Topics in sieve theory

Julia Stadlmann
Merton College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Trinity 2024

Für Daisy, Murli und Kassandra.

Acknowledgements

First and foremost, I would like to thank my supervisor, James Maynard, who was so tremendously helpful over the past 4 years. Thank you for many illuminating discussions, and for all the detailed feedback on my work. Thank you for answering so many of my questions, big and small, mathematical and administrative, and for your extraordinary patience.

I would also like to thank Günter Grayer, my Maths Olympiad teacher, who introduced me to the world of mathematics exactly half a life ago (= 13 years ago). Thank you for showing me the beauty of mathematics.

Further thanks go to the Merton tutors, Alex Scott, Minhyong Kim, Radek Erban and James Newton, who over my 8 years as undergraduate and graduate at Merton helped me in countless big and small ways. I am truly grateful for the wonderful and magical time I had at the college.

Beyond Merton, I am thankful to Ben Green for his very helpful advice, and to my fellow number theory students for many fun and interesting conversations. For making my time in Oxford memorable, I am thankful to Ella Caulfield and my friends at bell ringing and board games night.

I am also grateful to Promys Europe, which once upon a time changed the path of my life. Without Promys, I likely would have never set foot into this university. Most of all, I am thankful to the late Vicky Neale, whose dedication to her students I always greatly admired.

Finally, great thanks go to my family, my mother and father, my little sisters Sarah and Hannah, my grandparents, and my cousin Sophie.

> "Happiness can be found, even in the darkest of times, if one only remembers to turn on the light." — Albus Dumbledore

To my family: Thank you for allowing me to pursue my dreams. Thank you for countless long phone calls, for huge help with practical matters, and for your visits. Thank you also for sharing lots of fun and happiness.

Abstract

This thesis is concerned with the application of sieve theory to the study of gaps between primes. It focuses on techniques which combine sieves with methods from classical analytic number theory. We consider the following two questions: On average, how big are the squares of gaps between two consecutive primes less than x? How big is the smallest integer which appears infinitely often as a gap between m consecutive primes?

Concerning the first problem, we prove that the average size of the squares of differences between consecutive primes less than x is $O_\varepsilon(x^{0.23+\varepsilon})$ for any fixed $\varepsilon > 0$. This improves on a result of Peck [36], who gave bound $O_\varepsilon(x^{0.25+\varepsilon})$ in the place of $O_\varepsilon(x^{0.23+\varepsilon})$. Key ingredients of this work are Harman's sieve, Heath-Brown's mean value theorem for sparse Dirichlet polynomials and Heath-Brown's R^* bound.

Concerning the second problem, we prove that the primes below x are, on average, equidistributed in arithmetic progressions to smooth moduli of size up to $x^{1/2+1/40-\varepsilon}$. The exponent of distribution $\frac{1}{2}+\frac{1}{40}$ improves on a result of Polymath [37], who had previously obtained the exponent $\frac{1}{2}+\frac{7}{300}$. As a consequence, we improve results on intervals of bounded length which contain many primes, showing that

$$\liminf_{n\to\infty}(p_{n+m}-p_n) = O(\exp(3.8075m)).$$

The main new ingredient of our work is a modification of the q-van der Corput process. It allows us to exploit additional averaging for the exponential sums which appear in the Type I estimates of [37].

Contents

Chapter 1

Introduction

1.1 Sifting for primes

For millennia, mathematicians have been fascinated by prime numbers, by randomness and by surprising patterns in the distribution of primes. Which integers appear infinitely often as prime gaps? For primes of a given size, what is the maximal length of gaps between two consecutive primes, and how often do unusually long or short gaps occur? Primes are the subject of many famous conjectures. To name just a few, the twin prime conjecture suggests that 2 appears infinitely often as prime gap, and Legendre's conjecture suggests that primes below x are no more than $O(x^{1/2})$ apart. But despite mounting numerical evidence, the proofs of these old conjectures still seem impossibly far out of reach.

However, recent decades have seen the proof of several important partial results. Sieve methods have played a key role in many of these breakthroughs: both Chen's proof of an almost prime version of the Goldbach and twin prime conjectures [6] and Zhang's proof of bounded gaps between primes [45] made heavy use of tools from sieve theory.

The basic goal of sieve theory can be described as follows: Consider a set $\mathcal{A} \subseteq \mathbb{N}$. Suppose this set has a reasonably nice structure. That is, suppose we have good estimates for the number of elements of $\mathcal{A}$ which are divisible by d (when d is small). Then how many elements remain in $\mathcal{A}$ if we "sift out" integers with certain undesirable properties? For instance, can we obtain upper or lower bounds for the number of primes in $\mathcal{A}$? What about the number of twin primes in $\mathcal{A}$? If this problem is too difficult (as is often the case), can we at least get good estimates for the number of elements of $\mathcal{A}$ which have only few prime factors?

Examples of sets which may be targeted by sieve methods include short intervals $\mathcal{A} = [x, x + x^{\alpha}] \cap \mathbb{N}$ and values of polynomials $\mathcal{A} = \{f(m, n) : m, n \sim x\}$. The shorter the length of interval $[x, x+x^\alpha]$ or the higher the degree of polynomial $f(m,n)$, the harder it is to detect primes and almost primes in $\mathcal{A}$ – counting primes in sparse subsets of $\mathbb{N}$ is difficult. Success or failure also hinges on the extent of our understanding of the structure of set $\mathcal{A}$. Below we will now explain what kind of information about $\mathcal{A}$ is needed to make good use of sieves.

1.1.1 Sieve weights

Here we give an (informal) description of how to construct a basic sieve. Say we are given a set $\mathcal{A}$ and a property $P(n)$, and we simply want to prove that some element of $\mathcal{A}$ satisfies the property $P(n)$. That is, we want to show that $\sum_{n\in\mathcal{A}} 1_{P(n)} > 0$ (where $1_{P(n)}$ is the indicator function of property $P(n)$). However, it is often very difficult to estimate the sum $\sum_{n\in\mathcal{A}} 1_{P(n)}$ directly – for instance when $1_{P(n)}$ is the prime indicator function. Sieve weights now come to our aid.

First, we try to identify a collection of functions $f : \mathcal{A} \to \mathbb{R}$ for which good estimates for $\sum_{n\in\mathcal{A}} f(n)$ are available. Let us denote this collection by $\mathcal{C}$. If we are unable to give a good estimate for $\sum_{n\in\mathcal{A}} 1_{P(n)}$, we try to replace $1_{P(n)}$ by a linear combination of elements of $\mathcal{C}$. To be precise, the goal is to construct a weight function $w : \mathcal{A} \to \mathbb{R}$ with the following properties:

- (Property 0:) $w = \sum_{j\in J} c_j f_j$ for some $c_j \in \mathbb{R}$, $f_j \in \mathcal{C}$. Here $\sum_{n\in\mathcal{A}} f_j(n) \approx \alpha_j$.
- Property 1: $w(n)$ is a minorant for $1_{P(n)}$. That is, $1_{P(n)} \geq w(n)$ on $\mathcal{A}$.
- Property 2: $\sum_{n\in\mathcal{A}} w(n) \approx \sum_{j\in J} c_j\alpha_j > 0$.

Application: Suppose $w : \mathcal{A} \to \mathbb{R}$ with the above properties exist. By Property 1 and Property 2, $\sum_{n\in\mathcal{A}} 1_{P(n)} \geq \sum_{n\in\mathcal{A}} w(n) > 0$. This immediately implies that $1_{P(n)} = 1$ for some $n \in \mathcal{A}$. Then there exists an $n \in \mathcal{A}$ which satisfies $P(n)$, as desired. $\square$

Thus to prove $\{n \in \mathcal{A} : P(n) \text{ holds}\} \neq \emptyset$, the general strategy is to find elements of $\mathcal{C}$ and use these to build a function $w : \mathcal{A} \to \mathbb{R}$ with Properties 0, 1 and 2. The success (or failure) of this method depends crucially on the known elements of $\mathcal{C}$, which is the "information" known about summation over $\mathcal{A}$ (more on that below).

(In a sense, Property 0 is redundant: If $w = \sum_{j\in J} c_j f_j$ for some $f_j \in \mathcal{C}$, then we are able to give good estimates for $\sum_{n\in\mathcal{A}} c_j f_j(n)$ and hence also for $\sum_{n\in\mathcal{A}} w(n)$. But then w is itself an element of $\mathcal{C}$ and so there was no need to mention linear combinations. However, we included Property 0 to highlight the constructive nature of w: one first tries to identify many nice/simple functions f which can be summed accurately over $\mathcal{A}$, and one then uses these simple functions as building blocks to construct an approximation or minorant for $1_{P(n)}$.)

1.1.2 Decomposition of indicator functions

Intuitively, the bigger the collection of known elements of $\mathcal{C}$, the easier it should be to construct a function w with Properties 0, 1 and 2. However, when trying to identify elements of $\mathcal{C}$, it is certainly not helpful to aimlessly look at arbitrary functions. Instead, we should first think about which functions may be useful when trying to build an approximation or minorant for $1_{P(n)}$.

So far, our explanations have been quite abstract. We now remedy this by looking at a concrete choice of $P(n)$: Given some nice set $\mathcal{A}$, we want to show that $\mathcal{A}$ contains primes. That is, we want to approximate $\sum_{n\in\mathcal{A}} 1_{\mathbb{P}}(n)$. This sum is very hard to estimate directly, but luckily, there are several very useful decompositions of $1_{\mathbb{P}}(n)$ into more easily summable functions. Below we describe the two decompositions which are used extensively in this thesis.

1.1.2.1 Heath-Brown's identity

Heath-Brown's identity [18] (see Prop. 17.2 of [9]) states that for $K \in \mathbb{N}$ and $n \leq x$,

$$\Lambda(n) = \sum_{j=1}^{K} (-1)^{j+1} \binom{K}{j} \sum_{\substack{\prod_{i=1}^{2j} n_i = n \\ n_i \leq x^{1/K} \text{ for } i \leq j}} \mu(n_1)\dots\mu(n_j)\log(n_{2j}). \tag{1.1}$$

Approximating $1_{\mathbb{P}}(n)$ by $\frac{\Lambda(n)}{\log n}$ and substituting (1.1) for $\Lambda(n)$, we find that there exists a collection of convolutions $(\alpha_{1,\ell} \star \cdots \star \alpha_{2K,\ell})$, $\ell \in \mathbb{N}$, such that for $n \in [\frac{x}{2}, x)$

$$1_{\mathbb{P}}(n) \approx \sum_{\ell\in\mathbb{N}} c_\ell \, (\alpha_{1,\ell} \star \cdots \star \alpha_{2K,\ell})(n).$$

By dyadic decomposition, we see that the relevant convolutions $(\alpha_1 \star \cdots \star \alpha_{2K})$ can be chosen to all have the following nice properties: Every α_i is supported on some interval $[N_i, 2N_i]$ with $(N_i \leq x^{1/K}$ for $i \leq K)$ and $\prod_{i=1}^{2K} N_i \approx x$. Here $(\alpha_i \in \{\mu, 1\}$ on $[N_i, 2N_i]$ for $i \leq K)$ and $(\alpha_i \in \{1, \log\}$ on $[N_i, 2N_i]$ for $i > K)$.

In short, we can approximate $1_{\mathbb{P}}(n)$ by a linear combination of convolutions of functions which have small support or are slowly varying. This is very useful, as it is typically easier to estimate $\sum_{n\in\mathcal{A}}(\alpha \star \beta)(n)$ than $\sum_{n\in\mathcal{A}} 1_{\mathbb{P}}(n)$, particularly when the support of α is located in an interval of some convenient size (depending on $\mathcal{A}$) or when α is slowly varying and has large support. When searching for elements of $\mathcal{C}$, it is thus a good idea to focus on such convolutions.

1.1.2.2 The Buchstab identity

The Buchstab identity [13] is a simple, but important combinatorial identity. Writing

$$\psi(n,y) = \begin{cases} 1 & \text{if } p \mid n \to p \geq y, \\ 0 & \text{otherwise,} \end{cases}$$

the identity is given as follows: For $z > u$,

$$\psi(n,z) = \psi(n,u) \quad - \sum_{\substack{p\mid n \\ u \leq p < z}} \psi(\tfrac{n}{p}, p).$$

This statement is easily obtained from inclusion-exclusion. It provides us with a convenient way to rewrite $1_{\mathbb{P}}(n)$. For $n \in [\frac{x}{2}, x)$, we have $1_{\mathbb{P}}(n) = \psi(n, x^{1/2})$, and thus

$$1_{\mathbb{P}}(n) = \psi(n,u) \quad - \sum_{\substack{p_1\mid n \\ u \leq p_1 < x^{1/2}}} \psi(\tfrac{n}{p_1}, u) \quad + \sum_{\substack{p_1 p_2 \mid n \\ u \leq p_2 < p_1 < x^{1/2}}} \psi(\tfrac{n}{p_1 p_2}, p_2). \tag{1.2}$$

For nice sets $\mathcal{A}$, it is typically possible to give good estimates for sums $\sum_{n\in\mathcal{A}} \psi(n,u)$ and $\sum_{n\in\mathcal{A}} \sum_{p\mid n} \psi(\frac{n}{p_1}, u)$, provided that u is small compared to x (e.g. $u = x^{0.1}$) and provided that we have good estimates for $\sum_{n\in\mathcal{A}} (\alpha \star \beta)(n)$ when $\alpha = 1_{[N,2N]}$, $\mathrm{supp}(\beta) \subseteq [\frac{x}{N}, \frac{2x}{N}]$ and N is reasonably large (e.g. $N \in [x^{0.3}, x]$). This is a consequence of the fundamental lemma of sieve theory. Therefore, convolutions with slowly varying functions again play an important role: If $1_{[N,2N]} \star \beta \in \mathcal{C}$ for large N and β supported on $[\frac{x}{N}, \frac{2x}{N}]$, then $1_{\mathbb{P}}(n) \approx \sum_{j\in J} c_j f_j(n) + \sum \psi(\frac{n}{p_1 p_2}, p_2)$ for some $f_j \in \mathcal{C}$.

The final expression in (1.2), $\sum \psi(\frac{n}{p_1 p_2}, p_2)$, counts the number of choices of p_1 and p_2 for which we may write $n = p_1 p_2 (q_1 \dots q_r)$, where $p_1 > p_2 \geq u$ and $q_i \geq p_2$ for all i. The contribution of primes $p_1 \in [P_1, 2P_1]$, $p_2 \in [P_2, 2P_2]$ and $q_i \in [Q_i, 2Q_i]$ can be estimated using the following convolution:

$$\frac{1}{r!} \sum_{\substack{p_1 p_2 q_1 \dots q_r = n \\ p_i \in [P_i, 2P_i] \\ q_i \in [Q_i, 2Q_i]}} 1_{\mathbb{P}}(p_1) 1_{\mathbb{P}}(p_2) 1_{\mathbb{P}}(q_1) \dots 1_{\mathbb{P}}(q_r).$$

The problem of estimating $\sum_{n\in\mathcal{A}} 1_{\mathbb{P}}(n)$ has been reduced to working with sums $\sum_{n\in\mathcal{A}} (\alpha_1 \star \cdots \star \alpha_{r+2})(n)$, where $\alpha_i = 1_{\mathbb{P}}$ on $[N_i, 2N_i]$ and $\prod_{i=1}^{r+2} N_i \approx x$. Whether or not we can show $\alpha_1 \star \cdots \star \alpha_r \in \mathcal{C}$ generally depends on the values of N_i.

An important difference between the Buchstab identity and Heath-Brown's identity is the following: In our current construction, all functions α_i are non-negative. Therefore we may construct a minorant for $1_{\mathbb{P}}(n)$ by deleting the contribution of those P_i and

Q_i for which the corresponding convolution $(\alpha_1 \star \cdots \star \alpha_{r+2})(n)$ is not in $\mathcal{C}$. This way, we obtain a function w which satisfies Property 0 and Property 1 of Section 1.1.1. (However, if few elements of $\mathcal{C}$ are known, we must delete many p_i. The resulting minorant w may have $\sum_{n\in\mathcal{A}} w(n) \leq 0$ and will therefore not satisfy Property 2. It thus remains important to obtain a lot of information concerning summation over $\mathcal{A}$.)

The ideas described above are a fundamental ingredient of a sieve method known as Harman's sieve [14], which is used extensively in this thesis. More on that later.

1.1.3 Type I/II estimates (and beyond)

Section 1.1.2.1 and Section 1.1.2.2 highlighted that certain convolutions can be used to construct estimates and minorants for the prime indicator function. To detect primes in $\mathcal{A}$ – or to detect certain other prime–related properties $P(n)$ in $\mathcal{A}$ – we therefore seek estimates for $\sum_{n\in\mathcal{A}} f(n)$ when $f(n)$ is of the form

$$f(n) = \sum_{\substack{n=d_1\ldots d_k \\ d_i\in[D_i,2D_i]}} a^{(1)}_{d_1} \ldots a^{(k)}_{d_k}.$$

1.1.3.1 Useful terminology

The following terminology is commonly used to describe the available information:

Type I information (up to level D): Type I sums are of the form

$$\sum_{\substack{de\in\mathcal{A} \\ d<D}} a_d. \tag{1.3}$$

Here (a_d) are arbitrary coefficients, with only minor restrictions (e.g. $a_d = O(\tau(d))$). If good estimates for (1.3) are available for a given D and arbitrary coefficients (a_d), we have Type I information up to level D. Naturally, the bigger the size of D, the harder it is to obtain suitable estimates.

Type II information (in range $[A, B]$): Type II sums are of the form

$$\sum_{\substack{de\in\mathcal{A} \\ d\in[D,2D]}} a_d b_e. \tag{1.4}$$

Here both (a_d) and (b_e) are fairly arbitrary coefficients, say with $a_d = b_e = O(\tau(d))$. We say that we have Type II estimates on $[A, B]$ if good approximations to (1.4) are available for any choice of $D \in [A, B]$ and any coefficients (a_d) and (b_e).

Beyond Type II information: If no Type II estimates are available for some support $[D, 2D]$, one may instead consider further convolutions, such as

$$\sum_{\substack{def\in\mathcal{A}\\ d\in[D,2D]\\ e\in[E,2E]}} a_d b_e \qquad \text{or} \qquad \sum_{\substack{def\in\mathcal{A}\\ d\in[D,2D]\\ e\in[E,2E]}} a_d b_e c_f. \tag{1.5}$$

As more and more sequences are added to the convolution, it becomes increasingly complicated to decide which supports lead to suitable estimates, and the construction of associated weight functions w becomes more complicated. While inconvenient, this is sometimes the only way to obtain sufficient information to prove $\sum_{n\in\mathcal{A}} 1_{P(n)} > 0$. We will repeatedly use estimates beyond Type I and II in this thesis.

1.1.3.2 Significance of Type II information

In general, Type I information is much easier to obtain than Type II information, and classical sieve methods typically only make use of Type I information. Type I estimates have been employed very successfully in the search for products of few primes (so called almost primes) in well-understood sets $\mathcal{A}$. However, it is not possible to give non-trivial lower bounds for the number of primes in $\mathcal{A}$ if one solely uses Type I information. This is the parity barrier. To break the parity barrier and show that some set $\mathcal{A}$ contains primes, one must also obtain appropriate Type II information.

Obtaining Type II information can be very difficult, and techniques from many different areas of analytic number theory might make an appearance. Examples include bounds for exponential sums [37], bounds for Dirichlet polynomials and zeta functions [36], and lattice point counting arguments [31]. (The cited papers are provided as examples, rather than as an exhaustive list.)

The methods used to produce suitable Type II ranges, as well as the success or failure of these methods, hinge on the choice of set $\mathcal{A}$ and property $P(n)$. We will now discuss which sets $\mathcal{A}$ play a role in the proof of the two main results of this thesis, and how Type II estimates are obtained in these two cases.

1.2 Squares of prime gaps

To further our understanding of the primes, a common approach is to compare their occurrence with a probability distribution. Assuming the prime k-tuple conjecture, Gallagher [10] showed that the number of primes in short intervals follows a Poisson distribution, and so the lengths of prime gaps follow an exponential distribution. Considering moments of exponential distributions, we expect the following to be true:

$$M_k(x) = \sum_{p_n \le x} (p_{n+1} - p_n)^k \sim k! x (\log x)^{k-1}.$$

Empirical evidence can be found in [7]. However, currently a proof of the asymptotic $M_2(x) \sim 2x \log x$ seems impossibly far out of reach and even good upper bounds are hard to obtain. In 1979, Heath-Brown [16] proved that $M_2(x) = O_\varepsilon(x^{1.2\dot{7}+\varepsilon})$ and in 1996, Peck [36] obtained the bound $M_2(x) = O_\varepsilon(x^{1.25+\varepsilon})$. Improving on Heath-Brown's and Peck's results, in this thesis we show that

$$M_2(x) = \sum_{p_n \le x} (p_{n+1} - p_n)^2 = O_\varepsilon(x^{1.23+\varepsilon}).$$

1.2.1 Primes in short intervals

If the interval $[y, y + y^\alpha]$ contains no primes, then it contributes at least $y^{2\alpha}$ to the sum $\sum_{p_n \le x}(p_{n+1} - p_n)^2$. If α is large, prime free intervals of length y^α make a big contribution to $M_2(x)$. Thus in order to give a good upper bound on $M_2(x)$, we must show that only few intervals $[y, y + y^\alpha]$ contain no primes. To be precise, $M_2(x) = \sum_{p_n \le x}(p_{n+1} - p_n)^2 = O_\varepsilon(x^{1+\nu+\varepsilon})$ holds if the following statement is true:

Sufficient condition: The interval $[y, y + y^\alpha]$ contains primes for all but $O_\varepsilon(x^{1+\nu-\alpha+\varepsilon})$ choices of $y \in [x, 2x] \cap \mathbb{N}$. That is, for $\alpha \in [\nu, 1]$ and all but $O_\varepsilon(x^{1+\nu-\alpha+\varepsilon})$ values y,

$$\sum_{n \in [y, y+y^\alpha]} 1_{\mathbb{P}}(n) > 0.$$

This reduces the problem of bounding moments of prime gaps to detecting primes in short intervals $[y, y + y^\alpha]$. It is believed that for fixed $\alpha > 0$ and large y, the interval $[y, y + y^\alpha]$ always contains primes. However, even under assumption of the Riemann hypothesis, this problem is open for $\alpha \le 0.5$. In particular, Legendre's conjecture (the case $\alpha = 0.5$) remains open. Unconditionally, Baker, Harman and Pintz [1] proved that $[y, y + y^{0.525}]$ contains primes when y is large. We must still deal with the case $\alpha < 0.525$. Fortunately, we are allowed to have $O_\varepsilon(x^{1+\nu-\alpha+\varepsilon})$ bad choices of $y \in [x, 2x] \cap \mathbb{N}$, for which $[y, y + y^\alpha]$ does not contain primes.

In short, we seek Type II information for summation over $\mathcal{A}(y) = [y, y + y^\alpha]$. A special feature of our problem is that this Type II information is allowed to fail for $O_\varepsilon(x^{1+\nu-\alpha+\varepsilon})$ choices of $y \in [x, 2x] \cap \mathbb{N}$.

1.2.1.1 Obtaining Type I/II+ estimates

We try to estimate sums $\sum_{n\in[y,y+y^\alpha]} f(n)$, where $f(n)$ is of the form

$$f(n) = \sum_{\substack{n=d_1\dots d_k\\ d_i\in[D_i,2D_i]}} a^{(1)}_{d_1}\dots a^{(k)}_{d_k}.$$

Typically it is difficult to sum f over a short interval $[y, y+y^\alpha]$, but we expect the value of the sum to be proportional to the corresponding sum of f over the longer interval $[y, 2y]$. The latter is generally easier to evaluate. Luckily, sums over $[y, y+y^\alpha]$ can be related to sums over $[y, 2y]$ via use of Perron's theorem. We have

$$\sum_{n\in[y,y+y^\alpha]} f(n) \approx \frac{y^\alpha}{y}\sum_{n\in[y,2y]} f(n)$$

provided that $F(s) = \sum f(n)n^{-s} = \prod_{j=1}^k(\sum_{d\sim D_j} a^{(j)}_d d^{-s})$ satisfies

$$\left|\int_{y^{1-\alpha}}^{2y^{1-\alpha}} y^{it}F(1+it)\mathrm{d}t\right| = O_A\left(\frac{1}{(\log y)^A}\right). \tag{1.6}$$

Thus the problem of estimating $\sum_{n\in[y,y+y^\alpha]} f(n)$ can be reduced to bounding integrals of the corresponding Dirichlet polynomial $F(s) = \sum f(n)n^{-s}$.

To give good upper bounds on the LHS of (1.6), we use various Dirichlet polynomial bounds: Montgomery's mean value theorem, Huxley's large value estimate, Heath-Brown's R^* bound and Heath-Brown's mean value theorem for sparse Dirichlet polynomials all play important roles. In the case of Type I estimates, properties of the Riemann zeta function are also used.

The bound (1.6) does not need to hold for all choices of $y \in [x, 2x]$, $O_\varepsilon(x^{1+\nu-\alpha+\varepsilon})$ exceptions are allowed. This is important since both Heath-Brown's R^* bound and Heath-Brown's bound on sparse Dirichlet polynomials are designed to give upper bounds on the number of bad y. Observe that the allowed number of exceptions decreases as ν decreases. So if we aim for a smaller bound $M_2(x) = O_\varepsilon(x^{1+\nu+\varepsilon})$, (1.6) needs to hold for more y, which is harder to prove. For that reason, the amount of available Type II information decreases as ν decreases.

Which Dirichlet polynomial bounds work sufficiently well depends not just on the parameters $D_1, \dots, D_k$, but also on the interval length y^α. We combine bounds in many different ways, the computations are quite involved. The resulting Type I/II+ ranges are complicated and dependent on α, so we shall not state them in this introduction.

1.2.1.2 Improvements

Both Peck [36] and later Maynard [29] proved the bound $M_2(x) = O_\varepsilon(x^{1.25+\varepsilon})$. To obtain the bound $M_2(x) = O_\varepsilon(x^{1.23+\varepsilon})$, we improved on their work in two ways:

On the one hand, we made use of Heath-Brown's mean value theorem for sparse Dirichlet polynomials [20], which is a relatively new result and was not available to Peck and Maynard. It played a key role in improving estimates, and without this ingredient, our bound on $M_2(x)$ would have been considerably worse. However, this new Dirichlet polynomial bound on its own would have been insufficient to go below $x^{1.25}$ – there is a more fundamental, technical barrier at this point (see below).

Peck [36] and Maynard [29] both approached the problem of bounding squares of prime gaps via methods from classical analytic number theory. They used the Heath-Brown identity to approximate $1_{\mathbb{P}}(n)$ and subsequently needed to give good upper bounds on integrals of Dirichlet polynomials of many different lengths. The case $D_1 \approx D_2 \approx D_3 \approx D_4 \approx x^{1/4}$ (products of 4 Dirichlet polynomials of the same length) is the most difficult. It presented a barrier at $x^{1.25}$, which could not be removed via classical methods. By using sieve methods, and in particular the Buchstab identity and Harman's sieve, we are able to avoid some bad convolutions, such as the case $D_i \approx x^{1/4}$, and can go below $x^{1.25}$.

1.3 Interlude: Harman's sieve

We already briefly mentioned Harman's sieve [14] in Section 1.1.2.2, where we highlighted advantageous properties of the Buchstab identity. This sieve method is used extensively in this thesis, and plays a role in the proofs of both of our main results. So, as a brief interlude, we give an introduction to Harman's sieve via the example of squares of prime gaps.

1.3.1 Notation

For convenience's sake, we first introduce some standard notation from sieve theory: For $\mathcal{C} \subseteq \mathbb{N}$ and $z > 0$, we write

$$S(\mathcal{C}, z) = \#\{n \in \mathcal{C} : (n, p) = 1 \text{ for } p < z\}.$$

This is the number of elements of $\mathcal{C}$ not divisible by any prime less than z. In particular, the number of primes contained in the short interval $\mathcal{A}(y) = [y, y + y^\alpha]$ is

$S(\mathcal{A}(y),(2y)^{1/2})$. We write $\mathcal{C}_d = \{m \in \mathbb{N} : md \in \mathcal{C}\}$. The Buchstab identity reads

$$S(\mathcal{C},w) = S(\mathcal{C},z) - \sum_{z \le p < w} S(\mathcal{C}_p, p).$$

(If $n \in \mathcal{C}$ is counted by $S(\mathcal{C},z)$, but not by $S(\mathcal{C},w)$, then n has a prime factor p with $z \le p < w$ and all other prime factors of n are of size at least p.)

1.3.2 Buchstab iterations

The basic idea behind Harman's sieve is to compare the number of primes in some sets $\mathcal{A}$ and $\mathcal{B}$, where $\mathcal{A}$ is the set we are actually interested in and $\mathcal{B}$ is a related, but better understood set. (By better understood we mean that we know more about summation over $\mathcal{B}$.) To make such a comparison, the Buchstab identity is applied repeatedly. This allows us to instead compare the number of products of primes of a certain size in $\mathcal{A}$ and $\mathcal{B}$, which is often easier to accomplish.

We now carry out this process for $\mathcal{A}(y) = [y, y+y^\alpha]$, which is our set of interest, and for $\mathcal{B}(y) = [y, 2y]$, which is the related, better understood set.

$$\sum_{n \in \mathcal{A}(y)} 1_{\mathbb{P}}(n) - (y^{\alpha-1}) \sum_{n \in \mathcal{B}(y)} 1_{\mathbb{P}}(n) = S(\mathcal{A}(y), y^\delta) - (y^{\alpha-1}) S(\mathcal{B}(y), y^\delta) \tag{1.7}$$

$$- \sum_{y^\delta \le p_1 < (2y)^{0.5}} S(\mathcal{A}(y)_{p_1}, y^\delta) - (y^{\alpha-1}) S(\mathcal{B}(y)_{p_1}, y^\delta) \tag{1.8}$$

$$+ \sum_{y^\delta \le p_2 < p_1 < (2y)^{0.5}} S(\mathcal{A}(y)_{p_1 p_2}, p_2) - (y^{\alpha-1}) S(\mathcal{B}(y)_{p_1 p_2}, p_2). \tag{1.9}$$

In order to show that $\mathcal{A}(y)$ contains primes, we thus need to give suitable lower bounds for (1.7), (1.8) and (1.9).

1.3.3 Using Type I and II information

In Section 1.1.2.2 we observed that to evaluate $S(\mathcal{C},z) = \sum_{n \in \mathcal{C}} \psi(n,z)$, we must be able to give good estimates for $\sum_{n \in \mathcal{C}} f(n)$ whenever $f(n)$ is a convolution of some suitable functions. In the current set-up, we guess that $\sum_{n \in \mathcal{A}} f(n)$ can be approximated by $(\frac{\#\mathcal{A}}{\#\mathcal{B}}) \sum_{n \in \mathcal{B}} f(n)$. The goal is therefore to find many convolutions f for which the following holds true:

$$\sum_{n \in \mathcal{A}} f(n) = \left(\frac{\#\mathcal{A}}{\#\mathcal{B}}\right) \sum_{n \in \mathcal{B}} f(n) + \text{(small error)}.$$

In the case $\mathcal{A} = [y, y+y^\alpha]$ and $\mathcal{B} = [y, 2y]$, such a relationship is proved by bounding Dirichlet polynomials. We now give a list of relevant Dirichlet polynomials.

1.3.3.1 Corresponding Dirichlet polynomials

As mentioned in Section 1.2.1.1, we have $\sum_{n\in[y,y+y^\alpha]} f(n) \approx \frac{y^\alpha}{y}\sum_{n\in[y,2y]} f(n)$ provided that $F(s)=\sum f(n)n^{-s}$ satisfies

$$\left|\int_{y^{1-\alpha}}^{2y^{1-\alpha}} y^{it}F(1+it)\mathrm{d}t\right| = O_A\left(\frac{1}{(\log y)^A}\right). \tag{1.10}$$

Which convolutions and corresponding Dirichlet polynomials are of interest to us depends on whether we are looking at (1.7), (1.8) or (1.9). We split the summation ranges into dyadic intervals, so that $p_i\in[P_i,2P_i]$, and the corresponding Dirichlet polynomials also depend on P_i.

Dirichlet polynomials for (1.7): We must consider

$$F(s)=\prod_{j=1}^J S_j(s) \qquad \text{with} \qquad S_i(s)=\sum_{n\in[N_i,2N_i]}\frac{a_n^{(i)}}{n^s},$$

where $\prod_{i=1}^J N_i \approx x$ and ($N_j > x^\delta$ for at most one choice of j). Further, if there is such a long factor $S_j(s)$, we may assume it has all coefficients $a_n^{(j)}$ equal to 1. If (1.10) holds for all such choices of $F(s)$, then the RHS of (1.7) is $O(y^\alpha/(\log y)^2)$.

Dirichlet polynomials for (1.8) and $p_1\in[P_1,2P_1]$: We must consider

$$F(s)=\left(\sum_{n\in[P_1,2P_1]}\frac{1_{\mathbb{P}}(n)}{n^s}\right)\prod_{j=1}^J S_j(s) \qquad \text{with} \qquad S_i(s)=\sum_{n\in[N_i,2N_i]}\frac{a_n^{(i)}}{n^s},$$

where $\prod_{i=1}^J N_i \approx \frac{x}{P_1}$ and ($N_j > x^\delta$ for at most one choice of j), and the long factor has coefficients 1. (Additionally, if necessary, Heath-Brown's identity can be used to replace the Dirichlet polynomial $\sum_{n\in[P_1,2P_1]} 1_{\mathbb{P}}(n)n^{-s}$ by a product of Dirichlet polynomials which are short or have coefficients 1.) If (1.10) holds for all such choices of $F(s)$, then the contribution of $p_1\in[P_1,2P_1]$ to (1.8) is $O(y^\alpha/(\log y)^3)$.

Dirichlet polynomials for (1.9) and $p_1\in[P_1,2P_1]$ and $p_2\in[P_2,2P_2]$: We consider

$$F(s)=\left(\sum_{n\in[P_1,2P_1]}\frac{1_{\mathbb{P}}(n)}{n^s}\right)\left(\sum_{n\in[P_2,2P_2]}\frac{1_{\mathbb{P}}(n)}{n^s}\right)\prod_{j=1}^J S_j(s)$$

for $S_1(s),\dots,S_j(s)$ of total length $\frac{x}{P_1P_2}$. This time we cannot assume that all but one $S_j(s)$ are short – the lengths of the Dirichlet polynomials $S_j(s)$ are completely arbitrary. If (1.10) holds for all such choices of $F(s)$, then the contribution of $p_1\in[P_1,2P_1]$ and $p_2\in[P_2,2P_2]$ to (1.9) is $O(y^\alpha/(\log y)^4)$.

1.3.3.2 Good and bad products

We recall now the set-up of Section 1.2: Our aim is to prove $M_2(x) = O(x^{1.25-\delta})$ for some $\delta > 0$. To do so, we must show that $\mathcal{A}(y) = [y, y + y^\alpha]$ contains primes for all but $O(x^{1.25-\alpha-\delta})$ choices of $y \in [x, 2x] \cap \mathbb{N}$. Looking at the Buchstab iterations in Section 1.3.2 and at Section 1.3.3.1, we must prove that the bound (1.10) holds for suitable Dirichlet polynomials at all but at most $O(x^{1.25-\alpha-\delta})$ values of $y \in [x, 2x] \cap \mathbb{N}$.

Initially, we focus on very small δ. When $\delta = 0$ and $F(s)$ is a product of 4 Dirichlet polynomials of length $x^{0.25}$, (1.10) fails to hold too often. Increasing δ slightly, more problematic products appear. We need to avoid products $S_1(s)S_2(s)S_3(s)S_4(s)$ with $N_i \in [x^{0.25-\varepsilon(\delta)}, x^{0.25+\varepsilon(\delta)}]$. Here $\varepsilon(\delta)$ is an increasing function of δ with $\varepsilon(\delta) \to 0$ as $\delta \to 0$. For small δ, this is the only bad case – suitable bounds (1.10) are available for products $F(s) = \prod_{i=1}^J S_i(s)$, provided no factor $S_i(s)$ is too long (e.g. all factors are shorter than $x^{0.33}$) and provided $N_i \notin [x^{0.25-\varepsilon(\delta)}, x^{0.25+\varepsilon(\delta)}]$ for some $i \in \{1, \dots, J\}$.

Regarding Type I information, if $\delta > 0$ is very small and $F(s) = \prod_{i=1}^J S_i(s)$ has a factor $S_j(s)$ with ($N_j \geq x^{0.33}$ and all coefficients equal to 1), then (1.10) holds for all but at most $O(x^{1.25-\alpha-\delta})$ choices of $y \in [x, 2x] \cap \mathbb{N}$.

1.3.3.3 A minorant

Observe now that the Dirichlet polynomials corresponding to (1.7) and (1.8) all have many short factors or one fairly long factor with coefficients 1. In particular, we never run into the situation where we have four factors all of length close to $x^{0.25}$. For all but $O(x^{1.25-\alpha-\delta})$ choices of $y \in [x, 2x] \cap \mathbb{N}$, quantities (1.7) and (1.8) are $O(y^\alpha/(\log y)^2)$.

Looking at (1.9), we are in trouble when $P_1 \approx P_2 \approx x^{0.25}$, as there could be two factors $S_j(s)$ also of length $x^{0.25}$. We are also in trouble if $P_1 \approx x^{0.5}$ and $P_2 \approx x^{0.25}$: In order to produce Dirichlet polynomials of short length or with coefficients 1, Heath-Brown's identity must be applied to $\sum_{n \in [P_1, 2P_1]} 1_{\mathbb{P}}(n) n^{-s}$, which can produce two new factors of length $x^{0.25}$. All other choices of P_i correspond to suitable Type II estimates.

Comment: Following the exposition in Section 1.1, this information can be used to produce a good minorant $w(n)$ for $1_{\mathbb{P}}(n)$, which is given by

$$w(n) = 1_{\mathbb{P}}(n) \; - \sum_{\substack{p_1 p_2 | n \\ p_1 \in [x^{0.25-\varepsilon(\delta)}, x^{0.25+\varepsilon(\delta)}] \\ p_2 \in [x^{0.25-\varepsilon(\delta)}, x^{0.25+\varepsilon(\delta)}]}} \psi(\tfrac{n}{p_1 p_2}, p_2) \; - \sum_{\substack{p_1 p_2 | n \\ p_1 \in [x^{0.5-\varepsilon(\delta)}, x^{0.5+\varepsilon(\delta)}] \\ p_2 \in [x^{0.25-\varepsilon(\delta)}, x^{0.25+\varepsilon(\delta)}]}} \psi(\tfrac{n}{p_1 p_2}, p_2).$$

This choice of $w(n)$ satisfies Property 0 and Property 1. (Property 2, which is the criterion $\sum_{n \in \mathcal{A}} w(n) > 0$, must still be verified.)

However, instead of working directly with $w(n)$, we will continue to use the notation of Section 1.3.2, comparing $\mathcal{A}$ and $\mathcal{B}$. The difference between these two approaches is merely cosmetic, and we use both viewpoints in later chapters of this thesis.

1.3.4 Lower bounds

Using the Buchstab iterations of Section 1.3.2, and using our information concerning (1.7), (1.8) and (1.9), obtained in Section 1.3.3.3, we conclude that for all but at most $O(x^{1.25-\alpha-\delta})$ choices of $y \in [x, 2x] \cap \mathbb{N}$,

$$\sum_{n\in\mathcal{A}(y)} 1_{\mathbb{P}}(n) = (y^{\alpha-1}) \sum_{n\in\mathcal{B}(y)} 1_{\mathbb{P}}(n) + O\left(\frac{y^{\alpha}}{(\log y)^2}\right) \tag{1.11}$$

$$+ \sum_{\substack{y^{\delta}\leq p_2<p_1<(2y)^{0.5} \\ p_1\in[y^{0.25-\varepsilon(\delta)},y^{0.25+\varepsilon(\delta)}] \\ p_2\in[y^{0.25-\varepsilon(\delta)},y^{0.25+\varepsilon(\delta)}]}} S(\mathcal{A}(y)_{p_1p_2}, p_2) - (y^{\alpha-1})S(\mathcal{B}(y)_{p_1p_2}, p_2) \tag{1.12}$$

$$+ \sum_{\substack{y^{\delta}\leq p_2<p_1<(2y)^{0.5} \\ p_1\in[y^{0.5-\varepsilon(\delta)},y^{0.5+\varepsilon(\delta)}] \\ p_2\in[y^{0.25-\varepsilon(\delta)},y^{0.25+\varepsilon(\delta)}]}} S(\mathcal{A}(y)_{p_1p_2}, p_2) - (y^{\alpha-1})S(\mathcal{B}(y)_{p_1p_2}, p_2). \tag{1.13}$$

To give a lower bound on $\sum_{n\in\mathcal{A}(y)} 1_{\mathbb{P}}(n)$, we can simply delete the contribution of $\mathcal{A}(y)$ to (1.12) and (1.13), as this contribution is positive. On the other hand, $\mathcal{B}(y)$ is the long interval $[y, 2y]$. For that reason, it is easy to compute asymptotics for the sums in (1.11), (1.12) and (1.13) which involve $\mathcal{B}(y)$: By the prime number theorem, $\sum_{n\in\mathcal{B}(y)} 1_{\mathbb{P}}(n) = \frac{y(1+o(1))}{\log y}$. Using a generalized version of the PNT,

$$\sum_{\substack{p_1\in[y^{0.25-\varepsilon(\delta)},y^{0.25+\varepsilon(\delta)}] \\ p_2\in[y^{0.25-\varepsilon(\delta)},y^{0.25+\varepsilon(\delta)}]}} S(\mathcal{B}(y)_{p_1p_2}, p_2) \leq \frac{y(1+o(1))}{\log y} \int_{0.25-\varepsilon(\delta)}^{0.25+\varepsilon(\delta)} \int_{0.25-\varepsilon(\delta)}^{0.25+\varepsilon(\delta)} \frac{1}{\alpha_1\alpha_2^2}\, \mathrm{d}\alpha_1 \mathrm{d}\alpha_2,$$

$$\sum_{\substack{p_1\in[y^{0.5-\varepsilon(\delta)},y^{0.5+\varepsilon(\delta)}] \\ p_2\in[y^{0.25-\varepsilon(\delta)},y^{0.25+\varepsilon(\delta)}]}} S(\mathcal{B}(y)_{p_1p_2}, p_2) \leq \frac{y(1+o(1))}{\log y} \int_{0.5-\varepsilon(\delta)}^{0.5+\varepsilon(\delta)} \int_{0.25-\varepsilon(\delta)}^{0.25+\varepsilon(\delta)} \frac{1}{\alpha_1\alpha_2^2}\, \mathrm{d}\alpha_1 \mathrm{d}\alpha_2,$$

Since $\varepsilon(\delta) \to 0$ as $\delta \to 0$, we can find $\delta_0 > 0$ with

$$C(\delta_0) = 1 - \int_{0.25-\varepsilon(\delta_0)}^{0.25+\varepsilon(\delta_0)} \int_{0.25-\varepsilon(\delta_0)}^{0.25+\varepsilon(\delta_0)} \frac{\mathrm{d}\alpha_1 \mathrm{d}\alpha_2}{\alpha_1\alpha_2^2} - \int_{0.5-\varepsilon(\delta_0)}^{0.5+\varepsilon(\delta_0)} \int_{0.25-\varepsilon(\delta_0)}^{0.25+\varepsilon(\delta_0)} \frac{\mathrm{d}\alpha_1 \mathrm{d}\alpha_2}{\alpha_1\alpha_2^2} > 0.$$

Putting everything together, we have $\sum_{n\in\mathcal{A}(y)} 1_{\mathbb{P}}(n) \geq \frac{y^{\alpha}(1+o(1))}{\log y} C(\delta_0) > 0$ for all but at most $O(x^{1.25-\alpha-\delta_0})$ choices of $y \in [x, 2x] \cap \mathbb{N}$. Then $M_2(x) = O(x^{1.25-\delta_0})$. $\square$

1.3.5 Doing better

As shown above, Harman's sieve can be used to easily improve the bound $O(x^{1.25+\varepsilon})$ to $O(x^{1.25-\delta_0})$ for some very small $\delta_0 > 0$. However, we want to do much better than that, proving a bound $O_\varepsilon(x^{1.23+\varepsilon})$. This is much less straightforward:

As δ increases, many other bad products of Dirichlet polynomials appear. No longer are products of Dirichlet polynomials of equal length the only concern - conditions which identify bad cases become much more complicated and depend on the interval length y^α. Using just two Buchstab decompositions is no longer enough, we need to apply the Buchstab identity 4 or even 6 times, and compute the loss much more carefully. Since the available Type I/II+ information depends on α, we need to apply Harman's sieve not just once, but multiple times, considering different ranges of α.

In short, it is possible to prove $M_2(x) = O_\varepsilon(x^{1.23+\varepsilon})$, but will require much more work.

1.4 Bounded gaps between primes

One of the biggest open problems in number theory is the Hardy-Littlewood prime k-tuple conjecture. For a given k-tuple $(h_1, \dots, h_k)$, this conjecture predicts the asymptotic density of the natural numbers n for which all entries of $(n+h_1, \dots, n+h_k)$ are prime. A famous special case of the Hardy-Littlewood k-tuple conjecture is the twin prime conjecture:

$$\#\{n \le x : n \text{ and } n+2 \text{ prime}\} \sim 2 \prod_{p \ge 3} \left(1 - \frac{1}{(p-1)^2}\right) \int_2^x \frac{\mathrm{d}t}{(\log t)^2}.$$

A resolution of the prime k-tuple conjecture would have far reaching consequences, giving us very precise information about the distribution of prime numbers. For instance, the k-tuple conjecture tells us precisely which integers are expected to appear infinitely often as the difference between $m+1$ primes. Writing

$$H_m = \liminf_{n \to \infty} (p_{n+m} - p_n),$$

it has been conjectured that $H_1 = 2$, $H_2 = 6$, $H_3 = 8$, ..., and so on. Unfortunately, a resolution of these conjectures appears far out of reach of modern techniques. Only 10 years ago, no upper bounds on H_m were known at all. In 2013, Zhang [45] achieved a big breakthrough when he showed that $H_1 \le 70000000$. This was followed by work of Maynard [30] and Tao (unpublished), who proved that $H_m < \infty$ also holds for $m \ge 2$. Prior to our work, the best known upper bounds on H_m for $m \in \{1, \dots, 5\}$

were results of the Polymath project [38], while Baker and Irving [2] gave the bound $H_m = O(\exp(3.815m))$, best for large m. Improving on the results of Baker and Irving, in this thesis we show that

$$H_m = \liminf_{n\to\infty} = O(\exp(3.8075m)).$$

Improving on the results of Polymath, we also show $H_2 \leq 396504$, $H_3 \leq 24407016$, $H_4 \leq 1391051532$ and $H_5 \leq 77510685234$.

1.4.1 The GPY sieve

To detect integers n for which a tuple $(n+h_1, \dots, n+h_k)$ contains multiple primes, Goldston, Pintz and Yıldırım developed the so called GPY method [11]. The basic idea is as follows: Pick a tuple $\underline{h} = (h_1, \dots, h_k)$ with $h_1 < \dots < h_k$ and

$$\forall p \in \mathbb{P} : \{h_i \bmod p : 1 \leq i \leq k\} \neq \mathbb{Z}/p\mathbb{Z}.$$

Such tuples are called admissible. Choose some $\nu : \mathbb{N} \to [0, \infty)$ and define

$$N(x; \underline{h}, \nu) = \sum_{x \leq n \leq 2x} \nu(n) \left(\sum_{i=1}^{k} 1_{\mathbb{P}}(n+h_i) - m \right).$$

<u>Application</u>: Suppose that for some admissible tuple $\underline{h} = (h_1, \dots, h_k)$ and some weight function ν, we can show that $N(x; \underline{h}, \nu) > 0$ for all large x. Then if x is large, there exists an integer $n \in [x, 2x]$ with $(\sum_{i=1}^{k} 1_{\mathbb{P}}(n+h_i) - m) > 0$, so that $(n+h_1, \dots, n+h_k)$ contains at least $m+1$ primes. As x was arbitrary, this implies

$$H_m = \liminf_{n\to\infty} (p_{n+m} - p_n) \leq h_k - h_1.$$

1.4.1.1 A second sieve problem

The problem of bounding H_m has been reduced to the search for a suitable weight function $\nu(n)$. This is still a very difficult problem. In 2005, Goldston, Pintz and Yıldırım [11] used their GPY sieve to show $\liminf_{n\to\infty} \frac{(p_{n+1}-p_n)}{\log p_n} = 0$, and it took almost another decade until Zhang [45] managed to show $\liminf_{n\to\infty}(p_{n+1} - p_n) < \infty$.

It turns out that in order to construct suitable $\nu(n)$, one must have a good understanding of the behaviour of primes in arithmetic progressions. This, too, can be viewed as a sieve problem. Not the GPY sieve itself, but this secondary sieve problem is a focus of this thesis. We study the behaviour of primes in $\mathcal{A}(q) = \{n \in [1, x] : n \equiv a(q)\}$, where q is a product of primes less than x^δ. Such common differences q are called x^δ-smooth moduli. They played a key role in Zhang's proof [45].

