Turun yliopisto
University of Turku

TOPICS IN MULTIPLICATIVE NUMBER THEORY

Joni Teräväinen

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
Sarja – ser. AI osa – tom. 589 | Astronomica – Chemica – Physica – Mathematica | Turku 2018

Turun yliopisto
University of Turku

TOPICS IN MULTIPLICATIVE NUMBER THEORY

Joni Teräväinen

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
Sarja - ser. A I osa - tom. 589 | Astronomica - Chemica - Physica - Mathematica | Turku 2018

University of Turku

Faculty of Science and Engineering
Department of Mathematics and Statistics

Supervised by

Docent Kaisa Matomäki
Department of Mathematics and Statistics
University of Turku
Turku, Finland

Reviewed by

Assistant Professor Adam Harper
Mathematics Institute
University of Warwick
Coventry, England, UK

Associate Professor Dimitris Koukoulopoulos
Department of Mathematics and Statistics
Université de Montréal
Montreal, Canada

Opponent

Professor Ben Green
Mathematical Institute
University of Oxford
Oxford, England, UK

ISBN 978-951-29-7339-2 (PRINT)
ISBN 978-951-29-7340-8 (PDF)
ISSN 0082-7002 (PRINT)
ISSN 2343-3175 (PDF)
Painosalama Oy – Turku, Finland 2018

Abstract

This thesis is comprised of four articles in multiplicative number theory, a subfield of analytic number theory that studies questions related to prime numbers and multiplicative functions. A central principle in multiplicative number theory is that multiplicative structures, such as the primes or the values of a multiplicative function, should not correlate with additive structures of various types. The results in this thesis can be interpreted as instances of this principle.

In the first article, we consider the problem of finding almost primes in almost all short intervals, which is a natural approximation to the problem of finding primes in short intervals. We show that almost all intervals of nearly optimal length contain a product of exactly three primes. For products of exactly two primes, we improve a result of Harman. The proofs are based on careful analysis of Dirichlet polynomials related to almost primes.

The second article is about the Goldbach problem for a sparse subset of the primes. Vinogradov famously showed that any large odd number is the sum of three primes, so it is natural to study the same problem with the summands coming from a subset of the primes. Improving a result of Matomäki, we show that a special set of primes, consisting of primes representable as one plus the sum of two squares, satisfies the ternary Goldbach problem. We also establish a number of other additive results for this same set of primes. The proofs use sieve methods and transference principles for additive equations in primes.

We also study the Möbius function and its autocorrelations. A famous conjecture of Chowla asserts that products of shifts of the Möbius function should have mean zero. In the third article, together with T. Tao we settle a logarithmic version of this conjecture in all the cases involving an odd number of shifts. This complements Tao's earlier result that the two-point Chowla conjecture holds with logarithmic weights.

Lastly, in the fourth article, we study binary correlations of multiplicative functions with logarithmic weights. We prove an asymptotic formula for these correlations for a wide class of multiplicative functions, extending an earlier result of Tao. We then derive a number of applications regarding the largest prime factors of consecutive integers, including a logarithmic version of a conjecture of Erdős and Turán. Moreover, we prove a new estimate for character sums over reducible quadratic polynomials.

Tiivistelmä

Tämä väitöskirja koostuu neljästä artikkelista multiplikatiivisessa lukuteoriassa, joka on alkulukuja ja multiplikatiivisia funktioita tutkiva analyyttisen lukuteorian haara. Keskeinen periaate multiplikatiivisessa lukuteoriassa on, että multiplikatiivisten objektien (kuten alkulukujen tai multiplikatiivisten funktioiden arvojen) ei pitäisi korreloida additiivisten objektien kanssa. Tämän väitöskirjan tulokset voidaankin tulkita kyseisen periaatteen ilmentyminä.

Ensimmäisessä artikkelissa tarkastelemme melkein alkulukujen löytämistä melkein kaikilta lyhyiltä väleiltä; tämä on luonnollinen approksimaatio alkulukujen löytämiselle lyhyiltä väleiltä. Osoitamme, että melkein kaikki välit, joiden pituus on lähes optimaalisen lyhyt, sisältävät tasan kolmen alkuluvun tulon. Tasan kahden alkuluvun tulojen tapauksessa parannamme Harmanin tulosta. Todistukset perustuvat melkein alkulukuihin liitettyjen Dirichlet'n polynomien tarkkaan analysointiin.

Toinen artikkeli koskee Goldbach-ongelmaa eräälle harvalle osajoukolle alkulukuja. Vinogradov osoitti kuuluisassa työssään, että jokainen riittävän suuri pariton luku on kolmen alkuluvun summa, joten on luonnollista tarkastella vastaavaa ongelmaa alkulukujen osajoukoille. Parantaen Matomäen tulosta osoitamme, että vastaus ternääriseen Goldbach-oneglmaan on positiivinen niiden alkulukujen joukolle, jotka voidaan esittää ykkösen ja kahden neliöluvun summana. Osoitamme myös useita muita additiivisia tuloksia samalle alkulukujen osajoukolle. Todistukset käyttävät seulamenetelmiä sekä ns. traansferenssiperiaatteita additiivisille yhtälöille alkulukujen joukossa.

Tutkimme myös Möbiuksen funktiota ja sen autokorrelaatioita. Chowlan kuuluisa konjektuuri väittää, että Möbiuksen funktioiden translaatioiden tuloilla pitäisi olla keskiarvo nolla. Kolmannessa artikkelissa yhdessä T. Taon kanssa ratkaisemme logaritmisen version tästä konjektuurista kaikissa tapauksissa, joissa translaatioiden määrä on pariton. Tämä täydentää Taon aikaisempaa tulosta, jonka mukaan kahden pisteen Chowlan konjektuuri pätee logaritmisilla painoilla.

Lopuksi neljännessä artikkelissa tutkimme multiplikatiivisten funktioiden binäärisiä korrelaatioita logaritmisilla painoilla. Todistamme asymptoottisen kaavan näille korrelaatioille, joka pätee laajalle luokalle multiplikatiivisia funktioita ja parantaa Taon aikaisempaa tulosta. Johdamme sitten useita sovelluksia koskien peräkkäisten lukujen suurimpia alkutekijöitä – mukaan lukien logaritmisen version eräästä Erdősin ja Turánin konjektuurista. Lisäksi todistamme uuden arvion karakterisummille yli jaollisen toisen asteen polynomin arvojen.

Acknowledgments

I am sincerely grateful to my advisor, Prof. Kaisa Matomäki, for all the excellent advice and guidance she has given me throughout my PhD studies. Our regular meetings, the funding she has offered for my studies and travels, and her constant valuable feedback on my works have greatly progressed my career. It has been an honor learning the latest techniques in analytic number theory from one of the masters in the field.

I am grateful to the reviewers of this thesis, Prof. Adam Harper and Prof. Dimitris Koukoulopoulos, for spending a lot of their valuable time to carefully read my dissertation and for giving detailed comments on it. Equally, I am grateful to Prof. Ben Green for agreeing to act as the opponent at my doctoral defense.

I thank Prof. Terry Tao for all the fruitful collaboration projects we have had so far, for answering so many of my questions on additive combinatorics, and for hosting my visit to UCLA in April 2018. Working with him has been a great pleasure and has made me more mature as a mathematician.

I would like to thank the sources of funding of my research, namely the Vilho, Yrjö and Kalle Väisälä Fund of the Finnish Academy of Science and Letters, the UTUGS Graduate School of the University of Turku, and the Academy of Finland. The support from these grants has made my work and participation in various conferences possible.

I express my thanks to my colleagues at the universities of Oxford, KTH and Nancy for inviting me for research visits and for making my stays very enjoyable and scientifically productive. I extend my thanks to my number theory colleagues at Åbo Akademi for organizing a stimulating joint seminar. I also thank Alexander Mangerel for careful reading of my thesis.

All my colleagues at the Department of Mathematics and Statistics at the University of Turku deserve thanks for making the work environment pleasant, productive, and relaxed. Specifically, I thank Matti Vuorinen for all the help and friendship he has offered me, and for the lovely dinners I have had at Matti and Sinikka's home. In addition, I thank my fellow number theory students, Jori Merikoski and Juho Salmensuu, for our many interesting number-theoretic lunch discussions.

Special thanks go to all my friends for making my time as a graduate student so fun and enjoyable. Especially, I thank Jesse Jääsaari for the many fascinating discussions we have had about mathematics and other topics, and for great company during our many conference trips together. Likewise, I thank Otte Heinävaara for the countless hours we have spent solving and devising mathematical problems, and for the long, adventurous cycling trips we have taken through Finland.

I warmly thank my family for always supporting me and believing in me. Especially, I thank my parents, grandparents and siblings, whose help I can always count on and who have always encouraged me to pursue my dreams. Lastly, and most importantly, I thank my fiancée Parisa for unconditional love and support, and for bringing joy to the long days I spent working on this thesis.

Turku, June 2018
Joni Teräväinen

List of original publications

This thesis contains the following four publications.

[I] J. Teräväinen: Almost primes in almost all short intervals. *Math. Proc. Cambridge Philos. Soc.*, 161(2):247–281, 2016. DOI: 10.1017/S0305004116000232

[II] J. Teräväinen: The Goldbach problem for primes that are sums of two squares plus one. *Mathematika*, 64(1):20–70, 2018. DOI: 10.1112/S0025579317000341

[III] T. Tao and J. Teräväinen: Odd order cases of the logarithmically averaged Chowla conjecture. To appear in *J. Théor. Nombres Bordeaux*. arXiv: 1710.02112 [Math.NT]

[IV] J. Teräväinen: On binary correlations of multiplicative functions. *Forum Math. Sigma*, 6:e10, 41, 2018. DOI: 10.1017/fms.2018.10

The articles are included here with permission from the publishers.

Table of Contents

Summary

1. Notations and conventions

We collect here the notations used in the summary sections 2–6. The Articles [I]–[IV] have their own notation sections.

1.1. Sets

- $\mathbb{N}$ – the set of positive integers $\{1, 2, 3, \ldots\}$.
- $\mathbb{Z}$ – the set of all integers.
- $\mathbb{Z}_N$ – the ring of integers mod N.
- $\mathbb{P}$ – the set of prime numbers.
- Squarefree integers – integers $n \geq 1$ such that n is not divisible by p^2 for primes p.
- P_k – the set of integers with at most k prime factors, counting multiplicities.
- E_k – the set of integers with exactly k prime factors, counting multiplicities.
- $\mathbb{D}$ – the unit disk $\{z \in \mathbb{C} : \ |z| \leq 1\}$.
- $1_S(n)$ – the indicator function of a set S, equaling 1 if $n \in S$ and 0 otherwise.

1.2. Letters

- d, k, ℓ, m, n – positive integers.
- $p, p_1, p_2, \ldots$ – prime numbers.
- ε – an arbitrarily small positive constant.
- W – the product of primes in $[1, w]$ for some large w.

1.3. Arithmetic functions

- $\varphi(n)$ – the Euler function, giving the number of integers $1 \leq j \leq n$ coprime to n.
- $\Lambda(n)$ – the von Mangoldt function, which equals $\log p$ if $n = p^k$ for some prime p and some $k \geq 1$, and equals 0 if no such p exists.
- $\Omega(n)$ – number of prime factors of n, counted with multiplicities.
- $\lambda(n)$ – the Liouville function, given by $\lambda(n) := (-1)^{\Omega(n)}$.
- $\mu(n)$ – the Möbius function, given by $\mu(n) := \lambda(n) 1_{n \text{ squarefree}}$.
- $P^+(n)$ – the largest prime factor of n, with $P^+(1) := 1$.
- $\pi(x)$ – the number of prime numbers in $[1, x]$.
- Multiplicative function – a function $g : \mathbb{N} \to \mathbb{C}$ satisfying $g(mn) = g(m)g(n)$ whenever $m, n \in \mathbb{N}$ are coprime.
- $\lambda_d^{+,\mathrm{LIN}}$, $\lambda_d^{-,\mathrm{LIN}}$ – the upper and lower bound linear sieve weights. Given a level D and a sifting parameter z, they are equal to 1 for $d = 1$ and are equal to the Möbius function $\mu(d)$ for $d \geq 2$ belonging to the sets

  $$\mathcal{D}^{+,\mathrm{LIN}} := \{p_1 \cdots p_r \leq D : \ z > p_k > p_{k+1}, \ p_1 \cdots p_{2k-2} p_{2k-1}^3 \leq D \ \forall k \geq 1\},$$
  $$\mathcal{D}^{-,\mathrm{LIN}} := \{p_1 \cdots p_r \leq D : \ z > p_k > p_{k+1}, \ p_1 \cdots p_{2k-1} p_{2k}^3 \leq D \ \forall k \geq 1\}.$$

  For other values of d, they are equal to 0.
- $\lambda_d^{+,\mathrm{SEM}}$, $\lambda_d^{-,\mathrm{SEM}}$ – the upper and lower bound semilinear sieve weights. Given a level D and a sifting parameter z, they are equal to 1 for $d = 1$ and are equal to the Möbius function $\mu(d)$ for $d \geq 2$ belonging to the sets

$$\mathcal{D}^{+,\mathrm{SEM}} := \{p_1 \cdots p_r \le D : \ z > p_k > p_{k+1}, \ p_1 \cdots p_{2k-2} p_{2k-1}^2 \le D \ \forall k \ge 1\},$$
$$\mathcal{D}^{-,\mathrm{SEM}} := \{p_1 \cdots p_r \le D : \ z > p_k > p_{k+1}, \ p_1 \cdots p_{2k-1} p_{2k}^2 \le D \ \forall \ k \ge 1\}.$$

For other values of d, they are equal to 0.

1.4. Analysis

- $f(x) = o(g(x))$ – we have $\lim_{x\to\infty} f(x)/g(x) = 0$.
- $f(x) \ll g(x)$ – we have, for some constant C, $|f(x)| \le C|g(x)|$ for all large enough x.
- $f(x) \asymp g(x)$ – we have $f(x) \ll g(x)$ and $g(x) \ll f(x)$.
- $o_{\varepsilon\to 0}(1)$ – an unspecified function $f(\varepsilon)$ tending to 0 as $\varepsilon \to 0$.
- Almost all – for a proposition $P(n)$, we say that it holds for almost all $n \in \mathbb{N}$ if $|\{n \le X : \ P(n) \text{ fails}\}| = o(X)$. Analogously, we say that $P(n)$ holds for almost all even n if $|\{n \le X : \ n \equiv 0 \pmod 2, \ P(n) \text{ fails}\}| = o(X)$.
- $e(\alpha)$ – the additive character $e^{2\pi i\alpha}$.
- $\|x\|$ – the distance from $x \in \mathbb{R}$ to the nearest integer.
- $\sum_{p\in I}$ – summation over primes in I, whenever I is an interval.
- $\widehat{f}(\xi)$ – the discrete Fourier transform of $f : \mathbb{Z}_N \to \mathbb{C}$, given by
$$\widehat{f}(\xi) := \frac{1}{N} \sum_{n\in\mathbb{Z}_N} f(n) e\left(-\frac{\xi n}{N}\right).$$

1.5. Probability theory

- $\mathbf{P}(A)$ – the logarithmic probability of $A \subset [1, x]$, given by
$$\mathbf{P}(A) := \frac{\sum_{\substack{n\le x\\ n\in A}} \frac{1}{n}}{\sum_{n\le x} \frac{1}{n}}.$$
- $\mathbf{H}(\mathbf{X})$ – the entropy of a random variable $\mathbf{X}$ having a finite range $\mathcal{X}$. This is defined by
$$\mathbf{H}(\mathbf{X}) := \sum_{x\in\mathcal{X}} \mathbf{P}(\mathbf{X} = x) \log \frac{1}{\mathbf{P}(\mathbf{X} = x)}.$$
- $\mathbf{H}(\mathbf{X}, \mathbf{Y})$ – the joint entropy of two random variables $\mathbf{X}$ and $\mathbf{Y}$ with finite ranges $\mathcal{X}$ and $\mathcal{Y}$, respectively. This equals the entropy of the random variable $(\mathbf{X}, \mathbf{Y})$ that takes values in $\mathcal{X} \times \mathcal{Y}$.
- $\mathbf{H}(\mathbf{X}|\mathbf{Y})$ – the conditional entropy of $\mathbf{X}$ given $\mathbf{Y}$;
$$\mathbf{H}(\mathbf{X}|\mathbf{Y}) := \mathbf{H}(\mathbf{X}, \mathbf{Y}) - \mathbf{H}(\mathbf{Y}).$$
- $\mathbf{I}(\mathbf{X}, \mathbf{Y})$ – the mutual information between two random variables $\mathbf{X}$ and $\mathbf{Y}$;
$$\mathbf{I}(\mathbf{X}, \mathbf{Y}) = \mathbf{H}(\mathbf{X}) + \mathbf{H}(\mathbf{Y}) - \mathbf{H}(\mathbf{X}, \mathbf{Y}).$$

1.6. Miscellaneous

- (m,n) – the greatest common divisor of m and n.
- $\mathbb{D}(f,g;x)$ – the pretentious distance between f and g, defined in formula (6.1).
- $\mathcal{P}(z)$ – the product of all the primes in $[1,z)$.
- $\|a\|_{U^k(\mathbb{Z}_N)}$ the U^k Gowers norm of $a:\mathbb{Z}_N\to\mathbb{C}$, defined recursively as

$$\|a\|_{U^1(\mathbb{Z}_N)} := \left|\frac{1}{N}\sum_{n\in\mathbb{Z}_N} f(n)\right|, \quad \|a\|_{U^{k+1}(\mathbb{Z}_N)} := \left(\frac{1}{N}\sum_{t\in\mathbb{Z}_N} \|a\cdot\overline{a_t}\|_{U^k(\mathbb{Z}_N)}^{2^k}\right)^{1/2^{k+1}},$$

  where $a_t(n) := a(n+t)$.
- $\|a\|_{U^k[N]}$ – the U^k Gowers norm of $a:[1,N]\to\mathbb{C}$, defined by

$$\|a\|_{U^k[N]} := \frac{\|1_{[1,N]}\cdot a\|_{U^k(\mathbb{Z}_{2N+1})}}{\|1_{[1,N]}\|_{U^k(\mathbb{Z}_{2N+1})}}.$$

2. Introduction

Multiplicative number theory is an area of analytic number theory where one studies the distributional properties of multiplicative functions, prime numbers, and other sets possessing multiplicative structure. A fundamental principle in this area is that *multiplicative structures* (such as the primes, or the values of a multiplicative function) should behave independently of *additive structures* (such as intervals, additive equations, or arithmetic progressions). This gives rise to a number of conjectures, many of which are still open.

