouville function. Assuming the Riemann Hypothesis, we prove that $\int_X^{2X} \left| \sum_{x \leq n \leq x+h} \lambda(n) \right|^2 \mathrm{d}x \ll Xh(\log X)^6$, as $X \to \infty$, provided $h = h(X) \leq \exp\left(\sqrt{\left(\frac{1}{2} - o(1)\right) \log X \log\log X}\right)$. The proof uses a simple variation of the methods developed by Matomäki and Radziwiłł in their work on multiplicative functions in short intervals, as well as some standard results concerning smooth numbers.

1 Introduction

Let λ denote the Liouville function; that is, the completely multiplicative function defined by $\lambda(p) := -1$, for all primes p. It is well known that the prime number theorem (PNT) is equivalent to the fact that λ exhibits some cancellation in its partial sums; more precisely, that

$$\sum_{n \leq x} \lambda(n) = o(x),$$

as $x \to \infty$.

The *Möbius Randomness Law* (see page 338 of [3]) tells us that $\{\lambda(n)\}_n$ should behave like a sequence of independent random variables, taking on the values ± 1 with equal probability. As a result, we expect "square-root cancellation" in the partial sums

Received August 10, 2020; Revised February 15, 2021; Accepted February 24, 2021
Communicated by Prof. Sound Soundarajan

for the Liouville function; that is, for any $\epsilon > 0$,

$$\sum_{n \leq x} \lambda(n) \ll x^{\frac{1}{2}+\epsilon},$$

as $x \to \infty$. In fact, the above estimate is equivalent to the Riemann Hypothesis (RH); see Theorem 14.25(C) of [16]. Furthermore, it is possible to quantity ϵ in terms of x; see [12].

The Liouville function also exhibits cancellation in short intervals, provided that the length of the interval is sufficiently large. Motohashi [5] and Ramachandra [10], independently, proved that

$$\sum_{x \leq n \leq x+h} \lambda(n) = o(h),$$

provided $h > x^{\frac{7}{12}+\epsilon}$. Assuming RH, this can be improved to $h > x^{\frac{1}{2}}(\log x)^A$, for some suitable constant $A > 0$; see [4]. Recently, Matomäki and Teräväinen [8] improved Motohashi and Ramachandra's results to include the larger range $h > x^{\frac{11}{20}+\epsilon}$.

By relaxing the condition that our estimates hold for all short intervals, we can get results that hold for smaller values of h. In an unpublished work by Gao, it is shown that

$$\int_X^{2X} \Big| \sum_{x \leq n \leq x+h} \lambda(n) \Big|^2 \mathrm{d}x \ll Xh(\log X)^A,$$

assuming RH, for some large constant $A > 0$. Strictly speaking, Gao's result and those stated above his were initially proven for the Möbius function, μ, but the proofs extend to the Liouville function with little effort.

The preceding results should be compared with what is known for primes in short intervals. More precisely, an equivalent form of the PNT is given by

$$\sum_{n \leq x} \Lambda(n) = x + o(x),$$

where Λ denotes the von Mangoldt function (defined to be $\log p$ if n is a power of the prime p and 0 otherwise). Furthermore, RH is equivalent to the following estimate:

$$\sum_{n \leq x} \Lambda(n) = x + O(x^{\frac{1}{2}}(\log x)^2);$$

in particular, RH implies that

$$\sum_{x\leq n\leq x+h}\Lambda(n)=h+o(h),$$

provided $h>x^{\frac{1}{2}}(\log x)^{2+\epsilon}$.

Again restricting ourselves to results that hold almost everywhere, Selberg [11] proved that if RH holds, then

$$\int_X^{2X}\Big|\sum_{x\leq n\leq x+h}\Lambda(n)-h\Big|^2\mathrm{d}x\ll Xh(\log X)^2,$$

so that almost all short intervals contain the correct number of primes.

It is important to note that both Gao and Selberg obtain square-root cancellation in their estimates. In each case, they relate the short sum to a contour integral of some Dirichlet series and then shift this contour to the edge of the critical strip, picking up any poles along the way. For Selberg, the corresponding Dirichlet series is $\zeta'(s)/\zeta(s)$ and, assuming RH, the only pole will be at $s=1$. Near the half-line, we still need to consider the contributions from the non-trivial zeros, $\rho=1/2+i\gamma$, of $\zeta(s)$, as $\zeta'(s)/\zeta(s)$ is large for $s\approx\rho$. Fortunately, the behaviour of $\zeta'(s)/\zeta(s)$ for $s\approx\rho$ is well understood (it is even conjectured that the zeros of $\zeta(s)$ are simple). For Gao, the corresponding Dirichlet series is $\zeta(2s)/\zeta(s)$, which also has poles at $s=\rho$, but the residues in this case are equal to $\zeta(2\rho)/\zeta'(\rho)$. Since very little is known about $1/\zeta'(\rho)$, we need to proceed in an indirect manner. The key idea is to use a sum over primes to approximate $\zeta(s)$ and then avoid "clusters" of zeros of $\zeta(s)$ on the half-line: near these regions, $1/\zeta(s)$ is large and the contour is chosen so that $1/\zeta(s)$ is not too large, at least in some sense; see [4], to get an idea of Gao's proof.

We now turn our attention to the breakthrough work by Matomäki and Radziwiłł, where they relate the average value of 1-bounded multiplicative functions in short intervals to the corresponding average value in large intervals:

Theorem 1.1 ([7]). Let $f:\mathbb{N}\to[-1,1]$ be a multiplicative function and let $h=h(X)\to\infty$ arbitrarily slowly as $X\to\infty$. Then, for almost all $x\in[X,2X]$,

$$\frac{1}{h}\sum_{x\leq n\leq x+h}f(n)=\frac{1}{X}\sum_{X\leq n\leq 2X}f(n)+o(1),$$

with $o(1)$ not depending on f.

In the case of the Liouville function, Theorem 1.1 implies that

$$\sum_{x\leq n\leq x+h}\lambda(n)=o(h),$$

for almost all $x\in[X,2X]$; in particular,

$$\int_X^{2X}\Big|\sum_{x\leq n\leq x+h}\lambda(n)\Big|^2\mathrm{d}x=o(Xh^2),$$

for any $h=h(X)\to\infty$ as $X\to\infty$. Although Matomäki and Radziwiłł do not get square-root cancellation, their results are unconditional, hold for all h going to infinity, and avoid the complex analytic approach used by both Gao and Selberg; instead, they employ a clever decomposition of the corresponding Dirichlet polynomials, which is done by restricting to integers that have prime factors from certain convenient ranges. For a detailed account of the work in [7] restricted to the Liouville function, see [13].

In this paper, we apply a simple variation of the methods developed in [7] in order to prove the following:

Theorem 1.2. Assume RH. Then,

$$\int_X^{2X}\Big|\sum_{x\leq n\leq x+h}\lambda(n)\Big|^2\mathrm{d}x\ll Xh(\log X)^6,$$

provided $h=h(X)\leq\exp\left(\sqrt{\left(\frac{1}{2}-o(1)\right)\log X\log\log X}\right)$.

Note that Theorem 1.2 shows square-root cancellation for the Liouville function in almost all short intervals, provided $(\log X)^{6+\epsilon}\leq h\leq\exp\left(\sqrt{\left(\frac{1}{2}-o(1)\right)\log X\log\log X}\right)$. Of course, Theorem 1.2 gives an upper bound on the exceptional set of $x\in[X,2X]$ for which square-root cancellation does not hold (via Chebyshev's Inequality). By splitting the short sum into shorter intervals, Theorem 1.2 can also be used to obtain some cancellation for larger h; namely,

Corollary 1.3. Assume RH and suppose $h\leq X$. Then,

$$\int_X^{2X}\Big|\sum_{x\leq n\leq x+h}\lambda(n)\Big|^2\mathrm{d}x\ll Xh^2(\log X)^6\left(\frac{1}{h}+\frac{1}{H}\right),$$

where $H=H(X):=\exp\left(\sqrt{\left(\frac{1}{2}-o(1)\right)\log X\log\log X}\right)$.

2 Preliminaries

We present here a collection of standard results, which we use freely throughout our paper. The 1st result is a quantitative version of Perron's Formula, which serves as an approximation to the indicator function on $(1,\infty)$. The 2nd result relates the mean-square of a Dirichlet polynomial to its sum of squares and this will be our main tool in proving Theorem 1.2.

Lemma 2.1 (Quantitative version of Perron's Formula). Fix $\kappa > 0$. Then,

$$\frac{1}{2\pi i}\int_{\kappa-iT}^{\kappa+iT}\frac{y^s}{s}\mathrm{d}s = \begin{cases} 1 & \text{if } y>1 \\ \frac{1}{2} & \text{if } y=1 \\ 0 & \text{if } y<1 \end{cases} + \mathcal{O}\Big(\frac{y^\kappa}{\max\{1,T|\log y|\}}\Big),$$

uniformly for both $y > 0$ and $T > 0$.

Proof. See [15, Theorem II.2.3]. ■

Lemma 2.2 (Mean value theorem). For any sequence of complex numbers $\{a_n\}_n$, we have that

$$\int_0^T \Big|\sum_{1\le n\le N} a_n n^{it}\Big|^2 \mathrm{d}t = (T+\mathcal{O}(N))\sum_{1\le n\le N}|a_n|^2.$$

Proof. See [3, Theorem 9.1] ■

Remark 2.1. Notice that the main term in Lemma 2.2 corresponds to the contribution from the diagonal terms (as seen by expanding the square and integrating). Notice further that the mean value theorem (MVT) is exceptionally powerful when $N = o(T)$: in this case, the integral is bounded above by the contribution from the diagonal terms and this is best possible.

In their work on multiplicative functions in short intervals, Matomäki and Radziwiłł use the MVT in the range $T = o(N)$, which, together with the decomposition of their Dirichlet polynomials, provides a small saving in most cases; see Section 2.1 of [7]. In this paper, we are interested in the opposite range and decompose our Dirichlet polynomials into two sums, one of which has length $N = o(T)$ and another that can be handled trivially (at least for small h).

We need two more preliminary results. The 1st is a pointwise bound on $\sum_{X\leq n\leq 2X}\lambda(n)/n^{\frac{1}{2}+it}$, which allows us to remove the contribution from the small values of t in our average value (this can be thought of as the analogous result to Lemma 1 in [6]). The 2nd is an analogue of Lemma 2 in [7]: we need a pointwise bound on sums of the form $\sum_{P\leq p\leq 2P}1/p^{\frac{1}{2}+\frac{1}{\log X}+it}$, for large values of t.

Lemma 2.3. Assume RH. Then,

$$\sum_{X\leq n\leq 2X}\frac{\lambda(n)}{n^{\frac{1}{2}+it}}\ll_{\epsilon}(1+|t|)^{\epsilon}X^{\epsilon},$$

for all $\epsilon>0$ and all $t\in\mathbb{R}$.

Proof. By Lemma 3.12 of [16], we have that

$$\sum_{X\leq n\leq 2X}\frac{\lambda(n)}{n^{\frac{1}{2}+it}}=\frac{1}{2\pi i}\int_{2-iT}^{2+iT}\frac{\zeta(2s+1+2it)}{\zeta(s+\frac{1}{2}+it)}\frac{(2X)^{s}-X^{s}}{s}\mathrm{d}s+\mathcal{O}\Big(\frac{X^{2}}{T}\Big).$$

Given $\epsilon>0$ and assuming RH, the function $\zeta(2(s+\frac{1}{2}+it))/\zeta(s+\frac{1}{2}+it)$ is analytic for $\Re e(s)\geq\epsilon$. Shifting the contour to the edge of this region, we get that

$$\sum_{X\leq n\leq 2X}\frac{\lambda(n)}{n^{\frac{1}{2}+it}}=\frac{-1}{2\pi i}\Big(\int_{2+iT}^{\epsilon+iT}+\int_{\epsilon+iT}^{\epsilon-iT}+\int_{\epsilon-iT}^{2-iT}\Big)\Big(\frac{\zeta(2s+1+2it)}{\zeta(s+\frac{1}{2}+it)}\frac{(2X)^{s}-X^{s}}{s}\mathrm{d}s\Big)+\mathcal{O}\Big(\frac{X^{2}}{T}\Big).$$

Then, using the facts that $\zeta(s)\ll 1/(\Re(s)-1)$, for $\Re(s)>1$, and $1/\zeta(s)\ll|\Im(s)|^{\epsilon}$, for $\Re(s)\geq\frac{1}{2}+\epsilon$, (see 14.2.6 in [16]), we have that

$$\begin{aligned}\sum_{X\leq n\leq 2X}\frac{\lambda(n)}{n^{\frac{1}{2}+it}}&\ll_{\epsilon}(|t|+T)^{\epsilon}X^{\epsilon}\int_{-T}^{T}\frac{\mathrm{d}y}{\sqrt{\epsilon^{2}+y^{2}}}+(|t|+T)^{\epsilon}\frac{X^{2}}{T}\\&\ll_{\epsilon}(|t|+T)^{\epsilon}(X^{\epsilon}\log T+\frac{X^{2}}{T}).\end{aligned}$$

Taking $T=X^{2}$, this boils down to

$$\sum_{X\leq n\leq 2X}\frac{\lambda(n)}{n^{\frac{1}{2}+it}}\ll_{\epsilon}(|t|+X^{2})^{\epsilon}X^{\epsilon}\ll_{\epsilon}(1+|t|)^{\epsilon}X^{\epsilon},$$

as claimed. ∎

Lemma 2.4. Assume RH and suppose $P \leq X$. Then,

$$\sum_{P\leq p\leq 2P} \frac{1}{p^{\frac{1}{2}+\frac{1}{\log X}+it}} \ll \log X,$$

uniformly for $X^{\frac{1}{2}} \leq |t| \leq X$.

Proof. Note that

$$\sum_{P\leq p\leq 2P} \frac{1}{p^{\frac{1}{2}+\frac{1}{\log X}+it}} = \int_{1/\log X}^{1/2} \sum_{P\leq n\leq 2P} \frac{\Lambda(n)}{n^{\frac{1}{2}+\alpha+it}} \mathrm{d}\alpha + \mathcal{O}(1). \tag{1}$$

Applying Perron's Formula (Lemma 2.1) to the integrand, with $\kappa := 1/2-\alpha+1/\log X > 0$ and $T := P^{\frac{1}{2}}/2$, we have that

$$\sum_{P\leq n\leq 2P} \frac{\Lambda(n)}{n^{\frac{1}{2}+\alpha+it}} = \frac{1}{2\pi i}\int_{1/2-\alpha+1/\log X-iT}^{1/2-\alpha+1/\log X+iT} -\frac{\zeta'}{\zeta}(s+\frac{1}{2}+\alpha+it)\frac{(2P)^s-P^s}{s}\mathrm{d}s + \mathcal{O}\left(P^{-\alpha}\log X\right).$$

A standard contour shift argument, using Chapters 12 and 13 of [9], shows that this last expression is $\ll P^{-\alpha}(\log P)\log X$. Upon inserting this bound back into (1), and integrating over α, we get the desired result. Below, we provide some details for convenience to the reader.

Assuming RH and shifting the contour to the line $\Re(s) = -3/2-\alpha =: \sigma_0$, we have that

$$\sum_{P\leq n\leq 2P} \frac{\Lambda(n)}{n^{\frac{1}{2}+\alpha+it}} = \frac{-1}{2\pi i}\left(\int_{\kappa+iT}^{\sigma_0+iT} + \int_{\sigma_0+iT}^{\sigma_0-iT} + \int_{\sigma_0-iT}^{\kappa-iT}\right)\left(-\frac{\zeta'}{\zeta}(s+\frac{1}{2}+\alpha+it)\frac{(2P)^s-P^s}{s}\mathrm{d}s\right)$$
$$+ \sum_{\substack{\rho \\ |\gamma-t|\leq T}} \frac{(2P)^{-\alpha-it+i\gamma}-P^{-\alpha-it+i\gamma}}{-\alpha-it+i\gamma} + \mathcal{O}\left(P^{-\alpha}\log X\right),$$

where the sum is over the non-trivial zeros, $\rho = 1/2+i\gamma$, of ζ (counted with multiplicity). Notice that the sum over ρ corresponds to the residues of the integrand at $s = \rho-1/2-\alpha-it$. Notice further that we do not pick up the pole of ζ at 1, as $\Im(s)+t$ is bounded away from 0 (recall that $X^{\frac{1}{2}} \leq |t| \leq X$ with $T = P^{\frac{1}{2}}/2$ and $P \leq X$). Note: strictly speaking, we need to choose T in such a way that the horizontal line segments in the contour (shifted by t) are bounded away from ordinates γ of ζ; this can be done via Lemma 12.2 of [9].