1.4.2 Primes in arithmetic progressions

A classical result of Dirichlet states that the arithmetic progression $(a, a+q, a+2q, \dots)$ contains infinitely many primes when $(a, q) = 1$. More precisely, it is known that

$$\#\{p \le x : p \equiv a(q)\} \sim \frac{\#\{p \le x\}}{\phi(q)} \qquad \text{as} \qquad q \to \infty.$$

Asymptotically, the primes equidistribute over the residue classes $(\mathbb{Z}/q\mathbb{Z})^*$. To construct good weights $\nu(n)$ for an application of the GPY sieve, one needs strong upper bounds on the corresponding error terms $|\#\{p \le x : p \equiv a(q)\} - \frac{1}{\phi(q)}\#\{p \le x\}|$. To be precise, we need to prove the following statement for some $\theta > \frac{1}{2}$:

<u>Equidistribution estimate for smooth moduli</u>: There exists $\delta > 0$ with

$$\sum_{\substack{q \le x^\theta \\ q | P(x^\delta) \\ (q,a)=1}} \left| \sum_{\substack{n \le x \\ n \equiv a(q)}} 1_{\mathbb{P}}(n) - \frac{1}{\phi(q)} \sum_{n \le x} 1_{\mathbb{P}}(n) \right| \ll_A \frac{x}{(\log x)^A}. \tag{1.14}$$

Here $P(x^\delta) = \prod_{p < x^\delta} p$. The criterion $q \mid P(x^\delta)$ simply ensures that q is x^δ-smooth and squarefree. This restriction is key to allowing us to take $\theta > \frac{1}{2}$.

Zhang [45] proved (1.14) with $\theta = \frac{1}{2} + \frac{1}{584} - \varepsilon$. (Here $\varepsilon > 0$ is arbitrary, but δ and the implied constant depend on ε.) Polymath [37] obtained $\theta = \frac{1}{2} + \frac{7}{300} - \varepsilon$. We succeed in proving the equidistribution estimate (1.14) with $\theta = \frac{1}{2} + \frac{1}{40} - \varepsilon$. Larger values of θ allow for shorter admissible tuples $\underline{h}$ in the GPY sieve, and thus correspond to stronger bounds on H_m.

1.4.2.1 Type II information

In Section 1.1 we discussed at length how Type I and II information may be used to detect primes in some nice set $\mathcal{A}$. Here, the sieve problem takes a more complicated shape: It is already well known that $\mathcal{A}(q) = \{n \le x : n \equiv a(q)\}$ contains primes – but we desire much more accurate information about the number of primes in $\mathcal{A}(q)$. Nonetheless, the concept of Type I/II+ estimates is still extremely useful.

In this set-up, Type II estimates look as follows:

$$\sum_{\substack{q \le x^\theta \\ q | P(x^\delta) \\ (q,a)=1}} \left| \sum_{n \equiv a(q)} (\alpha \star \beta)(n) - \frac{1}{\phi(q)} \sum_{(n,q)=1} (\alpha \star \beta)(n) \right| \ll \frac{x}{(\log x)^A}, \tag{1.15}$$

where $MN \asymp x$ and α and β are supported on $[N, 2N]$ and $[M, 2M]$, respectively.

The Type II estimates (1.15) will be applied to convolutions arising from an application of Heath-Brown's identity. We must establish (1.15) for all $N \in [x^{0.4}, x^{0.6}]$, as otherwise convolutions of 5 sequences of equal support cause problems. Polymath showed (1.15) for $N \in [x^{0.4}, x^{0.6}]$ and $\theta = \frac{1}{2} + \frac{7}{300} - \varepsilon$, we instead take $\theta = \frac{1}{2} + \frac{1}{40} - \varepsilon$.

Here is not the place to give a detailed description of how Type II estimates (1.15) are obtained - the process is quite technical. Linnik's dispersion method and completion of sums are used to reduce the problem to bounding exponential sums. To bound exponential sums, various techniques are used: the q-van der Corput method, Weil's bound for Kloosterman sums and Deligne's work on the Riemann hypothesis over finite fields all make an appearance in the Polymath paper [37]. To finish the introduction of this thesis, we just very briefly indicate how we improve on the exponential sum estimates of Polymath.

1.4.2.2 Exponential sums

To prove Type II estimates (1.15) for $\theta = \frac{1}{2} + 2\omega$ and $N \in [x^{0.4}, x^{0.6}]$, we must show the following: For some D with $1 \ll D \ll N$ and for $Y \approx x^{8\omega}$, $m \approx x^{1+8\omega}/(ND)$ and B with $(B, md) = 1$,

$$\sum_{d \asymp D} \left| \sum_{n \asymp N} \sum_{\substack{y \asymp Y \\ (y,m)=1}} e_{md}\left(\frac{By}{n + Cd}\right) \right| \ll \frac{DN}{x^{\epsilon}}. \tag{1.16}$$

To bound the exponential sum given in (1.16), Polymath fixed y and applied the so called q-van der Corput method to d and n, partitioning according to the value of n mod d. We use the special form of the exponential phases in (1.16) to introduce a variant of the q-van der Corput method, which is based on partitions according to the value of y/n mod d. In Polymath's work, the sum over y was bounded trivially. By incorporating y into our computations, we get smaller diagonal terms, additional averaging and better results. This idea is described more carefully in Section 3.1.1.

Comment: To conclude the introduction, we remark that Chapter 2 is based on the arXiv preprint [41] and Chapter 3 is based on the arXiv preprint [42].

Chapter 2

The mean square gap between primes

2.1 Introduction

Let $p_1, p_2, \ldots$ denote the sequence of primes and write $\pi(x) = \#\{n : p_n \leq x\}$. We study the average size of the squares of gaps between consecutive primes less than x, given by $\frac{1}{\pi(x)} \sum_{p_n \leq x} (p_{n+1} - p_n)^2$. Under assumption of the Riemann hypothesis, Selberg [39] first showed that

$$\sum_{p_n \leq x} (p_{n+1} - p_n)^2 \ll x(\log x)^3.$$

Assuming the Lindelöf hypothesis, Yu [44] proved that $\sum_{p_n \leq x} (p_{n+1} - p_n)^2 = O_\varepsilon(x^{1+\varepsilon})$ for every $\varepsilon > 0$, implying that the average size of $(p_{n+1} - p_n)^2$ with $p_n \leq x$ is $O_\varepsilon(x^\varepsilon)$.

A first unconditional result was given by Heath-Brown in [15]. He showed that

$$\sum_{p_n \leq x} (p_{n+1} - p_n)^2 \ll_\varepsilon x^{1+\nu+\varepsilon} \tag{2.1}$$

holds with $\nu = 1/3 = 0.\dot{3}$. In [16] he improved this bound to $\nu = 5/18 = 0.2\dot{7}$. The problem was further studied by Peck [36] and later Maynard [29], who both obtained (2.1) with $\nu = 1/4 = 0.25$. New phenomena occur when $\nu < 0.25$ and handling these issues is the key technical innovation of our work. We prove the following result:

Theorem 2.1. *For any $\varepsilon > 0$, we have*

$$\sum_{p_n \leq x} (p_{n+1} - p_n)^2 \ll_\varepsilon x^{1.23+\varepsilon}.$$

At the core of our proof is a parametric version of Harman's sieve [14]. To obtain the information required for the application of this sieve, we rely on a complicated combination of bounds on large values of Dirichlet polynomials, in particular Heath-Brown's R^* bound [29] and Heath-Brown's sparse mean value theorem [21].

2.1.1 Key ideas

First we recall the argument of Peck. Given some value of τ and a prime gap $p_{n+1}-p_n$ with $p_{n+1}-p_n > 6x/\tau$ and $p_n \leq 2x$, we note that $y \in (p_n, (p_n+p_{n+1})/2)$ implies that the interval $[y, y+y/\tau]$ contains no primes. This differs significantly from the expected number of primes in an interval of length y/τ and, by Perron's formula, roughly implies that

$$\left|\int_{\tau}^{2\tau} y^{it} \sum_{p<2x} p^{-it}\,\mathrm{d}t\right|$$

is unusually large. We can show this happens rarely by combining L^∞, L^2 and L^4 bounds over $y \in [x, 2x]$. To get adequate bounds we apply a combinatorial decomposition to the primes and reduce the situation to obtaining suitable mean value estimates for products of Dirichlet polynomials. Doing this for all relevant values of τ gives result (2.1) with $\nu = 0.25$.

Unfortunately, this argument breaks down when $\nu < 1/4$, as there is a discontinuity in the estimates of Peck [36] and Maynard [29]. By refining the underlying estimates, most of the problematic terms can be handled, but products of 4 Dirichlet polynomials of length roughly $x^{1/4}$ cannot be treated using existing mean value theorems. (Limitations correspond to the 3/4–line, just like in many similar problems.) To handle this issue, we instead count primes with a minorant chosen via Harman's sieve, allowing us to exclude some small number of bad terms. In principle, this new argument should then be sufficient to give (2.1) with $\nu = 1/4 - \delta$ for some very small $\delta > 0$.

However, as ν decreases further away from 1/4, we quickly encounter many new bad factor lengths, which cannot be treated using the approaches of Peck [36] and Maynard [29]. (Here issues are also no longer only clustered around the 3/4–line, there are many bad Dirichlet polynomial sizes.) To handle many of these new bad terms, we introduce a new estimate based on a mean value theorem of Heath-Brown for sparse Dirichlet polynomials, which combines well with the previous techniques, particularly Heath-Brown's R^* bound. Which terms still cause issues depends on the parameter τ and so we require a parametric version of Harman's sieve. In this parametric version of the sieve, the retained terms, which must have no bad sizes, must depend on τ in a suitably concrete way that we can verify certain integral computations for all relevant values of τ and thus can ensure that the minorant constructed produces primes. It is handling these difficulties where our key new ideas appear.

2.1.2 Proof outline

To prove Theorem 2.1, we must show that the interval $[y, y + y/\tau]$ with $\tau \leq x^{0.77}$ contains primes for most choices of $y \in [x, 3x] \cap \mathbb{N}$ – no more than $O(\tau x^{0.23+\varepsilon})$ exceptions are allowed. When τ is large, $[y, y + y/\tau]$ is a short interval and we lack good estimates for $\pi(y + y/\tau) - \pi(y)$, the number of primes it contains. However, we guess that for most choices of y, the number of primes in short interval $[y, y + y/\tau]$ is roughly proportional to that in the longer interval $[y, y + y/x^b]$, where b is a small constant. This larger interval is known to contain $(1 + o(1))\frac{y}{x^b \log x}$ primes. Hence if y satisfies

$$(\pi(y + y/\tau) - \pi(y)) - \left(\frac{x^b}{\tau}\right)(\pi(y + y/x^b) - \pi(y)) > -\frac{\alpha y}{\tau \log x} \tag{2.2}$$

for some fixed $\alpha \in (0, 1)$, then $[y, y + y/\tau]$ contains primes. To show that such an inequality indeed holds for all but $O(\tau x^{0.23+\varepsilon})$ choices of y, we use the Buchstab identity, an inclusion–exclusion argument, which lets us compare the number of primes in $[y, y + y/\tau]$ and $[y, y + y/x^b]$ by comparing the number of multiples of products of primes of certain sizes. These multiples sometimes provide additional structure in computations, making them easier to count than the primes itself. This process of comparison of multiples of certain primes in a short and in a long interval is known as Harman's sieve. Our application of Harman's sieve and proof of Theorem 2.1 are split into three propositions and one lemma, each corresponding to one section of this chapter. The precise formulation of these results can be found in Section 2.2, but we first give a rough outline of the main arguments. For clarity of exposition, details are suppressed.

In Section 2.3 we compare certain sums over $[y, y+y/\tau]$ and $[y, y+y/x^b]$. We consider

$$\sum_{\substack{k_1 \dots k_r \in [y, y+y/\tau] \\ k_i \in (N_i, 2N_i]}} \left(\prod_{1 \leq i \leq r} a_{k_i}^{(i)}\right) - \frac{x^b}{\tau} \sum_{\substack{k_1 \dots k_r \in [y, y+y/x^b] \\ k_i \in (N_i, 2N_i]}} \left(\prod_{1 \leq i \leq r} a_{k_i}^{(i)}\right), \tag{2.3}$$

where $a_n^{(i)} = 1_{\mathbb{P}}(n)$ (the prime indicator function) or $a_n^{(i)} = 1$ whenever N_i is large. We also consider corresponding Dirichlet polynomials $F(s) = \prod_{i=1}^{j} S_i(s)$ which look roughly like $\prod_{i=1}^r \sum_{k_i\in(N_i,2N_i]} a_{k_i}^{(i)} k_i^{-s}$, except factors of the form $\sum_{k_i\in(N_i,2N_i]} 1_{\mathbb{P}}(k_i)k_i^{-s}$ are replaced by their combinatorial decomposition into smooth or short factors via Heath–Brown's identity. By Perron's formula, (2.3) is $O_A(\frac{x}{\tau(\log x)^A})$ unless there is some corresponding Dirichlet polynomial $F(s) = \prod_{i=1}^j S_i(s)$ for which the integral

$$\frac{1}{\tau}\left|\int_{x^b}^{\tau} y^{1+it}F(1+it)\,\mathrm{d}t\right| \tag{2.4}$$

fails to be bounded by $\frac{x}{\tau(\log x)^A}$. Denoting by B_i the length of Dirichlet polynomial $S_i(s)$, we split the domain of integration of (2.4) into $O((\log x)^{j+1})$ regions, considering separately the contribution of $\bigcup_{n\in\mathcal{S}_1}[n,n+1]$ where $\mathcal{S}_1$ is the set of $n\in[T,2T]\cap\mathbb{N}\subseteq[x^b,\tau]$ for which each factor $S_i(1+it)$ is of size between $B_i^{\sigma_i-1}$ and $2B_i^{\sigma_i-1}$ on $[n,n+1]$ and for which $F(1+it)$ has a total size of about $x^{\sigma-1}$ on $[n,n+1]$. If $\#\mathcal{S}_1 < x^{1-\sigma-\varepsilon}$, this region contributes only little to (2.4) for every y and can be ignored. If instead $\#\mathcal{S}_1 \geq x^{1-\sigma-\varepsilon}$, we wish to show that the contribution of $\bigcup_{n\in\mathcal{S}_1}[n,n+1]$ to (2.4) exceeds $\frac{x}{\tau(\log x)^{A+j+1}}$ for only $O(\tau x^{0.23+\varepsilon})$ values $y\in[x,3x]\cap\mathbb{N}$. Taking a second moment over $y\in[x,3x]$,

$$\int_x^{3x}\left|\sum_{n\in\mathcal{S}_1}\int_n^{n+1} y^{1+it}F(1+it)\,\mathrm{d}t\right|^2\mathrm{d}y,$$

we get the desired bound on the number of bad y whenever $\#\mathcal{S}_1 < \tau x^{1-2\sigma+0.23}$. We denote $\#\mathcal{S}_1$ by $R(F,(S_i),(\sigma_i),T)$. Alternatively, by taking a fourth moment, we find that we also have a suitable bound if the number of quadruples $n_1,\dots,n_4\in\mathcal{S}_1$ with $n_1+n_2=n_3+n_4$ is bounded above by $\tau x^{3-4\sigma+0.23}$. We denote this number of quadruples by $R^*(F,(S_i),(\sigma_i),T)$. For a third approach, we denote by $\mathcal{J}$ the set of $m\in[1,\tau]\cap\mathbb{N}$ for which the contribution of $\bigcup_{n\in\mathcal{S}_1}[n,n+1]$ to (2.4) with $y=\lceil 3x/\tau\rceil m$ exceeds $\frac{x}{\tau(\log x)^{A+j+1}}$. Effectively, we sample bad y, leaving large gaps to ensure that $\#\mathcal{J}\leq\tau$. Multiplying by some $\xi_m\in\{z\in\mathbb{C}:|z|=1\}$ to remove absolute value signs and summing over $\mathcal{J}$, we find

$$\#\mathcal{J}\leq(\log x)^{A+j+2}\left|\sum_{n\in\mathcal{S}_1}\int_{n+1}^{n}\left(\sum_{m\in\mathcal{J}}\xi_m m^{it}\right)F(1+it)\,\mathrm{d}t\right|.$$

We have introduced a second Dirichlet polynomial $G(s)=\sum_{m\in\mathcal{J}}\overline{\xi_m}m^{-s}$. If the integral on the RHS is small, there are only few bad y. We split up the sum, restricting our attention to $n\in\mathcal{S}_2$, where $\mathcal{S}_2$ is the set of $n\in[T,2T]$ for which the size of $G(it)$ is contained in a prescribed dyadic interval $(L^\gamma,2L^\gamma]$. We find that the desired bound $\#\mathcal{J}=O(\tau^2x^{0.23-1})$ holds if $\#(\mathcal{S}_1\cap\mathcal{S}_2)\leq(\#\mathcal{J})^{7/8-\gamma}\tau^{1/4}x^{7/8-\sigma+0.23/8}$ or

$\#(\mathcal{S}_1 \cap \mathcal{S}_2) \leq (\#\mathcal{J})^{1-\gamma} x^{1-\sigma}$. We write $\#(\mathcal{S}_1 \cap \mathcal{S}_2) = Q(F, G, (S_i), (\sigma_i), \gamma, T)$ and have reduced the problem of obtaining a good bound of $O(\frac{x}{\tau(\log x)^A})$ on (2.2) to showing that $R(F, (S_i), (\sigma_i), T)$, $R^*(F, (S_i), (\sigma_i), T)$ and $Q(F, G, (S_i), (\sigma_i), \gamma, T)$ satisfy one of five inequality conditions for every choice of (σ_i), γ and T. The formal statement of this result is Proposition 2.1, which is proved in Section 2.3.

In Section 2.4 we then give various bounds for R, R^* and Q. Heath-Brown's bound on sparse Dirichlet polynomials plays a key role: it is used to replace the condition $Q \leq \max\{(\#\mathcal{J})^{7/8-\gamma} \tau^{1/4} x^{7/8-\sigma+0.23/8}, (\#\mathcal{J})^{1-\gamma} x^{1-\sigma}\}$ by an alternative condition $R \leq \min\{\tau^{-1/4} M^{2\beta-1} x^{7/4-2\sigma+\nu/4}, M^{2\beta-2} x^{2-2\sigma}\}$ where $M = \prod_{i\in I} B_i^{m_i}$ and $M^\beta = \prod_{i \in I} B_i^{m_i \sigma_i}$ for some $I \subseteq \{1, \dots, j\}$ and $m_i \in \mathbb{N}$. We also use Heath-Brown's R^* bound to replace condition $R^* \leq \tau x^{3-4\sigma+0.23}$ by an alternative condition $R \leq \min_{(U,V,W,X,Y,Z)\in\mathcal{Z}} \tau^U M^{V\beta - W} x^{X - Y\sigma + 0.23Z}$, where $\mathcal{Z}$ is a set of nine tuples (U, V, W, X, Y, Z) specified in Section 2.4. This reduces the problem of bounding (2.3) to obtaining good upper bounds on $R(F, (S_i), (\sigma_i), T)$, the number of intervals $[n, n+1]$ on which Dirichlet polynomials $S_i(s)$ are of size about $B_i^{\sigma_i - 1}$. We then use results such as Montgomery's mean value estimate and Huxley's large values estimate to obtain such bounds. These are very sensitive to the lengths B_i of Dirichlet polynomials $S_i(s)$. We show that we get adequate bounds on R provided that some combination of factors $S_i(s)$ is of a convenient length (depending on τ). The formal statement of this result is Proposition 2.2, proved in Section 2.4.

In Section 2.5 we use our bounds on the differences (2.3) to prove that if $P_1, \dots, P_r$ are in certain good ranges, then there exists $\mathcal{I} \subseteq [x, 3x] \cap \mathbb{N}$ with $\#\mathcal{I} = O(\tau x^{0.23+\varepsilon})$ such that $y \in ([x, 3x] \cap \mathbb{N}) \setminus \mathcal{I}$ implies

$$\sum_{\substack{k_1 \dots k_r m \in [y, y+y/\tau] \\ k_i \in (P_i, 2P_i] \\ (m,p)=1 \text{ for } p<z(p_k)}} 1_{\mathbb{P}}(k_1) \dots 1_{\mathbb{P}}(k_r) - \frac{x^b}{\tau} \sum_{\substack{k_1 \dots k_r m \in [y, y+y/x^b] \\ k_i \in (P_i, 2P_i] \\ (m,p)=1 \text{ for } p<z(p_k)}} 1_{\mathbb{P}}(k_1) \dots 1_{\mathbb{P}}(k_r) \approx 0. \tag{2.5}$$

Here z is a function of p_k with $z(p_k) = p_k$ or $z(p_k) = x^\beta$ for some small β. This allows us to compare how many elements of $[y, y + y/\tau]$ and $[y, y + y/x^b]$ are of the form $p_1 \dots p_r m$ with $p_i \in (P_i, 2P_i]$ and m not divisible by any prime less than $z(p_k)$. The formal statement of this result is Lemma 2.1.

Finally, recall that we are looking to bound $(\pi(y+\frac{y}{\tau}) - \pi(y)) - (\frac{x^b}{\tau})(\pi(y+\frac{y}{x^b}) - \pi(y))$ from below. We may use the Buchstab identity, an inclusion–exclusion argument, to rewrite the difference $(\pi(y+\frac{y}{\tau}) - \pi(y)) - (\frac{x^b}{\tau})(\pi(y+\frac{y}{x^b}) - \pi(y))$ as a sum of terms which look like the LHS of (2.5) weighted by $(-1)^r$. From Lemma 2.1 we know that the LHS of (2.5) is close to zero for most y when $P_1, \ldots, P_r$ are in certain good ranges. On the other hand, when $P_1, \ldots, P_r$ are bad and r is even, then the LHS of (2.5), weighted by $(-1)^r = 1$, can be bounded below by discarding the contribution of short interval $[y, y + y/\tau]$ and estimating the sum over long interval $[y, y + y/x^b]$ asymptotically. This is possible because counting multiples of certain primes in a long interval is easy.

If we choose our Buchstab decomposition so that the overall bound on the bad $[y, y + y/x^b]$–sums does not exceed $\frac{\alpha y}{x^b \log x}$ for some fixed $\alpha \in (0, 1)$, then this gives the desired lower bound (2.2) for all but $O(\tau x^{0.23+\varepsilon})$ choices of $y \in [x, 3x] \cap \mathbb{N}$, which in turn proves Theorem 2.1. In practice, finding suitable Buchstab iterations is a long and calculation-heavy process. These computations, reformulated as seeking a suitable minorant for $1_{\mathbb{P}}(n)$, can be found Section 2.6. Since the good ranges of P_i vary with τ, we apply Harman's sieve six separate times. The formal statement of the results of Section 2.6 is Proposition 2.3.

2.1.3 Notation

We will use the following notation from sieve theory: For $\mathcal{C} \subseteq \mathbb{N}$ and $z > 0$,

$$S(\mathcal{C}, z) = \#\{n \in \mathcal{C} : (n, P(z)) = 1\} \quad \text{where} \quad P(z) = \prod_{p<z} p.$$

We also let $V(z) = \prod_{p<z}(1 - 1/p)$ and $\mathcal{C}_d = \{n \in \mathbb{N} : nd \in \mathcal{C}\}$. For any $n \in \mathbb{N}$ and any $z > 0$, we take

$$\psi(n, z) = \begin{cases} 1 & \text{if } p \mid n \Rightarrow p \geq z \\ 0 & \text{otherwise.} \end{cases}$$

In particular, $\sum_{md \in \mathcal{C}} \psi(m, z) = S(\mathcal{C}_d, z)$.

$1_{\mathcal{C}}$ is the indicator function of $\mathcal{C}$ and $\mathbb{P}$ is the set of primes. We write $n \sim N$ to mean $n \in (N, 2N] \cap \mathbb{N}$.

Throughout this chapter, we will work with a fixed, but arbitrary constant $\varepsilon > 0$ and with quantities $\nu = 0.23$, $a \in [0.475 - \varepsilon, 0.77 - \varepsilon]$, $\tau = x^a$, $b = 1/10^5$, $\varepsilon_1 = \varepsilon/10^5$, $T_0 = \tau x^{\varepsilon_1}$ and $T_1 = \gamma x^{b/2 - \varepsilon_1}$, where $\gamma \in [1, 2]$ is chosen such that $\log(T_0/T_1)/\log 2 \in \mathbb{Z}$, allowing for a nice dyadic decomposition of $[T_1, T_0]$. We compare sets

$$\mathcal{A}(y) = \left[y, y + \frac{y}{\tau}\right] \cap \mathbb{N} \quad \text{and} \quad \mathcal{B}(y) = \left[y, y + \frac{y}{x^b}\right] \cap \mathbb{N}.$$

Throughout all our proofs, we will assume without loss of generality that x is very large and ε is small.

2.2 Key propositions

In this section we now introduce important definitions, formally state our key propositions and lemma and then show that they combine to give a proof of Theorem 2.1.

2.2.1 About Proposition 2.1

Our first proposition relates bounds on differences between sums over $\mathcal{A}(y)$ and $\mathcal{B}(y)$ to values of Dirichlet polynomials. Given $\varepsilon > 0$ and $\tau \in [x^{0.475-\varepsilon}, x^{0.77-\varepsilon}]$, we consider

$$\sum_{n \in \mathcal{A}(y)} a_n - \frac{x^b}{\tau} \sum_{n \in \mathcal{B}(y)} a_n,$$

where (a_n) is a sequence with the following properties: There exist $K \in \mathbb{N}$, $N_i \in [\frac{1}{2}, \infty)$, $r_1, r_2 \in \mathbb{N}$ with $0 \leq r_1 \leq r_2 \leq 10^5$ and $a_n^{(i)} \in \mathbb{C}$ with $a_n^{(i)} = O(\log(n+3))$ for $i > r_1$ such that

$$a_n = \sum_{\substack{k_0 k_1 \dots k_{r_2} = n \\ k_i \sim N_i}} \left(1_{\mathbb{N}}(k_0) \prod_{1 \leq i \leq r_1} 1_{\mathbb{P}}(k_i) \prod_{r_1 < i \leq r_2} a_{k_i}^{(i)} \right).$$

Here we require that $2^{-r_2} x \leq \prod_{i=0}^{r_2} N_i \leq 6x$ and $N_i < x^{1/K}$ for $i > r_1$ and $N_i \geq x^{1/K}$ for $1 \leq i \leq r_1$. We additionally assume that r_1, $N_0, \dots, N_{r_2}$ and $(a_n^{(i)})$ satisfy one of the following three options:

(i) $r_1 \geq 2$ or $N_0 \geq x^{0.95}$.

(ii) $r_1 = 1$ and $N_0 \geq x^{1/\log\log x}$.

(iii) $r_1 \in \{0, 1\}$ and there exists $i > r_1$ with $a_n^{(i)} = 1_{\mathbb{P}}(n)$ for all $n \in \mathbb{N}$ and $N_i \geq x^{1/\log\log x}$.

We will associate with (a_n) a collection of Dirichlet polynomials with various nice properties. In particular, the factors of these Dirichlet polynomials will be of the form described below.

Definition 2.1. *For a given $T \in [1, x]$ and a Dirichlet polynomial $H(s) = \sum_{n \sim N} \frac{h_n}{n^s}$, we say $H(s)$ has property $(Z1)$ with respect to T if there exists $w \in [-T/2, T/2]$ with*

$$H(s + iw) = \sum_{n \sim N} \frac{1}{n^s} \qquad \text{or} \qquad H(s + iw) = \sum_{n \sim N} \frac{\log n}{n^s}. \tag{Z1}$$

We say $H(s)$ has property $(Z2)$ with respect to T if there exists $w \in [-T/2, T/2]$ with

$$H(s+iw) = \sum_{k=2}^{\infty} \sum_{p^k \sim N} \frac{1}{kp^{ks}}. \tag{Z2}$$

We say that $H(s)$ has property (D) with respect to T if there exists $w \in [-T/2, T/2]$ such that $H(s+iw)$ equals one of the following five Dirichlet polynomials:

$$H(s+iw) \in \left\{ \sum_{n \sim N} \frac{1}{n^s}, \quad \sum_{n \sim N} \frac{\log n}{n^s}, \quad \sum_{k=2}^{\infty} \sum_{p^k \sim N} \frac{1}{kp^{ks}}, \quad \sum_{n \sim N} \frac{1_{\mathbb{P}}(n)}{n^s}, \quad \sum_{n \sim N} \frac{\mu(n)}{n^s} \right\}. \tag{D}$$

For (a_n) with corresponding $(N_0, \ldots, N_{r_1})$ and K, we consider the following sets of Dirichlet polynomials:

Definition 2.2. *$\mathcal{F}^*(T, J, K)$ is the set of Dirichlet polynomials $F(s)$ with*

$$F(s) = \prod_{i=1}^{j} S_i(s)$$

for some $j \leq J$ and some $S_i(s) = \sum_{n \sim B_i} b_n^{(i)} n^{-s}$ which satisfy:

1. *$b_n^{(i)} \in \mathbb{C}$ with $b_n^{(i)} = O(\log n)$.*
2. *$B_i \geq \log x$ and $2^{-J} x \leq \prod_{i=1}^{j} B_i \leq 2^J x$.*
3. *If $B_\ell \geq x^{2/K}$, then $S_\ell(s)$ has property $(Z1)$ or $(Z2)$ with respect to T.*
4. *At least one of the following three conditions is satisfied:*
   - (A) *There exists $\ell \in \{1, \ldots, j\}$ with $B_\ell \geq x^{2/K}$ and there exists $i_0 \neq \ell$ such that $B_{i_0} \geq x^{\frac{1}{\log \log x}}$ and such that $S_{i_0}(s)$ has property (D) with respect to T.*
   - (B) *For every $i \in \{1, \ldots, j\}$, $B_i < x^{2/K}$, and there exists $i_0 \in \{1, \ldots, j\}$ such that $B_{i_0} \geq x^{\frac{1}{\log \log x}}$ and such that $S_{i_0}(s)$ has property (D) with respect to T.*
   - (C) *There exists $\ell \in \{1, \ldots, j\}$ with $B_\ell \geq x^{0.95}$.*

$\mathcal{F}((N_0, \ldots, N_{r_1}), T, J, K)$ is the set of $F(s) \in \mathcal{F}^*(T, J, K)$ for which factorisation $\prod_{i=1}^{j} S_i(s)$ can be chosen so that the following three conditions are all satisfied:

(i) *There exist disjoint subsets $X_1, \ldots, X_{r_1}$ of $\{1, \ldots, j\}$ with*

$$(\log x)^{-2K-1} N_\ell \leq \prod_{i \in X_\ell} B_i \leq (\log x)^{2K+1} N_\ell \quad \text{for} \quad 1 \leq \ell \leq r_1.$$

(ii) *If $N_0 \geq \log x$, there exists $i_0 \in \{1, \ldots, j\} \setminus (\cup_{\ell=1}^{r_1} X_\ell)$ with $S_{i_0}(s) = \sum_{k \sim N_0} k^{-s}$.*

(iii) *If $i \notin \bigcup_{\ell=1}^{r_1} X_\ell$ and $i \neq i_0$, then $B_i < x^{1/K}$.*

$\mathcal{G}(\tau, \varepsilon)$ is the set of Dirichlet polynomials with the following property: If $G(s) \in \mathcal{G}(\tau, \varepsilon)$, then $G(s) = \sum_{n \in \mathcal{J}} \xi_n n^{-s}$ for some $\mathcal{J} \subseteq [1, \tau x^{\varepsilon_1}]$ with $\#\mathcal{J} \geq \tau^2 x^{-0.77}$ and some $\xi_n \in \mathbb{C}$ with $|\xi_n| = 1$.

Next we give the formal definition of quantities R, R^* and Q, first introduced in Section 2.1.2.

Definition 2.3. *Let $c = 1 + 1/\log x$. For Dirichlet polynomials $F(s) = \prod_{i=1}^{j} S_i(s) = \prod_{i=1}^{j} (\sum_{n \sim B_i} \frac{b_n^{(i)}}{n^s})$, $G(s) = \sum_{n \in \mathcal{J}} \frac{\xi_n}{n^s}$ and $(\sigma_i) \in \mathbb{R}^j$, $\gamma \in \mathbb{R}$ and $T \in [T_1, T_0]$, we define:*

$$\mathcal{S}_1(F, (S_i), (\sigma_i)) = \bigcap_{i=1}^{j} \Big\{ n \in \mathbb{N} : \sup_{t \in [n, n+1]} |S_i(c + it)| \in (B_i^{-c+\sigma_i}, 2B_i^{-c+\sigma_i}] \Big\},$$

$$R(F, (S_i), (\sigma_i), T) = \#(\mathcal{S}_1(F, (S_i), (\sigma_i)) \cap [T, 2T]), \tag{2.6}$$

$$\mathcal{S}_1^*(F, (S_i), (\sigma_i)) = \left\{ (n_1, n_2, n_3, n_4) \in \mathcal{S}_1(F, (S_i), (\sigma_i))^4 : n_1 + n_2 = n_3 + n_4 \right\},$$

$$R^*(F, (S_i), (\sigma_i), T) = \#(\mathcal{S}_1^*(F, (S_i), (\sigma_i)) \cap [T, 2T]^4), \tag{2.7}$$

$$\mathcal{S}_2(G, \gamma) = \Big\{ n \in \mathbb{N} : \sup_{t \in [n, n+1]} |G(it)| \in ((\#\mathcal{J})^\gamma, 2(\#\mathcal{J})^\gamma] \Big\},$$

$$Q(F, G, (S_i), (\sigma_i), \gamma, T) = \#(\mathcal{S}_1(F, (S_i), (\sigma_i)) \cap \mathcal{S}_2(G, \gamma) \cap [T, 2T]). \tag{2.8}$$

We are now ready to state Proposition 2.1.

Proposition 2.1. *Let $\varepsilon > 0$, $\tau \in [x^{0.475 - \varepsilon}, x^{0.77 - \varepsilon}]$ and $A \in \mathbb{N}$. Suppose that (a_n) is as described at the start of Section 2.2.1, with corresponding constants r_1, r_2, $(N_0, \ldots, N_{r_2})$ and K. Let $J = A + 10^7 K$.*

Suppose each choice of $T \in [T_1, T_0]$, $F(s) = \prod_{i=1}^{j} S_i(s) = \prod_{i=1}^{j} (\sum_{n \sim B_i} \frac{b_n^{(i)}}{n^s}) \in \mathcal{F}((N_0, \ldots, N_{r_1}), T, J, K)$, $G(s) = \sum_{n \in \mathcal{J}} \xi_n n^{-s} \in \mathcal{G}(\tau, \varepsilon)$, $\gamma \in (-\infty, 1)$ and (σ_i) with $\sigma_i \in \mathbb{R}$ and $\sigma = \frac{\sum_{i=1}^{j} \log(B_i) \sigma_i}{\sum_{i=1}^{j} \log(B_i)}$ satisfies at least one of the following five conditions:

$$R(F, (S_i), (\sigma_i), T) \leq \frac{x^{1-\sigma}}{(\log x)^{2J}}, \tag{C1}$$

$$R(F, (S_i), (\sigma_i), T) \leq \tau x^{1.23 - 2\sigma + 2\varepsilon_1}, \tag{C2}$$

$$R^*(F, (S_i), (\sigma_i), T) \leq \tau x^{3.23 - 4\sigma + 2\varepsilon_1}, \tag{C3}$$

$$Q(F, G, (S_i), (\sigma_i), \gamma, T) \leq (\#\mathcal{J})^{7/8 - \gamma} \tau^{1/4} x^{7/8 - \sigma + 0.23/8 + \varepsilon_1}, \tag{C4.A}$$

$$Q(F, G, (S_i), (\sigma_i), \gamma, T) \leq (\#\mathcal{J})^{1 - \gamma} x^{1 - \sigma - \varepsilon_1}. \tag{C4.B}$$

Then there exists $\mathcal{I} \subseteq [x, 3x] \cap \mathbb{N}$ *such that* $\#\mathcal{I} = O(\tau x^{0.23+\varepsilon/2})$ *and such that* $y \notin ([x, 3x] \cap \mathbb{N}) \setminus \mathcal{I}$ *implies*

$$\left| \sum_{n \in \mathcal{A}(y)} a_n - \frac{x^b}{\tau} \sum_{n \in \mathcal{B}(y)} a_n \right| \leq \frac{x}{\tau (\log x)^A}.$$

Here the constant implied by big O notation only depends on A, K *and* ε.

2.2.2 About Proposition 2.2

The second proposition gives conditions on the factor lengths of $F(s) = \prod_{i=1}^{j} S_i(s)$ which ensure that one of (C1), (C2), (C3), (C4.A) or (C4.B) holds. For a fixed $\varepsilon > 0$, $\varepsilon_1 = \varepsilon/10^5$ and a given a, we define a quantity $\chi_0(a)$ and regions $\chi_1(a)$, $\chi_2(a)$ and $\chi_3(a)$ as follows:

$$\chi_0(a) = \begin{cases} 0.290 - \varepsilon_1 & \text{if } a \leq 0.53, \\ 0.315 - \varepsilon_1 & \text{if } a \in (0.53, 0.545], \\ 0.335 - \varepsilon_1 & \text{if } a \in (0.545, 0.57], \\ 0.330 - \varepsilon_1 & \text{if } a \in (0.57, 0.59], \\ 0.330 - \varepsilon_1 & \text{if } a \in (0.59, 0.61], \\ 0.320 - \varepsilon_1 & \text{if } a > 0.61. \end{cases} \tag{2.9}$$

$$\widetilde{\chi}_1(a) = \begin{cases} [0.290 - \varepsilon_1, 0.360 + \varepsilon_1] & \text{if } a \leq 0.53, \\ [0.315 - \varepsilon_1, 0.345 + \varepsilon_1] \cup [0.427 - \varepsilon_1, 0.474 + \varepsilon_1] & \text{if } a \in (0.53, 0.545], \\ [0.400 - \varepsilon_1, 0.475 + \varepsilon_1] & \text{if } a \in (0.545, 0.57], \\ [0.380 - \varepsilon_1, 0.420 + \varepsilon_1] & \text{if } a \in (0.57, 0.59], \\ [0.365 - \varepsilon_1, 0.420 + \varepsilon_1] & \text{if } a \in (0.59, 0.61], \\ [0.355 - \varepsilon_1, 0.420 + \varepsilon_1] & \text{if } a > 0.61. \end{cases}$$

$$\chi_1(a) = \widetilde{\chi}_1(a) \cup (1 - \widetilde{\chi}_1(a)) = \widetilde{\chi}_1(a) \cup \{1 - \ell : \ell \in \widetilde{\chi}_1(a)\}. \tag{2.10}$$

$$\chi_2(a) = \begin{cases} \chi_1(a) & \text{if } a \leq 0.53, \\ [0.405 - \varepsilon_1, 0.485 + \varepsilon_1] \cup [0.515 - \varepsilon_1, 0.595 + \varepsilon_1] & \text{if } a \in (0.53, 0.545], \\ \chi_1(a) & \text{if } a \in (0.545, 0.57], \\ [0.380 - \varepsilon_1, 0.455 + \varepsilon_1] \cup [0.545 - \varepsilon_1, 0.620 + \varepsilon_1] & \text{if } a \in (0.57, 0.59], \\ [0.365 - \varepsilon_1, 0.435 + \varepsilon_1] \cup [0.565 - \varepsilon_1, 0.635 + \varepsilon_1] & \text{if } a \in (0.59, 0.61], \\ \chi_1(a) & \text{if } a > 0.61. \end{cases} \tag{2.11}$$

$$\chi_3(a) = \begin{cases} \chi_1(a) & \text{if } a \le 0.53, \\ [0.285-\varepsilon_1, 0.375+\varepsilon_1] \cup [0.625-\varepsilon_1, 0.715+\varepsilon_1] & \text{if } a \in (0.53, 0.545], \\ \chi_1(a) & \text{if } a \in (0.545, 0.57], \\ [0.315-\varepsilon_1, 0.420+\varepsilon_1] \cup [0.580-\varepsilon_1, 0.685+\varepsilon_1] & \text{if } a \in (0.57, 0.59], \\ [0.330-\varepsilon_1, 0.420+\varepsilon_1] \cup [0.580-\varepsilon_1, 0.670+\varepsilon_1] & \text{if } a \in (0.59, 0.61], \\ \chi_1(a) & \text{if } a > 0.61. \end{cases} \tag{2.12}$$

The following result then holds for $\chi_0(a)$, $\chi_1(a)$, $\chi_2(a)$ and $\chi_3(a)$ as given in (2.9), (2.10), (2.11) and (2.12):

Proposition 2.2. *Let $\varepsilon > 0$, $a \in [0.475-\varepsilon, 0.77-\varepsilon]$, $\tau = x^a$, $K = 2000$, $J > 10^7 K$ and $T \in [T_1, T_0]$. Let $\mathcal{F}^*(T, J, K)$ and $\mathcal{G}(\tau, \varepsilon)$ be as described in Definition 2.2 and suppose that $F(s) = \prod_{i=1}^{j} S_i(s) = \prod_{i=1}^{j} \sum_{n \sim B_i} b_n^{(i)} n^{-s} \in \mathcal{F}^*(T, J, K)$ and $G(s) \in \mathcal{G}(\tau, \varepsilon)$.*

Write $\ell_k = \frac{\log(B_k)}{\sum_{i=1}^{j} \log(B_i)}$ for $1 \le k \le j$. Suppose one of the following options holds:

1. *There exists $k \in \{1, \ldots, j\}$ with $\ell_k \ge \chi_0(a)$.*
2. *There exists $I_1 \subseteq \{1, \ldots, j\}$ with $\sum_{i \in I_1} \ell_i \in \chi_1(a)$.*
3. *There exist $I_2 \subseteq \{1, \ldots, j\}$ with $\sum_{i \in I_2} \ell_i \in \chi_2(a)$ and $I_3 \subseteq \{1, \ldots, j\}$ with $\sum_{i \in I_3} \ell_i \in \chi_3(a)$.*

Then there exists a large constant C, dependent only on J and ε, such that for $x \ge C$ and for every $\gamma \in (-\infty, 1)$ and (σ_i) with $\sigma_i \in \mathbb{R}$ and $\sigma = \frac{\sum_{i=1}^{j} \log(B_i)\sigma_i}{\sum_{i=1}^{j} \log(B_i)}$, one of the five conditions (C1)*,* (C2)*,* (C3)*,* (C4.A) *or* (C4.B)*, described in Proposition* 2.1*, holds.*

2.2.3 About Lemma 2.1

We combine Proposition 2.1 and 2.2 to compare sifted sets. Criteria concerning the size of Dirichlet polynomials are now replaced by purely combinatorial conditions.

Definition 2.4. *Let $\beta \in [0.01, 0.15]$, $r \in \{0, \ldots, 5\}$ and $\beta \le \ell_i^* \le 0.5 + \varepsilon$, $\ell_i^* \ge \ell_{i+1}^*$ and $\ell_1^* + \cdots + \ell_r^* \le 0.75$. We denote by $\Xi^\star(\ell_1^*, \ldots, \ell_r^*, \beta)$ the set of finite sequences $\{\ell_i\}_{i=1}^{j}$ with $\ell_1, \ldots, \ell_j \in [0,1]$, $\sum_{i=1}^{j} \ell_i = 1$ and $j \le 10^{20}$ for which there exist disjoint subsets $X_1, \ldots, X_r$ of $\{1, \ldots, j\}$ with*

$$\sum_{i \in X_s} \ell_i \in \left[\ell_s^* - \frac{\varepsilon}{10^{100}}, \ell_s^* + \frac{\varepsilon}{10^{100}}\right]$$

for $s \le r$ and $\ell_i \le \beta + 10^{-100}\varepsilon$ for all but at most one $i \in \{1,\dots,j\} \setminus \bigcup_{i=1}^r X_r$.

Let $r \in \{2,4,6\}$ and $0.01 - \varepsilon \le \ell_i^* \le 0.5+\varepsilon$, $\ell_i^* \ge \ell_{i+1}^*$ and $\ell_1^* + \dots + \ell_r^* \le 0.99$. We denote by $\Xi^{\star\star}(\ell_1^*,\dots,\ell_r^*)$ the set of finite sequences $\{\ell_i\}_{i=1}^j$ with $\ell_1,\dots,\ell_j \in [0,1]$, $\sum_{i=1}^j \ell_i = 1$ and $j \le 10^{20}$ for which there exist disjoint subsets $X_1,\dots,X_r$ of $\{1,\dots,j\}$ with

$$\sum_{i \in X_s} \ell_i \in \left[\ell_s^* - \frac{\varepsilon}{10^{100}}, \ell_s^* + \frac{\varepsilon}{10^{100}}\right] \qquad \text{for } s \le r.$$

Suppose $\{\ell_i\}_{i=1}^j \in \Xi^\star(\ell_1^*,\dots,\ell_r^*,\beta)$ or $\{\ell_i\}_{i=1}^j \in \Xi^{\star\star}(\ell_1^*,\dots,\ell_r^*)$. We then say that $\{\ell_i\}_{i=1}^j$ satisfies one of the options (1), (2) or (3) for a given $a \in [0.475-\varepsilon, 0.77-\varepsilon]$ if one of the following conditions holds:

(1) *There exists $k \in \{1,\dots,j\}$ with $\ell_k \ge \chi_0(a)$.*

(2) *There exists $I_1 \subseteq \{1,\dots,j\}$ with $\sum_{i\in I_1} \ell_i \in \chi_1(a)$.*

(3) *There exist $I_2 \subseteq \{1,\dots,j\}$ with $\sum_{i\in I_2} \ell_i \in \chi_2(a)$ and $I_3 \subseteq \{1,\dots,j\}$ with $\sum_{i\in I_3} \ell_i \in \chi_3(a)$.*

We have the following relationships between sifted sets:

Lemma 2.1. *Let $\varepsilon > 0$. Let $a \in [0.475-\varepsilon, 0.77-\varepsilon]$ and set $\tau = x^a$.*

Let $\beta \in [0.01, 0.15]$, $r \in \{0,\dots,5\}$ and $P_i = x^{\ell_i^*}$ with $\beta \le \ell_i^* \le 0.5+\varepsilon$, $P_i \ge P_{i+1}$ and $P_1 \dots P_r \le x^{0.75}$. Suppose that every $\{\ell_i\}_{i=1}^j \in \Xi^\star(\ell_1^*,\dots,\ell_r^*,\beta)$ satisfies one of the options (1), (2) or (3) for the given a.

Then there exists $\mathcal{I} \subseteq [x,3x] \cap \mathbb{N}$ such that $\#\mathcal{I} = O(\tau x^{0.23+3\varepsilon/4})$ and such that $y \in ([x,3x]\cap\mathbb{N}) \setminus \mathcal{I}$ implies

$$\left| \sum_{p_i \sim P_i} S(\mathcal{A}(y)_{p_1\dots p_r}, x^\beta) - \frac{x^b}{\tau} \sum_{p_i\sim P_i} S(\mathcal{B}(y)_{p_1\dots p_r}, x^\beta) \right| \le \left(\frac{(\log\log x)^{O(1)} y}{\tau (\log x)^{2+r}} \right).$$

Alternatively, let $r \in \{2,4,6\}$ and $P_i = x^{\ell_i^*}$ with $0.01-\varepsilon \le \ell_i^* \le 0.5+\varepsilon$, $P_i \ge P_{i+1}$ and $P_1 \dots P_r \le x^{0.99}$. Suppose that every $\{\ell_i\}_{i=1}^j \in \Xi^{\star\star}(\ell_1^*,\dots,\ell_r^*)$ satisfies one of the options (1), (2) or (3) for the given a.

Then there again exists $\mathcal{I} \subseteq [x,3x]\cap\mathbb{N}$ such that $\#\mathcal{I} = O(\tau x^{0.23+3\varepsilon/4})$ and $y \in ([x,3x]\cap\mathbb{N})\setminus\mathcal{I}$ implies

$$\left| \sum_{p_i \sim P_i} S(\mathcal{A}(y)_{p_1\dots p_r}, p_r) - \frac{x^b}{\tau} \sum_{p_i\sim P_i} S(\mathcal{B}(y)_{p_1\dots p_r}, p_r) \right| \le \left(\frac{(\log\log x)^{O(1)} y}{\tau (\log x)^{2+r}} \right).$$

Here the constants implied by big O notation depend only on ε.