One instance of this principle is that the primes are expected to be distributed somewhat uniformly on very short intervals, such as $[x, x + (\log x)^c]$ with $c > 1$ and $x \in \mathbb{N}$ large. Under the Riemann hypothesis, Selberg proved that for $c = 2 + \varepsilon$ this holds, at least for *almost all* such intervals. A famous conjecture of Cramér [8] asserts that for $c = 2 + \varepsilon$ there should be a prime on $[x, x + (\log x)^c]$ for large x, even without any exceptional intervals, but this is not known even conditional on the Riemann hypothesis. In Article [I], we show, improving the work of Harman [35], that almost all intervals $[x, x + (\log x)^{3.51}]$ contain a product of exactly two primes, and almost all intervals $[x, x + (\log x)^{1+\varepsilon}]$ contain a product of exactly three primes for any $\varepsilon > 0$, the latter result being nearly optimal. Numbers that are the product of exactly two or three primes can be thought of as approximations to the primes, and they have a rigid multiplicative structure in particular. It turns out that these *almost primes* (discussed in detail in Section 3) have more flexibility than the primes, and this is what enables us to prove much stronger results about them.

Another conjecture that combines multiplicative and additive structures is the binary Goldbach conjecture, dating from 1742 and stating that every even $n \geq 4$ is the sum of two primes. This remains an important open problem. On the other hand, the ternary version of the problem, to the effect that every odd number $n \geq 7$ is the sum of three primes, was proved by Vinogradov [108] in 1937 for all sufficiently large integers, and by Helfgott [43] in 2013 in the remaining cases. The essence of many problems of Goldbach-type is showing that the primes do not correlate with "additive sets" (such as the so-called Bohr sets, defined in Section 4.2). The binary Goldbach conjecture has been known since the 1930s to be true for almost all even n, so a natural question to examine is whether the conjecture remains true in almost all cases when one only uses summands coming from a specific subset of the primes. Improving a result of Matomäki [70], we show in Article [II] that this question has a positive answer for primes represented by the polynomial $x^2 + y^2 + 1$; this subset of the primes has also been studied in several other contexts [49], [115], [69] and is an example of a sparse subset of the primes (that is, it has relative density 0 within the primes). Continuing with the theme of additive problems in the primes, we show that the primes of the form $x^2 + y^2 + 1$ also contain infinitely many three-term arithmetic progressions, and that the numbers αp, where α is a fixed irrational and

p runs through such primes, are “well-distributed” modulo 1.

Turning our attention from primes to multiplicative functions, we also study the Liouville function, $\lambda(n)$, whose value is determined by the parity of the number of prime factors of n. Chowla [7] posed in the 1960s the famous conjecture that the Liouville function should behave independently along strings of consecutive integers, taking any sequence of $+1$s and -1s with equal probability. More precisely, Chowla's conjecture can be written as the statement that

$$\frac{1}{x}\sum_{n\leq x}\lambda(n+h_1)\cdots\lambda(n+h_k)=o(1)$$

for any fixed $k\geq 1$ and distinct $h_1,\ldots,h_k\in\mathbb{N}$. Thus, for example, the probability that the Liouville function takes value $+1$ at both n and $n+1$ should be $\frac{1}{4}$, the product of the individual probabilities of the events $\lambda(n)=1$ and $\lambda(n+1)=1$ (which have probability $\frac{1}{2}$). During the last few years, there has been a lot of research activity surrounding Chowla's conjecture, and several approximations to the conjecture have been proved (see Section 5 for descriptions of them). In particular, Tao [98] showed in 2015 that the two-point case of Chowla's conjecture holds with logarithmic weights, in the sense that

$$\frac{1}{\log x}\sum_{n\leq x}\frac{\lambda(n+h_1)\lambda(n+h_2)}{n}=o(1)$$

for any distinct $h_1,h_2\in\mathbb{N}$. In Article [III], jointly with Tao, we consider the higher order cases and show that for odd values of k the k-point Chowla conjecture holds with logarithmic weights. Our proof uses combinatorial tools, such as the theory of Gowers norms, and is independent of and simpler than our earlier proof of the same result in [100].

The Liouville function is an archetypal example of a multiplicative function, so it is natural to believe that also shifts of more general multiplicative functions are independent of each other under suitable assumptions. This was made precise by Elliott [11] in the 1990s; he conjectured that one has the discorrelation estimate

$$\frac{1}{x}\sum_{n\leq x}g_1(n+h_1)\cdots g_k(n+h_k)=o(1),$$

whenever $g_1,\ldots,g_k$ are multiplicative functions that take values in the unit disc, $h_1,\ldots,h_k\in\mathbb{N}$ are distinct shifts, and one of the functions g_j is *non-pretentious* in a suitable sense (we elaborate on this in Section 6). In 2015, Tao [98] proved that Elliott's conjecture holds for $k=2$ with logarithmic weights, in the sense that

$$\frac{1}{\log x}\sum_{n\leq x}\frac{g_1(n+h_1)g_2(n+h_2)}{n}=o(1)$$

under the same assumptions. In Article [IV], we show that Tao's result on the two-point logarithmic Elliott conjecture can be extended to a wider class of real-valued multiplicative functions (with a main term in the asymptotic). This wider class turns out to contain many functions of interest, such as indicator functions related to *smooth numbers* (see Section 6 for details). Making use of this, we prove a logarithmic version of a conjecture of Erdős and Turán [94] on the largest prime factors of n and $n+1$. We also show that certain other sets constructed from multiplicative functions behave independently at n and $n+1$, as one would expect from the heuristic discussed above.

The structure of this thesis is as follows. In Sections 3, 4, 5 and 6, we introduce the topics of the articles [I], [II], [III] and [IV], respectively, and give a wealth of references to the literature on these and related questions. This is followed by the original publications in the same order. The preprint versions of these publications can also be found on the arXiv.org preprint server.

3. Almost primes in very short intervals

In article [I], we study the problem of finding almost primes in almost all short intervals. Since almost primes (defined subsequently) are an approximation to prime numbers, we begin with an overview of conjectures and results on primes in short intervals.

3.1. **Heuristics and conjectures for primes in short intervals**

The prime number theorem, a cornerstone in classical analytic number theory, states that the number of primes $\pi(x)$ up to x satisfies the asymptotic relation

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

Interpreted probabilistically, this means that an integer $n \leq x$ chosen uniformly at random is prime with probability $(1 + o(1))/(\log x)$. Based on this, H. Cramér [8] introduced in the 1930s the heuristic model that the indicator function $1_{\mathbb{P}}(n)$ of primes should behave for $n \leq x$ like a random variable $\mathbf{X}_n \in \{0, 1\}$ that equals 1 with probability $1/\log x$. Moreover, he made the strong assumption that the $\mathbf{X}_n$ are jointly independent of each other; this property of course does not hold for the primes as such (since both n and $n + 1$ cannot be prime for $n \geq 3$), but it serves as a good approximation in various problems[1]. Cramér then deduced from basic probability theory that if the $\mathbf{X}_n$ are as above, then the sum $\sum_{x-\lambda \log x \leq k < x} \mathbf{X}_k$ is Bernoulli distributed with mean λ, and further that the Bernoulli distribution is very closely approximated by the Poisson distribution with the same mean λ (in the regime where $\lambda > 0$ is fixed and $x \to \infty$). Thus, if the model of the primes as the random variables $\mathbf{X}_n$ is adequate, the primes follow the Poisson distribution in short intervals, in the sense that

$$\frac{1}{x}|\{n \leq x : \ \pi(n + \lambda \log x) - \pi(n) = k\}| = (1 + o(1))e^{-\lambda}\frac{\lambda^k}{k!} \tag{3.1}$$

for any fixed $\lambda > 0$ and $k \in \mathbb{N}$. There is strong evidence in support of (3.1), as Gallagher [20] showed that it would follow from a certain uniform version of the widely believed Hardy–Littlewood prime tuples conjecture (for the non-uniform version, see Subsection 5.1). From (3.1) one can deduce many further (yet unproved) properties of the primes in short intervals; in particular, letting λ grow slowly with x and taking $k = 0$, (3.1) would imply that, for any function $\psi(x) \to \infty$ as $x \to \infty$, we have

$$\pi(x + \psi(x)\log x) - \pi(x) \geq 1 \tag{3.2}$$

for almost all $x \in \mathbb{N}$. By a more careful analysis of the tails of the Poisson distribution, one could similarly infer the stronger statement

$$\pi(x + \psi(x)\log x) - \pi(x) = (1 + O(\varepsilon))\psi(x) \tag{3.3}$$

[1] There are more elaborate versions of Cramér's model that take into account local obstructions; see [23].

for almost all $x \in \mathbb{N}$, for any for any fixed $\varepsilon > 0$ and any given function $\psi(x) \to \infty$ with $\psi(x) \le x$ (and with the implied constant in the $O(\cdot)$ notation being absolute). Moreover, under the assumption that (3.1) holds uniformly for $\lambda \le C \log x$ with $C \ge 1$, the right-hand side of (3.1) is $\ll 1/x$, so one ends up with the following conjecture[2].

3.1. **Conjecture** (Cramér's conjecture)**.** There exists a constant $C > 0$ such that the interval $[x, x + C(\log x)^2]$ contains a prime for all large enough x.

It seems that Cramér's conjecture is out of reach even under the Riemann hypothesis. Nevertheless, Selberg [91] showed that the Riemann hypothesis implies a version of Cramér's conjecture for *almost all* x.

3.2. **Theorem** (Selberg)**.** *Assume the Riemann hypothesis. Then, for any function* $\psi(x)$ *tending to infinity as* $x \to \infty$*, almost all intervals* $[x, x + \psi(x)(\log x)^2]$ *contain a prime.*

When it comes to the existence of primes in all short intervals, the best statement known under the Riemann hypothesis is that $[x, x + x^{1/2} \log x]$ contains a prime for all large x; see [90]. This remains very far from intervals of polylogarithmic length. Let us mention in passing that Cramér's model also gives probabilistic evidence for the Riemann hypothesis; namely, if one redefines the random variables $\mathbf{X}_n$ slightly to take the value 1 with probability $1/\log n$, then one can use basic properties of random walks to show that for any fixed $\varepsilon > 0$ we have

$$\sum_{n \le x} \mathbf{X}_n - \int_2^x \frac{dt}{\log t} \ll x^{1/2+\varepsilon}$$

with probability 1, and the corresponding statement for $1_{\mathbb{P}}(n)$ in place of $\mathbf{X}_n$ is well-known to be equivalent to the Riemann hypothesis.

The above indicates that results (whether conditional or unconditional) one can prove about primes in almost all intervals tend to be considerably stronger than what can be proved about primes in all short intervals. One fact that complicates the study of primes in all short intervals is that there are actually some short intervals where the primes notably deviate from their typical behavior. Namely, Maier [67] showed in 1985 in a seminal work that, given any $C > 0$, there is a constant $\eta(C) > 0$ such that for an infinite sequence of $x \in \mathbb{N}$ we have

$$\pi(x + (\log x)^C) - \pi(x) > (1 + \eta(C))(\log x)^{C-1},$$

and an analogous statement holds with the inequality reversed and $1 - \eta(C)$ in place of $1 + \eta(C)$. This is however not in contradiction with the Cramér model, as that model only predicts how the primes should behave on typical intervals, instead of

[2]The above heuristic in fact suggests that $C = 1$ in Conjecture 3.1. There is however some reason to doubt this choice of C, since a more refined version of the Cramér heuristic due to Granville [23], which takes into account the local distribution of primes, predicts that $C \ge 2e^{-\gamma} = 1.12\ldots$. It is nevertheless generally believed that there is a constant C such that Conjecture 3.1 is true.

all intervals (at least if λ is fixed in (3.1)). In particular, the asymptotic (3.3) is believed to be true on almost all intervals, and in [I] we prove some analogues of (3.2) for the counting function of almost primes, with $\psi(x)$ a very slowly growing function, such as $\psi(x) = (\log x)^{\varepsilon}$.

3.2. **Primes in all short intervals**

The problem of detecting primes in short intervals has attracted wide interest in analytic number theory over several decades; see for instance [38], [116], [48, Chapter 12] for treatises on this topic. There are still many open conjectures in this topic, including the Cramér conjecture (Conjecture 3.1) mentioned above. A much more approachable problem than Conjecture 3.1 is that of finding a real number $\theta \in (0, 1)$ as small as possible such that every interval $[x, x+x^{\theta}]$ with x large enough contains a prime number. It is expected that any $\theta > 0$ is admissible, as would follow from Conjecture 3.1. The first result in this direction is Hoheisel's result [46] from 1930, with $\theta = 1 - \delta$ for some small $\delta > 0$ (he had $\delta = \frac{1}{33000}$). The exponent θ was improved several times during the following decades by various authors, by using results on the theory of the Riemann zeta function and in particular zero density estimates for it. In 1972, Huxley [47] proved that any $\theta > \frac{7}{12}$ is admissible, and this was slightly improved to $\theta = \frac{7}{12}$ by Heath-Brown [40]. All of the above mentioned results in fact provide an asymptotic of the form

$$\pi(x + x^{\theta}) - \pi(x) = (1 + o(1))\frac{x^{\theta}}{\log x}, \tag{3.4}$$

where $\pi(x)$ is the number of primes up to x; furthermore, when it comes to asymptotics of the type (3.4), the result $\theta = \frac{7}{12}$ is still the best one known.

Subsequent authors have considered the problem of obtaining lower bounds of the correct order of magnitude for the number of primes in an interval, meaning estimates of the form

$$\pi(x + x^{\theta}) - \pi(x) \gg \frac{x^{\theta}}{\log x}. \tag{3.5}$$

To achieve such bounds, one can utilize sieve methods in addition to zero density estimates for the Riemann zeta function to obtain stronger results than for the problem (3.4). Such improvements were achieved for instance in [51], [42], [86], [53], [2], and the best result to date is that of Baker, Harman and Pintz [3], who reached $\theta = 0.525$.

The exponent $\theta = \frac{1}{2}$ certainly appears to be the limit of all known methods; as we mentioned, even under the Riemann hypothesis it is only known that (3.4) is true for all $\theta > \frac{1}{2}$. For the same conclusion $\theta > \frac{1}{2}$, it would suffice to assume the density hypothesis[3], which is implied by the Riemann hypothesis. In conclusion, conjectures related to the Riemann zeta function do not seem to enable getting information on primes in all intervals of polylogarithmic length (unlike Conjecture 3.1).

3.3. **Primes in almost all short intervals**

Since known results on the problem of primes in all short intervals remain far from what is being conjectured, it is worthwhile to consider the problem of primes in *almost all* short intervals. Naturally, we say that almost all intervals $[x, x + y(x)]$ contain a prime if

$$|\{n \leq X :\ [n, n + y(n)] \cap \mathbb{P} = \emptyset\}| = o(X). \tag{3.6}$$

Several authors have obtained much better results for (3.6) than for the problem of finding primes in all short intervals. Regarding asymptotics for primes in almost all short intervals, the best result is Huxley's [47] with $y(x) = x^{\theta}$ and $\theta > \frac{1}{6}$ in (3.6). When one gives up asymptotics, one can again obtain much better results, as was done in [36], [112], [54], and most recently by Jia in [55] with $\theta > \frac{1}{20}$.

The natural barrier for the known methods is $\theta > 0$; in particular, one is still far from reaching unconditionally intervals that are as short as in Conjecture 3.1. In analogy with the case of all short intervals, $y(x) = x^{\theta}$ for all $\theta > 0$ in (3.6) would follow from the density hypothesis (and thus also from the Riemann hypothesis), but has not been attained without resorting to such conjectures.

Nevertheless, if one assumes the full strength of the Riemann hypothesis, then Selberg's result (Theorem 3.2) nearly establishes Conjecture 3.1 in almost all cases. Since Gallagher [20] proved that the Poisson distribution property (3.1) of the primes holds under a uniform version of the Hardy–Littlewood prime tuples conjecture, by the discussion of Subsection 3.1 even the optimally short intervals $[x, x+\psi(x)\log x]$ contain a prime almost always under the uniform Hardy–Littlewood conjecture. Heath-Brown [39] showed that the same result can be obtained by assuming the Riemann hypothesis and a suitable uniform version of the pair correlation conjecture for the zeroes of the Riemann zeta function. Needless to say, proving any of these hypotheses seems to be out of reach for all known methods.

If we seek unconditional results in almost all short intervals that are of similar length as in Theorem 3.2, we must relax the notion of primes somewhat. This leads to the study of almost primes in short intervals.

---

[3] This hypothesis states that if $\sigma \in [\frac{1}{2}, 1]$ and $N(\sigma, T)$ is the number of zeros of the Riemann zeta function in the rectangle $[\sigma, 1] \times [-T, T]$ of the complex plane, then $N(\sigma, T) \ll T^{2(1-\sigma)+\varepsilon}$ for any fixed $\varepsilon > 0$.

3.4. **Almost primes in almost all short intervals**
We define two classes of almost primes, the E_k *numbers*

$$E_k = \{n \in \mathbb{N} : \ \Omega(n) = k\}$$

and the P_k *numbers*

$$P_k = \{n \in \mathbb{N} : \ \Omega(n) \leq k\},$$

where $\Omega(n)$ is the number of prime factors of n, counted with multiplicities. Then we trivially have the relations $\mathbb{P} \subset P_k$, $E_k \subset P_k$, and $\mathbb{P} = E_1 = P_1 \setminus \{1\}$. Many of the questions of interest for the set $\mathbb{P}$ of primes have also been investigated for these sets of almost primes, often with significantly better unconditional results. For works that study analogues of classical questions on the primes for the E_k numbers see [21], [35], and for P_k numbers see [5], [81].

There are several reasons why the sets E_k and P_k can be viewed as good approximations[4] to the set $\mathbb{P}$. An obvious reason is of course that the E_k and P_k numbers have only a bounded number of prime factors. In addition, if we denote by $\pi_k(x)$ and $\pi_k^*(x)$ the counting functions of E_k and P_k numbers up to x, respectively, then it is a classical result of Landau (see [102, Section II.6.1]) that we have

$$\pi_k(x) = (1+o(1))\pi_k^*(x) = (1+o(1))\frac{x}{\log x} \cdot \frac{(\log\log x)^{k-1}}{(k-1)!} \tag{3.7}$$

for fixed k, so the sets E_k and P_k have nearly the same density $1/(\log x)$ on $[1, x]$ as $\mathbb{P}$ has. In article [I] and in many earlier works, one actually considers numbers $p_1 \cdots p_k \leq x$ with the constraint $P_i \leq p_i \leq P_i^c$ for $i \leq k-1$ for some suitably chosen $P_i \leq x$ and $c > 1$, and it is not difficult to show that such numbers have cardinality $\asymp_c \frac{x}{\log x}$ up to x, just like the primes.

Another reason for the abundance of results on P_k numbers in analytic number theory is that they are exactly the kind of numbers detected by sieve methods. Indeed, sieve methods typically produce numbers $n \leq x$ with no prime factors $p \leq x^c$ for some $c < \frac{1}{2}$, which then means that $n \in P_{\lceil 1/c \rceil - 1}$. Here we see however an important contrast between the E_k and P_k numbers, namely that the E_k numbers (just like the primes) cannot be produced using only classical combinatorial sieves. Indeed, the notorious parity problem in sieve theory, first discovered by Selberg (and discussed for example in [19, Chapter 16]), states that classical combinatorial sieves cannot distinguish numbers with an odd and even number of prime factors from each other. As E_k numbers have exactly k prime factors, they cannot be distinguished from E_{k+1} numbers in such a manner (and so in particular, primes and P_2 numbers cannot be distinguished). Due to this, many results are significantly weaker for E_k numbers than for P_k numbers, and the E_k numbers are a much closer

[4]Of course, the sets $\mathbb{P}$ and E_k are disjoint for $k > 1$, but so are for instance $\mathbb{P}$ and $\{2p : p \in \mathbb{P}\}$, yet they have essentially the same distributional properties.

approximation to the primes, since the sets $\mathbb{P}$ and E_k are both subject to the parity problem. In problems involving E_k numbers in short intervals one therefore wants to make use of the theory of Dirichlet polynomials (and in particular, the theory of the Riemann zeta function), and the advantage compared to primes is that E_k numbers offer more variables to work with in these Dirichlet polynomial bounds, a substantial benefit in proving various estimates.