To deal with the vertical integral, we use the fact that $\zeta'(s)/\zeta(s) \ll \log(|s|+1)$ uniformly for $\Re(s) \leq -1$, provided s is not close to an even integer (i.e., provided we are

not close to a trivial zero of ζ); see Lemma 12.4 of [9]. Using the above and bounding the vertical integral trivially produces an admissible error.

For the horizontal integrals, write $s = \sigma \pm iT$ and note that

$$\left|\int_{\sigma_0 \pm iT}^{\kappa \pm iT} \frac{\zeta'}{\zeta}(\sigma + \frac{1}{2} + \alpha + i(t \pm T)) \frac{(2P)^{\sigma \pm iT} - P^{\sigma \pm iT}}{\sigma \pm iT} \mathrm{d}\sigma\right| \ll \frac{P^{\frac{1}{2}-\alpha}}{T} \int_{-1}^{1+1/\log X} \left|\frac{\zeta'}{\zeta}(\sigma + i(t \pm T))\right| \mathrm{d}\sigma$$
$$\ll P^{-\alpha} \int_{-1}^{1+1/\log X} \left|\frac{\zeta'}{\zeta}(\sigma + i(t \pm T))\right| \mathrm{d}\sigma,$$

which follows by making the change of variables $\sigma \mapsto \sigma - 1/2 - \alpha$, taking a pointwise bound on the integrand, and recalling that $\kappa = 1/2 - \alpha - it$, $\sigma_0 = -3/2 - \alpha$, and $T = P^{\frac{1}{2}}/2$. From the functional equation, it suffices to consider the integral over $\sigma \in [1/2, 1 + 1/\log X]$ alone, which we break into three intervals, according to the bounds one can get for ζ'/ζ:

1. By Lemma 12.2 of [9], we know that $\frac{\zeta'}{\zeta}(\sigma + i(t \pm T)) \ll (\log X)^2$, uniformly for $-1 \leq \sigma \leq 2$; in particular,

$$P^{-\alpha} \int_{1/2}^{1/2+1/\log X} \left|\frac{\zeta'}{\zeta}(\sigma + i(t \pm T))\right| \mathrm{d}\sigma \ll P^{-\alpha} \log X.$$

2. By Lemma 13.20 of [9], we know that

$$\frac{\zeta'}{\zeta}(\sigma + i(t \pm T)) = \sum_{\substack{\rho \\ |\gamma - (t \pm T)| \leq 1/\log\log((|t \pm T|+4)}} \frac{1}{\sigma - 1/2 + i(t \pm T - \gamma)} + \mathcal{O}(\log X),$$

uniformly for $|\sigma - 1/2| \leq 1/\log\log(|t \pm T| + 4)$. Using the trivial bound

$$\frac{1}{|\sigma - 1/2 + i(t \pm T - \gamma)|} \leq \frac{1}{|\sigma - 1/2|},$$

and integrating over σ, we then have that

$$P^{-\alpha} \int_{1/2+1/\log X}^{1/2+1/\log\log(|t \pm T|+4)} \left|\frac{\zeta'}{\zeta}(\sigma + i(t \pm T))\right| \mathrm{d}\sigma \ll P^{-\alpha} \log X,$$

which follows from the fact that the integral over σ is bounded by $\log\log X$ and from the fact that there are $\ll \log X/\log\log X$ zeros in each window of length $1/\log\log X$ (see Lemma 13.19 of [9]).

3. Finally, for $1/2 + 1/\log\log(|t \pm T| + 4) \le \sigma \le 1 + 1/\log X$, we use the fact that

$$\frac{\zeta'}{\zeta}(\sigma + i(t \pm T)) \ll (\log X)^{2-2\sigma} \log\log X,$$

uniformly for σ in the desired range (see Corollary 13.14 of [9]), which yields

$$P^{-\alpha}\int_{1/2+1/\log\log(|t\pm T|+4)}^{1+1/\log X}\left|\frac{\zeta'}{\zeta}(\sigma+i(t\pm T))\right| d\sigma \ll P^{-\alpha}(\log\log X)\int_{1/2+1/\log\log(|t\pm T|+4)}^{1+1/\log X}(\log X)^{2-2\sigma} d\sigma$$
$$\ll P^{-\alpha}\log X.$$

In other words, a standard contour shift yields the following:

$$\begin{aligned}\sum_{P\le n\le 2P}\frac{\Lambda(n)}{n^{\frac{1}{2}+\alpha+it}} &\ll P^{-\alpha}\sum_{\substack{\rho\\ |\gamma-t|\le P^{\frac{1}{2}}/2}}\left|\frac{2^{-\alpha-it+i\gamma}-1}{-\alpha-it+i\gamma}\right| + P^{-\alpha}\log X\\ &\ll P^{-\alpha}\sum_{\substack{\rho\\ |\gamma-t|\le P^{\frac{1}{2}}/2}}\frac{1}{1+|t-\gamma|} + P^{-\alpha}\log X\\ &\ll P^{-\alpha}(\log P)\log X,\end{aligned}$$

where the last line follows from the fact that there are $\ll \log X$ zeros in each interval of length 1 (counted with multiplicity) and recalling that the sum over ρ comes from the contribution of the residues in our contour integral. Upon substituting the last bound into (1), and integrating over α, we obtain the desired result. ■

Remark 2.2. The proof of Theorem 1.2 can easily be adapted to a more general setting. For an arbitrary multiplicative function f, all we need is an analogue to Lemma 2.4. Essentially, we are looking for square-root cancellation in the corresponding sum over primes:

$$\sum_{P\le p\le 2P} f(p)p^{it} \ll P^{\epsilon}\Big(\sum_{P\le p\le 2P}\big(f(p)\big)^2\Big)^{\frac{1}{2}}.$$

For example, the above estimate is known to hold, assuming RH, for coefficients of automorphic forms and for multiplicative functions of the form $\mu(n)\lambda_\pi(n)$ or $\lambda_\pi(n)$, where μ is the Möbius function and where the $\lambda_\pi(n)$'s are the coefficients of an automorphic representation π.

3 Initial Reductions

In this section, we reduce our problem to that of bounding the mean square of a Dirichlet polynomial. We begin with the following standard lemma, which essentially follows from Perron's Formula (together with a few other tricks):

Lemma 3.1 (Plancherel). For $1 \leq h \leq X$,

$$\frac{1}{X}\int_{X}^{2X}\Big|\frac{1}{h}\sum_{x\leq n\leq x+h}\lambda(n)\Big|^2\mathrm{d}x \ll \frac{1}{X}\int_{0}^{X/h}\Big|\sum_{X\leq n\leq 4X}\frac{\lambda(n)}{n^{\frac{1}{2}+it}}\Big|^2\mathrm{d}t + \max_{T>X/h}\frac{1}{hT}\int_{T}^{2T}\Big|\sum_{X\leq n\leq 4X}\frac{\lambda(n)}{n^{\frac{1}{2}+it}}\Big|^2\mathrm{d}t.$$

Proof. See [7, Lemma 14] (equivalently, [6, Lemma 4]). ■

Using Lemma 2.3, we can remove the contribution from the small values of t in the RHS of Lemma 3.1: for any $\epsilon > 0$,

$$\frac{1}{X}\int_{0}^{X^{\frac{1}{2}}}\Big|\sum_{X\leq n\leq 4X}\frac{\lambda(n)}{n^{\frac{1}{2}+it}}\Big|^2\mathrm{d}t \ll_{\epsilon} X^{\epsilon-\frac{1}{2}},$$

which follows by bounding the integrand pointwise. Thus, Theorem 1.2 follows from Lemma 3.1 once we show that

$$\frac{1}{hT}\int_{X^{\frac{1}{2}}}^{T}\Big|\sum_{X\leq n\leq 4X}\frac{\lambda(n)}{n^{\frac{1}{2}+it}}\Big|^2\mathrm{d}t \ll \frac{(\log X)^6}{h}, \tag{2}$$

for $T \geq X/h$, where we have assumed, w.l.o.g., that $h \leq X^{\frac{1}{2}}$ (this is done to avoid the degenerate case when $T = X/h \leq X^{1/2}$).

For $T > X$, the MVT (Lemma 2.2) immediately gives the desired bound:

$$\frac{1}{hT}\int_{X^{\frac{1}{2}}}^{T}\Big|\sum_{X\leq n\leq 4X}\frac{\lambda(n)}{n^{\frac{1}{2}+it}}\Big|^2\mathrm{d}t \ll \frac{1}{hT}(T+X)\sum_{X\leq n\leq 4X}\frac{1}{n} \ll \frac{1}{h}.$$

It remains to prove that (2) holds for $X/h \leq T \leq X$. Our main tool is still the MVT, but this is only winning if we can shorten the length our sum: we should think of $T \approx X/h$ and recall that the MVT is best possible if the length of the corresponding Dirichlet polynomial is of size $N \approx T$; see Remark 2.1. Since our Dirichlet polynomial has length X, our goal is to split the sum over $X \leq n \leq 4X$ into two sums, one of which has length $\approx X/h$ and another that can be handled separately. This is done by considering the integers $X \leq n \leq 4X$, which have a prime factor $p > h$: by writing such integers as $n = pm$, and then factoring out the sum over p, the sum over $m \approx X/p$ that remains

has length $\ll X/h$; using the MVT on the sum over m is now winning. In Section 5, we deal with the remaining integers, all of whose primes factors are $\leq h$. Fortunately for us, there are few of these so-called h-smooth integers, at least for small h, so that the MVT can be applied directly to give us what we want. Getting square-root cancellation for larger h will require some new ideas; this is an ongoing work.

4 Integers With Large Prime Factors

For the integers $X \leq n \leq 4X$, which have at least one prime factor $p > h$, we have the following:

Proposition 4.1. Assume RH. If $X/h \leq T \leq X$, then

$$\frac{1}{hT}\int_{X^{\frac{1}{2}}}^{T}\Big|\sum_{\substack{X\leq n\leq 4X\\ \exists p>h:p|n}}\frac{\lambda(n)}{n^{\frac{1}{2}+it}}\Big|^2\mathrm{d}t \ll \frac{(\log X)^6}{h}.$$

Proof. Note that

$$\sum_{\substack{X\leq n\leq 4X\\ \exists p|n:p>h}}\frac{\lambda(n)}{n^{\frac{1}{2}+it}} = \sum_{h<p\leq 4X}\frac{\lambda(p)}{p^{\frac{1}{2}+it}}\sum_{X/p\leq m\leq 4X/p}\frac{1}{m^{\frac{1}{2}+it}}\frac{\lambda(m)}{\#\{q \text{ prime} : q>h, q|m\}+\mathbb{1}_{p\nmid m}},$$

where $\#\{q \text{ prime} : q>h, q|m\}+\mathbb{1}_{p\nmid m}$ counts the number of ways $X \leq n \leq 4X$ can be factored as $n=mp$ with $p>h$. In the published version of [7], the term $\mathbb{1}_{p\nmid m}$ appears as the constant 1, but this was corrected in later versions of their paper. In any case, we will see that this misprint does not affect their argument. As Matomäki and Radziwiłł remark in [7], the above decomposition is analogous to Buchstab's identity, which is a variant of Ramaré's identity; see Section 17.3 of [1].

Our goal now is to remove the dependence on $\mathbb{1}_{p\nmid m}$, which is done by splitting the inner sum into those m for which $p\nmid m$ and $p|m$, respectively:

$$\begin{aligned}\sum_{h<p\leq 4X}\frac{\lambda(p)}{p^{\frac{1}{2}+it}}\sum_{X/p\leq m\leq 4X/p}\frac{\lambda(m)}{m^{\frac{1}{2}+it}}\frac{1}{\#\{q>h : q|m\}+\mathbb{1}_{p\nmid m}}&\\ = \sum_{h<p\leq 4X}\frac{\lambda(p)}{p^{\frac{1}{2}+it}}\sum_{\substack{X/p\leq m\leq 4X/p\\ p\nmid m}}\frac{1}{m^{\frac{1}{2}+it}}\frac{\lambda(m)}{\#\{q>h : q|m\}+1}&\\ + \sum_{h<p\leq 4X}\frac{\lambda(p)}{p^{\frac{1}{2}+it}}\sum_{\substack{X/p\leq m\leq 4X/p\\ p|m}}\frac{1}{m^{\frac{1}{2}+it}}\frac{\lambda(m)}{\#\{q>h : q|m\}},&\end{aligned}$$

with q varying over the set of primes. This can be simplified further by adding and subtracting all m for which $p|m$ to the 1st term and setting

$$a_m := \frac{-\lambda(m)}{\#\{q > h : q|m\} + 1}, \; b_m := \frac{-\lambda(m)}{\#\{q > h : q|m\}(\#\{q > h : q|m\} + 1)},$$

which yields

$$\begin{aligned}\sum_{h<p\leq 4X} \frac{-1}{p^{\frac{1}{2}+it}} \sum_{X/p\leq m\leq 4X/p} \frac{\lambda(m)}{m^{\frac{1}{2}+it}} \frac{1}{\#\{q > h : q|m\} + \mathbb{1}_{p\nmid m}} &= \sum_{h<p\leq 4X} \frac{1}{p^{\frac{1}{2}+it}} \sum_{X/p\leq m\leq 4X/p} \frac{a_m}{m^{\frac{1}{2}+it}} + \sum_{h<p\leq 4X} \frac{1}{p^{\frac{1}{2}+it}} \sum_{\substack{X/p\leq m\leq 4X/p \\ p|m}} \frac{b_m}{m^{\frac{1}{2}+it}} \\ &= \sum_{h<p\leq 4X} \frac{1}{p^{\frac{1}{2}+it}} \sum_{X/p\leq m\leq 4X/p} \frac{a_m}{m^{\frac{1}{2}+it}} + \sum_{h<p\leq 4X} \frac{1}{p^{1+2it}} \sum_{X/p^2\leq m\leq 4X/p^2} \frac{b_{mp}}{m^{\frac{1}{2}+it}},\end{aligned}$$

and where the last line follows by writing $m = mp$ in the 2nd double sum. Thus,

$$\begin{aligned}\frac{1}{hT}\int_{X^{\frac{1}{2}}}^{T} \Big| \sum_{\substack{X\leq n\leq 4X \\ \exists p|n : p>h}} \frac{\lambda(n)}{n^{\frac{1}{2}+it}} \Big|^2 \mathrm{d}t &\ll \frac{1}{hT}\int_{X^{\frac{1}{2}}}^{T} \Big| \sum_{h<p\leq 4X} \frac{1}{p^{\frac{1}{2}+it}} \sum_{X/p\leq m\leq 4X/p} \frac{a_m}{m^{\frac{1}{2}+it}} \Big|^2 \mathrm{d}t \\ &\quad + \frac{1}{hT}\int_{X^{\frac{1}{2}}}^{T} \Big| \sum_{h<p\leq 4X} \frac{1}{p^{1+2it}} \sum_{X/p^2\leq m\leq 4X/p^2} \frac{b_{mp}}{m^{\frac{1}{2}+it}} \Big|^2 \mathrm{d}t.\end{aligned}$$

Applying the MVT (Lemma 2.2) to the 2nd integral, we see that

$$\begin{aligned}&\frac{1}{hT}\int_{X^{\frac{1}{2}}}^{T} \Big| \sum_{h<p<4X} \frac{1}{p^{1+2it}} \sum_{X/p^2\leq m\leq 4X/p^2} \frac{b_{mp}}{m^{\frac{1}{2}+it}} \Big|^2 \mathrm{d}t \\ &\qquad \ll \frac{1}{hT}(T+X) \sum_{h<p<4X} \frac{1}{p^2} \sum_{X/p^2\leq m\leq 4X/p^2} \frac{1}{m} \ll \frac{1}{h},\end{aligned}$$

recalling that $X/h \leq T \leq X$.

For the remaining integral, we wish to separate the variables p and m, so that we may apply a pointwise bound to the sum over p. We begin by splitting the sum over

p into dyadic intervals

$$\begin{aligned}\frac{1}{hT}\int_{X^{\frac{1}{2}}}^{T}\Big|\sum_{h<p\le 4X}\frac{1}{p^{\frac{1}{2}+it}}\sum_{X/p\le m\le 4X/p}\frac{a_m}{m^{\frac{1}{2}+it}}\Big|^2\mathrm{d}t&\\ =\frac{1}{hT}\int_{X^{\frac{1}{2}}}^{T}\Big|\sum_{j=\lfloor\log_2 h\rfloor}^{\lfloor\log_2 4X\rfloor-1}\sum_{2^j<p\le 2^{j+1}}\frac{1}{p^{\frac{1}{2}+it}}\sum_{\substack{X/2^{j+1}\le m\le X/2^{j-1}\\ X/p\le m\le 4X/p}}\frac{a_m}{m^{\frac{1}{2}+it}}\Big|^2\mathrm{d}t&\\ \ll\frac{(\log X)^2}{hT}\max_{\log_2 h\le j\le\log_2(4X)}\int_{X^{\frac{1}{2}}}^{T}\Big|\sum_{2^j<p\le 2^{j+1}}\frac{1}{p^{\frac{1}{2}+it}}\sum_{\substack{X/2^{j+1}\le m\le X/2^{j-1}\\ X/p\le m\le 4X/p}}\frac{a_m}{m^{\frac{1}{2}+it}}\Big|^2\mathrm{d}t,&\end{aligned}$$

where $\log_2$ denotes the base 2 logarithm function and where the last line follows by taking the absolute value inside the sum over j, noting that there are $\ll \log X/h \le \log X$ such dyadic intervals.