2.2.4 About Proposition 2.3

The final proposition concerns minorants of the prime indicator function, which are constructed via Harman's sieve, using as input the comparison of sifted sets from Lemma 2.1. We work with the following regions:

Definition 2.5. *Let $\mathcal{R}^*(a)$ denote the set of tuples $(\ell_1^*, \dots, \ell_r^*, \beta)$ with $\beta \in [0.01, 0.15]$, $r \in \{0, \dots, 5\}$, $\ell_1^*, \dots, \ell_r^* \in [\beta, 0.5+\varepsilon]$, $\ell_i^* \geq \ell_{i+1}^*$ and $\ell_1^* + \cdots + \ell_r^* \leq 0.75$ and for which every $\{\ell_i\}_{i=1}^j \in \Xi^\star(\ell_1^*, \dots, \ell_r^*)$ satisfies one of the options (1), (2) or (3) for a.*

Let $\mathcal{R}^{**}(a)$ denote the set of tuples $(\ell_1^*, \dots, \ell_r^*)$ which have $r \in \{2,4,6\}$, $\ell_1^*, \dots, \ell_r^* \in [0.01-\varepsilon, 0.5+\varepsilon]$, $\ell_i^* \geq \ell_{i+1}^*$ and $\sum_{i=1}^r \ell_i^* \leq 0.99$ and for which every $\{\ell_i\}_{i=1}^j \in \Xi^{\star\star}(\ell_1^*, \dots, \ell_r^*)$ satisfies (1), (2) or (3) for a.

Proposition 2.3. *Let $\varepsilon > 0$ and $a \in [0.475-\varepsilon, 0.77-\varepsilon]$. If x is sufficiently large, there exists a function $\rho : [x, 6x] \cap \mathbb{N} \to \mathbb{R}$ with the following properties:*

(α) *For any $n \in [x, 6x] \cap \mathbb{N}$, $\rho(n) \leq 1_{\mathbb{P}}(n)$.*

(β) *ρ can be written as a sum $\rho = \sum_{t=1}^s \rho_t$, where each $\rho_t : [x, 6x] \cap \mathbb{N} \to \mathbb{R}$ satisfies*

$$\rho_t(n) = (-1)^r \sum_{\substack{n = p_1 \dots p_r m \\ p_i \sim x^{\ell_i^*}}} \psi(m, x^\beta) \ \text{ for some } r \in \{0, \dots, 5\},\ (\ell_1^*, \dots, \ell_r^*, \beta) \in \mathcal{R}^*(a),$$

$$\text{or} \quad \rho_t(n) = \sum_{\substack{n = p_1 \dots p_r m \\ p_i \sim x^{\ell_i^*}}} \psi(m, p_r) \ \text{ for some } r \in \{2,4,6\} \text{ and } (\ell_1^*, \dots, \ell_r^*) \in \mathcal{R}^{**}(a).$$

Here only $O((\log x)^u)$ choices of t shall have $r = u$.

(γ) *For any $y \in [x, 3x] \cap \mathbb{N}$,*

$$\sum_{n \in \mathcal{B}(y)} (1_{\mathbb{P}}(n) - \rho(n)) \leq \frac{0.99999y}{x^b \log x}.$$

The required size of x depends only on ε.

2.2.5 Final steps

We now show how to deduce Theorem 2.1 from Lemma 2.1 and Proposition 2.3. Recall that our goal is to prove bound $\sum_{p_n \leq x} (p_{n+1} - p_n)^2 \ll_\varepsilon x^{1.23+\varepsilon}$ for every $\varepsilon > 0$. By dyadic decomposition, it is sufficient to show that for any $\tau > 0$,

$$\sum_{\substack{x \leq p_n \leq 2x \\ 6x/\tau \leq (p_{n+1}-p_n) \leq 12x/\tau}} (p_{n+1} - p_n)^2 \ll_\varepsilon x^{1.23+\varepsilon}. \tag{2.13}$$

This statement holds trivially when $\tau \geq x^{0.77-\varepsilon}$. A result of Baker, Harman and Pintz [1] also tells us that $(p_{n+1} - p_n) \ll x^{0.525}$ and so (2.13) holds when $\tau \leq x^{0.475-\varepsilon}$. Hence we restrict our attention to τ with $x^{0.475-\varepsilon} \leq \tau \leq x^{0.77-\varepsilon}$. Following an argument of Peck [36], we observe that

$$\sum_{\substack{x \leq p_n \leq 2x \\ 6x/\tau \leq (p_{n+1}-p_n) \leq 12x/\tau}} (p_{n+1} - p_n)^2 \ll \frac{x}{\tau} \sum_{\substack{x \leq p_n \leq 2x, \\ p_{n+1}-p_n \geq 6x/\tau}} \# \left\{ y \in \mathbb{N} : y \in \left(p_n, \frac{p_n + p_{n+1}}{2} \right) \right\}.$$

If $y \in (p_n, (p_n + p_{n+1})/2)$ and $p_n \leq 2x$, then $y \leq 3x$ by Bertrand's postulate and $y + y/\tau \leq y + 3x/\tau$. If $p_{n+1} - p_n \geq 6x/\tau$, then also $y + y/\tau \leq y + (p_{n+1} - p_n)/2 < p_{n+1}$. Hence $[y, y + y/\tau]$ is free of primes. To obtain bound (2.13) it thus suffices to show that there are only $O(\tau x^{0.23+\varepsilon})$ integers $y \in \mathbb{N} \cap [x, 3x]$ for which $[y, y + y/\tau]$ contains no primes.

We count the number of primes in short interval $\mathcal{A}(y) = [y, y+y/\tau] \cap \mathbb{N}$ by comparing it to the number of primes in the longer interval $\mathcal{B}(y) = [y, y+y/x^b] \cap \mathbb{N}$, where $b = 10^{-5}$. Using Proposition 2.3 with $\tau = x^a$, there is a minorant $\rho(n) = \sum_{t=1}^{s} \rho_t(n)$ of the prime indicator function $1_{\mathbb{P}}(n)$ which satisfies properties (α), (β) and (γ). Using (α) and (γ), we obtain lower bound

$$\begin{aligned} \sum_{n \in \mathcal{A}(y)} 1_{\mathbb{P}}(n) - \frac{x^b}{\tau} \sum_{n \in \mathcal{B}(y)} 1_{\mathbb{P}}(n) &\geq \sum_{n \in \mathcal{A}(y)} \rho(n) - \frac{x^b}{\tau} \sum_{n \in \mathcal{B}(y)} \rho(n) + (1_{\mathbb{P}}(n) - \rho(n)) \qquad (2.14) \\ &\geq \sum_{t=1}^{s} \left(\sum_{n \in \mathcal{A}(y)} \rho_t(n) - \frac{x^b}{\tau} \sum_{n \in \mathcal{B}(y)} \rho_t(n) \right) - \frac{0.99999y}{\tau \log x}. \end{aligned}$$

Recall that $S(\mathcal{C}_d, z) = \sum_{dm \in \mathcal{C}} \psi(m, z)$ for any $\mathcal{C} \subseteq \mathbb{N}$, $d \in \mathbb{N}$ and $z > 0$. By property (β) we thus have that $\sum_{n \in \mathcal{A}(y)} \rho_t(n) - (x^b/\tau) \sum_{n \in \mathcal{B}(y)} \rho_t(n)$ equals one of

$$\begin{aligned} &(-1)^r \left(\sum_{p_i \sim P_i} S(\mathcal{A}(y)_{p_1 \ldots p_r}, x^\beta) - \frac{x^b}{\tau} \sum_{p_i \sim P_i} S(\mathcal{B}(y)_{p_1 \ldots p_r}, x^\beta) \right) \\ \text{or} \quad &\left(\sum_{p_i \sim P_i} S(\mathcal{A}(y)_{p_1 \ldots p_r}, p_r) - \frac{x^b}{\tau} \sum_{p_i \sim P_i} S(\mathcal{B}(y)_{p_1 \ldots p_r}, p_r) \right), \end{aligned}$$

where $P_i = x^{\ell_i^*}$ and where every $\{\ell_i\}_{i=1}^{j} \in \Xi^\star(\ell_1^*, \ldots, \ell_r^*, \beta)$ satisfies one of the options (1), (2) or (3) (when $z = x^\beta$) or every $\{\ell_i\}_{i=1}^{j} \in \Xi^{\star\star}(\ell_1^*, \ldots, \ell_r^*)$ satisfies one of the options (1), (2) or (3) (when $z = p_r$). Here $\Xi^\star(\ell_1^*, \ldots, \ell_r^*, \beta)$, $\Xi^{\star\star}(\ell_1^*, \ldots, \ell_r^*)$ and (1), (2) and (3) (with $a = \log_x(\tau)$) are as given in Definition 2.4.

Hence by Lemma 2.1 there exists for each $t \in \{1, \dots, s\}$ a set $\mathcal{I}_t$ such that $\#\mathcal{I}_t = O(\tau x^{0.23+3\varepsilon/4})$ and such that every $y \in ([x, 3x] \cap \mathbb{N}) \setminus \mathcal{I}_t$ satisfies

$$\sum_{n \in \mathcal{A}(y)} \rho_t(n) - (x^b/\tau) \sum_{n \in \mathcal{B}(y)} \rho_t(n) \geq -\frac{(\log\log x)^{O(1)} y}{\tau (\log x)^{2+r}}.$$

However, as part of property (β) we also have that only $O((\log x)^u)$ choices of t have $r = u$. Thus when $y \in ([x, 3x] \cap \mathbb{N}) \setminus (\cup_{t=1}^s \mathcal{I}_t)$, also $\sum_{t=1}^s (\sum_{n \in \mathcal{A}(y)} \rho_t(n) - (x^b/\tau) \sum_{n \in \mathcal{B}(y)} \rho_t(n)) \geq -(\log\log x)^{O(1)} y/(\tau(\log x)^2)$. Rearranging (2.14), we find that $\mathcal{I} = \bigcup_{t=1}^s \mathcal{I}_t$ has $\#\mathcal{I} = O(\tau x^{0.23+\varepsilon})$ and $y \in ([x, 3x] \cap \mathbb{N}) \setminus \mathcal{I}$ implies

$$\sum_{n \in \mathcal{A}(y)} 1_{\mathbb{P}}(n) \geq \frac{x^b}{\tau} \sum_{n \in \mathcal{B}(y)} 1_{\mathbb{P}}(n) - \frac{0.999999y}{\tau \log x} = \frac{(0.000001 + o(1))y}{\tau \log x}. \tag{2.15}$$

In the final inequality we used that $\mathcal{B}(y)$ is an interval of length y/x^b, where $y \in [x, 3x]$ and b is very small, and thus must contain $(1 + o(1))y/(x^b \log x)$ primes. Inequality (2.15) implies that $\mathcal{A}(y)$ contains primes, provided that x is sufficiently large. We have shown that there are only $O(\tau x^{0.23+\varepsilon})$ integers $y \in \mathbb{N} \cap [x, 3x]$ for which $[y, y + y/\tau]$ contains no primes. As discussed above, this implies Theorem 2.1.

Hence to complete the proof of Theorem 2.1, it remains to show that Proposition 2.1 and Proposition 2.2 are true (see Section 2.3 and Section 2.4, respectively), to derive Lemma 2.1 from Proposition 2.1 and Proposition 2.2 (see Section 2.5) and to construct minorants which satisfy the requirements of Proposition 2.3 (see Section 2.6).

2.3 From sums to Dirichlet polynomials

The goal of this section is to prove Proposition 2.1, which roughly states that

$$\sum_{\substack{k_0 \dots k_{r_2} \in \mathcal{A}(y) \\ k_i \sim N_i}} \left(1_{\mathbb{N}}(k_0) \prod_{1 \leq i \leq r_1} 1_{\mathbb{P}}(k_i) \prod_{i > r_1} a_{k_i}^{(i)} \right) - \frac{x^b}{\tau} \sum_{\substack{k_0 \dots k_{r_2} \in \mathcal{B}(y) \\ k_i \sim N_i}} \left(1_{\mathbb{N}}(k_0) \prod_{1 \leq i \leq r_1} 1_{\mathbb{P}}(k_i) \prod_{i > r_1} a_{k_i}^{(i)} \right) \tag{2.16}$$

is $O_A(\frac{x}{\tau(\log x)^A})$ for most $y \in \mathbb{N} \cap [x, 3x]$ and that the number of y for which (2.16) is larger can be bounded by counting how often the values of corresponding Dirichlet polynomials are of a certain size.

Throughout Section 2.3 we work with a fixed $\varepsilon > 0$ and $a \in [0.475 - \varepsilon, 0.77 - \varepsilon]$ and set $\tau = x^a$. We also recall that $\varepsilon_1 = \varepsilon/10^5$, $T_0 = \tau x^{\varepsilon_1}$ and $T_1 = \gamma x^{b/2-\varepsilon_1}$, where $\gamma \in [1, 2]$ is chosen such that $\log(T_0/T_1)/\log 2 \in \mathbb{Z}$ and $[T_1, T_0]$ can be split into dyadic intervals.

2.3.1 Comparison lemma

We begin with a simple consequence of Perron's theorem.

Lemma 2.2. *Let $A \in \mathbb{N}$ and $J \in \mathbb{N}$. Suppose (a_n) is a complex sequence with $|a_n| \leq x^{o(1)}$. Suppose also that $a_n = 0$ unless $2^{-J}x \leq n \leq 2^J x$. Write $F(s) = \sum a_n n^{-s}$. Then for any $y \in [x, 3x]$,*

$$\sum_{n\in\mathcal{A}(y)} a_n - \frac{x^b}{\tau}\sum_{n\in\mathcal{B}(y)} a_n = O\biggl(\biggl|\int_{c+iT_1}^{c+iT_0} y^s C_\tau(s)F(s)\,\mathrm{d}s\biggr|\biggr) + O\biggl(\frac{x}{\tau(\log x)^A}\biggr)$$
$$+ O\biggl(\frac{x^b}{\tau}\biggl|\int_{c+iT_1}^{c+iT_0} y^s C_{x^b}(s)F(s)\,\mathrm{d}s\biggr|\biggr),$$

where $c = 1 + 1/\log x$ and $C_z(s) = ((1+1/z)^s - 1)/s$.
The constants implied by big O notation only depend on A, J and ε.

Proof. Let $I(z) = [y, y + y/z] \cap \mathbb{N}$. We consider $z \in \{\tau, x^b\}$. By Perron's theorem, since $a_n = O(x^{\varepsilon_1/5})$,

$$\begin{aligned}\sum_{n\in I(z)} a_n &= \frac{1}{2\pi i}\int_{c-iT_0}^{c+iT_0} y^s C_z(s)F(s)\,\mathrm{d}s + O\biggl(\frac{x^{1+\varepsilon_1/2}}{T_0}\biggr) \\ &= \frac{1}{2\pi i}\int_{c-iT_1}^{c+iT_1} y^s C_z(s)F(s)\,\mathrm{d}s + O\biggl(\biggl|\int_{c+iT_1}^{c+iT_0} y^s C_z(s)F(s)\,\mathrm{d}s\biggr|\biggr) + O\left(\frac{x}{\tau x^{\varepsilon_1/2}}\right).\end{aligned} \tag{2.17}$$

We further note that

$$\frac{(y+Y)^s - y^s}{s} - y^{s-1}Y = \int_y^{y+Y} t_1^{s-1} - y^{s-1}\,\mathrm{d}t_1 = (s-1)\int_y^{y+Y}\int_y^{t_1} t_2^{s-2}\,\mathrm{d}t_2\,\mathrm{d}t_1$$

and hence obtain $((y+Y)^s - y^s)/s = y^{s-1}Y + O(|s-1|y^{\mathrm{Re}(s)-2}Y^2)$ for $Y = y/z$ with $z \in \{\tau, x^b\}$. Since $a_n = O(x^{\varepsilon_1/5})$ and $T_1 = \gamma x^{b/2-\varepsilon_1}$ with $\gamma \in [1,2]$, we thus have

$$\begin{aligned}\frac{1}{2\pi i}\int_{c-iT_1}^{c+iT_1} y^s C_z(s)F(s)\,\mathrm{d}s &= \frac{1}{2\pi i}\int_{c-iT_1}^{c+iT_1} \frac{y^s}{z}F(s)\,\mathrm{d}s + O\biggl(\frac{T_1^2 x^{1+\varepsilon_1}}{z^2}\biggr) \\ &= \frac{1}{2\pi i z}\int_{c-iT_1}^{c+iT_1} y^s F(s)\,\mathrm{d}s + O\biggl(\frac{x}{z x^{\varepsilon_1}}\biggr).\end{aligned} \tag{2.18}$$

Combining (2.17) and (2.18), once with $z = \tau$ and once with $z = x^b$, and multiplying the latter equation by x^b/τ, we obtain Lemma 2.2. □

Applying Lemma 2.2 to the choice of (a_n) given at the start of Section 2.2.1, we then have to work with error terms $O\left(\left|\int_{c+iT_1}^{c+iT_0} y^s C_z(s)F(s)\,\mathrm{d}s\right|\right)$, where

$$F(s) = \left(\sum_{k_0\sim N_0}\frac{1}{k_0^s}\right)\prod_{i=1}^{r_1}\left(\sum_{k_i\sim N_i}\frac{1_{\mathbb{P}}(k_i)}{k_i^s}\right)\prod_{i>r_1}\left(\sum_{k_i\sim N_i}\frac{a_{k_i}^{(i)}}{k_i^s}\right).$$

But $F(s)$ has long factors with coefficients $1_{\mathbb{P}}(n)$. We do not have good Dirichlet polynomial bounds in this case - instead we would rather work with short factors or factors with smooth coefficients 1 or $\log n$. To fix this problem, we decompose $1_{\mathbb{P}}(n)$ via Heath-Brown's identity.

2.3.2 Heath-Brown's identity

Heath-Brown's identity states the following:

Lemma 2.3. *For $n \leq 6x$ and $K \in \mathbb{N}$*

$$\Lambda(n) = \sum_{j=1}^{K} (-1)^{j+1} \binom{K}{j} \sum_{\substack{\prod_{i=1}^{2j} n_i = n \\ n_i \leq (6x)^{1/K} \text{ for } i \leq j}} \mu(n_1) \dots \mu(n_j) \log(n_{2j}).$$

Proof. This follows from equation (6) of [18]. See also Proposition 17.2 of [9]. □

Corollary 2.4. *Let $z \in \{\tau, x^b\}$, $J, K \in \mathbb{N}$ and $T \in [T_1, T_0]$.*
Suppose $F(s) = \prod_{i=1}^{r} S_i(s)$ where $r \leq J$ and $S_i(s) = \sum_{n \sim N_i} a_n^{(i)} n^{-s}$ for some $a_n^{(i)} \in \mathbb{C}$ with $a_n^{(i)} = O(\tau_J(n) \log(n+3))$ and some $N_i \geq 1/2$ with $2^{-J} x \leq \prod_{i=1}^{j} N_i \leq 2^J x$.
Suppose $S_1(s) = \sum_{n \sim N_1} 1_{\mathbb{P}}(n) n^{-s}$ with $N_1 \geq x^{1/K}$.

Denote by S_k the set of $n \in \mathbb{N}$ with $n = p^k$ for some prime p. Denote by $\mathcal{M}(N, j, K)$ the set of $(M_i)_{i=1}^{2j}$ with $M_i \in \left\{ \frac{(6x)^{1/K}}{2^m} : 1 \leq m \leq \frac{\log((6x)^{1/K})}{\log 2} + 1 \right\}$ for $i \leq j$ and $M_i \in \left\{ \frac{N}{2^m} : 0 \leq m \leq \frac{\log N}{\log 2} + 1 \right\}$ for $i > j$ and $2^{-2K} N \leq \prod_{i=1}^{2j} M_i \leq 2^{2K} N$.

Then there exists $U_1 \in (N_1, 2N_1]$ with

$$\left| \int_{c+iT}^{c+i2T} y^s C_z(s) F(s) \, \mathrm{d}s \right| = O\left(\left| \int_{c+iT}^{c+i2T} y^s C_z(s) \sum_{n \sim N_1} \left(\sum_{k=2}^{\infty} \frac{1_{S_k}(n)}{k n^s} \right) \prod_{i>2} S_i(s) \, \mathrm{d}s \right| \right. \tag{2.19}$$
$$\left. + \sum_{j=1}^{K} \sum_{(M_i) \in \mathcal{M}(N_1, j, K)} \left| \int_{c+iT}^{c+i2T} y^s C_z(s) \left(\sum_{n \sim N_1} \frac{b_n^{(j,(M_i))}}{n^s} \right) \prod_{i>2} S_i(s) \, \mathrm{d}s \right| \right)$$

where

$$b_n^{(j,(M_i))} = \sum_{\substack{\prod_{i=1}^{2j} m_i = n \\ m_i \sim M_i \\ \prod_{i=1}^{2j} m_i \in (N_1, U_1]}} \left(\log(m_{2j}) \prod_{i=1}^{j} \mu(m_i) \prod_{i=j+1}^{2j-1} 1_{\mathbb{N}}(m_i) \right).$$

Here the constant implied by big O notation depends only on K.

Proof. By some rearrangements and partial summation, we have

$$\sum_{n\sim N_1}\frac{1_{\mathbb{P}}(n)}{n^s}=\sum_{n\sim N_1}\frac{\Lambda(n)}{n^s\log n}-\sum_{k=2}^{\infty}\sum_{\substack{n\sim N_1\\ n=p^k\text{ for a prime }p}}\frac{1}{kn^s}$$

$$=\frac{1}{\log(2N_1)}\sum_{n\sim N_1}\frac{\Lambda(n)}{n^s}+\int_{N_1}^{2N_1}\frac{1}{t(\log t)^2}\sum_{\substack{n\sim N_1\\ n\le t}}\frac{\Lambda(n)}{n^s}\,\mathrm{d}t-\sum_{n\sim N_1}\left(\sum_{k=2}^{\infty}\frac{1_{S_k}(n)}{kn^s}\right)$$

Using Heath-Brown's identity, stated in Lemma 2.3, then

$$\sum_{n\sim N_1}\frac{1_{\mathbb{P}}(n)}{n^s}=\frac{1}{\log(2N_1)}\sum_{j=1}^{K}(-1)^{j+1}\binom{K}{j}\sum_{n\sim N_1}\sum_{\substack{\prod_{i=1}^{2j}m_i=n\\ m_i\le(6x)^{1/K}\text{ for }i\le j}}\frac{\mu(m_1)\dots\mu(m_j)\log(m_{2j})}{(m_1\dots m_{2j})^s}$$

$$+\int_{N_1}^{2N_1}\frac{1}{t(\log t)^2}\sum_{j=1}^{K}(-1)^{j+1}\binom{K}{j}\sum_{\substack{n\sim N_1\\ n\le t}}\sum_{\substack{\prod_{i=1}^{2j}m_i=n\\ m_i\le(6x)^{1/K}\text{ for }i\le j}}\frac{\mu(m_1)\dots\mu(m_j)\log(m_{2j})}{(m_1\dots m_{2j})^s}\,\mathrm{d}t$$

$$-\sum_{n\sim N_1}\left(\sum_{k=2}^{\infty}\frac{1_{S_k}(n)}{kn^s}\right).$$

We now substitute this identity into $I=\left|\int_{c+iT}^{c+i2T}y^sC_z(s)F(s)\,\mathrm{d}s\right|$ and split sums over $m_1,\dots,m_{2j}$ up by restricting the ranges of m_i to $m_i\sim M_i$. The quantities M_i satisfy $M_i\in\left\{\frac{(6x)^{\frac{1}{K}}}{2^m}:1\le m\le\frac{\log((6x)^{\frac{1}{K}})}{\log 2}+1\right\}$ for $i\le j$ and $M_i\in\left\{\frac{N_1}{2^m}:0\le m\le\frac{\log(N_1)}{\log 2}+1\right\}$ for $i>j$. The following bound holds:

$$I\ll\sum_{j=1}^{K}\sum_{(M_i)\in\mathcal{M}(N_1,j,K)}\left|\int_{c+iT}^{c+i2T}y^sC_z(s)\sum_{\substack{n\sim N_1\\ \prod_{i=1}^{2j}m_i=n\\ m_i\sim M_i}}\left(\frac{\log(m_{2j})\prod_{i=1}^{j}\mu(m_i)}{(m_1\dots m_{2j})^s}\right)\prod_{i>2}S_i(s)\,\mathrm{d}s\right| \tag{2.20}$$

$$+\int_{N_1}^{2N_1}\frac{1}{t(\log t)^2}\sum_{j=1}^{K}\sum_{(M_i)\in\mathcal{M}}\left|\int_{c+iT}^{c+i2T}y^sC_z(s)\sum_{\substack{n\in(N_1,t]\\ \prod_{i=1}^{2j}m_i=n\\ m_i\sim M_i}}\left(\frac{\log(m_{2j})\prod_{i=1}^{j}\mu(m_i)}{(m_1\dots m_{2j})^s}\right)\prod_{i>2}S_i(s)\,\mathrm{d}s\right|\mathrm{d}t \tag{2.21}$$

$$+\left|\int_{c+iT}^{c+i2T}y^sC_z(s)\sum_{n\sim N_1}\left(\sum_{k=2}^{\infty}\frac{1_{S_k}(n)}{kn^s}\right)\prod_{i>2}S_i(s)\,\mathrm{d}s\right|. \tag{2.22}$$

Now choose $U_1 \in (N_1, 2N_1]$ so that

$$\sum_{j=1}^{K} \sum_{(M_i)\in\mathcal{M}(N_1,j,K)} \left| \int_{c+iT_1}^{c+iT_0} y^s C_z(s) \sum_{\substack{n\in(N_1,U_1] \\ \prod_{i=1}^{2j} m_i = n \\ m_i \sim M_i}} \left(\frac{\log(m_{2j}) \prod_{i=1}^{j} \mu(m_i)}{(m_1 \dots m_{2j})^s} \right) \prod_{i>2} S_i(s) \, \mathrm{d}s \right|$$

is maximized. Then the RHS of (2.20), (2.21) and (2.22) each is bounded above by the RHS of (2.19). □

2.3.3 Removal of inequality conditions

Applying Lemma 2.2 to the choice of (a_n) given at the start of Section 2.2.1 and repeatedly applying Corollary 2.4 to the error terms, we eventually end up working with bounds $O\left((\log x)^C \left|\int_{c+iT_1}^{c+iT_0} y^s C_z(s) F(s) \, \mathrm{d}s\right|\right)$ where $F(s)$ factorises as $F(s) = \prod_{i=1}^{j} S_i(s)$ for some $S_i(s)$ which are either short or have property (Z1) or (Z2) (with $w = 0$) or which are of the form

$$S_i(s) = \sum_{\substack{\prod_{k=1}^{2j_i} m_{i,k} \in (N_i, U_i] \\ m_{i,k} \sim M_{i,k}}} \left(\frac{\log(m_{i,2j_i})}{m_{i,2j_i}^s} \prod_{k=1}^{j_i} \frac{\mu(m_{i,k})}{m_{i,k}^s} \prod_{k=j_i+1}^{2j_i-1} \frac{1_{\mathbb{N}}(m_{i,k})}{m_{i,k}^s} \right). \tag{2.23}$$

If we could omit the condition $\prod_{k=1}^{2j_i} m_{i,k} \in (N_i, U_i]$, then $S_i(s)$ would factorise very nicely. The goal of this section is thus to remove certain inequality conditions which prevent further factorisation of a Dirichlet polynomial. More precisely, we will prove the following result:

Lemma 2.5. *Let $F(s) = \prod_{i=1}^{j} S_i(s)$ where $j \le J$ and where $S_i(s) = \sum_{n\sim N_i} a_n^{(i)} n^{-s}$ for some $a_n^{(i)} \in \mathbb{C}$ with $a_n^{(i)} = O(\tau_J(n) \log(n+3))$ and some $N_i \ge 1/2$ with $2^{-J}x \le \prod_{i=1}^{j} N_i \le 2^J x$.*

Suppose there exist $\ell \in \{1, \dots, j\}$, $U_\ell \in (N_\ell, 2N_\ell]$, $b_n^{(i)} \in \mathbb{C}$ with $b_n^{(i)} = O(\log(n+3))$ for $1 \le i \le r \le J$ and (M_i) with $2^{-J}x^{1/J} \le 2^{-J}N_\ell \le \prod_{i=1}^{r} M_i \le 2^J N_\ell$ such that

$$a_n^{(\ell)} = \sum_{\substack{n=k_1\dots k_r \\ k_i \sim M_i \\ k_1\dots k_r \in (N_\ell, U_\ell]}} b_{k_1}^{(1)} \dots b_{k_r}^{(r)}.$$

Let $T \in [T_1, T_0]$. Suppose also that there exists $i_0 \neq \ell$ such that $N_{i_0} \ge x^{1/\log\log x}$ and such that $S_{i_0}(s)$ satisfies property (D) with respect to T.

Let $z \in \{\tau, x^b\}$, $C_z(s) = ((1+1/z)^s - 1)/s$ and $c_0 = 1/\log x$. Then

$$\left|\int_{c+iT}^{c+i2T} y^s C_z(s) F(s)\,\mathrm{d}s\right|$$
$$= O\left(\int_{-\frac{T}{4}}^{\frac{T}{4}}\int_{-\frac{T}{4}}^{\frac{T}{4}} \frac{1}{|\frac{c_0}{4}+it_1|}\frac{1}{|\frac{c_0}{4}+it_2|}\left|\int_{c+iT}^{c+i2T} y^s C_z(s) F^*(s;t_1-t_2)\,\mathrm{d}s\right| \mathrm{d}t_1\mathrm{d}t_2\right)$$
$$+ O\left(\frac{x}{z(\log x)^A}\right),$$
$$\text{where } F^*(s;t) = \prod_{n=1}^{r}\left(\sum_{k_n\sim M_n} b_{k_n}^{(n)} k_n^{-s-it}\right)\prod_{i\neq \ell} S_i(s).$$

The constants implied by big O notation depend only on J and ε.

2.3.3.1 Large values of Dirichlet polynomials

We first record upper bounds on the size of Dirichlet polynomials which satisfy property (D). These results will be very important when we bound error terms in the proof of Lemma 2.5 and will also make multiple other appearances in Section 2.4.

Lemma 2.6. *Let $T \in [T_1, T_0]$ and $t \in [T, 2T]$. Suppose $S(s) = \sum_{n\sim N} a_n n^{-s}$ for some $N \in [x^{1/\log\log x}, 6x]$ and $S(s)$ has property (D) with respect to T. Let $c = 1 + 1/\log x$ and $\mu = (\log x)^{-7/10}$. Then $|S(c+it)| \ll N^{-\mu}$ and, for $D > 0$,*

$$|S(c+it)| \ll_D \frac{1}{(\log x)^D}.$$

Proof. Since $S(s)$ has property (D) with respect to T, there exists $w \in [-T/2, T/2]$ with $S(s+iw) = \sum_{n\sim N} b_n n^{-s}$ such that $b_n = 1$ or $b_n = \log n$ or $b_n = \mu(n)$ or $b_n = 1_{\mathbb{P}}(n)$ or $b_n = \frac{\Lambda(n)}{\log n} - 1_{\mathbb{P}}(n)$ for all n.

We first consider the case where

$$S(c+it) = \sum_{n\sim N}\frac{1}{n^{c+i(t-w)}},\quad S(c+it) = \sum_{n\sim N}\frac{\log n}{n^{c+i(t-w)}}\quad \text{or}\quad S(c+it) = \sum_{n\sim N}\frac{\mu(n)}{n^{c+i(t-w)}}.$$

Note $t - w \in \left[\frac{T}{2}, \frac{5T}{2}\right]$. Let $\mu_1 = \frac{(\log x)^{-2/3}}{(\log\log x)^{1/3}}$. By the proof of Lemma 5.5 of [29],

$$|S(c+it)| \ll (\log x)^3 (N^{-2C_0\mu_1} + T_1^{-1})$$

for a suitable constant $C_0 > 0$. This is a consequence of the Vinogradov-Korobov estimate. Then since $N \geq x^{1/\log\log x}$, also $N^{-C_0\mu_1} \leq x^{-C_0((\log x)^{-2/3}(\log\log x)^{-4/3})} \leq x^{-3(\log\log x)(\log x)^{-1}} = (\log x)^{-3}$ and so $|S(c+it)| \ll N^{-C_0\mu_1}$. Since $\mu \leq C_0\mu_1$ for large x, further $|S(c+it)| \ll N^{-\mu}$.

Now say $S(c+it) = \sum_{n\sim N} 1_{\mathbb{P}}(n) n^{-c-i(t-w)}$. Lemma 1.5 of [14] and $N \geq x^{\frac{1}{\log\log x}}$ give

$$|S(c+it)| \ll \exp\left(-\frac{\log N}{\log(5T/2)^{7/10}}\right) + \frac{(\log x)^3}{T} \ll N^{-(\log x)^{-7/10}} = N^{-\mu}.$$

Finally, if $S(c+it) = \sum_{n\sim N}(\frac{\Lambda(n)}{\log n} - 1_{\mathbb{P}}(n))n^{-c-i(t-w)} = \sum_{k=2}^{\infty}\sum_{p^k\sim N}(\frac{1}{k})n^{-c-i(t-w)}$, we have $|S(c+it)| \ll (\log N)N^{-1/2} \ll N^{-\mu}$. We conclude the proof with observation

$$N^{-\mu} \ll \exp\left(-\frac{(\log x)^{3/10}}{\log\log x}\right) \ll_D \exp\left(-D\log\log x\right) = \frac{1}{(\log x)^D}. \qquad \square$$

2.3.3.2 Comparison of integrals of Dirichlet polynomials

Now we give a proof of Lemma 2.5.

Proof of Lemma 2.5. Recall that we wish to remove condition $k_1 \dots k_r \in (N_\ell, U_\ell]$ from $S_\ell(s)$. To do so, we use the following identity, a precursor to Perron's formula:

$$\frac{1}{2\pi i}\int_{(c_0-iT)/4}^{(c_0+iT)/4} \frac{y^w}{w}\,\mathrm{d}w = H(y) + O\left(\frac{y^{c_0/4}}{T|\log y|}\right) \quad \text{where} \quad H(y) = \begin{cases} 1 & \text{if } y > 1, \\ 0 & \text{if } 0 < y < 1. \end{cases} \tag{2.24}$$

We first apply (2.24) with $y = (\lfloor U_\ell \rfloor + 1/2)/(\prod_{i=1}^r k_i)$. For $k_i \sim M_i$ we have $\prod_{i=1}^r k_i \in [2^{-2J}N_\ell, 2^{2J}N_\ell]$, while $U_\ell \in [N_\ell, 2N_\ell]$. So $y \in [2^{-3J}, 1 - 1/(2\lfloor U_\ell\rfloor + 2)]$ or $y \in [1 + 1/(2\lfloor U_\ell \rfloor), 2^{3J}]$ and $y^{c_0/4} = O(1)$. Combining error terms $O(1/(T|\log y|)$ via standard summation arguments, for $s = c+it$,

$$\begin{aligned} S_\ell(s) &= \sum_{\substack{k_i \sim M_i \\ k_1\dots k_r \in (N_\ell, U_\ell]}} \frac{b_{k_1}^{(1)}\dots b_{k_r}^{(r)}}{k_1^s \dots k_r^s} \\ &= \frac{1}{2\pi i}\sum_{\substack{k_i\sim M_i \\ k_1\dots k_r > N_\ell}} \frac{b_{k_1}^{(1)}\dots b_{k_r}^{(r)}}{k_1^s\dots k_r^s}\left\{\int_{(c_0-iT)/4}^{(c_0+iT)/4} \frac{(\lfloor U_\ell\rfloor + 1/2)^w}{w(\prod_{i=1}^r k_i)^w}\,\mathrm{d}w \right. \\ &\qquad\qquad \left. + O\left(\left|T\log\left(\frac{\lfloor U_\ell \rfloor + 1/2}{\prod_{i=1}^r k_i}\right)\right|^{-1}\right)\right\} \\ &= \frac{1}{2\pi i}\int_{(c_0-iT)/4}^{(c_0+iT)/4} \frac{(\lfloor U_\ell\rfloor + 1/2)^w}{w}\left(\sum_{\substack{k_i\sim M_i \\ k_1\dots k_r > N_\ell}} \frac{b_{k_1}^{(1)}\dots b_{k_r}^{(r)}}{k_1^{s+w}\dots k_r^{s+w}}\right)\mathrm{d}w \qquad (2.25) \\ &+ O\left(\sum_{k_l\sim M_l} \frac{(\log x)^J}{TM_1\dots M_r} + \sum_{i=1}^r \sum_{\substack{k_l\sim M_l \\ l\neq i}} \frac{M_i(\log x)^{J+1}}{TM_1\dots M_r}\right). \qquad (2.26) \end{aligned}$$

For fixed $t_1 \in [-T/4, T/4]$ we define Dirichlet polynomial

$$S_\ell^{(1)}(s;t_1) = \sum_{\substack{k_i \sim M_i \\ k_1\dots k_r > N_\ell}} \frac{b_{k_1}^{(1)} \dots b_{k_r}^{(r)}}{(k_1 \dots k_r)^{s+c_0/4+it_1}}.$$

Substituting (2.25) and (2.26) for $S_\ell(s)$ and changing the order of integration, then

$$\begin{aligned}
&\left| \int_{c+iT}^{c+i2T} y^s C_z(s) F(s) \, \mathrm{d}s \right| \\
&= O\left(\left| \int_{c+iT}^{c+i2T} \int_{-\frac{T}{4}}^{\frac{T}{4}} \frac{(\lfloor U_\ell \rfloor + 1/2)^{\frac{c_0}{4}+it_1}}{(\frac{c_0}{4} + it_1)} y^s C_z(s) S_\ell^{(1)}(s;t_1) \prod_{i \neq \ell} S_i(s) \, \mathrm{d}t_1 \, \mathrm{d}s \right| \right) \\
&+ O\left(\int_{c+iT}^{c+i2T} \left| y^s C_z(s) \prod_{i \neq \ell} S_i(s) \left(\frac{(\log x)^{J+1}}{T} \right) \right| \mathrm{d}s \right) \\
&= O\left(\int_{-\frac{T}{4}}^{\frac{T}{4}} \frac{1}{|\frac{c_0}{4} + it_1|} \left| \int_{c+iT}^{c+i2T} y^s C_z(s) S_\ell^{(1)}(s;t_1) \prod_{i \neq \ell} S_i(s) \, \mathrm{d}s \right| \mathrm{d}t_1 \right) && (2.27) \\
&+ O\left(\int_{c+iT}^{c+i2T} \left| y^s C_z(s) \prod_{i \neq \ell} S_i(s) \left(\frac{(\log x)^{J+1}}{T} \right) \right| \mathrm{d}s \right). && (2.28)
\end{aligned}$$

To treat the error term (2.28), we recall that there exists $i_0 \neq \ell$ such that $N_{i_0} \geq x^{1/\log\log x}$ and $S_{i_0}(s)$ satisfies property (D) with respect to T. By Lemma 2.6, then

$$|S_{i_0}(s)| = \left| \sum_{n \sim N_{i_0}} \frac{a_n^{(i_0)}}{n^{c+it}} \right| \ll_D \frac{1}{(\log x)^D} \tag{2.29}$$

for any $D > 0$ and $t \in [T, 2T]$. Also, $|y^s| = O(x)$ and $C_z(s) = O(1/z)$. Since $a_n^{(i)} = O(\tau_J(n)\log(n+3))$, we have $|S_i(c+it)| = O(\log(8N_i)^J)$ by Shiu's theorem [40]. Taking $D = J^2 + J + A + 1$ in (2.29), error term (2.28) is thus bounded by $O(\frac{x}{z(\log x)^A})$.

The main term (2.27) still involves $S_\ell^{(1)}(s;t_1)$, which is a sum over $k_i \sim M_i$ with restriction $k_1 \dots k_r > N_\ell$. Repeating the above argument, but choosing $y = \frac{\prod_{i=1}^r k_i}{\lfloor N_\ell \rfloor + \frac{1}{2}}$, we also remove this condition and upper bound (2.27) by $O(\frac{x}{z(\log x)^A})$ plus

$$O\left(\int_{-\frac{T}{4}}^{\frac{T}{4}} \int_{-\frac{T}{4}}^{\frac{T}{4}} \prod_{j=1}^{2} \frac{1}{|\frac{c_0}{4} + it_j|} \left| \int_{c+iT}^{c+i2T} y^s C_z(s) \prod_{n=1}^{r} \left(\sum_{k_n \sim M_n} \frac{b_{k_n}^{(n)}}{k_n^{s+i(t_1-t_2)}} \right) \prod_{i \neq \ell} S_i(s) \, \mathrm{d}s \right| \mathrm{d}\mathbf{t} \right). \quad \square$$

2.3.4 Summary and sampling

Next we combine our results from Section 2.3.1, Section 2.3.2 and Section 2.3.3, giving a first bound on the number of $y \in \mathbb{N} \cap [x, 3x]$ for which difference (2.16) is large.

This bound involves the integral of a Dirichlet polynomial which factorises nicely into short factors and long factors with smooth coefficients and another Dirichlet polynomial which has no more than T_0 non-zero coefficients. Both of these properties will be very important for further bounds in Section 2.4.

Lemma 2.7. *Let $A \in \mathbb{N}$ and $K \in \mathbb{N}$ and $0 \leq r_1 \leq r_2 \leq 10^5$. Let $J \geq 10^6 K$. Let $N_i \in [\frac{1}{2}, \infty)$ with $2^{-r_2}x \leq \prod_{i=0}^{r_2} N_i \leq 6x$, $N_i < x^{1/K}$ for $i > r_1$ and $N_i \geq x^{1/K}$ for $1 \leq i \leq r_1$. Let $a_n^{(i)} \in \mathbb{C}$ with $a_n^{(i)} = O(\log(n+3))$ and consider (a_n) with*

$$a_n = \sum_{\substack{k_0 k_1 \ldots k_{r_2} = n \\ k_i \sim N_i}} \left(1_{\mathbb{N}}(k_0) \prod_{1 \leq i \leq r_1} 1_{\mathbb{P}}(k_i) \prod_{i > r_1} a_{k_i}^{(i)} \right).$$

Suppose also that one of the following three options holds:

(i) $r_1 \geq 2$ *or* $N_0 \geq x^{0.95}$.

(ii) $r_1 = 1$ *and* $N_0 \geq x^{1/\log\log x}$.

(iii) $r_1 \in \{0, 1\}$ *and* $\exists i > r_1$ *with* $a_n^{(i)} = 1_{\mathbb{P}}(n)$ *for all* $n \in \mathbb{N}$ *and* $N_i \geq x^{1/\log\log x}$.

Let $\mathcal{I}$ denote the set of $y \in \mathbb{N} \cap [x, 3x]$ for which

$$\left| \sum_{n \in \mathcal{A}(y)} a_n - \frac{x^b}{\tau} \sum_{n \in \mathcal{B}(y)} a_n \right| > \frac{x}{\tau(\log x)^A}.$$

Then if $x \geq C$ and $\#\mathcal{I} > 10\tau x^{0.23}$, there exist $F \in \mathcal{F}((N_i), T, J, K)$, $G(s) = \sum_{n \in \mathcal{J}} \frac{\xi_n}{n^s} \in \mathcal{G}(\tau, \varepsilon)$, $z \in \{\tau, x^b\}$, $T \in [T_1, T_0]$ and $u \in \{1, \ldots, (2\log x)^{10^6 K}\}$ with

$$\left| \int_T^{2T} \left(\frac{C_z(c+it)}{H^{-it}u^{c+it}} \right) \overline{G(it)} F(c+it) \mathrm{d}t \right| \geq \frac{\#\mathcal{J}}{z(\log x)^{A + 5 \cdot 10^6 K}}$$

and $\#\mathcal{I} \leq 10(x/T_0)\#\mathcal{J}$.

Here the required size of C depends only on A, J, K and ε.

Recall here that sets $\mathcal{F}((N_i), T, J, K)$ and $\mathcal{G}(\tau, \varepsilon)$ were defined in Definition 2.2.

Proof. First note that $y \in \mathcal{I}$ if and only if $y \in \mathbb{N} \cap [x, 3x]$ and

$$\mathrm{Diff}(\tau, (a_n), y) = \left| \sum_{n \in \mathcal{A}(y)} a_n - \frac{x^b}{\tau} \sum_{n \in \mathcal{B}(y)} a_n \right| > \frac{x}{\tau(\log x)^A}.$$

Set $H = \lceil 3x/T_0 \rceil$ and let $\mathcal{I}^*$ be the set of $y \in H\mathbb{Z} \cap [x, 3x]$ with $\mathrm{Diff}(\tau, (a_n), y) > x/(\tau(\log x)^{A+1})$. Consider $Hm + d$, where $Hm \in H\mathbb{Z} \cap [x, 3x]$ and $|d| \leq H - 1$. Using the J-fold divisor bound of Shiu [40], noting that $a_n = O(\tau_{K+1}(n)(\log x)^K)$ and recalling $H \leq 4x/(\tau x^{\varepsilon_1})$, we have

$$\begin{aligned}
&|\mathrm{Diff}(\tau, (a_n), Hm) - \mathrm{Diff}(\tau, (a_n), Hm + d)| \\
&\ll \sum_{n \in \mathcal{A}(Hm) \triangle \mathcal{A}(Hm+d)} |a_n| + \frac{x^b}{\tau} \sum_{n \in \mathcal{B}(Hm) \triangle \mathcal{B}(Hm+d)} |a_n| \\
&\ll \frac{x}{\tau x^{\varepsilon_1}} (\log x)^{2K} \ll \frac{x}{\tau (\log x)^{A+1}}.
\end{aligned}$$

Hence if $Hm + d \in \mathcal{I}$, also $\mathrm{Diff}(\tau, (a_n), Hm) > x/(\tau(\log x)^{A+1})$ and $Hm \in \mathcal{I}^*$. Conversely, $Hm \notin \mathcal{I}^*$ implies $Hm + d \notin \mathcal{I}$ for $|d| \leq H - 1$ and thus $\#\mathcal{I} \leq 2H\#\mathcal{I}^*$. Write $\mathcal{J} = \{m \in \mathbb{N} : Hm \in \mathcal{I}^*\} \subseteq [1, T_0]$.

For each $y \in \mathcal{I}^*$, we now apply Corollary 2.4 to (a_n) (with $A+2$ in the place A). The resulting error terms involve integrals over $[c + iT_1, c + iT_0]$, but T_1 was chosen so that interval $(T_1, T_0]$ can be split exactly into $O(\log x)$ intervals $(T, 2T]$. If x is sufficiently large, there must exist $z_y \in \{\tau, x^b\}$ and $T_y \in \{\frac{T_0}{2^n} : 1 \leq n \leq \frac{\log(T_0/T_1)}{\log 2}\}$ with

$$\frac{(x/\tau)}{(\log x)^{A+3}} < \frac{z_y}{\tau} \left| \int_{c+iT_y}^{c+i2T_y} y^s C_{z_y}(s) \Bigg(\sum_{k_0 \sim N_0} \frac{1_{\mathbb{N}}(k_0)}{k_0^s} \Bigg) \prod_{i=1}^{r_1} \Bigg(\sum_{k_i \sim N_i} \frac{1_{\mathbb{P}}(k_i)}{k_i^s} \Bigg) \prod_{i > r_1} \Bigg(\sum_{k_i \sim N_i} \frac{a_{k_i}^{(i)}}{k_i^s} \Bigg) \mathrm{d}s \right|. \tag{2.30}$$

Next we alternate between applying Corollary 2.4 and Lemma 2.5 to the RHS of (2.30), removing coefficients $1_{\mathbb{P}}(k_i)$ for $1 \leq i \leq r_1$. The first application of Lemma 2.5 is possible because options (i), (ii) and (iii) imply that the relevant Dirichlet polynomial has a factor $\sum_{n \sim N_0} \frac{1}{n^s}$ with $N_0 \geq x^{\frac{1}{\log \log x}}$ or $\sum_{n \sim N_i} \frac{1_{\mathbb{P}}(n)}{n^s}$ with $N_i \geq x^{\frac{1}{\log \log x}}$. This factor satisfies property (D) with respect to T (with $w = 0$). Each application of Lemma 2.5 gives another factor with property (D), allowing for more applications of Lemma 2.5. By Corollary 2.4 and Lemma 2.5, $\frac{x}{z_y (\log x)^{A+4}}$ is bounded by

$$\sum_{\substack{I \in \mathcal{P}(r_1) \\ j_k \in \{1, \ldots, K\} \\ (M_{k,i}) \in \mathcal{M}(N_k, j_k, K)}} \int_{-\frac{T}{4}}^{\frac{T}{4}} \cdots \int_{-\frac{T}{4}}^{\frac{T}{4}} \prod_{k \in I} \frac{1}{|\frac{c_0}{4} + it_k|} \frac{1}{|\frac{c_0}{4} + is_k|} \left| \int_{c+iT_y}^{c+i2T_y} y^s C_{z_y}(s) F^*(s; \mathbf{t} - \mathbf{s}, I, (j_k), (M_{k,i}))\, \mathrm{d}s \right| \mathrm{d}\mathbf{t}\mathrm{d}\mathbf{s},$$

where $F^*(s; \mathbf{t}, I, \ldots) = \Bigg(\sum_{k_0 \sim N_0} \frac{1}{k_0^s} \Bigg) \prod_{\substack{k=1 \\ k \in I}}^{r_1} S^*(s; \mathbf{t}, j_k, (M_{k,i})) \prod_{\substack{k=1 \\ k \notin I}}^{r_1} \sum_{n \sim N_k} \Bigg(\sum_{\ell=2}^{\infty} \frac{1_{S_\ell}(n)}{\ell n^s} \Bigg) \prod_{i > r_1} \Bigg(\sum_{k_i \sim N_i} \frac{a_{k_i}^{(i)}}{k_i^s} \Bigg)$

and $S^*(s; \mathbf{t}, \ldots) = \prod_{i=1}^{j_k} \Bigg(\sum_{n_i \sim M_{k,i}} \frac{\mu(n_i)}{n_i^{s + i\sum_{\ell \in I} t_\ell}} \Bigg) \prod_{i=j_k+1}^{2j_k - 1} \Bigg(\sum_{n_i \sim M_{k,i}} \frac{1}{n_i^{s + i\sum_{\ell \in I} t_\ell}} \Bigg) \Bigg(\sum_{n_{2j_k} \sim M_{k,2j_k}} \frac{\log(n_{2j_k})}{n_i^{s + i\sum_{\ell \in I} t_\ell}} \Bigg).$

Here $\mathcal{P}(r_1)$ is the power set of $\{1,\dots,r_1\}$, $\mathcal{M}(N,j,K)$ is as described in Corollary 2.4 and we integrate with respect to variables t_i and s_i with $i\in I$, using $\mathbf{t}$ as shorthand for $(t_i)_{i\in I}$ and d$\mathbf{t}$ as shorthand for $\prod_{i\in I} d\mathbf{t_i}$. To save space, $\dots$ was used once as shorthand for $(j_k),(M_{k,i})$ and once for j_k, $(M_{k,i})$.