Concerning P_k numbers, there are very satisfactory short interval results. Notably, Friedlander and Iwaniec [19, Chapters 6 and 11] proved that, for any function $\psi(x)$ tending to infinity with x, almost all intervals $[x, x+\psi(x)\log x]$ contain a P_4 number. They also hinted how to obtain the same result for P_3 numbers. Mikawa proved in turn that almost all intervals $[x, x+(\log x)^{5+\varepsilon}]$ contain a P_2 number. As both proofs are based on classical sieve methods, they are not applicable to the corresponding question for E_k numbers.

It is nevertheless the case that considerably stronger short interval results have been obtained for the E_k numbers than for the primes. Motohashi [83] proved that, for any $\varepsilon > 0$, almost all intervals $[x, x + x^{\varepsilon}]$ contain an E_2 number[5]. Soon after that, Wolke [114] improved this to almost all intervals $[x, x+(\log x)^c]$ for some large constant c (he had $c = 5\cdot 10^6$). This was the first result for E_2 numbers that involved intervals of merely polylogarithmic length, as in Conjecture 3.1 and Theorem 3.2. Harman [35] then gave a reasonable value of c, namely $c = 7 + \varepsilon$ for any $\varepsilon > 0$. In Article [I], we improve the exponent $7 + \varepsilon$ for E_2 numbers to 3.51 and obtain a nearly optimal result for E_3 numbers.

3.3. **Theorem** (Article [I]). *(a) Almost all intervals* $[x, x + (\log x)^{1+\varepsilon}]$ *contain an* E_3 *number, for any fixed* $\varepsilon > 0$.
(b) Almost all intervals $[x, x + (\log x)^{3.51}]$ *contain an* E_2 *number.*

The result for E_3 numbers is close to optimal, since by (3.7) there exists a positive proportion of intervals $[x, x+(\log x)(\log\log x)^{-2}]$ with no E_3 numbers in them. When it comes to E_2 numbers, significantly improving the exponent 3.51 appears difficult, since even under the density hypothesis the method used in [I] would only improve the exponent to $3 + \varepsilon$ (for this, see [I, Remark 10]).

As a matter of fact, we prove the following quantitative version of Theorem 3.3, where the prime factors of the E_3 and E_2 numbers that we detect are of specific sizes.

3.4. **Theorem** (Article [I]). *Let* $\varepsilon > 0$ *be small but fixed. Let* $X \geq 1$ *be large enough. Define the parameters* $P_1 = (\log\log X)^{6+10\sqrt{\varepsilon}}$, $P_2 = (\log X)^{\varepsilon^{-2}}$ *and* $P_1' = (\log X)^{2.51}$.

[5] In the works [83], [114], [35], the numbers under consideration are E_2 numbers, although the wording "P_2 numbers" is used there for lack of better terminology.

Then, for $P_1 \log X \leq h \leq X$ we have

$$\frac{1}{X}\int_X^{2X}\Bigg|\frac{1}{h}\sum_{\substack{x\leq p_1p_2p_3\leq x+h\\ p_1\in[P_1,P_1^{1+\varepsilon}]\\ p_2\in[P_2,P_2^{1+\varepsilon}]}}1-\frac{1}{X}\sum_{\substack{X\leq p_1p_2p_3\leq 2X\\ p_1\in[P_1,P_1^{1+\varepsilon}]\\ p_2\in[P_2,P_2^{1+\varepsilon}]}}1\Bigg|^2 dx = o\left(\frac{1}{(\log X)^2}\right), \tag{3.8}$$

and for $P_1' \log X \leq h' \leq X$

$$\frac{1}{h'}\sum_{\substack{x\leq p_1p_2\leq x+h'\\ p_1\in[P_1',(P_1')^{1+\varepsilon}]}}1 \geq \frac{\delta}{X}\sum_{\substack{X\leq p_1p_2\leq 2X\\ p_1\in[P_1',(P_1')^{1+\varepsilon}]}}1 \tag{3.9}$$

for some small absolute constant $\delta > 0$ and almost all $x \in [X, 2X]$.

By the prime number theorem and a simple application of Chebyshev's inequality, one can show that Theorem 3.4 indeed implies Theorem 3.3. We remark that in [I] we also find E_k numbers on intervals whose lengths approach $\log x$ as k grows. More precisely, almost all intervals $[x, x+(\log x)(\log_{k-1} x)^{C_k}]$ contain an E_k number for some constant $C_k > 0$. Subsequently, Goudout [22] considered E_k numbers in almost all short intervals $[x, x + h_k(x)]$ uniformly in the k aspect. He gave optimal results for $k \asymp \log\log x$ and nearly optimal results for $5 \leq k \leq \log\log x$.

3.5. **Proof ideas for products of three primes**

As in many previous works on primes and almost primes in short intervals, we reduce proving (3.8), and hence Theorem 3.3(a), to the study of Dirichlet polynomials.

More precisely, we use Perron's formula and a Parseval-type inequality (which utilizes the mean square present in (3.8)) to essentially reduce (3.8) to the corresponding bound for Dirichlet polynomials:

$$\int_{X^{0.01}}^{X/h}|F(1+it)|^2\,dt = o\left(\frac{1}{(\log X)^2}\right), \quad \text{where } F(s) := \sum_{\substack{X\leq p_1p_2p_3\leq 2X\\ p_1\in[P_1,P_1^{1+\varepsilon}]\\ p_2\in[P_2,P_2^{1+\varepsilon}]}}(p_1p_2p_3)^{-s}; \tag{3.10}$$

strictly speaking, we also need to consider the integral over other intervals than $[X^{0.01}, X/h]$, but this turns out to be the most difficult regime. See [I, Lemma 1, formula (4)] for a more precise version of (3.10). Reducing a problem about short intervals to Dirichlet polynomials is advantageous, because the sum $F(s)$ now runs over a long interval and we can make use of various pointwise, mean value and large values estimates for Dirichlet polynomials to estimate the mean square of $F(s)$.

To effectively estimate these Dirichlet polynomials, we incorporate the method that Matomäki and Radziwiłł [74] developed in 2015 for analyzing multiplicative functions in very short intervals to the setting of almost primes in almost all short intervals. Matomäki and Radziwiłł proved, as a special case of their breakthrough on multiplicative functions, that the Möbius function $\mu(n)$ has mean $o(1)$ on almost all intervals $[x, x+\psi(x)]$, for any $\psi(x)$ tending to infinity with x. This result suggests that their method in [74] might imply something about primes or almost primes in almost all short intervals, as well. However, for the case of primes the method is not amenable, as a vital element of the proof is a combinatorial factorization identity available for multiplicative functions (the Ramaré identity, [74, Formula (9)]). For the primes there certainly is no analogous identity[6]. The indicator function of those E_k numbers that we will consider, on the other hand, does have a useful factorization, owing to the constraints for their prime factors in Theorem 3.4. Using this, (3.10) roughly speaking takes the factorized form

(3.11)
$$\int_{X^{0.01}}^{X/h} |P_1(1+it)P_2(1+it)P_3(1+it)|^2\,dt = o\left(\frac{1}{(\log X)^2}\right), \quad P_j(s) := \sum_{P_j \le p \le P_j^{1+\varepsilon}} p^{-s}$$

and $P_3 := X/P_1P_2$ and the sum $P_3(s)$ is over a dyadic interval. Above we have separated the contribution of each of the variables p_i and can estimate the polynomials corresponding to different variables in different ways.

An estimate of the shape (3.11) is our goal in the proof of Theorem 3.3(a), but a number of aspects of the Matomäki–Radziwiłł method require modifications when working with E_k numbers; in particular, one needs to obtain logarithmic savings in places where $o(1)$ savings would suffice for multiplicative functions (for instance, in [I, Lemma 4]). This is due to the fact that the E_k numbers are a sparse set, of density roughly $1/(\log x)$ up to x. Additionally, there is a part of the proof ([I, Proposition 3]), where we need a product of three Dirichlet polynomials of "significant length", in order to apply a $L^2 - L^\infty$ bound to the mean square of their product (if we had only one or two polynomials, we could not afford to apply a pointwise bound to one of them; this is reminiscent of the differences in difficulty between binary and ternary problems in applications of the circle method; see [107, Chapter 3]). We do obtain three Dirichlet polynomials when dealing with E_3 numbers, but two of them are of minuscule length[7] (reflecting the fact that we want to minimize the length of the intervals on which we detect E_3 numbers). We go around this issue by applying Heath-Brown's identity [52, Chapter 13] to decompose one of the "long" Dirichlet polynomials into a product of either two "zeta sums" or three "prime-factored polynomials" (for these concepts, see [I, Section 1.2, Section 2.5]). We then employ a result of Watt [111] (which generalizes the fourth moment bound of the Riemann zeta function) to deal with mean squares of the resulting zeta sums,

[6] It is the case that the indicator function of the primes can be "factorized" into a Dirichlet convolution, by means of Vaughan's or Heath-Brown's identities [52, Chapter 13], but these factorizations are not nearly flexible enough.

[7] The lengths of the Dirichlet polynomials involved will be roughly $(\log\log x)^6$, $(\log x)^{\varepsilon^{-2}}$ and $x(\log x)^{-\varepsilon^{-2}}$, the first two of which are too short for pointwise bounds.

whereas the mean square of the product of three prime-factored polynomials can be dealt with the mentioned $L^2 - L^\infty$ approach. These are the main ingredients we use for the E_3 part of Theorem 3.3.

3.6. **Proof ideas for products of two primes**

For obtaining good results about E_2 numbers, we make use of all the above-mentioned ideas, as well as some additional ones. One could readily apply the strategy we used for E_3 numbers to obtain the exponent $5+\varepsilon$ for E_2 numbers (see [I, Section 4.1]; in the case of the exponent $5+\varepsilon$, we would even get an asymptotic formula on almost all short intervals for the number of E_2 numbers with prime factors in certain ranges, as in the E_3 case). The fact that we have only two variables to work with in the case of E_2 numbers appears to make improving the exponent hard. However, we can apply the principle of Harman's sieve to gain more flexibility. Firstly, we can increase the number of variables by applying the Buchstab identity, a number-theoretic form of the inclusion-exclusion identity. This identity allows us to decompose

$$S_h(x) := \sum_{\substack{x \le p_1 p_2 \le x+h \\ P_1' \le p_1 \le (P_1')^{1+\varepsilon}}} 1,$$

for any choice of $1 \le w < \sqrt{x}$, as

$$\begin{aligned} S_h(x) &= \sum_{\substack{x \le p_1 n \le x+h \\ P_1' \le p_1 \le (P_1')^{1+\varepsilon} \\ (n, \mathcal{P}(w))=1 \\ n>1}} 1 - \sum_{\substack{x \le p_1 q_1 n \le x+h \\ P_1' \le p_1 \le (P_1')^{1+\varepsilon} \\ w \le q_1 < \sqrt{x} \\ (n, \mathcal{P}(q_1))=1 \\ n>1}} 1 \\ &= \sum_{\substack{x \le p_1 n \le x+h \\ P_1' \le p_1 \le (P_1')^{1+\varepsilon} \\ (n, \mathcal{P}(w))=1 \\ n>1}} 1 - \sum_{\substack{x \le p_1 q_1 n \le x+h \\ P_1' \le p_1 \le (P_1')^{1+\varepsilon} \\ w \le q_1 < \sqrt{x} \\ (n, \mathcal{P}(w))=1 \\ n>1}} 1 + \sum_{\substack{x \le p_1 q_1 q_2 n \le x+h \\ P_1' \le p_1 \le (P_1')^{1+\varepsilon} \\ w \le q_2 < q_1 < \sqrt{x} \\ (n, \mathcal{P}(q_2))=1 \\ n>1}} 1 \\ &:= \Sigma_1(h) - \Sigma_2(h) + \Sigma_3(h). \end{aligned}$$

We take here $w = X^{\eta(X)}$ for a suitable function $\eta(X)$ tending to 0. (In particular, w is small enough for the fundamental lemma of sieve theory [19, Chapter 6] to be applicable). As we will see later, the first two sums $\Sigma_1(h), \Sigma_2(h)$ are asymptotically equal to their dyadic counterparts $\frac{h}{X}\Sigma_1(X)$ and $\frac{h}{X}\Sigma_2(X)$, respectively, for almost all $x \in [X, 2X]$. For the sum $\Sigma_3(h)$, however, we are not able to prove an asymptotic unless we impose some additional conditions on the sizes of the variables q_i. Let $\Sigma_3'(h)$ be the part of $\Sigma_3(h)$ that we can evaluate asymptotically (to be asymptotic to the normalized dyadic version of the same sum), and let $\Sigma_3''(h) \ge 0$ be the rest (the part that we can evaluate is expressed precisely in [I, Subsection 6.3] as the sums $\Sigma_3^{(1)}(h)$ and $\Sigma_3^{(2)}(h)$). Then, for almost all $x \in [X, 2X]$, we get

$$\begin{aligned}\frac{1}{h}S_h(x) &= \frac{1}{h}(\Sigma_1(h) - \Sigma_2(h) + \Sigma_3(h)) \\ &= \frac{1}{X}\Sigma_1(X) - \frac{1}{X}\Sigma_2(X) + \frac{1}{h}\Sigma_3'(h) + \frac{1}{h}\Sigma_3''(h) + o\left(\frac{1}{\log X}\right) \\ &\geq \frac{1}{X}(\Sigma_1(X) - \Sigma_2(X) + \Sigma_3'(X)) + o\left(\frac{1}{\log X}\right) \\ &= \frac{1}{X}S_X(X) - \frac{1}{X}\Sigma_3''(X) + o\left(\frac{1}{\log X}\right).\end{aligned}$$

Hence, proving (3.9) has been reduced to establishing the mentioned asymptotics for $\Sigma_1(h), \Sigma_2(h)$ and $\Sigma_3'(h)$, and additionally to showing that $\Sigma_3''(X) \leq (1-2\delta)S_X(X)$ for some fixed $\delta > 0$. Clearly, the part $\Sigma_3'(h)$ of $\Sigma_3(h)$ that we can evaluate must be large enough for this upper bound to hold. It turns out that if we look for E_2 numbers on intervals of length $[x, x + (\log x)^c]$, then we should take $P_1' = (\log x)^{c-1}$ in the definitions of $S_h(x)$ and $\Sigma_i(h)$, and the smaller P_1' is, the harder $\Sigma_3(h)$ is to estimate. We can give by the prime number theorem an asymptotic for $\Sigma_3''(X)$ (since the sum is over a dyadic interval) in terms of multidimensional "Buchstab integrals" [38, Chapter 3], [I, Section 6.3.3], and we compute that if $c = 3.51$ above, the sum $\Sigma_3''(X)$ is indeed smaller than the main term $S_X(X)$. We are thus left with showing asymptotics for $\Sigma_1(h), \Sigma_2(h), \Sigma_3'(h)$

The proofs of the asymptotics of the sums $\Sigma_1(h), \Sigma_2(h)$ and $\Sigma_3'(h)$ follow the same strategy as in the E_3 case, but make use of some additional inputs. We reduce the problem to the setting of Dirichlet polynomials, so that the aim is to prove that (3.10) holds for the Dirichlet polynomials $F(s)$ that correspond to $\Sigma_1(h), \Sigma_2(h)$ and $\Sigma_3'(h)$. By applying a simple sieve to $\Sigma_1(h)$ and $\Sigma_2(h)$, they become *type I sums* (meaning a sum having a long, unrestricted integer variable), and therefore we can employ Watt's mean value theorem as in the E_3 case to handle them.

The sum $\Sigma_3'(h)$, in turn, is a *type II sum* (it has several variables of substantial length, but these variables come with weights), and is more difficult to estimate. However, we have restricted the sizes of the variables in a suitable manner in this sum, making asymptotic evaluation possible. We utilize the ideas from the E_3 case together with the theory of exponent pairs [I, Section 5.1] and better large values theorems for Dirichlet polynomials [I, Lemma 7] to obtain the bound (3.10) for the Dirichlet polynomial corresponding to $\Sigma_3'(h)$, and this then implies that $\Sigma_3'(h)$ has the desired asymptotic.

We have now outlined the main strategy for proving Theorem 3.3; the details can be found in [I].

4. The Goldbach problem for primes of a special form

4.1. The Goldbach conjectures

The Goldbach conjectures, proposed by Goldbach in a letter to Euler in 1742, are some of the most influential and well-known problems in analytic number theory. The ternary Goldbach conjecture asserts that every odd integer $n \geq 7$ is the sum of three primes. The binary Goldbach conjecture in turn claims that every even integer $n \geq 4$ can be written as the sum of two primes; this is still unsolved. The binary Goldbach conjecture is evidently stronger than the ternary one, since if $n = p_1 + p_2$ is a sum of two primes, then $n+3 = p_1+p_2+3$ is a sum of three primes.

The ternary conjecture was settled in all but finitely many cases by Vinogradov [108] in 1937 in a work that redefined the Hardy–Littlewood circle method.

4.1. **Theorem** (Vinogradov)**.** *Every large enough odd integer n can be written as $n = p_1 + p_2 + p_3$ with $p_1, p_2, p_3 \in \mathbb{P}$.*

For a modern proof of Theorem 4.1, see [107, Chapter 3]. It took until 2013 before Theorem 4.1 was extended to all $n \geq 7$; this was achieved by Helfgott [43], by introducing new ideas both on the analytic and numerical sides. Although the binary analogue of Vinogradov's result has resisted all attempts to a full resolution, shortly after Vinogradov's proof it was shown independently by Chudakov, van der Corput and Estermann that we have the following approximation (see [107, Chapter 3]).

4.2. **Theorem** (Almost all cases of binary Goldbach)**.** *Almost all even integers n can be expressed as $n = p_1 + p_2$, where $p_1, p_2 \in \mathbb{P}$.*

Here and in what follows, by "almost all" we mean that the number of exceptional even integers $n \leq N$ is $o(N)$.

Given that one has such an approximation to the binary Goldbach conjecture, one may contemplate a number of refinements, such as strengthening the bound for the number of exceptions

$$E(X) := |\{n \leq X : \ n \equiv 0 \pmod 2, \ n \text{ not a sum of two primes}\}|.$$

This question was considered most notably by Montgomery and Vaughan [82], who obtained $E(X) \ll X^{1-\delta}$ for some fixed $\delta > 0$. The bound was improved by Chen and Pan [6], Li [64] and Lu [66], among others, the last of whom holds the record $\delta = 0.121$. In a somewhat different direction, one can try to minimize $\theta > 0$ such that every interval $[X, X + X^{\theta}]$ with X large contains a sum of two primes. This was investigated in [84], [85], [63], among others, and in Jia's work [56], where the best known result $\theta = \frac{7}{108} + \varepsilon$ was obtained.