Using Lemma 2.1 with $\kappa = 1/\log X$, we can remove the condition that $X \le mp \le 4X$:

$$\mathbb{1}_{X\le mp\le 4X}=\frac{1}{2\pi i}\int_{\kappa-iY}^{\kappa+iY}\frac{(4X)^s-X^s}{(mp)^s}\frac{\mathrm{d}s}{s}+\mathcal{O}\Big(\frac{1/(mp)^{\kappa}}{\max\{1,Y|\log X/mp|\}}\Big),$$

which yields

$$\begin{aligned}\frac{1}{hT}\int_{X^{\frac{1}{2}}}^{T}\Big|\sum_{h<p\le 4X}\frac{1}{p^{\frac{1}{2}+it}}\sum_{X/p\le m\le 4X/p}\frac{a_m}{m^{\frac{1}{2}+it}}\Big|^2\mathrm{d}t&\\ \ll\frac{(\log X)^2}{hT}\int_{X^{\frac{1}{2}}}^{T}\Big|\int_{\kappa-iY}^{\kappa+iY}\frac{(4X)^s-X^s}{s}\sum_{2^j<p\le 2^{j+1}}\frac{1}{p^{s+\frac{1}{2}+it}}\sum_{X/2^{j+1}\le m\le X/2^{j-1}}\frac{a_m}{m^{s+\frac{1}{2}+it}}\mathrm{d}s\Big|^2\mathrm{d}t&\\ +\frac{(\log X)^2}{h}\Big(\frac{X^{\frac{1}{2}}\log X}{Y}\Big)^2,&\end{aligned}$$

for some $\log_2 h \le j \le \log_2(4X)$

We can now use Minkowski's Inequality for integrals (see Section A.1 of [14]) to change the order of integration:

$$\begin{aligned}\int_{X^{\frac{1}{2}}}^{T}\Big|\int_{\kappa-iY}^{\kappa+iY}\frac{(4X)^s-X^s}{s}\sum_{2^j<p\le 2^{j+1}}\frac{1}{p^{s+\frac{1}{2}+it}}\sum_{X/2^{j+1}\le m\le X/2^{j-1}}\frac{a_m}{m^{s+\frac{1}{2}+it}}\mathrm{d}s\Big|^2\mathrm{d}t&\\ \ll\Big(\int_{\kappa-iY}^{\kappa+iY}\Big(\int_{X^{\frac{1}{2}}}^{T}\Big|\frac{(4X)^s-X^s}{s}\sum_{2^j<p\le 2^{j+1}}\frac{1}{p^{s+\frac{1}{2}+it}}\sum_{X/2^{j+1}\le m\le X/2^{j-1}}\frac{a_m}{m^{s+\frac{1}{2}+it}}\Big|^2\mathrm{d}t\Big)^{\frac{1}{2}}\mathrm{d}s\Big)^2.&\end{aligned}$$

Then, by taking $Y = X^{\frac{1}{2}}/2$, we can apply Lemma 2.4 to bound the sum over p; this yields the upper bound

$$\begin{aligned}
&\ll (\log X)^2\Bigg(\int_{\kappa-iY}^{\kappa+iY}\frac{1}{|s|}\Bigg(\int_{X^{\frac{1}{2}}}^{T}\Big|\sum_{X/2^{j+1}\leq m\leq X/2^{j-1}}\frac{a_m}{m^{s+\frac{1}{2}+it}}\Big|^2\mathrm{d}t\Bigg)^{\frac{1}{2}}\mathrm{d}s\Bigg)^2\\
&\ll (\log X)^4(T+X/h)\sum_{X/2^{j+1}\leq m\leq X/2^{j-1}}\frac{1}{m}\\
&\ll (\log X)^4(T+X/h),
\end{aligned}$$

where the 2nd to the last line follows from the MVT, recalling that $j \geq \log h/\log 2$, and where the additional powers of $\log X$ come from the bound

$$\int_{\kappa-iY}^{\kappa+iY}\frac{1}{|s|}\mathrm{d}s \ll \log X,$$

recalling that $\kappa = 1/\log X$ and $Y = X^{\frac{1}{2}}/2$.

Therefore,

$$\frac{1}{hT}\int_{X^{\frac{1}{2}}}^{T}\Big|\sum_{\substack{X\leq n\leq 4X\\ \exists p>h: p|n}}\frac{\lambda(n)}{n^{\frac{1}{2}+it}}\Big|^2\mathrm{d}t \ll \frac{(\log X)^6}{h},$$

provided that $X/h \leq T \leq X$, which is the desired result. ■

Remark 4.1. In the proof of Proposition 4.1, we used a contour integral to separate the variables p and m, which lost two factors of $\log X$. As the referee remarked, introducing a smoothing at the beginning of our argument makes the separation of variables lossless: after Lemma 3.1 (Plancherel), we are working with smooth sums and we can use the rapid decay of the Mellin Transform to ensure that the integral of $1/|s|$, which produced the original factors of $\log X$, converges. This would yield the upper bound $\ll Xh(\log X)^4$ in Theorem 1.2. If we truly wish to be pedantic, our current methods yield the upper bound $\ll Xh((\log X)(\log X/h))^2$, where the $\log X$ factors come from the bound on the sum over primes (Lemma 2.4) and where the $\log X/h$ factors come from splitting the sum over primes $h < p \leq 4X$ into dyadic intervals.

To complete the proof of Theorem 1.2, it remains to consider the h-smooth integers. The next section is dedicated to this task.

5 Smooth Integers

For the integers $X \leq n \leq 4X$ all of whose prime factors are $\leq h$, our goal is to obtain the following estimate:

$$\frac{1}{hT}\int_{X^{\frac{1}{2}}}^{T}\Big|\sum_{\substack{X\leq n\leq 4X\\ p|n\Rightarrow p\leq h}}\frac{\lambda(n)}{n^{\frac{1}{2}+it}}\Big|^2\mathrm{d}t \ll \frac{1}{h},$$

for $X/h \leq T \leq X$ and h as large as possible. This will be accomplished using the MVT (Lemma 2.2), together with some standard results concerning smooth numbers.

To begin, fix $\epsilon > 0$ and let $\Psi(x,y)$ denote the number of y-smooth integers up to x. By writing $x = y^u$, we have that

$$\Psi(x,y) = xu^{-(1+o(1))},$$

uniformly in the range $u \leq y^{1-\epsilon}$, as both y and u tend to infinity; see Corollary 1.3 of [2], for example. In particular,

$$\Psi(x,y) \ll \frac{x}{y},$$

provided $y \leq \exp\left(\sqrt{\left(\frac{1}{2}-o(1)\right)\log x\log\log x}\right)$. Using the above, together with the MVT, we then have that

$$\begin{aligned}\frac{1}{hT}\int_{X^{\frac{1}{2}}}^{T}\Big|\sum_{\substack{X\leq n\leq 4X\\ p|n\Rightarrow p\leq h}}\frac{\lambda(n)}{n^{\frac{1}{2}+it}}\Big|^2\mathrm{d}t &\ll \frac{1}{hT}(T+X)\sum_{\substack{X\leq n\leq 4X\\ p|n\Rightarrow p\leq h}}\frac{1}{n}\\ &\ll \frac{1}{hT}(T+X)\frac{\Psi(4X,h)}{X}\\ &\ll \frac{1}{h},\end{aligned}$$

for $h \leq \exp\left(\sqrt{\left(\frac{1}{2}-o(1)\right)\log X\log\log X}\right)$, recalling that $X/h \leq T \leq X$. In other words,

Proposition 5.1. Suppose $h = h(X) \to \infty$ as $X \to \infty$, then

$$\frac{1}{hT}\int_{X^{\frac{1}{2}}}^{T}\Big|\sum_{\substack{X\leq n\leq 4X\\ p|n\Rightarrow p\leq h}}\frac{\lambda(n)}{n^{\frac{1}{2}+it}}\Big|^2\mathrm{d}t \ll \frac{1}{h},$$

provided $h \leq \exp\left(\sqrt{\left(\frac{1}{2}-o(1)\right)\log X\log\log X}\right)$.

Combining Propositions 4.1 and 5.1 yields (2), which, together with the initial reductions from Section 3, yields Theorem 1.2.

Acknowledgments

The author would like to thank Maksym Radziwiłł for suggesting this problem and for his continuous support throughout the research phase of this paper, Andrew Granville for simplifying the argument originally presented in Section 5, and the anonymous referee for carefully reviewing this manuscript and whose comments have not only improved the overall exposition of this paper but also strengthened the main theorem. In addition to this, the author thanks Renaud Alie, Alison Inglis, Oleksiy Klurman, Jiajun Mai, Sacha Mangerel, Kaisa Matomäki, Andrei Shubin, Joni Teräväinen, and Peter Zenz for many helpful discussions and comments. Finally, some of this work was conducted while the author was visiting CalTech; he is grateful for their hospitality.

References

[1] Friedlander, J. and H. Iwaniec. *Opera de Cribro*, vol. 57 of American Mathematical Society Colloquium Publications. Providence, RI: American Mathematical Society, 2010.

[2] Hildebrand, A. and G. Tenenbaum. "Integers without large prime factors." *J. Théor. Nombres Bordeaux* 5, no. 2 (1993): 197–211.

[3] Iwaniec, H. and E. Kowalski. *Analytic Number Theory*, vol. 53 of American Mathematical Society Colloquium Publications. Providence, RI: American Mathematical Society, 2004.

[4] Maier, H. and H. L. Montgomery. "The sum of the Möbius function." *Bull. London Math. Soc.* 41, no. 2 (2009): 213–26.

[5] Motohashi, Y. "On the sum of the Möbius function in a short segment." *Proc. Japan Acad.* 52, no. 9 (1976): 477–9.

[6] Matomäki, K. and M. Radziwiłł. "A note on the Liouville function in short intervals." (2015): preprint arXiv:1502.02374.

[7] Matomäki, K. and M. Radziwiłł. "Multiplicative functions in short intervals." *Ann. Math. (2)* 183, no. 3 (2016): 1015–56.

[8] Matomäki, K. and J. Teräväinen. "On the Möbius function in all short intervals." (2019): preprint arXiv:1911.09076.

[9] Montgomery, H. L. and R. C. Vaughan. *Multiplicative Number Theory I: Classical Theory*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2006.

[10] Ramachandra, K. "Some problems of analytic number theory." *Acta Arith.* 31, no. 4 (1976): 313–24.

[11] Selberg, A. "On the normal density of primes in small intervals, and the difference between consecutive primes." *Arch. Math. Naturvid.* 47, no. 6 (1943): 87–105.

[12] Soundararajan, K. "Partial sums of the Möbius function." *J. Reine Angew. Math.* 2009 (2009): 141–52.

[13] Soundararajan, K. "The Liouville function in short intervals [after Matomäki and Radziwiłł]." (2016): preprint arXiv:1606.08021.

[14] Stein, E. M. *Singular Integrals and Differentiability Properties of Functions (PMS-30)*. Princeton University Press, 1970.

[15] Tenenbaum, G. *Introduction to Analytic and Probabilistic Number Theory*, vol. 163 of Graduate Studies in Mathematics. Providence, RI: American Mathematical Society, 2015.

[16] Titchmarsh, E. C. and D. R. Heath-Brown. *The Theory of the Riemann Zeta-Function*. Oxford: Clarendon Press, 1986.

Article II

[Chi21b] Jake Chinis. Siegel Zeros and Sarnak's Conjecture. Preprint, arXiv: math/2105.14653. Submitted to: *Adv. in Math.*, 2021

SIEGEL ZEROS AND SARNAK'S CONJECTURE

JAKE CHINIS

Abstract. Assuming the existence of Siegel zeros, we prove that there exists an increasing sequence of positive integers for which Chowla's Conjecture on k-point correlations of the Liouville function holds. This extends work of Germán and Kátai, where they studied the case $k = 2$ under identical hypotheses.

An immediate corollary, which follows from a well-known argument due to Sarnak, is that Sarnak's Conjecture on Möbius disjointness holds. More precisely, assuming the existence of Siegel zeros, there exists a subsequence of the natural numbers for which the Liouville function is asymptotically orthogonal to any sequence of topological entropy zero.

1. Introduction

The Liouville function, λ, is the completely multiplicative function defined by $\lambda(p) := -1$ for all primes p. As such, $\lambda(n) = 1$ if n has an even number of prime factors, counted with multiplicity, and $\lambda(n) = -1$ otherwise. An important problem in analytic number theory is to understand the asymptotic behaviour of the partial sums $\sum_{n\leq x} \lambda(n)$. Viewing $\{\lambda(n)\}_n$ as a sequence of independent random variables, each taking the value ± 1 with equal probability, we might expect that the partial sums of λ exhibit some cancellation. Indeed, the Prime Number Theorem (PNT) is equivalent to the fact

$$\sum_{n\leq x} \lambda(n) = o(x),$$

as $x \to \infty$. The Riemann Hypothesis (RH) is equivalent to the fact that the partial sums of λ exhibit "square-root cancellation:" for any $\epsilon > 0$,

$$\sum_{n\leq x} \lambda(n) \ll x^{1/2+\epsilon},$$

as $x \to \infty$. A famous conjecture due to Chowla [Cho65] states similar estimates hold for *multiple correlations* of λ:

Conjecture 1.1 (Chowla's Conjecture)**.** *For any k distinct integers $h_1, \dots, h_k$,*

$$\sum_{n\leq x} \lambda(n + h_1) \cdots \lambda(n + h_k) = o_k(x),$$

as $x \to \infty$, where we set $\lambda(n) = 0$ for integers $n \leq 0$ and where the implied constant may depend on $h_1, \dots, h_k$.

As stated earlier, the case $k = 1$ is equivalent to the PNT. Chowla's Conjecture remains open for $k > 1$ and the case $k = 2$ is believed to be as difficult as the Twin Prime Conjecture; see [Hil86].

Remark 1.1. *There are analogues of Chowla's Conjecture over finite function fields. We refer the reader to the work of Sawin and Shusterman* [SS18]*, where they study the relationship between Chowla's Conjecture and the Twin Prime Conjecture over finite function fields. It is interesting to note that Sawin and Shusterman use the same general idea that we use in this paper (and that was also used in* [GK10]*); namely, they approximate the Möbius function by a Dirichlet character.*

Although Chowla's Conjecture still seems to be out of reach, there has been much progress towards partial results of this conjecture, as well as proofs of variations of Chowla's original claim. For instance, Harman, Pintz, and Wolke [HPW85] proved that

$$-\frac{1}{3}+\mathcal{O}\Big(\frac{\log x}{x}\Big)\leq \frac{1}{x}\sum_{n\leq x}\lambda(n)\lambda(n+1)\leq 1-\mathcal{O}_{\epsilon}\Big(\frac{1}{\log^{7+\epsilon}x}\Big),$$

for all $\epsilon>0$. This was subsequently improved by Matomäki and Radziwiłł [MR16] to

$$\frac{1}{x}\left|\sum_{n\leq x}\lambda(n)\lambda(n+1)\right|\leq 1-\delta,$$

for some explicit constant $\delta>0$ and all x sufficiently large.

Concerning weaker versions of Chowla's Conjecture, Matomäki, Radziwiłł, and Tao [MRT15] averaged over the parameters $h_1,\dots,h_k$ and showed that, for any $k\in\mathbb{N}$ and any $10\leq h\leq x$,

$$\sum_{1\leq h_1,\dots,h_k\leq h}\left|\sum_{n\leq x}\lambda(n+h_1)\cdots\lambda(n+h_k)\right|\ll k\left(\frac{\log\log h}{\log h}+\frac{1}{\log^{1/3000}x}\right)h^k x,$$

thus establishing an *averaged* form of Chowla's Conjecture.

Tao [Tao16] made progress towards a *logarithmically averaged* version of Chowla's Conjecture by showing that

$$\sum_{n\leq x}\frac{\lambda(n)\lambda(n+1)}{n}=o(\log x)$$

as $x\to\infty$. Following up on this, Tao and Teräväinen [TT18, TT19] were able to establish a logarithmically averaged version of Chowla's Conjecture for odd k-point correlations; that is, for any odd $k\in\mathbb{N}$ and any integers $h_1,\dots,h_k$,

$$\sum_{n\leq x}\frac{\lambda(n+h_1)\cdots\lambda(n+h_k)}{n}=o(\log x)$$

as $x\to\infty$. Recently, Helfgott and Radziwiłł [HR21] improved the bounds obtained by Tao in [Tao16] and Tao and Teräväinen in [TT19] for $k=2$.