The number of possible combinations of I, (j_k) and $(M_{k,i})$ in the above sum is $O((\log x)^{10^6K})$. There are two choices for z_y and $O(\log x)$ choices for T_y. Summing over all $y\in\mathcal{I}^*$, we thus find that there exist $I\subseteq\{1,\dots,r_1\}$, $j_k\in\{1,\dots,K\}$, $(M_{k,i})\in\mathcal{M}(N_k,j_k,K)$ (for $k\in I$), $z\in\{\tau,x^b\}$ and $T\in[T_1,T_0]$ with $\frac{x\#\mathcal{I}^*}{z(\log x)^{A+10^6K+6}}$

$$\leq \sum_{y\in\mathcal{I}^*}\int_{-\frac{T}{4}}^{\frac{T}{4}}\dots\int_{-\frac{T}{4}}^{\frac{T}{4}}\prod_{k\in I}\frac{1}{|\frac{c_0}{4}+it_k|}\frac{1}{|\frac{c_0}{4}+is_k|}\left|\int_{c+iT}^{c+i2T} y^sC_z(s)F^*(s;\mathbf{t}-\mathbf{s},I,(j_k),(M_{k,i}))\mathrm{d}s\right|\mathrm{d}\mathbf{t}\mathrm{d}\mathbf{s}.$$

Now choose $t_i,s_i\in[-T/4,T/4]$ such that $\mathbf{u}=(t_i-s_i)_{i\in I}$ maximises

$$\sum_{y\in\mathcal{I}^*}\left|\int_{c+iT}^{c+i2T} y^sC_z(s)F^*(s;\mathbf{u},I,(j_k),(M_{k,i}))\,\mathrm{d}s\right|.$$

Recalling that $c_0=1/\log x$ and $\#I\leq r_1\leq 10^5$, we get

$$\frac{x\#\mathcal{I}^*}{z(\log x)^{A+10^6K+10^6}}\leq\sum_{y\in\mathcal{I}^*}\left|\int_{c+iT}^{c+i2T} y^sC_z(s)F^*(s;\mathbf{u},I,(j_k),(M_{k,i}))\,\mathrm{d}s\right|. \tag{2.31}$$

In our definition of $F^*(s;\mathbf{u},I,(j_k),(M_{k,i}))$, we wrote $F^*(s)$ as a product $\prod_{i=1}^j S_i(s)$ where $j\leq 10^6K$ and $S_i(s)=\sum_{n\sim B_i} b_n^{(i)}n^{-s}$ with $b_n^{(i)}=O(\log(n+3))$ and $2^{-J}x\leq\prod_{i=1}^j B_i\leq 2^Jx$. The factorisation was chosen so that $B_\ell\geq x^{2/K}$ implies that $S_\ell(s)$ has property $(Z1)$ or $(Z2)$ with respect to T and so that there exist disjoint subsets $X_1,\dots,X_{r_1}$ of $\{1,\dots,j\}$ with $\#X_\ell\leq 2K$ and

$$2^{-2K}N_\ell\leq\prod_{i\in X_\ell}B_i\leq 2^{2K}N_\ell\quad\text{for}\quad 1\leq\ell\leq r_1,$$

and so that one of the conditions (A), (B) or (C) holds. However, we might have factors $S_i(s)$ with $B_i<\log x$ and so $F^*(s;\mathbf{u},I,(j_k),(M_{k,i}))$ is not necessarily in $\mathcal{F}((N_i),T,J,K)$.

Hence write $E=\{i\in\{1,\dots,j\}:B_i<\log x\}$ and define

$$F_1^*(s;\mathbf{u},I,(j_k),(M_{k,i}))=\prod_{i\in\{1,\dots,j\}\setminus E}S_i(s)\quad\text{and}\quad F_2^*(s;\mathbf{u},I,(j_k),(M_{k,i}))=\prod_{i\in E}S_i(s).$$

There still exist disjoint subsets $Y_1, \dots, Y_{r_1}$ of $\{1, \dots, j\} \setminus E$ with

$$(\log x)^{-2K-1} N_\ell \le \prod_{i \in Y_\ell} B_i \le (\log x)^{2K+1} N_\ell \quad \text{for} \quad 1 \le \ell \le r_1$$

and now $F_1^*(s; \mathbf{u}, I, (j_k), (M_{k,i})) \in \mathcal{F}((N_i), T, J, K)$ for $J \ge 10^6 K$.

On the other hand, write $F_2^*(s) = \sum d_n n^{-s}$. We have $d_n = 0$ for $n \ge (2\log x)^{10^6 K}$ and $d_n = O(\tau_{10^6 K}(n)(\log n)^{10^6 K})$ for $n < (2\log x)^{10^6 K}$, so that $d_n = (\log x)^{o(1)}$. Continuing on from (2.31), there is some $u \in \{1, \dots, (2\log x)^{10^6 K}\}$ with

$$\frac{x \# \mathcal{I}^*}{z(\log x)^{A+2\cdot 10^6 K + 10^6 + 1}} \le \sum_{y \in \mathcal{I}^*} \left| \int_{c+iT}^{c+i2T} y^s u^{-s} C_z(s) F_1^*(s; \mathbf{u}, I, (j_k), (M_{k,i}))\, \mathrm{d}s \right|. \tag{2.32}$$

Write $F(s) = F_1^*(s; \mathbf{u}, I, (j_k), (M_{k,i})) \in \mathcal{F}((N_i), T, J, K)$. Now recall that $m \in \mathcal{J} \subseteq [1, T_0]$ if and only if $Hm \in \mathcal{I}^*$. Suppose first that $\#\mathcal{J} < \tau^2 x^{-0.77}$. Then $\#\mathcal{I} \le 2H\#\mathcal{J} < 10\tau x^{0.23}$. Hence we now assume $\#\mathcal{J} \ge \tau^2 x^{-0.77}$. For each $m \in \mathcal{J}$ choose $\xi_m \in \mathcal{C}$ with $|\xi_m| = 1$ and

$$\overline{\xi_m} \left(\int_{c+iT}^{c+i2T} (Hm)^s u^{-s} C_z(s) F(s)\, \mathrm{d}s \right) = \left| \int_{c+iT}^{c+i2T} (Hm)^s u^{-s} C_z(s) F(s)\, \mathrm{d}s \right|.$$

Then set $G(s) = \sum_{m \in \mathcal{J}} \xi_m m^{-s} \in \mathcal{G}(\tau, \varepsilon)$. Since $Hm \in [x, 3x]$, inequality (2.32) gives

$$\frac{\#\mathcal{I}^*}{z(\log x)^{A+2\cdot 10^6 K + 10^6 + 2}} \le \int_T^{2T} \left(\frac{C_z(c+it)}{H^{-it} u^{c+it}} \right) \overline{G(it)} F(c+it)\, \mathrm{d}t. \qquad \square$$

2.3.5 Introduction of R, R^* and Q

Now we complete the proof of Proposition 2.1. Let $\mathcal{I}$ denote the set of $y \in \mathbb{N} \cap [x, 3x]$ for which

$$\left| \sum_{n \in \mathcal{A}(y)} a_n - \frac{x^b}{\tau} \sum_{n \in \mathcal{B}(y)} a_n \right| > \frac{x}{\tau (\log x)^A}.$$

Recall that we wish to show $\#\mathcal{I} = O(\tau x^{0.23+\varepsilon/2})$ for (a_n) as given in Section 2.2.1. By Lemma 2.7, it is enough to show the following: If $F(s) = \prod_{i=1}^{j} S_i(s) \in \mathcal{F}((N_i), T, A + 10^7 K, K)$, $G(s) = \sum_{n \in \mathcal{J}} \xi_n n^{-s} \in \mathcal{G}(\tau, \varepsilon)$, $T \in [T_1, T_0]$ and $u \in \{1, \dots, (2\log x)^{10^6 K}\}$ satisfy

$$\left| \int_T^{2T} \left(\frac{C_z(c+it)}{H^{-it} u^{c+it}} \right) \overline{G(it)} F(c+it) \mathrm{d}t \right| \ge \frac{\#\mathcal{J}}{z(\log x)^{A+5\cdot 10^6 K}}$$

and $F(s)$, $G(s)$ and T satisfy one of the conditions (C1), (C2), (C3), (C4.A) or (C4.B) of Proposition 2.1 for every choice of $\gamma \in (-\infty, 1)$ and $\sigma_i \in \mathbb{R}$ and x is large, then $\#\mathcal{J} \leq \tau^2 x^{0.23-1+\varepsilon/2}$. This implies $10(x/T_0)\#\mathcal{J} \leq \tau x^{0.23+\varepsilon/2}$ and hence $\#\mathcal{I} \leq \tau^2 x^{0.23-1+\varepsilon/2}$.

Proof of Proposition 2.1. We write $J = A + 10^7 K$ and consider $F(s) = \prod_{i=1}^{j} S_i(s) \in \mathcal{F}((N_i)_{i=0}^{r_1}, T, J, K)$, $G(s) = \sum_{n\in\mathcal{J}} \xi_n n^{-s} \in \mathcal{G}(\tau, \varepsilon)$, $z \in \{\tau, x^b\}$, $T \in [T_1, T_0]$ and $u \in \{1, \ldots, (2\log x)^{2K^2}\}$ which satisfy

$$\left|\int_T^{2T} D_z(c+it)\overline{G(it)}F(c+it)\mathrm{d}t\right| \geq \frac{\#\mathcal{J}}{z(\log x)^{A+5\cdot 10^6 K}} \text{ where } D_z(c+it) = \frac{C_z(c+it)}{H^{-it}u^{c+it}}. \tag{2.33}$$

We assume that $F(s)$, $G(s)$ and T satisfy one of the conditions (C1), (C2), (C3), (C4.A) or (C4.B) of Proposition 2.1 for each choice of (σ_i) and γ. Recall here that we defined $\mathcal{S}_1(F, (S_i), (\sigma_i))$, $R(F, (S_i), (\sigma_i), T)$, $\mathcal{S}_1^*(F, (S_i), (\sigma_i))$, $R^*(F, (S_i), (\sigma_i), T)$, $\mathcal{S}_2(G, \gamma)$ and $Q(F, G, (S_i), (\sigma_i), \gamma, T)$ in Definition 2.3.

Since $S_i(s) = \sum_{n\sim B_i} b_n^{(i)} n^{-s}$ with $b_n^{(i)} = O(\log n)$, $|S_i(c+it)| \leq \log(2B_i)B_i^{-c+1}$. So if $\sigma_\ell \geq 1 + \varepsilon_1$ for some ℓ, then $\mathcal{S}_1(F, (S_i), (\sigma_i)) = \emptyset$. Let

$$\Theta_i = \left\{\frac{-2\log x + k\log 2}{\log B_i} : 1 \leq k < \frac{2\log x}{\log 2} + \frac{(1+\varepsilon_1)\log B_i}{\log 2}\right\}.$$

Since $|G(it)| \leq \#\mathcal{J}$, $\mathcal{S}_2(G, \gamma) = \emptyset$ when $\gamma \geq 1$. Write

$$\Theta_L = \left\{\frac{-2\log x + k\log 2}{\log(\#\mathcal{J})} : 1 \leq k < \frac{2\log x}{\log 2} + \frac{\log(\#\mathcal{J})}{\log 2}\right\},$$

$$\mathcal{M}(F, (S_i), G) = \bigcup_{k=1}^{j} \bigcup_{\substack{\sigma_i \in \mathbb{R} \text{ for } i\neq k \\ \sigma_k \leq -\frac{2\log x}{\log N_i}}} \mathcal{S}_1(F, (S_i), (\sigma_i)) \quad \cup \bigcup_{\gamma \leq -\frac{2\log x}{\log(\#\mathcal{J})}} \mathcal{S}_2(G, \gamma).$$

Sets $\{\mathcal{M}(F, (S_i), G)\} \cup \{\mathcal{S}_1(F, (S_i), (\sigma_i)) \cap \mathcal{S}_2(G, \gamma) : \sigma_i \in \Theta_i, \gamma \in \Theta_L\}$ partition $\mathbb{N} \cap [T, 2T]$. By (2.33),

$$\frac{\#\mathcal{J}}{z(\log x)^{A+5\cdot 10^6 K}} \leq \sum_{\substack{(\sigma_i):\sigma_i\in\Theta_i \\ \gamma\in\Theta_L}} \left| \sum_{\substack{n\in\mathcal{S}_1(F,(S_i),(\sigma_i)) \\ n\in\mathcal{S}_2(G,\gamma) \\ n\in[T,2T]}} \int_n^{n+1} D_z(c+it)\overline{G(it)}F(c+it)\mathrm{d}t\right| \tag{2.34}$$

$$+ \sum_{\substack{n\in\mathcal{M}(F,(S_i),G) \\ n\in[T,2T]}} \int_n^{n+1} \left|D_z(c+it)\overline{G(it)}F(c+it)\right| \mathrm{d}t \tag{2.35}$$

$$+ \int_{[T,T+1]\cup[2T,2T+1]} \left|D_z(c+it)\overline{G(it)}F(c+it)\right| \mathrm{d}t. \tag{2.36}$$

We have $D_z(c+it) = O(1/z)$, $|G(it)| \le \#\mathcal{J}$ and $S_i(c+it) = O(\log x)$. Recall that $F(s)$ also has factor $S_\ell(s)$ which satisfies property (D) with respect to T. Using Lemma 2.6, $|S_\ell(c+it)| \ll (\log x)^{-(A+5\cdot 10^6 K+J+1)}$ and so $|F(c+it)| \ll (\log x)^{-(A+5\cdot 10^6 K+1)}$. So (2.36) is $O(\#\mathcal{J}/(z(\log x)^{A+5\cdot 10^6 K+1}))$. If $n \in \mathcal{M}(F,(S_i),G)$, we also have $S_\ell(c+it) = O(1/x^2)$ for all $t \in [n, n+1]$ and at least one choice of ℓ or $G(it) = O(1/x^2)$ for all $t \in [n, n+1]$. Hence the contribution of (2.35) is also $\ll T_0(\#\mathcal{J})(\log x)^J/(x^2 z) \ll (\#\mathcal{J})/(z(\log x)^{A+5\cdot 10^6 K+1})$. Hence we focus on the RHS of (2.34). Since $\#\Theta_i = O(\log x)$ and $\#\Theta_L = O(\log x)$, there exist (σ_i) with $\sigma_i \in \Theta_i$ and $\gamma \in \Theta_L$ such that

$$\left| \sum_{n \in \mathcal{S}_1(F,(S_i),(\sigma_i)) \cap \mathcal{S}_2(G,\gamma) \cap [T,2T]} \int_n^{n+1} D_z(c+it)\overline{G(it)}F(c+it)\mathrm{d}t \right| \ge \frac{\#\mathcal{J}}{z(\log x)^{A+5\cdot 10^6 K+J+2}}. \tag{2.37}$$

Set $\sigma = \frac{\sum_{i=1}^j \log(B_i)\sigma_i}{\sum_{i=1}^j \log(B_i)}$, so that $(\prod_{i=1}^j B_i)^\sigma = \prod_{i=1}^j B_i^{\sigma_i}$.

Recall that we assumed that $F(s)$, $G(s)$ and T satisfy one of the conditions (C1), (C2), (C3), (C4.A) or (C4.B) of Proposition 2.1 for any given $\gamma \in (-\infty, 1)$ and (σ_i). We now split our argument into four different cases, depending on which of these conditions holds for the choice of γ and (σ_i) which satisfies (2.37). The goal is to show that $\#\mathcal{J} \le \tau^2 x^{0.23-1+\varepsilon/2}$.

Case 1: Condition (C1) holds. Suppose $F(s)$, (σ_i) and T satisfy both (2.37) and condition (C1) of Proposition 2.1. Recalling $D_z(c+it) = O(\frac{1}{z})$ and $|G(it)| \le \#\mathcal{J}$, using the size of $F(c+it)$ prescribed by $\mathcal{S}_1(F,(S_i),(\sigma_i))$ (namely $O(x^{\sigma-c})$) and rearranging, the LHS of (2.37) is certainly of size

$$O\left(\frac{x^{\sigma-c}R(F,(S_i),(\sigma_i),T)\#\mathcal{J}}{z}\right). \tag{2.38}$$

But $R(F,(S_i),(\sigma_i),T) \le \frac{x^{1-\sigma}}{(\log x)^{2J}}$ by (C1). Recall that $J = A + 10^7 K$. Hence (2.38) is less than $\frac{\#\mathcal{J}}{z(\log x)^{A+5\cdot 10^6 K+J+3}}$. This contradicts (2.37) when $\#\mathcal{J} > 0$. Hence $\#\mathcal{J} = 0$.

Case 2: Condition (C2) holds. Suppose $F(s)$, (σ_i) and T satisfy both (2.37) and condition (C2) of Proposition 2.1. By the Cauchy-Schwarz inequality and Montgomery's mean value estimate for Dirichlet polynomials (see Theorem 7.3 of [34]),

$$\sum_{n \in \mathcal{S}_1(F,(S_i),(\sigma_i)) \cap \mathcal{S}_2(G,\gamma) \cap [T,2T]} \int_n^{n+1} \left| D_z(c+it)\overline{G(it)}F(c+it) \right| \mathrm{d}t$$

$$\le \left(\int_T^{2T} |G(it)|^2 \, \mathrm{d}t \right)^{1/2} \left(\sum_{n \in \mathcal{S}_1(F,(S_i),(\sigma_i)) \cap \mathcal{S}_2(G,\gamma) \cap [T,2T]} \int_n^{n+1} |D_z(c+it)F(c+it)|^2 \, \mathrm{d}t \right)^{1/2}$$

$$\ll (T_0 \#\mathcal{J})^{1/2} R(F,(S_i),(\sigma_i),T)^{1/2} (x^{\sigma-c}) z^{-1}.$$

(When applying Montgomery's mean value estimate, we used that $\mathcal{J} \subseteq [1, T_0]$.) Comparing with (2.37), we get $\#\mathcal{J} \le \tau x^{2\sigma-2+2\varepsilon_1} R(F, (S_i), (\sigma_i), T)$. But by (C2) we have $R(F, (S_i), (\sigma_i), T) \le \tau x^{1.23-2\sigma+2\varepsilon_1}$ and so $\#\mathcal{J} \le \tau^2 x^{0.23-1+4\varepsilon_1}$.

Case 3: Condition (C3) holds. Suppose $F(s)$, (σ_i) and T satisfy (2.37) and condition (C3). Let

$$K(y) = \sum_{\substack{n\in\mathcal{S}_1(F,(S_i),(\sigma_i)) \\ n\in\mathcal{S}_2(G,\gamma) \\ n\in[T,2T]}} \int_n^{n+1} y^{it} u^{-(c+it)} C_z(c+it) F(c+it)\,\mathrm{d}t.$$

$K(w) = K(y) - \int_y^w K'(u)\mathrm{d}u$, $\frac{1}{H}\int_y^{y+H} K(w)\mathrm{d}w = K(y) - \frac{1}{H}\int_y^{y+H} K'(u)(y+H-u)\mathrm{d}u$. So

$$\begin{aligned}
\frac{\#\mathcal{J}}{z(\log x)^{A+5\cdot10^6K+J+2}} &\le \left| \sum_{\substack{n\in\mathcal{S}_1(F,(S_i),(\sigma_i)) \\ n\in\mathcal{S}_2(G,\gamma) \\ n\in[T,2T]}} \int_n^{n+1} D_z(c+it)\overline{G(it)}F(c+it)\mathrm{d}t \right| \\
&\le \sum_{m\in\mathcal{J}} \left| \sum_{\substack{n\in\mathcal{S}_1(F,(S_i),(\sigma_i)) \\ n\in\mathcal{S}_2(G,\gamma) \\ n\in[T,2T]}} \int_n^{n+1} (Hm)^{it} u^{-(c+it)} C_z(c+it) F(c+it)\mathrm{d}t \right| \\
&= \sum_{m\in\mathcal{J}} |K(Hm)| \le \frac{1}{H}\sum_{m\in\mathcal{J}} \int_{Hm}^{Hm+H} |K(w)|\mathrm{d}w + \sum_{m\in\mathcal{J}} \int_{Hm}^{Hm+H} |K'(w)|\mathrm{d}w \\
&\le \frac{(H\#\mathcal{J})^{3/4}}{H} \left(\int_x^{4x} |K(w)|^4 \mathrm{d}w \right)^{1/4} + (H\#\mathcal{J})^{3/4} \left(\int_x^{4x} |K'(w)|^4\mathrm{d}w \right)^{1/4}. \qquad (2.39)
\end{aligned}$$

The purpose of these estimates was to remove $G(s)$ and to introduce a factor w^{it} with w varying over all $w \in [x, 4x]$. This change is needed for the introduction of $R^*(F, (S_i), (\sigma_i), T)$. To be precise, we have

$$\int_x^{4x} |K(w)|^4\mathrm{d}w = \int_x^{4x} \left| \sum_{\substack{n\in\mathcal{S}_1(F,(S_i),(\sigma_i)) \\ n\in\mathcal{S}_2(G,\gamma)\cap[T,2T]}} \int_n^{n+1} w^{it} \left(\frac{C_z(c+it)}{u^{c+it}} \right) F(c+it)\,\mathrm{d}t \right|^4 \mathrm{d}w, \quad (2.40)$$

$$\int_x^{4x} |K'(w)|^4\mathrm{d}w = \int_x^{4x} \left| \sum_{\substack{n\in\mathcal{S}_1(F,(S_i),(\sigma_i)) \\ n\in\mathcal{S}_2(G,\gamma)\cap[T,2T]}} \int_n^{n+1} w^{it-1} \left(\frac{tC_z(c+it)}{u^{c+it}} \right) F(c+it)\,\mathrm{d}t \right|^4 \mathrm{d}w. \quad (2.41)$$

An upper bound for (2.40) and (2.41) was given in the proof of Lemma 4.2 of [29]: The LHS of (2.40) is $O(x^{-3+4\sigma+\varepsilon_1} z^{-4} R^*(F, (S_i), (\sigma_i), T))$, while the LHS of (2.41) is $O(\tau^4 x^{-7+4\sigma+5\varepsilon_1} z^{-4} R^*(F, (S_i), (\sigma_i), T))$. But (C3) holds and $R^*(F, (S_i), (\sigma_i), T) \le$

$\tau x^{3.23-4\sigma+2\varepsilon_1}$. Hence we have the upper bound

$$(H\#\mathcal{J})^{3/4}\left(\frac{1}{H}\left(\int_x^{4x}|K(w)|^4\mathrm{d}w\right)^{1/4}+\left(\int_x^{4x}|K'(w)|^4\mathrm{d}w\right)^{1/4}\right)\ll\frac{(\#\mathcal{J})^{3/4}}{(z^4\tau^{-2}x^{0.77-4\varepsilon_1})^{1/4}} \tag{2.42}$$

and by combining (2.39) and (2.42) we find that $\#\mathcal{J}\leq\tau^2x^{0.23-1+\varepsilon/2}$, provided x is sufficiently large.

Case 4: Condition (C4.A) or (C4.B) holds. Finally, suppose $F(s)$, (σ_i), $G(s)$, γ and T satisfy (2.37) and one of condition (C4.A) or condition (C4.B) of Proposition 2.1. Using the sizes of F and G prescribed by $\mathcal{S}_1(F,(S_i),(\sigma_i))\cap\mathcal{S}_2(G,\gamma)$,

$$\left|\sum_{\substack{n\in\mathcal{S}_1(F,(S_i),(\sigma_i))\\ n\in\mathcal{S}_2(G,\gamma)\\ n\in[T,2T]}}\int_n^{n+1}D_z(c+it)\overline{G(it)}F(c+it)\mathrm{d}t\right|\ll\left((\#\mathcal{J})^\gamma x^{\sigma-1}\right)\frac{Q(F,G,(S_i),(\sigma_i),\gamma,T)}{z}.$$

Hence by (2.37), $(\#\mathcal{J})^{1-\gamma}<x^{\sigma-1+\varepsilon_1}Q(F,G,(S_i),(\sigma_i),\gamma,T)$.

If (C4.A) holds, then $Q(F,G,(S_i),(\sigma_i),\gamma,T)\leq(\#\mathcal{J})^{7/8-\gamma}\tau^{1/4}x^{7/8-\sigma+0.23/8+\varepsilon_1}$. So

$$\#\mathcal{J}<\left(\frac{x^{\sigma-1+\varepsilon_1}Q(F,G,(S_i),(\sigma_i),\gamma,T)}{(\#\mathcal{J})^{7/8-\gamma}}\right)^8\leq\tau^2x^{0.23-1+\varepsilon/2}. \tag{2.43}$$

Alternatively, if (C4.B) holds, then $Q(F,G,(S_i),(\sigma_i),\gamma,T)\leq(\#\mathcal{J})^{1-\gamma}x^{1-\sigma-\varepsilon_1}$. In particular,

$$1<\frac{x^{\sigma-1+\varepsilon_1}Q(F,G,(S_i),(\sigma_i),\gamma,T)}{(\#\mathcal{J})^{1-\gamma}}\leq 1. \tag{2.44}$$

This gives a contradiction when $\#\mathcal{J}\neq 0$. So $\#\mathcal{J}=0\leq\tau^2x^{0.23-1+\varepsilon/2}$.

Hence each of the conditions (C1), (C2), (C3), (C4.A) and (C4.B) implies $\#\mathcal{J}\leq\tau^2x^{0.23-1+\varepsilon/2}$ and $10(x/T_0)\#\mathcal{J}\leq\tau x^{0.23+\varepsilon/2}$. Continuing on from Lemma 2.7, this completes our proof of Proposition 2.1. □

2.4 Values of Dirichlet polynomials

The goal of this section is to prove Proposition 2.2. We will give various bounds on $R(F,(S_i),(\sigma_i),T)$, $R^*(F,(S_i),(\sigma_i),T)$ and $Q(F,G,(S_i),(\sigma_i),\gamma,T)$, dependent on the factorisation of $F(s)$.

Let $K = 2000$ and $J > 10^7 K$. Throughout Section 2.4 we assume for given τ, ε and $T \in [T_1, T_0]$ that $F(s) = \prod_{i=1}^{j} S_i(s)$, where

$$S_i(s) = \sum_{n \sim N_i} \frac{a_n^{(i)}}{n^s}$$

for some $j \leq J$ and some $a_n^{(i)} \in \mathbb{C}$ with $a_n^{(i)} = O(\log n)$. Further, $N_i \geq \log x$ for all $i \in \{1, \ldots, j\}$ and $2^{-J} x \leq \prod_{i=1}^{j} N_i \leq 2^J x$. Additionally, we have the following criteria:

Definition 2.6. *Let $F(s) = \prod_{i=1}^{j} S_i(s)$ be as described at the start of Section 2.4, with $N_i < x^{0.95}$ for all i.*

Assume that $S_i(s)$ has property (Z1) or (Z2) with respect to T whenever $N_i \geq x^{2/K}$. Assume also that there exists $\ell \in \{1, \ldots, j\}$ with $N_\ell \geq x^{2/K}$ and some $i_0 \neq \ell$ for which $N_{i_0} \geq x^{1/\log\log x}$ and for which $S_{i_0}(s)$ has property (D) with respect to T.

Then we say that $F(s) = \prod_{i=1}^{j} S_i(s)$ is a Dirichlet polynomial of Type A at T.

Here (D), (Z1) and (Z2) are as described in Definition 2.1.

Definition 2.7. *Let $F(s) = \prod_{i=1}^{j} S_i(s)$ be as described at the start of Section 2.4, with $N_i < x^{2/K}$ for all i.*

Assume that for some $i_0 \in \{1, \ldots, j\}$, $S_{i_0}(s)$ has property (D) with respect to T and $N_{i_0} \geq x^{1/\log\log x}$.

Then we say $F(s) = \prod_{i=1}^{j} S_i(s)$ is a Dirichlet polynomial of Type B at T.

Definition 2.8. *Let $F(s) = \prod_{i=1}^{j} S_i(s)$ be as described at the start of Section 2.4.*

Assume that for some $\ell \in \{1, \ldots, j\}$, $S_\ell(s)$ has property (Z1) or (Z2) with respect to T and $N_\ell \geq x^{0.95}$.

Then we say $F(s) = \prod_{i=1}^{j} S_i(s)$ is a Dirichlet polynomial of Type C at T.

Throughout Section 2.4 we only work with $F(s)$ which are of Type A, B or C at T. Note in particular that all choices of $F(s) \in \mathcal{F}^*(T, J, K)$, described in Definition 2.2, satisfy one of the above criteria.

We also assume that $G(s) = \sum_{n \in \mathcal{J}} \xi_n n^{-s}$, where $\mathcal{J} \subseteq [1, T_0]$ with $\#\mathcal{J} \geq \tau^2 x^{-0.77}$ and $|\xi_n| = 1$.

We work with $\gamma \in (-\infty, 1)$ and (σ_i) with $\sigma_i \in \mathbb{R}$ and corresponding $\sigma = \frac{\sum_{i=1}^{j} \log(N_i)\sigma_i}{\sum_{i=1}^{j} \log(N_i)}$. If x is large and $\sigma_j > 1 + \varepsilon_1$ for some j, then $R(F, (S_i), (\sigma_i), T) = 0$. So we assume $\sigma_i \le 1 + \varepsilon_1$ for all i. If $\sigma < 0.6$, then

$$R(F, (S_i), (\sigma_i), T) \le T \le \tau x^{\varepsilon_1} \le \tau x^{1.23 - 2\sigma + 2\varepsilon_1}$$

and condition (C2) of Proposition 2.1 holds. Hence we restrict our attention to (σ_i) with $\sigma \in [0.6, 1 + \varepsilon_1]$.

We set $N = \prod_{i=1}^{j} N_i$ and write $N_i = N^{\ell_i}$, so that $\sum_{i=1}^{j} \ell_i = 1$ and $\sum_{i=1}^{j} \ell_i \sigma_i = \sigma$.

2.4.1 Standard bounds on R

We now consider a non-empty set $I \subseteq \{1, \dots, j\}$ and let $\{k_i : i \in I\}$ be a corresponding set of positive integers, with k_i bounded by an absolute constant. For simplicity's sake, we assume that $k_i \le J$. We set $M = \prod_{i \in I} N_i^{k_i}$ and $M^\beta = \prod_{i \in I} N_i^{k_i \sigma_i}$, so that

$$R(F, (S_i), (\sigma_i), T) \le \# \Big\{ n \in \mathbb{Z} \cap [T, 2T] : \sup_{t \in [n, n+1]} \prod_{i \in I} |S_i(c + it)|^{k_i} \in [M^{-c+\beta}, 2^{J^2} M^{-c+\beta}] \Big\}.$$

For the set on the RHS, concerning the size of Dirichlet polynomial $\prod_{i \in I} S_i(s)^{k_i}$, a number of standard bounds are available. We will frequently use the following bounds:

Lemma 2.8. *Uniformly for $T_1 \le T \le T_0$ and for $\delta > 0$*

$$R(F, (S_i), (\sigma_i), T) \ll_\delta x^\delta M^{2-2\beta} + x^\delta T_0 M^{1-2\beta}, \tag{2.45}$$

$$R(F, (S_i), (\sigma_i), T) \ll_\delta x^\delta M^{2-2\beta} + x^\delta T_0 M^{4-6\beta}, \tag{2.46}$$

$$R(F, (S_i), (\sigma_i), T) \ll_\delta x^\delta M^{2-2\beta} + x^\delta T_0 M^{11-14\beta}. \tag{2.47}$$

Result (2.45) is a consequence of Theorem 7.3 of [34] and is known as Montgomery's mean value estimate. (2.46) is a consequence of Equation 2.9 of [23], Huxley's large values estimate. All three bounds can also be deduced from Lemma 8.2 of [25].

Next we list simple bounds on $R(F, (S_i), (\sigma_i), T)$ which require $F(s)$ to be of Type A, B or C at T. Here it is important that long factors of $F(s)$ satisfy (Z1) or (Z2).

2.4.1.1 Long factors

Dirichlet polynomials of Type C are easy to deal with, as the next lemma shows.

Lemma 2.9. *Suppose $F(s) = \prod_{i=1}^{j} S_i(s)$ is a Dirichlet polynomial of Type C at T, where $T \in [T_1, T_0]$. Then for large x and for (σ_i) with corresponding $\sigma \ge 0.6$ we have*

$$R(F, (S_i), (\sigma_i), T) = 0.$$

Proof. Let $S_\ell(s)$ denote the factor of $F(s)$ which has $N_\ell \geq x^{0.95}$. Suppose first that $S_\ell(s)$ satisfies (Z2). Then $|S_\ell(c+it)| \ll (\log N_\ell) N_\ell^{-1/2} \ll x^{-0.47}$. Now suppose that $S_\ell(s)$ satisfies (Z1). By van der Corput's method of exponential sums, taking for instance $l = 2$ on page 37 of [29], we have for $t \in [T, 2T]$,

$$|S_\ell(c+it)| \ll_\delta \frac{T^{1/2+\delta}}{N_\ell}.$$

But we are working with $T \leq T_0 = \tau x^{\varepsilon_1} \leq x^{1-\nu}$ and hence $\nu = 0.23$ and $N_\ell \geq x^{0.95}$ ensure again that $|S_\ell(c+it)| \ll x^{-0.47}$. Our general assumptions on $F(s)$ also guarantee $|S_i(c+it)| \ll_\delta x^\delta$. So

$$|F(c+it)| = |S_\ell(c+it)| \left(\prod_{i \neq \ell} |S_i(c+it)| \right) \ll x^{-0.45}.$$

In particular, $R(F, (S_i), (\sigma_i), T) = 0$ for $\sigma \geq 0.6$ and x large. □

Very good bounds on $R(F, (S_i), (\sigma_i), T)$ are also available when a Dirichlet polynomial of Type A has a particularly long factor.

Lemma 2.10. *Suppose $F(s) = \prod_{i=1}^{j} S_i(s)$ is as described at the start of Section 2.4. Let $T \in [T_1, T_0]$.*
Suppose $S_\ell(s)$ has property (Z1) or (Z2) with respect to T. Suppose also that there is some $i_0 \neq \ell$ for which $S_{i_0}(s)$ has property (D) with respect to T and for which $N_{i_0} \geq x^{1/\log\log x}$.
Then if $x \geq C$, one of the following two inequalities holds:

$$R(F, (S_i), (\sigma_i), T) \leq \frac{x^{1-\sigma}}{(\log x)^B} \quad \text{or}$$
$$R(F, (S_i), (\sigma_i), T) \leq x^\delta \min\left\{ T N_\ell^{(2-4\sigma_\ell)}, T^2 N_\ell^{(6-12\sigma_\ell)} \right\}$$

Here C is a large constant dependent only on $J \in \mathbb{N}$, $B > 0$ and $\delta > 0$.

Proof. We begin by assuming that $S_\ell(s)$ satisfies property (Z1) with respect to T.

It follows from small modifications of the proof of Lemma 5.5 of [29] that either

$$R(F, (S_i), (\sigma_i), T) \ll_\delta x^{\delta/2} \min\left\{ T N_\ell^{(2-4\sigma_\ell)}, T^2 N_\ell^{(6-12\sigma_\ell)} \right\},$$

in which case we are done, or

$$T \ll (\log x)^4 N_\ell^{1-\sigma_\ell} \ll (\log x)^4 x^{1-\sigma} \prod_{i \neq \ell} N_i^{\sigma_i - 1}. \tag{2.48}$$

(This is (i) and (ii) on page 22 of [29].)

Since $S_{i_0}(s)$ has property (D) and $N_{i_0} \geq x^{1/\log\log x}$, Lemma 2.6 gives

$$|S_{i_0}(c+it)| \leq \frac{C_{B,J}}{(\log x)^{B+5+J}} = N_{i_0}^{-\frac{(B+5+J)\log\log x}{\log(N_{i_0})} + \frac{\log(C_{B,J})}{\log(N_{i_0})}}$$

for $t \in [T, 2T]$, where $C_{B,J}$ is a suitable large constant, dependent on B and J. Hence $R(F,(S_i),(\sigma_i),T) = 0$ when $\sigma_{i_0} > c - \frac{(B+5+J)\log\log x}{\log(N_{i_0})} + \frac{\log(C_{B,J})}{\log(N_{i_0})}$. But $|S_i(c+it)| \ll \log(N_i)$ for all other i. By (2.48) we must have $R(F,(S_i),(\sigma_i),T) = 0$ or

$$R(F,(S_i),(\sigma_i),T) \leq T \ll_{B,J} (\log x)^{4+J} x^{1-\sigma} \frac{1}{(\log x)^{B+5+J}}.$$

This concludes the proof for the case where $S_\ell(s)$ has property (Z1).

Now suppose that $S_\ell(s)$ has property (Z2) with respect to T, so that $S_\ell(c+it) = O((\log N_\ell) N_\ell^{1/2-c})$. This implies that for $\sigma_\ell > \frac{1}{2} + \frac{2\log\log x}{\log N_\ell}$ and x large, we have $R(F,(S_i),(\sigma_i),T) = 0$. On the other hand, if $\sigma_\ell \leq \frac{1}{2} + \frac{2\log\log x}{\log N_\ell}$, then trivially

$$R(F,(S_i),(\sigma_i),T) \leq T \leq x^\delta \min\left\{T N_\ell^{(2-4\sigma_\ell)}, T^2 N_\ell^{(6-12\sigma_\ell)}\right\}. \qquad \square$$

2.4.1.2 Very large sigma

We also observe that extremely large values of σ occur only rarely.

Lemma 2.11. *Suppose $F(s) = \prod_{i=1}^{j} S_i(s)$ is of Type A or Type B at T. Suppose (σ_i) corresponds to σ with $1 - 10^{-500} \leq \sigma \leq 1 + \varepsilon_1$. Then for $B > 0$,*

$$R(F,(S_i),(\sigma_i),T) \ll_B \frac{x^{1-\sigma}}{(\log x)^B}.$$

Proof. First we suppose that $F(s)$ is a Dirichlet polynomial of Type A. Then $F(s)$ has a factor $S_\ell(s)$ with $N_\ell \geq x^{2/K}$ which satisfies (Z1) or (Z2). If $S_\ell(s)$ has property (Z1), the computations of Case 1B on page 37 and 38 of [29], using van der Corput's method of exponential sums, apply in slightly amended form with $k = 1000$ and give $R(F,(S_i),(\sigma_i),T) = 0$ when x is large and $\sigma \geq 1 - 10^{-500}$.

If $S_\ell(s)$ has property (Z2), then $R(F,(S_i),(\sigma_i),T) = 0$ whenever $\sigma_\ell > \frac{1}{2} + \frac{2\log\log x}{\log N_\ell}$. But if $\sigma_\ell \leq \frac{1}{2} + \frac{2\log\log x}{\log N_\ell}$, then $\sum_{i=1}^{j} \ell_i = 1$, $\sum_{i=1}^{j} \ell_i\sigma_i = \sigma$, $\sigma_i \leq 1 + \varepsilon_1$ and $N_\ell \geq x^{0.001}$ certainly give $\sigma < 1 - 10^{-500}$.

Now suppose that $F(s)$ is a Dirichlet polynomial of Type B. There exists i_0 such that $S_{i_0}(s)$ has property (D) and such that $N_{i_0} \geq x^{1/\log\log x}$. By Lemma 2.6, then

$$|F(c+it)| \ll (\log x)^J x^{-(\log x)^{-7/10}(\log\log x)^{-1}}$$

and we may work with $\sigma \leq 1 - (\log x)^{-8/10}$. Since $\nu = 0.23$, a few small numerical changes to the computations of Case 1A on page 37 of [29] give

$$R(F, (S_i), (\sigma_i), T) \ll T_0^{5(1-\sigma)/4} (\log x)^C$$

for a very large constant C, dependent only on J. However, $T_0^{5/4} \leq x^{1-0.01}$ and $-(1-\sigma) \leq -(\log x)^{-8/10}$.

$$R(F, (S_i), (\sigma_i), T) \ll x^{(1-\sigma) - 0.01(\log x)^{-8/10}} (\log x)^C \ll_B \frac{x^{1-\sigma}}{(\log x)^B}$$

for $1 - 10^{-500} \leq \sigma \leq 1 - (\log x)^{-8/10}$, as proposed. $\square$

2.4.1.3 Simple general bounds

Finally, we also give some bounds on $R(F, (S_i), (\sigma_i), T)$ which do not depend on the lengths N_i of factors $S_i(s)$. Often they are convenient for simple computations.

Lemma 2.12. *Suppose $F(s) = \prod_{i=1}^{j} S_i(s)$ is of Type A or B at T, where $T \in [T_1, T_0]$. Let $I \subseteq \{1, \ldots, j\}$ and suppose $\sum_{i \in I} \ell_i \geq 0.01$ and $\sigma_I = (\sum_{i \in I} \ell_i \sigma_i)/(\sum_{i \in I} \ell_i) \geq 0.6$. Let $B > 0$ and $\delta > 0$. Suppose*

$$R(F, (S_i), (\sigma_i), T) > \frac{x^{1-\sigma}}{(\log x)^B}.$$

Then the following four bounds hold:

$$R(F, (S_i), (\sigma_i), T) \ll T_0^{(3-3\sigma_I)/(2-\sigma_I)+\delta} \quad \text{if} \quad \sigma_I \geq 6/10, \tag{2.49}$$
$$R(F, (S_i), (\sigma_i), T) \ll T_0^{(3-3\sigma_I)/(3\sigma_I-1)+\delta} \quad \text{if} \quad \sigma_I \geq 6/10, \tag{2.50}$$
$$R(F, (S_i), (\sigma_i), T) \ll T_0^{(3-3\sigma_I)/(10\sigma_I-7)+\delta} \quad \text{if} \quad 7/10 < \sigma_I \leq 25/28, \tag{2.51}$$
$$R(F, (S_i), (\sigma_i), T) \ll T_0^{(4-4\sigma_I)/(4\sigma_I-1)+\delta} \quad \text{if} \quad \sigma_I \geq 25/28. \tag{2.52}$$

Here the implied constants depend on B and δ.

Proof. This statement follows from the proof of Lemma 5.6 of [29]. Key in that proof was the observation that bounds (2.49), (2.50), (2.51) and (2.52) hold if there exist $I_2 \subseteq I$ and $k \in \mathbb{N}$ with $(\prod_{i \in I_2} N^{\ell_i})^k \in [Y^{1/2}, Y]$ and $\sum_{i \in I_2} \ell_i \sigma_i \geq \sigma_I (\sum_{i \in I_2} \ell_i)$ and $k = O(1)$, where

$$Y = \min\{T_0^{3/(8-4\sigma_I)}, T_0^{3/(12\sigma_I-4)}, T_0^{3/(40\sigma_I-28)}, T_0^{3/(4\sigma_I-1)}\}.$$

Since $T_0 \geq x^{0.47}$ and $\sigma_I \in [0.6, 1+\varepsilon_1]$, we got $Y \geq x^{0.11}$.

If $F(s)$ is a Dirichlet polynomial of Type B, all $S_i(s)$ have $N_i < x^{0.001}$ and we can easily choose I_2 such that $\prod_{i\in I_2} N^{\ell_i} \in [x^{0.01}, Y]$ and $\sum_{i\in I_2} \ell_i\sigma_i \geq \sigma_I(\sum_{i\in I_2} \ell_i)$ are satisfied.

On the other hand, in the presence of longer factors, the proof of Lemma 5.6 of [29] additionally required that the conclusion of Lemma 2.10 holds whenever $i \in I$ and $N_i \geq Yx^{-0.01} \geq x^{0.1}$. But if $F(s)$ is of Type A, then $N_i \geq x^{0.1}$ implies that $S_i(s)$ satisfies property (Z1) or (Z2), so that Lemma 2.10 is indeed applicable. Hence Lemma 2.12 follows from the proof of (134), (135), (136) and (137) of Lemma 5.6 of [29]. □

We often use the following version of Lemma 2.12:

Corollary 2.13. *Suppose $F(s) = \prod_{i=1}^{j} S_i(s)$ is of Type A or B at T, where $T \in [T_1, T_0]$.*
Let $I \subseteq \{1, \ldots, j\}$ and suppose $\sum_{i\in I} \ell_i \geq 0.01$. Write $\sigma_I = (\sum_{i\in I} \ell_i\sigma_i)/(\sum_{i\in I} \ell_i)$.
Let $B > 0$. Suppose that $\sigma_I \geq \sigma$ or $\sum_{i\in I} \ell_i \geq 0.2$. Then one of the following three statements holds:

$$R(F, (S_i), (\sigma_i), T) \leq \frac{x^{1-\sigma}}{(\log x)^B}, \tag{2.53}$$

$$R(F, (S_i), (\sigma_i), T) \leq \tau x^{1-2\sigma+\nu+2\varepsilon_1}, \tag{2.54}$$

$$R(F, (S_i), (\sigma_i), T) \leq \min\left\{x^{2\varepsilon_1}\tau^{\frac{(3-3\sigma_I)}{(2-\sigma_I)}}, x^{2\varepsilon_1}\tau^{\frac{(3-3\sigma_I)}{(3\sigma_I-1)}}\right\} \quad \text{and} \quad \sigma_I \geq 0.35. \tag{2.55}$$

Let $B(a, \sigma)$ be a non-negative function dependent on $a \in [0.475 - \varepsilon, 0.77 - \varepsilon]$ and $\sigma \in [0.6, 1-10^{-500}]$. If (2.55) holds, we also have $R(F, (S_i), (\sigma_i), T) \leq x^{B(a,\sigma)}$, provided

$$\sigma_I \geq \min\left\{\frac{3a - 2B(a, \sigma) + 4\varepsilon_1}{3a - B(a, \sigma) + 2\varepsilon_1}, \frac{3a + B(a, \sigma) - 2\varepsilon_1}{3a + 3B(a, \sigma) - 6\varepsilon_1}\right\}. \tag{2.56}$$

Finally, if $\sigma_I > 0.7$, we also have (2.53) *or $R(F, (S_i), (\sigma_i), T) \leq x^{B(a,\sigma)}$ whenever*

$$\sigma_I \geq \max\left\{\frac{3a + 7B(a, \sigma) - 14\varepsilon_1}{3a + 10B(a, \sigma) - 20\varepsilon_1}, \frac{4a + B(a, \sigma) - 2\varepsilon_1}{4a + 4B(a, \sigma) - 8\varepsilon_1}\right\}. \tag{2.57}$$

Proof. Assume $R(F, (S_i), (\sigma_i), T) > x^{1-\sigma}(\log x)^{-B}$. We always take $\sigma \geq 0.6$, so if $\sigma_I \geq \sigma$, (2.55) follows directly from Lemma 2.12. (We use that $(3 - 3x)/(3x - 1)$ and $(3 - 3x)/(2 - x)$ are bounded by 1.5 on $[0.6, 1]$ to replace T_0 by τ.) If $\sigma_I < \sigma$, Lemma 2.12, applied with $\{1, \ldots, j\}$ in the place of I, gives

$$R(F, (S_i), (\sigma_i), T) \leq \min\left\{x^{2\varepsilon_1}\tau^{(3-3\sigma)/(2-\sigma)}, x^{2\varepsilon_1}\tau^{(3-3\sigma)/(3\sigma-1)}\right\}. \tag{2.58}$$

But $(3-3x)/(2-x)$ is decreasing on $x<2$ and $(3-3x)/(3x-1)$ is decreasing on $x \geq 0.35$. Thus if $0.35 \leq \sigma_I < \sigma$, then (2.58) implies (2.55). Finally, if $\sigma_I < 0.35$ and $\sum_{i \in I} \ell_i \geq 0.2$, then $I' = \{1, \ldots, j\} \setminus I$ satisfies

$$\sigma_{I'} = \frac{\sigma - \sum_{i \in I} \ell_i \sigma_i}{1 - \sum_{i \in I} \ell_i} \geq \frac{\sigma - 0.35 \cdot 0.2}{1 - 0.2}.$$

In this case, Lemma 2.12, applied with I' in the place of I, gives

$$R(F, (S_i), (\sigma_i), T) \leq \min \left\{ x^{2\varepsilon_1} \tau^{\frac{(3.2625-3.75\sigma)}{(2.0875-1.25\sigma)}}, x^{2\varepsilon_1} \tau^{\frac{(3.2625-3.75\sigma)}{(3.75\sigma-1.2625)}} \right\}. \tag{2.59}$$

For $a \in [0.475 - \varepsilon, 0.77 - \varepsilon]$ and $\sigma \geq 0.6$, the RHS of (2.59) is less than

$$\max\{\tau x^{1-2\sigma+0.23+2\varepsilon_1}, x^{1-\sigma-\varepsilon_1}\}.$$

Now suppose we have (2.55). If $B(a,\sigma) > a + \varepsilon_1$, $R(F, (S_i), (\sigma_i), T) \leq T_0 < x^{B(a,\sigma)}$ holds trivially. Thus we assume $0 \leq B(a,\sigma) \leq a + \varepsilon_1$. Inequality $x^{2\varepsilon_1} \tau^{(3-3\sigma_I)/(2-\sigma_I)} \leq x^{B(a,\sigma)}$ is satisfied if

$$\sigma_I \geq \frac{3a - 2B(a,\sigma) + 4\varepsilon_1}{3a - B(a,\sigma) + 2\varepsilon_1},$$

while inequality $x^{2\varepsilon_1} \tau^{(3-3\sigma_I)/(3\sigma_I-1)} \leq x^{B(a,\sigma)}$ holds if

$$\sigma_I \geq \frac{3a + B(a,\sigma) - 2\varepsilon_1}{3a + 3B(a,\sigma) - 6\varepsilon_1}.$$

Finally, if $\sigma_I > 0.7$, then $R(F, (S_i), (\sigma_i), T) \leq \max\{x^{2\varepsilon_1} \tau^{\frac{(3-3\sigma_I)}{(10\sigma_I-7)}}, x^{2\varepsilon_1} \tau^{\frac{(4-4\sigma_I)}{(4\sigma_I-1)}}\}$ by bounds (2.51) and (2.52) of Lemma 2.12. The RHS of this inequality is at most $x^{B(a,\sigma)}$ if (2.57) holds. $\square$

2.4.2 Heath-Brown's R^* bound

We now have plenty of bounds on R and move on to R^*. We use the following result of Heath-Brown [17], stated here like in Lemma 5.4 of [29]:

Lemma 2.14 (Heath-Brown's R^* bound)**.** *Let $F(s) = \prod_{i=1}^{j} S_i(s)$ be as described at the start of Section 2.4. Let $I \subseteq \{1, \ldots, j\}$ and $k_i \in \mathbb{N} \cap [1, J]$. Set $M = \prod_{i \in I} N_i^{k_i}$ and $M^\beta = \prod_{i \in I} N_i^{k_i \sigma_i}$. For $T \in [T_1, T_0]$,*

$$R^*(F, (S_i), (\sigma_i), T) \ll_\delta x^\delta M^{1-2\beta} (RM + R^2 + R^{\frac{5}{4}} T_0^{1/2})^{\frac{1}{2}} (R^* M + R^4 + R (R^*)^{\frac{3}{4}} T_0^{1/2})^{\frac{1}{2}},$$

where $R = R(F, (S_i), (\sigma_i), T)$ and $R^* = R^*(F, (S_i), (\sigma_i), T)$.