In Article [II], we had a different generalization of Theorems 4.1 and 4.2 in mind, namely a version of the problem where only a specific subset of the primes are allowed as summands.

4.2. **The Goldbach problem for subsets of the primes**
When looking for a problem which is more challenging than the ternary Goldbach problem, but (hopefully) more manageable than the binary Goldbach problem, the following problem naturally arises.

4.3. **Problem** (Ternary Goldbach for subsets of the primes)**.** *Let $\mathcal{P} \subset \mathbb{P}$ be a given interesting subset of the primes. Is it the case that all large enough odd n can be represented as $n = p_1 + p_2 + p_3$ with $p_1, p_2, p_3 \in \mathcal{P}$?*

Whether Problem 4.3 has a positive or negative answer crucially depends on the distribution of the set $\mathcal{P}$ in arithmetic progressions and more general *Bohr sets*. These are sets of the form

$$\bigcup_{i \leq m} \{n \in \mathbb{N} : \|\alpha_i n\| \leq \eta_i\} \tag{4.1}$$

with $\alpha_i \in \mathbb{R}$, $\eta_i \in (0,1)$. If we take $\alpha_i \in \mathbb{Q}$, we see that arithmetic progressions are a special case of Bohr sets. Now, if for example $\mathcal{P} = \{p \in \mathbb{P} : \ p \equiv 1 \pmod 3\}$, then only integers of the form $n \equiv 0 \pmod 3$ can be represented as a sum of three primes from $\mathcal{P}$. Similarly, if $\mathcal{P} = \{p \in \mathbb{P} : \ \|\sqrt{2}p\| < \frac{1}{10}\}$, then every n representable as a sum of three primes from $\mathcal{P}$ satisfies $\|\sqrt{2}n\| < \frac{3}{10}$, a property that fails for a positive proportion of odd n. In light of these examples where the answer to Problem 4.3 is negative, we would like the set $\mathcal{P}$ studied in Problem 4.3 to contain a fair proportion of elements from each Bohr set of the form (4.1).

It is only in recent years that interesting subcases of Problem 4.3 have been solved. In 2014, Shao [92] showed that Problem 4.3 has an affirmative answer for any subset $\mathcal{P} \subset \mathbb{P}$ of relative lower density[8] greater than $\frac{5}{8}$. Perhaps surprisingly, this is optimal when taking only the density into consideration: the subset

$$\mathcal{P} := \{p \in \mathbb{P} : \ p \equiv 1, 2, 4, 7, 13 \pmod{15}\}$$

has density $\frac{5}{8}$ and, by simple modular arithmetic, sums of three of its elements are never $\equiv 14 \pmod{15}$. Matomäki and Shao [77] considered Problem 4.3 for significantly sparser but specific subsets of the primes. Their subsets of interest are the *Chen primes* and the *bounded gap primes*[9]. Chen primes are primes p such that $p+2$ has at most two prime factors; the infinitude of such primes was proved by Chen [5] in 1973. The bounded gap primes are primes p such that the interval $[p, p+C]$ contains at least two primes for some large, fixed C, and their infinitude was proved in the celebrated work of Zhang [117] in 2013 and in a more general form by Maynard [80] and Tao (unpublished) in 2014.

---

[8]We define the relative lower density of $B \subset A$ with respect to A as $\liminf_{N\to\infty} \frac{|B \cap [1,N]|}{|A \cap [1,N]|}$.

[9]The latter set does not have a standardized name.

Problem 4.3 should be compared to the problem of locating three-term arithmetic progressions in the set $\mathcal{P}$, which also involves studying a linear equation in the primes.

4.4. **Problem** (Three-term arithmetic progressions in subsets of the primes)**.** *Let $\mathcal{P} \subset \mathbb{P}$ be a given interesting subset of the primes. Is it the case that $\mathcal{P}$ contains infinitely many solutions to $p_1 + p_3 = 2p_2$ with $p_1, p_2, p_3 \in \mathcal{P}$ distinct?*

As with Problem 4.3, a number of special cases of Problem 4.4 have been solved. Concerning this, Green [25] proved Roth's theorem for the primes[10], stating that any subset of the primes of positive relative upper density[11] contains infinitely many non-constant three-term arithmetic progressions. This was famously generalized to k-term arithmetic progressions by Green and Tao [27]. In another work, Green and Tao [26] showed that the Chen primes satisfy Roth's theorem.

The approach that Green and Green–Tao developed for this type of problems is called a *transference principle*, as it allows one to transfer information (such as Roth's theorem) from dense subsets of the integers to sparse ones (such as the primes) under suitable conditions. Intuitively speaking, the principle says that if $A \subset [1, N]$ is a set with $|A| = \delta N$ and $\delta = \delta(N) > 0$, then A contains many three-term arithmetic progressions, provided that the normalized indicator $\delta^{-1}1_A(n)$ has a *pseudorandom majorant* (that is, a majorizing function $\nu(n)$ that has mean $\asymp 1$ and has small Fourier coefficients) and that $\delta^{-1}1_A(n)$ is "Fourier bounded" (that is, its Fourier transform has small L^r norm for $r > 2$). This version of the transference principle is specific to the translation-invariant linear equation $x + z = 2y$, and does as such not apply to the setting of Problem 4.3. Indeed, the set $\{p \in \mathbb{P} : \ \|\sqrt{2}p\| < \frac{1}{10}\}$ is an example of a set that contains an abundance of arithmetic progressions (since it is a positive relative density set of the primes) but, as mentioned earlier, has no solutions to $n = p_1 + p_2 + p_3$ for many odd n. Therefore, to deal with Problem 4.3, one needs a different version of the transference principle, which takes into account the distribution of $\mathcal{P}$ in Bohr sets (an example of which is the fractional part set above); we shall discuss this later in this section.

4.3. Statements of results

In Article [II], we study Problem 4.3 for the specific subset

$$\mathscr{P} := \{p \in \mathbb{P} : \ p = x^2 + y^2 + 1, x, y \in \mathbb{Z}\},$$

consisting of primes representable as values of the polynomial $x^2 + y^2 + 1$. There are a number of reasons why the set $\mathscr{P}$ is interesting. Firstly, it is perhaps the simplest non-trivial example of a sparse subset of the primes consisting of the values of a multivariate polynomial. When it comes to single-variable polynomials,

[10]The theorem is named so, since Roth [88] proved that positive upper density subsets of the integers contain infinitely many non-trivial three-term arithmetic progressions.

[11]The relative upper density of a set $A \subset B$ with respect to $B \subset \mathbb{N}$ is $\limsup\limits_{N \to \infty} \dfrac{|B \cap [1, N]|}{|A \cap [1, N]|}$.

only degree one polynomials have been proved to produce infinitely many primes (by Dirichlet's theorem, under a coprimality condition), so one should turn to multivariate polynomials for interesting unconditional results. Concerning irreducible binary quadratic forms $ax^2+bxy+cy^2$, which are some of the simplest multivariate polynomials, the Chebotarev density theorem can be used to characterize when such a form represents infinitely many primes. When the form does represent infinitely many primes, Chebotarev's theorem implies that the relative density of such primes is $\frac{1}{2}$. Therefore, binary quadratic forms do not produce sparse subsets of the primes.

We mention that there are also some interesting higher degree multivariate polynomials that are known to represent infinitely many primes. Friedlander and Iwaniec [18] showed that the polynomial x^2+y^4 takes infinitely many prime values, Heath-Brown [41] showed that x^3+2y^3 has the same property, and Maynard [79] showed this property for an infinite class of more general polynomials called norm forms. The sets of primes corresponding to these polynomials have cardinalities $\ll X^{1-\delta}$ up to X for some $\delta>0$, so they are certainly sparse subsets of the primes. Since primes represented by these polynomials have not been studied in arithmetic progressions to large moduli, the Goldbach problem appears formidable for them.

The set $\mathscr{P}$ of primes represented by x^2+y^2+1 is also a sparse subset of the primes; an application of Selberg's sieve provides the bound $|\mathscr{P}\cap[1,N]|\ll N(\log N)^{-\frac{3}{2}}$ (to see this, note that if $m\in\mathscr{P}\cap[\frac{N}{2},N]$, then $(m,\prod_{p\le z}p)=(\frac{m-1}{k^2},\prod_{p\le z,p\equiv 3 \pmod 4}p)=1$ for $z=N^{0.01}$ and for some $k\in\mathbb{N}$). It is known that $\mathscr{P}$ is infinite, a result first shown by Linnik [65] in 1960, using his dispersion method. Later, a sieve-theoretic proof of this was given by Iwaniec [50], making use of the linear and semilinear sieves. Iwaniec's proof also established the matching lower bound $|\mathscr{P}\cap[1,N]|\gg N(\log N)^{-\frac{3}{2}}$. Subsequently, various properties of the set $\mathscr{P}$ have been investigated; in particular, it has been studied over short intervals [115], [69], and variants of the Goldbach problem have been studied for this set. In 2008, Matomäki [70] showed that almost all even $n\not\equiv 2 \pmod 6$ can be expressed as $n=p+q$ with $p\in\mathscr{P}$ and $q\in\mathbb{P}$ a generic prime. Next, Tolev [104] gave an asymptotic formula for the representations of such n as $n=p+q$ with $p\in\mathscr{P}$, $q\in\mathbb{P}$, again for almost all n. In another work [105], he considered the corresponding ternary Goldbach problem and showed that every large enough odd n can be written as $n=p+q+r$ with $p,q\in\mathscr{P}$ and $r\in\mathbb{P}$. We strengthen these results by solving Problem 4.3 for the set $\mathscr{P}$.

4.5. **Theorem** (Article [II]). *Every large enough odd n can be represented as $n=p_1+p_2+p_3$ with $p_1,p_2,p_3\in\mathscr{P}$.*

We also improve Matomäki's result [70] by settling almost all cases of the binary Goldbach problem for $\mathscr{P}$.

4.6. **Theorem** (Article [II]). *Almost all even integers $n\not\equiv 5,8 \pmod 9$ can be represented as $n=p_1+p_2$ with $p_1,p_2\in\mathscr{P}$.*

Note that the condition $n \not\equiv 5, 8 \pmod 9$ is necessary, since in the complementary case the fact that $p \equiv 1, 2, 5$ or $8 \pmod 9$ for primes $p = x^2 + y^2 + 1 \neq 3$ shows that p_1 or p_2 equals 3 in the representation $n = p_1 + p_2$, in which case we could only represent $\ll N(\log N)^{-1}$ integers up to N.

We also investigate Problem 4.4 in Article [II], again for the specific set $\mathscr{P}$. We are able to resolve this problem and, more generally, to prove Roth's theorem for $\mathscr{P}$.

4.7. **Theorem** (Article [II]). *The set $\mathscr{P}$ contains infinitely many non-trivial three-term arithmetic progressions. More generally, any subset of*

$$\mathscr{P}^* := \{p \in \mathbb{P} : \ p = x^2 + y^2 + 1, \ x, y \text{ coprime}\}$$

having positive relative upper density contains infinitely many non-trivial three-term arithmetic progressions.

We remark that subsequently Sun and Pan [95] generalized this result by proving that the set $\mathscr{P}$ contains arbitrarily long arithmetic progressions.

One more topic considered in Article [II] is the distribution of irrational multiples of primes belonging to the subset $\mathscr{P}$. For the whole set of primes, such results take the form:

(4.2)

$$\text{For all } \alpha \in \mathbb{R} \setminus \mathbb{Q}, \ \beta \in \mathbb{R} \text{ we have } \|\alpha p + \beta\| < p^{-\theta} \text{ for infinitely many } p \in \mathbb{P},$$

where θ is a constant whose value we are attempting to maximize. The first result in this direction was that of Vinogradov [109] with $\theta = \frac{1}{5} - \varepsilon$. This was improved several times, notably by Vaughan [106] to $\theta = \frac{1}{4} - \varepsilon$, by Harman [37] to $\theta = \frac{3}{10}$, and by Jia [57] to $\theta = \frac{9}{28}$. In the special case $\beta = 0$, the record is Matomäki's result [72] with $\theta = \frac{1}{3} - \varepsilon$.

The problem (4.2) has also been studied for Chen primes in [71], [93], Piatetski-Shapiro primes in [32], and Gaussian primes in [1]. Here we obtain the first result concerning (4.2) for the subset $\mathscr{P}$ of the primes.

4.8. **Theorem** (Article [II]). *Let $\varepsilon > 0$. For any $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ and $\beta \in \mathbb{R}$, we have*

$$\|\alpha p + \beta\| < p^{-\frac{1}{80}+\varepsilon}$$

for infinitely many $p \in \mathscr{P}$.

This establishes that the elements of $\mathscr{P}$ are somewhat uniformly distributed in Bohr sets. We remark that the exponent $\frac{1}{80}$ in Theorem 4.8 could be improved by a more careful analysis in [II, Sections 8-9]; we however confined ourselves to showing that one can get some positive, explicit exponent.

4.4. **Method of proof**
Our proofs of Theorems 4.5 and 4.6 do not apply the classical circle method, but rather a transference principle for ternary equations. Let us first describe why the traditional circle method approach is not applicable to the set $\mathscr{P}$.

The starting point of Vinogradov's proof of Theorem 4.1, and many subsequent developments of the circle method, is the reduction of the problem to analyzing exponential sums via the identity

$$|\{(p_1,p_2,p_3) \in \mathbb{P}^3 :\ N = p_1+p_2+p_3\}| = \int_0^1 S(\alpha)^3 e(-N\alpha)\, d\alpha,\ S(\alpha) := \sum_{p\leq N} e(\alpha p).$$

This identity is seen to hold by expanding out $S(\alpha)^3$ and applying the orthogonality identity

$$1_{n=0} = \int_0^1 e(n\alpha)\, d\alpha.$$

One then examines separately the *major arc* case $\alpha \in \mathfrak{M}$, where α is close to a rational number with small denominator, and the opposite *minor arc* case $\alpha \in \mathfrak{m} := [0,1] \setminus \mathfrak{M}$. For $\alpha \in \mathfrak{M}$, the sum $S(\alpha)$ is often "large", and one can evaluate it asymptotically by the Siegel–Walfisz theorem. For $\alpha \in \mathfrak{m}$, in turn, one expects $S(\alpha)$ to be "small", and this can be proved with the help of Vaughan's identity, which transforms sums over primes to bilinear sums. We refer to [107, Chapter 3] for details of the method.

If we applied the same strategy to Theorem 4.5, we would run into trouble, since we do not have a good understanding of

$$S_{\mathscr{P}}(\alpha) := \sum_{\substack{p\leq N \\ p\in\mathscr{P}}} e(\alpha p),$$

neither in the major arc nor in the minor arc case. In the major arcs, the problem is that we only have upper and lower bounds for $|\mathscr{P} \cap [1,N]|$ that are off by a constant factor, and hence we have no asymptotic even for $S_{\mathscr{P}}(0)$. In the minor arc case, the difficulty is that no analogue of Vaughan's identity is known for the indicator function $1_{\mathscr{P}}(n)$. For these reasons, the classical circle method is not the right line of attack for Theorem 4.5. We mention though that if one studies the ternary Goldbach problem with *two* of the three prime variables coming from the subset $\mathscr{P}$, then the circle method coupled with sieve methods is applicable; see [70], [105].

The proofs of Theorems 4.5 and 4.6 are instead based on a transference type principle of Matomäki and Shao [77, Theorem 2.3]. Roughly speaking, the principle says that if N is large and a set $A \subset [1,N]$ with $|A| = \delta N$ obeys, for some fixed $\delta_0 > 0$ (and small enough $\eta = \eta(\delta_0)$) the properties

(i) (well-distribution in Bohr sets) $|A \cap (B-t)| \geq \delta_0 |A||B|/N$ for $t \in [N/4, N/2]$ and all Bohr sets[12] B with $|B| \geq \eta N$;
(ii) (Fourier boundedness) $\sum_{\xi \in \mathbb{Z}_N} |\delta^{-1}\widehat{1_A}(\xi)|^{\frac{5}{2}} \leq \delta_0^{-1}$;
(iii) (Non-sparseness) $|A \cap [0.1N, 0.4N]| \geq \delta_0 \cdot \delta N$,

then there exist $a_1, a_2, a_3 \in A$ such that $N = a_1 + a_2 + a_3$. The actual formulation of the transference-type principle is somewhat more involved; we refer to [77, Theorem 2.3] for the details. The principle can also be generalized to work for almost all cases of binary problems; see [II, Proposition 2.1] for this. We also note that Matomäki, Maynard and Shao [73] developed a different transference-type principle for Goldbach-type problems; this version allowed them to improve the exponent for the Goldbach problem with almost equal variables [73, Theorem 1.1].

We will apply the transference principle essentially to

$$A := \{n \leq N : \; Wn + b \in \mathscr{P}\} \tag{4.3}$$

with

$$\delta \asymp \left(\frac{W}{\varphi(W)}\right)^{\frac{3}{2}} (\log N)^{-\frac{3}{2}},$$

where $(b, W) = 1$ and $W = \prod_{p \leq w} p$ for some large, fixed w. This "W-trick" of restricting to primes in a residue class is necessary to guarantee the well-distribution of A in arithmetic progressions (which are a special case of Bohr sets).

Intuitively, condition (i) of the transference principle guarantees that A contains a fair proportion of each Bohr set (which, as we indicated in Subsection 4.2, is necessary); condition (ii) is related to the existence of a pseudorandom majorant[13]; and condition (iii) says that A is not too concentrated on certain subintervals.

The main task in the proofs of Theorems 4.5 and 4.6 is then verifying the conditions (i)–(iii) of the transference principle for the specific set A given by (4.3). Condition (iii) is the simplest to check and follows with minor modifications from Iwaniec's proof of the infinitude of $\mathscr{P}$. Condition (ii), the Fourier boundedness condition, is closely related to the restriction theory of the primes, a topic studied by Green [25] and Green–Tao [26]. To obtain (ii), we roughly speaking want to construct a function $\beta : \mathbb{N} \to \mathbb{R}_{\geq 0}$ that enjoys the majorization property $\delta^{-1}1_A(n) \leq C\beta(n)$, has mean value $\asymp 1$ (so that β is essentially a probability measure on $[1, N]$), and has a Fourier expansion that is of "low enough complexity". In other words, β is a pseudorandom majorant for 1_A in a suitable sense. Then, under these conditions,

[12] We defined Bohr sets in formula (4.1).
[13] In [II, Section 4], we show that the existence of a suitable pseudorandom majorant implies (ii), and then we construct such a majorant.