Remark 1.2. *Note that the preceding results follow from Chowla's Conjecture either immediately or by partial summation. Furthermore, the same results can be stated for the Möbius function, μ, in place of λ, via the identity $\mu(n)=\sum_{d^2|n}\mu(d)\lambda\left(\frac{n}{d^2}\right)$.*

In this paper, we are concerned with the relationship between the Liouville function and Siegel zeros. Our ultimate aim is to extend the work of Germán and Kátai [GK10], where they studied 2-point correlations of the Liouville assuming the existence of Siegel zeros:

Theorem 1.1. [GK10, Theorem 2] *Let $\{q_\ell\}_\ell$ be an increasing sequence of positive integers with corresponding sequence of real primitive characters $\{\chi_\ell \pmod{q_\ell}\}_\ell$. Suppose that $L(s,\chi_\ell)$ has a Siegel zero $\beta_\ell:=1-\frac{1}{\eta_\ell\log q_\ell}$ with $\eta_\ell>\exp\exp(30)$ for all $\ell\in\mathbb{N}$. Then, there exists a constant $c>0$ and a function $\varepsilon(x)\to 0$ as $x\to\infty$ such that*

$$\frac{1}{x}\left|\sum_{n\leq x}\lambda(n)\lambda(n+1)\right|\leq\frac{c}{\log\log\eta_\ell}+\varepsilon(x),$$

uniformly for $x\in[q_\ell^{10},q_\ell^{(\log\log\eta_\ell)/3}]$.

The key to the work of Germán and Kátai is to approximate λ by χ_ℓ on "large" primes via the completely multiplicative function λ_r defined by

$$\lambda_r(p) := \begin{cases} \lambda(p) & \text{if } p \le r, \\ \chi_\ell(p) & \text{if } p > r, \end{cases}$$

for some well-chosen parameter $r = r_\ell$. Then, using similar ideas as Heath-Brown in his work on Siegel zeros and the Twin Prime Conjecture [HB83], Germán and Kátai show that the 2-point correlations of λ are well approximated by the 2-point correlations of λ_r, along a subsequence. The added benefit to this approach is that we can now use sieve theory, together with the definition of λ_r, to relate the 2-point correlations of λ to some character sum, which is known to be small. Following this same line of reasoning, we prove the corresponding result for (general) k-point correlations:

Theorem 1.2. *Let $\{q_\ell\}_\ell$ be an increasing sequence of positive integers with corresponding sequence of real primitive characters $\{\chi_\ell \pmod{q_\ell}\}_\ell$. Suppose that $L(s, \chi_\ell)$ has a Siegel zero $\beta_\ell := 1 - \frac{1}{\eta_\ell \log q_\ell}$ with $\eta_\ell > \exp\exp(30)$ for all $\ell \in \mathbb{N}$. Then, for any distinct (positive) integers $h_1, \dots, h_k$, there exists a constant $c_k = c(h_1, \dots, h_k) > 0$ such that*

$$\frac{1}{x}\left|\sum_{n \le x} \lambda(n+h_1)\cdots\lambda(n+h_k)\right| \le \frac{c_k}{(\log\log\eta_\ell)^{1/2}(\log\eta_\ell)^{1/12}},$$

uniformly for $x \in [q_\ell^{10}, q_\ell^{(\log\log\eta_\ell)/3}]$.

Remark 1.3. *Since $\eta_\ell \to \infty$ as $\ell \to \infty$ (see Section 2), Theorem 1.2 thus establishes Chowla's Conjecture along a subsequence, assuming the existence of Siegel zeros. In fact, one should think of Theorem 1.1 as the multiplicative analogue of Heath-Brown's result on Twin Primes and Siegel zeros* [HB83]*, while Theorem 1.2 is the multiplicative analogue of the Hardy–Littlewood k-tuples conjecture (which remains open, even when assuming the existence of Siegel zeros).*

Remark 1.4. *Notice that Theorem 1.2 is an improvement on Theorem 1.1 in two respects: first, we can handle general k-point correlations (as opposed to the case where $k = 2$); further, we have an exponential improvement in our bounds (which follows from using a different version of the Fundamental Lemma of Sieve Theory and from taking a different choice of r than those used in* [GK10]*; see Appendix A and the end of Section 4).*

Note further that the work in [GK10] *deals only with $h_1 = 0$ and $h_1 = 1$, but the proof extends easily to general h_1, h_2 (with minor modifications). The main difficulty in going from the case $k = 2$ to general k-point correlations lies in being able to parametrize integer solutions of the following system of linear equations, in the unknowns $x_0, x_1, \dots x_k$, for some integers $a_0, a_1, \dots, a_k$:*

$$\begin{cases} a_1 x_1 = a_0 x_0 + h_1, \\ \vdots \\ a_k x_k = a_0 x_0 + h_k. \end{cases}$$

It is to verify that, if this system is solvable, then the solutions are given by $x_i = x_i^* + m\,\mathrm{lcm}(a_0, a_1, \dots, a_k)/a_i$, where $(x_0^*, x_1^*, \dots, x_k^*)$ is one particular solution and $m \in \mathbb{Z}$ (essentially generalizing Bezout's Identity to k equations). From there, we need to bound character sums evaluated at the polynomial $f(n) := (x_0^* + na_0^*)\cdots(x_k^* + na_k^*)$, where $a_i^* := \mathrm{lcm}(a_0, \dots, a_k)/a_i$, as n varies over one complete residue class modulo q_ℓ.

Fortunately for us, these character sums exhibit squareroot cancellation via Weil's Bound, provided that f is not a square. For more details, see Appendices B and C.

1.1. **Sarnak's Conjecture.** We should think of the previous results as instances of the so-called "Möbius Randomness Law," which states that the values of λ (or μ) are random enough so that the twisted sums $\sum_{n\leq x}\lambda(n)a_n$ should be small for any "reasonable" sequence of complex numbers $\{a_n\}_n$; see [IK04, Section 13]. A famous conjecture due to Sarnak characterizes one such family of "reasonable" sequences as those which are *deterministic*:

Definition 1.1. *Given a bounded sequence $f:\mathbb{N}\to\mathbb{C}$, its **topological entropy** is equal to the least exponent σ for which the set*

$$\{(f(n+1),f(n+2),\dots,f(n+m))\}_{n=1}^{\infty}\subset\mathbb{C}^m$$

*can be covered by $\mathcal{O}(\exp(\sigma m+o(m)))$ balls of radius ϵ (in the ℓ^∞ metric), for any fixed $\epsilon>0$, as $m\to\infty$. In the case where $\sigma=0$, we say that f is **deterministic**.*

Conjecture 1.2 (Sarnak's Conjecture)**.** *Let $f:\mathbb{N}\to\mathbb{C}$ be a deterministic sequence. Then,*

$$\sum_{n\leq x}\lambda(n)f(n)=o_f(x),$$

as $x\to\infty$.

Although Sarnak's Conjecture has yet to be resolved, there are many instances for which the conjecture holds. For example, in the case where f is constant, Sarnak's Conjecture is equivalent to the PNT; in the case where f is periodic, it is equivalent to the PNT in arithmetic progressions. For a more thorough survey on various instances for which Sarnak's Conjecture holds, see [FKPL18, KPL20].

By a purely combinatorial argument due to Sarnak [Sar11], we know that Chowla's Conjecture implies Sarnak's Conjecture; as a result, Theorem 1.2 yields the following:

Corollary 1.1. *Let $f:\mathbb{N}\to\mathbb{C}$ be a deterministic sequence. Under the hypotheses of Theorem 1.2,*

$$\sum_{n\leq x}\lambda(n)f(n)=o_f(x),$$

for $x\in[q_\ell^{10},q_\ell^{(\log\log\eta_\ell)/3}]$.

Proof. The proof follows Sarnak's argument verbatim, the details of which can be found on Tao's blog[1]. For further work on the relationship between Chowla's Conjecture and Sarnak's Conjecture, see [AKPLdlR17, GKL18, GLdlR20]. □

Remark 1.5. *The proof that Chowla implies Sarnak holds equally well for any other uncorrelated function (as opposed to the Liouville function). The key is to get precise estimates on the number of integers $n\leq x$ such that $|\frac{1}{m}\sum_{i=1}^m c_i\lambda(n+i)|\geq\epsilon$ for fixed m and epilon >0 and for arbitrary 1-bounded complex numbers $c_1,\dots,c_m$. One accomplishes this feat via the method of moments/Chebyshev's Inequality:*

$$\{n\leq x:|\frac{1}{m}\sum_{i=1}^m c_i\lambda(n+i)|\geq\epsilon\}\leq\frac{1}{(\epsilon m)^k}\sum_{n\leq x}\left|\sum_{i=1}^m c_i\lambda(n+i)\right|^k,$$

[1] https://terrytao.wordpress.com/2012/10/14/the-chowla-conjecture-and-the-sarnak-conjecture/

for some large, even integer k to be optimized later. Upon expanding the k-th power, one obtains an upper bound of the form

$$\frac{1}{(\epsilon m)^k} \sum_{1\leq 1_i,\dots i_k\leq m} \left| \sum_{n\leq x} \lambda(n+i_1)\cdots\lambda(n+i_k) \right|,$$

recalling that $|c_i| \leq 1$ for all i.

For the indices i_j which appear an odd number of times, we may apply Chowla's Conjecture directly to the summand and this is enough. It then remains to count the number of indices which always occur an even number of times; for these, a simple counting argument and trivial bounds suffice.

The remaining part of the proof amounts to replacing these constants c_i with $f(n+i)$, where we use the fact that our sequence is deterministic, so that, up to balls of radius ϵ, one should think of the terms $f(n+i)$ as "constant."

For an alternative proof set in the context of ergodic theory, see [AKPLdlR17].

1.2. **Outline.** Our paper is split as follows: in Section 2, we give a brief introduction on Siegel zeros and the work of Heath-Brown on counting the number of primes p such that $\chi(p) = 1$; in Section 3, we use the work of Germán–Kátai/Heath-Brown to relate the k-point correlations of λ to those of λ_r; from there, we use some estimates on character sums, sieve theory, and elementary number theory to prove Theorem 1.2; Appendices A, B, and C contain the relevant background information on sieve theory, character sums, and Diophantine equations, which we use freely in the proof of Theorem 1.2.

2. Siegel zeros and primes p such that $\chi(p) = 1$

In this section, we provide a brief introduction to Siegel zeros, culminating in the work of Heath-Brown on primes p such that $\chi(p) = 1$. To begin, we must first discuss zero-free regions of Dirichlet L-functions associated to Dirichlet characters $\chi \pmod q$; we follow Chapter 12 of [Kou20]:

Theorem 2.1. *Let $q \geq 3$ and set $Z_q(s) := \prod_{\chi \ (mod\ q)} L(s,\chi)$. Then, there is an absolute constant $c > 0$ such that the region $\Re(s) \geq 1 - \frac{c}{\log(q\tau)}$, where $\tau = \max\{1, |\Im(s)|\}$, contains at most one zero of Z_q. Furthermore, if this exceptional zero exists, then it is necessarily a real, simple zero of Z_q, say $\beta_1 \in [1 - c/\log q, 1]$, and there is a real, non-principal character $\chi_1 \ (mod\ q)$ such that $L(\beta_1, \chi_1) = 0$.*

Proof. See [Kou20, Theorem 12.3], for example. □

We call the character χ_1 in Theorem 2.1 an *exceptional character* and its zero, β_1, is the associated *exceptional zero*, or *Siegel/Landau–Siegel zero*. Note that this exceptional character depends on the choice of absolute constant and that this relationship implies some interesting facts:

(1) If we have one exceptional character, then we actually have infinitely many exceptional characters: if we had only finitely many exceptional characters $\chi_i \pmod{q_i}$, we could set $c' := \frac{1}{2}\min_i\{(1-\beta_i)\log q_i\}$ and we would then have that

$$1 - \frac{c}{\log q_i} \leq \beta_i < 1 - \frac{c'}{\log q_i}$$

for all i; in particular, replacing c with c' in Theorem 2.1, we no longer have any exceptional zeros.

(2) Similarly, we can take c to be arbitrarily small: if there are no exceptional zeros for c small enough, then we are done.

Thus, when we talk about Siegel zeros/exceptional characters, we are actually talking about an infinite sequence of real, primitive Dirichlet characters $\{\chi_\ell \pmod{q_\ell}\}_{\ell=1}^{\infty}$ for which $L(s,\chi_\ell)$ has a real zero

$$\beta_\ell = 1 - o_{\ell\to\infty}\left(\frac{1}{\log q_\ell}\right), \tag{2.1}$$

and such that no product $\chi_\ell\chi_{\ell'}$ is principal for any $\ell \neq \ell'$. Using Siegel's Theorem, we can quantify the rate of convergence in Equation (2.1):

Theorem 2.2 (Siegel)**.** *Let $\epsilon > 0$. Then, there is a constant $c(\epsilon) > 0$, which cannot be computed effectively, such that $L(\sigma,\chi) \neq 0$ for $\sigma > 1 - c(\epsilon)q^{-\epsilon}$ and for all real, non-principal Dirichlet characters $\chi \pmod q$.*

Proof. See [Kou20, Theorem 12.10], for example. □

In particular,

$$\eta_\ell := ((\beta_\ell - 1)\log q_\ell)^{-1} \ll q_\ell, \tag{2.2}$$

as $\ell \to \infty$. In fact, one could show that $\eta_\ell \ll$ any fixed power of q_ℓ, but the above is all we need for our purposes.

Now that we know exactly what we mean by Siegel zeros/exceptional characters, we can consider consequences of their existence. For example, Heath-Brown [HB83] showed, under similar hypotheses to Theorem 1.2, that the existence of Siegel zeros implies the Twin Prime Conjecture. Recently, Granville [Gra20] used the existence of Siegel zeros to study problems in sieve theory, such as improving (conditionally) lower bounds on the longest gaps between primes. For our purposes, we are interested in the following lemma, due to Germán and Kátai, which is a variation of Lemma 3 in [HB83]:

Lemma 2.1 ([GK10])**.** *Let $\{\chi_\ell \pmod{q_\ell}\}_\ell$ denote a sequence of exceptional characters with corresponding Siegel zero*

$$\beta_\ell = 1 - \frac{1}{\eta_\ell \log q_\ell},$$

with $\eta_\ell > \exp(\exp(30))$. Then,

$$\sum_{\substack{p\le x\\ \chi_\ell(p)=1}} \frac{\log p}{p} \ll \exp\left(\frac{\log x}{\log q_\ell}\right)(\log q_\ell)(\log \eta_\ell)^{-1/2},$$

uniformly for $x \in [q_\ell^{10}, q_\ell^{(\log\log \eta_\ell)/3}]$.

Remark 2.1. *Note that the upper bound in Lemma 2.1 is worse than that in Lemma 3 of* [HB83]*, but the range of admissible x is larger: Lemma 3 of* [HB83] *yields the upper bound $\ll (\log q_\ell)(\log \eta_\ell)^{-1/2}$, uniformly for $x \in [q_\ell^{250}, q_\ell^{500}]$ (so that Lemma 2.1 recovers Heath-Brown's result when x is restricted to the interval $[q_\ell^{250}, q_\ell^{500}]$).*

With Lemma 2.1 in tow, we can now approximate λ by a Dirichlet character on large primes; this is done in the next section. The key idea is to notice that Lemma 2.1 essentially states the following: the existence of Siegel zeros implies that $\chi_\ell(p) = -1$ more often than $\chi_\ell(p) = 1$, so that $\chi_\ell \approx \lambda$.

3. Going from λ to λ_r

From now on, we fix a character $\chi_\ell \pmod{q_\ell}$, so that we may drop the dependence on ℓ. Using Lemma 2.1, we can relate the k-point correlations of the Liouville function to the k-point correlations of the completely

multiplicative function λ_r, which is defined by

$$\lambda_r(p) := \begin{cases} \lambda(p) & \text{if } p \le r \\ \chi(p) & \text{if } p > r, \end{cases} \tag{3.1}$$

where $r := x^{1/((\log\log\eta)^{1/2}(\log\eta)^{1/12})}$. The details can be found in pages 48-50 of [GK10]; we reproduce their results here, for convenience/completeness[2]:

Lemma 3.1. *Suppose $h_1, \dots, h_k$ are distinct (positive) integers. Set $\lambda(n;k) := \lambda(n+h_1)\cdots\lambda(n+h_k)$ and define $\lambda_r(n;k)$ in the same way. Then, under the assumptions of Theorem 1.2,*

$$\frac{1}{x}\sum_{n\le x}\lambda(n;k) = \frac{1}{x}\sum_{n\le x}\lambda_r(n;k) + \mathcal{O}_k\left(\frac{1}{(\log\log\eta)^{1/2}(\log\eta)^{1/12}}\right),$$

uniformly for $x \in [q^{10}, q^{(\log\log\eta)/3}]$.