To obtain a good bound on R^*, we thus again need to bound R. In particular, recalling that condition (C3) of Proposition 2.1 requires $R^*(F,(S_i),(\sigma_i),T)) \leq \tau x^{3-4\sigma+\nu+2\varepsilon_1}$, we desire the following bounds:

Corollary 2.15. *Let $F(s) = \prod_{i=1}^{j} S_i(s)$ be as described at the start of Section 2.4. Suppose the following condition (C3*) holds:*

> *There exist $I \subseteq \{1,\ldots,j\}$ and $k_i \in \mathbb{N} \cap [1,J]$ such that $M = \prod_{i\in I} N_i^{k_i}$ and $M^\beta = \prod_{i\in I} N_i^{k_i\sigma_i}$ satisfy each of the following nine inequalities:* (C3*)

$$R(F,(S_i),(\sigma_i),T) \leq \tau M^{4\beta-4} x^{3-4\sigma+\nu+\varepsilon_1/2}, \tag{2.60}$$
$$R(F,(S_i),(\sigma_i),T) \leq \tau^{1/2} M^{2\beta-3/2} x^{3/2-2\sigma+(1/2)\nu+\varepsilon_1/2}, \tag{2.61}$$
$$R(F,(S_i),(\sigma_i),T) \leq \tau^{2/5} M^{(16/5)\beta-12/5} x^{12/5-(16/5)\sigma+(4/5)\nu+\varepsilon_1/2}, \tag{2.62}$$
$$R(F,(S_i),(\sigma_i),T) \leq \tau^{2/5} M^{(4/5)\beta-3/5} x^{6/5-(8/5)\sigma+(2/5)\nu+\varepsilon_1/2}, \tag{2.63}$$
$$R(F,(S_i),(\sigma_i),T) \leq \tau^{1/3} M^{(2/3)\beta-1/3} x^{1-(4/3)\sigma+(1/3)\nu+\varepsilon_1/2}, \tag{2.64}$$
$$R(F,(S_i),(\sigma_i),T) \leq \tau^{2/7} M^{(16/21)\beta-8/21} x^{8/7-(32/21)\sigma+(8/21)\nu+\varepsilon_1/2}, \tag{2.65}$$
$$R(F,(S_i),(\sigma_i),T) \leq \tau^{3/8} M^{2\beta-3/2} x^{15/8-(5/2)\sigma+(5/8)\nu+\varepsilon_1/2}, \tag{2.66}$$
$$R(F,(S_i),(\sigma_i),T) \leq \tau^{1/4} M^{(4/3)\beta-2/3} x^{5/4-(5/3)\sigma+(5/12)\nu+\varepsilon_1/2}, \tag{2.67}$$
$$R(F,(S_i),(\sigma_i),T) \leq \tau^{1/9} M^{(16/9)\beta-8/9} x^{5/3-(20/9)\sigma+(5/9)\nu+\varepsilon_1/2}. \tag{2.68}$$

Then $R^*(F,(S_i),(\sigma_i),T) \leq \tau x^{3-4\sigma+\nu+2\varepsilon_1}$.

Proof. We rearrange the bound $R^* \ll_\delta x^\delta M^{1-2\beta}(RM + R^2 + R^{5/4}T_0^{1/2})^{1/2}(R^*M + R^4 + R(R^*)^{3/4}T_0^{1/2})^{1/2}$, given in Lemma 2.14. For a fixed $\delta > 0$ and all sufficiently large x, one of the following nine inequalities holds:

$$R^*(F,(S_i),(\sigma_i),T) \leq x^\delta M^{4-4\beta} R, \tag{2.69}$$
$$R^*(F,(S_i),(\sigma_i),T) \leq x^\delta M^{3-4\beta} R^2, \tag{2.70}$$
$$R^*(F,(S_i),(\sigma_i),T) \leq x^\delta M^{3-4\beta} R^{5/4} T_0^{1/2}, \tag{2.71}$$
$$R^*(F,(S_i),(\sigma_i),T) \leq x^\delta M^{3/2-2\beta} R^{5/2}, \tag{2.72}$$
$$R^*(F,(S_i),(\sigma_i),T) \leq x^\delta M^{1-2\beta} R^3, \tag{2.73}$$
$$R^*(F,(S_i),(\sigma_i),T) \leq x^\delta M^{1-2\beta} R^{21/8} T_0^{1/4}, \tag{2.74}$$
$$R^*(F,(S_i),(\sigma_i),T) \leq x^\delta M^{12/5-(16/5)\beta} R^{8/5} T_0^{2/5}, \tag{2.75}$$
$$R^*(F,(S_i),(\sigma_i),T) \leq x^\delta M^{8/5-(16/5)\beta} R^{12/5} T_0^{2/5}, \tag{2.76}$$
$$R^*(F,(S_i),(\sigma_i),T) \leq x^\delta M^{8/5-(16/5)\beta} R^{9/5} T_0^{4/5}, \tag{2.77}$$

where $R = R(F, (S_i), (\sigma_i), T)$. However, if $R^*(F, (S_i), (\sigma_i), T) \le x^\delta M^{A_1 - B_1\beta} R^{C_1} T_0^{D_1}$ with δ sufficiently small compared to ε_1, then $R^*(F, (S_i), (\sigma_i), T) \le \tau x^{3-4\sigma+\nu+2\varepsilon_1}$ whenever

$$R(F, (S_i), (\sigma_i), T) \le \tau^{(1-D_1)/C_1} M^{(B_1/C_1)\beta - A_1/C_1} x^{3/C_1 - (4/C_1)\sigma + \nu/C_1 + \varepsilon_1/2}. \tag{2.78}$$

Applying (2.78) to inequalities (2.69) – (2.77), we obtain inequalities (2.60) – (2.68).

□

2.4.3 Heath-Brown's Q bound

To bound Q, we use recent work of Heath-Brown [21]. It provides good bounds on long Dirichlet polynomials with few non-zero coefficients:

Lemma 2.16. *Let $G(s) = \sum_{n\in\mathcal{J}} \xi_n n^{-s}$ with $\mathcal{J} \subseteq [1, T_0]$ and $|\xi_n| = 1$.*
Let $F(s) = \sum_{n\le N} a_n n^{-s}$ with $a_n \in \mathbb{C}$. Write $L = \#\mathcal{J}$. Then

$$\int_0^{T_0} \left|\sum_{m\in\mathcal{J}} \frac{\xi_m}{m^{it}}\right|^2 \left|\sum_{n\le N} \frac{a_n}{n^{it}}\right|^2 dt \ll_\delta (N^2L^2 + (NT_0)^\delta (NLT_0 + NL^{7/4}T_0^{3/4})) \max_n |a_n|^2$$

holds for any fixed $\delta > 0$.

This is Proposition 1 of [21]. This result was also used by Heath-Brown in [20] to give a bound on the squares of gaps between smooth numbers and in [21] to improve a bound on the number of very long prime gaps. Simultaneously to our work, the latter bound was further improved by Järviniemi [26], who showed that at most $O(x^{0.57+\varepsilon})$ intevals $[y, y + y^{1/2}]$ with $y \in [x, 2x] \cap \mathbb{N}$ do not contain primes. Recently, Matomäki and Teräväinen [28] also used Heath Brown's sparse Dirichlet polynomial bound to improve a bound on the length of intervals which almost surely contain products of exactly two primes.

We use Lemma 2.16 to deduce the following bound on $Q(F, G, (S_i), (\sigma_i), \gamma, T)$:

Corollary 2.17. *Let $F(s) = \prod_{i=1}^{j} S_i(s)$ and $G(s)$ both be as described at the start of Section 2.4. In particular, $G(s) = \sum_{m\in\mathcal{J}} \xi_m m^{-s}$ with $L = \#\mathcal{J} \ge \tau^2 x^{-0.77}$.*
Let $I \subseteq \{1, \ldots, j\}$ and $k_i \in \mathbb{N} \cap [1, J]$. Set $M = \prod_{i\in I} N_i^{k_i}$ and $M^\beta = \prod_{i\in I} N_i^{k_i\sigma_i}$. Then for every $\delta > 0$,

$$Q(F, G, (S_i), (\sigma_i), \gamma, T) \ll_\delta x^\delta (M^{2-2\beta} L^{2-2\gamma} + M^{1-2\beta} L^{7/4-2\gamma} T_0^{3/4}).$$

Proof. Let $\mathcal{S}_1(F,(S_i),(\sigma_i))$ and $\mathcal{S}_2(G,\gamma)$ be as given in Definition 2.3. For each $k \in \mathcal{S}_1 \cap \mathcal{S}_2 \cap [T,2T]$, select $t_k^{(i)} \in [k,k+1]$ and $t_k^{(G)} \in [k,k+1]$ with $|S_i(c+it_k^{(i)})| \geq N_i^{-c+\sigma_i}$ and $|G(it_k^{(G)})| \geq L^\gamma$. By partial summation we find that for $t \in [k,k+1]$

$$\begin{aligned} S_i(c+it_k^{(i)}) &= \sum_{n\sim N_i} a_n^{(i)} n^{-c-it_k^{(i)}} \\ &= (2N_i)^{-c-i(t_k^{(i)}-t)} \sum_{n\sim N_i} a_n^{(i)} n^{-it} + O\left(\frac{1}{N_i^2}\int_{N_i}^{2N_i}\left|\sum_{n<y} a_n^{(i)} n^{-it}\right| \mathrm{d}y\right). \end{aligned}$$

We obtain a similar expression for $G(it_k^{(G)})$. Multiplying these together, we find

$$\begin{aligned} L^\gamma M^{-c+\beta} &\leq \left|G(it_k^{(G)})\prod_{i\in I} S_i(c+it_k^{(i)})^{k_i}\right| &\quad (2.79)\\ &\ll \left(\left|\sum_{m\in\mathcal{J}} \xi_m m^{-it}\right| + \int_1^{T_0}\frac{1}{y}\left|\sum_{\substack{m<y\\ m\in\mathcal{J}}}\xi_m m^{-it}\right|\mathrm{d}y\right) \\ &\quad \cdot \prod_{i\in I}\left(\frac{1}{N_i}\left|\sum_{n\sim N_i} a_n^{(i)} n^{-it}\right| + \frac{1}{N_i^2}\int_{N_i}^{2N_i}\left|\sum_{n<y} a_n^{(i)} n^{-it}\right|\mathrm{d}y\right)^{k_i} &\quad (2.80) \end{aligned}$$

and thus work with summands of the form

$$\left(\prod_{i\in I}\frac{1}{N_i^{k_i}}\prod_{i=1}^r\frac{1}{N_{i_j}}\right)\int_{N_{i_1}}^{2N_{i_1}}\cdots\int_{N_{i_r}}^{2N_{i_r}}\left|\sum_{m\in\mathcal{J}}\frac{\xi_m}{m^{it}}\right|\prod_{j=1}^r\left|\sum_{n<y_j}\frac{a_n^{(i_j)}}{n^{it}}\right|\prod_{j=r+1}^l\left|\sum_{n\sim N_i}\frac{a_n^{(i_j)}}{n^{it}}\right|\mathrm{d}y_1\ldots\mathrm{d}y_r$$

or similar products which involve $(1/y)\sum_{m<y}\xi_m m^{-it}$ integrated with respect to y rather than the full Dirichlet polynomial $\sum \xi_m m^{-it}$. Here $i_1,\ldots,i_l$ are such that each $i\in I$ appears exactly k_i times.

Next, we integrate (2.80) over $t\in[k,k+1]$ and sum k over all elements of $\mathcal{S}_1\cap\mathcal{S}_2\cap[T,2T]$. On the LHS we get $L^\gamma M^{-c+\beta}Q(F,G,(S_i),(\sigma_i),\gamma,T)$, while on the RHS we are now looking at summands of the form

$$\left(\prod_{i\in I}\frac{1}{N_i^{k_i}}\prod_{i=1}^r\frac{1}{N_{i_j}}\right)\int_{N_{i_1}}^{2N_{i_1}}\cdots\int_{N_{i_r}}^{2N_{i_r}}\sum_{k\in\mathcal{S}}\int_k^{k+1}\left|\sum_{m\in\mathcal{J}}\frac{\xi_m}{m^{it}}\right|\left|\sum_{n\leq 2^\alpha M}\frac{b_n(\mathbf{y})}{n^{it}}\right|\mathrm{d}t\,\mathrm{d}y_1\ldots\mathrm{d}y_r, \tag{2.81}$$

$$\text{where}\ \sum_{n\leq 2^\alpha M} b_n(\mathbf{y})n^{-it} = \prod_{j=1}^r\left(\sum_{n<y_j}a_n^{(i_j)}n^{-it}\right)\prod_{j=r+1}^l\left(\sum_{n\sim N_i}a_n^{(i_j)}n^{-it}\right),$$

or similar expressions in which $\sum\xi_m m^{-it}$ is replaced by $(1/y)\sum_{m<y}\xi_m m^{-it}$ and integrated over $[1,T_0]$. Here $\alpha=\sum_{i\in I}k_i$ and $|b_n(\mathbf{y})|\ll_\delta n^\delta$ for $\delta>0$ and $\mathcal{S}=$

$\mathcal{S}_1(F,(S_i),(\sigma_i)) \cap \mathcal{S}_2(G,\gamma) \cap [T,2T]$. Focusing only on the innermost integral, the Cauchy-Schwarz inequality gives

$$\sum_{k\in\mathcal{S}} \int_k^{k+1} \left|\sum_{\substack{m<y\\ m\in\mathcal{J}}} \frac{\xi_m}{m^{it}}\right| \left|\sum_{n\le 2^\alpha M} \frac{b_n(\mathbf{y})}{n^{it}}\right| \mathrm{d}t \le Q^{1/2}\left(\int_0^{2T_0} \left|\sum_{\substack{m<y\\ m\in\mathcal{J}}} \frac{\xi_m}{m^{it}}\right|^2 \left|\sum_{n\le 2^\alpha M} \frac{b_n(\mathbf{y})}{n^{it}}\right|^2 \mathrm{d}t\right)^{1/2},$$

where $Q = Q(F,G,(S_i),(\sigma_i),\gamma,T)$. Lemma 2.16 applies to the RHS and hence (2.81) is bounded above by $O_\delta(Q^{1/2}x^\delta(1/M)(M^2L^2+MLT_0+ML^{7/4}T_0^{3/4})^{1/2})$. Returning to (2.79), which we integrated over $[k,k+1]$, $k\in\mathcal{S}$, we have $L^\gamma M^{-c+\beta}Q$ on the LHS and a finite number of multiple integrals of the form given in (2.81) on the RHS. So

$$Q(F,G,(S_i),(\sigma_i),\gamma,T)L^{2\gamma}M^{-2c+2\beta} \ll_\delta x^\delta M^{-2}(M^2L^2+MLT_0+ML^{7/4}T_0^{3/4}).$$

Since $\tau > x^{0.475}$ and $L > \tau^2 x^{-0.77}$, we have $L > \tau^{1/3}$ and $MLT_0 < ML^{7/4}T_0^{3/4}$. Then

$$Q(F,G,(S_i),(\sigma_i),\gamma,T) \ll_\delta x^\delta(M^{2-2\beta}L^{2-2\gamma} + M^{1-2\beta}L^{7/4-2\gamma}T_0^{3/4})$$

for $\delta > 0$. This is what we wanted to show. $\square$

We now use Corollary 2.17 to replace conditions (C4.A) and (C4.B) by bounds on R.

Corollary 2.18. *Let $F(s) = \prod_{i=1}^j S_i(s)$ and $G(s) = \sum_{m\in\mathcal{J}} \xi_m m^{-s}$ be as described at the start of Section* 2.4. *Suppose the following condition* (C4*) *holds:*

> *There exist $I \subseteq \{1,\ldots,j\}$ and $k_i \in \mathbb{N}\cap[1,J]$ such that $M = \prod_{i\in I} N_i^{k_i}$ and $M^\beta = \prod_{i\in I} N_i^{k_i\sigma_i}$ satisfy each of the following two inequalities:* (C4*)

$$R(F,(S_i),(\sigma_i),T) \le \tau^{-1/4}M^{2\beta-1}x^{7/4-2\sigma+\nu/4-4\varepsilon_1}, \tag{2.82}$$
$$R(F,(S_i),(\sigma_i),T) \le M^{2\beta-2}x^{2-2\sigma-4\varepsilon_1}. \tag{2.83}$$

Then $Q(F,G,(S_i),(\sigma_i),\gamma,T) \le \max\{(\#\mathcal{J})^{7/8-\gamma}\tau^{1/4}x^{7/8-\sigma+\nu/8+\varepsilon_1}, (\#\mathcal{J})^{1-\gamma}x^{1-\sigma-\varepsilon_1}\}$.

Proof. Corollary 2.17 tells us that one of the following inequalities holds:

$$Q(F,G,(S_i),(\sigma_i),\gamma,T) \le x^{\varepsilon_1}M^{1-2\beta}(\#\mathcal{J})^{7/4-2\gamma}T_0^{3/4}, \tag{2.84}$$
$$Q(F,G,(S_i),(\sigma_i),\gamma,T) \le x^{\varepsilon_1}M^{2-2\beta}(\#\mathcal{J})^{2-2\gamma}. \tag{2.85}$$

Note that $Q = Q(F,G,(S_i),(\sigma_i),\gamma,T) \le R = R(F,(S_i),(\sigma_i),T)$. First assume we have (2.84). Then

$$Q(F,G,(S_i),(\sigma_i),\gamma,T) \le Q^{1/2}R^{1/2} \le x^{\varepsilon_1/2}M^{1/2-\beta}(\#\mathcal{J})^{7/8-\gamma}T_0^{3/8}R^{1/2}. \tag{2.86}$$

But the RHS of (2.86) is at most $(\#\mathcal{J})^{7/8-\gamma}\tau^{1/4}x^{7/8-\sigma+\nu/8+\varepsilon_1}$ when

$$R(F,(S_i),(\sigma_i),T) \le \tau^{-1/4}M^{2\beta-1}x^{7/4-2\sigma+\nu/4+\varepsilon_1/4}.$$

Now assume we have (2.85), so that

$$Q(F,G,(S_i),(\sigma_i),\gamma,T) \le Q^{1/2}R^{1/2} \le x^{\varepsilon_1/2}M^{1-\beta}(\#\mathcal{J})^{1-\gamma}R^{1/2}. \tag{2.87}$$

But the RHS of (2.87) is at most $(\#\mathcal{J})^{1-\gamma}x^{1-\sigma-\varepsilon_1}$ when

$$R(F,(S_i),(\sigma_i),T) \le M^{2\beta-2}x^{2-2\sigma-4\varepsilon_1}. \qquad \square$$

2.4.4 Computations

Let $F(s)=\prod_{i=1}^{j} S_i(s)$, $G(s)$, (σ_i) and γ be as described at the start of Section 2.4 with $F(s)$ of Type A, B or C at T. Recall that $S_i(s)=\sum_{n\sim N_i} a_n^{(i)}n^{-s}$ with $N=\prod_{i=1}^{j} N_i$ and $N_i=N^{\ell_i}$. We now consider $F(s)$ which satisfy one of the following three criteria:

(1) There exists $k\in\{1,\dots,j\}$ with $\ell_k \ge \chi_0(a)$.

(2) There exists $I_1 \subseteq \{1,\dots,j\}$ with $\sum_{i\in I_1}\ell_i \in \chi_1(a)$.

(3) There exist $I_2, I_3 \subseteq \{1,\dots,j\}$ with $\sum_{i\in I_2}\ell_i \in \chi_2(a)$ and $\sum_{i\in I_3}\ell_i \in \chi_3(a)$.

Here $\chi_0(a)$, $\chi_1(a)$, $\chi_2(a)$ and $\chi_3(a)$ are as given in (2.9), (2.10), (2.11) and (2.12).

To prove Proposition 2.2, we need to show that one of the inequalities (C1), (C2), (C3), (C4.A) or (C4.B) of Proposition 2.1 holds. If $F(s)$ is of Type C at T, then Lemma 2.9 immediately tells us that $R(F,(S_i),(\sigma_i),T)=0$ and these bounds are trivial. Hence we focus on $F(s)$ of Type A or B. We use Corollary 2.15 and Corollary 2.18 to replace (C3), (C4.A) and (C4.B) by conditions (C3*) and (C4*). Write

$$\begin{aligned}\mathcal{Z} =\big\{&(1,4,4,3,4,1),\ (1/2,2,3/2,3/2,2,1/2),\ (2/5,16/5,12/5,12/5,16/5,4/5),\\ &(2/5,4/5,3/5,6/5,8/5,2/5),\ (1/3,2/3,1/3,1,4/3,1/3),\\ &(2/7,16/21,8/21,8/7,32/21,8/21),\ (3/8,2,3/2,15/8,5/2,5/8),\\ &(1/4,4/3,2/3,5/4,5/3,5/12),\ (1/9,16/9,8/9,5/3,20/9,5/9)\big\}.\end{aligned} \tag{2.88}$$

The proof of Proposition 2.2 is complete if we can show for large x and given $F(s)=\prod_{i=1}^{j} S_i(s)$ and (σ_i) that there exist $I\subseteq\{1,\dots,j\}$ and $k_i\in\mathbb{N}\cap[1,J]$ such that

$M = \prod_{i\in I} N_i^{k_i}$ and $M^\beta = \prod_{i\in I} N_i^{k_i\sigma_i}$ satisfy one of the following four conditions:

$$R(F,(S_i),(\sigma_i),T) \le x^{1-\sigma}(\log x)^{-2J}, \tag{C1}$$
$$R(F,(S_i),(\sigma_i),T) \le \tau x^{1-2\sigma+\nu+2\varepsilon_1}, \tag{C2}$$
$$R(F,(S_i),(\sigma_i),T) \le \min_{(U,V,W,X,Y,Z)\in\mathcal{Z}} \tau^U M^{V\beta-W} x^{X-Y\sigma+Z\nu+\varepsilon_1}, \tag{C3*}$$
$$R(F,(S_i),(\sigma_i),T) \le \min\{\tau^{-1/4}M^{2\beta-1}x^{7/4-2\sigma+\nu/4-4\varepsilon_1}, M^{2\beta-2}x^{2-2\sigma-4\varepsilon_1}\}. \tag{C4*}$$

To accomplish this goal we use the many bounds on R introduced in Section 2.4.1.

2.4.4.1 Long factors

When $F(s) = \prod_{i=1}^{j} S_i(s)$ is of Type B at T, all its factors $S_i(s)$ are very short. However, if it is of Type A, there are some longer factors and these have property (Z1) or (Z2). We now show that one of (C1), (C2) or (C4*) holds if one of these factors is very long.

Lemma 2.19. *Let $F(s) = \prod_{i=1}^{j} S_i(s)$ be as described at the start of Section 2.4. Recall $S_i(s) = \sum_{n\sim N_i} a_n^{(i)} n^{-s}$ and $N_i = N^{\ell_i}$ with $x \ll N = \prod_{i=1}^{j} N_i \ll x$. Let $T \in [T_1, T_0]$ and suppose that $F(s)$ is of Type A at T. Write $\tau = x^a$. Suppose there exists $k \in \{1,\dots,j\}$ such that one of the following three options holds:*

1. *$\ell_k \ge 0.335 - \varepsilon_1$ and $a \le 0.57$,*
2. *$\ell_k \ge 0.330 - \varepsilon_1$ and $a \in [0.57, 0.61]$,*
3. *$\ell_k \ge 0.320 - \varepsilon_1$ and $a \ge 0.61$.*

Then if $x \ge C$, one of the following three inequalities holds:

$$R(F,(S_i),(\sigma_i),T) \le x^{1-\sigma}(\log x)^{-2J}, \tag{2.89}$$
$$R(F,(S_i),(\sigma_i),T) \le \tau x^{1-2\sigma+\nu+2\varepsilon_1}, \tag{2.90}$$
$$R(F,(S_i),(\sigma_i),T) \le \min\{\tau^{-1/4}N_k^{2\sigma_k-1}x^{7/4-2\sigma+\nu/4-4\varepsilon_1}, N_k^{2\sigma_k-2}x^{2-2\sigma-4\varepsilon_1}\}. \tag{2.91}$$

Here C is a large constant dependent only on J and $\varepsilon_1 = \varepsilon/10^5$.

This completes the proof of Proposition 2.2 when $a > 0.545$ and $F(s) = \prod_{i=1}^{j} S_i(s)$ has $\ell_k \ge \chi_0(a)$ for some k. For smaller values of a we still need to cover the case $\ell_k \in [\chi_0(a), 0.335]$ in a later lemma.

Proof. If $\sigma \geq 1 - 10^{-500}$, we are done by Lemma 2.11. Thus we consider $\sigma \in [0.6, 1 - 10^{-500}]$. Since $F(s)$ is of Type A at T and $S_k(s)$ has $\ell_k \geq 0.32 - \varepsilon_1$, $S_k(s)$ has property (Z1) or (Z2) with respect to T. Hence Lemma 2.10 applies. We assume $R(F, (S_i), (\sigma_i), T) > x^{1-\sigma}(\log x)^{-2J}$, as otherwise (2.89) holds. Taking $B = 2J$ in Lemma 2.10, we then get

$$R(F, (S_i), (\sigma_i), T) \leq x^{4\varepsilon_1} \min\left\{\tau N_k^{(2-4\sigma_k)}, \tau^2 N_k^{(6-12\sigma_k)}\right\}. \tag{2.92}$$

Write $B(a, \sigma) = \max\{1 - \sigma - \varepsilon_1, a + 1 - 2\sigma + \nu + 2\varepsilon_1\}$. If one of the inequalities

$$N_k^{\sigma_k} \geq \left(\frac{x^{4\varepsilon_1}\tau N_k^2}{x^{B(a,\sigma)}}\right)^{1/4} \quad \text{or} \quad N_k^{\sigma_k} \geq \left(\frac{x^{4\varepsilon_1}\tau^2 N_k^6}{x^{B(a,\sigma)}}\right)^{1/12} \tag{2.93}$$

is satisfied, (2.92) gives $R(F, (S_i), (\sigma_i), T) \leq x^{B(a,\sigma)} = \max\{x^{1-\sigma-\varepsilon_1}, x^{a+1-2\sigma+\nu+2\varepsilon_1}\}$. Suppose first that $\sigma_k \geq \sigma$. Then (2.93) holds if $\ell_k \geq \frac{(a/4 - B(a,\sigma)/4 + 2\varepsilon_1)}{(\sigma - 1/2)}$. The RHS of this inequality is at most 0.5 whenever $a \in [0.47, 0.77]$ and $\sigma \in [0.6, 1 - 10^{-500}]$. Hence we may assume $1 - \ell_k \geq 0.2$ or $\sigma_k < \sigma$ and $\sigma_I \geq \sigma$ for $I = \{1, \ldots, j\} \setminus \{k\}$. Either way, Corollary 2.13 is applicable with $I = \{1, \ldots, j\} \setminus \{k\}$. We thus have (2.89) or (2.90) if

$$\sigma_I \geq \min\left\{\frac{3a - 2B(a,\sigma) + 4\varepsilon_1}{3a - B(a,\sigma) + 2\varepsilon_1}, \frac{3a + B(a,\sigma) - 2\varepsilon_1}{3a + 3B(a,\sigma) - 6\varepsilon_1}\right\}.$$

So one of (2.89) or (2.90) holds whenever one of the following four inequalities holds:

$$\ell_k \sigma_k \geq \frac{a}{4} + \frac{\ell_k}{2} - \frac{B(a,\sigma)}{4} + 2\varepsilon_1, \tag{2.94}$$

$$\ell_k \sigma_k \geq \frac{a}{6} + \frac{\ell_k}{2} - \frac{B(a,\sigma)}{12} + 2\varepsilon_1, \tag{2.95}$$

$$(1 - \ell_k)\sigma_I \geq (1 - \ell_k)\left(\frac{3a + B(a,\sigma) - 2\varepsilon_1}{3a + 3B(a,\sigma) - 6\varepsilon_1}\right) = (1 - \ell_k) C_1(a, \sigma), \tag{2.96}$$

$$(1 - \ell_k)\sigma_I \geq (1 - \ell_k)\left(\frac{3a - 2B(a,\sigma) + 4\varepsilon_1}{3a - B(a,\sigma) + 2\varepsilon_1}\right) = (1 - \ell_k) C_2(a, \sigma). \tag{2.97}$$

Note $\ell_k \sigma_k + (1 - \ell_k)\sigma_I = \sigma$. Adding ((2.94) and (2.96)) or ((2.94) and (2.97)) or ((2.95) and (2.96)), we find that one of (2.94), (2.95), (2.96) or (2.97) certainly holds whenever one of

$$\sigma \geq \left(\tfrac{a}{4} - \tfrac{B(a,\sigma)}{4} + C_1(a, \sigma) + 2\varepsilon_1\right) + \left(\tfrac{1}{2} - C_1(a, \sigma)\right)\ell_k, \tag{2.98}$$

$$\sigma \geq \left(\tfrac{a}{4} - \tfrac{B(a,\sigma)}{4} + C_2(a, \sigma) + 2\varepsilon_1\right) + \left(\tfrac{1}{2} - C_2(a, \sigma)\right)\ell_k, \tag{2.99}$$

$$\sigma \geq \left(\tfrac{a}{6} - \tfrac{B(a,\sigma)}{12} + C_1(a, \sigma) + 2\varepsilon_1\right) + \left(\tfrac{1}{2} - C_1(a, \sigma)\right)\ell_k, \tag{2.100}$$

is satisfied. Note that the coefficient of ℓ_k in (2.98) and (2.100) is negative if $-3a/2 + B/2 - \varepsilon_1 < 0$. This is certainly true for $\sigma \in [0.6, 1]$. The coefficient of ℓ_k in (2.99) is negative if $-3a/2 + 3B/2 - 3\varepsilon_1 < 0$ and hence if $B(a,\sigma) < a + 2\varepsilon_1$. But if $B(a,\sigma) \geq a + 2\varepsilon_1$, then certainly $R(F, (S_i), (\sigma_i), T) \leq T_0 = x^{a+\varepsilon_1} \leq x^{B(a,\sigma)}$. Thus we may assume that all coefficients of ℓ_k are negative. Rearranging, we get that one of (2.98), (2.99) or (2.100) holds whenever $\ell_k \geq \min\{X_1(a,\sigma), X_2(a,\sigma), X_3(a,\sigma)\}$, where

$$\begin{aligned}
X_1(a,\sigma) &= \left(\sigma - \tfrac{a}{4} + \tfrac{B(a,\sigma)}{4} - C_1(a,\sigma) - 2\varepsilon_1\right)\left(\tfrac{1}{2} - C_1(a,\sigma)\right)^{-1}, \\
X_2(a,\sigma) &= \left(\sigma - \tfrac{a}{4} + \tfrac{B(a,\sigma)}{4} - C_2(a,\sigma) - 2\varepsilon_1\right)\left(\tfrac{1}{2} - C_2(a,\sigma)\right)^{-1}, \\
X_3(a,\sigma) &= \left(\sigma - \tfrac{a}{6} + \tfrac{B(a,\sigma)}{12} - C_1(a,\sigma) - 2\varepsilon_1\right)\left(\tfrac{1}{2} - C_1(a,\sigma)\right)^{-1}.
\end{aligned}$$

Since condition $\ell_k \geq \min\{X_1(a,\sigma), X_2(a,\sigma), X_3(a,\sigma)\}$ only involves the two variables a and σ, it is easy to check explicitly that $\min\{X_1(a,\sigma), X_2(a,\sigma), X_3(a,\sigma)\} \leq 0.3 - \varepsilon_1$ if $a \geq 0.65$. For smaller a we still need to do a bit more work. Using (2.92), we note that (2.91) holds whenever

$$x^{4\varepsilon_1} \tau N_k^{(2-4\sigma_k)} \leq \min\{\tau^{-1/4} N_k^{2\sigma_k - 1} x^{7/4 - 2\sigma + \nu/4 - 4\varepsilon_1}, N_k^{2\sigma_k - 2} x^{2 - 2\sigma - 4\varepsilon_1}\}. \tag{2.101}$$

Rearranging a bit, we find that (2.101) is satisfied if both inequalities

$$\ell_k \sigma_k \geq \frac{5a}{24} - \frac{7}{24} + \frac{\sigma}{3} - \frac{\nu}{24} + 2\varepsilon_1 + \frac{\ell_k}{2}, \tag{2.102}$$

$$\ell_k \sigma_k \geq \frac{a}{6} - \frac{1}{3} + \frac{\sigma}{3} + 2\varepsilon_1 + \frac{2\ell_k}{3} \tag{2.103}$$

hold. Adding ((2.102) and (2.96)) and ((2.103) and (2.96)), we are done if the following is true:

$$\sigma \geq \left(\frac{5a}{24} - \frac{7}{24} + \frac{\sigma}{3} - \frac{\nu}{24} + 2\varepsilon_1 + C_1(a,\sigma)\right) + \left(\frac{1}{2} - C_1(a,\sigma)\right)\ell_k, \tag{2.104}$$

$$\sigma \geq \left(\frac{a}{6} - \frac{1}{3} + \frac{\sigma}{3} + 2\varepsilon_1 + C_1(a,\sigma)\right) + \left(\frac{2}{3} - C_1(a,\sigma)\right)\ell_k. \tag{2.105}$$

The coefficient of ℓ_k in (2.105) is only non-negative if $B(a,\sigma) \geq a + 2\varepsilon_1$ and hence we may again assume that all coefficients of ℓ_k are negative. Rearranging, (2.91) holds if $\ell_k \geq \max\{Y_1(a,\sigma), Y_2(a,\sigma)\}$ where

$$\begin{aligned}
Y_1(a,\sigma) &= \frac{\left(\sigma - \left(\frac{5a}{24} - \frac{7}{24} + \frac{\sigma}{3} - \frac{\nu}{24} + 2\varepsilon_1 + C_1(a,\sigma)\right)\right)}{\left(\frac{1}{2} - C_1(a,\sigma)\right)}, \\
Y_2(a,\sigma) &= \frac{\left(\sigma - \left(\frac{a}{6} - \frac{1}{3} + \frac{\sigma}{3} + 2\varepsilon_1 + C_1(a,\sigma)\right)\right)}{\left(\frac{2}{3} - C_1(a,\sigma)\right)}.
\end{aligned}$$

In summary, we have one of (2.89), (2.90) or (2.91) whenever

$$\ell_k \geq \min\{X_1(a,\sigma),X_2(a,\sigma),X_3(a,\sigma),\max\{Y_1(a,\sigma),Y_2(a,\sigma)\}\}.$$

Evaluating the expression $\min\{X_1(a,\sigma), X_2(a,\sigma), X_3(a,\sigma), \max\{Y_1(a,\sigma), Y_2(a,\sigma)\}\}$ explicitly, we see that the quantity is strictly less than 0.335 when $a \leq 0.57$, strictly less than 0.33 when $a \in [0.57, 0.61]$ and strictly less than 0.32 when $a > 0.61$. $\square$

2.4.4.2 Large τ

Next we deal with particularly large values of τ.

Lemma 2.20. *Let $F(s) = \prod_{i=1}^{j} S_i(s)$ be as described at the start of Section 2.4. Recall $S_i(s) = \sum_{n\sim N_i} a_n^{(i)} n^{-s}$ and $N_i = N^{\ell_i}$ with $x \ll N = \prod_{i=1}^{j} N_i \ll x$. Let $T \in [T_1, T_0]$ and suppose that $F(s)$ is of Type A or B at T. Write $\tau = x^a$. Suppose $a \geq 0.685$ and there exists $I \subseteq \{1, \dots, j\}$ such that*

$$\sum_{i\in I} \ell_i \in [0.35 - \varepsilon_1, 0.48 + \varepsilon_1] \cup [0.52 - \varepsilon_1, 0.65 + \varepsilon_1].$$

Then if $x \geq C$, one of the following two inequalities holds:

$$R(F, (S_i), (\sigma_i), T) \leq x^{1-\sigma}(\log x)^{-2J}, \tag{2.106}$$

$$R(F, (S_i), (\sigma_i), T) \leq \tau x^{1-2\sigma+\nu+2\varepsilon_1}. \tag{2.107}$$

Here C is a large constant dependent only on J and ε_1, and ε_1 is assumed to be small.

Lemma 2.20 completes the proof of Proposition 2.2 when $a \geq 0.685$.

Proof. If $\sigma \geq 1 - 10^{-500}$, we are done by Lemma 2.11. Thus we consider $\sigma \in [0.6, 1 - 10^{-500}]$. Suppose $R(F, (S_i), (\sigma_i), T) > x^{1-\sigma}(\log x)^{-2J}$. Write $B(a, \sigma) = \max\{1 - \sigma - \varepsilon_1, a + 1 - 2\sigma + \nu + 2\varepsilon_1\}$. By Corollary 2.13 we have $R(F, (S_i), (\sigma_i), T) \leq x^{B(a,\sigma)}$ if

$$\sigma \geq \min\left\{\frac{3a - 2B(a,\sigma) + 4\varepsilon_1}{3a - B(a,\sigma) + 2\varepsilon_1}, \frac{3a + B(a,\sigma) - 2\varepsilon_1}{3a + 3B(a,\sigma) - 6\varepsilon_1}\right\}. \tag{2.108}$$

This inequality condition can be rewritten in terms of four simple quadratic equations in σ. Solving these quadratic equations, we find that (2.108) holds whenever $a \in [0.685, 0.77]$ and $\sigma \in [0.6, 0.88]$.

For $\sigma > 0.88$, we use bound (2.47) of Lemma 2.8: For $\ell_I = \sum_{i\in I} \ell_i$ and $\sigma_I = \frac{\sum_{i\in I} \ell_i \sigma_i}{\ell_I}$,

$$R(F, (S_i), (\sigma_i), T) \leq \max\left\{x^{\ell_I(2-2\sigma_I)+2\varepsilon_1}, x^{a+\ell_I(11-14\sigma_I)+2\varepsilon_1}\right\}.$$

We thus have $R(F,(S_i),(\sigma_i),T) \leq x^{B(a,\sigma)}$ if

$$\sigma_I \geq \frac{1}{\ell_I} \max\left\{\ell_I - \frac{B(a,\sigma)}{2},\ \frac{11}{14}\ell_I + \frac{a}{14} - \frac{B(a,\sigma)}{14}\right\} + \varepsilon_1. \tag{2.109}$$

However, for $\ell_I \in [0.35-\varepsilon_1, 0.48+\varepsilon_1]$, the RHS of (2.109) is at most $\max\{\sigma, 0.92\}$. Hence we may assume $\sigma_I < \max\{\sigma, 0.92\}$, so that $\frac{(\sigma-\ell_I\sigma_I)}{(1-\ell_I)} \geq \min\{\sigma, 2(\sigma-0.46)\} \geq 0.84$ for $\sigma \geq 0.88$. But Corollary 2.13, applied with $K = \{1,\ldots,j\} \setminus I$ in the place of I, then gives $R(F,(S_i),(\sigma_i),T) \leq x^{B(a,\sigma)}$ if

$$\sigma_K \geq \max\left\{\frac{3a+7B(a,\sigma)-14\varepsilon_1}{3a+10B(a,\sigma)-20\varepsilon_1}, \frac{4a+B(a,\sigma)-2\varepsilon_1}{4a+4B(a,\sigma)-8\varepsilon_1}\right\} = F(a,\sigma). \tag{2.110}$$

Multiplying (2.109) by ℓ_I and (2.110) by $(1-\ell_I)$ and adding up, we are hence done if

$$\ell_I \leq \frac{1}{1-F(a,\sigma)}\left(\sigma + \frac{B(a,\sigma)}{2} - F(a,\sigma) - \varepsilon_1\right) \quad \text{and}$$
$$\ell_I \geq \frac{1}{11/14 - F(a,\sigma)}\left(\sigma - \frac{a}{14} + \frac{B(a,\sigma)}{14} - F(a,\sigma) - \varepsilon_1\right).$$

Both inequalities are satisfied if $a \geq 0.685$, $\sigma \geq 0.88$ and $\ell_I \in [0.35-\varepsilon_1, 0.48+\varepsilon_1]$. □

2.4.4.3 Small τ

Next we work with particularly small τ, say $\tau \leq x^{0.53}$, or large σ, say $\sigma \geq a+\nu-\varepsilon_1$. Here condition (C4*), which was derived using $Q(F,G,(S_i),(\sigma_i),\gamma,T)$, is very useful.

Lemma 2.21. *Let $F(s) = \prod_{i=1}^{j} S_i(s)$ be as described at the start of Section 2.4. Recall $S_i(s) = \sum_{n\sim N_i} a_n^{(i)} n^{-s}$ and $N_i = N^{\ell_i}$. Let $T \in [T_1, T_0]$ and suppose that $F(s)$ is of Type A or B at T. Write $\tau = x^a$.*

Suppose there exists $I \subseteq \{1,\ldots,j\}$ such that I, a and σ satisfy one of the following five options:

(1) $a \in [0.47, 0.53]$ *and* $\sum_{i\in I} \ell_i \in [0.29-\varepsilon_1, 0.36+\varepsilon_1] \cup [0.64-\varepsilon_1, 0.71+\varepsilon_1]$,

(2) $a \in [0.53, 0.545]$ *and* $\sum_{i\in I} \ell_i \in [0.315-\varepsilon_1, 0.345+\varepsilon_1] \cup [0.655-\varepsilon_1, 0.685+\varepsilon_1]$,

(3) $a \in [0.53, 0.545]$, $\sum_{i\in I} \ell_i \in [0.285-\varepsilon_1, 0.375+\varepsilon_1] \cup [0.625-\varepsilon_1, 0.715+\varepsilon_1]$ *and* $\sigma \geq a+\nu-\varepsilon_1$,

(4) $a \in [0.53, 0.685]$ *and* $\sigma \geq \min\{a+\nu-\varepsilon_1, 0.85\}$ *and*

$$\sum_{i\in I} \ell_i \in \left[1-m(a,\sigma)+10^{-100}, 1-M(a,\sigma)-10^{-100}\right], \tag{2.111}$$

where $M(a,\sigma)$ and $m(a,\sigma)$ are given by:

$$M(a,\sigma) = \min\left\{\frac{3a}{6\sigma-2}, \max\left\{\frac{3a}{20\sigma-14}, \frac{2a}{4\sigma-1}\right\}\right\}, \tag{2.112}$$

$$m(a,\sigma) = \max\{m_1(a,\sigma), \min\{m_2(a,\sigma), m_3(a,\sigma)\}\}, \tag{2.113}$$

$$m_1(a,\sigma) = \min_{s\in\{\sigma,\sigma_1^\circ\}}\left\{\frac{a-3-\nu}{4(1-2s)} + \frac{a(3-3s)}{(3s-1)(1-2s)}\right\}, \tag{2.114}$$

$$m_2(a,\sigma) = \min_{s\in\{\sigma,\sigma_2^\circ\}}\left\{\frac{a-3-\nu}{4(1-2s)} + \frac{a(3-3s)}{(10s-7)(1-2s)}\right\}, \tag{2.115}$$

$$m_3(a,\sigma) = \min_{s\in\{\sigma,\sigma_2^\circ\}}\left\{\frac{a-3-\nu}{4(1-2s)} + \frac{a(4-4s)}{(4s-1)(1-2s)}\right\} \tag{2.116}$$

$$\textit{and} \quad \sigma_1^\circ = \min\left\{1,\, a+\tfrac{1}{3}\right\} \quad \textit{and} \quad \sigma_2^\circ = \max\left\{\tfrac{3a}{10}+\tfrac{7}{10},\, a+\tfrac{1}{4}\right\},$$

(5) $a \in [0.53, 0.685]$ *and σ satisfies* $\sigma \geq \min\left\{a+\frac{1}{3}, \max\left\{\frac{3a}{10}+\frac{7}{10}, a+\frac{1}{4}\right\}\right\} + 10^{-100}$.

Write $M = \prod_{i\in I} N_i$ and $M^\beta = \prod_{i\in I} N_i^{\sigma_i}$. Then if $x \geq C$, one of the following two inequalities holds:

$$R(F,(S_i),(\sigma_i),T) \leq x^{1-\sigma}(\log x)^{-2J}, \tag{2.117}$$

$$R(F,(S_i),(\sigma_i),T) \leq \min\{\tau^{-1/4}M^{2\beta-1}x^{7/4-2\sigma+\nu/4-4\varepsilon_1}, M^{2\beta-2}x^{2-2\sigma-4\varepsilon_1}\}. \tag{2.118}$$

Here C is a large constant dependent only on J and ε_1, and ε_1 is assumed to be small.

In particular, Lemma 2.21 completes the proof of Proposition 2.2 for $a \leq 0.53$ and also combines with Lemma 2.19 to fully cover the case $\ell_k \geq \chi_0(a)$ when $a \leq 0.545$.

Proof. We first consider an arbitrary $a \in [0.47, 0.685]$. If $\sigma \geq 1 - 10^{-500}$, we have $R(F,(S_i),(\sigma_i),T) \leq x^{1-\sigma}(\log x)^{-2J}$ by Lemma 2.11. Thus we consider $\sigma \in [0.6, 1-10^{-500}]$ and also assume $R(F,(S_i),(\sigma_i),T) > x^{1-\sigma}(\log x)^{-2J}$. (Otherwise (2.117) holds.) We ask when (2.118) is satisfied.

Case 1: $\tau^{-1/4}M^{2\beta-1}x^{7/4-2\sigma+\nu/4-4\varepsilon_1} \geq M^{2\beta-2}x^{2-2\sigma-4\varepsilon_1}$.