[26, Proposition 4.2] implies a *restriction estimate*

$$\left(\sum_{\xi\in\mathbb{Z}_N}\left|\frac{1}{N}\sum_{n\le N}a_n\beta(n)e\left(-\frac{\xi n}{N}\right)\right|^r\right)^{1/r}\le C_r\left(\frac{1}{N}\sum_{n\le N}|a_n|^2\beta(n)\right)^{1/2}$$

for all complex numbers a_n and fixed $r>2$. Taking

$$a_n=\begin{cases}\frac{\delta^{-1}1_A(n)}{\beta(n)}, & \beta(n)\ne 0\\ 0, \quad \beta(n)=0,\end{cases}$$

(so that $|a_n|\ll 1$) we deduce, in particular, that

$$\sum_{\xi\in\mathbb{Z}_N}|\delta^{-1}\widehat{1_A}(\xi)|^{\frac{5}{2}}\ll\left(\frac{1}{N}\sum_{n\le N}\beta(n)\right)^{5/4}\ll 1.$$

Naturally, we still need to prove that such a pseudorandom majorant β exists. It turns out that the Selberg upper bound sieve does the job, and to see this we closely follow a paper of Ramaré and Ruzsa [87].

The majority of the proofs of Theorems 4.5 and 4.6 is then devoted to proving condition (i), well-distribution in Bohr sets. In other words, we wish to analyze sums of the form

$$\sum_{\substack{n\le N\\ n\in\mathscr{P}}}1_B(n),$$

where B is a Bohr set (or a smoothed version thereof). By applying a weighted form of the linear and semilinear sieves (as developed in [II, Section 6], following Iwaniec's work in [49]), we reduce the problem to showing that the count of primes in Bohr sets has a good enough level of distribution. More precisely, we want to find levels of distribution $\rho_1,\rho_2\in(0,1)$ as large as possible, such that the following holds. For a set $\mathcal{L}\subset\mathbb{N}$ of "bilinear type" (in the sense that it consists of integers having a certain type of factorization), we have

$$\sum_{d\le N^{\rho_1}}\lambda_d^{+,\mathrm{LIN}}\sum_{\substack{\ell\le N^{0.9}\\ \ell\in\mathcal{L}}}\Bigg(\sum_{\substack{N\le n\le 2N\\ n=\ell p+1\\ n\equiv 0 \pmod d}}1_B(n)-\frac{1}{\varphi(d)}\sum_{N\le n\le 2N}\frac{1_B(n)}{\ell\log\frac{n}{\ell}}\Bigg)\ll\frac{N}{(\log N)^{100}}, \tag{4.4}$$

and

$$\sum_{d\le N^{\rho_2}}\lambda_d^{-,\mathrm{SEM}}\Bigg(\sum_{\substack{N\le p\le 2N\\ p\equiv 1 \pmod d}}1_B(p)-\frac{1}{\varphi(d)}\sum_{N\le p\le 2N}1_B(p)\Bigg)\ll\frac{N}{(\log N)^{100}}, \tag{4.5}$$

where B is a Bohr set (or a smoothed version of it), $\lambda_d^{+,\mathrm{LIN}}$ are the upper bound linear sieve weights with level $D_1=N^{\rho_1}$ and sifting parameter $z_1=N^{1/5}$ and $\lambda_d^{-,\mathrm{SEM}}$

are the lower bound semilinear sieve weights with level $D_2 = N^{\rho_2}$ and sifting parameter $z_2 = N^{1/3-\varepsilon}$ (see Section 1 for a precise definition of sieve weights and [II, Hypothesis 6.4] for the exact, slightly more complicated statements of interest).

By expanding $1_B(p)$ as a finite Fourier series (and a small error), we then need to bound the Bombieri–Vinogradov type averages (4.4) and (4.5) with an additive character $e(\alpha p)$ in place of $1_B(p)$. In the case of the linear sieve weights (which have the well-factorability property defined in [19, Chapter 12]), we manage to obtain the good value $\rho_1 = \frac{1}{2} - \varepsilon$ for the level of distribution by following [71]. The semilinear sieve weights, however, are not well-factorable, and the level of distribution $\rho_2 = \frac{1}{3} - \varepsilon$ obtained for general weights in [103, Lemma 1] is not good enough for our purposes. We therefore prove a combinatorial factorization of semilinear sieve weights [II, Lemma 9.2] by following the principle of Harman's sieve [38, Chapter 3], and this enables us to show that the weights $\lambda_d^{-,\mathrm{SEM}}$ have "enough flexibility" in their factorizations as Dirichlet convolutions. This amount of flexibility allows us to achieve the better level of distribution $\rho_2 = \frac{3}{7} - \varepsilon$, which is good enough for our needs. The details of the proof can be found in [II].

5. On the logarithmic Chowla conjecture

5.1. **Chowla's conjecture**

The Liouville function, a fundamentally important function in multiplicative number theory, is defined as $\lambda(n) := (-1)^{\Omega(n)}$, where $\Omega(n)$ is the number of prime factors of the integer n counted with multiplicities. This function is closely related to the more well-known Möbius function, given by $\mu(n) := (-1)^{\Omega(n)} \cdot 1_{n \text{ squarefree}}$; in particular, they are both multiplicative functions having value -1 at the primes. The distribution of the Liouville function (or equally well of the Möbius function) appears highly random (like a series of coin flips) and in particular, consecutive values of the Liouville function should be asymptotically independent of each other. This was formalized by Chowla [7] in 1965 as the following assertion.

5.1. **Conjecture** (Chowla's conjecture)**.** For any $k \geq 1$ and any distinct shifts $h_1, \dots, h_k \in \mathbb{N}$, we have

$$\frac{1}{x}\sum_{n \leq x} \lambda(n+h_1)\cdots\lambda(n+h_k) = o(1). \tag{5.1}$$

The conjecture can be interpreted as stating that shifted products of the Liouville function have mean 0. Alternatively, the conjecture can be stated in the following equivalent form from which it is clearer that it is a statement about the independence of simultaneous values of the Liouville function.

5.2. **Conjecture** (Chowla's conjecture, sign pattern formulation)**.** For any $k \geq 1$, any signs $\varepsilon_1, \dots, \varepsilon_k \in \{-1, +1\}$, and any distinct shifts $h_1, \dots, h_k \in \mathbb{N}$, we have

$$\lim_{x\to\infty} \frac{1}{x}|\{n \leq x : \ \lambda(n+h_1) = \varepsilon_1, \dots, \lambda(n+h_k) = \varepsilon_k\}| = 2^{-k}.$$

To see that Conjectures 5.1 and 5.2 are indeed equivalent, one can simply substitute $\lambda(n+h_i) = \varepsilon_i(2 \cdot 1_{\lambda(n+h_i)=\varepsilon_i} - 1)$ into (5.1) and expand the product.

We remark that Conjecture 5.1 could be generalized to the assertion that

$$\frac{1}{x}\sum_{n \leq x} \lambda(a_1 n+h_1)\cdots\lambda(a_k n+h_k) = o(1), \tag{5.2}$$

whenever the non-degeneracy condition $a_i h_j \neq a_j h_i$ for $1 \leq i < j \leq k$ is fulfilled. One could also formulate Conjecture 5.1 with the Möbius function in place of the Liouville function; one can show by elementary sieve theory that such a conjecture would still follow from (5.2). Conjecture 5.2, however, takes a more complicated form for the Möbius function, as for example the events $\mu(n) = 1$, $\mu(n+1) = 1$, $\mu(n+2) = 1$ and $\mu(n+3) = 1$ are not independent (at most three of them can hold simultaneously, since one of the values is 0).

Conjectures 5.1 and 5.2 resemble the famous Hardy-Littlewood prime tuples conjecture [34], and they can be thought of as simpler analogues of it.

5.3. **Conjecture** (Hardy-Littlewood prime tuples conjecture)**.** Let $h_1, \dots, h_k \in \mathbb{N}$ be distinct integers. Then the von Mangoldt function $\Lambda(n)$ has the correlation asymptotics

$$\frac{1}{x}\sum_{n\le x}\Lambda(n+h_1)\cdots\Lambda(n+h_k)=\mathfrak{S}(h_1,\dots,h_k)+o(1),$$

with $\mathfrak{S}(h_1,\dots,h_k)$ an effectively computable constant satisfying $\mathfrak{S}(h_1,\dots,h_k)>0$ if and only if the polynomial $(n+h_1)\cdots(n+h_k)$ has no fixed prime divisor.

A connection between the Chowla and Hardy-Littlewood conjectures is hinted by the identity

$$\Lambda(n)=\sum_{d\mid n}\mu(d)\log\frac{n}{d},$$

which binds together the Möbius and von Mangoldt functions. Nevertheless, one would need a strong error term of the form $O((\log x)^{-A})$ on the right-hand side of (5.2) to be able to have a rigorous implication from Conjecture 5.1 to Conjecture 5.3. None of the current progress on Chowla's conjecture (for $k\ge 2$) has produced such good error terms.

In its original form, Chowla's conjecture is open for all $k\ge 2$. The simplest $k=1$ case

$$\frac{1}{x}\sum_{n\le x}\lambda(n)=o(1)$$

can be shown to be equivalent to the prime number theorem and, more generally,

$$\frac{1}{x}\sum_{n\le x}\lambda(an+h)=o(1)$$

is equivalent to the prime number theorem in arithmetic progressions.

Despite this lack of progress on the original conjecture, starting from 2015 there has been major progress on different variants of Chowla's conjecture. Matomäki and Radziwiłł [74] proved, while showing cancellation in very short averages of multiplicative functions, that, for any $h\in\mathbb{N}$,

$$\limsup_{x\to\infty}\left|\frac{1}{x}\sum_{n\le x}\lambda(n)\lambda(n+h)\right|\le 1-\delta(h)$$

for some $\delta(h)>0$. This was the first nontrivial progress towards the two-point Chowla conjecture (for the odd order cases, the analogous result was proved by Elliott [11]). Soon after this result, Matomäki, Radziwiłł and Tao [75] showed that

Chowla's conjecture (as well as the more general Elliott's conjecture, discussed in Chapter 6) holds on average over the shifts, in the sense that

$$\frac{1}{H(x)^k}\sum_{h_1,\dots,h_k\le H(x)}\left|\frac{1}{x}\sum_{n\le x}\lambda(n+h_1)\cdots\lambda(n+h_k)\right| = o(1) \tag{5.3}$$

for any $H(x)\le x$ tending to infinity with x. If one could take $H(x)$ bounded, one would of course obtain Chowla's conjecture. The result (5.3) was generalized by Frantzikinakis [15] to averages where the shifts are given by independent multivariate polynomials.

Another interesting approximation to Chowla's conjecture is obtained by adding weights to the conjecture. The logarithmic weights $\frac{1}{n}$ are a fruitful choice, since they have the property that

$$\frac{\sum_{x/2\le n\le x}\frac{1}{n}}{\sum_{n\le x}\frac{1}{n}} = o(1),$$

or in other words that the measure of the interval $[\frac{x}{2},x]$ is small. We will see the usefulness of this in Subsection 5.3

In this direction, Tao [98] made a breakthrough by settling the two-point case of Chowla's conjecture with logarithmic weights.

5.4. **Theorem** (Tao). *For any distinct $h_1,h_2\in\mathbb{N}$, we have*

$$\frac{1}{\log x}\sum_{n\le x}\frac{\lambda(n+h_1)\lambda(n+h_2)}{n} = o(1). \tag{5.4}$$

In fact, Tao proved an analogous approximation to Elliott's conjecture (see Section 6) from which (5.4) follows as a special case.

In light of the result (5.4), it is natural to study in detail the logarithmic variant of Chowla's conjecture.

5.5. **Conjecture** (The logarithmic Chowla conjecture). For any $k\ge 1$ and any distinct shifts $h_1,\dots,h_k\in\mathbb{N}$, we have

$$\frac{1}{\log x}\sum_{n\le x}\frac{\lambda(n+h_1)\cdots\lambda(n+h_k)}{n} = o(1).$$

Tao's result [98] is the $k=2$ case of this. In [100] and [III], Tao and the author settled Conjecture 5.5 for all odd k. Therefore, the cases $k=1,2,3,5,7,9\dots$ of the conjecture are now known, whereas the even cases $k\ge 4$ remain open.

5.6. **Theorem** (Tao-T., [100], [III]). *Let $k \geq 1$ be odd and $a_1, \dots, a_k, h_1, \dots, h_k \in \mathbb{N}$. Then we have*

$$\frac{1}{\log x} \sum_{n \leq x} \frac{\lambda(a_1 n + h_1) \cdots \lambda(a_k n + h_k)}{n} = o(1). \tag{5.5}$$

There is no need to assume a non-degeneracy condition on a_i and h_i here, since such a condition makes a difference only for even k.

Our first proof of Theorem 5.6 in [100] utilized deep results of Leibman [62] and Le [61] from ergodic theory, as well as the theory of nilsequences. On the other hand, the proof in [100] gave a general structural theorem for correlations of multiplicative functions, of which Theorem 5.6 is a special case.

The second proof, which we present in [III], proceeds along rather different lines, since after applying the so-called *entropy decrement argument* from [98, Section 3], we do not employ ergodic theory machinery, but use combinatorial results instead. The proof via this method turns out to be both shorter and simpler; in the proof given in [100], the case of the Liouville function was not significantly easier than the case of arbitrary multiplicative functions. Using this combinatorial proof, it would in addition be possible to obtain quantitative error bounds for the right hand side of (5.5); however, these error terms were not analyzed in [III], due to the fact that the error bounds would be very weak[14].

Since Chowla's conjecture can be stated as a claim about the sign patterns of the Liouville function, it is natural that Theorem 5.6 also implies something about sign patterns. We showed in [100] that Theorem 5.6, together with the two-point result and some additional considerations, gives the following.

5.7. **Theorem** (Tao-T., [100], [III]). *Let $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4 \in \{-1, +1\}^4$. Then we have*

$$\lim_{x \to \infty} \frac{1}{x} |\{n \leq x : \ \lambda(n+1) = \varepsilon_1, \lambda(n+2) = \varepsilon_2, \lambda(n+3) = \varepsilon_3\}| = \frac{1}{8}$$

and

$$\liminf_{x \to \infty} \frac{1}{x} |\{n \leq x : \ \lambda(n+1) = \varepsilon_1, \lambda(n+2) = \varepsilon_2, \lambda(n+3) = \varepsilon_3, \lambda(n+4) = \varepsilon_4\}| > 0.$$

We also proved the analogous results for the Möbius function[15]. These improve the result of Matomäki, Radziwiłł and Tao [76] on sign patterns of length 3, as well as Tao's result on sign patterns in [98, Corollary 1.7].

---

[14]For the $k = 2$ case of (5.5), Tao's method [98] gives an error term of the form $O((\log \log \log x)^{-c})$ for some $c > 0$. For $k \geq 3$, we expect even worse error terms.

[15]In the case of $\mu(n)$, we of course need to exclude from the four-point result those sign patterns $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4) \in \{-1, 0, +1\}^4$ which cannot occur for trivial reasons, and in the three point result the density of the set is some function of $\varepsilon_i \in \{-1, 0, +1\}$ instead of $\frac{1}{8}$.

5.2. **Connections to other conjectures**

Chowla's conjecture can be viewed as one instance of the *Möbius randomness law* from [52, p. 338], a heuristic stating that the Möbius function (or the Liouville function) should behave randomly. Another manifestation of this heuristic is a conjecture of Sarnak [89], which states that $\mu(n)$ does not correlate with any bounded sequence of "low complexity". The complexity of a sequence $a : \mathbb{N} \to \mathbb{C}$ is measured in terms of its *topological entropy*, which is the infimum of all $\sigma > 0$ such that sets of the form $\{a(n), a(n+1), \ldots, a(n+k-1)\} \subset \mathbb{C}^k$ can be covered with $\leq \exp((\sigma + o(1))k)$ balls of any fixed radius. With this definition, Sarnak's conjecture takes the form below.

5.8. **Conjecture** (Sarnak)**.** Let $a : \mathbb{N} \to \mathbb{C}$ be a bounded sequence of topological entropy 0. Then we have

$$\frac{1}{x}\sum_{n \leq x} \mu(n)a(n) = o(1).$$

Sarnak's conjecture has been extensively studied in the ergodic theory literature, and many important special cases have been verified; see [14] for a survey. In the ergodic theory literature, one usually assumes in Conjecture 5.8 the (equivalent) condition for the sequence a that it can be written as $a(n) = F(T^n X)$ for (X, T) a *topological dynamical system* of zero topological entropy and $F : X \to \mathbb{C}$ continuous. Here we will not work with the dynamical systems definition, and instead refer to [89] for its details.

It was already observed by Sarnak that his conjecture would follow from Chowla's conjecture. In [99], Tao strengthened the connection between the two conjectures by showing that their logarithmic forms are equivalent (that is, Conjecture 5.5 is equivalent to Conjecture 5.8 with $(1/(\log x))\sum_{n \leq x} \mu(n)a(n)/n$ in place of $(1/x)\sum_{n \leq x} \mu(n)a(n)$). He also showed that both of these conjectures are equivalent to the yet unproved "logarithmic local Gowers uniformity of the Liouville function", which can be thought of as a short exponential sum estimate for the Liouville function and contains as the simplest case the Matomäki-Radziwiłł theorem [74]. Further works that lie at the intersection of the Sarnak and Chowla conjectures include [10] and [17]. In the latter, Frantzikinakis and Host verify many new cases of the logarithmic Sarnak conjecture, and as a byproduct obtain also a "minor arc" Chowla-type result

$$\frac{1}{\log x}\sum_{n \leq x} \lambda(n+h_1)\cdots\lambda(n+h_k)\frac{e(\alpha n)}{n} = o(1)$$

for any $k \geq 1$ and any fixed irrational α. In Article [III], however, we do not make progress on Sarnak's conjecture, since it is the even order cases of Chowla's conjecture that are needed in the proof that Chowla's conjecture implies Sarnak's. It seems therefore that the even order cases lie deeper, and indeed in [100, Remark

1.7] we observed that the $2k$-point case of the logarithmic Chowla conjecture (with dilations as in (5.2)) implies the $(k+1)$-point case.

5.3. **Proof ideas**

The proof of the odd order cases of the logarithmic Chowla conjecture in [III] starts with the averaging over small primes and entropy decrement arguments devised in [98] to deal with the two-point case of the conjecture. The averaging argument enables us to replace a correlation average over n with a double average over n and p for p belonging to some small scale, thus offering more flexibility. More precisely, if we define

$$f_x(a) := \frac{1}{\log x}\sum_{n\le x}\frac{\lambda(n+a)\cdots\lambda(n+ak)}{n}$$

(assuming for simplicity that $a_1=\cdots=a_k=1$ and $h_j=j$ in Theorem 5.6), then the multiplicativity property $\lambda(pn)=-\lambda(n)$ for all primes p allows us to write

$$f_x(1) = -\frac{1}{\log x}\sum_{p\le n'\le px}\frac{\lambda(n'+p)\cdots\lambda(n'+pk)}{n'}p1_{p|n'}+o(1) \tag{5.6}$$

for any prime p and for odd k (for even k we would have a $+$ sign). Taking averages over p, we get the identity

$$f_x(1) = -\frac{m}{2^m}\sum_{2^m\le p<2^{m+1}}\frac{1}{\log x}\sum_{p\le n'\le px}\frac{\lambda(n'+p)\cdots\lambda(n'+pk)}{n'}p1_{p|n'}+O(\varepsilon)$$

for $\varepsilon^{-1}\le m\le \log\log x$. Since logarithmic averages are slowly varying, we can replace the average over $p\le n'\le px$ with an average over $n\le x$ (this is the benefit of logarithmic averaging). Thus we have

$$f_x(1) = -\frac{m}{2^m}\sum_{2^m\le p<2^{m+1}}\frac{1}{\log x}\sum_{n\le x}\frac{\lambda(n+p)\cdots\lambda(n+pk)}{n}p1_{p|n}+O(\varepsilon). \tag{5.7}$$

for $\varepsilon^{-1}\le m\le \log\log x$.