Proof. Note that

$$\begin{aligned}\sum_{n\le x}\lambda(n;k) &= \sum_{n\le x}\Big(\lambda(n;k) \pm \lambda_r(n+h_1)\lambda(n+h_2)\cdots\lambda(n+h_k)\Big) \\ &= \sum_{n\le x}\Big(\lambda(n+h_1) - \lambda_r(n+h_1)\Big)\lambda(n+h_2)\cdots\lambda(n+h_k) + \sum_{n\le x}\lambda_r(n+h_1)\lambda(n+h_2)\cdots\lambda(n+h_k).\end{aligned}$$

Continuing by induction, we have that

$$\begin{aligned}\left|\sum_{n\le x}\Big(\lambda(n;k) - \lambda_r(n;k)\Big)\right| &\le \sum_{i=1}^{k}\sum_{n\le x}|\lambda(n+h_i) - \lambda_r(n+h_i)| \\ &= k\sum_{n\le x}|\lambda(n) - \lambda_r(n)| + \mathcal{O}_k(1),\end{aligned}$$

where the last line follows from the "approximate translation-invariance" of the partial sums, noting that the error term depends on $h_1, \dots, h_k$. To bound $\sum_{n\le x}|\lambda(n;k) - \lambda_r(n;k)|$, recall the definition of λ_r from Equation (3.1) and note that

$$\begin{aligned}\sum_{n\le x}|\lambda(n) - \lambda_r(n)| &= \sum_{n\le x}\left|\prod_{p^\alpha||n}\lambda(p^\alpha) - \prod_{\substack{p^\alpha||n\\ p\le r}}\lambda(p^\alpha)\prod_{\substack{p^\alpha||n\\ p>r}}\chi(p^\alpha)\right| \\ &= \sum_{n\le x}\left|\prod_{\substack{p^\alpha||n\\ p\le r}}\lambda(p^\alpha)\left(\prod_{\substack{p^\alpha||n\\ p>r}}\lambda(p^\alpha) - \prod_{\substack{p^\alpha||n\\ p>r}}\chi(p^\alpha)\right)\right| \\ &= \sum_{n\le x}\left|\prod_{\substack{p^\alpha||n\\ p>r}}\lambda(p^\alpha) - \prod_{\substack{p^\alpha||n\\ p>r}}\chi(p^\alpha)\right|.\end{aligned}$$

Then, using the fact that

$$\left|\prod_i x_i - \prod_i y_i\right| \le \sum_i |x_i - y_i|,$$

[2]In [GK10], the authors take : $r = x^{1/\log\log\eta}$, which produces an error of size $1/\log\log\eta$ in Theorem 1.1. Making a different choice of r and using a different version of the FLST allows us to obtain better estimates; see the very end of Section 4 for why the choice of $r = x^{1/((\log\log\eta)^{1/2}(\log\eta)^{1/12})}$ was made/is optimal.

for all $x_i, y_i \in \{-1, 0, 1\}$, we have that

$$\begin{aligned}\sum_{n\leq x}\left|\prod_{\substack{p^\alpha||n\\p>r}}\lambda(p^\alpha)-\prod_{\substack{p^\alpha||n\\p>r}}\chi(p^\alpha)\right| &\leq \sum_{n\leq x}\sum_{\substack{p^\alpha||n\\p>r}}|\lambda(p^\alpha)-\chi(p^\alpha)|\\ &\leq \sum_{\substack{p^\alpha\leq x\\p>r}}|\lambda(p^\alpha)-\chi(p^\alpha)|\sum_{\substack{n\leq x\\p^\alpha||n}}1\\ &\ll x\sum_{\substack{p^\alpha\leq x\\p>r}}\frac{|\lambda(p^\alpha)-\chi(p^\alpha)|}{p^\alpha}\\ &= x\sum_{r<p\leq x}\frac{|\lambda(p)-\chi(p)|}{p}+x\sum_{\substack{p^\alpha\leq x\\p>r\\\alpha\geq 2}}\frac{|\lambda(p^\alpha)-\chi(p^\alpha)|}{p^\alpha}.\end{aligned}$$

The sum over the higher prime powers can be bounded trivially:

$$x\sum_{\substack{p^\alpha\leq x\\p>r\\\alpha\geq 2}}\frac{|\lambda(p^\alpha)-\chi(p^\alpha)|}{p^\alpha}\ll x\sum_{p>r}\sum_{\alpha\geq 2}\frac{1}{p^\alpha}\ll\frac{x}{r}.$$

For the sum over the primes, recall that $\lambda(p)=-1$, which yields:

$$\begin{aligned}x\sum_{r<p\leq x}\frac{|\lambda(p)-\chi(p)|}{p} &= x\sum_{\substack{r<p\leq x\\\chi(p)=1}}\frac{2}{p}+x\sum_{\substack{r<p\leq x\\\chi(p)=0}}\frac{1}{p}\\ &\ll\frac{x}{\log r}\sum_{\substack{r<p\leq x\\\chi(p)=1}}\frac{\log p}{p}+\frac{x}{\log r}\sum_{\substack{r<p\leq x\\\chi(p)=0}}\frac{\log p}{p},\end{aligned}$$

where the last line follows from the fact that $\log p/\log r>1$ for all $p>r$. Since $\chi(p)=0$ iff $p|q$, the sum over primes p such that $\chi(p)=0$ can be bounded above by

$$\begin{aligned}\frac{x}{\log r}\sum_{\substack{r<p\leq x\\\chi(p)=0}}\frac{\log p}{p} &\ll\frac{x}{\log r}\left(\sum_{p\leq\log q}\frac{\log p}{p}+\sum_{\substack{\log q<p\leq x\\p|q}}\frac{\log p}{\log q}\right)\\ &\ll\frac{x}{\log r}(\log\log q+1)\\ &\ll\frac{x\log\log q}{\log r}.\end{aligned}$$

For the remaining sum, we can apply Lemma 2.1, which yields:

$$\frac{x}{\log r}\sum_{\substack{r<p\leq x\\\chi(p)=1}}\frac{\log p}{p}\ll\frac{x}{\log r}\exp\left(\frac{\log x}{\log q}\right)(\log q)(\log\eta)^{-1/2}.$$

Recalling that $x\in[q^{10},q^{(\log\log\eta)/3}]$, $r=x^{1/((\log\log\eta)^{1/2}(\log\eta)^{1/12})}$, and $\eta\ll q$, the total error is then bounded above by

$$\frac{x}{\log r}\left(\exp\left(\frac{\log x}{\log q}\right)(\log q)(\log\eta)^{-1/2}+\log\log q\right)+\frac{x}{r}\ll\frac{x}{(\log\log\eta)^{1/2}(\log\eta)^{1/12}},$$

which follows from the fact that

$$\frac{1}{\log r}\exp\left(\frac{\log x}{\log q}\right)(\log q)(\log\eta)^{-1/2} = \frac{(\log\log\eta)^{1/2}(\log\eta)^{1/12}}{\log x}\exp\left(\frac{\log x}{\log q}\right)(\log q)(\log\eta)^{-1/2}$$

is an increasing function of x (for $x \geq q$), whose maximum on the interval $[q^{10}, q^{(\log\log\eta)/3}]$ will occur at $x = q^{(\log\log\eta)/3}$. In any case, we then have that

$$\frac{1}{x}\sum_{n\leq x}\lambda(n;k) = \frac{1}{x}\sum_{n\leq x}\lambda_r(n;k) + \mathcal{O}_k\left(\frac{1}{(\log\log\eta)^{1/2}(\log\eta)^{1/12}}\right),$$

as claimed. □

From Lemma 3.1, it now suffices to bound the k-point correlations of λ_r in order to prove Theorem 1.2; the next section is dedicated to this task.

4. Proof of Theorem 1.2

In this section, we complete the proof of Theorem 1.2 by bounding the k-point correlations of λ_r. Our main tool is the Fundamental Lemma of Sieve Theory (Lemma A.1), but before we can use this, we must first control the so-called "level of distribution" of the sieve; this is done with some preliminary sieving, by removing integers with "small" prime factors:

Lemma 4.1 (Controlling the level of distribution)**.** *Let $r < x$ and suppose $A_x \to \infty$ as $x \to \infty$. Then*[3]*:*

$$\#\{n \leq x : \prod_{\substack{p^\alpha || n \\ p\leq r}} p^\alpha > r^{A_x}\} \ll \frac{x}{A_x}.$$

Proof. This follows from Chebyshev's Inequality. □

Using Lemma 4.1, we can now restrict ourselves to integers $n \leq x$ such that the r-smooth parts of $n + h_i$ are all bounded above by r^{A_x}, for some A_x going to infinity slowly enough with respect to both x and η:

$$\sum_{n\leq x}\lambda_r(n;k) = \sum_{\substack{n\leq x \\ \prod_{\substack{p^\alpha||(n+h_i) \\ p\leq r}} p^\alpha \leq r^{A_x}}} \lambda_r(n;k) + \mathcal{O}\left(\frac{x}{A_x}\right).$$

Note: we will eventually choose $A_x \asymp_k (\log\log\eta)^{1/2}(\log\eta)^{1/12}$, which produces an admissible error in Theorem 1.2. For simplicity, we assume that $h_1 = 0$ and relabel the remaining indices: this amounts to shifting the sum over n by h_1 (which incurs an error of $\mathcal{O}_k(1)$), so that $h_i := h_i - h_1$ for $i = 2, 3, \ldots, k$. Relabeling the indices as $i = 1, 2, \ldots, k$, it suffices to bound

$$\sum_{\substack{n\leq x \\ \prod_{\substack{p^\alpha||(n+h_i) \\ p\leq r}} p^\alpha \leq r^{A_x}}} \lambda(n)\lambda_r(n+h_1)\cdots\lambda_r(n+h_k),$$

where we should think of k as $k-1$, with $h_0 = 0$.

Next, write n as a_0b_0, where a_0 is the r-smooth part of n and where b_0 is the r-sifted part. Doing the same procedure for $n + h_i$, $i = 1, \ldots, k$, we have that $a_ib_i = a_0b_0 + h_i$ and, in order for this system to be

[3]In [GK10], the bound x/A_x is simply written as $o(x)$, which is where this $\varepsilon(x)$ function comes from in Theorem 1.1. Keeping track of this error and then optimising the choice of r is how we obtain the improvements in Theorem 1.2.

solvable, it is necessary that $(a_i, a_j)|(h_i - h_j)$ for all $i \neq j$, recalling that $h_0 = 0$. Then:

$$\sum_{\substack{n\leq x\\ \prod_{p^\alpha||(n+h_i)}\, p^\alpha\leq r^{A_x}\\ p\leq r}} \lambda_r(n)\lambda_r(n+h_1)\cdots\lambda_r(n+h_k) = \sum_{\substack{a_0,a_1,\dots,a_k\leq r^{A_x}\\ p|a_i\Rightarrow p\leq r\\ (a_i,a_j)|(h_i-h_j)}} \lambda(a_0)\lambda(a_1)\cdots\lambda(a_k) \sum_{\substack{b_0\leq x/a_0\\ a_ib_i=a_0b_0+h_i\\ p|b_i\Rightarrow p>r}} \chi(b_0)\chi(b_1)\cdots\chi(b_k),$$

which follows from the definition of λ_r and after writing each $n + h_i$ as a product of its r-smooth and its r-sifted parts.

The key now is to parametrize the b_i's and to notice that if the system

$$\begin{cases} a_1b_1 = a_0b_0 + h_1,\\ \vdots\\ a_kb_k = a_0b_0 + h_k, \end{cases}$$

is solvable in the unknowns $b_0, b_1, \dots, b_k$, then the solutions are given by

$$b_i = b_i^* + m\frac{\operatorname{lcm}(a_0, a_1, \dots, a_k)}{a_i} =: b_i^* + ma_i^*,$$

where $(b_0^*, b_1^*, \dots, b_k^*)$ is one particular solution to the system and where m ranges over all integers; see Appendix C. Furthermore, we can take the b_i^*'s to be positive and minimal, in the sense that $b_i^* > 0$ for all i, while $b_i^* - a_i^* < 0$ for at least one i (this allows us to restrict ourselves to non-negative integers m and makes it so that $0 < b_i^* < a_i^*$ for at least one i, both of which are required to be completely rigorous in the next step). Then, using the fact that χ is periodic modulo q, we can write our sum as

$$\sum_{\substack{a_0,a_1,\dots,a_k\leq r^{A_x}\\ p|a_i\Rightarrow p\leq r\\ (a_i,a_j)|(h_i-h_j)}} \lambda(a_0)\lambda(a_1)\cdots\lambda(a_k) \sum_{\substack{b_0\leq x/a_0\\ a_ib_i=a_0b_0+h_i\\ p|b_i\Rightarrow p>r}} \chi(b_0)\chi(b_1)\cdots\chi(b_k) \tag{4.1}$$

$$= \sideset{}{^*}\sum_{\substack{a_0,a_1,\dots,a_k\leq r^{A_x}\\ p|a_i\Rightarrow p\leq r\\ (a_i,a_j)|(h_i-h_j)}} \lambda(a_0)\lambda(a_1)\cdots\lambda(a_k) \sum_{n=0}^{q-1} \chi(b_0^* + na_0^*)\chi(b_1^* + na_1^*)\cdots\chi(b_k^* + na_k^*) \tag{4.2}$$

$$\times \left(\#\left\{ m \leq \frac{x}{q\operatorname{lcm}(a_0, a_1, \dots, a_k)} : \left(\prod_{i=0}^{k} (b_i^* + na_i^* + mqa_i^*), \prod_{p\leq r} p \right) = 1 \right\} + \mathcal{O}_k(1) \right), \tag{4.3}$$

where the last factor counts the number of solutions which fall into each congruence class modulo q and where $\sideset{}{^*}\sum$ indicates that we are only summing over the a_i's for which the system is solvable. The conditions on a_i which make the system solvable are determined by the Smith Normal Form of the system; see Appendix C. For our purposes, we only care about the necessary condition $(a_i, a_j)|(h_i - h_j)$: this will allow us to control $\operatorname{lcm}(a_0, a_1, \dots, a_k)$, which will be needed later on in the proof. Note: the Big-O term comes from the fact that we are really looking at $b_0 = b_0^* + na_0^* + mqa_0^* \leq x/a_0$ (so that $m \leq x/(qa_0a_0^*) - b_0^*/qa_0^* - n/q$, with $0 \leq n \leq q - 1$).

Now, the Fundamental Lemma of Sieve Theorem (FLST, Lemma A.1) can be used to count the number of solutions which fall into each congruence class. The starting point for this is to get an asymptotic estimate for the number of such m which fall into the arithmetic progression 0 (mod d) for $d|\prod_{p\leq r} p$. So, let

$$\nu(d) := \#\left\{ m \in \mathbb{Z}/d\mathbb{Z} : \prod_{i=0}^{k} (b_i^* + na_i^* + mqa_i^*) \equiv 0 \pmod{d} \right\}$$

and note that

$$\#\left\{m \leq \frac{x}{q\,\mathrm{lcm}(a_0,a_1,\ldots,a_k)} : \prod_{i=0}^{k}(b_i^* + na_i^* + mqa_i^*) \equiv 0 \pmod{d}\right\} = \frac{x}{q\,\mathrm{lcm}(a_0,a_1,\ldots,a_k)}\frac{\nu(d)}{d} + \mathcal{O}(\nu(d)).$$

By the Chinese Remainder Theorem, $\nu(d)$ is completely determined by $\nu(p)$, for $p \leq r$. Moreover, $\nu(p)$ is equal to the number of distinct solutions $m \pmod p$ to any of the following linear congruences:

$$\begin{cases} (qa_0^*)m \equiv -(b_0^* + na_0^*) \pmod p \\ (qa_1^*)m \equiv -(b_1^* + na_1^*) \pmod p \\ \vdots \\ (qa_k^*)m \equiv -(b_k^* + na_k^*) \pmod p. \end{cases}$$

In order to obtain precise estimates for $\nu(p)$, we consider various possibilities depending on whether or not qa_i^* is invertible modulo p. For starters, we restrict the sum over n to the sum over n such that $(b_i^* + na_i^*, q) = 1$ for all i; otherwise, $\chi(b_i^* + na_i^*) = 0$, which contributes nothing to Equation 4.1. We consider the following scenarios:

(1) $p|q$: In the case where $p|q$, there are no solutions because this would require that $p|b_i^* + na_i^*$ for at least one i, contrary to our hypothesis that $(b_i^* + na_i^*, q) = 1$ for all i; i.e., $\nu(p) = 0$ when $p|q$.