Write $\ell_I = \sum_{i\in I}\ell_i$ and $\frac{\sum_{i\notin I}\ell_i\sigma_i}{(1-\ell_I)} = \alpha$, so that $N^\sigma = (N/M)^\alpha M^\beta = N^{(1-\ell_I)\alpha}N^{\ell_I\beta}$. Since $x \ll N \ll x$ and $M = N^{\ell_I}$, further $M^{2\beta-2}x^{2-2\sigma-4\varepsilon_1} \geq x^{(2-2\alpha)(1-\ell_I)-5\varepsilon_1}$, provided x is sufficiently large.

Suppose $\alpha \geq \sigma$. By Lemma 2.12, applied with $\{1,\ldots,j\}\setminus I$ in the place of I, we have

$$R(F,(S_i),(\sigma_i),T) \leq x^{2\varepsilon_1}\min\left\{\tau^{(3-3\alpha)/(2-\alpha)}, \tau^{(3-3\alpha)/(3\alpha-1)}\right\}. \tag{2.119}$$

Hence $R \le x^{(2-2\alpha)(1-\ell_I)-5\varepsilon_1}$ if $\frac{2}{3}(1-\ell_I)\max\{(2-\alpha),(3\alpha-1)\} - a - C^*\varepsilon_1 \ge 0$, where C^* is a very large absolute constant. (Here we assumed that $\alpha < 1 - 10^{-1000}$. This assumption is valid as (2.119) immediately gives (2.117) if $\alpha > 1 - 10^{-1000}$ and $\sigma \le 1 - 10^{-500}$.) Rearranging, we are done if

$$\ell_I \le 1 - \min\left\{\frac{3a}{4-2\alpha}, \frac{3a}{6\alpha-2}\right\} - C^{**}\varepsilon_1.$$

C^{**} is another large constant. On the other hand, if $\alpha < \sigma$, then $x^{(2-2\alpha)(1-\ell_I)-5\varepsilon_1} \ge x^{(2-2\sigma)(1-\ell_I)-5\varepsilon_1}$. Lemma 2.12, applied with $\{1,\ldots,j\}$ in the place of I, tells us that

$$R(F,(S_i),(\sigma_i),T) \le x^{2\varepsilon_1}\min\left\{\tau^{(3-3\sigma)/(2-\sigma)}, \tau^{(3-3\sigma)/(3\sigma-1)}\right\}. \tag{2.120}$$

By the same computation as above, with σ in the place of α, $R(F,(S_i),(\sigma_i),T) \le x^{(2-2\sigma)(1-\ell_I)-5\varepsilon_1}$ if

$$\ell_I \le 1 - \min\left\{\frac{3a}{4-2\sigma}, \frac{3a}{6\sigma-2}\right\} - C^{**}\varepsilon_1.$$

Hence we have (2.118) if $\ell_I \le 1 - \min\{\frac{3a}{4-2\gamma}, \frac{3a}{6\gamma-2}\} - C^{**}\varepsilon_1$ holds for every $\gamma \in [\sigma, 1-10^{-1000}]$.

Case 1A: Now suppose that $a \in [0.47, 0.545]$.

Note that $1 - \min\{\frac{3a}{4-2\gamma}, \frac{3a}{6\gamma-2}\} - C^{**}\varepsilon_1$ is strictly greater than 0.36 if $a \le 0.53$. It is also strictly greater than 0.345 if $a \in [0.53, 0.545]$. Further, for ε sufficiently small and $\sigma \ge \sigma^*$, quantity $1-\min\{\frac{3a}{4-2\gamma}, \frac{3a}{6\gamma-2}\} - C^{**}\varepsilon_1$ is greater than $1 - 3a/(6\sigma^*-2) - 10^{-100}$. If $a \in [0.53, 0.545]$ and $\sigma^* = a + \nu - \varepsilon_1$, then $1 - 3a/(6\sigma^*-2) - 10^{-100}$ is greater than 0.375. This covers options (1), (2) and (3) for Case 1.

Case 1B: Now consider $a \in [0.53, 0.685]$ and $\sigma \ge \min\{a+\nu-\varepsilon_1, 0.85\}$.

Inequality $\ell_I \le 1 - \min\{\frac{3a}{4-2\gamma}, \frac{3a}{6\gamma-2}\} - C^{**}\varepsilon_1$ holds for all $\gamma \in [\sigma, 1-10^{-1000}]$ if $\ell_I \le 1 - 3a/(6\sigma-2) - 10^{-100}$ and in that case (2.118) is satisfied. On the other hand, Lemma 2.12 also gives

$$R(F,(S_i),(\sigma_i),T) \le x^{2\varepsilon_1}\max\left\{\tau^{(3-3\alpha)/(10\alpha-7)}, \tau^{(4-4\alpha)/(4\alpha-1)}\right\}. \tag{2.121}$$

We still desire $R(F,(S_i),(\sigma_i),T) \le x^{(2-2\alpha)(1-\ell_I)-5\varepsilon_1}$ and (2.121) provides this bound whenever

$$\ell_I \le 1 - \max\left\{\frac{3a}{20\alpha-14}, \frac{2a}{4\alpha-1}\right\} - C^{**}\varepsilon_1. \tag{2.122}$$

The RHS of (2.122) is increasing in α, so if $\alpha \geq \sigma$, we only need to assume that this inequality holds with $\alpha = \sigma$ to obtain (2.118). For $\alpha < \sigma$ we replace α by σ in the application of Lemma 2.12 and arrive at the same conclusion. In summary, upper bound $\ell_I \leq 1 - M(a,\sigma) - 10^{-100}$ suffices to give (2.117) or (2.118) in the case $\sigma \geq \min\{a+\nu-\varepsilon_1, 0.85\}$ and $\tau^{-1/4}M^{2\beta-1}x^{7/4-2\sigma+\nu/4-4\varepsilon_1} \geq M^{2\beta-2}x^{2-2\sigma-4\varepsilon_1}$. This covers option (4) for Case 1.

Case 2: $\tau^{-1/4}M^{2\beta-1}x^{7/4-2\sigma+\nu/4-4\varepsilon_1} < M^{2\beta-2}x^{2-2\sigma-4\varepsilon_1}$.

Note $\tau^{-1/4}M^{2\beta-1}x^{7/4-2\sigma+\nu/4-4\varepsilon_1} \geq x^{(1-2\alpha)(1-\ell_I)-a/4+3/4+\nu/4-5\varepsilon_1}$. If $\alpha \geq \sigma \geq 0.6$, we apply Lemma 2.12 with $\{1,\dots,j\}\backslash I$ in the place of I to obtain (2.119) and additionally note $x^{1-\sigma}(\log x)^{-2J} \geq x^{1-\alpha}(\log x)^{-2J}$. If $\alpha < \sigma$, we note $x^{(1-2\alpha)(1-\ell_I)-a/4+3/4+\nu/4-5\varepsilon_1} \geq x^{(1-2\sigma)(1-\ell_I)-a/4+3/4+\nu/4-5\varepsilon_1}$ and apply Lemma 2.12 with $\{1,\dots,j\}$ in the place of I to obtain (2.120).

Either way, we have one of (2.117) or (2.118) if

$$x^{2\varepsilon_1}\min\left\{\tau^{(3-3\gamma)/(2-\gamma)}, \tau^{(3-3\gamma)/(3\gamma-1)}\right\} \leq \max\left\{x^{(1-2\gamma)(1-\ell_I)-a/4+3/4+\nu/4-5\varepsilon_1}, x^{1-\gamma-\varepsilon_1}\right\}$$

for every $\gamma \in [\sigma, 1-10^{-1000}]$. Rearranging, we require one of the following two inequalities to hold:

$$\ell_I \geq 1 - \frac{a-3-\nu}{4(1-2\gamma)} - \max\left\{\frac{a(3-3\gamma)}{(2-\gamma)(1-2\gamma)}, \frac{a(3-3\gamma)}{(3\gamma-1)(1-2\gamma)}\right\} + 35\varepsilon_1, \tag{2.123}$$

$$\gamma \geq a + \frac{1}{3} + C^{**}\varepsilon_1. \tag{2.124}$$

Case 2A: Again we first consider $a \in [0.47, 0.545]$. The RHS of (2.123) is strictly less than 0.29 if $a \leq 0.53$ and $\gamma \leq a + 0.34$. It is also strictly less than 0.315 if $a \in [0.53, 0.545]$ and $\gamma \leq a + 0.34$. For $a \in [0.53, 0.545]$ and $\gamma \in [a+\nu-\varepsilon_1, a+0.34]$, it is less than 0.285. This covers options (1), (2) and (3) for Case 2.

Case 2B: For $a \in [0.53, 0.685]$ and $\sigma \geq \min\{a+\nu-\varepsilon_1, 0.85\}$, the maximum of the RHS of (2.123) over $\gamma \in [\sigma, \min\{1-10^{-1000}, a+1/3+C^{**}\varepsilon_1\}]$ is attained at one of the endpoints of the interval. Hence, for ε_1 sufficiently small, (2.123) holds if

$$\ell_I \geq 1 - \min_{s\in\{\sigma,\sigma_1^\circ\}}\left\{\frac{a-3-\nu}{4(1-2s)} + \frac{a(3-3s)}{(3s-1)(1-2s)}\right\} + 10^{-100}. \tag{2.125}$$

Alternatively, using Lemma 2.12 with bounds (2.51) and (2.52) rather than (2.49) and (2.50), we also have one of (2.117) or (2.118) for $\sigma \geq \min\{a+\nu-\varepsilon_1, 0.85\}$ if

$$x^{2\varepsilon_1}\max\left\{\tau^{(3-3\gamma)/(10\gamma-7)}, \tau^{(4-4\gamma)/(4\gamma-1)}\right\} \leq \max\left\{x^{(1-2\gamma)(1-\ell_I)-a/4+3/4+\nu/4-5\varepsilon_1}, x^{1-\gamma-\varepsilon_1}\right\}$$

for every $\gamma \in [\sigma, 1-10^{-1000}]$. Rearranging, we require one of the following two inequalities to hold:

$$\ell_I \geq 1 - \frac{a-3-\nu}{4(1-2\gamma)} - \min\left\{\frac{a(3-3\gamma)}{(10\gamma-7)(1-2\gamma)}, \frac{a(4-4\gamma)}{(4\gamma-1)(1-2\gamma)}\right\} + 35\varepsilon_1, \quad (2.126)$$

$$\gamma \geq \max\left\{\frac{3a}{10} + \frac{7}{10}, a + \frac{1}{4}\right\} + C^{**}\varepsilon_1. \quad (2.127)$$

The RHS of (2.126) also attains its maximum over $\gamma \in [\sigma, \sigma_2^\circ + C^{**}\varepsilon_1]$ at one of the endpoints of the interval and this gives the lower bound $1 - \min\{m_2(a,\sigma), m_3(a,\sigma)\} + 10^{-100}$ for the size of ℓ_I, covering option (4) for Case 2.

Case C: Finally, let $a \in [0.53, 0.685]$ and $\sigma \geq \min\{a + \frac{1}{3}, \max\{\frac{3a}{10} + \frac{7}{10}, a + \frac{1}{4}\}\} + 10^{-100}$.

Here we take $I = \emptyset$, so that $\tau^{-1/4}M^{2\beta-1}x^{7/4-2\sigma+\nu/4-4\varepsilon_1} < M^{2\beta-2}x^{2-2\sigma-4\varepsilon_1}$ and we are back in Case 2. But every $\gamma \in [\sigma, 1-10^{-1000}]$ satisfies one of (2.124) or (2.127). This concludes the treatment of option (5) and the proof of Lemma 2.21. $\square$

2.4.4.4 Large σ

Having dealt with large and small τ, we now focus on large σ and midsized τ.

Lemma 2.22. *Let $F(s) = \prod_{i=1}^{j} S_i(s)$ be as described at the start of Section 2.4. Recall $S_i(s) = \sum_{n\sim N_i} a_n^{(i)} n^{-s}$ and $N_i = N^{\ell_i}$. Let $T \in [T_1, T_0]$ and suppose that $F(s)$ is of Type A or B at T. Write $\tau = x^a$.*
Finally, suppose we have one of the following two options:

(1) $a \in [0.545, 0.57]$ *and* $\sigma \geq a + \nu - \varepsilon_1$,

(2) $a \in [0.53, 0.545]$, $\sigma \geq a + \nu - \varepsilon_1$ *and* $\exists I$ *with* $\sum_{i\in I} \ell_i \in [0.427 - \varepsilon_1, 0.474 + \varepsilon_1]$.

Let $\mathcal{Z}$ be as defined in (2.88). Then if $x \geq C$, there exist $K \subseteq \{1, \ldots, j\}$ and $r \leq J$ such that $M = \prod_{i\in K} N_i^r$ and $M^\beta = \prod_{i\in K} N_i^{r\sigma_i}$ satisfy one of the following three inequalities:

$$R(F, (S_i), (\sigma_i), T) \leq x^{1-\sigma}(\log x)^{-2J}, \quad (2.128)$$

$$R(F, (S_i), (\sigma_i), T) \leq \min\{\tau^{-1/4}M^{2\beta-1}x^{7/4-2\sigma+\nu/4-4\varepsilon_1}, M^{2\beta-2}x^{2-2\sigma-4\varepsilon_1}\}, \quad (2.129)$$

$$R(F, (S_i), (\sigma_i), T) \leq \min_{(U,V,W,X,Y,Z)\in\mathcal{Z}} \tau^U M^{V\beta-W} x^{X-Y\sigma+Z\nu+\varepsilon_1}. \quad (2.130)$$

Here C is a large constant dependent only on J and ε_1, and ε_1 is assumed to be small.

Proof. We first work with an arbitrary $a \in [0.53, 0.57]$ and $\sigma \geq a + \nu - \varepsilon_1$. If $\sigma \geq 1 - 10^{-500}$, we are done by Lemma 2.11. Thus we consider $\sigma \in [0.6, 1 - 10^{-500}]$. We also assume $R(F, (S_i), (\sigma_i), T) > x^{1-\sigma}(\log x)^{-2J}$. Recall $F(s) = \prod_{i=1}^{j} S_i(s)$ with $S_i(s) = \sum_{n \sim N_i} a_n^{(i)} n^{-s}$, $N_i = N^{\ell_i}$ and $x \ll N \ll x$. Combining factors of $F(s)$ which have $\ell_i < 0.0001$, we note that $\{1, \dots, j\}$ can be partitioned into sets $U_1, \dots, U_k$ with:

(I) $\ell_s^* = \sum_{i \in U_s} \ell_i < 0.0001$ holds for at most one element s of $\{1, \dots, k\}$,

(II) $\ell_s^* = \sum_{i \in U_s} \ell_i \geq 0.0002$ implies that set U_s is a singleton,

(III) $\sigma_s^* = (\sum_{i \in U_s} \ell_i \sigma_i)/(\sum_{i \in U_s} \ell_i)$ are such that $\sigma_1^* \geq \cdots \geq \sigma_k^*$.

Since $F(s)$ is assumed to be of Type A or B, we have $k \geq 2$. Note that $\sigma_1^* \geq \sigma$.

Case 1: Long factors. We first assume that $\ell_1^* \geq 0.249$ or ($\ell_1^* < 0.0001$ and $\ell_2^* \geq 0.2489$).

Since $\sigma_2^* \geq \sigma_s^*$ for $s \geq 2$, the latter option implies $\sigma_2^* \geq (\sigma - \ell_1^* \sigma_1^*)/(1 - \ell_1^*) \geq \sigma - 0.0002$. For both options we have some $S_i(s)$ with $\ell_i \geq 0.2489$ and $\sigma_i \geq \sigma - 0.0002$. Then $F(s)$ must be of Type A and $S_i(s)$ must have property (Z1) or (Z2) with respect to T, so that Lemma 2.10 applies and

$$R(F, (S_i), (\sigma_i), T) \leq x^{4\varepsilon_1} \tau^2 N_i^{(6-12\sigma_i)} \leq x^{2a + 0.2489(6 - 12\sigma + 0.0024) + 4\varepsilon_1}. \tag{2.131}$$

The RHS of (2.131) is bounded above by $x^{1-\sigma-\varepsilon_1}$ if $a \in [0.53, 0.57]$ and $\sigma \in [a + \nu + 0.027, 1 - 10^{-500}]$. On the other hand, if additionally $\ell_i > 0.27$, we have

$$R(F, (S_i), (\sigma_i), T) \leq x^{4\varepsilon_1} \tau^2 N_i^{(6-12\sigma_i)} \leq x^{2a + 0.27(6 - 12\sigma + 0.0024) + 4\varepsilon_1}. \tag{2.132}$$

and the RHS of (2.132) is bounded above by $x^{1-\sigma-\varepsilon_1}$ for all $\sigma \in [a + \nu - \varepsilon_1, 1 - 10^{-500}]$. Now the only case left to consider is $a \in [0.53, 0.57]$, $\sigma \in [a + \nu - \varepsilon_1, a + \nu + 0.027]$ and $\ell_i \in [0.248, 0.27]$. Here we take $M = N_i$ in condition (2.129). We require

$$R(F, (S_i), (\sigma_i), T) \leq \min\{\tau^{-1/4} N_i^{2\sigma_i - 1} x^{7/4 - 2\sigma + \nu/4 - 4\varepsilon_1}, N_i^{2\sigma_i - 2} x^{2 - 2\sigma - 4\varepsilon_1}\}. \tag{2.133}$$

But the RHS of (2.133) is greater than $x^{(2-2\sigma) - \ell_i(2 - 2\sigma + 0.0004) + \ell_i - (a + 1 - \nu)/4 - 5\varepsilon_1}$. On the other hand, by bound (2.46) of Lemma 2.8, with $I = \{i\}$ and $k_i = 2$, we have

$$\begin{aligned} R(F, (S_i), (\sigma_i), T) &\leq x^{2\varepsilon_1} \max\left\{N_i^{2(2-2\sigma_i)}, \tau N_i^{2(4-6\sigma_i)}\right\} \\ &\leq \max\left\{x^{2\ell_i(2-2\sigma+0.0004)+3\varepsilon_1}, x^{a+2\ell_i(4-6\sigma+0.0012)+3\varepsilon_1}\right\}. \end{aligned} \tag{2.134}$$

The RHS of (2.134) is less than $x^{(2-2\sigma)-\ell_i(2-2\sigma+\frac{1}{2500})+\ell_i-\frac{(a+1-\nu)}{4}-5\varepsilon_1}$ for $\ell_i \in [0.2489, 0.27]$ and $\sigma \in [a+\nu-\varepsilon_1, a+\nu+0.027]$. We indeed have (2.128) or (2.130) if $\ell_1^* \geq 0.249$ or ($\ell_1^* < 0.0001$ and $\ell_2^* \geq 0.2489$).

Case 2: Short factors. We now assume that $\ell_1^* \in [0.0001, 0.249]$ or ($\ell_1^* < 0.0001$ and $\ell_2^* < 0.2489$).

We can choose $U = U_1$ or $U = U_1 \cup U_2$ such that $S(s) = \prod_{i\in U} S_i(s)$ satisfies $\ell_U = \sum_{i\in U} \ell_i \in [0.0001, 0.249]$ and $\sigma_U = (\sum_{i\in U} \ell_i\sigma_i)/\ell_U \geq \sigma$. Take $m \in \mathbb{N}$ such that $2^m\ell_U \in [0.1245, 0.249]$. We now describe three different approaches to obtaining regions in which one of (2.128), (2.129) or (2.130) holds.

Approach 1: Use of (C1). Choose $r_1 \in \{2, 3\}$. By bound (2.46) of Lemma 2.8 with $I = U$ and $k_i = 2^m r_1$, we have $R(F, (S_i), (\sigma_i), T) \leq x^{1-\sigma-\varepsilon_1}$ whenever

$$\max\{r_1 2^m \ell_U(2-2\sigma_U), a + r_1 2^m \ell_U(4-6\sigma_U)\} + 2\varepsilon_1 \leq 1-\sigma-\varepsilon_1.$$

But we assumed $\sigma_U \geq \sigma$. Thus the inequality certainly holds when

$$\frac{-1+\sigma+a+3\varepsilon_1}{r_1(6\sigma-4)} \leq 2^m\ell_U \leq \frac{1-\sigma-3\varepsilon_1}{r_1(2-2\sigma)}.$$

Write $S = \{(\sigma, \ell) : \sigma \in [a+\nu-\varepsilon_1, 1-10^{-500}], \ell \in [0.1245, 0.249]\}$ and

$$S_r(a) = \left\{(\sigma, \ell) \in S : \frac{-1+\sigma+a+3\varepsilon_1}{r(6\sigma-4)} \leq \ell \leq \frac{1-\sigma-3\varepsilon_1}{r(2-2\sigma)}\right\}.$$

If $(\sigma, 2^m\ell_U) \in S_r(a)$ for some $r \in \{2, 3\}$, then (2.128) holds.

Approach 2: Use of (C4). Note $R \leq \min\{\tau^{-\frac{1}{4}} M^{2\beta-1} x^{\frac{7}{4}-2\sigma+\frac{\nu}{4}-4\varepsilon_1}, M^{2\beta-2}x^{2-2\sigma-4\varepsilon_1}\}$ with $M = \prod_{i\in U} N_i^{2^m r_2}$ if $R \leq x^{-5\varepsilon_1}\min\{x^{r_2 2^m\ell_U(2\sigma-1)+\frac{7}{4}-\frac{a}{4}-2\sigma+\frac{\nu}{4}}, x^{r_2 2^m \ell_U(2\sigma-2)+2-2\sigma}\}$. Applying once more Lemma 2.8 with $I = U$ and $k_i = 2^m r_1$, this condition is certainly satisfied whenever

$$\ell(r_1(2-2\sigma) + r_2(1-2\sigma)) \leq 7/4 - a/4 - 2\sigma + \nu/4 - 7\varepsilon_1, \tag{2.135}$$

$$\ell(r_1(4-6\sigma) + r_2(1-2\sigma)) \leq 7/4 - 5a/4 - 2\sigma + \nu/4 - 7\varepsilon_1, \tag{2.136}$$

$$\ell(r_1+r_2)(2-2\sigma) \leq 2-2\sigma-7\varepsilon_1, \tag{2.137}$$

$$\ell(r_1(4-6\sigma) + r_2(2-2\sigma)) \leq 2-2\sigma-a-7\varepsilon_1 \tag{2.138}$$

all hold for $\ell = 2^m\ell_U$. Denote by $S^Q_{r_1,r_2}(a)$ the region of S in which inequalities (2.135), (2.136), (2.137) and (2.138) are all satisfied. If $(\sigma, 2^m\ell_U) \in S_{r_1,r_2}(a)$, then (2.129) holds.

We compute that the union $S_2(a) \cup S_3(a) \cup S_{3,2}^Q(a) \cup S_{4,3}^Q(a)$ covers $S \cap \{(\sigma, \ell) : \ell < 0.2 - C^*\varepsilon_1\}$. Here C^* is a large absolute constant, not dependent on any variables. Hence we only need to consider $\ell \geq 0.2 - C^*\varepsilon_1$. For that case, we also make use of condition (2.130):

Approach 3: Use of (C3). Let $r_2 \in \mathbb{N}$ and $M = \prod_{i \in U} N_i^{2^m r_2}$. (2.130) holds if

$$R(F, (S_i), (\sigma_i), T) \leq x^{r_2 2^m \ell_U (V\sigma - W) + aU + X - Y\sigma + Z\nu}$$

for every $(U, V, W, X, Y, Z) \in \mathcal{Z}$. Applying Lemma 2.8 with $I = U$ and $k_i = 2^m r_1$, this is true if

$$\begin{aligned} &\ell(r_1(2 - 2\sigma) + r_2(W - V\sigma)) \leq aU + X - Y\sigma + Z\nu - 2\varepsilon_1 \quad \text{and} \\ &\ell(r_1(4 - 6\sigma) + r_2(W - V\sigma)) \leq a(U - 1) + X - Y\sigma + Z\nu - 2\varepsilon_1 \end{aligned} \tag{2.139}$$

for $\ell = 2^m \ell_U$ and every $(U, V, W, X, Y, Z) \in \mathcal{Z}$. Denote by $S_{r_1, r_2}^*(a)$ the intersection over $\mathcal{Z}$ of the regions of S cut out by (2.139).

We now distinguish between option (1) and (2) from the statement of the lemma.

Case 2A: Treatment of option (1). Suppose $a \in [0.545, 0.57]$. Careful computations tell us that

$$S \cap \{(\ell, \sigma) : \ell \geq 0.2 - C^*\varepsilon_1\} \subseteq S_2(a) \cup S_{2,2}^*(a) \cup S_{3,2}^*(a).$$

Hence we have one (2.128), (2.129) or (2.130) for any choice of $2^m \ell_U \in [0.1245, 0.249]$ and this concludes the proof for option (1).

Case 2B: Treatment of option (2). Suppose $a < 0.545$ and suppose there exists $I \subseteq \{1, \ldots, j\}$ with $\sum_{i \in I} \ell_i \in [0.427 - \varepsilon_1, 0.474 + \varepsilon_1]$. Now $S_2(a) \cup S_{2,2}^*(a) \cup S_{3,2}^*(a) \cup S_{2,1}^Q(a)$ only covers the region

$$\begin{aligned} &(S \cap \{(\ell, \sigma) : \ell \geq 0.2 - C^*\varepsilon_1\}) \setminus S^\circ \\ &\text{where } S^\circ = \{(\sigma, \ell) : \sigma \in [a + \nu - \varepsilon_1, a + \nu + 0.025], \ell \in [0.2 - C^*\varepsilon_1, 0.215]\}. \end{aligned}$$

We consider the case $2^m \ell_U \in [0.2 - C^*\varepsilon_1, 0.215]$ separately when $\sigma \leq a + \nu + 0.025$.

Suppose first that $m \geq 1$. Note $2^{m-1}\ell_U \in [0.1 - C^*\varepsilon_1/2, 0.1075]$. We take $M = \prod_{i \in U} N_i^{2^{m-1} r_2}$ in (2.129) and note that (2.135), (2.136), (2.137) and (2.138) are satisfied if $r_1 = 5$, $r_2 = 4$, $\ell \in [0.1 - C^*\varepsilon_1/2, 0.1075]$, $a \in [0.53, 0.545]$ and $\sigma \in [a + \nu - \varepsilon_1, a + \nu + 0.025]$. This gives (2.129) for $m \geq 1$ and thus we now assume $m = 0$, so that $\ell_U \in [0.2 - C^*\varepsilon_1, 0.215]$. Either $U = U_1$ and U_1 is a singleton with $\ell_1^* \in [0.2 - C^*\varepsilon_1, 0.215]$ or $U = U_1 \cup U_2$ and $\ell_2^* \in [0.2 - 0.0001 - C^*\varepsilon_1, 0.215]$, $\sigma_2^* \geq \sigma - 0.0002$ and U_2 is a singleton. Thus we may assume that there exists $S_k(s)$ with $\ell_k \in [0.2 - 0.0001 - C^*\varepsilon_1, 0.215]$ and $\sigma_k \geq \sigma - 0.0002$.

Recall that there exists I with $\sum_{i \in I} \ell_i \in [0.427 - \varepsilon_1, 0.474 + \varepsilon_1]$. Suppose first that $k \notin I$. Then $\ell_k + \sum_{i \in I} \ell_i \in [0.627 - 0.0001 - 2C^*\varepsilon_1, 0.7 + \varepsilon_1]$ and

$$\sum_{i \notin I \cup \{k\}} \ell_i \in [0.3 - \varepsilon_1, 0.375 + \varepsilon_1].$$

But for $a \in [0.53, 0.545]$, this sum then satisfies option (3) of Lemma 2.21. Hence by Lemma 2.21 we have (2.128) or (2.129) for $\sigma \geq a + \nu - \varepsilon_1$. So suppose instead that $k \in I$. Assume first that $\sum_{i \in I} \ell_i \sigma_i < \sigma$. Then we replace the partition of $\{1, \ldots, j\}$ into $U_1, \ldots, U_k$, carried out at the start of our proof, by a partition of $\{1, \ldots, j\} \setminus I$ with the same properties (I), (II) and (III). Since $\sum_{i \notin I} \ell_i \sigma_i \geq \sigma$, all arguments still work the same way and give one of (2.128), (2.129) or (2.130) unless there is some $k_2 \in \{1, \ldots, j\} \setminus I$ with $\ell_{k_2} \in [0.2 - 0.0001 - C^*\varepsilon_1, 0.215]$. But then $\ell_{k_2} + \sum_{i \in I} \ell_i \in [0.627 - 0.0001 - 2C^*\varepsilon_1, 0.7 + \varepsilon_1]$ and we are again done by Lemma 2.21. Hence we may finally assume that $\sum_{i \in I} \ell_i \sigma_i \geq \sigma$. Taking $M = \prod_{i \in I} N_i$ in (2.130) and noting that conditions (2.139) are satisfied for each $(U, V, W, X, Y, Z) \in \mathcal{Z}$, $a \in [0.53, 0.545]$, $\sigma \in [a + \nu - \varepsilon_1, a + \nu + 0.025]$ and $\ell \in [0.427 - \varepsilon_1, 0.474 + \varepsilon_1]$ when $r_1 = r_2 = 1$, we conclude the proof. $\square$

2.4.4.5 Midsized τ, small σ

Now we move on to midsized τ and small σ. Here condition (C3*), which was derived using $R^*(F, (S_i), (\sigma_i), T)$, is used extensively. However, (C4*) also makes another appearance, together with all lemmas of Section 2.4.1 – this is the most complicated computation.

Lemma 2.23. *Let $F(s) = \prod_{i=1}^{j} S_i(s)$ be as described at the start of Section 2.4. Recall $S_i(s) = \sum_{n \sim N_i} a_n^{(i)} n^{-s}$ and $N_i = N^{\ell_i}$. Let $T \in [T_1, T_0]$ and suppose that $F(s)$ is of Type A or B at T. Write $\tau = x^a$.*
Finally, suppose there exists $I \subseteq \{1, \ldots, j\}$ such that I and a satisfy one of the following conditions:

(1) $\sum_{i \in I} \ell_i \in [0.405 - \varepsilon_1, 0.485 + \varepsilon_1] \cup [0.515 - \varepsilon_1, 0.595 + \varepsilon_1]$ *and* $a \in [0.53, 0.545]$,

(2) $\sum_{i \in I} \ell_i \in [0.400 - \varepsilon_1, 0.475 + \varepsilon_1] \cup [0.525 - \varepsilon_1, 0.600 + \varepsilon_1]$ *and* $a \in [0.545, 0.57]$,

(3) $\sum_{i \in I} \ell_i \in [0.380 - \varepsilon_1, 0.455 + \varepsilon_1] \cup [0.545 - \varepsilon_1, 0.620 + \varepsilon_1]$ *and* $a \in [0.57, 0.59]$,

(4) $\sum_{i\in I}\ell_i \in [0.365-\varepsilon_1, 0.435+\varepsilon_1]\cup[0.565-\varepsilon_1, 0.635+\varepsilon_1]$ *and* $a\in[0.59, 0.61]$,

(5) $\sum_{i\in I}\ell_i \in [0.355-\varepsilon_1, 0.420+\varepsilon_1]\cup[0.580-\varepsilon_1, 0.645+\varepsilon_1]$ *and* $a\in[0.61, 0.685]$.

Write $M=\prod_{i\in I}N_i$ *and* $M^\beta=\prod_{i\in I}N_i^{\sigma_i}$. *Let* $\mathcal{Z}$ *be as given in* (2.88). *Then if* $x\ge C$ *and* (σ_i) *has corresponding* $\sigma\le a+\nu-\varepsilon_1$, *one of the following four inequalities holds:*

$$R(F,(S_i),(\sigma_i),T)\le x^{1-\sigma}(\log x)^{-2J}, \tag{2.140}$$

$$R(F,(S_i),(\sigma_i),T)\le \tau x^{1-2\sigma+\nu+2\varepsilon_1}, \tag{2.141}$$

$$R(F,(S_i),(\sigma_i),T)\le \min\{\tau^{-1/4}M^{2\beta-1}x^{7/4-2\sigma+\nu/4-4\varepsilon_1}, M^{2\beta-2}x^{2-2\sigma-4\varepsilon_1}\}, \tag{2.142}$$

$$R(F,(S_i),(\sigma_i),T)\le \min_{(U,V,W,X,Y,Z)\in\mathcal{Z}}\tau^U M^{V\beta-W}x^{X-Y\sigma+Z\nu+\varepsilon_1}. \tag{2.143}$$

Here C *is a large constant dependent only on* J *and* ε_1, *and* ε_1 *is assumed to be small.*

Lemma 2.23 combines with Lemma 2.22 to complete the proof of Proposition 2.2 when $a\in(0.545, 0.57]$. It also combines with Lemma 2.21 and Lemma 2.22 to complete the proof of Proposition 2.2 when $a\in(0.53, 0.545]$. For $a\in(0.57, 0.685]$, it covers the case $\sigma\le a+\nu-\varepsilon_1$, explaining our choice of $\chi_2(a)$, while $\sigma>a+\nu-\varepsilon_1$ is treated in a later lemma.

Proof. We assume $R(F,(S_i),(\sigma_i),T)>x^{1-\sigma}(\log x)^{-2J}$. Suppose $(U,V,W,X,Y,Z)\in\mathcal{Z}$ is such that

$$\tau^U M^{V\beta-W}x^{X-Y\sigma+Z\nu+\varepsilon_1}=\min_{(U_1,V_1,W_1,X_1,Y_1,Z_1)\in\mathcal{Z}}\tau^{U_1}M^{V_1\beta-W_1}x^{X_1-Y_1\sigma+Z_1\nu+\varepsilon_1}. \tag{2.144}$$

Write $\ell_I=\sum_{i\in I}\ell_i$ and $(\sum_{i\notin I}\ell_i\sigma_i)/(1-\ell_I)=\alpha$, so that $N^\sigma=(N/M)^\alpha M^\beta=N^{(1-\ell_I)\alpha}N^{\ell_I\beta}$. Since $x\ll N\ll x$ and $M=N^{\ell_I}$, then $\tau^U M^{V\beta-W}x^{X-Y\sigma+Z\nu+\varepsilon_1}\ge x^{(1-\ell_I)(W-V\alpha)+aU+(X-W)-(Y-V)\sigma+Z\nu}$. By Lemma 2.8 with $\{1,\dots,j\}\setminus I$ in the place of I and $k_i=1$, we have

$$R(F,(S_i),(\sigma_i),T)\le x^{2\varepsilon_1}\max\left\{x^{(1-\ell_I)(2-2\alpha)},\min\left\{\tau x^{(1-\ell_I)(1-2\alpha)},\tau x^{(1-\ell_I)(4-6\alpha)}\right\}\right\}.$$

Hence if $\max\{(1-\ell_I)(2-2\alpha), a+(1-\ell_I)(\gamma-\delta\alpha)\}+2\varepsilon_1\le(1-\ell_I)(W-V\alpha)+aU+(X-W)-(Y-V)\sigma+Z\nu$ holds for some $(\gamma,\delta)\in\{(1,2),(4,6)\}$, then (2.143) is satisfied. Rearranging, we need the following:

$$(V-2)(1-\ell_I)\alpha\le(1-\ell_I)(W-2)+aU+(X-W)-(Y-V)\sigma+Z\nu-2\varepsilon_1, \tag{2.145}$$

$$(V-\delta)(1-\ell_I)\alpha\le(1-\ell_I)(W-\gamma)+a(U-1)+(X-W)-(Y-V)\sigma+Z\nu-2\varepsilon_1. \tag{2.146}$$

We will now derive many conditions which ensure that (2.145) and (2.146) hold.

Step 1: Treatment of (2.146) with $\delta = 6$.

Since $(U, V, W, X, Y, Z) \in \mathcal{Z}$, we have $V < 6$. Hence for $\delta = 6$, (2.146) holds whenever

$$(1 - \ell_I)\alpha \geq \frac{a(U-1) + (X-4) - (Y-V)\sigma + Z\nu - 2\varepsilon_1}{(V-6)} - \frac{(W-4)}{(V-6)}\ell_I. \tag{2.147}$$

On the other hand, taking $B(a,\sigma) = a + 1 - 2\sigma + \nu + 2\varepsilon_1$, Corollary 2.13 tells us that (2.141) holds if

$$\ell_I\beta \geq \ell_I \min\left\{\left(\frac{3a + B(a,\sigma) - 2\varepsilon_1}{3a + 3B(a,\sigma) - 6\varepsilon_1}\right), \left(\frac{3a - 2B(a,\sigma) + 4\varepsilon_1}{3a - B(a,\sigma) + 2\varepsilon_1}\right)\right\} = \ell_I E(a,\sigma). \tag{2.148}$$

Adding (2.147) and (2.148), we have (2.141) or ((2.146) with $\delta = 6$) if

$$\sigma \geq \frac{a(U-1) + (X-4) - (Y-V)\sigma + Z\nu - 2\varepsilon_1}{(V-6)} - \left(\left(\frac{W-4}{V-6}\right) - E(a,\sigma)\right)\ell_I. \tag{2.149}$$

But $\frac{(W-4)}{(V-6)} - E(a,\sigma) \leq \frac{14}{19} - \frac{(4a+1-3/2+\nu)}{(6a+3-9/2+3\nu)} < 0$ whenever $(U, V, W, X, Y, Z) \in \mathcal{Z}$, $a \in [0.53, 0.68]$ and $\sigma \in [0.75, 1]$. Rearranging, (2.149) becomes

$$\sigma \geq 0.75 \quad \text{and} \quad \ell_I \leq \Upsilon_6(a, \sigma; (U, V, W, X, Y, Z)) \tag{2.150}$$

$$\text{where } \Upsilon_6(a, \sigma; (U, V, W, X, Y, Z)) = \frac{a(U-1) + (X-4) - (Y-6)\sigma + Z\nu - 2\varepsilon_1}{(W-4) - (V-6)E(a,\sigma)}.$$

Step 2: Treatment of (2.145) and ((2.146) with $\delta = 2$).

We now split into three cases, depending on whether $V = 2$, $V > 2$ or $V < 2$.

Case 1: $V = 2$. Here (U, V, W, X, Y, Z) is $(\frac{1}{2}, 2, \frac{3}{2}, \frac{3}{2}, 2, \frac{1}{2})$ or $(\frac{3}{8}, 2, \frac{3}{2}, \frac{15}{8}, \frac{5}{2}, \frac{5}{8})$.

Then (2.145) holds if $\ell_I(W-2) \leq aU + (X-2) - (Y-V)\sigma + Z\nu - 2\varepsilon_1$. But $W = 3/2$ and so we need

$$\ell_I \geq -2aU - 2(X-2) + 2(Y-V)\sigma - 2Z\nu + 4\varepsilon_1.$$

Further, (2.146) with $\delta = 2$ holds if $(W-1)\ell_I \leq a(U-1) + (X-1) - (Y-V)\sigma + Z\nu - 2\varepsilon_1$ and so if

$$\ell_I \leq 2a(U-1) + 2(X-1) - 2(Y-V)\sigma + 2Z\nu - 4\varepsilon_1.$$

Write $\mathcal{Z}_1 = \{(1/2, 2, 3/2, 3/2, 2, 1/2), (3/8, 2, 3/2, 15/8, 5/2, 5/8)\}$ and define

$$\Lambda_*(a, \sigma; (U, V, W, X, Y, Z)) = -2aU - 2(X-2) + 2(Y-V)\sigma - 2Z\nu + 4\varepsilon_1,$$

$$\Upsilon_2(a, \sigma; (U, V, W, X, Y, Z)) = 2a(U-1) + 2(X-1) - 2(Y-V)\sigma + 2Z\nu - 4\varepsilon_1,$$

for $(U,V,W,X,Y,Z) \in \mathscr{Z}_1$.

Case 2: $V > 2$. Here (U,V,W,X,Y,Z) is $(1,4,4,3,4,1)$ or $(\frac{2}{5},\frac{16}{5},\frac{12}{5},\frac{12}{5},\frac{16}{5},\frac{4}{5})$.

Then (2.145) and ((2.146) with $\delta = 2$) hold if the following inequalities are satisfied, respectively:

$$(1-\ell_I)\alpha \le \frac{aU + (X-2) - (Y-V)\sigma + Z\nu - 2\varepsilon_1}{(V-2)} - \frac{(W-2)}{(V-2)}\ell_I, \tag{2.151}$$

$$(1-\ell_I)\alpha \le \frac{a(U-1) + (X-1) - (Y-V)\sigma + Z\nu - 2\varepsilon_1}{(V-2)} - \frac{(W-1)}{(V-2)}\ell_I. \tag{2.152}$$

On the other hand, by Corollary 2.13 with $\{1,\dots,j\}\setminus I$ in the place of I, $R \le x^{B(a,\sigma)}$ if

$$(1-\ell_I)\alpha \ge (1-\ell_I)\min\left\{\left(\frac{3a+B(a,\sigma)-2\varepsilon_1}{3a+3B(a,\sigma)-6\varepsilon_1}\right),\left(\frac{3a-2B(a,\sigma)+4\varepsilon_1}{3a-B(a,\sigma)+2\varepsilon_1}\right)\right\}. \tag{2.153}$$

Let $E(a,\sigma) = \min\left\{\left(\frac{3a+B(a,\sigma)-2\varepsilon_1}{3a+3B(a,\sigma)-6\varepsilon_1}\right),\left(\frac{3a-2B(a,\sigma)+4\varepsilon_1}{3a-B(a,\sigma)+2\varepsilon_1}\right)\right\}$. (2.151) and (2.153) combined then give that one of (2.140), (2.141) or (2.145) holds whenever

$$\frac{aU + (X-2) - (Y-V)\sigma + Z\nu - 2\varepsilon_1}{(V-2)} - E(a,\sigma) \ge \left(\frac{(W-2)}{(V-2)} - E(a,\sigma)\right)\ell_I. \tag{2.154}$$

If $(U,V,W,X,Y,Z) = (1,4,4,3,4,1)$, then $(W-2)/(V-2) - E(a,\sigma) = 1 - E(a,\sigma)$ is positive for all $a \in [0.53, 0.77]$ and $\sigma \in [0.6, 1-10^{-500}]$. If $(U,V,W,X,Y,Z) = (2/5, 16/5, 12/5, 12/5, 16/5, 4/5)$, then $(W-2)/(V-2) - E(a,\sigma) = 1/3 - E(a,\sigma)$ is instead negative for all $a \in [0.53, 0.77]$ and $\sigma \in [0.6, 1-10^{-500}]$. Thus to satisfy (2.154), we require

$$\frac{2a + 2 + 4\nu - 6E(a,\sigma) - 10\varepsilon_1}{6\,(1/3 - E(a,\sigma))} \le \ell_I \le \frac{a + 1 + \nu - 2E(a,\sigma) - 2\varepsilon_1}{2\,(1 - E(a,\sigma))}.$$

We proceed similarly for (2.146): (2.152) and (2.153) combined give that one of (2.140), (2.141) or ((2.146) with $\delta = 2$) holds whenever

$$\frac{a(U-1) + (X-1) - (Y-V)\sigma + Z\nu - 2\varepsilon_1}{(V-2)} - E(a,\sigma) \ge \left(\frac{(W-1)}{(V-2)} - E(a,\sigma)\right)\ell_I.$$

But $(W-1)/(V-2)$ is either $3/2$ or $7/6$ and $(W-2)/(V-2) - E(a,\sigma)$ is positive for all $a \in [0.53, 0.77]$ and $\sigma \in [0.6, 1-10^{-500}]$ in both cases. So we require

$$\ell_I \le \frac{a(U-1) + (X-1) + Z\nu - (V-2)E(a,\sigma) - 2\varepsilon_1}{(W-1) - (V-2)E(a,\sigma)}.$$

Next, a second approach to lower bounds on $(1-\ell_I)\alpha$: So far, we have made use of Lemma 2.12, but if we instead use Lemma 2.8, we have $R(F,(S_i),(\sigma_i),T) \le x^{B(a,\sigma)}$ if

$$\max\{(1-\ell_I)(2-2\alpha), a+\min\{(1-\ell_I)(1-2\alpha),(1-\ell_I)(4-6\alpha)\}\}+2\varepsilon_1 \le B(a,\sigma). \tag{2.155}$$

We have $\xi a+(1-\ell_I)(\lambda-\mu\alpha)+2\varepsilon_1 \le B(a,\sigma)$ if and only if $(1-\ell_I)\alpha \ge (1-\ell_I)(\frac{\lambda}{\mu}) - \frac{B(a,\sigma)}{\mu} + \frac{\xi a}{\mu} + \frac{2\varepsilon_1}{\mu}$. Combining this inequality with (2.151) and (2.152), we need

$$\left(\frac{(W-2)}{(V-2)} - \left(\frac{\lambda}{\mu}\right)\right)\ell_I \le \frac{aU+(X-2)+Z\nu-2\varepsilon_1}{(V-2)} + \frac{B(a,\sigma)}{\mu} - \frac{\xi a}{\mu} - \left(\frac{\lambda}{\mu}\right) - \frac{2\varepsilon_1}{\mu}, \tag{2.156}$$

$$\left(\frac{(W-1)}{(V-2)} - \left(\frac{\lambda}{\mu}\right)\right)\ell_I \le \frac{a(U-1)+(X-1)+Z\nu-2\varepsilon_1}{(V-2)} + \frac{B(a,\sigma)}{\mu} - \frac{\xi a}{\mu} - \left(\frac{\lambda}{\mu}\right) - \frac{2\varepsilon_1}{\mu}. \tag{2.157}$$

for both $((\xi,\lambda,\mu)=(0,2,2)$ and $(\xi,\lambda,\mu)=(1,1,2))$ or for both $((\xi,\lambda,\mu)=(0,2,2)$ and $(\xi,\lambda,\mu)=(1,4,6))$. Then one of ((2.140), (2.141) or (2.145)) and one of ((2.140), (2.141) or (2.146)) is satisfied.

But $\frac{W-2}{V-2}$ is either 1 or $\frac{1}{3}$, while $\frac{W-1}{V-2}$ is either $\frac{3}{2}$ or $\frac{7}{6}$ and $\frac{\lambda}{\mu}$ is 1, $\frac{1}{2}$ or $\frac{2}{3}$. We first consider the case $\frac{W-2}{V-2}=1$, where $(U,V,W,X,Y,Z)=(1,4,4,3,4,1)$.

If $(\xi,\lambda,\mu)=(0,2,2)$, then $\frac{W-2}{V-2}-\frac{\lambda}{\mu}=0$. Since $B(a,\sigma)=a+1-2\sigma+\nu+2\varepsilon_1$, (2.156) becomes $\sigma \le a+\nu-\varepsilon_1$.

If $(\xi,\lambda,\mu)\in\{(1,1,2),(1,4,6)\}$, then $\frac{W-2}{V-2}-\frac{\lambda}{\mu}$ is positive and (2.156) becomes

$$\ell_I \le \left(1-\frac{\lambda}{\mu}\right)^{-1}\left(\frac{a+1+\nu-2\varepsilon_1}{2} + \frac{B(a,\sigma)}{\mu} - \frac{a}{\mu} - \frac{\lambda}{\mu} - \frac{2\varepsilon_1}{\mu}\right),$$

For the case $(W-2)/(V-2)=1/3$, we note that $\frac{(W-2)}{(V-2)}-\frac{\lambda}{\mu}$ is certainly negative. We have $(U,V,W,X,Y,Z)=(2/5,16/5,12/5,12/5,16/5,4/5)$ and (2.156) becomes

$$\ell_I \ge \left(\frac{1}{3}-\frac{\lambda}{\mu}\right)^{-1}\left(\frac{2a+2+4\nu-10\varepsilon_1}{6} + \frac{B(a,\sigma)}{\mu} - \frac{\xi a}{\mu} - \frac{\lambda}{\mu} - \frac{2\varepsilon_1}{\mu}\right).$$

On the other hand, $\frac{(W-1)}{(V-2)}-\frac{\lambda}{\mu}$ is certainly positive and (2.157) rearranges to

$$\ell_I \le \left(\frac{(W-1)}{(V-2)}-\frac{\lambda}{\mu}\right)^{-1}\left(\frac{a(U-1)+(X-1)+Z\nu-2\varepsilon_1}{(V-2)} + \frac{B(a,\sigma)}{\mu} - \frac{\xi a}{\mu} - \frac{\lambda}{\mu} - \frac{2\varepsilon_1}{\mu}\right).$$

Finally, for yet another treatment of (2.146), recall from Step 1 that we have (2.146) with $\delta=6$ if

$$(1-\ell_I)\alpha \ge \frac{a(U-1)+(X-4)-(Y-V)\sigma+Z\nu-2\varepsilon_1}{(V-6)} - \frac{(W-4)}{(V-6)}\ell_I. \tag{2.158}$$

Combining (2.158) and (2.152), we have (2.146) with $\delta = 2$ or $\delta = 6$ if

$$\ell_I \leq \left(\frac{(W-4)}{(V-6)} - \frac{(W-1)}{(V-2)}\right)^{-1} \left(\frac{4a(U-1) + 4X - 3V + 2 + 4Z\nu - 8\varepsilon_1}{(V-2)(V-6)}\right).$$

We summarize these conditions as follows:

For $(U, V, W, X, Y, Z) = (1, 4, 4, 3, 4, 1)$, we have one of (2.140), (2.141) or (2.145) if $\sigma \leq a + \nu - \varepsilon_1$ and

$$\ell_I \leq \max\left\{\left(\frac{a + 1 + \nu - 2E(a,\sigma) - 2\varepsilon_1}{2\,(1 - E(a,\sigma))}\right), \left(\max_{(\lambda,\mu)\in\{(1,2),(4,6)\}} \left(1 - \frac{\lambda}{\mu}\right)^{-1} \left(\frac{a+1+\nu-2\varepsilon_1}{2} + \frac{B(a,\sigma)}{\mu} - \frac{a}{\mu} - \frac{\lambda}{\mu} - \frac{2\varepsilon_1}{\mu}\right)\right)\right\}. \tag{2.159}$$

Denote the RHS of (2.159) by $\Upsilon_*(a, \sigma; (1, 4, 4, 3, 4, 1))$. Write $\mathscr{Z}_{2,1} = \{(1, 4, 4, 3, 4, 1)\}$.