We wish to replace the factor $p1_{p|n'}$ with its average value $1+O(\varepsilon)$ in order to get a bilinear sum over n' and p, for which there are many tools available. This is enabled by Tao's entropy decrement argument [98] (with refinements in [96], [100], [III]), which draws ideas from probability and information theory to show that this replacement can be done for "almost all" scales m in (5.7).

We elaborate on this part of the argument. Firstly, by using the approximate translation invariance of averages, (5.7) becomes

(5.8)

$$f_x(1) = -\frac{1}{\log x}\sum_{n\le x}\frac{m}{2^{2m}}\sum_{2^m\le p<2^{m+1}}\sum_{j\le 2^m}\frac{\lambda(n+j+p)\cdots\lambda(n+j+pk)}{n}p1_{p|n+j}+O(\varepsilon)$$

which is a more convenient form to work with. The task is then to show that, for most choices of the scale m, the sign pattern $\mathbf{X}_m(n) := (\lambda(n), \lambda(n+1), \dots, \lambda(n+2^{m+2}-1))$ and the divisibility conditions $\mathbf{Y}_m(n) := (n \mod p)_{2^m \le p < 2^{m+1}}$ behave essentially independently (with respect to the logarithmic probability on $[1, x]$).

In the language of information theory, we thus want to show that if $\mathbf{X}_m$ and $\mathbf{Y}_m$ are interpreted as random variables, the *entropy*[16] $\mathbf{H}(\mathbf{Y}_m)$ is essentially the same as the *conditional entropy* $\mathbf{H}(\mathbf{Y}_m|\mathbf{X}_m)$ (the two entropies are equal if $\mathbf{Y}_m$ and $\mathbf{X}_m$ are independent, so a small difference between them amounts to near independence). In other words, we want the *mutual information*

$$\mathbf{I}(\mathbf{X}_m, \mathbf{Y}_m) := \mathbf{H}(\mathbf{Y}_m) - \mathbf{H}(\mathbf{Y}_m|\mathbf{X}_m) \tag{5.9}$$

to be small for most m; more precisely, it should be of size $\varepsilon^{10} \cdot 2^m/m$, whereas the trivial upper bound is $\le \mathbf{H}(\mathbf{Y}_m) \ll 2^m$. As mentioned above, mutual information reflects how close two random variables are to being independent (in particular, the information is maximal when one of the two random variables is a deterministic function of the other).

By applying inequalities from information theory, and an insightful pigeonholing argument, Tao showed in [98] that one can indeed bound (5.9) by $\le \varepsilon^{10} \cdot 2^m/m$, not for all scales m, but for infinitely many m. In [III, Section 3], we need a refinement of this, to the effect that if $\mathcal{M}(x, \varepsilon)$ is the set of scales $m \le \log\log x$ for which (5.9) is $> \varepsilon^{10} 2^m/m$, then

$$\sum_{m \in \mathcal{M}(x,\varepsilon)} \frac{1}{m} \ll \varepsilon^{-20},$$

say. In paricular, the set of suitable scales has logarithmic density 1. We refer to [96], [III, Proposition 4.3] for the details[17].

After applying the entropy decrement argument, we know that we can replace in (5.7) the factor $p 1_{p|n}$ with $1 + O(\varepsilon)$ for all $m \le \log\log x$ outside a set whose sum of reciprocals over $[1, \log\log x]$ is $\ll \varepsilon^{-20}$. In particular, we can average logarithmically over different scales m to reach

$$f_x(1) = -\frac{1}{\log_2 H_2 - \log_2 H_1} \sum_{H_1 \le p \le H_2} \frac{1}{p} \frac{1}{\log x} \sum_{n \le x} \frac{\lambda(n+p) \cdots \lambda(n+kp)}{n} + O(\varepsilon) \tag{5.10}$$

for $H_j = H_j(x)$ tending to infinity slowly enough and $H_1(x)$ growing slowly enough in terms of $H_2(x)$. Here $\log_2 x$ is the second iterate of $\log x$.

[16]For the definitions of entropy and other related notions from information theory, see Section 1.
[17]For technical reasons, those works deal with a more general notion of information, namely *conditional mutual information.*

We can apply the same argument again to the right-hand side of (5.10), finding

$$f_x(1) = +\frac{1}{\log_2 H_2 - \log_2 H_1} \sum_{H_1 \le p_1 \le H_2} \frac{1}{p_1} \frac{1}{\log_2 H_4 - \log_2 H_3} \sum_{H_3 \le p_2 \le H_4} \frac{1}{p_2} \frac{1}{\log x} \sum_{n \le x} \frac{\lambda(n + p_1 p_2) \cdots \lambda(n + k p_1 p_2)}{n} + O(\varepsilon), \tag{5.11}$$

where $H_1 < H_2 < H_3 < H_4$ and $H_j(x)$ grows slowly enough in terms of $H_{j+1}(x)$, and $H_4(x)$ tends to infinity slowly. Crucially, we have a $+$ sign in (5.11) and a $-$ sign in (5.10); this allows us to break the symmetry of the correlations.

We can easily replace the averages over primes with averages over the integers weighted by the von Mangoldt function $\Lambda(d)$, so (5.10) (with H_1 and H_2 replaced with H_3 and H_4) and (5.11) take the forms

$$f_x(1) = -\frac{1}{\log_2 H_4 - \log_2 H_3} \sum_{H_3 \le d \le H_4} \frac{\Lambda(d)}{d \log d} \frac{1}{\log x} \sum_{n \le x} \frac{\lambda(n+d) \cdots \lambda(n+kd)}{n} + O(\varepsilon) \tag{5.12}$$

and

$$f_x(1) = +\frac{1}{\log_2 H_2 - \log_2 H_1} \sum_{H_1 \le d_1 \le H_2} \frac{\Lambda(d_1)}{d_1 \log d_1} \frac{1}{\log_2 H_4 - \log_2 H_3} \sum_{H_3 \le d_2 \le H_4} \frac{\Lambda(d_2)}{d_2 \log d_2} \frac{1}{\log x} \sum_{n \le x} \frac{\lambda(n + d_1 d_2) \cdots \lambda(n + k d_1 d_2)}{n} + O(\varepsilon), \tag{5.13}$$

respectively.

We now encounter multilinear averages of the form

$$\frac{1}{N^2} \sum_{d \le N} \sum_{n \le N} \theta(d) f_1(n+d) \cdots f_k(n+kd), \tag{5.14}$$

where $f_1, \dots, f_k : \mathbb{N} \to \mathbb{C}$ are some functions with $|f_i| \le 1$ and $\theta : \mathbb{N} \to \mathbb{C}$ is some other function (in this case a normalized version of $\Lambda(d)$). The expression (5.14) thus counts patterns of the form $(d, n+d, \dots, n+kd)$ with weights. Such averages have been widely studied both in the additive combinatorics and the ergodic theory literature (see for instance [101, Chapter 11]), and by a version of the so-called generalized von Neumann theorem [III, Lemma 5.2], it turns out that one has the bound[18]

$$\left|\frac{1}{N^2}\sum_{d\le N}\sum_{n\le N}\theta(d)f_1(n+d)\cdots f_k(n+kd)\right|\le C_k\|\theta\|_{U^k[N]}+o(1),$$

where $\|\theta\|_{U^k[N]}$ is the U^k *Gowers norm* of θ on $[1,N]$ (see [101, Chapter 11]). Thus, analyzing (5.12) and (5.13) has been reduced to understanding the Gowers norm of $\Lambda(Wn+b)-1$, where $W=\prod_{p\le w}p$, $(b,W)=1$, and w is a large constant.

It is known that the W-tricked von Mangoldt function has negligible Gowers norms; this was proved by Green, Tao and Ziegler in a series of breakthroughs [28],[29],[30], [31]. Therefore, we can remove the von Mangoldt function weight both in the average (5.12) and the average (5.13), after splitting the sums into residue classes (mod W). This leads, after some considerations, to

(5.15)

$$\begin{aligned} f_x(1) &= \frac{W}{\varphi(W)}\frac{1}{\log_2 H_4-\log_2 H_3}\sum_{\substack{H_3\le d\le H_4\\(d,W)=1}}\frac{1}{d\log d\log x}\sum_{n\le x}\frac{\lambda(n+d)\cdots\lambda(n+dk)}{n}+O(\varepsilon)\\ &= -f_x(1)+O(\varepsilon).\end{aligned}$$

Importantly, $f_x(1)$ appears with different signs in (5.15), so $f_x(1)=O(\varepsilon)$, after which we can send $\varepsilon\to 0$. This concludes the sketch of the proof; for the full proof, see [III].

[18] As is shown for example in [101, Chapter 11], the weighted arithmetic progression patterns $(n+d,\dots,n+kd)$ are controlled by the U^{k-1} Gowers norm, but the pattern $(d,n+d,\dots,n+kd)$ in (5.14) has "complexity" one higher, and should thus be controlled by the U^k Gowers norm.

6. Binary correlations of multiplicative functions and applications

Article [IV] is concerned with binary correlations of multiplicative functions with logarithmic averaging. Before stating the results of that article, we review what is known and conjectured on correlations of multiplicative functions.

6.1. **Correlations of multiplicative functions**

A function $g : \mathbb{N} \to \mathbb{C}$ is called multiplicative if it satisfies $g(mn) = g(m)g(n)$ whenever $m, n \in \mathbb{N}$ are coprime. In what follows, we will restrict attention to 1-bounded multiplicative functions, that is, multiplicative functions taking values in the unit disk $\mathbb{D} := \{z \in \mathbb{C} : \ |z| \leq 1\}$, since much less is known about the behavior of unbounded multiplicative functions.

A fundamental notion in multiplicative number theory is the *pretentious distance* $\mathbb{D}(f, g; x)$ between two multiplicative functions $f, g : \mathbb{N} \to \mathbb{D}$, introduced by Granville and Soundararajan [24]. This quantity is defined as

$$\mathbb{D}(f, g; x) := \left(\sum_{p \leq x} \frac{1 - \mathrm{Re}(f(p)\overline{g(p)})}{p} \right)^{1/2}, \tag{6.1}$$

and it is a pseudometric[19] and, heuristically, if f and g "behave similarly" (when it comes to their mean values or correlations), then the distance between them is "small".

The Dirichlet characters $\chi(n)$ and the Archimedean characters n^{it} are important classes of 1-bounded multiplicative functions, and although their complexity is relatively low in the sense that $\chi(n)$ is periodic and n^{it} is slowly varying, one usually wants to exclude these functions when studying mean values or correlations of multiplicative functions, as these two classes of functions exhibit different behavior from other functions in this context. One thus classifies 1-bounded multiplicative functions as either
(i) pretentious, in the sense that $\mathbb{D}(g, \chi(n)n^{it}; \infty) < \infty$ for some Dirichlet character χ and some $t \in \mathbb{R}$,
or
(ii) non-pretentious, in the sense that $\mathbb{D}(g, \chi(n)n^{it}; \infty) = \infty$ for all Dirichlet characters χ and all $t \in \mathbb{R}$.

By the zero-free region for the Dirichlet L-functions, the Liouville function $\lambda(n)$ from Section 5 is non-pretentious, whereas any multiplicative function $f : \mathbb{N} \to \mathbb{D}$ with $f(p) \neq 1$ for only finitely many primes p is an example of a pretentious function (one can take $\chi \equiv 1$, $t = 0$ in (i)).

[19] This means that it satisfies the axioms of a metric, excluding the property that $\mathrm{d}(x, y) = 0 \Rightarrow x = y$.

The mean values

$$\frac{1}{x}\sum_{n\le x} g(n) \tag{6.2}$$

of multiplicative functions are connected to many topics of interest in multiplicative number theory, including the prime number theorem and its generalizations, sieve methods, and probabilistic number theory. The asymptotics of these mean values are described by a theorem of Halász [33] from the 1960s (generalizing a theorem of Wirsing [113] from the real-valued case), and the result demonstrates the need for distinguishing pretentious and non-pretentious functions from each other.

6.1. **Theorem** (Halász). *Let $g:\mathbb{N}\to\mathbb{D}$ be a 1-bounded multiplicative function. Then*
(i) If there exists $t\in\mathbb{R}$ such that $\mathbb{D}(g,n^{it};\infty)<\infty$, we have

$$\frac{1}{x}\sum_{n\le x} g(n) = (1+o(1))\frac{x^{it}}{1+it}\prod_p\left(1-\frac{1}{p}\right)\left(1+\frac{g(p)}{p^{1+it}}+\frac{g(p^2)}{p^{2(1+it)}}+\cdots\right).$$

(ii)If no such t exists, we have

$$\frac{1}{x}\sum_{n\le x} g(n) = o(1).$$

For a proof of the theorem, see [102, Section III.4]. Among other things, Theorem 6.1 implies that if $g:\mathbb{N}\to[-1,1]$ is real-valued, then the mean value of g always exists (that is, (6.2) converges as $x\to\infty$).

We wish to understand the much more general *correlation averages* of bounded multiplicative functions $g_1,\ldots,g_k:\mathbb{N}\to\mathbb{D}$, defined as

$$\frac{1}{x}\sum_{n\le x} g_1(n+h_1)\cdots g_k(n+h_k), \tag{6.3}$$

where $h_1,\ldots,h_k\in\mathbb{N}$ are fixed, distinct integers. These correlations have a number of applications; most notably, in the case of the Liouville function showing that the correlations are small reduce to the celebrated Chowla conjecture, discussed in Section 5 and in particular, gives information on the sign patterns of the Liouville function, studied in [76], [98], [100], [17]. In a very different and surprising direction, Tao [97] used his breakthrough on two-point correlations to settle the Erdős discrepancy problem [13] in combinatorics. There are further applications to discrepancy of multiplicative functions in [58], rigidity theorems for multiplicative functions in [60], and to distribution laws of additive functions in [12]. In article [IV], we give further applications, discussed in Subsection 6.2.

The central conjecture pertaining to (6.3) is that of Elliott [11], [12] from the 1990s. His conjecture states that, in the case where at least one of $g_1,\ldots,g_k:\mathbb{N}\to\mathbb{D}$ is non-pretentious, distinct shifts of the functions g_j should behave independently of each other.

6.2. **Conjecture** (Elliott). Let $k \geq 1$ and let $g_1, \dots, g_k : \mathbb{N} \to \mathbb{D}$ be 1-bounded multiplicative functions and $h_1, \dots, h_k \in \mathbb{N}$ distinct shifts. Then we have

$$\frac{1}{x}\sum_{n \leq x} g_1(n+h_1)\cdots g_k(n+h_k) = o(1) \tag{6.4}$$

unless for all $1 \leq j \leq k$ there exists a Dirichlet character χ_j such that

$$\liminf_{x\to\infty} \inf_{|t|\leq x} \mathbb{D}(g_j, \chi_j(n)n^{it}; x) < \infty.$$

The formulation above takes into account the observation in [75, Appendix B] that the original conjecture in [11], [12] has to be slightly modified in the complex-valued case. As in the case of Halász's theorem (Theorem 6.1), the property (6.4) often fails in the pretentious case; for example

$$\frac{1}{x}\sum_{n\leq x} n^{it}(n+1)^{-it} = 1 + o(1) \quad \text{and} \quad \frac{1}{x}\sum_{n\leq x} \chi_3(n)\chi_3(n+1) = -\frac{1}{3} + o(1),$$

where χ_3 is the real non-principal Dirichlet character modulo 3. On the other hand, a theorem of Klurman [58, Theorem 1.3] gives a formula for (6.3) in the case where $g_1, \dots, g_k$ are fixed pretentious functions.

In the form presented above, Conjecture 6.2 is open for all $k \geq 2$, whereas the $k = 1$ case follows from Theorem 6.1. However, several variants of (6.4) have been established in the last few years. In particular, Matomäki, Radziwiłł and Tao [75] showed that Elliott's conjecture holds on average over the shifts $h_1, \dots, h_k$. Tao [98] made another breakthrough by proving the binary case $k = 2$ of Elliott's conjecture with logarithmic averaging.

6.3. **Theorem** (Tao). *Let $g_1, g_2 : \mathbb{N} \to \mathbb{D}$ be 1-bounded multiplicative functions and $h_1 \neq h_2$ natural numbers. Then we have*

$$\frac{1}{\log x}\sum_{n\leq x} \frac{g_1(n+h_1)g_2(n+h_2)}{n} = o(1)$$

unless for both $j \in \{1, 2\}$ there exists a Dirichlet character χ_j such that

$$\liminf_{x\to\infty} \inf_{|t|\leq x} \mathbb{D}(g_j, \chi_j(n)n^{it}; x) < \infty.$$

For many purposes, this logarithmic averaging is acceptable; see [97], [58], [60] for some applications. In [100] we generalized Theorem 6.3 to the higher order cases $k \geq 3$, under an additional non-pretentiousness assumption on the product of the functions involved.

6.4. **Theorem** (Tao-T., [100]). *Let $k \geq 1$ and let $g_1, \dots, g_k : \mathbb{N} \to \mathbb{D}$ be 1-bounded multiplicative functions and $h_1, \dots, h_k \in \mathbb{N}$ natural numbers. Then we have*

$$\frac{1}{\log x} \sum_{n \leq x} \frac{g_1(n+h_1)\cdots g_k(n+h_k)}{n} = o(1)$$

unless there exists a Dirichlet character χ for which the product $g_1 \cdots g_k$ weakly pretends *to be χ, in the sense that $\mathbb{D}(g_1 \cdots g_k, \chi; x)^2 = o(\log\log x)$.*

We applied this result to settle the odd order cases of the logarithmically averaged Chowla conjecture; see Section 5.

Let us also mention a different line of study to Elliott's conjecture, namely two-dimensional variants of it. The two-dimensional version of Elliott's conjecture states that

$$\frac{1}{x^2} \sum_{d \leq x} \sum_{n \leq x} g_1(n+dh_1)\cdots g_k(n+dh_k) = o(1),$$

given the assumptions of Conjecture 6.2. This was proved by Frantzikinakis and Host in [16], and further works on two-dimensional correlations include those of Matthiesen [78] and Klurman–Mangerel [59].

6.2. The result and its applications

In Article [IV], we generalize Tao's result on the binary logarithmic Elliott conjecture, but in a different direction than in [100], where higher order correlations were considered. Namely, we show that for a large class of real-valued multiplicative functions $g_1, g_2 : \mathbb{N} \to [-1, 1]$ we can give an asymptotic formula for their correlation (and typically the asymptotic formula has a nonzero main term). The class of functions we consider is defined as follows.