(2) $p \nmid qa_0a_1\ldots a_k$: In the case where $p \nmid q$ and $p \nmid a_i$ for any i, we have that qa_i^* is invertible modulo p for all i; in particular, each equation in the system produces exactly one solution. If, in addition, we have that $p > \max_i\{h_i\}$, then we have that $\nu(p) = k+1$. To see this, let $m_i := -(qa_i^*)^{-1}(b_i^* + na_i^*) \pmod p$ for all i and note that $m_i = m_j$ iff $a_i^*b_j^* = a_j^*b_i^* \pmod p$. Then, using the fact that $a_i^* = \mathrm{lcm}(a_0,a_1,\ldots,a_k)/a_i$, together with the fact that both $\mathrm{lcm}(a_0,a_1,\ldots,a_k)$ and a_i are invertible (mod p), we have that $m_i = m_j$ iff $a_ib_i^* = a_jb_j^* \pmod p$. Then, recalling our definition of the b_i^*'s, we have that $m_i = m_j$ iff $p|(h_i - h_j)$. Since $1 \leq |h_i - h_j| \leq \max_i\{h_i\}$, it follows that the m_i's are distinct for $p > \max_i\{h_i\}$, which implies that $\nu(p) = k+1$ for such p. In other words, the a_i's are relatively prime on the primes $p > \max_i\{h_i\}$ and this makes it so that there are exactly $k+1$ solutions to our system of linear congruences, provided p is large enough. For the smaller primes, we content ourselves with the fact that $\nu(p) \leq p$.

(3) $p \nmid q, p|a_i$ (for at least one i): As mentioned above, the a_i's are pairwise relatively prime on the primes $p > \max_i\{h_i\}$; that is, if $p|a_i$ for at least one i and if $p > \max_i\{h_i\}$, then $p|a_i$ for exactly one i, say a_{i_0}. Moreover, this implies that $p \nmid a_{i_0}^*$ and that $p|a_i^*$ for all $i \neq i_0$, which again follows from the fact that the a_i's are pairwise relatively prime on the primes $> \max_i\{h_i\}$. In particular, there is exactly one solution for the i_0-th equation (as $a_{i_0}^*$ is invertible modulo p) and there are either no solutions for the other equations or p solutions, depending on whether or not $p|b_i^*$ for some $i \neq i_0$. To summarize the case where $p \nmid q, p|a_{i_0}$, we have precisely one of the following: (for $p > \max_i\{h_i\}$) either $\nu(p) = 1$ or $\nu(p) = p$ and these situations occur if $p \nmid b_i^*$ for all $i \neq i_0$ or $p|b_i^*$ for some $i \neq i_0$, respectively.

There are a few key points to notice from the above analysis. First, note that $\nu(p)$ is independent of n: this is clear if $p > \max_i\{h_i\}$ or if $p|q$, but even in the case where $p \leq \max_i\{h_i\}$ and $p \nmid q$, we either have that $p|a_i^*$ for some i (so that we either have no solutions or p solutions for the i-th equation) or $p \nmid a_i^*$ for some i (in which case, two solutions m_i, m_j are equal iff $a_i^*b_j^* = a_j^*b_i^*$, so that $\nu(p)$ is still independent of n). Next, we can simply restrict the sum over the a_i's so that $\nu(p) = \nu(p; a_0,\ldots,a_k) \neq p$ for any p; in the case where

$\nu(p) = p$, the sum over the b_i's is 0, as all s are such that $(qa_i^*)s \equiv -(b_i^* + na_i^*) \pmod p$, for some i (which implies that there is no m such that $\left(\prod_{i=0}^k (b_i^* + na_i^* + mqa_i^*), \prod_{p\le r} p\right) = 1$), and there is nothing to prove.

After verifying that $\nu(p)$ satisfies Axioms A.1 and A.2, the FLST (Lemma A.1) then yields the following:

$$\#\left\{m \le \frac{x}{q\operatorname{lcm}(a_0,a_1,\dots,a_k)} : \left(\prod_{i=0}^k (b_i^* + na_i^* + mqa_i^*), \prod_{p\le r} p\right) = 1\right\}$$
$$= (1+\mathcal{O}_k(u^{-u/2}))\frac{x}{q\operatorname{lcm}(a_0,a_1,\dots,a_k)}\prod_{\substack{p\le r\\ p\nmid q}}\left(1-\frac{\nu(p)}{p}\right) + \mathcal{O}\left(\sum_{\substack{d\le r^u\\ d\mid \prod_{p\le r} p}} \nu(d)\right),$$

uniformly for $u \ge 1$.

Plugging the above back into Equation 4.1, we are left to bound

$$\sideset{}{^*}\sum_{\substack{a_0,a_1,\dots,a_k\le r^{A_x}\\ p\mid a_i\Rightarrow p\le r\\ (a_i,a_j)\mid(h_i-h_j)}} \lambda(a_0)\lambda(a_1)\cdots\lambda(a_k)\sum_{n=0}^{q-1}\chi(b_0^*+na_0^*)\chi(b_1^*+na_1^*)\cdots\chi(b_k^*+na_k^*)$$
$$\times(1+\mathcal{O}_k(u^{-u/2}))\frac{x}{q\operatorname{lcm}(a_0,a_1,\dots,a_k)}\prod_{\substack{p\le r\\ p\nmid q}}\left(1-\frac{\nu(p)}{p}\right) + \mathcal{O}\left(\sum_{\substack{d\le r^u\\ d\mid \prod_{p\le r} p}} \nu(d)\right),$$

which we break into three parts, according to the three summands in the last factor.

4.1. **Bounding the error term containing the sum over $d \le r^u$.** Our goal in this subsection is to choose u (the level of distribution) as large as possible, while minimizing the "error" term containing the sum over $d \le r^u$. To begin, recall that $\nu(p) \le \min\{k+1, p\}$; in particular, $\nu(d) \le (k+1)^{\omega(d)}$ for all $d\mid\prod_{p\le r}$, where $\omega(d)$ counts the number of prime divisors of d. If we let τ_κ denote the κ-th divisor function (which counts the number of representations an integer has as a product of κ integers), it is clear that $(k+1) \le \tau_{k+1}(p)$ for all p, so that $(k+1)^{\omega(d)} \le \tau_{k+1}(d)$. Hence,

$$\sum_{\substack{d\le r^u\\ d\mid\prod_{p\le r} p}} \nu(d) \le \sum_{d\le r^u}\tau_{k+1}(d) \ll_k r^u(u\log r)^{k+1},$$

which follows from the average order of τ_κ; see Exercise 3.10 in [Kou20]. The total contribution to Equation (4.1) is then bounded by

$$\ll_k qr^{(k+1)A_x}\cdot r^u(u\log r)^{k+1} \ll x^{1/2},$$

as $x\to\infty$, provided $A_x \le ((\log\log\eta)^{1/2}(\log\eta)^{1/12})/(10(k+1))$ and $u \le ((\log\log\eta)^{1/2}(\log\eta)^{1/12})/10$, say, recalling that $r = x^{1/((\log\log\eta)^{1/2}(\log\eta)^{1/12})}$ with $q^{10}\le x$.

4.2. **Bounding the error from the "main term" in the FLST.** For the main term, we use Lemma B.1, with $f(x) = (b_0^*+a_0^*x)(b_1^*+a_1^*x)\cdots(b_k+a_k^*x)$, to bound the character sum:

$$\sum_{n=0}^{q-1}\chi(b_0^*+na_0^*)\chi(b_1^*+na_1^*)\cdots\chi(b_k^*+na_k^*) \ll_k q^{1/2+\epsilon};$$

the key is to note that f is not a square modulo p for any prime $p > \max_i\{h_i\}$, which follows from the fact that $a_j^* b_i^* = a_i^* b_j^*$ iff $p|(h_i - h_j)$. Hence,

$$\frac{x}{q}\sideset{}{^*}\sum_{\substack{a_0,a_1,\dots,a_k\le r^{A_x}\\ p|a_i\Rightarrow p\le r\\ (a_i,a_j)|(h_i-h_j)}}\frac{\lambda(a_0)\lambda(a_1)\cdots\lambda(a_k)}{\operatorname{lcm}(a_0,a_1,\dots,a_k)}\prod_{\substack{p\le r\\ p\nmid q}}\left(1-\frac{\nu(p)}{p}\right)\sum_{n=0}^{q-1}\chi(b_0^*+na_0^*)\chi(b_1^*+na_1^*)\cdots\chi(b_k+na_k^*)$$

$$\ll_k \frac{x}{q^{1/2-\epsilon}}\sum_{\substack{a_0,a_1,\dots,a_k\le r^{A_x}\\ p|a_i\Rightarrow\le r\\ (a_i,a_j)|i-j}}\frac{1}{\operatorname{lcm}(a_0,a_1,\dots,a_k)}\prod_{\substack{\max_i\{h_i\}<p\le r\\ p\nmid q\\ p\nmid a_0a_1\cdots a_k}}\left(1-\frac{k+1}{p}\right)\prod_{\substack{\max_i\{h_i\}<p\le r\\ p\nmid q\\ p|a_0a_1\cdots a_k}}\left(1-\frac{1}{p}\right),$$

where we have removed the contribution from the small prime factors (which is $\ll_k 1$), where we have bounded the sum over the a_i's trivially, and where we have split the product over the primes according to the value of $\nu(p)$.

Remark 4.1. *Here is a more detailed approach to showing that f is not a square modulo p for all but finitely-many p: we need to consider two cases, depending on whether or not $p|a_i$, for some i. So, suppose $p > \max_i\{h_i\}$ and that $p\nmid a_i$ for any i; then $p\nmid a_i^*$ for any i, so that a_i^* is invertible modulo p for all i and that f can only be a square if $b_i^*(a_i^*)^{-1} \equiv b_j^*(a_j^*)^{-1}\ (mod\ p)$ for at least some $i\neq j$, but this occurs, as we say before, iff $p|(h_i-h_j)$ which cannot occur for $p > \max_i\{h_i\}$. In the case where $p|a_i$ for some i, then $p|a_i$ for exactly one i (otherwise, by the condition that $(a_i,a_j)|(h_i-h_j)$, we would get a contradiction); in particular, $p|a_j^*$ for all $j\neq i$, so that $f(x)\equiv c(b_i^*+a_i^*x)\ (mod\ p)$, for some $c\in\mathbb{Z}/p\mathbb{Z}$ and where $p\nmid a_i^*$ and f is clearly not a square.*

Next, using the fact that

$$\operatorname{lcm}(a_0,a_1,\dots,a_k)\ge\prod_{i<j}(a_i,a_j)^{-1}a_0a_1\cdots a_k,$$

together with the fact that $(a_i,a_j)|(h_i-h_j)$, we can bound our sum by

$$\ll_k \frac{x}{q^{1/2-\epsilon}}\sum_{\substack{a_0,a_1,\dots,a_k\le r^{A_x}\\ p|a_i\Rightarrow\le r\\ (a_i,a_j)|(h_i-h_j)}}\frac{1}{a_0a_1\cdots a_k}\prod_{\substack{\max_i\{h_i\}<p\le r\\ p\nmid q\\ p\nmid a_0a_1\cdots a_k}}\left(1-\frac{k+1}{p}\right)\prod_{\substack{\max_i\{h_i\}<p\le r\\ p\nmid q\\ p|a_0a_1\cdots a_k}}\left(1-\frac{1}{p}\right).$$

Again recalling that the a_i's are pairwise relatively prime on the primes $p > \max_i\{h_i\}$, we can actually separate the variables in our sum, so that

$$\frac{1}{a_0a_1\cdots a_k}\prod_{\substack{\max_i\{h_i\}<p\le r\\ p\nmid q\\ p\nmid a_0a_1\cdots a_k}}\left(1-\frac{k+1}{p}\right)\prod_{\substack{\max_i\{h_i\}<p\le r\\ p\nmid q\\ p|a_0a_1\cdots a_k}}\left(1-\frac{1}{p}\right)$$

$$=\prod_{i=0}^{k}\frac{1}{a_i}\prod_{\substack{\max_i\{h_i\}<p\le r\\ p\nmid q\\ p\nmid a_i}}\left(1-\frac{k+1}{p}\right)\prod_{\substack{\max_i\{h_i\}<p\le r\\ p\nmid q\\ p|a_i}}\left(1-\frac{1}{p}\right).$$

To deal with the products over the primes, we can multiply and divide by "complements" of the given divisibility conditions, the idea being that we want to "complete" our products to keep only divisibility

conditions (as opposed to saying $p \nmid q$, for example):

$$\prod_{\substack{\max_i\{h_i\}<p\le r\\ p\nmid q\\ p\nmid a_i}}\left(1-\frac{k+1}{p}\right)\prod_{\substack{\max_i\{h_i\}<p\le r\\ p\nmid q\\ p\mid a_i}}\left(1-\frac{1}{p}\right)$$

$$=\prod_{\max_i\{h_i\}<p\le r}\left(1-\frac{k+1}{p}\right)\prod_{\substack{\max_i\{h_i\}<p\le r\\ p\mid q}}\left(1-\frac{k+1}{p}\right)^{-1}\prod_{\substack{\max_i\{h_i\}<p\le r\\ p\mid a_i}}\left(1-\frac{k+1}{p}\right)^{-1}\prod_{\substack{\max_i\{h_i\}<p\le r\\ p\mid q\\ p\mid a_i}}\left(1-\frac{k+1}{p}\right)$$

$$\times\prod_{\substack{\max_i\{h_i\}<p\le r\\ p\mid a_i}}\left(1-\frac{1}{p}\right)\prod_{\substack{\max_i\{h_i\}<p\le r\\ p\mid q\\ p\mid a_i}}\left(1-\frac{1}{p}\right)^{-1}$$

$$\le\prod_{\max_i\{h_i\}<p\le r}\left(1-\frac{k+1}{p}\right)\prod_{\substack{\max_i\{h_i\}<p\le r\\ p\mid q}}\left(1-\frac{k+1}{p}\right)^{-1}\prod_{\substack{\max_i\{h_i\}<p\le r\\ p\mid a_i}}\left(\frac{p-1}{p-(k+1)}\right),$$

where we have combined the factors for $p|a_i$ and where we have trivially bounded the product of the factors over $p|q, a_i$ by ≤ 1.

Putting all of this together, we have that

$$\sum_{\substack{a_0,a_1,\dots,a_k\le r^{A_x}\\ p\mid a_i\Rightarrow\le r\\ (a_i,a_j)\mid(h_i-h_j)}}\frac{1}{a_0a_1\cdots a_k}\prod_{\substack{\max_i\{h_i\}<p\le r\\ p\nmid q\\ p\nmid a_0a_1\cdots a_k}}\left(1-\frac{k+1}{p}\right)\prod_{\substack{\max_i\{h_i\}<p\le r\\ p\nmid q\\ p\mid a_0a_1\cdots a_k}}\left(1-\frac{1}{p}\right)$$

$$\ll\prod_{\max_i\{h_i\}<p\le r}\left(1-\frac{k+1}{p}\right)\prod_{\substack{\max_i\{h_i\}<p\le r\\ p\mid q}}\left(1-\frac{k+1}{p}\right)^{-1}\left(\sum_{\substack{a\le r^{A_x}\\ p\mid a\Rightarrow p\le r}}\frac{1}{a}\prod_{\substack{\max_i\{h_i\}<p\le r\\ p\mid a}}\left(\frac{p-1}{p-(k+1)}\right)\right)^{k+1}.$$

Finally, we can bound the sum over a by an Euler product, as the corresponding summand is a multiplicative function supported over the r-smooth integers:

$$\sum_{\substack{a\le r^{A_x}\\ p\mid a\Rightarrow p\le r}}\frac{1}{a}\prod_{\substack{\max_i\{h_i\}<p\le r\\ p\mid a}}\left(\frac{p-1}{p-(k+1)}\right)\le\prod_{p\le k+1}\left(1-\frac{1}{p}\right)^{-1}\prod_{\max_i\{h_i\}<p\le r}\left(1+\frac{p-1}{p-(k+1)}\left(\frac{1}{p}+\frac{1}{p^2}+\dots\right)\right)$$

$$\ll_k\prod_{\max_i\{h_i\}<p\le r}\left(1+\frac{1}{p-(k+1)}\right).$$

Therefore, the total contribution from the "main term" in the FLST is bounded above by

$$\ll_k\frac{x}{q^{1/2-\epsilon}}\prod_{\max_i\{h_i\}<p\le r}\left(1-\frac{k+1}{p}\right)\prod_{\substack{\max_i\{h_i\}<p\le r\\ p\mid q}}\left(1-\frac{k+1}{p}\right)^{-1}\prod_{\max_i\{h_i\}<p\le r}\left(1+\frac{1}{p-(k+1)}\right)^{k+1}$$

$$\ll_k\frac{x}{q^{1/2-\epsilon}}\prod_{\substack{\max_i\{h_i\}<p\le r\\ p\mid q}}\left(1-\frac{k+1}{p}\right)^{-1},$$

where the last line follows from the fact that

$$\prod_{\max_i\{h_i\}<p\le r}\left(1-\frac{k+1}{p}\right)\left(1+\frac{1}{p-(k+1)}\right)^{k+1}=\prod_{\max_i\{h_i\}<p\le r}\left(1+\mathcal{O}\left(\frac{1}{p^2}\right)\right)=\mathcal{O}(1),$$

which in turn follows from the Binomial Theorem. Hence, the main term is bounded above by

$$\frac{x}{q^{1/2-\epsilon}}\prod_{\substack{\max_i\{h_i\}<p\le r\\ p\mid q}}\left(1-\frac{k+1}{p}\right)^{-1}\ll_k \frac{x\log^{k+1}q}{q^{1/2-\epsilon}},$$

where the first line follows by taking exp-log of the product and using the Taylor series expansion of log and which produces an admissible error, after recalling that $x\le q^{(\log\log\eta)/3}$.