Write $\mathscr{Z}_{2,2} = \{(2/5, 16/5, 12/5, 12/5, 16/5, 4/5)\}$. For $(U, V, W, X, Y, Z) \in \mathscr{Z}_{2,2}$, we have one of (2.140), (2.141) or (2.145) if ℓ_I is greater than or equal to

$$\min\left\{\left(\frac{2a + 2 + 4\nu - 6E(a,\sigma) - 10\varepsilon_1}{6\,(1/3 - E(a,\sigma))}\right), \max\left\{\left(-\frac{2}{3}\right)^{-1} \left(\frac{2a + 2 + 4\nu - 10\varepsilon_1}{6} + \frac{B(a,\sigma)}{2} - 1 - \varepsilon_1\right), \min_{(\lambda,\mu)\in\{(1,2),(4,6)\}} \left(\frac{1}{3} - \frac{\lambda}{\mu}\right)^{-1}\left(\frac{2a+2+4\nu-10\varepsilon_1}{6} + \frac{B(a,\sigma)}{\mu} - \frac{a}{\mu} - \frac{\lambda}{\mu} - \frac{2\varepsilon_1}{\mu}\right)\right\}\right\}. \tag{2.160}$$

Denote the RHS of (2.160) by $\Lambda_*(a, \sigma; (2/5, 16/5, 12/5, 12/5, 16/5, 4/5))$.

For any $(U, V, W, X, Y, Z) \in \mathscr{Z}_2 = \mathscr{Z}_{2,1} \cup \mathscr{Z}_{2,2}$, we have one of (2.140), (2.141) or ((2.146) with $\delta = 2$ or $\delta = 6$) if ℓ_I is less than or equal to

$$\max\left\{\left(\frac{a(U-1) + (X-1) + Z\nu - (V-2)E(a,\sigma) - 2\varepsilon_1}{(W-1) - (V-2)E(a,\sigma)}\right), \left(\frac{(W-4)}{(V-6)} - \frac{(W-1)}{(V-2)}\right)^{-1} \left(\frac{4a(U-1) + 4X - 3V + 2 + 4Z\nu - 8\varepsilon_1}{(V-2)(V-6)}\right), \min\left\{\left(\frac{(W-1)}{(V-2)} - 1\right)^{-1}\left(\frac{a(U-1) + (X-1) + Z\nu - 2\varepsilon_1}{(V-2)} + \frac{B(a,\sigma)}{2} - 1 - \varepsilon_1\right), \left(\max_{\{(1,2),(4,6)\}} \frac{\frac{(a(U-1)+(X-1)+Z\nu-2\varepsilon_1)}{(V-2)} + (B(a,\sigma) - a - \lambda - 2\varepsilon_1)/\mu}{(W-1)/(V-2) - \lambda/\mu}\right)\right\}\right\}. \tag{2.161}$$

Denote the RHS of (2.161) by $\Upsilon_2(a,\sigma;(U,V,W,X,Y,Z))$ when $(U,V,W,X,Y,Z) \in \mathcal{Z}_2$.

Case 3: $V < 2$. This case covers the remaining elements of $\mathcal{Z}$, namely $(\frac{2}{5},\frac{4}{5},\frac{3}{5},\frac{6}{5},\frac{8}{5},\frac{2}{5})$, $(\frac{1}{3},\frac{2}{3},\frac{1}{3},1,\frac{4}{3},\frac{1}{3})$, $(\frac{2}{7},\frac{16}{21},\frac{8}{21},\frac{8}{7},\frac{32}{21},\frac{8}{21})$, $(\frac{1}{4},\frac{4}{3},\frac{2}{3},\frac{5}{4},\frac{5}{3},\frac{5}{12})$ and also $(\frac{1}{9},\frac{16}{9},\frac{8}{9},\frac{5}{3},\frac{20}{9},\frac{5}{9})$. Write $\mathcal{Z}_3 = \mathcal{Z} \setminus (\mathcal{Z}_1 \cup \mathcal{Z}_2)$.

Then (2.145) and ((2.146) with $\delta = 2$) hold if the following inequalities are satisfied, respectively:

$$(1-\ell_I)\alpha \geq \frac{aU + (X-2) - (Y-V)\sigma + Z\nu - 2\varepsilon_1}{(V-2)} - \frac{(W-2)}{(V-2)}\ell_I, \tag{2.162}$$

$$(1-\ell_I)\alpha \geq \frac{a(U-1) + (X-1) - (Y-V)\sigma + Z\nu - 2\varepsilon_1}{(V-2)} - \frac{(W-1)}{(V-2)}\ell_I. \tag{2.163}$$

On the other hand, Corollary 2.13 tells us that $R(F,(S_i),(\sigma_i),T) \leq x^{B(a,\sigma)}$ whenever

$$\ell_I\beta \geq \ell_I \min\left\{\left(\frac{3a + B(a,\sigma) - 2\varepsilon_1}{3a + 3B(a,\sigma) - 6\varepsilon_1}\right), \left(\frac{3a - 2B(a,\sigma) + 4\varepsilon_1}{3a - B(a,\sigma) + 2\varepsilon_1}\right)\right\} = \ell_I E(a,\sigma). \tag{2.164}$$

Summing up (2.162) and (2.164), we have one of (2.140), (2.141) or (2.145) if

$$\sigma \geq \frac{aU + (X-2) - (Y-V)\sigma + Z\nu - 2\varepsilon_1}{(V-2)} + \left(E(a,\sigma) - \frac{(W-2)}{(V-2)}\right)\ell_I.$$

But $E(a,\sigma) - (W-2)/(V-2) \leq 1 - 7/6 < 0$ and so this rearranges to

$$\ell_I \geq \left(E(a,\sigma) - \frac{(W-2)}{(V-2)}\right)^{-1}\left(\sigma - \left(\frac{aU + (X-2) - (Y-V)\sigma + Z\nu - 2\varepsilon_1}{(V-2)}\right)\right). \tag{2.165}$$

Summing up (2.163) and (2.164), we certainly have one of (2.140), (2.141) or ((2.146) with $\delta = 2$) if

$$\sigma \geq \frac{a(U-1) + (X-1) - (Y-V)\sigma + Z\nu - 2\varepsilon_1}{(V-2)} + \left(E(a,\sigma) - \frac{(W-1)}{(V-2)}\right)\ell_I. \tag{2.166}$$

But $E(a,\sigma) - (W-1)/(V-2) \geq (3a - 2B + 4\varepsilon_1)/(3a - B + 2\varepsilon_1) - 1/2 > 0$ whenever $B(a,\sigma) < a + 2\varepsilon_1$. Of course, if $B(a,\sigma) \geq a + 2\varepsilon_1$, we automatically have $R(F,(S_i),(\sigma_i),T) \leq T_0 = x^{a+\varepsilon_1} < x^{B(a,\sigma)}$. So we may assume $B(a,\sigma) < a + 2\varepsilon_1$ and $E(a,\sigma) - (W-2)/(V-2) > 0$. Thus (2.166) rearranges to

$$\ell_I \leq \left(E(a,\sigma) - \frac{(W-1)}{(V-2)}\right)^{-1}\left(\sigma - \left(\frac{a(U-1) + (X-1) - (Y-V)\sigma + Z\nu - 2\varepsilon_1}{(V-2)}\right)\right). \tag{2.167}$$

For $a \geq 0.64$ and $\Delta = (U, V, W, X, Y, Z) \in \mathcal{Z}_3$, we now simply set

$$\Lambda_*(a,\sigma;\Delta) = \left(E(a,\sigma) - \frac{(W-2)}{(V-2)}\right)^{-1}\left(\sigma - \frac{aU + (X-2) - (Y-V)\sigma + Z\nu - 2\varepsilon_1}{(V-2)}\right),$$

$$\Upsilon_2(a,\sigma;\Delta) = \left(E(a,\sigma) - \frac{(W-1)}{(V-2)}\right)^{-1}\left(\sigma - \frac{a(U-1) + (X-1) - (Y-V)\sigma + Z\nu - 2\varepsilon_1}{(V-2)}\right).$$

For $a < 0.64$, we also mix in condition (2.142): It is satisfied if

$$R(F,(S_i),(\sigma_i),T) \leq x^{(2\beta-2)\ell_I} x^{2-2\sigma-5\varepsilon_1} \min\{x^{\ell_I-(1+a-\nu)/4}, 1\}. \tag{2.168}$$

If $a < 0.64$ then $\frac{(1+a-\nu)}{4} < 0.3525$ and $\ell_I \geq 0.355 - \varepsilon_1$ gives $\min\{x^{\ell_I-(1+a-\nu)/4}, 1\} = 1$. Applying Lemma 2.8, we have (2.168) whenever

$$\xi a + (\lambda - \mu\beta)\ell_I + 2\varepsilon_1 \leq (2\beta - 2)\ell_I + 2 - 2\sigma - 5\varepsilon_1$$

for both $((\xi,\lambda,\mu) = (0,2,2)$ and $(\xi,\lambda,\mu) = (1,1,2))$ or $((\xi,\lambda,\mu) = (0,2,2)$ and $(\xi,\lambda,\mu) = (1,4,6))$ or $((\xi,\lambda,\mu) = (0,2,2)$ and $(\xi,\lambda,\mu) = (1,11,14))$. The inequality rearranges to

$$\beta\ell_I \geq \frac{1}{\mu+2}\left(\xi a - (2-2\sigma) + (\lambda+2)\ell_I + 7\varepsilon_1\right). \tag{2.169}$$

Summing up (2.162) and (2.169) and rearranging, one of (2.142) or (2.145) holds if

$$\ell_I \geq \left(1 - \frac{W-2}{V-2}\right)^{-1}\left(-\frac{aU + (X-2) - (Y-V)\sigma + Z\nu - 2\varepsilon_1}{V-2} + \frac{2 + 2\sigma - 7\varepsilon_1}{4}\right) \quad \text{and} \tag{2.170}$$

$$\ell_I \geq \left(\frac{\lambda+2}{\mu+2} - \frac{W-2}{V-2}\right)^{-1}\left(-\frac{aU + (X-2) - (Y-V)\sigma + Z\nu - 2\varepsilon_1}{V-2} + \frac{2 + \mu\sigma - a - 7\varepsilon_1}{2+\mu}\right) \tag{2.171}$$

for some $(\lambda,\mu) \in \{(1,2),(4,6),(11,14)\}$. Here we used that $\frac{(\lambda+2)}{(\mu+2)} - \frac{(W-2)}{(V-2)} \leq 1 - \frac{7}{6} < 0$. For $a < 0.64$ and $(U,V,W,X,Y,Z) \in \mathcal{Z}_3$, set $\Lambda_*(a,\sigma;(U,V,W,X,Y,Z))$ equal to

$$\begin{aligned}
&\min\Bigg\{\left(E(a,\sigma) - \frac{(W-2)}{(V-2)}\right)^{-1}\left(\sigma - \left(\frac{aU + (X-2) - (Y-V)\sigma + Z\nu - 2\varepsilon_1}{(V-2)}\right)\right),\\
&\quad\max\Bigg\{\left(1 - \frac{W-2}{V-2}\right)^{-1}\left(-\frac{aU + (X-2) - (Y-V)\sigma + Z\nu - 2\varepsilon_1}{V-2} + \frac{2 + 2\sigma - 7\varepsilon_1}{4}\right),\\
&\qquad\min\Bigg\{\left(\frac{3}{4} - \frac{W-2}{V-2}\right)^{-1}\left(-\frac{aU + (X-2) - (Y-V)\sigma + Z\nu - 2\varepsilon_1}{V-2} + \frac{2 + 2\sigma - a - 7\varepsilon_1}{4}\right),\\
&\qquad\left(\frac{6}{8} - \frac{W-2}{V-2}\right)^{-1}\left(-\frac{aU + (X-2) - (Y-V)\sigma + Z\nu - 2\varepsilon_1}{V-2} + \frac{2 + 6\sigma - a - 7\varepsilon_1}{8}\right),\\
&\qquad\left(\frac{13}{16} - \frac{W-2}{V-2}\right)^{-1}\left(-\frac{aU + (X-2) - (Y-V)\sigma + Z\nu - 2\varepsilon_1}{V-2} + \frac{2 + 14\sigma - a - 7\varepsilon_1}{16}\right)\Bigg\}\Bigg\}\Bigg\}.
\end{aligned}$$

One of (2.140), (2.141), (2.142) or (2.145) holds if $\ell_I \geq \Lambda_*(a,\sigma;(U,V,W,X,Y,Z))$.

Instead summing up (2.163) and (2.169) and rearranging, we have one of (2.142) or (2.146) if both

$$\ell_I \leq \left(1-\frac{W-1}{V-2}\right)^{-1}\left(-\frac{a(U-1)+(X-1)-(Y-V)\sigma+Z\nu-2\varepsilon_1}{V-2}+\frac{2+2\sigma-7\varepsilon_1}{4}\right) \quad \text{and} \tag{2.172}$$

$$\ell_I \leq \left(\frac{\lambda+2}{\mu+2}-\frac{W-1}{V-2}\right)^{-1}\left(-\frac{a(U-1)+(X-1)-(Y-V)\sigma+Z\nu-2\varepsilon_1}{V-2}+\frac{2+\mu\sigma-a-7\varepsilon_1}{2+\mu}\right). \tag{2.173}$$

Let $\Upsilon_2(a,\sigma;(U,V,W,X,Y,Z)) = \max\{(2.167), \min\{(2.172), \max(2.173)\}\}$, where the terms (2.167), (2.172) and (2.173) are short-hand for the RHS of the corresponding inequalities and the innermost maximum is taken over $(\lambda,\mu) \in \{(1,2),(4,6),(11,14)\}$. One of (2.140), (2.141), (2.142) or (2.146) holds if $a < 0.64$ and $0.355 - \varepsilon_1 \leq \ell_I \leq \Upsilon_2(a,\sigma;(U,V,W,X,Y,Z))$.

Step 3: Treatment of large a.

We are now done with the discussion of (2.145) and (2.146), which combine to give (2.143). But for large a it is sometimes better to take a direct approach, working only towards conditions (2.140) and (2.141). Recall that we already know from (2.155) in Step 2, Case 2 and (2.164) in Step 2, Case 3 that $R(F,(S_i),(\sigma_i),T) \leq x^{B(a,\sigma)}$ whenever one of the following three inequalities is satisfied:

$$(1-\ell_I)\alpha \geq \max\left\{(1-\ell_I) - \frac{B(a,\sigma)}{2} + \varepsilon_1, (1-\ell_I)\left(\frac{1}{2}\right) - \frac{B(a,\sigma)}{2} + \frac{a}{2} + \varepsilon_1\right\},$$
$$(1-\ell_I)\alpha \geq \max\left\{(1-\ell_I) - \frac{B(a,\sigma)}{2} + \varepsilon_1, (1-\ell_I)\left(\frac{2}{3}\right) - \frac{B(a,\sigma)}{6} + \frac{a}{6} + \frac{2\varepsilon_1}{6}\right\},$$
$$\ell_I\beta \geq \ell_I \min\left\{\left(\frac{3a+B(a,\sigma)-2\varepsilon_1}{3a+3B(a,\sigma)-6\varepsilon_1}\right), \left(\frac{3a-2B(a,\sigma)+4\varepsilon_1}{3a-B(a,\sigma)+2\varepsilon_1}\right)\right\} = \ell_I E(a,\sigma).$$

Adding up and rearranging, we have (2.140) or (2.141) if $\sigma \geq 0.75$ and

$$\ell_I \geq \Lambda_E(a,\sigma) = \frac{1}{(1-E(a,\sigma))}\left(1 - \frac{B(a,\sigma)}{2} - \sigma + \varepsilon_1\right),$$
$$\ell_I \leq \Upsilon_{E,2}(a,\sigma) = \max\left\{\left(\frac{1-B(a,\sigma)+a-2\sigma+2\varepsilon_1}{1-2E(a,\sigma)}\right), \left(\frac{4-B(a,\sigma)+a-6\sigma+6\varepsilon_1}{4-6E(a,\sigma)}\right)\right\}.$$

(The assumption $\sigma \geq 0.75$ ensured that $\frac{2}{3} - E(a,\sigma) \leq \frac{2}{3} - E(a,0.75) \leq \frac{2}{3} - E(0.68,0.75) < 0$.) For $\sigma \geq 0.65$ we still have (2.140) or (2.141) if

$$\Lambda_E(a,\sigma) \leq \ell_I \leq \Upsilon_{E,1}(a,\sigma) = \left(\frac{1-B(a,\sigma)+a-2\sigma+2\varepsilon_1}{1-2E(a,\sigma)}\right).$$

Alternatively, we can also use option (4) of Lemma 2.21: It gives (2.140) or (2.142) if $\sigma \geq \min\{a+\nu-\varepsilon_1, 0.85\}$ and $\Lambda_Q(a,\sigma) \leq \ell_I \leq \Upsilon_Q(a,\sigma)$, where

$$\Upsilon_Q(a,\sigma) = 1 - \min\left\{\frac{3a}{6\sigma-2}, \max\left\{\frac{3a}{20\sigma-14}, \frac{2a}{4\sigma-1}\right\}\right\} - 10^{-100} \quad \text{and}$$
$$\Lambda_Q(a,\sigma) = 1 - \max\{m_1(a,\sigma), \min\{m_2(a,\sigma), m_3(a,\sigma)\}\} + 10^{-100} \quad \text{where}$$
$$m_1(a,\sigma) = \min\left\{\frac{a-3-\nu}{4(1-2\sigma)} + \frac{a(3-3\sigma)}{(3\sigma-1)(1-2\sigma)}, \frac{a-3-\nu}{4(1-2\sigma_1^\circ)} + \frac{a(3-3\sigma_1^\circ)}{(3\sigma_1^\circ-1)(1-2\sigma_1^\circ)}\right\},$$
$$m_2(a,\sigma) = \min\left\{\frac{a-3-\nu}{4(1-2\sigma)} + \frac{a(3-3\sigma)}{(10\sigma-7)(1-2\sigma)}, \frac{a-3-\nu}{4(1-2\sigma_2^\circ)} + \frac{a(3-3\sigma_2^\circ)}{(10\sigma_2^\circ-7)(1-2\sigma_2^\circ)}\right\},$$
$$m_3(a,\sigma) = \min\left\{\frac{a-3-\nu}{4(1-2\sigma)} + \frac{a(4-4\sigma)}{(4\sigma-1)(1-2\sigma)}, \frac{a-3-\nu}{4(1-2\sigma_2^\circ)} + \frac{a(4-4\sigma_2^\circ)}{(4\sigma_2^\circ-1)(1-2\sigma_2^\circ)}\right\}$$
$$\text{and} \quad \sigma_1^\circ = \min\left\{1, a+\frac{1}{3}\right\} \quad \text{and} \quad \sigma_2^\circ = \max\left\{\frac{3a}{10}+\frac{7}{10}, a+\frac{1}{4}\right\}.$$

Summary: We assume $\ell_I \in [0.355-\varepsilon_1, 0.645+\varepsilon_1]$, $a \in [0.53, 0.685]$ and $\sigma \in [0.6, a+\nu-\varepsilon_1]$. If $B(a,\sigma) \geq a+2\varepsilon_1$, we immediately have $R(F,(S_i),(\sigma_i),T) < x^{B(a,\sigma)}$. Thus we also assume $\sigma > (1+\nu)/2$.

At least one of (2.140), (2.141), (2.142) or ((2.145) and (2.146) with $\delta = 2$ or $\delta = 6$) holds if $\sigma \leq a+\nu-\varepsilon_1$ and

$$\max_{\triangle \in \mathcal{Z}_1 \cup \mathcal{Z}_{2,2} \cup \mathcal{Z}_3} \Lambda_*(a,\sigma;\triangle) \leq \ell_I \leq \min\left\{\min_{\triangle \in \mathcal{Z}_{2,1}} \Upsilon_*(a,\sigma;\triangle), \min_{\triangle \in \mathcal{Z}} \Upsilon_2(a,\sigma;\triangle)\right\}. \quad (2.174)$$

We also have one of (2.140), (2.141), (2.142) or ((2.145) and (2.146)) if $0.75 \leq \sigma \leq a+\nu-\varepsilon_1$ and

$$\max_{\triangle \in \mathcal{Z}_1 \cup \mathcal{Z}_{2,2} \cup \mathcal{Z}_3} \Lambda_*(a,\sigma;\triangle) \leq \ell_I \leq \min\left\{\min_{\triangle \in \mathcal{Z}_{2,1}} \Upsilon_*(a,\sigma;\triangle), \min_{\triangle \in \mathcal{Z}} \max_{s \in \{2,6\}} \Upsilon_s(a,\sigma;\triangle)\right\}. \quad (2.175)$$

One of (2.140) or (2.141) holds if ($\sigma \geq 0.65$ and $\Lambda_E(a,\sigma) \leq \ell_I \leq \Upsilon_{E,1}(a,\sigma)$) or ($\sigma \geq 0.75$ and $\Lambda_E \leq \ell_I \leq \Upsilon_{E,2}(a,\sigma)$). One of (2.140) or (2.142) holds if $\sigma \geq \min\{a+\nu-\varepsilon_1, 0.85\}$ and $\Lambda_Q(a,\sigma) \leq \ell_I \leq \Upsilon_Q(a,\sigma)$.

But if (2.145) and (2.146) are satisfied for some $(\gamma,\delta) \in \{(1,2),(4,6)\}$, then (2.143) holds and we are done. So all we have got left to do is to give explicit lower and upper bounds on the minima and maxima in (2.174) and (2.175). These expressions may appear complicated at first, but quantities $\Upsilon_*(a,\sigma;\triangle)$, $\Upsilon_2(a,\sigma;\triangle)$, $\Upsilon_6(a,\sigma;\triangle)$ and $\Lambda_*(a,\sigma;\triangle)$ individually are easy to bound for a and σ contained in some fixed intervals. In particular, we numerically compute the following:

$$\max_{\triangle\in\mathcal{Z}_1\cup\mathcal{Z}_{2,2}\cup\mathcal{Z}_3}\Lambda_*(a,\sigma;\triangle)\le\begin{cases}0.405-\varepsilon_1 & \text{if } a\in[0.53,0.545],\\ 0.400-\varepsilon_1 & \text{if } a\in[0.545,0.57],\\ 0.380-\varepsilon_1 & \text{if } a\in[0.57,0.59],\\ 0.365-\varepsilon_1 & \text{if } a\in[0.59,0.61],\end{cases} \tag{2.176}$$

$$\min\left\{\min_{\triangle\in\mathcal{Z}_{2,1}}\Upsilon_*(a,\sigma;\triangle),\min_{\triangle\in\mathcal{Z}}\Upsilon_2(a,\sigma;\triangle)\right\}\ge\begin{cases}0.485+\varepsilon_1 & \text{if } a\in[0.53,0.545],\\ 0.475+\varepsilon_1 & \text{if } a\in[0.545,0.57],\\ 0.455+\varepsilon_1 & \text{if } a\in[0.57,0.59],\\ 0.435+\varepsilon_1 & \text{if } a\in[0.59,0.61],\end{cases} \tag{2.177}$$

$$\min\{\min_{\triangle\in\mathcal{Z}_{2,1}}\Upsilon_*(a,\sigma;\triangle),\min_{\triangle\in\mathcal{Z}}\max_{s\in\{2,6\}}\Upsilon_s(a,\sigma;\triangle)\}\ge\begin{cases}0.485+\varepsilon_1 & \text{if } a\in[0.53,0.545],\\ 0.475+\varepsilon_1 & \text{if } a\in[0.545,0.57],\\ 0.455+\varepsilon_1 & \text{if } a\in[0.57,0.59],\\ 0.435+\varepsilon_1 & \text{if } a\in[0.59,0.61],\end{cases} \tag{2.178}$$

where in (2.176) we additionally assume $\sigma\in[(1+\nu)/2,a+\nu]$, in (2.177) we additionally assume $\sigma\le 0.75$ and in (2.178) we additionally assume $\sigma\in[0.75,a+\nu]$.

For $a\in[0.61,0.685]$, $\sigma\in[0.6,a+\nu]$ and $\ell_I\in[0.355-\varepsilon_1,0.42+\varepsilon_1]$, there are instances in which (2.174) and (2.175) fail to hold, particularly when (σ is close to 0.75 and ℓ_I is close to 0.42) or (σ is close to $a+\nu$ and ℓ_I is close to 0.355). But these cases are instead easily covered by (($\sigma\ge 0.65$ and $\Lambda_E\le\ell_I\le\Upsilon_{E,1}$) or ($\sigma\ge 0.75$ and $\Lambda_E\le\ell_I\le\Upsilon_{E,2}$)) or ($\sigma>0.85$ and $\Lambda_Q(a,\sigma)\le\ell_I\le\Upsilon_Q(a,\sigma)$), respectively. □

2.4.4.6 Midsized τ, large σ

Most parts of the proof of Lemma 2.23 are actually also applicable when $\sigma\ge a+\nu-\varepsilon_1$. In particular, combining the proof of Lemma 2.23 with Lemma 2.21, we get the following result, which is very important for the treatment of midsized τ and large σ:

Lemma 2.24. *Let $F(s)=\prod_{i=1}^{j}S_i(s)$ be as described at the start of Section 2.4. Recall $S_i(s)=\sum_{n\sim N_i}a_n^{(i)}n^{-s}$ and $N_i=N^{\ell_i}$. Let $T\in[T_1,T_0]$ and suppose that $F(s)$ is of Type A or B at T. Write $\tau=x^a$.*
Finally, suppose there exists $I\subseteq\{1,\dots,j\}$ such that I and a satisfy one of the following conditions:

(1) $\sum_{i\in I}\ell_i\in[0.315-\varepsilon_1,0.420+\varepsilon_1]\cup[0.580-\varepsilon_1,0.685+\varepsilon_1]$ *and* $a\in[0.57,0.59]$,

(2) $\sum_{i\in I}\ell_i\in[0.330-\varepsilon_1,0.420+\varepsilon_1]\cup[0.580-\varepsilon_1,0.670+\varepsilon_1]$ *and* $a\in[0.59,0.61]$,

(3) $\sum_{i\in I}\ell_i \in [0.355-\varepsilon_1, 0.420+\varepsilon_1]\cup[0.580-\varepsilon_1, 0.645+\varepsilon_1]$ *and* $a\in[0.61, 0.685]$.

Write $M=\prod_{i\in I}N_i$ *and* $M^\beta=\prod_{i\in I}N_i^{\sigma_i}$. *Let* $\mathcal{Z}$ *be as given in* (2.88). *Then if* $x\geq C$ *and* (σ_i) *has corresponding* $\sigma > a+\nu-\varepsilon_1$, *one of the following four inequalities holds:*

$$R(F,(S_i),(\sigma_i),T)\leq x^{1-\sigma}(\log x)^{-2J}, \tag{2.179}$$

$$R(F,(S_i),(\sigma_i),T)\leq \tau x^{1-2\sigma+\nu+2\varepsilon_1}, \tag{2.180}$$

$$R(F,(S_i),(\sigma_i),T)\leq \min\{\tau^{-1/4}M^{2\beta-1}x^{7/4-2\sigma+\nu/4-4\varepsilon_1}, M^{2\beta-2}x^{2-2\sigma-4\varepsilon_1}\}, \tag{2.181}$$

$$R(F,(S_i),(\sigma_i),T)\leq \min_{(U,V,W,X,Y,Z)\in\mathcal{Z}}\tau^U M^{V\beta-W}x^{X-Y\sigma+Z\nu+\varepsilon_1}. \tag{2.182}$$

Here C *is a large constant dependent only on* J *and* ε_1, *and* ε_1 *is assumed to be small.*

Lemma 2.24 explains our choice of $\chi_3(a)$ and completes the proof of Proposition 2.2 when $a\in(0.57, 0.685]$.

Proof. If $\sigma\geq\max\{\frac{3a}{10}+\frac{7}{10}, a+\frac14\}+10^{-100}$, we are done by option (5) of Lemma 2.21. Thus we only consider $\sigma\in[a+\nu-\varepsilon_1, \max\{\frac{3a}{10}+\frac{7}{10}, a+\frac14\}+10^{-100}]$. We also assume $R(F,(S_i),(\sigma_i),T) > x^{1-\sigma}(\log x)^{-2J}$. Suppose $(U,V,W,X,Y,Z)\in\mathcal{Z}$ is such that

$$\tau^U M^{V\beta-W}x^{X-Y\sigma+Z\nu+\varepsilon_1} = \min_{(U_1,V_1,W_1,X_1,Y_1,Z_1)\in\mathcal{Z}}\tau^{U_1}M^{V_1\beta-W_1}x^{X_1-Y_1\sigma+Z_1\nu+\varepsilon_1}.$$

Write $\ell_I=\sum_{i\in I}\ell_i$ and $\frac{\sum_{i\notin I}\ell_i\sigma_i}{(1-\ell_I)}=\alpha$, so that $N^\sigma=(N/M)^\alpha M^\beta=N^{(1-\ell_I)\alpha}N^{\ell_I\beta}$ and $\tau^U M^{V\beta-W}x^{X-Y\sigma+Z\nu+\varepsilon_1}\geq x^{(1-\ell_I)(W-V\alpha)+aU+(X-W)-(Y-V)\sigma+Z\nu}$. By Lemma 2.8,

$$R(F,(S_i),(\sigma_i),T)\leq x^{2\varepsilon_1}\max\left\{x^{(1-\ell_I)(2-2\alpha)}, \tau x^{(1-\ell_I)(4-6\alpha)}\right\}.$$

Rearranging, we thus have (2.182) whenever

$$(V-2)(1-\ell_I)\alpha\leq(1-\ell_I)(W-2)+aU+(X-W)-(Y-V)\sigma+Z\nu-2\varepsilon_1, \tag{2.183}$$

$$(V-6)(1-\ell_I)\alpha\leq(1-\ell_I)(W-4)+a(U-1)+(X-W)-(Y-V)\sigma+Z\nu-2\varepsilon_1. \tag{2.184}$$

In the proof of Lemma 2.23 we carefully derived conditions which ensure that (2.183) and (2.184) hold. We assumed $\sigma\leq a+\nu-\varepsilon_1$ and took $B(a,\sigma)=a+1-2\sigma+\nu+2\varepsilon_1$, but all but one of our arguments work just as well with $\sigma\in[a+\nu-\varepsilon_1, \max\{\frac{3a}{10}+\frac{7}{10}, a+\frac14\}+10^{-100}]$ and $B(a,\sigma)=1-\sigma-\varepsilon_1$. The exception is the treatment of (2.183) with $(U,V,W,X,Y,Z)=(1,4,4,3,4,1)$, for which we now give a different approach.

Note that (2.183) holds with $(U,V,W,X,Y,Z)=(1,4,4,3,4,1)$ whenever

$$(1-\ell_I)\alpha\leq\frac{a+1+\nu-2\varepsilon_1}{2}-\ell_I. \tag{2.185}$$

So suppose instead that $\alpha > (\frac{a+1+\nu-2\varepsilon_1}{2} - \ell_I)/(1-\ell_I)$. For $\ell_I < 0.5$, this gives $\alpha \geq 0.6$. By Lemma 2.12 with $\{1, \dots, j\} \setminus I$ in the place of I, then

$$R(F, (S_i), (\sigma_i), T) \leq x^{2\varepsilon_1} T^{(3-3\alpha)/(3\alpha-1)}. \tag{2.186}$$

Assume now that $\ell_I > 0.365$. Then $\ell_I > (a + 1 - \nu)/4$ for $a \leq 0.685$. Condition (2.181) is satisfied if

$$R(F, (S_i), (\sigma_i), T) \leq x^{(1-\ell_I)(2-2\alpha)-5\varepsilon_1}. \tag{2.187}$$

Combining (2.186) and (2.187) and rearranging, we thus require

$$\frac{2}{3}(1 - \ell_I)(3\alpha - 1) - a - C^* \varepsilon_1 \geq 0. \tag{2.188}$$

Here C^* is a large absolute constant. (We assumed $\alpha < 0.99$ as otherwise (2.179) follows trivially from (2.186) and $\sigma \leq \max\{3a/10 + 7/10, a + 1/4\} + 10^{-100} \leq 0.95$.) But $\alpha > (\frac{a+1+\nu-2\varepsilon_1}{2} - \ell_I)/(1 - \ell_I)$ and so

$$\frac{2}{3}(1 - \ell_I)(3\alpha - 1) - a - C^* \varepsilon_1 > \left(\frac{1}{3} + \nu - 2\varepsilon_1 - C^* \varepsilon_1\right) - \frac{4}{3}\ell_I.$$

Therefore we have (2.181) if $\ell_I \leq 1/4 + 3\nu/4 - C^{**}\varepsilon_1$. It suffices to take $\ell_I \leq 0.42 + \varepsilon_1$.

The requirement $0.365 \leq \ell_I \leq 0.42 + \varepsilon_1$ replaces conditions $\sigma \leq a + \nu - \varepsilon_1$ and $\ell_I \leq \Upsilon_*(a, \sigma; (1, 4, 4, 3, 4, 1))$ of the proof of Lemma 2.23. Leaving the rest of that proof unaltered, we thus have one of (2.179), (2.180), (2.181) or (2.182) if $\sigma \in [a + \nu - \varepsilon_1, \max\{3a/10 + 7/10, a + 1/4\} + 1)^{-100}]$ and

$$\max\left\{\max_{\triangle \in \mathcal{Z}_1 \cup \mathcal{Z}_{2,2} \cup \mathcal{Z}_3} \Lambda_*(a, \sigma; \triangle), 0.365\right\} \leq \ell_I \leq \min\left\{\min_{\triangle \in \mathcal{Z}} \Upsilon_6(a, \sigma; \triangle), 0.42 + \varepsilon_1\right\}. \tag{2.189}$$

Here $\Lambda_*(a, \sigma; \triangle)$ and $\Upsilon_6(a, \sigma; \triangle)$ are as defined in the proof of Lemma 2.23, except $B(a, \sigma) = 1 - \sigma - \varepsilon_1$. Alternatively, the proof of Lemma 2.23 also gives (2.179) or (2.181) whenever

$$\Lambda_Q(a, \sigma) \leq \ell_I \leq \Upsilon_Q(a, \sigma). \tag{2.190}$$

Computing the values of $\Lambda_*(a, \sigma; \triangle)$, $\Upsilon_6(a, \sigma; \triangle)$, $\Lambda_Q(a, \sigma)$ and $\Upsilon_Q(a, \sigma)$, we find that one of (2.189) or (2.190) holds for any $\sigma \in [a + \nu - \varepsilon_1, \max\{3a/10 + 7/10, a + 1/4\} + 1)^{-100}]$ provided we have ($a \in [0.57, 0.59]$ and $\ell_I \in [0.315 - \varepsilon_1, 0.420 + \varepsilon_1]$) or ($a \in [0.59, 0.61]$ and $\ell_I \in [0.33 - \varepsilon_1, 0.420 + \varepsilon_1]$) or ($a \in [0.61, 0.685]$ and $\ell_I \in [0.355 - \varepsilon_1, 0.420 + \varepsilon_1]$). $\square$

2.5 Comparison of sifted sets

The goal of this section is to prove Lemma 2.1. For certain $P_1, \dots, P_r$, we show

$$\left| \sum_{p_i \sim P_i} S(\mathcal{A}(y)_{p_1 \dots p_r}, z) - \frac{x^b}{\tau} \sum_{p_i \sim P_i} S(\mathcal{B}(y)_{p_1 \dots p_r}, z) \right| \leq \left(\frac{(\log\log x)^{O(1)} y}{\tau (\log x)^{2+r}} \right).$$

Here we use the following notation from sieve theory: For $\mathcal{C} \subseteq \mathbb{N}$ and $z > 0$, we write $P(z) = \prod_{p<z} p$. Then $S(\mathcal{C}, z) = \#\{n \in \mathcal{C} : (n, P(z)) = 1\}$. We also write $V(z) = \prod_{p<z}(1 - 1/p)$ and $\mathcal{C}_d = \{n \in \mathbb{N} : nd \in \mathcal{C}\}$.

2.5.1 Results so far

Lemma 2.1 will be derived from Proposition 2.1 and Proposition 2.2. These two propositions can be summarized as follows:

Corollary 2.25. *Let $\varepsilon > 0$ and $a \in [0.475 - \varepsilon, 0.77 - \varepsilon]$. Set $\tau = x^a$. Let $K = 2000$ and $0 \leq r_1 \leq r_2 \leq 10^5$. Let $N_i \in [\frac{1}{2}, \infty)$ with $2^{-r_2} x \leq \prod_{i=0}^{r_2} N_i \leq 6x$ and $N_i < x^{1/K}$ for $i > r_1$ and $N_i \geq x^{1/K}$ for $1 \leq i \leq r_1$. Let $a_n^{(i)} \in \mathbb{C}$ with $a_n^{(i)} = O(\log(n+3))$ for $i > r_1$. Then consider*

$$a_n = \sum_{\substack{k_0 k_1 \dots k_{r_2} = n \\ k_i \sim N_i}} \left(1_{\mathbb{N}}(k_0) \prod_{1 \leq i \leq r_1} 1_{\mathbb{P}}(k_i) \prod_{i > r_1} a_{k_i}^{(i)} \right).$$

Suppose also that one of the following three options holds:

(i) *$r_1 \geq 2$ or $N_0 \geq x^{0.95}$.*

(ii) *$r_1 = 1$ and $N_0 \geq x^{1/\log\log x}$.*

(iii) *$r_1 \in \{0, 1\}$ and there exists $i > r_1$ with $a_n^{(i)} = 1_{\mathbb{P}}(n)$ for all n and $N_i \geq x^{\frac{1}{\log\log x}}$.*

Let $A \in \mathbb{N}$. Write $N_i = x^{\ell_i^*}$. Denote by $\Xi(\ell_0^*, \dots, \ell_r^*)$ the set of finite sequences $\{\ell_i\}_{i=1}^j$ with $\ell_1, \dots, \ell_j \in [0, 1]$, $\sum_{i=1}^j \ell_i = 1$ and $j \leq A + 10^7 K$ for which there exist disjoint subsets $X_1, \dots, X_{r_1}$ of $\{1, \dots, j\}$ with

$$\sum_{i \in X_s} \ell_i \in \left[\ell_s^* - \frac{\varepsilon}{10^{100}}, \ell_s^* + \frac{\varepsilon}{10^{100}} \right] \quad \text{for} \quad 1 \leq s \leq r_1$$

and with $\ell_i \leq 2/K = 0.001$ for all but at most one $i \in \{1, \dots, j\} \setminus \bigcup_{s=1}^{r_1} X_s$.

Let $\chi_0(a)$, $\chi_1(a)$, $\chi_2(a)$ and $\chi_3(a)$ be as given in equations (2.9), (2.10), (2.11) *and* (2.12) *of Section* 2.2. *Suppose every $\{\ell_i\}_{i=1}^j \in \Xi(\ell_1^*, \dots, \ell_r^*)$ satisfies one of the following three options:*

(1) *There exists* $k \in \{1,\dots,j\}$ *with* $\ell_k \geq \chi_0(a)$.

(2) *There exists* $I_1 \subseteq \{1,\dots,j\}$ *with* $\sum_{i\in I_1} \ell_i \in \chi_1(a)$.

(3) *There exist* $I_2, I_3 \subseteq \{1,\dots,j\}$ *with* $\sum_{i\in I_2} \ell_i \in \chi_2(a)$ *and* $\sum_{i\in I_3} \ell_i \in \chi_3(a)$.

Then there exists $\mathcal{I} \subseteq [x,3x]\cap\mathbb{N}$ *such that* $\#\mathcal{I} = O(\tau x^{0.23+\varepsilon/2})$ *and such that* $y \notin ([x,3x]\cap\mathbb{N})\setminus\mathcal{I}$ *implies*

$$\left|\sum_{n\in\mathcal{A}(y)} a_n - \frac{x^b}{\tau}\sum_{n\in\mathcal{B}(y)} a_n\right| \leq \frac{x}{\tau(\log x)^A}.$$

Here the constant implied by big O notation only depends on A *and* ε.

Proof. Here (a_n) is as given at the start of Section 2.2.1. Let $F(s) = \prod_{i=1}^j S_i(s) \in \mathcal{F}((N_i), T, A + 10^7K, K)$ with $S_i(s) = \sum_{n\sim B_i} b_n^{(i)} n^{-s}$, $B = \prod_{i=1}^j B_i$ and $B_i = B^{\ell_i}$. There exist disjoint sets $X_1,\dots,X_{r_1} \subseteq \{1,\dots,j\}$ with $B_i < x^{1/K}$ for all but at most one $i \in \{1,\dots,j\}\setminus\bigcup_{\ell=1}^{r_1} X_\ell$ and

$$(\log x)^{-2K-1}N_\ell \leq \prod_{i\in X_\ell} B_i \leq (\log x)^{2K+1}N_\ell \quad \text{for} \quad 1\leq \ell\leq r_1.$$

Hence we must have $\{\ell_i\}_{i=1}^j \in \Xi(\ell_1^*,\dots,\ell_r^*)$. But then one of the options (1), (2) or (3) of Proposition 2.2 holds. Hence by Proposition 2.2, every choice of $F(s)$, $G(s)$, (σ_i) and γ described in Proposition 2.1 satisfies one of (C1), (C2), (C3), (C4.A) or (C4.B). But by Proposition 2.1 the set $\mathcal{I}$ of $y \in [x,3x]\cap\mathbb{N}$ with

$$\left|\sum_{n\in\mathcal{A}(y)} a_n - \frac{x^b}{\tau}\sum_{n\in\mathcal{B}(y)} a_n\right| > \frac{x}{\tau(\log x)^A}$$

then has $\#\mathcal{I} = O(\tau x^{0.23+\frac{\varepsilon}{2}})$, where the implied constant only depends on A and ε. □

2.5.2 Fundamental lemma

We now show how Corollary 2.25 can be used to compare $\mathcal{S}(\mathcal{A}(y),w)$ and $\mathcal{S}(\mathcal{B}(y),w)$ when w is very small. This, too, will be a key ingredient in the proof of Lemma 2.1.

Lemma 2.26. *Let* $y\in[x,3x]\cap\mathbb{N}$, $K_2\in\mathbb{N}$, $A>0$ *and* $w=\alpha x^{\frac{1}{\log\log x}}$ *for some* $\alpha\in[1,2]$. *Let* (c_s) *be a sequence with* $|c_s|\leq \log x$ *and* $c_s = 0$ *when* $(s,P(w))\neq 1$ *or* $s \geq x^{1-b-\frac{2}{K_2}}$.

Then there exist sequences $(\xi_d^{(1)})$ *and* $(\xi_d^{(2)})$*, dependent on* x*, such that* $|\xi_d^{(i)}| \leq 1$ *for* $i = 1, 2$ *and* $\xi_d^{(i)} = 0$ *for* $d \geq x^{1/K_2}$ *and such that the following is true:*

$$E_2 \leq \sum_s c_s\, S(\mathcal{A}(y)_s, w) - \frac{x^b}{\tau} \sum_s c_s\, S(\mathcal{B}(y)_s, w) \leq E_1, \tag{2.191}$$

where
$$E_1 = \sum_{\substack{dsn \in \mathcal{A}(y) \\ d|P(w)}} c_s\, \xi_d^{(1)} - \frac{x^b}{\tau} \sum_{\substack{dsn \in \mathcal{B}(y) \\ d|P(w)}} c_s\, \xi_d^{(1)} + O\biggl(\frac{x}{\tau(\log x)^A}\biggr), \tag{2.192}$$

$$E_2 = \sum_{\substack{dsn \in \mathcal{A}(y) \\ d|P(w)}} c_s\, \xi_d^{(2)} - \frac{x^b}{\tau} \sum_{\substack{dsn \in \mathcal{B}(y) \\ d|P(w)}} c_s\, \xi_d^{(2)} + O\biggl(\frac{x}{\tau(\log x)^A}\biggr). \tag{2.193}$$

Here the implied constants only depend on the choice of A *and* K_2*.*

Proof. We take $\delta(p) = 1$, $k = 1$, $s = (\frac{1}{K_2})(\frac{\log \alpha}{\log x} + \frac{1}{\log\log x})^{-1}$ and $y = x^{1/K_2}$ in Theorem 7 of [35]. We apply the theorem twice for each $q = s$, once with $A = \mathcal{A}(y)$, $X = \frac{y}{\tau}$, $R_d = |\mathcal{A}(y)| - \frac{y}{(\tau d)}$ and $\nu = 1$, and once with $A = \mathcal{B}(y)$, $X = \frac{y}{x^b}$, $R_d = |\mathcal{B}(y)| - \frac{y}{(x^b d)}$ and $\nu = 2$.

$$\begin{aligned} S(\mathcal{A}(y)_s, w) \leq{}& \frac{y}{\tau s} V(w) + O\left(\frac{y}{\tau s} \exp\left(-\frac{\log\log x \log\log\log x}{8K_2}\right)\right) \\ &+ \sum_{\substack{d|P(w) \\ d<x^{1/K_2}}} \xi_d^{(1)} \left(|\mathcal{A}(y)_{ds}| - \frac{y}{\tau ds}\right), \end{aligned} \tag{2.194}$$

$$\begin{aligned} S(\mathcal{B}(y)_s, w) \geq{}& \frac{y}{x^b s} V(w) + O\left(\frac{y}{x^b s} \exp\left(-\frac{\log\log x \log\log\log x}{8K_2}\right)\right) \\ &+ \sum_{\substack{d|P(w) \\ d<x^{1/K_2}}} \xi_d^{(2)} \left(|\mathcal{B}(y)_{ds}| - \frac{y}{x^b ds}\right). \end{aligned} \tag{2.195}$$

Now we note that $||\mathcal{B}(y)_{ds}| - y/(x^b ds)|$ is bounded by 1 and that

$$\begin{aligned} |\mathcal{A}(y)_{ds}| - \frac{y}{\tau ds} &= |\mathcal{A}(y)_{ds}| - \frac{x^b}{\tau}|\mathcal{B}(y)_{ds}| + \left(\frac{x^b}{\tau}\right)\left(|\mathcal{B}(y)_{ds}| - \frac{x}{x^b ds}\right) \\ &= |\mathcal{A}(y)_{ds}| - \frac{x^b}{\tau}|\mathcal{B}(y)_{ds}| + O\biggl(\frac{x^b}{\tau}\biggr). \end{aligned}$$

Substituting the identity into (2.194) and summing (2.194) and (2.195) over s, we get

$$\begin{aligned} &\sum_s c_s S(\mathcal{A}(y)_s, w) - \frac{x^b}{\tau} \sum_s c_s S(\mathcal{B}(y)_s, w) \\ &\leq O\biggl(\frac{x}{\tau}(\log x)^{-\log\log\log x/(8K_2)} \sum_{s \leq x} \frac{1}{s} + \sum_s |c_s| \frac{x^{b+1/K_2}}{\tau}\biggr) \end{aligned}$$

$$+\sum_s \sum_{d|P(w)} c_s \xi_d^{(1)} \left(|\mathcal{A}(y)_{ds}| - \frac{x^b}{\tau} |\mathcal{B}(y)_{ds}| \right)$$
$$= \sum_s \sum_{d|P(w)} c_s \xi_d^{(1)} \left(|\mathcal{A}(y)_{ds}| - \frac{x^b}{\tau} |\mathcal{B}(y)_{ds}| \right) + O\left(\frac{x}{\tau (\log x)^A} \right).$$

Here we used that $c_s = 0$ for $s \geq x^{1-b-2/K_2}$. By the same argument,

$$\sum_s c_s S(\mathcal{A}(y)_s, w) - \frac{x^b}{\tau} \sum_s c_s S(\mathcal{B}(y)_s, w) \geq \sum_s \sum_{d|P(w)} c_s \xi_d^{(2)} \left(|\mathcal{A}(y)_{ds}| - \frac{x^b}{\tau} |\mathcal{B}(y)_{ds}| \right),$$

where we omitted an error term of $O(x/(\tau(\log x)^A))$ on the RHS. □

2.5.3 Buchstab's identity and small prime factors

We now prove the first half of Lemma 2.1.