6.5. **Definition** (Uniformity assumption). Let $x \geq 1$, $1 \leq Q \leq x$ and $\eta > 0$. For a function $g : \mathbb{N} \to \mathbb{D}$, denote $g \in \mathcal{U}(x, Q, \eta)$ if we have the estimate

$$\left| \frac{1}{x} \sum_{\substack{x \leq n \leq 2x \\ n \equiv a \pmod{q}}} g(n) - \frac{1}{qx} \sum_{x \leq n \leq 2x} g(n) \right| \leq \frac{\eta}{q} \quad \text{for all} \quad 1 \leq a \leq q \leq Q.$$

From Halász's theorem we see (as was observed in [IV, Remark 1.3]) that if $g : \mathbb{N} \to \mathbb{D}$ is non-pretentious in the sense that $\inf_{|t| \leq x} \mathbb{D}(g, \chi(n)n^{it}; x) \geq \varepsilon^{-10}$ for all Dirichlet characters χ of modulus $\leq \varepsilon^{-10}$ (and with $\varepsilon > 0$ small), then $g \in \mathcal{U}(x, \varepsilon^{-1}, \varepsilon)$ for $x \geq x_0(\varepsilon)$. This means that the class of functions in Definition 6.5 is larger than the class of real-valued functions considered in Conjecture 6.2 or in [98]. Very importantly, Definition 6.5 allows the multiplicative function g to depend on the summation length x, as will be the case in our applications. One can for example show that if $\alpha \in (0, 1)$ is given, then the indicator of smooth numbers[20] $g(n) := 1_{n \text{ is } x^{\alpha}-\text{smooth}}$ is a multiplicative function satisfying $g \in \mathcal{U}(x, \varepsilon^{-1}, \varepsilon)$ for $x \geq x_0(\varepsilon, \alpha)$, although g

[20]We say that n is y–smooth (also called y-friable) if n has no prime factor larger than y.

pretends to be 1 on the interval $[1, x]$.

The main result in [IV] then states that if $g_1, g_2 : \mathbb{N} \to [-1, 1]$ are two multiplicative functions, possibly depending on x, and g_1 is uniformly distributed at scale x in the sense of Definition 6.5, then the shifts of g_1 and g_2 are independent of each other.

6.6. **Theorem** (Article [IV]). *Let a small real number $\varepsilon > 0$, a fixed integer shift $h \neq 0$, and a function $\omega : \mathbb{R}_{\geq 1} \to \mathbb{R}_{\geq 1}$ with $1 \leq \omega(X) \leq \log(3X)$ and $\omega(X) \xrightarrow{X\to\infty} \infty$ be given. Let $x \geq x_0(\varepsilon, h, \omega)$. Then, for any multiplicative functions $g_1, g_2 : \mathbb{N} \to [-1, 1]$ with $g_1 \in \mathcal{U}(x, \varepsilon^{-1}, \varepsilon)$, we have*

(6.5)
$$\frac{1}{\log \omega(x)} \sum_{\frac{x}{\omega(x)} \leq n \leq x} \frac{g_1(n)g_2(n+h)}{n} = \left(\frac{1}{x} \sum_{x \leq n \leq 2x} g_1(n)\right)\left(\frac{1}{x} \sum_{x \leq n \leq 2x} g_2(n)\right) + o_{\varepsilon\to 0}(1).$$

Here $o_{\varepsilon\to 0}(1)$ denotes some function that tends uniformly to 0 as $\varepsilon \to 0$. Note also that even if $h < 0$ in Theorem 6.6, $g_2(n + h)$ is still well-defined, as the function $x_0(\cdot)$ above can be chosen to be large enough, so that $\frac{x}{\omega(x)} > h$ for $x \geq x_0(\varepsilon, h, \omega)$.

As mentioned, Theorem 6.6 contains the real-valued case of Tao's result [98] and shows that g_1 and g_2 are discorrelated in the sense that the correlation of g_1 and g_2 is the product of their mean values. In the complex-valued case, Theorem 6.6 does not hold as such, as is seen by taking g_1 and g_2 to be suitable Archimedean characters (such as $g_1(n) = g_2(n) = n^{it}$ with $t \neq 0$). It would nevertheless be possible to generalize it to the case where g_1 and g_2 take values in roots of unity of a fixed order.

Theorem 6.6 could also be generalized to the case of functions that are uniformly distributed only in coprime residue classes, instead of all residue classes as in Definition 6.5. However, this would significantly complicate the main term on the right-hand side of (6.5) and make it dependent on the shift h (as is seen by considering the simple example $g_1(n) = g_2(n) = 1_{n\equiv 1 \pmod 2}$). Therefore, we do not pursue this generalization.

The utility of Theorem 6.6 lies in its uniformity over the choice of the functions g_1, g_2. For example, the theorem can be applied to the interesting cases
(i) $g(n) = 1_{n \text{ is } x^{\alpha}-\text{smooth}}$
and
(ii) $g(n) = \chi_Q(n)$ where χ_Q is a real non-principal character $\pmod Q$ with $Q = Q(x) \leq x^{4-\varepsilon}$ cube-free[21] (so Q can be very large in terms of x).
In the case of (i), the result of [98] is clearly not applicable, and also in case (ii), for

[21]We say that n is cube-free if $p^3 \nmid n$ for all primes p.

all we know, it could be that[22] $\mathbb{D}(\chi_Q, 1; x) \ll 1$, in which case [98] does not apply. The range $Q \le x^{4-\varepsilon}$ in (ii) is the same as in a celebrated estimate of Burgess [4], a special case of which implies that

$$\sum_{n \le x} \chi_Q(n) = o(x), \tag{6.6}$$

uniformly for $Q \le x^{4-\varepsilon}$ cube-free. We note that (6.6) is not enough to exclude the (unlikely) scenario that $\chi_Q(p) = 1$ for all $p \le x^{\varepsilon}$.

We mentioned earlier the result of Klurman [58, Theorem 1.3] that gives an asymptotic formula for the correlations (6.3) in the case where all the functions $g_1, \ldots, g_k$ are fixed and pretentious. Nevertheless, this asymptotic cannot be applied to (i) or (ii), since both of these two functions depend on x in a very essential way, whereas in [58] it is necessary that the functions are (almost) independent of x (in fact, the asymptotic formula in [58, Theorem 1.3] does not predict the correct asymptotic for the autocorrelations of the functions in (i) or (ii)).

As the examples (i) and (ii) indicate, Theorem 6.6 should yield new results on consecutive smooth (friable) numbers and quadratic residues. We confirm this in [IV]. Define the function $P^+(n)$ that outputs the largest prime factor of $n \in \mathbb{N}$, with the convention that $P^+(1) = 1$. Then n is y–smooth if and only if $P^+(n) \le y$. The distribution of smooth numbers is well-understood (see [45] for a survey), but much more elusive is the simultaneous distribution of two or more consecutive smooth numbers. Related to this, Erdős and Turán [94] posed the following problem.

6.7. **Conjecture.** The asymptotic density[23] of the set

$$\{n \in \mathbb{N} : \ P^+(n) < P^+(n+1)\}$$

exists and equals $\frac{1}{2}$.

By applying Theorem 6.6 to the indicator function of x^{α}–smooth numbers for various α, and doing some additional deductions, we were able to prove a logarithmic variant of this conjecture.

6.8. **Theorem** (Article [IV]). *The logarithmic density[24] of the set*

$$\{n \in \mathbb{N} : \ P^+(n) < P^+(n+1)\}$$

exists and equals $\frac{1}{2}$.

---

[22]It is a well-known conjecture, due to Vinogradov, that, for any $\varepsilon > 0$ and any $Q \ge Q_0(\varepsilon)$, there is a quadratic nonresidue modulo Q on $[1, Q^{\varepsilon}]$. But this is open, and if it fails, then χ_Q pretends to be 1 on $[1, Q]$.

[23]We define the asymptotic density of $A \subset \mathbb{N}$ as $\lim_{x \to \infty} \frac{1}{x} \sum_{n \le x, n \in A} 1$, whenever the limit exists. The upper and lower asymptotic densities are defined analogously with $\limsup$ and $\liminf$.

[24]We define the logarithmic density of $A \subset \mathbb{N}$ as $\lim_{x \to \infty} \frac{1}{\log x} \sum_{n \le x, n \in A} \frac{1}{n}$, whenever the limit exists. The upper and lower logarithmic densities are defined analogously with $\limsup$ and $\liminf$.

We mention in passing that we also proved in [IV] some generalizations of Theorem 6.8, including a logarithmic version of a conjecture of Erdős and Pomerance ([IV, Theorem 1.12]).

We can also say something about *asymptotic* densities of sets related to two consecutive smooth numbers. In this case, we are not able to determine the precise value of the density, but we can at least show that the lower density is positive.

6.9. **Theorem** (Article [IV]). *Let $a, b, c, d \in (0,1)$ be real numbers with $a < b$ and $c < d$. Then the set*

$$\{n \in \mathbb{N}: \ n^a < P^+(n) < n^b, \ n^c < P^+(n+1) < n^d\}$$

has positive asymptotic lower density.

This theorem implies a result of Hildebrand [44], which is the special case $(a, b) = (c, d)$ (Hildebrand also considers more general "stable sets", in addition to sets of smooth numbers). Our theorem also reproves a recent result of Wang [110, Théorème 2] on the truncated largest prime factor $P_y^+(n) := \max\{p \leq y: \ p \mid n\}$ at two consecutive integers. This result states that, if $a \in (0,1)$ is fixed, then $P_{x^a}^+(n) < P_{x^a}^+(n+1)$ for a positive lower density of integers $n \leq x$.

Another source for applications of Theorem 6.6 is the collection of real non-principal Dirichlet characters whose modulus $Q(x)$ grows moderately fast in terms of x. A fundamental result of Burgess [4] from 1963 says, among other things, that if χ_Q is a non-principal Dirichlet character of cube-free[25] modulus $Q = Q(x)$, and $\varepsilon > 0$ is fixed, then

$$\sum_{n \leq x} \chi_Q(n) = o(x), \tag{6.7}$$

uniformly for $Q \leq x^{4-\varepsilon}$. The range of Q here is still the best one known up to the ε in the exponent.

By employing the Burgess bound (6.7), we can show that if χ_Q is as above with $Q \leq x^{4-\varepsilon}$ cube-free, then the uniformity assumption $\chi_Q \in \mathcal{U}(x, \eta^{-1}, \eta)$ holds for $x \geq x_0(\eta, \varepsilon)$; see [IV, Section 4]. Therefore, Theorem 6.6 implies a result on the sums of χ_Q along reducible quadratics $n(n+h)$.

6.10. **Theorem** (Article [IV]). *Let a small number $\varepsilon > 0$, a fixed integer $h \neq 0$, and a function $1 \leq \omega(X) \leq \log(3X)$ tending to infinity be given. For $x \geq x_0(\varepsilon, h, \omega)$, let $Q = Q(x) \leq x^{4-\varepsilon}$ be a cube-free natural number with $Q(x) \xrightarrow{x \to \infty} \infty$. Then, the real primitive Dirichlet character χ_Q modulo Q satisfies*

$$\frac{1}{\log \omega(x)} \sum_{\frac{x}{\omega(x)} \leq n \leq x} \frac{\chi_Q(n(n+h))}{n} = o(1). \tag{6.8}$$

[25]We say that Q is cube-free if it is not divisible by the cube of any prime.

Moreover, if Q is as before and QNR denotes quadratic nonresidue[26], *we have*

$$\frac{1}{\log x}\sum_{\substack{n\leq x\\ n,n+1\,\mathrm{QNR}\pmod{Q}}}\frac{1}{n}=\frac{1}{4}\prod_{p\mid Q}\left(1-\frac{2}{p}\right)+o(1) \tag{6.9}$$

and

$$\frac{1}{x}|\{n\leq x:\ n \text{ and } n+1 \text{ QNR} \pmod{Q}\}|\gg\prod_{p\mid Q}\left(1-\frac{2}{p}\right). \tag{6.10}$$

We remark that the well-known Weil bound [52, Theorem 11.23] for character sums would give, for prime values of Q, the estimate (6.8) only in the smaller range $Q=o(\frac{x^2}{\log x})$.

Lastly, we employ Theorem 6.6 to study the number of large prime factors of consecutive integers. For $y\geq 1$, define the truncated count of prime factors as $\omega_{>y}(n):=|\{p>y:\ p\mid n\}|$. It is natural to conjecture that the numbers of large prime factors (say $>n^{\varepsilon}$) of n and $n+1$ are independent. Choosing in Theorem 6.6 multiplicative functions of the form $z^{\omega_{>x^a}(n)}$ with $z\in[-1,1]$, and using a generating function argument, we show that this independence property indeed holds, at least in the logarithmic sense.

6.11. **Theorem** (Article [IV]). *Let $a,b\in(0,1)$ be real numbers and $0\leq k<\frac{1}{a}$, $0\leq \ell<\frac{1}{b}$ integers. Then, if $\delta(\cdot)$ stands for logarithmic density, we have*

$$\begin{aligned}&\delta(\{n\in\mathbb{N}:\ \omega_{>n^a}(n)=k,\ \omega_{>n^b}(n+1)=\ell\})\\ &=\delta(\{n\in\mathbb{N}:\ \omega_{>n^a}(n)=k\})\cdot\delta(\{n\in\mathbb{N}:\ \omega_{>n^b}(n)=\ell\}).\end{aligned}$$

Moreover, the set $\{n\in\mathbb{N}:\ \omega_{>n^a}(n)=k,\ \omega_{>n^b}(n+1)=\ell\}$ has positive asymptotic lower density.

Theorem 6.11 in a sense complements the result of Daboussi–Sárközy [9] and Mangerel [68], which states that if $\omega_{<y}(n)=|\{p<y:\ p\mid n\}|$ is the count of the *small* prime factors of n, then we have the independence of small primes property

$$\frac{1}{x}\sum_{n\leq x}(-1)^{\omega_{<x^{\varepsilon}}(n)}(-1)^{\omega_{<x^{\varepsilon}}(n+1)}=o_{\varepsilon\to 0}(1). \tag{6.11}$$

In comparison, Theorem 6.11 implies among other things the independence of large primes property

$$\frac{1}{\log x}\sum_{n\leq x}\frac{(-1)^{\omega_{>x^{\varepsilon}}(n)}(-1)^{\omega_{>x^{\varepsilon}}(n+1)}}{n}=o_{\varepsilon\to 0}(1). \tag{6.12}$$

The methods used to prove (6.11) and (6.12) are however completely different, the proof of (6.11) being based on sieve theory.

[26]We say that n is a quadratic nonresidue $\pmod{Q}$ if $\chi_Q(n)=-1$.

6.3. **Proof sketch for the main result**

The proof of Theorem 6.6 makes use of the ideas Tao [98] developed for his proof of Theorem 6.3; these are an averaging over small primes argument and the entropy decrement argument, also discussed in Section 5.

The averaging over small primes works as follows. Suppose for simplicity that g_1, g_2 are completely multiplicative and take only values ± 1. Then, for any prime $p \le \log \omega(x)$, we have

(6.13)
$$\begin{aligned}\frac{1}{\log \omega(x)} \sum_{\frac{x}{\omega(x)} \le n \le x} \frac{g_1(n)g_2(n+h)}{n} &= \frac{g_1g_2(p)}{\log \omega(x)} \sum_{\frac{x}{\omega(x)} \le n \le x} \frac{g_1(pn)g_2(pn+ph)}{n} \\ &= \frac{g_1g_2(p)}{\log \omega(x)} \sum_{\frac{x}{\omega(x)} \le n' \le x} \frac{g_1(n')g_2(n'+ph)}{n'} p1_{p|n'} + O(\varepsilon)\end{aligned}$$

where we wrote $n' = pn$ and used the fact that the average is a logarithmic one. We can then sum (6.13) over p to conclude that

$$\begin{aligned}&\frac{1}{\log \omega(x)} \sum_{\frac{x}{\omega(x)} \le n \le x} \frac{g_1(n)g_2(n+h)}{n} \\ &= \frac{m}{2^m} \sum_{2^m \le p < 2^{m+1}} g_1(p)g_2(p) \frac{1}{\log \omega(x)} \sum_{\frac{x}{\omega(x)} \le n' \le x} \frac{g_1(n')g_2(n'+ph)}{n'} p1_{p|n'} + O(\varepsilon),\end{aligned} \tag{6.14}$$

where $\varepsilon^{-1} \le m \le \log\log \omega(x)$. By the entropy decrement argument, developed by Tao in [98] and based on inequalities from information theory, we can replace $p1_{p|n'}$ with its average value $1 + O(\varepsilon)$ for some suitable, large $m = m(\varepsilon)$. The advantage gained is that now (6.14) becomes a bilinear average

$$\frac{m}{2^m} \sum_{2^m \le p < 2^{m+1}} g_1(p)g_2(p) \frac{1}{\log \omega(x)} \sum_{\frac{x}{\omega(x)} \le n \le x} \frac{g_1(n)g_2(n+ph)}{n} + o(1),$$

where n and p have been decoupled. This enables us to apply the circle method. In the same spirit as in [98], the circle method gives the anticipated asymptotic for this sum, provided that we prove the short exponential sum bound[27]

$$\sup_{\alpha \in \mathbb{R}} \frac{1}{x} \int_x^{2x} \left| \frac{1}{H} \sum_{y \le n \le y+H} (g_1(n) - \delta_1)e(n\alpha) \right| dy = o_{\varepsilon \to 0}(1), \tag{6.15}$$

where δ_1 is the mean value of g_1 on $[x, 2x]$ and $H \asymp 2^{(1+O(\varepsilon))m}$ with $m = m(\varepsilon)$ large. This estimate deviates from what was used in [98], since there (6.15) was used in the non-pretentious case covered by a result of Matomäki, Radziwiłł and Tao [75,

[27] In reality, we need to consider the integral of the exponential sum over more general intervals $[y, 2y]$ with $\frac{x}{\omega(x)} \le y \le x$.

Theorem 1.7]. The case where g_1 is uniformly distributed in the sense of Definition 6.5 is not addressed in [75], but can be dealt with using the tools employed there.

The proof of (6.15) naturally splits into the *major arc* case, where α is close to a rational number with small denominator, and the opposite *minor arc* case. In the minor arc case, we can ignore the constant term δ_1 in (6.15) and then follow the argument in [75, Section 3], as that is based solely on the multiplicativity and boundedness of g_1.

In the major arc case, in contrast, we plainly need to use the uniform distribution property of g_1, as the result fails for example for Dirichlet characters, which are not equidistributed. If α is on a major arc, then $e(n\alpha)$ is essentially periodic, so we may make it essentially constant by splitting n into residue classes. Then we end up with the need to prove that

$$\frac{1}{x}\int_x^{2x}\left|\frac{1}{H}\sum_{\substack{y\leq n\leq y+H\\ n\equiv b \pmod{q}}} g_1(n)-\frac{1}{qH}\sum_{y\leq n\leq y+H} g_1(n)\right| dy=\frac{o_{\varepsilon\to 0}(1)}{q} \tag{6.16}$$

uniformly for $1\leq b\leq q\leq \varepsilon^{-1}$. Here we used the fact that δ_1 is the mean of g_1 also in arithmetic progressions of modulus $\leq \varepsilon^{-1}$.