4.3. **Dealing with the error from $u^{-u/2}$.** For the error containing the term $u^{-u/2}$, we have that Equation (4.1) is bounded above by

$$\ll_k u^{-u/2}\frac{x}{q}\sum_{\substack{a_0,a_1,\dots,a_k\le r^{A_x}\\ p\mid a_i\Rightarrow p\le r\\ (a_i,a_j)\mid(h_i-h_j)}}\frac{1}{\operatorname{lcm}(a_0,a_1,\dots,a_k)}\sum_{n=0}^{q-1}|\chi(b_0^*+na_0^*)\chi(b_1^*+na_1^*)\cdots\chi(b_k^*+na_k^*)|\prod_{\substack{p\le r\\ p\nmid q}}\left(1-\frac{\nu(p)}{p}\right)$$

$$=u^{-u/2}\frac{x}{q}\sum_{\substack{a_0,a_1,\dots,a_k\le r^{A_x}\\ p\mid a_i\Rightarrow p\le r\\ (a_i,a_j)\mid(h_i-h_j)}}\frac{1}{\operatorname{lcm}(a_0,a_1,\dots,a_k)}\prod_{\substack{p\le r\\ p\nmid q}}\left(1-\frac{\nu(p)}{p}\right)\sum_{\substack{n=0\\ \left(\prod_{i=0}^k(b_i^*+na_i^*),q\right)=1}}^{q-1}1.$$

To estimate the sum over n, we use the fact that

$$\sum_{d\mid n}\mu(n)=\begin{cases}1 \text{ if } n=1\\ 0 \text{ otherwise,}\end{cases}$$

which yields:

$$\begin{aligned}\sum_{\substack{n=0\\ \left(\prod_{i=0}^k(b_i^*+na_i^*),q\right)=1}}^{q-1}1&=\sum_{n=0}^{q-1}\sum_{d\mid\left(\prod_i(b_i^*+na_i^*),q\right)}\mu(d)\\ &=\sum_{d\mid q}\sum_{\substack{n=0\\ d\mid\prod_i(b_i^*+na_i^*)}}^{q-1}\mu(d)\\ &=q\sum_{d\mid q}\frac{N(d)}{d},\end{aligned}$$

where $N(d)$ is the number of $n\in\mathbb{Z}/d\mathbb{Z}$ such that $\prod_{i=0}^{k+1}(b_i^*+na_i^*)\equiv 0 \pmod d$, $d\mid q$. By the CRT, $N(d)$ is multiplicative, so that the sum can be written as

$$q\sum_{d\mid q}\frac{\mu(d)N(d)}{d}=q\prod_{p\mid q}\left(1-\frac{N(p)}{p}\right)$$

and it remains to compute $N(p)$; we consider various cases, depending on whether or not a_i^* is invertible (mod p), noting that $N(p)$ is equal to the number of $n \pmod p$ such that $na_i^*\equiv -b_i^* \pmod p$, for any i.

(1) $p\nmid a_0a_1\cdots a_k$: In this case, all the a_i^* are invertible modulo p so that exactly one n satisfies the given congruence in the i-th equation. Assuming further that $p>\max_i\{h_i\}$, we get $k+1$ distinct

solutions as $(a_i^*)^{-1}b_i^* = (a_j^*)^{-1}b_j^* \pmod p$ iff $p|(h_i - h_j)$; that is, $N(p) = k+1$ if $p > \max_i\{h_i\}$, with $1 \le N(p) \le p$ otherwise.

(2) $p|a_0a_1\cdots a_k$: Similarly, $N(p) = 1$ if $p > \max_i\{h_i\}$ with $N(p) \le p$ otherwise. The idea here is that the a_i's are pairwise relatively prime on the primes $p > \max_i\{h_i\}$ so that exactly one a_i is divisible by p if $p|a_0\cdots a_k$, say a_{i_0}; in particular, we get exactly one solution at level i_0 and the other congruences are solvable iff $b_i^* \equiv 0 \pmod p$ for some i. In the latter case, we have that $N(p) = p$ and the product is 0, so we may assume that $p \nmid b_i^*$ for all i.

Thus, the second error term is bounded above by

$$\begin{aligned}
&\ll_k u^{-u/2}x \sum_{\substack{a_0,a_1,\dots,a_k \le r^{A_x}\\ p|a_i \Rightarrow p\le r\\ (a_i,a_j)|(h_i-h_j)}} \frac{1}{\operatorname{lcm}(a_0,a_1,\dots,a_k)}\\
&\qquad\times \prod_{\substack{\max_i\{h_i\}<p\le r\\ p\nmid q\\ p\nmid a_0a_1\cdots a_k}}\left(1-\frac{k+1}{p}\right)\prod_{\substack{\max_i\{h_i\}<p\le r\\ p\nmid q\\ p|a_0a_1\cdots a_k}}\left(1-\frac{1}{p}\right)\prod_{\substack{\max_i\{h_i\}<p\\ p|q\\ p\nmid a_0a_1\cdots a_k}}\left(1-\frac{k+1}{p}\right)\prod_{\substack{\max_i\{h_i\}<p\\ p|q\\ p|a_0a_1\cdots a_k}}\left(1-\frac{1}{p}\right)\\
&\ll_k u^{-u/2}x \sum_{\substack{a_0,a_1,\dots,a_k \le r^{A_x}\\ p|a_i \Rightarrow p\le r\\ (a_i,a_j)|(h_i-h_j)}} \frac{1}{a_0a_1\cdots a_k}\prod_{\substack{\max_i\{h_i\}<p\le r\\ p\nmid a_0a_1\cdots a_k}}\left(1-\frac{k+1}{p}\right)\prod_{\substack{\max_i\{h_i\}<p\le r\\ p|a_0a_1\cdots a_k}}\left(1-\frac{1}{p}\right)\\
&\ll_k u^{-u/2}x \prod_{\max_i\{h_i\}<p\le r}\left(1-\frac{k+1}{p}\right)\sum_{\substack{a_0,a_1,\dots,a_k \le r^{A_x}\\ p|a_i \Rightarrow p\le r\\ (a_i,a_j)|(h_i-h_j)}} \frac{1}{a_0a_1\cdots a_k}\prod_{\substack{\max_i\{h_i\}<p\le r\\ p|a_0a_1\cdots a_k}}\left(1-\frac{k+1}{p}\right)^{-1}\prod_{\substack{\max_i\{h_i\}<p\le r\\ p|a_0a_1\cdots a_k}}\left(1-\frac{1}{p}\right).
\end{aligned}$$

Again using the fact that the a_i's are pairwise relatively on the primes $p > \max_i\{h_i\}$, we can separate the variables in the sum over the a_i's:

$$\begin{aligned}
&\sum_{\substack{a_0,a_1,\dots,a_k \le r^{A_x}\\ p|a_i \Rightarrow p\le r\\ (a_i,a_j)|(h_i-h_j)}} \frac{1}{a_0a_1\cdots a_k}\prod_{\substack{\max_i\{h_i\}<p\le r\\ p|a_0a_1\cdots a_k}}\left(1-\frac{k+1}{p}\right)^{-1}\prod_{\substack{\max_i\{h_i\}<p\le r\\ p|a_0a_1\cdots a_k}}\left(1-\frac{1}{p}\right)\\
&= \sum_{\substack{a_0,a_1,\dots,a_k \le r^{A_x}\\ p|a_i \Rightarrow p\le r\\ (a_i,a_j)|(h_i-h_j)}}\left(\prod_{i=0}^{k}\frac{1}{a_i}\prod_{\substack{\max_i\{h_i\}<p\le r\\ p|a_i}}\left(1-\frac{k+1}{p}\right)^{-1}\left(1-\frac{1}{p}\right)\right)\\
&\le \left(\sum_{\substack{a\le r^{A_x}\\ p|a\Rightarrow p\le r}}\frac{1}{a}\prod_{\substack{\max_i\{h_i\}<p\le r\\ p|a}}\left(\frac{p-1}{p-(k+1)}\right)\right)^{k+1}\\
&\le \left(\prod_{p\le\max_i\{h_i\}}\left(1-\frac{1}{p}\right)^{-1}\prod_{\max_i\{h_i\}<p\le r}\left(1+\frac{p-1}{p-(k+1)}\left(\frac{1}{p}+\frac{1}{p^2}+\dots\right)\right)\right)^{k+1}.
\end{aligned}$$

Therefore, the total contribution from the secondary error term is

$$\begin{aligned}&\ll_k u^{-u/2}x \prod_{\max_i\{h_i\}<p\leq r}\left(1-\frac{k+1}{p}\right)\left(1+\frac{1}{p-(k+1)}\right)^{k+1}\\&\ll_k u^{-u/2}x,\end{aligned}$$

noting that this is the same product over the primes that we encountered earlier and which we saw was $\ll_k 1$. Taking $u \asymp (\log\log\eta)^{1/2}(\log\eta)^{1/12}$ produces an admissible error and thus establishes Theorem 1.2.

4.4. **Fitting the pieces.** In this section, we want to say a few words which justify our choice of parameters for r, A_x, and u. Setting $r = x^{1/\alpha}$, we need to choose the largest possible A_x and u for which the following errors are minimized, while optimizing α:

$$\begin{cases}\frac{x}{\log r}\left(\exp\left(\frac{\log x}{\log q}\right)(\log q)(\log\eta)^{-1/2}\right)+\log\log q\Big)+\frac{x}{r} & \text{(error from the proof of Lemma 3.1)},\\ \frac{x}{A_x} & \text{(error from Lemma 4.1)},\\ qr^{(k+1)A_x+u}(u\log r)^{k+1} & \text{(error from Section 4.1)},\\ \frac{x\log^{k+1}q}{q^{1/2-\epsilon}} & \text{(error from Section 4.2)},\\ xu^{-u/2} & \text{(error from Section 4.3)}.\end{cases}$$

For the error from Section 4.1, we can get power savings by taking $(k+1)A_x = u = C\alpha$, for $C > 0$ sufficiently small. The error from Lemmas 3.1 and 4.1, after normalizing by x, are of the form $\alpha/((\log\log\eta)(\log\eta)^{1/6})$ and $1/\alpha$, respectively, which yields the optimal choice of $\alpha := (\log\log\eta)^{1/2}(\log\eta)^{1/12}$, thus establishing Theorem 1.2.

5. Acknowledgements

The author would like to thank Maksym Radziwiłł for suggesting this problem and for his continuous support throughout the research phase of this paper. The author would also like to thank Dimitris Koukoulopoulos, Mariusz Lemańczyk, Patrick Meisner, James Rickards, and Peter Zenz for many helpful discussions and comments. Finally, some of this work was conducted while the author was visiting CalTech; he is grateful for their hospitality.

References

[AKPLdlR17] Houcein El Abdalaoui, Joanna Kułaga-Przymus, Mariusz Lemańczyk, and Thierry de la Rue. The Chowla and the Sarnak conjectures from ergodic theory point of view. *Discrete and Continuous Dynamical Systems*, 37:2899–2944, 2017.

[Bur63] David A. Burgess. On character sums and L-series, II. *Proceedings of the London Mathematical Society*, s3-13(1):524–536, 1963.

[Cho65] Sarvadaman Chowla. *The Riemann Hypothesis and Hilbert's Tenth Problem*, volume 4 of *Mathematics and its Applications*. Gordon and Breach Science Publishers, New York-London-Paris, 1965.

[FKPL18] Sébastien Ferenczi, Joanna Kułaga-Przymus, and Mariusz Lemańczyk. *Sarnak's Conjecture – what's new. In: Ferenczi, S., Kułaga-Przymus, J., Lemańczyk, M. (eds) Ergodic Theory and Dynamical Systems in their Interactions with Arithmetics and Combinatorics - Lecture Notes in Mathematics, vol 2213*. Springer International Publishing, 2018.

[GK10] László Germán and Imre Kátai. Multiplicative functions at consecutive integers. *Lithuanian Mathematical Journal*, 50(1):43–53, 2010.

[GKL18] Alexander Gomilko, Dominik Kwietniak, and Mariusz Lemańczyk. *Sarnak's Conjecture Implies the Chowla Conjecture Along a Subsequence. In: Ferenczi S., Kułaga-Przymus J., Lemańczyk M. (eds) Ergodic Theory and*

Dynamical Systems in their Interactions with Arithmetics and Combinatorics - Lecture Notes in Mathematics, vol 2213. Springer, Cham, 2018.

[GLdlR20] Alexander Gomilko, Mariusz Lemańczyk, and Thierry de la Rue. Möbius orthogonality in density for zero entropy dynamical systems. *Pure and Applied Functional Analysis*, 5:1357–1376, 2020.

[Gra20] Andrew Granville. Sieving intervals and siegel zeros. Preprint, arXiv: math/2010.01211, 2020.

[HB83] David Rodney Heath-Brown. Prime twins and Siegel zeros. *Proc. Lond. Math.*, 47(3):193–224, 1983.

[Hil86] Adolf Hildebrand. Multiplicative functions at consecutive integers. *Math. Proc. Cambridge Philos. Soc.*, 100(2):229–236, 1986.

[HPW85] Glyn Harman, Jànos Pintz, and Dieter Wolke. A note on the Möbius and Liouville functions. *Stud. Sci. Math. Hung.*, 20:295–299, 1985.

[HR21] Harald A. Helfgott and Maksym Radziwiłł. Expansion, divisibility, and parity. Preprint, arXiv: math/2103.06853, 2021.

[IK04] Henryk Iwaniec and Emmanuel Kowalski. *Analytic number theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.

[Kou20] Dimitris Koukoulopoulos. *The Distribution of Prime Numbers*. Graduate Studies in Mathematics 203. American Mathematical Society, 2020.

[KPL20] Joanna Kułaga-Przymus and Mariusz Lemańczyk. Sarnak's conjectures from the ergodic theory point of view. To appear in Encyclopedia of Complexity and Systems Science, arXiv: math/2009.04757, 2020.

[MR16] Kaisa Matomäki and Maksym Radziwiłł. Multiplicative functions in short intervals. *Annals of Mathematics*, 183(3):1015–1056, 2016.

[MRT15] Kaisa Matomäki, Maksym Radziwiłł, and Terence Tao. An averaged form of Chowla's conjecture. *Algebra and Number Theory*, 9(9):2167–2196, 2015.

[MV07] Hugh L. Montgomery and Robert C. Vaughan. *Multiplicative Number Theory I: Classical Theory*. Cambridge studies in advanced mathematics. Cambrigde University Press, 1st edition, 2007.

[Sar11] Peter Sarnak. Three lectures on the Möbius function, randomness and dynamics, 2011. Lecture notes, http://publications.ias.edu/sarnak/.

[Sch76] Wolfgang M. Schmidt. *Equations over Finite Fields An Elementary Approach*. Lecture Notes in Mathematics. Springer-Berlin-Heidelberg, 1st edition, 1976.

[SS18] Will Sawin and Mark Shusterman. On the Chowla and twin primes conjectures over $\mathbb{F}_q[t]$. Preprint, arXiv:math/1808.04001, 2018.

[Tao16] Terence Tao. The logarithmically averaged Chowla and Elliott conjectures for two-point correlations. *Forum of Mathematics, Pi*, 4(8):1–36, 2016. doi:10.1017/fmp.2016.6.

[TT18] Terence Tao and Joni Teräväinen. Odd order cases of the logarithmically averaged Chowla conjecture. *Journal de Théorie de Nombres de Bordeaux*, 30(3):997–1015, 2018.

[TT19] Terence Tao and Joni Teräväinen. The structure of logarithmically averaged correlations of multiplicative functions, with applications to the Chowla and Elliott conjectures. *Duke Math J.*, 168(11):1977–2027, 2019. doi:10.1215/00127094-2019-0002.

Department of Mathematics and Statistics, McGill University, 805 Sherbrooke St. W., Montreal, QC H3A 2K6, Canada

Email address: iakov.chinis@gmail.com

Appendix A. The Fundamental Lemma of Sieve Theory

In this section, we present the Axioms of Sieve Theory, culminating in the Fundamental Lemma of Sieve Theory (FLST). We use the ideas presented here in order to move from a character sum over r-sifted integers to a character sum over one complete residue class (see Section 4). We follow Chapters 18 and 19 of [Kou20] and start off by listing some notation and the appropriate hypotheses needed for the FLST.

Let $\mathcal{A}$ denote a finite set of integers and let $\mathcal{P}$ denote a finite set of primes. We are interested in counting the number of elements of $\mathcal{A}$ which are relatively prime to $\mathcal{P}$; that is, we are interested in bounding the

quantity

$$S(\mathcal{A},\mathcal{P}) := \#\{a \in \mathcal{A} : (a,\mathcal{P}) = 1\},$$

where $(a,\mathcal{P}) = 1$ means that a has no prime factors in $\mathcal{P}$. In order to do so, it suffices to look at elements of $a \in \mathcal{A}$ which are divisible by $d|\prod_{p\in\mathcal{P}} p$; see Examples 18.1-18.6 in [Kou20]. More precisely, we are interested in having an asymptotic estimate for

$$\mathcal{A}_d := \#\{a \in \mathcal{A} : a \equiv 0 \pmod{d}\},$$

so we assume the following:

Axiom A.1. *There exists a multiplicative function ν, a parameter X, and a sequence of remainders $(r_d)_{d|\mathcal{P}}$ such that*

$$\mathcal{A}_d = \frac{\nu(d)}{d}X + r_d \quad \textit{for all } d|\mathcal{P}$$

and

$$\nu(p) < p \quad \textit{for all } p \in \mathcal{P}.$$

Remark A.1. *Note that $d|\mathcal{P}$ is shorthand for $d|\prod_{p\in\mathcal{P}} p$.*

We should think of $\nu(p)$ as the number of residue classes modulo p we are "removing" in order to capture elements of our set $\mathcal{A}$ which are prime to $\mathcal{P}$. As such, we want some sort of control over $\nu(p)$:

Axiom A.2. *There are constants $\kappa, k \geq 0$ and $\epsilon \in (0,1]$ such that*

$$\sum_{p\in\mathcal{P}\cap[1,\omega]} \frac{\nu(p)\log p}{p} = \kappa \log\omega + \mathcal{O}(1) \quad \textit{for all } \omega \leq \max\mathcal{P}$$

and

$$\nu(p) \leq \min\{k, (1-\epsilon)p\} \quad \textit{for all } p \in \mathcal{P}.$$

Assuming that Axioms A.1 and A.2 hold, we are able to compute $S(\mathcal{A},\mathcal{P})$:

Lemma A.1 (The Fundamental Lemma of Sieve Theory)**.** *Suppose $\mathcal{A}$ and $\mathcal{P}$ satisfy Axioms A.1 and A.2 for some constants $\kappa, k \geq 0$, $\epsilon \in (0,1]$, and let $y = \max\mathcal{P}$. Then:*

$$S(\mathcal{A},\mathcal{P}) = (1 + \mathcal{O}_{\kappa,k,\epsilon}(u^{-u/2}))X\prod_{p\in\mathcal{P}}\left(1 - \frac{\nu(p)}{p}\right) + \mathcal{O}\left(\sum_{d\leq y^u, d|\mathcal{P}} |r_d|\right),$$

uniformly for $u \geq 1$.

Proof. See [Kou20, Chapters 18-19]. □

Remark A.2. *We call $D = y^u$ the **level of distribution** of the sieve. As Koukoulopoulos remarks in his book, the level of distribution is "a measure of how well we can control the distribution of $\mathcal{A}$ among the progressions $0 \pmod d$," with $d|\mathcal{P}$. In order to control the level of distribution, we often use a "preliminary sieve," which removes integers with smaller prime factors and then use another sieve to remove larger primes. This is exactly what we do in the proof of Theorem 1.2 by sieving out the integers whose r-smooth part is "large."*

Appendix B. Character Sums

The key to the work of Germán and Kátai is to approximate the Liouville function by a real, primitive Dirichlet character $\chi \pmod q$ on "large" primes, so that, with the help of some sieve theory, we can change

our problem of bounding k-point correlations of λ to one of bounding character sums with a polynomial argument, which are well understood. For our purposes, we need to bound character sums of the form

$$\sum_{n \pmod q} \chi(f(n)),$$

where f is some polynomial with integer coefficients which can be factored into distinct linear factors, with q equal to the conductor/modulus of the real, primitive character χ. There are various instances of these types of bounds when the conductor q is a prime, dating back to the work of Weil on the Riemann Hypothesis over finite function fields; see, [Bur63], for example, and also [Sch76] for an elementary approach to understanding curves over finite fields. In [GK10], they look at $f(x) = x(x+1)$, it which case it is known that the above character sum is exactly equal to -1. For general f, we have the following, due to Weil:

Lemma B.1 (Weil)**.** *Let χ be a Dirichlet character modulo p of order $d|(p-1)$. If $f \in \mathbb{Z}[x]$ is not a d-th power modulo p (i.e., $f(x) \not\equiv cg(x)^d \pmod p$ identically for any $c \in \mathbb{Z}$ and any $g \in \mathbb{Z}[x]$) and if f has m distinct roots modulo p, then:*

$$\left| \sum_{n \pmod p} \chi(f(n)) \right| \leq (m-1)p^{1/2},$$

where the sum runs over an entire residue class modulo the prime p.

Proof. See [Sch76, Theorem 2C' (pg. 43)] (or even [IK04, Theorem 11.23]/[MV07, Lemma 9.25]). □

Our goal is to apply Lemma B.1 with χ a real, primitive Dirichlet character modulo q, with q not necessarily a prime. Fortunately for us, all such characters have conductor $q = 2^j m$, where $j \leq 3$ and where m is an odd, squarefree integer; see [MV07, Section 9.3], for example. Furthermore, the Chinese Remainder Theorem allows us to write each $n \pmod q$ uniquely as

$$n = a_1 \frac{q}{p_1^{\alpha_1}} + \cdots + a_s \frac{q}{p_s^{\alpha_s}},$$

for any $q = p_1^{\alpha_1} \cdots p_s^{\alpha_s}$, with a_i varying over a complete residue class modulo $p_i^{\alpha_i}$. In particular, for characters χ modulo q such that $\chi = \chi_1 \cdots \chi_s$ with χ_i a character modulo $p_i^{\alpha_i}$,

$$\begin{aligned}\sum_{n \pmod q} \chi(f(n)) &= \sum_{a_1,\ldots,a_s} \prod_{i=1}^{s} \chi_i \left(f\left(a_1 \frac{q}{p_1^{\alpha_1}} + \cdots + a_s \frac{q}{p_s^{\alpha_s}} \right) \right) \\ &= \prod_{i=1}^{s} \sum_{a_i \pmod{p_i^{\alpha_i}}} \chi_i \left(f\left(a_i \frac{q}{p_i^{\alpha_i}} \right) \right),\end{aligned}$$

where the last line follows from the fact that χ_i is periodic with period $p_i^{\alpha_i}$. Then, using the fact that every real, primitive character $\chi \pmod q$ can be written uniquely as $\chi = \chi_1 \cdots \chi_s$ with each χ_i being real and primitive, Lemma B.1 implies that

$$\sum_{n \pmod q} \chi(f(n)) \ll (\deg(f) - 1)^s q^{1/2},$$

provided f is not a square modulo p for all but finitely many primes p. Setting $N = \deg(f) - 1$ and noting that $\omega(q) = s$, we then have that $N^{\omega(q)} \leq \tau_N(q) \ll q^\epsilon$, for any $\epsilon > 0$, so that

$$\sum_{n \pmod q} \chi(f(n)) \ll q^{1/2+\epsilon},$$

for any real, primitive Dirichlet character modulo q, provided f is not a square modulo p for all but finitely-many p.

Appendix C. Parametrization

In this section, we briefly discuss how to use the Smith Normal Form of a matrix in order to solve a system of Diophantine equations. Our ultimate goal is to apply this technique in order to show that the solutions of the following system of integer equations

$$\begin{cases} a_1 b_1 = a_0 b_0 + h_1, \\ \vdots \\ a_k b_k = a_0 b_0 + h_k, \end{cases}$$

in the unknowns $b_0, b_1, \ldots, b_k$, are given by

$$b_i = b_i^* + m\frac{\mathrm{lcm}(a_0, a_1, \ldots, a_k)}{a_i} =: b_i^* + m a_i^*,$$

where $(b_0*, b_1^*, \ldots, b_k^*)$ is one particular solution of the system and where m ranges over all the integers.

Remark C.1. *In the case where $k = 1$ and $(a_0, a_1) = 1$, Bezout's Lemma tells us that the b_i's can be parametrized as*

$$b_i = b_i^* + m\frac{a_1 a_2}{a_i},$$

where (b_0^*, b_1^*) is one particular solution of the system and where m ranges over all integers. The proof of Bezout's Lemma follows from the Chinese Remainder Theorem; it is very likely that the proof generalizes, but we prefer to use a more direct method. Also, it is clear that such b_i's generate a set of solutions and seems likely that one could show that all solutions must be of the form above. In any case, the SNF gives us a versatile tool to handle more general cases.

Let A be an $m \times n$ matrix with integer entries and consider the system $AX = C$, for a given integer matrix C. Then, there exist invertible matrices U and V with integer entries such that $B := UAV$ is (almost) diagonal: in general, B may not be a square matrix, but the non-diagonal entries will be zero. We call B the *Smith Normal Form* of A and finding the matrices U and V amounts to using limited versions of the elementary row and column operations which preserve integer entries: since we are looking for invertible matrices U and V with integer entries, we must ensure that whatever operations we apply to the matrix A will preserve our integer entries. What is important here is that solving $AX = C$ in the integers is equivalent to making the change of variable $Y = V^{-1}X$ and solving the system $BY = D$, where $D := UC$. In particular, the original system will have integer solutions iff $b_{ii}y_i = d_i$ for all i (where D is a column matrix and d_i is the entry in row i). This last system is then solvable over the integers iff $b_{ii}|d_i$ whenever $b_{ii} \neq 0$ and $d_i = 0$

whenever $b_{ii} = 0$, in which case,

$$X = V \begin{bmatrix} \frac{b_{11}}{d_1} \\ \frac{b_{22}}{d_2} \\ \vdots \\ \frac{b_{kk}}{d_k} \\ f_{k+1} \\ \vdots \\ f_n \end{bmatrix},$$

where the b_{ii}'s are arranged so that $b_{ii} \neq 0$ for all $i = 1, \ldots, k$ and where $f_{k+1}, \ldots, f_n$ are arbitrary integers (representing the $n - k$ free variables).

The above decomposition hinges on our ability to find invertible matrices U and V such that UAV is diagonal. We illustrate how to do this for the following system of equations in the unknowns $b_0, b_1, \ldots, b_k$,

$$\begin{cases} a_1 b_1 = a_0 b_0 + h_1, \\ \vdots \\ a_k b_k = a_0 b_0 + h_k, \end{cases}$$

proving each solution $(b_0, b_1, \ldots, b_k)$ can be parametrized as

$$b_i = b_i^* + m a_i^*,$$

where $(b_0^*, b_1^*, \ldots, b_k^*)$ is one particular solution of the system (assuming that the system is solvable) and where

$$a_i^* := \frac{\operatorname{lcm}(a_0, a_1, \ldots, a_k)}{a_i},$$

for $i = 0, 1, \ldots, k$.

To show the above, we simply use the algorithm which produces the SNF of A. So, let

$$A := \begin{bmatrix} a_0 & -a_1 & 0 & \ldots & 0 & 0 \\ 0 & a_1 & -a_2 & \ldots & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & \ldots & a_{k-1} & -a_k \end{bmatrix},$$

$$X := \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_k, \end{bmatrix}$$

and

$$C := \begin{bmatrix} -h_1 \\ h_1 - h_2 \\ \vdots \\ h_{k-1} - h_k \end{bmatrix}.$$

Our goal it to solve the system $AX = C$ over the integers. To begin, let $d_{0,1} := (a_0, a_1)$. Then, there exist integers $x_{0,1}, y_{0,1}$ such that $d_{0,1} = a_0 x + a_1 y$; in particular, $1 = x_{0,1} a_0 / d_{0,1} + y_{0,1} a_1 / d_{0,1}$. Embedding

this information into a $(k+2)\times(k+2)$ identity matrix V_1, we can get the following:

$$AV_1 = \begin{bmatrix} a_0 & -a_1 & 0 & \dots & 0 & 0 \\ 0 & a_1 & -a_2 & \dots & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & \dots & a_{k-1} & -a_k \end{bmatrix} \begin{bmatrix} x_{0,1} & a_1/d_{0,1} & 0 & \dots & 0 \\ -y_{0,1} & a_0/d_{0,1} & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

$$= \begin{bmatrix} d_{0,1} & 0 & 0 & \dots & 0 & 0 \\ -a_1 y & a_0 a_1/d_{0,1} & -a_2 & \dots & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & \dots & a_{k-1} & -a_k \end{bmatrix}.$$

To get zeros below the leading variable, consider the $(k+1)\times(k+1)$ matrix U_1 defined by

$$U_1 := \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ a_1 y & d_{0,1} & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

and note that

$$U_1 A V_1 = \begin{bmatrix} d_{0,1} & 0 & 0 & \dots & 0 & 0 \\ 0 & a_0 a_1 & -d_{0,1} a_2 & \dots & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & \dots & a_{k-1} & -a_k \end{bmatrix}.$$

We continue in the same manner: let $d_{01,2} = (a_0 a_1, d_{0,1} a_2) = d_{0,1}(a_0 a_1/d_{0,1}, a_2)$, then there exist integers $x_{01,2}, y_{01,2}$ such that $d_{01,2} = a_0 a_1 x_{01,2} + d_{0,1} a_2 y_{01,2}$. Embedding this information in another $(k+2)\times(k+2)$ identity matrix V_2, we have that

$$U_1 A V_1 V_2 = \begin{bmatrix} d_{0,1} & 0 & 0 & \dots & 0 & 0 \\ 0 & a_0 a_1 & -d_{0,1} a_2 & \dots & 0 & 0 \\ 0 & 0 & a_2 & \dots & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & \dots & a_{k-1} & -a_k \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & x_{01,2} & a_2 d_{0,1}/d_{01,2} & \dots & 0 \\ 0 & -y_{01,2} & a_0 a_1/d_{01,2} & \dots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

$$= \begin{bmatrix} d_{0,1} & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & d_{01,2} & 0 & 0 & \dots & 0 & 0 \\ 0 & -a_2 y_{01,2} & a_0 a_1 a_2/d_{01,2} & -a_3 & \dots & 0 & 0 \\ \vdots & & & & & & \\ 0 & 0 & 0 & 0 & \dots & a_{k-1} & -a_k \end{bmatrix}$$

To get zeros below the leading variable, consider the $(k+1)\times(k+1)$ matrix U_2 defined by

$$U_2 := \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & -a_2 y_{01,2} & d_{01,2} & 0 & \dots & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

and note that

$$U_2U_1AV_1V_2 = \begin{bmatrix} d_{0,1} & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & d_{01,2} & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & a_0a_1a_2 & -a_3d_{01,2} & \dots & 0 & 0 \\ \vdots & & & & & & \\ 0 & 0 & 0 & 0 & \dots & a_{k-1} & -a_k \end{bmatrix}$$

Continuing by induction, we see that the Smith Normal Form of the matrix A has diagonal entries $d_{01\cdots j,j+1}$, for $j = 0, 1, \dots, k-1$, which are defined recursively by

$$d_{01\cdots j,j+1} := \gcd(a_0a_1\cdots a_j, a_{j+1}d_{01\cdots j-1,j})$$

with

$$d_{0,1} := \gcd(a_0, a_1).$$

What is important to note here is that the SNF of A has the maximal rank; in particular, if the SNF satisfies some nice divisibility properties in relation to the matrix $D = UC$, we get infinitely-many solutions which are parametrized by exactly one free variable (because we have full rank, but the matrix A has dimensions $k\times(k+1)$). Furthermore, the solutions will then be given by

$$X = V\begin{bmatrix} \frac{b_{11}}{d_1} \\ \frac{b_{22}}{d_2} \\ \vdots \\ \frac{b_{kk}}{d_k} \\ f_{k+1} \end{bmatrix}$$

and all that is left for us to do is to compute the last column of the matrix V. A quick calculation shows that the last column of V has entry

$$\frac{a_0a_1\cdots a_{i-1}a_{i+1}\cdots a_k}{d_{01\cdots k-1,k}}$$

in row i, for $i = 0, 1, \dots, k$ and that this is equal to

$$\frac{\operatorname{lcm}(a_0, a_1, \dots, a_k)}{a_i},$$

as claimed.