The estimate (6.16) follows for $q=1$ from the Matomäki–Radziwiłł theorem [74] (since g_1 is real-valued), and it turns out that for $q>1$, by expanding $1_{n\equiv b \pmod{q}}$ in terms of characters, we can use the complex-valued case of that theorem from [75, Appendix A] together with some pretentious distance estimates. This then leads to the desired conclusion (6.16). For the proof in its entirety, we refer to [IV].

References

[1] S. Baier. A note on Diophantine approximation with Gaussian primes. *ArXiv e-prints*, September 2016.

[2] R. C. Baker and G. Harman. The difference between consecutive primes. *Proc. London Math. Soc. (3)*, 72(2):261–280, 1996.

[3] R. C. Baker, G. Harman, and J. Pintz. The difference between consecutive primes. II. *Proc. London Math. Soc. (3)*, 83(3):532–562, 2001.

[4] D. A. Burgess. On character sums and L-series. II. *Proc. London Math. Soc. (3)*, 13:524–536, 1963.

[5] J. R. Chen. On the representation of a larger even integer as the sum of a prime and the product of at most two primes. *Sci. Sinica*, 16:157–176, 1973.

[6] J. R. Chen and C. D. Pan. The exceptional set of Goldbach numbers. I. *Sci. Sinica*, 23(4):416–430, 1980.

[7] S. Chowla. *The Riemann hypothesis and Hilbert's tenth problem.* Mathematics and Its Applications, Vol. 4. Gordon and Breach Science Publishers, New York-London-Paris, 1965.

[8] H. Cramér. On the order of magnitude of the difference between consecutive prime numbers. *Acta Arith.*, 2:23–46, 1936.

[9] H. Daboussi and A. Sárközy. On the correlation of the truncated Liouville function. *Acta Arith.*, 108(1):61–76, 2003.

[10] H. El Abdalaoui, J. Kułaga-Przymus, M. Lemańczyk, and T. de la Rue. The Chowla and the Sarnak conjectures from ergodic theory point of view. *Discrete Contin. Dyn. Syst.*, 37(6):2899–2944, 2017.

[11] P. D. T. A. Elliott. On the correlation of multiplicative functions. *Notas Soc. Mat. Chile*, 11(1):1–11, 1992.

[12] P. D. T. A. Elliott. On the correlation of multiplicative and the sum of additive arithmetic functions. *Mem. Amer. Math. Soc.*, 112(538):viii+88, 1994.

[13] P. Erdős. Some unconventional problems in number theory. In *Journées Arithmétiques de Luminy (Colloq. Internat. CNRS, Centre Univ. Luminy, Luminy, 1978)*, volume 61 of *Astérisque*, pages 73–82. Soc. Math. France, Paris, 1979.

[14] S. Ferenczi, J. Kułaga-Przymus, and M. Lemańczyk. Sarnak's Conjecture – what's new. *ArXiv e-prints*, October 2017.

[15] N. Frantzikinakis. An averaged Chowla and Elliott conjecture along independent polynomials. *Int. Math. Res. Not. IMRN*, 2018(12):3721–3743, 2018.

[16] N. Frantzikinakis and B. Host. Asymptotics for multilinear averages of multiplicative functions. *Math. Proc. Cambridge Philos. Soc.*, 161(1):87–101, 2016.

[17] N. Frantzikinakis and B. Host. The logarithmic Sarnak conjecture for ergodic weights. *Ann. of Math. (2)*, 187(3):869–931, 2018.

[18] J. Friedlander and H. Iwaniec. The polynomial $X^2 + Y^4$ captures its primes. *Ann. of Math. (2)*, 148(3):945–1040, 1998.

[19] J. Friedlander and H. Iwaniec. *Opera de cribro*, volume 57 of *American Mathematical Society Colloquium Publications.* American Mathematical Society,

Providence, RI, 2010.
[20] P. X. Gallagher. On the distribution of primes in short intervals. *Mathematika*, 23(1):4–9, 1976.
[21] D. A. Goldston, S. W. Graham, J. Pintz, and C. Y. Yıldırım. Small gaps between almost primes, the parity problem, and some conjectures of Erdős on consecutive integers. *Int. Math. Res. Not. IMRN*, 2011(7):1439–1450, 2011.
[22] É. Goudout. Lois locales de la fonction ω dans presque tous les petits intervalles. *Proc. Lond. Math. Soc. (3)*, 115(3):599–637, 2017.
[23] A. Granville. Harald Cramér and the distribution of prime numbers. *Scand. Actuar. J.*, 1995(1):12–28, 1995. Harald Cramér Symposium (Stockholm, 1993).
[24] A. Granville and K. Soundararajan. Large character sums: pretentious characters and the Pólya-Vinogradov theorem. *J. Amer. Math. Soc.*, 20(2):357–384, 2007.
[25] B. Green. Roth's theorem in the primes. *Ann. of Math. (2)*, 161(3):1609–1636, 2005.
[26] B. Green and T. Tao. Restriction theory of the Selberg sieve, with applications. *J. Théor. Nombres Bordeaux*, 18(1):147–182, 2006.
[27] B. Green and T. Tao. The primes contain arbitrarily long arithmetic progressions. *Ann. of Math. (2)*, 167(2):481–547, 2008.
[28] B. Green and T. Tao. Linear equations in primes. *Ann. of Math. (2)*, 171(3):1753–1850, 2010.
[29] B. Green and T. Tao. The Möbius function is strongly orthogonal to nilsequences. *Ann. of Math. (2)*, 175(2):541–566, 2012.
[30] B. Green and T. Tao. The quantitative behaviour of polynomial orbits on nilmanifolds. *Ann. of Math. (2)*, 175(2):465–540, 2012.
[31] B. Green, T. Tao, and T. Ziegler. An inverse theorem for the Gowers $U^{s+1}[N]$-norm. *Ann. of Math. (2)*, 176(2):1231–1372, 2012.
[32] V. Z. Guo. Piatetski-Shapiro primes in a Beatty sequence. *J. Number Theory*, 156:317–330, 2015.
[33] G. Halász. Über die Mittelwerte multiplikativer zahlentheoretischer Funktionen. *Acta Math. Acad. Sci. Hungar.*, 19:365–403, 1968.
[34] G. H. Hardy and J. E. Littlewood. Some problems of 'Partitio numerorum'; III: On the expression of a number as a sum of primes. *Acta Math.*, 44(1):1–70, 1923.
[35] G. Harman. Almost-primes in short intervals. *Math. Ann.*, 258(1):107–112, 1981/82.
[36] G. Harman. Primes in short intervals. *Math. Z.*, 180(3):335–348, 1982.
[37] G. Harman. On the distribution of αp modulo one. *J. London Math. Soc. (2)*, 27(1):9–18, 1983.
[38] G. Harman. *Prime-detecting sieves*, volume 33 of *London Mathematical Society Monographs Series*. Princeton University Press, Princeton, NJ, 2007.

[39] D. R. Heath-Brown. Gaps between primes, and the pair correlation of zeros of the zeta function. *Acta Arith.*, 41(1):85–99, 1982.
[40] D. R. Heath-Brown. The number of primes in a short interval. *J. Reine Angew. Math.*, 389:22–63, 1988.
[41] D. R. Heath-Brown. Primes represented by $x^3 + 2y^3$. *Acta Math.*, 186(1):1–84, 2001.
[42] D. R. Heath-Brown and H. Iwaniec. On the difference between consecutive primes. *Invent. Math.*, 55(1):49–69, 1979.
[43] H. A. Helfgott. The ternary Goldbach problem. *ArXiv e-prints*, January 2015.
[44] A. Hildebrand. On a conjecture of Balog. *Proc. Amer. Math. Soc.*, 95(4):517–523, 1985.
[45] A. Hildebrand and G. Tenenbaum. Integers without large prime factors. *J. Théor. Nombres Bordeaux*, 5(2):411–484, 1993.
[46] G. Hoheisel. Primzahlprobleme in der Analysis. *Sitz. Preuss. Akad. Wiss.*, 33:580–588, 1930.
[47] M. N. Huxley. On the difference between consecutive primes. *Invent. Math.*, 15:164–170, 1972.
[48] A. Ivić. *The Riemann zeta-function: theory and applications*. Dover Publications, Inc., Mineola, NY, 2003.
[49] H. Iwaniec. Primes of the type $\phi(x, y) + A$ where ϕ is a quadratic form. *Acta Arith.*, 21:203–234, 1972.
[50] H. Iwaniec. The half dimensional sieve. *Acta Arith.*, 29(1):69–95, 1976.
[51] H. Iwaniec and M. Jutila. Primes in short intervals. *Ark. Mat.*, 17(1):167–176, 1979.
[52] H. Iwaniec and E. Kowalski. *Analytic number theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.
[53] H Iwaniec and J. Pintz. Primes in short intervals. *Monatsh. Math.*, 98(2):115–143, 1984.
[54] C. H. Jia. Difference between consecutive primes. *Sci. China Ser. A*, 38(10):1163–1186, 1995.
[55] C. H. Jia. Almost all short intervals containing prime numbers. *Acta Arith.*, 76(1):21–84, 1996.
[56] C. H. Jia. On the exceptional set of Goldbach numbers in a short interval. *Acta Arith.*, 77(3):207–287, 1996.
[57] C. H. Jia. On the distribution of αp modulo one. II. *Sci. China Ser. A*, 43(7):703–721, 2000.
[58] O. Klurman. Correlations of multiplicative functions and applications. *Compos. Math.*, 153(8):1622–1657, 2017.
[59] O. Klurman and A. P. Mangerel. Effective Asymptotic Formulae for Multilinear Averages of Multiplicative Functions. *ArXiv e-prints*, August 2017.
[60] O. Klurman and A. P. Mangerel. Rigidity Theorems for Multiplicative Functions. *ArXiv e-prints*, July 2017.

[61] A. N. Le. Nilsequences and Multiple Correlations along Subsequences. *ArXiv e-prints*, August 2017.

[62] A. Leibman. Nilsequences, null-sequences, and multiple correlation sequences. *Ergodic Theory Dynam. Systems*, 35(1):176–191, 2015.

[63] H. Z. Li. Goldbach numbers in short intervals. *Sci. China Ser. A*, 38(6):641–652, 1995.

[64] H. Z. Li. The exceptional set of Goldbach numbers. II. *Acta Arith.*, 92(1):71–88, 2000.

[65] Ju. V. Linnik. An asymptotic formula in an additive problem of Hardy-Littlewood. *Izv. Akad. Nauk SSSR Ser. Mat.*, 24:629–706, 1960.

[66] W. C. Lu. Exceptional set of Goldbach number. *J. Number Theory*, 130(10):2359–2392, 2010.

[67] H. Maier. Primes in short intervals. *Michigan Math. J.*, 32(2):221–225, 1985.

[68] A. P. Mangerel. On the Bivariate Erdős-Kac Theorem and Correlations of the Möbius Function. *ArXiv e-prints*, December 2016.

[69] K. Matomäki. Prime numbers of the form $p = m^2 + n^2 + 1$ in short intervals. *Acta Arith.*, 128(2):193–200, 2007.

[70] K. Matomäki. The binary Goldbach problem with one prime of the form $p = k^2 + l^2 + 1$. *J. Number Theory*, 128(5):1195–1210, 2008.

[71] K. Matomäki. A Bombieri-Vinogradov type exponential sum result with applications. *J. Number Theory*, 129(9):2214–2225, 2009.

[72] K. Matomäki. The distribution of αp modulo one. *Math. Proc. Cambridge Philos. Soc.*, 147(2):267–283, 2009.

[73] K. Matomäki, J. Maynard, and X. Shao. Vinogradov's theorem with almost equal summands. *Proc. Lond. Math. Soc. (3)*, 115(2):323–347, 2017.

[74] K. Matomäki and M. Radziwiłł. Multiplicative functions in short intervals. *Ann. of Math. (2)*, 183(3):1015–1056, 2016.

[75] K. Matomäki, M. Radziwiłł, and T. Tao. An averaged form of Chowla's conjecture. *Algebra Number Theory*, 9(9):2167–2196, 2015.

[76] K. Matomäki, M. Radziwiłł, and T. Tao. Sign patterns of the Liouville and Möbius functions. *Forum Math. Sigma*, 4:e14, 44, 2016.

[77] K. Matomäki and X. Shao. Vinogradov's three primes theorem with almost twin primes. *Compos. Math.*, 153(6):1220–1256, 2017.

[78] L. Matthiesen. Linear correlations of multiplicative functions. *ArXiv e-prints*, June 2016.

[79] J. Maynard. Primes represented by incomplete norm forms. *ArXiv e-prints*, July 2015.

[80] J. Maynard. Small gaps between primes. *Ann. of Math. (2)*, 181(1):383–413, 2015.

[81] H. Mikawa. Almost-primes in arithmetic progressions and short intervals. *Tsukuba J. Math.*, 13(2):387–401, 1989.

[82] H. L. Montgomery and R. C. Vaughan. The exceptional set in Goldbach's problem. *Acta Arith.*, 27:353–370, 1975. Collection of articles in memory of

Juriĭ Vladimirovič Linnik.
[83] Y. Motohashi. A note on almost-primes in short intervals. *Proc. Japan Acad. Ser. A Math. Sci.*, 55(6):225–226, 1979.
[84] A. Perelli and J. Pintz. On the exceptional set for the $2k$-twin primes problem. *Compositio Math.*, 82(3):355–372, 1992.
[85] A. Perelli and J. Pintz. On the exceptional set for Goldbach's problem in short intervals. *J. London Math. Soc. (2)*, 47(1):41–49, 1993.
[86] J. Pintz. On primes in short intervals. II. *Studia Sci. Math. Hungar.*, 19(1):89–96, 1984.
[87] O. Ramaré and I. Z. Ruzsa. Additive properties of dense subsets of sifted sequences. *J. Théor. Nombres Bordeaux*, 13(2):559–581, 2001.
[88] K. F. Roth. On certain sets of integers. *J. London Math. Soc.*, 28:104–109, 1953.
[89] P. Sarnak. Mobius randomness and dynamics. *Not. S. Afr. Math. Soc.*, 43(2):89–97, 2012.
[90] L. Schoenfeld. Sharper bounds for the Chebyshev functions $\theta(x)$ and $\psi(x)$. II. *Math. Comp.*, 30(134):337–360, 1976.
[91] A. Selberg. On the normal density of primes in small intervals, and the difference between consecutive primes. *Arch. Math. Naturvid.*, 47(6):87–105, 1943.
[92] X. Shao. A density version of the Vinogradov three primes theorem. *Duke Math. J.*, 163(3):489–512, 2014.
[93] S.-Y. Shi. On the distribution of αp modulo one for primes p of a special form. *Osaka J. Math.*, 49(4):993–1004, 2012.
[94] V. T. Sós. Turbulent years: Erdős in his correspondence with Turán from 1934 to 1940. In *Paul Erdős and his mathematics, I (Budapest, 1999)*, volume 11 of *Bolyai Soc. Math. Stud.*, pages 85–146. János Bolyai Math. Soc., Budapest, 2002.
[95] Y.-C. Sun and H. Pan. The Green-Tao theorem for primes of the form $x^2 + y^2 + 1$. *ArXiv e-prints*, August 2017.
[96] T. Tao. Special cases of Shannon entropy. `https://terrytao.wordpress.com/2017/03/01/special-cases-of-shannon-entropy/`.
[97] T. Tao. The Erdős discrepancy problem. *Discrete Anal.*, pages Paper No. 1, 29, 2016.
[98] T. Tao. The logarithmically averaged Chowla and Elliott conjectures for two-point correlations. *Forum Math. Pi*, 4:e8, 36, 2016.
[99] T. Tao. Equivalence of the logarithmically averaged Chowla and Sarnak conjectures. In *Number theory—Diophantine problems, uniform distribution and applications*. Springer, Cham, 2017.
[100] T. Tao and J. Teräväinen. The structure of logarithmically averaged correlations of multiplicative functions, with applications to the Chowla and Elliott conjectures. *ArXiv e-prints*, August 2017.
[101] T. Tao and V. Vu. *Additive combinatorics*, volume 105 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006.

[102] G. Tenenbaum. *Introduction to analytic and probabilistic number theory*, volume 163 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, third edition, 2015. Translated from the 2008 French edition by Patrick D. F. Ion.

[103] T. L. Todorova and D. I. Tolev. On the distribution of αp modulo one for primes p of a special form. *Math. Slovaca*, 60(6):771–786, 2010.

[104] D. I. Tolev. The binary Goldbach problem with arithmetic weights attached to one of the variables. *Acta Arith.*, 142(2):169–178, 2010.

[105] D. I. Tolev. The ternary Goldbach problem with arithmetic weights attached to two of the variables. *J. Number Theory*, 130(2):439–457, 2010.

[106] R. C. Vaughan. On the distribution of αp modulo 1. *Mathematika*, 24(2):135–141, 1977.

[107] R. C. Vaughan. *The Hardy-Littlewood method*, volume 125 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, second edition, 1997.

[108] I. M. Vinogradov. A new method in analytic number theory (Russian). *Tr. Mat. Inst. Steklova*, 10:5–122, 1937.

[109] I. M. Vinogradov. *The Method of Trigonometric Sums in the Theory of Numbers*. Translated from the Russian, revised and annotated by K. F. Roth and A. Davenport. Wiley-Interscience, 1954.

[110] Z. Wang. Sur les plus grands facteurs premiers d'entiers consécutifs. *Mathematika*, 64(2):343–379, 2018.

[111] N. Watt. Kloosterman sums and a mean value for Dirichlet polynomials. *J. Number Theory*, 53(1):179–210, 1995.

[112] N. Watt. Short intervals almost all containing primes. *Acta Arith.*, 72(2):131–167, 1995.

[113] E. Wirsing. Das asymptotische Verhalten von Summen über multiplikative Funktionen. II. *Acta Math. Acad. Sci. Hungar.*, 18:411–467, 1967.

[114] D. Wolke. Fast-Primzahlen in kurzen Intervallen. *Math. Ann.*, 244(3):233–242, 1979.

[115] J. Wu. Primes of the form $p = 1 + m^2 + n^2$ in short intervals. *Proc. Amer. Math. Soc.*, 126(1):1–8, 1998.

[116] C. Y. Yıldırım. A survey of results on primes in short intervals. In *Number theory and its applications (Ankara, 1996)*, volume 204 of *Lecture Notes in Pure and Appl. Math.*, pages 307–343. Dekker, New York, 1999.

[117] Y. Zhang. Bounded gaps between primes. *Ann. of Math. (2)*, 179(3):1121–1174, 2014.

Annales Universitatis Turkuensis

Turun yliopisto
University of Turku

Painosalama Oy, Turku, Finland 2018

ISBN 978-951-29-7339-2 (PRINT)
ISBN 978-951-29-7340-8 (PDF)
ISSN 0082-7002 (PRINT) | ISSN 2343-3175 (PDF)