Lemma 2.27. *Consider $\tau = x^a$ with $x^{19/40-\varepsilon} \leq \tau \leq x^{0.77-\varepsilon}$. Let $\beta \in [0.01, 0.15]$. Let $r \in \{0, \dots, 5\}$ and $P_i = x^{\ell_i^*}$ with $\beta \leq \ell_i^* \leq 0.5+\varepsilon_1$, $P_i \geq P_{i+1}$ and $P_1 \dots P_r \leq x^{0.75}$.*

Denote by $\Xi^\star(\ell_1^*, \dots, \ell_r^*, \beta)$ the set of finite sequences $\{\ell_i\}_{i=1}^j$ with $\ell_1, \dots, \ell_j \in [0,1]$, $\sum_{i=1}^j \ell_i = 1$ and $j \leq 10^{20}$ for which $\exists$ disjoint sets $X_1, \dots, X_r \subseteq \{1, \dots, j\}$ with

$$\sum_{i \in X_s} \ell_i \in \left[\ell_s^* - \frac{\varepsilon}{10^{100}}, \ell_s^* + \frac{\varepsilon}{10^{100}} \right]$$

for $s \leq r$ and $\ell_i \leq \beta + 10^{-100}\varepsilon$ for all but at most one $i \in \{1, \dots, j\} \setminus \bigcup_{i=1}^r X_r$. Suppose that $\{\ell_i\}_{i=1}^j \in \Xi^\star(\ell_1^*, \dots, \ell_r^*, \beta)$ implies that one of the following options holds:

1. *There exists $k \in \{1, \dots, j\}$ with $\ell_k \geq \chi_0(a)$.*
2. *There exists $I_1 \subseteq \{1, \dots, j\}$ with $\sum_{i \in I_1} \ell_i \in \chi_1(a)$.*
3. *There exist $I_2, I_3 \subseteq \{1, \dots, j\}$ with $\sum_{i \in I_2} \ell_i \in \chi_2(a)$ and $\sum_{i \in I_3} \ell_i \in \chi_3(a)$.*

Here $\chi_0(a)$, $\chi_1(a)$, $\chi_2(a)$ and $\chi_3(a)$ are as described in (2.9), (2.10), (2.11) *and* (2.12), *respectively.*

Then $\exists\, \mathcal{I} \subseteq [x, 3x] \cap \mathbb{N}$ such that $\#\mathcal{I} = O(\tau x^{0.23+3\varepsilon/4})$ and $y \in ([x, 3x] \cap \mathbb{N}) \setminus \mathcal{I}$ implies

$$\left| \sum_{p_i \sim P_i} S(\mathcal{A}(y)_{p_1 \dots p_r}, x^\beta) - \frac{x^b}{\tau} \sum_{p_i \sim P_i} S(\mathcal{B}(y)_{p_1 \dots p_r}, x^\beta) \right| \leq \left(\frac{(\log\log x)^{O(1)} y}{\tau (\log x)^{2+r}} \right). \quad (2.196)$$

Here the constants implied by big O notation only depend on ε.

The key ingredients of the proof of this lemma are Corollary 2.25, Lemma 2.26 and the Buchstab identity. Lemma 2.27 is related to Theorem 3.1 of Harman's book [14], but the sequences (a_n) allowed in Corollary 2.25 are of a significantly more restrictive form than those considered in Theorem 3.1 and thus we include a sketch of the proof.

Proof. We first prove the following lower bound:

$$(-1)^r\left(\sum_{p_i\sim P_i} S(\mathcal{A}(y)_{p_1\dots p_r},x^\beta)-\frac{x^b}{\tau}\sum_{p_i\sim P_i} S(\mathcal{B}(y)_{p_1\dots p_r},x^\beta)\right)\geq -\left(\frac{(\log\log x)^{O(1)}y}{\tau(\log x)^{2+r}}\right). \tag{2.197}$$

Write $D=\lfloor\log(x^\beta/x^{1/\log\log x})/\log 2\rfloor$ and $w=2^{-D}x^\beta$. We repeatedly apply the Buchstab identity $S(\mathcal{C},z)=S(\mathcal{C},w)-\sum_{w\leq p<z}S(\mathcal{C}_p,p)$, starting with the LHS of (2.197). At intermediate steps, we look at

$$(-1)^k\sum_{\substack{p_\ell\sim P_\ell\\ p_k<p_{k-1}}} S(\mathcal{A}(y)_{p_1\dots p_k},p_k)-\frac{x^b}{\tau}S(\mathcal{B}(y)_{p_1\dots p_k},p_k), \tag{2.198}$$

where $P_i\in\{2^{-d}x^\beta:1\leq d\leq D\}$. The condition $p_k<p_{k-1}$ is only relevant when $P_k=P_{k-1}$. If $P_k<P_{k-1}$ and $p_1\dots p_{k-1}\leq x^{0.999}$, we decompose (2.198) further via the Buchstab identity. If $P_k=P_{k-1}$ and k is even, we bound (2.198) trivially from below by discarding the contribution of the $\mathcal{A}(y)$–terms and only recording the $\mathcal{B}(y)$–terms. If $P_k=P_{k-1}$ and k is odd, we instead remove the condition $p_k<p_{k-1}$ from the $\mathcal{A}(y)$–terms, making their contribution more negative, and also add and subtract the corresponding $\mathcal{B}(y)$–terms. For k odd, we then decompose (2.198) further, but the restriction $p_k<p_{k-1}$ is omitted.

Overall, we obtain the following lower bound on the LHS of (2.197):

$$(-1)^r\sum_{p_i\sim P_i\text{ for }i\leq r} S(\mathcal{A}(y)_{p_1\dots p_r},w)-\frac{x^b}{\tau}S(\mathcal{B}(y)_{p_1\dots p_r},w) \tag{2.199}$$

$$+\sum_{k=r+1}^{\lfloor\log\log x\rfloor}\sideset{}{^*}\sum_{P_{r+1},\dots,P_k}(-1)^k\sum_{\substack{p_\ell\sim P_\ell\\ p_1p_2p_3\dots p_{k-1}\leq x^{0.999}}} S(\mathcal{A}(y)_{p_1\dots p_k},w)-\frac{x^b}{\tau}S(\mathcal{B}(y)_{p_1\dots p_k},w) \tag{2.200}$$

$$+\sum_{k=r+3}^{\lfloor\log\log x\rfloor}\sideset{}{^*}\sum_{P_{r+1},\dots,P_k}(-1)^k\sum_{\substack{p_\ell\sim P_\ell\\ p_1p_2p_3\dots p_{k-2}\leq x^{0.999}\\ p_1p_2p_3\dots p_{k-1}>x^{0.999}}} S(\mathcal{A}(y)_{p_1\dots p_k},p_k)-\frac{x^b}{\tau}S(\mathcal{B}(y)_{p_1\dots p_k},p_k) \tag{2.201}$$

$$-\frac{x^b}{\tau}\sum_{k=r+2}^{\lfloor\log\log x\rfloor}\sideset{}{^*}\sum_{P_{r+1},\dots,P_{k-1}}\sum_{\substack{p_\ell\sim P_\ell\\ p_k\sim P_{k-1}\\ p_1p_2p_3\dots p_{k-2}\leq x^{0.999}}} S(\mathcal{B}(y)_{p_1\dots p_k},p_k), \tag{2.202}$$

where sums $\sum^*_{P_{r+1},\dots,P_j}$ range over all choices of $P_i \in \{\frac{x^\beta}{2^d} : 1 \le d \le D\}$ with $P_{r+1} \ge P_{r+2} \ge \cdots \ge P_j$ and $P_i > P_{i+1}$ for odd i. Recall here that $P_1, \dots, P_r$ are fixed.

We first bound the contribution of (2.202). Since $\mathcal{B}(y)$ is an interval of length about x/x^b, we can use standard results, such as [19], to estimate $\sum S(\mathcal{B}(y)_{p_1\dots p_k}, p_k)$ for p_i ranging over some intervals, provided we impose condition $\frac{x}{x^b p_1\dots p_k} \ge (\frac{x}{p_1\dots p_k})^{7/12}$. Since $b \le 10^{-5}$, we only require say $p_1p_2p_3\dots p_k \le x^{0.9999}$. We split (2.202) into $(2.202)_1$+$(2.202)_2$, where $(2.202)_1$ is the component of (2.202) with $p_1p_2p_3\dots p_k \le x^{0.9999}$ and $(2.202)_2$ is the component with $p_1p_2p_3\dots p_k > x^{0.9999}$.

We fix k and $\mathcal{M} \subseteq \{r+1, \dots, k-1\} \cap \mathbb{N}^{\text{odd}}$, set $\widehat{m} = \#\mathcal{M}$ and denote by $\sum^{**}_{P_1,\dots,P_{k-1}}$ the sum over $P_1, \dots, P_{k-1}$ which appear in $\sum^*_{P_1,\dots,P_{k-1}}$ and additionally satisfy $P_i = P_{i-1}$ if and only if $i \in \mathcal{M}$. Then

$$\sum^{**}_{P_{r+1},\dots,P_{k-1}} \sum_{\substack{p_\ell \sim P_\ell \\ p_k \sim P_{k-1} \\ p_1p_2p_3\cdots p_{k-2} \le x^{0.999} \\ p_1p_2p_3\cdots p_k \le x^{0.9999}}} S(\mathcal{B}(y)_{p_1\dots p_k}, p_k) \le \frac{1}{(k-1-r-2\widehat{m})!} \sum_{\substack{p_1\dots p_k \le x^{0.9999}, \\ p_\ell \sim P_\ell \text{ for } \ell \le r \\ w \le p_\ell \le x^\beta \text{ for } \ell > r,\, \ell \notin \mathcal{M}\cup\{k\} \\ p_\ell \le 2p_{\ell-1} \text{ for } \ell \in \mathcal{M}\cup\{k\} \\ p_\ell \ge (1/2)p_{\ell-1} \text{ for } \ell \in \mathcal{M}\cup\{k\}}} S(\mathcal{B}(y)_{p_1\dots p_k}, p_k), \tag{2.203}$$

where we used that in $\sum^{**}$, $P_{r+1} \ge \cdots \ge P_{k-1}$ with $P_i = P_{i-1}$ for only $\widehat{m}$ choices of i, and removed the double counting introduced by ignoring this condition by dividing by $(k-1-r-2\widehat{m})!$.

Let $\gamma(x) = \frac{\log 2}{\log x}$. For $i \in \mathcal{M} \cup \{k\}$ we set $\beta_i^+ = \alpha_{i-1} + \gamma(x)$ and $\beta_i^- = \alpha_{i-1} - \gamma(x)$. For $i \notin \mathcal{M} \cup \{1, \dots, r,\} \cup \{k\}$ we set $\beta_i^+ = \beta$ and $\beta_i^- = \frac{1}{\log\log x}$. For $i \in \{1, \dots r\}$ we set $\beta_i^+ = \ell_i^* + \gamma(x)$ and $\beta_i^- = \ell_i^*$. For $p_1 \dots p_k \le x^{0.9999}$, standard computations give

$$\begin{aligned} \text{RHS of (2.203)} &\le \frac{y(1+o(1))}{x^b(k-1-r-2\widehat{m})!} \int_{\beta_1^-}^{\beta_1^+} \cdots \int_{\beta_k^-}^{\beta_k^+} \frac{1}{\alpha_1 \dots \alpha_{k-1}\alpha_k^2 \log x} \mathrm{d}\alpha_k \dots \mathrm{d}\alpha_1 \\ &\le \frac{C^* y}{x^b(k-1-r-2\widehat{m})!} \frac{(\log\log\log x)^{k-1-r-\widehat{m}} (\log\log x)^{\widehat{m}+2}}{(\log x)^{\widehat{m}+2+r}}, \end{aligned}$$

where C^* is a large constant. Summing over all choices of $k \le \log\log x$ and $\mathcal{M} \subseteq \{r+1, \dots, k-1\} \cap \mathbb{N}^{\text{odd}}$, $(2.202)_1$ is thus bounded below by $-y(\log\log x)^5/(\tau(\log x)^{2+r})$.

Next we consider $(2.202)_2$, where $p_1p_2\dots p_k > x^{0.9999}$. Since we also got $p_1 \dots p_{k-2} \le x^{0.999}$, the relevant products have $p_1 \ge p_2 \ge \cdots \ge p_{k-1} \ge x^{0.0001}$. This implies that any $n \in [x, 6x]$ with m prime factors greater or equal to $x^{0.0001}$ is counted at most m^{k-1} times by quantities $S(\mathcal{B}(y)_{p_1\dots p_k}, p_k)$ present in $(2.202)_2$. Here we must have $m \le 10001$ and $k \le 10001$ and hence we count any single $n \in [x, 6x]$ at most $O(1)$

times. For a lower bound on $(2.202)_2$, it thus suffices to count the elements of $\mathcal{B}(y)$ which have two prime factors p_{k-1} and p_k which are very close together and also have prime factors $p_i \sim P_i$ for $i \le r$. Using again [19], simple computations give a lower bound of $-\frac{y(\log\log x)^3}{(\tau(\log x)^{2+r})}$ and overall we find that (2.202) $\ge -\frac{2y(\log\log x)^5}{(\tau(\log x)^{2+r})}$.

To treat (2.201), we fix some k, $t_0 = r < t_1 < \cdots < t_s = k-2$ and $Q_\ell \in \{2^{-d}x^\beta : 1 \le d \le D\}$ for $\ell \in \{t_0+1, t_1, t_1+1, t_2, \dots, t_{s-1}+1, t_s, k-1, k\}$ and denote by

$$\mathcal{T}(t_0, \dots, t_s, Q_{t_0+1}, Q_{t_1}, Q_{t_1+1}, Q_{t_2}, \dots, Q_{t_{s-1}+1}, Q_{t_s}, Q_{k-1}, Q_k, k)$$

the set of $(P_\ell)_{\ell=r+1}^k \in \{2^{-d}x^\beta : 1 \le d \le D\}^{k-r}$ with the following properties:

1. $P_\ell = Q_\ell$ whenever $\ell = t_{i-1}+1$ or $\ell = t_i$ for $1 \le i \le s$ or $\ell = k-1$ or $\ell = k$.
2. $P_{r+1} \ge \cdots \ge P_k$ and $P_\ell > P_{\ell+1}$ for odd ℓ.
3. $P_{t_i+1} \dots P_{t_{i+1}-1} < x^{0.00005}$ and $P_{t_i+1} \dots P_{t_{i+1}} \ge x^{0.00005}$ for $0 \le i \le s-2$.
4. $P_{t_{s-1}+1} \dots P_{t_s} \le x^{0.0001}$ if $t_s - t_{s-1} \ge 2$.

We can partition the set of $(P_\ell)_{\ell=r+1}^k$ which are under consideration in $\sum^*_{P_{r+1},\dots,P_k}$ in (2.201) via such $\mathcal{T}(t_0, \dots, t_s, Q_{t_0+1}, Q_{t_1}, Q_{t_1+1}, Q_{t_2}, \dots, Q_{t_{s-1}+1}, Q_{t_s}, Q_{k-1}, Q_k, k)$. Here $k \le \log\log x$, $s \le 10^5$, and there are at most $\log\log x$ choices for each t_i and $\log x$ choices for each Q_i. Only $(\log x)^{10^6}$ sets are needed to complete the partition.

For a $\mathcal{T}(t_0, \dots, t_s, Q_{t_0+1}, Q_{t_1}, Q_{t_1+1}, Q_{t_2}, \dots, Q_{t_{s-1}+1}, Q_{t_s}, Q_{k-1}, Q_k, k)$ of interest, note

$$\sum_{(P_\ell)_{\ell=r+1}^k \in \mathcal{T}} \sum_{\substack{p_\ell \sim P_\ell \\ p_1 \dots p_{k-2} \le x^{0.999} \\ p_1 \dots p_{k-1} > x^{0.999}}} S(\mathcal{C}_{p_1 \dots p_k}, p_k) = \sum_{\substack{n_1 \dots n_{r+s+2} \in \mathcal{C} \\ n_i \sim P_i \text{ for } i \le r \\ n_{r+s+1} \sim Q_{k-1} \\ n_1 \dots n_{r+s} \le x^{0.999} \\ n_1 \dots n_{r+s+1} > x^{0.999}}} a_{n_1}^{(1)} \dots a_{n_{r+s+2}}^{(r+s+2)} \tag{2.204}$$

where $a_n^{(i)} = 1_{\mathbb{P}}(n) \quad \text{for} \quad i \in \{1, \dots, r, r+s+1\}$,

$$a_n^{(r+i+1)} = \sideset{}{^*}\sum_{\substack{P_{t_i+1},\dots,P_{t_{i+1}} \\ P_{t_i+1} \dots P_{t_{i+1}-1} < x^{0.00005} \\ P_{t_i+1} \dots P_{t_{i+1}} \ge x^{0.00005} \\ P_{t_i+1} = Q_{t_i+1},\, P_{t_{i+1}} = Q_{t_{i+1}}}} \sum_{\substack{p_{t_i+1} \dots p_{t_{i+1}} = n \\ p_\ell \sim P_\ell}} 1 \quad \text{for} \quad i \in \{0, \dots, s-2\},$$

$$a_n^{(r+s)} = \sideset{}{^*}\sum_{\substack{P_{t_{s-1}+1},\dots,P_{t_s} \\ P_{t_{s-1}+1} \dots P_{t_s} \le x^{0.0001} \text{ if } t_s - t_{s-1} \ge 2 \\ P_{t_{s-1}+1} = Q_{t_{s-1}+1},\, P_{t_s} = Q_{t_s}}} \sum_{\substack{p_{t_{s-1}+1} \dots p_{t_s} = n \\ p_\ell \sim P_\ell}} 1,$$

$$a_n^{(r+s+2)} = \sum_{p_k \sim Q_k} 1_{\mathbb{N}}\left(\frac{n}{p_k}\right) \cdot 1_{\{1\}}\left(\gcd\left(\frac{n}{p_k}, P(p_k)\right)\right).$$

Here $a_n^{(i)} = O(\log n)$. We split (2.204) up by restricting the summation ranges to $n_i \sim N_i$ for $r+1 \leq i \leq r+s$ and $n_{r+s+2} \sim N_{r+s+2}$. The properties of $\mathcal{T}$ ensure that we only need to consider $(x^{0.00005}/2) \leq N_i \leq x^{\beta}$ for $r+1 \leq i \leq r+s$, while $N_{r+s+2} \leq x^{0.0001}$. Additionally, $a_n^{(i)} = 1_{\mathbb{P}}(n)$ if $N_i \geq x^{0.0001}$. Set

$$a_n = \sum_{\substack{n_1 \dots n_{r+s+2} = n \\ n_i \sim P_i \text{ for } i \leq r \\ n_i \sim N_i \text{ for } r+1 \leq i \leq r+s \\ n_{r+s+1} \sim Q_{k-1} \\ n_1 \dots n_{r+s} \leq x^{0.999} \\ n_1 \dots n_{r+s+1} > x^{0.999}}} a_{n_1}^{(1)} \dots a_{n_{r+s+2}}^{(r+s+2)}.$$

We split (2.201) into expressions $(-1)^k (\sum_{n \in \mathcal{A}(y)} a_n - (\frac{x^b}{\tau}) \sum_{n \in \mathcal{B}(y)} a_n)$ with a_n as described above. Conditions $n_1 \dots n_{r+s} \leq x^{0.999}$, $n_1 \dots n_{r+s+1} > x^{0.999}$ are only relevant if $P_1 \dots P_r N_{r+1} \dots N_{r+s} \in [\frac{x^{0.999}}{2^{r+s}}, x^{0.999}]$ or $P_1 \dots P_r N_{r+1} \dots N_{r+s} Q_{k-1} \in [\frac{x^{0.999}}{2^{r+s+1}}, x^{0.999}]$. In that case, if k is even, we bound $(-1)^k (\sum_{n \in \mathcal{A}(y)} a_n - (\frac{x^b}{\tau}) \sum_{n \in \mathcal{B}(y)} a_n)$ below by discarding the contribution of the sum over $\mathcal{A}(y)$ and bounding below the contribution of long interval $\mathcal{B}(y)$. To do so, we use standard integral computations, very similar to the ones used in the treatment of (2.202). In particular, for given $N_{r+1}, \dots, N_{r+s-1}$, the choice of N_{r+s} is severely restricted by requirements $P_1 \dots P_r N_{r+1} \dots N_{r+s} \in [\frac{x^{0.999}}{2^{r+s}}, x^{0.999}]$ or $P_1 \dots P_r N_{r+1} \dots N_{r+s} Q_{k-1} \in [\frac{x^{0.999}}{2^{r+s+1}}, x^{0.999}]$ (effectively giving us an extra factor of $1/\log x$). We get a total lower bound of $-y(\log\log x)^5/(\tau(\log x)^{2+r})$ for the contribution of problematic terms with even k. If k is odd, we instead remove conditions $n_1 \dots n_{r+s} \leq x^{0.999}$ and $n_1 \dots n_{r+s+1} > x^{0.999}$ from the $\mathcal{A}(y)$ term, making it smaller, and from the $\mathcal{B}(y)$ term, making it larger. What we have added to the $\mathcal{B}(y)$ term is again easy to bound from above, as interval $\mathcal{B}(y)$ is long. In total, the error introduced by removal of these conditions for even k is bounded below by $-y(\log\log x)^5/(\tau(\log x)^{2+r})$.

Hence we may exclusively work with (a_n) for which the conditions $n_1 \dots n_{r+s} \leq x^{0.999}$ and $n_1 \dots n_{r+s+1} > x^{0.999}$ have been removed or are not relevant. But then (a_n) is of the form given in Corollary 2.25. The corresponding $\{\ell_i\}_{i=1}^{j} \in \Xi$ are all also contained in $\Xi^*(\ell_1^*, \dots, \ell_r^*, \beta)$ and so satisfy one of the options (1), (2) or (3). Hence Corollary 2.25 applies. We apply it once for each relevant choice of $\mathcal{T}$ and N_i – no more than $(\log x)^{10^7}$ applications are needed in total. We find that there is some $\mathcal{I}_1 \subseteq [x, 3x] \cap \mathbb{N}$ with $\#\mathcal{I}_1^* \ll (\log x)^{10^7} \tau x^{\nu+\varepsilon/2}$ such that $y \in ([x, 3x] \cap \mathbb{N}) \setminus \mathcal{I}_1$ implies that (2.201) is at least $-y(\log\log x)^6/(\tau(\log x)^{2+r})$.

We apply Lemma 2.26 to (2.199), obtaining upper and lower bounds E_1 and E_2, where

$$E_s = \sum_{\substack{dp_1\dots p_r n \in \mathcal{A}(y)\\ p_i\sim P_i \text{ for } i\le r\\ d|P(w)}} \xi_d^{(s)} - \frac{x^b}{\tau}\sum_{\substack{dp_1\dots p_r n \in \mathcal{B}(y)\\ p_i\sim P_i \text{ for } i\le r\\ d|P(w)}} \xi_d^{(s)} + O\left(\frac{x}{\tau(\log x)^{100}}\right), \tag{2.205}$$

and where $\xi_d^{(s)} = 0$ for $d \geq x^{0.0001}$ and $s \in \{1,2\}$. Decomposing into sums

$$\sum_{\substack{n_0\dots n_r d\in\mathcal{A}(y)\\ d|P(w)\\ n_0\sim N,\, n_i\sim P_i,\, d\sim D}} 1_{\mathbb{N}}(n_0)1_{\mathbb{P}}(n_1)\dots 1_{\mathbb{P}}(n_r)\xi_d^{(s)} - \frac{x^b}{\tau}\sum_{\substack{n_0\dots n_r d\in\mathcal{B}(y)\\ d|P(w)\\ n_0\sim N,\, n_i\sim P_i,\, d\sim D}} 1_{\mathbb{N}}(n_0)1_{\mathbb{P}}(n_1)\dots 1_{\mathbb{P}}(n_r)\xi_d^{(s)} \tag{2.206}$$

with $x/2^{r+2} \leq NP_1\dots P_r D \leq 6x$ and $D \leq x^{0.0001}$, so that $N \geq x^{0.9999-\ell_1^*-\dots-\ell_r^*}$, we see that Corollary 2.25 is applicable with option (i) or (ii). The corresponding $\{\ell_i\}_{i=1}^{j} \in \Xi$ are again contained in $\Xi^*(\ell_1^*,\dots,\ell_r^*,\beta)$. Hence the assumptions of Lemma 2.27 together with Corollary 2.25 imply that there is some $\mathcal{I}_2 \subseteq [x,3x]\cap\mathbb{N}$ with $\#\mathcal{I}_2 \ll \tau x^{\nu+\varepsilon/2}$ such that $y \in ([x,3x]\cap\mathbb{N})\setminus\mathcal{I}_2$ implies that quantity (2.199) is $O(x/(\tau(\log x)^{100}))$.

To treat (2.200), we simply combine the techniques used in the treatment of (2.199) and (2.201). To be a little more precise, we decompose (2.200) using a partition very similar to the previously described $\mathcal{T}$, splitting the expression into $\leq (\log x)^{10^6}$ parts. To each of the resulting parts we then apply Lemma 2.26 and the resulting upper and lower bounds can be described by expressions similar to (2.205), except now k rather than r prime factors appear in the sum and these are combined using $\mathcal{T}$ like in (2.204). Corollary 2.25 is applicable to the resulting expressions and gives us some $\mathcal{I}_3 \subseteq [x,3x]\cap\mathbb{N}$ with $\#\mathcal{I}_3 \ll (\log x)^{10^7}\tau x^{\nu+\varepsilon/2}$ such that $y \in ([x,3x]\cap\mathbb{N})\setminus\mathcal{I}_3$ implies that quantity (2.200) is at least $-y(\log\log x)^6/(\tau(\log x)^{2+r})$.

Combining our bounds on (2.199), (2.200), (2.201) and (2.202), we now have that $\mathcal{I}^* = \mathcal{I}_1\cup\mathcal{I}_2\cup\mathcal{I}_3$ satisfies $\#\mathcal{I}^* \ll \tau x^{\nu+3\varepsilon/4}$ and whenever $y \in ([x,3x]\cap\mathbb{N})\setminus\mathcal{I}^*$,

$$(-1)^r\left(\sum_{p_i\sim P_i} S(\mathcal{A}(y)_{p_1\dots p_r}, x^\beta) - \frac{x^b}{\tau}\sum_{p_i\sim P_i} S(\mathcal{B}(y)_{p_1\dots p_r}, x^\beta)\right) \geq -\left(\frac{(\log\log x)^7 y}{\tau(\log x)^{2+r}}\right).$$

This assumes that x is large, but the required size of x depends only on ε.

The corresponding upper bound

$$(-1)^r\left(\sum_{p_i\sim P_i} S(\mathcal{A}(y)_{p_1\dots p_r}, x^\beta) - \frac{x^b}{\tau}\sum_{p_i\sim P_i} S(\mathcal{B}(y)_{p_1\dots p_r}, x^\beta)\right) \leq \left(\frac{(\log\log x)^7 y}{\tau(\log x)^{2+r}}\right)$$

can be derived in almost exactly the same way, we only need to swap the words "odd" and "even" or "lower" and "upper" in various places. We omit its proof here. □

2.5.4 Products of larger primes

We now prove the second half of Lemma 2.1.

Lemma 2.28. *Consider $\tau = x^a$ with $x^{19/40-\varepsilon} \le \tau \le x^{0.77-\varepsilon}$. Let $r \in \{2,4,6\}$. Let $P_i = x^{\ell_i^*}$ with $0.01 - \varepsilon \le \ell_i^* \le 0.5 + \varepsilon$, $P_i \ge P_{i+1}$ and $P_1 \dots P_r \le x^{0.99}$. Let $\Xi^{\star\star}(\ell_1^*, \dots, \ell_r^*)$ be as described in Definition 2.4. Suppose that every $\{\ell_i\}_{i=1}^{j} \in \Xi^{\star\star}(\ell_1^*, \dots, \ell_r^*)$ satisfies one of the options (1), (2) or (3).*

Then $\exists \mathcal{I} \subseteq [x, 3x] \cap \mathbb{N}$ such that $\#\mathcal{I} = O(\tau x^{0.23+3\varepsilon/4})$ and $y \in ([x,3x] \cap \mathbb{N}) \setminus \mathcal{I}$ implies

$$\left| \sum_{p_i \sim P_i} S(\mathcal{A}(y)_{p_1 \dots p_r}, p_r) - \frac{x^b}{\tau} \sum_{p_i \sim P_i} S(\mathcal{B}(y)_{p_1 \dots p_r}, p_r) \right| \le \left(\frac{(\log\log x)^{O(1)} y}{\tau (\log x)^{2+r}} \right). \quad (2.207)$$

Here the constants implied by big O notation depend only on ε.

Proof. Again we first derive the lower bound

$$\sum_{p_i \sim P_i} S(\mathcal{A}(y)_{p_1 \dots p_r}, p_r) - \frac{x^b}{\tau} \sum_{p_i \sim P_i} S(\mathcal{B}(y)_{p_1 \dots p_r}, p_r) \ge - \left(\frac{(\log\log x)^{O(1)} y}{\tau (\log x)^{2+r}} \right). \quad (2.208)$$

We split the LHS of (2.208) into sums

$$\frac{1}{k!} \left(\sum_{\substack{n_i \sim P_i \\ m_i \sim M_i \\ m_i > n_r \\ \prod n_i \prod m_i \in \mathcal{A}(y)}} \prod_{i=1}^{r} 1_{\mathbb{P}}(n_i) \prod_{i=1}^{k} 1_{\mathbb{P}}(m_i) - \frac{x^b}{\tau} \sum_{\substack{n_i \sim P_i \\ m_i \sim M_i \\ m_i > n_r \\ \prod n_i \prod m_i \in \mathcal{B}(y)}} \prod_{i=1}^{r} 1_{\mathbb{P}}(n_i) \prod_{i=1}^{k} 1_{\mathbb{P}}(m_i) \right), \quad (2.209)$$

where we consider $1 \le k \le (1 - \sum_{i=1}^{r} \ell_i^*)/\ell_r^*$ and $M_i \in \{2^d P_r : 0 \le d \le D\}$ for $D = \lceil \log(x/P_r)/\log 2 \rceil$. The factor $\frac{1}{k!}$ is included to account for double counting, which arises from rearrangements of M_i.

Double counting is still an issue when $M_i = M_j$ for some $i \ne j$. However, the contribution of such terms to the LHS of (2.208) is small: We discard the $\mathcal{A}$-term in (2.209) and bound below the $\mathcal{B}$-term, noting that it only involves $n \in \mathcal{B}(y)$ which have two prime factors q_1, q_2 with $q_1, q_2 \sim M_i$ as well as prime factors $p_\ell \sim P_\ell$ for $1 \le \ell \le r$. Since $\mathcal{B}(y)$ is an interval of length y/x^b and we multiply by x^b/τ, standard computations then give a bound of $O(y/(\tau(\log x)^{3+r}))$. Summing over the $O(\log x)$ choices for M_i then gives a total bound of $O(y/(\tau(\log x)^{2+r}))$. Hence we can ignore the contribution of case $M_i = M_j$.

Similarly, the condition $m_i > n_r$ is only relevant whenever $M_i = P_r$ for some i and this case also only contributes $O(y/(\tau(\log x)^{r+2}))$. For $M_i \neq P_r$, (2.209) is of the form described in Corollary 2.25 with option (i) and its corresponding Ξ is contained in $\Xi^{\star\star}(\ell_1^*, \ldots, \ell_r^*)$. Hence our assumptions ensure that Corollary 2.25 is applicable and we obtain that there exists some $\mathcal{I}^* \cap [x, 3x] \subseteq \mathbb{N}$ with $\#\mathcal{I}^* = O(\tau x^{\nu+3\varepsilon/4})$ such that every $y \in ([x, 3x] \cap \mathbb{N}) \setminus \mathcal{I}^*$ satisfies (2.208).

The corresponding upper bound $((\log\log x)^{O(1)}y)/(\tau(\log x)^{2+r})$ is derived in almost exactly the same way and so we omit its proof. $\square$

2.6 Harman's sieve

The goal of this section is to prove Proposition 2.3. We construct a minorant of the prime indicator function with various nice properties, linked to a comparison of sums over sets $\mathcal{A}(y)$ and $\mathcal{B}(y)$. This approach is based on Harman's sieve [14]. Relatedly, Harman's sieve was also used by Baker, Harman and Pintz [1] to improve an upper bound on the lengths of prime gaps. It was also used by Matomäki [27] and later Islam [24] to improve a bound on the number of very long prime gaps. Simultaneously to our work, these results were improved by Järviniemi [26], who showed that $\sum_{p_n \leq x,\, p_{n+1}-p_n \geq x^{1/2}} (p_{n+1} - p_n) \ll_\varepsilon x^{0.57+\varepsilon}$.

2.6.1 Preparation

Before embarking on the proof of Proposition 2.3, we introduce a few more definitions.

2.6.1.1 Good* and good** functions

Recall that our ultimate aim is to show that $\mathcal{A}(y) = [y, y + y/\tau]$ contains primes for all but $O(\tau x^{0.23+\varepsilon})$ choices of $y \in [x, 3x] \cap \mathbb{N}$. We write $\tau = x^a$.

Write $X = (7x)^{1/2}$. Fix some $a \in [0.475 - \varepsilon, 0.77 - \varepsilon]$. Throughout Section 2.6 we say that $\Psi : [x, 6x] \cap \mathbb{N} \to \mathbb{R}$ is good* if it can be written as

$$\Psi(n) = (-1)^r \sum_{\substack{n = p_1 \ldots p_r m \\ x^\beta \leq p_r < \cdots < p_1 < X \\ \mathcal{G}}} \psi(m, x^\beta)$$

for some $\beta \in [0.01, 0.15]$, $r \in \{0, \ldots, 5\}$ and a set of conditions $\mathcal{G}$, with $\#\mathcal{G} \leq 100$ and each $g \in \mathcal{G}$ of the form $\sum_{i \in I_g} k_{(g,i)} \log_x(p_i) \in [a_g, b_g]$ for some $I_g \subseteq \{1, \ldots, r\}$, $k_{(g,i)} \in \mathbb{N}$ and $a_g, b_g \in [0, 1]$. Additionally, we require that condition $(0.5+\varepsilon \geq \ell_1^* \geq \cdots \geq \ell_r^* \geq \beta$

and $\sum_{i\in I_g} k_{(g,i)}\ell_i^* \in [a_g - \frac{\varepsilon_1}{2}, b_g + \frac{\varepsilon_1}{2}]$ for all $g \in \mathcal{G}$) implies $(\ell_1^*, \ldots, \ell_r^*, \beta) \in \mathcal{R}^\star(a)$. (We also allow $[a_g, b_g]$ to be replaced by $(a_g, b_g]$, $[a_g, b_g)$ or (a_g, b_g).)

We say a function $\Psi : [x, 6x] \cap \mathbb{N} \to \mathbb{R}$ is good** if it can be written as

$$\Psi(n) = \sum_{\substack{n=p_1\ldots p_r m \\ p_r<\cdots<p_1 \\ \mathcal{G}}} \psi(m, p_r)$$

for some $r \in \{2, 4, 6\}$ and a set of conditions $\mathcal{G}$, with $\#\mathcal{G} \leq 100$ and $g \in \mathcal{G}$ of the form $\sum_{i\in I_g} k_{(g,i)} \log_x(p_i) \in [a_g, b_g]$ for some $I_g \subseteq \{1, \ldots, r\}$, $k_{(g,i)} \in \mathbb{N}$ and $a_g, b_g \in [0, 1]$. Additionally, we require that $(0.5 + \varepsilon \geq \ell_1^* \geq \cdots \geq \ell_r^* \geq 0.01 - \varepsilon$ and $\sum_{i\in I_g} k_{(g,i)}\ell_i^* \in [a_g - \frac{\varepsilon_1}{2}, b_g + \frac{\varepsilon_1}{2}]$ for all $g \in \mathcal{G}$) implies $(\ell_1^*, \ldots, \ell_r^*) \in \mathcal{R}^{\star\star}(a)$.

Here $\psi(n, z)$, $\mathcal{R}^\star(a)$ and $\mathcal{R}^{\star\star}(a)$ are as given in Section 2.1.3 and Definition 2.5, respectively. In particular, for any $n \in \mathbb{N}$ and $z > 0$,

$$\psi(n, z) = \begin{cases} 1 & \text{if } p \mid n \Rightarrow p \geq z \\ 0 & \text{otherwise.} \end{cases}$$

Specific subsets of $\mathcal{R}^\star(a)$ and $\mathcal{R}^{\star\star}(a)$ are computed later on, in Section 2.6.2.

2.6.1.2 Key strategy

Fix some a. Suppose $\exists\ \Psi_1, \ldots \Psi_{s_1} : [x, 6x] \cap \mathbb{N} \to \mathbb{R}$ and $\Theta_1, \ldots, \Theta_{s_2} : [x, 6x] \cap \mathbb{N} \to [0, \infty)$ such that $s_1, s_2 \leq 100$, such that each Ψ_i is good* or good**, and such that

$$1_{\mathbb{P}}(n) = \sum_{i=1}^{s_1} \Psi_i(n) + \sum_{i=1}^{s_2} \Theta_i(n) \qquad \text{and} \qquad \sum_{i=1}^{s_2} \sum_{n\in\mathcal{B}(y)} \Theta_i(n) \leq \frac{0.9999y}{x^b \log x}. \tag{2.210}$$

We now show that Proposition 2.3 then holds for a.

Proof: We split each $\Psi_i(n)$ up, restricting the range of p_j to $p_j \sim P_j$. If $\Psi_i(n)$ is good*, there are some $\beta \in [0.01, 0.15]$, $r \in \{0, \ldots, 5\}$ and $\mathcal{G}$ such that $\Psi_i(n)$ can be written as a sum of $O((\log x)^r)$ functions $\rho_{i,t} : [x, 6x] \cap \mathbb{N} \to \mathbb{R}$ of the form

$$\rho_{i,t}(n) = (-1)^r \sum_{\substack{n=p_1\ldots p_r m,\, p_j\sim x^{\ell_j^*} \\ x^\beta\leq p_r<\cdots<p_1<X \\ \mathcal{G}}} \psi(m, x^\beta).$$

Here $\ell_1^*, \ldots, \ell_r^*$ depend on t and satisfy $x^{\ell_j^*} = P_j \in \{2^d x^\beta : 0 \leq d \leq \log_2(\frac{X}{x^\beta})\}$. (For $d = 0$ we may take $p_j \in [P_j, 2P_j]$ rather than $p_j \in (P_j, 2P_j]$ to ensure that the case $p_j = x^\beta$ is covered. Any of our results concerning comparisons of $\sum_{p_j} \mathcal{S}(\mathcal{A}(y)_{p_1\ldots p_r}, z)$ and $\sum_{p_j} \mathcal{S}(\mathcal{B}(y)_{p_1\ldots p_r}, z)$ are still valid with this change.)

If $\sum_{j\in I_g} k_{(g,i)}[\ell_j^*, \ell_j^* + \log_x(2)] \cap [a_g, b_g] = \emptyset$ for some $g \in \mathcal{G}$ or $\ell_j^* < \ell_{j+1}^*$ for some $j \in \{1, \dots, r\}$, then $\rho_{i,t}(n) = 0$ for all n and we can discard $\rho_{i,t}$. Denote by U_i the set of t for which $\rho_{i,t}$ has $\sum_{j\in I_g} k_{(g,i)}[\ell_j^*, \ell_j^* + \log_x(2)] \subseteq [a_g, b_g]$ for all $g \in \mathcal{G}$ and for which $\ell_1^* + \log_x(2) < \log_x(X)$ and $\ell_1^* > \dots > \ell_r^*$. If r is even, denote by W_i the set of t for which $\rho_{i,t}$ has $(\sum_{j\in I_g} k_{(g,i)}[\ell_j^*, \ell_j^* + \log_x(2)] \not\subseteq [a_g, b_g]$ for some $g \in \mathcal{G}$ or $\ell_j^* = \ell_{j+1}^*$ for some $j \in \{1, \dots, r\}$ or $\ell_1^* + \log_x(2) \geq \log_x(X))$ and for which $\ell_1^* \geq \dots \geq \ell_r^*$. If r is odd, instead denote this set by V_i. If r is even, write $V_i = \emptyset$. If r is odd, write $W_i = \emptyset$. For $t \in V_i$, set

$$\widetilde{\rho}_{i,t}(n) = (-1) \sum_{\substack{n=p_1\dots p_r m \\ p_j \sim x^{\ell_j^*}}} \psi(m, x^\beta) \qquad \text{and} \qquad \widehat{\rho}_{i,t}(n) = \rho_{i,t}(n) - \widetilde{\rho}_{i,t}(n).$$

Observe that $\widehat{\rho}_{i,t}(n)$ counts the number of ways in which n can be written as $n = p_1 \dots p_r m$ with $p_j \sim x^{\ell_j^*}$ and ($p_j \leq p_{j+1}$ for some j, g not satisfied for some $g \in \mathcal{G}$ or $p_1 \geq X$). In particular, $\widehat{\rho}_{i,t}(n)$ is non-negative.

If $\Psi_i(n)$ is good**, we proceed similarly. There are some $r \in \{2, 4, 6\}$ and some conditions $\mathcal{G}$ such that $\Psi_i(n)$ can be written as a sum of $O((\log x)^r)$ functions $\rho_{i,t} : [x, 6x] \cap \mathbb{N} \to \mathbb{R}$ of the form

$$\rho_{i,t}(n) = \sum_{\substack{n=p_1\dots p_r m,\, p_j \sim x^{\ell_j^*} \\ p_r < \dots < p_1 \\ \mathcal{G}}} \psi(m, p_r).$$

We denote by U_i the set of t for which $\rho_{i,t}$ has $\sum_{j\in I_g} k_{(g,i)}[\ell_j^*, \ell_j^* + \log_x(2)] \subseteq [a_g, b_g]$ for all $g \in \mathcal{G}$ and for which $\ell_1^* > \dots > \ell_r^*$. We denote by W_i the set of t for which $\rho_{i,t}$ has $(\sum_{j\in I_g} k_{(g,i)}[\ell_j^*, \ell_j^* + \log_x(2)] \not\subseteq [a_g, b_g]$ for some $g \in \mathcal{G}$ or $\ell_j^* = \ell_{j+1}^*$ for some $j \in \{1, \dots, r\}$) and for which $\sum_{j\in I_g} k_{(g,i)}[\ell_j^*, \ell_j^* + \log_x(2)] \cap [a_g, b_g] \neq \emptyset$ and $\ell_1^* \geq \dots \geq \ell_r^*$. We also write $V_i = \emptyset$.

We have ensured that every $\rho_{i,t}(n)$, $t \in U_i$, and $\widetilde{\rho}_{i,t}(n)$, $t \in V_i$, can be written as

$$(-1)^r \sum_{\substack{n=p_1\dots p_r m \\ p_j \sim x^{\ell_j^*}}} \psi(m, x^\beta) \quad \text{for some } r \in \{0, \dots, 5\} \text{ and } (\ell_1^*, \dots, \ell_r^*, \beta) \in \mathcal{R}^*(a),$$

$$\text{or} \quad \sum_{\substack{n=p_1\dots p_r m \\ p_i \sim x^{\ell_i^*}}} \psi(m, p_r) \quad \text{for some } r \in \{2, 4, 6\} \text{ and } (\ell_1^*, \dots, \ell_r^*) \in \mathcal{R}^{**}(a).$$

Hence every $\rho_{i,t}$ with $t \in U_i$ and every $\widetilde{\rho}_{i,t}$ with $t \in V_i$ is of the form given in property (β) of Proposition 2.3. Furthermore, the prime indicator function $1_{\mathbb{P}}(n)$ has been decomposed as

$$1_{\mathbb{P}}(n) = \sum_{i=1}^{s_1}\sum_{t\in U_i}\rho_{i,t}(n)+\sum_{i=1}^{s_1}\sum_{t\in V_i}\widetilde{\rho}_{i,t}(n)+\sum_{i=1}^{s_1}\sum_{t\in V_i}\widehat{\rho}_{i,t}(n)+\sum_{i=1}^{s_1}\sum_{t\in W_i}\rho_{i,t}(n)+\sum_{i=1}^{s_2}\Theta_i(n).$$

But every $\widehat{\rho}_{i,t}$ with $t \in V_i$, every $\rho_{i,t}$ with $t \in W_i$ and every Θ_i is non-negative and so the function $\rho(n) = \sum_{i=1}^{s_1}\sum_{t\in U_i}\rho_{i,t}(n) + \sum_{i=1}^{s_2}\sum_{t\in V_i}\widetilde{\rho}_{i,t}(n)$ is a minorant of $1_{\mathbb{P}}(n)$ (property (α)) and decomposes into functions which satisfy property (β) of Proposition 2.3. Futhermore, we have simple estimates

$$\sum_{n\in\mathcal{B}(y)}\rho_{i,t}(n) \le \sum_{n\in\mathcal{B}(y)}\sum_{\substack{n=p_1\dots p_r m\\ p_j\sim x^{\ell_j^*}}}\psi(m,z) = \sum_{p_j\sim P_j} S(\mathcal{B}(y)_{p_1\dots p_r}, z) = O\left(\frac{y}{x^b(\log x)^{r+1}}\right),$$

where $P_j = x^{\ell_j^*}$ and $z \in \{x^\beta, p_r\}$. Similarly, $\sum_{n\in\mathcal{B}(y)}\widehat{\rho}_{i,t}(n) = O(\frac{y}{x^b(\log x)^{r+1}})$. But since $t \in V_i \cup W_i$ requires $\sum_{j\in I_g}[\ell_j^*, \ell_j^* + \log_x(2)] \cap [a_g, b_g] \neq \emptyset$ and $\sum_{j\in I_g}[\ell_j^*, \ell_j^* + \log_x(2)] \not\subseteq [a_g, b_g]$ for some $g \in \mathcal{G}$ or $\ell_j^* = \ell_{j+1}^*$ for some $j \in \{1, \dots, r\}$, we have $\#(\bigcup_{i=1}^r V_i \cup \bigcup_{i=1}^r W_i) = O((\log x)^{r-1})$. Summing up,

$$\sum_{n\in\mathcal{B}(y)}(1_{\mathbb{P}}(n) - \rho(n)) = \sum_{i=1}^{s_1}\sum_{t\in V_i}\widehat{\rho}_{i,t}(n) + \sum_{i=1}^{s_1}\sum_{t\in W_i}\rho_{i,t}(n) + \sum_{i=1}^{s_2}\Theta_i(n) \le \frac{0.99999y}{x^b\log x}.$$

This is property (γ). Hence to prove Proposition 2.3, it suffices to show that we can write $1_{\mathbb{P}}(n)$ as a sum $1_{\mathbb{P}}(n) = \sum_{i=1}^{s_1}\Psi_i(n) + \sum_{i=1}^{s_2}\Theta_i(n)$ with $s_1, s_2 \le 100$ and each $\Psi_i : [x, 6x] \cap \mathbb{N} \to \mathbb{R}$ good* or good** and $\Theta_i : [x, 6x] \cap \mathbb{N} \to [0, \infty)$ such that $\sum_{i=1}^{s}\sum_{n\in\mathcal{B}(y)}\Theta_i(n) \le \frac{0.9999y}{x^b\log x}$, just like described in (2.210). $\square$

(Note: On a few occasions, we will also run into functions of the form

$$\Psi(n) = -\sum_{\substack{n=p_1\dots p_r m\\ x^\beta\le p_r<\dots<p_1<X\\ \mathcal{G}}}\psi(m, p_r), \tag{2.211}$$

where $\mathcal{G}$ is a set of conditions (with $g \in \mathcal{G}$ of the form $\sum_{i\in I_g} k_{(g,i)}\log_x(p_i) \in [a_g, b_g]$) such that the inequalities ($0.5 + \varepsilon \ge \ell_1^* \ge \dots \ge \ell_r^* \ge 0.01 - \varepsilon$ and $\sum_{i\in I_g} k_{(g,i)}\ell_i^* \in [a_g - \frac{\varepsilon_1}{2}, b_g + \frac{\varepsilon_1}{2}]$ for all $g \in \mathcal{G}$) imply $(\ell_1^*, \dots, \ell_r^*, \beta) \in \mathcal{R}^\star(a)$ for $\beta = \ell_r^*$.

Restricting each p_j to $p_j \sim x^{\ell_j^*}$, the function $\Psi(n)$ can be decomposed into a sum $\sum_t \rho_t(n) + \rho_t^*(n)$ with

$$\rho_t(n) + \rho_t^*(n) = \sum_{\substack{n=p_1\dots p_r m,\, p_j\sim x^{\ell_j^*}\\ x^\beta\le p_r<\dots<p_1<X\\ \mathcal{G}}}(-1)\psi(m, x^{\ell_r^*}) + \sum_{\substack{n=p_1\dots p_r m,\, p_j\sim x^{\ell_j^*}\\ x^\beta\le p_r<\dots<p_1<X\\ \mathcal{G}}}(\psi(m, x^{\ell_r^*}) - \psi(m, p_r)),$$

where $\rho_t(n)$ denotes the first big sum and $\rho_t^*(n)$ denotes the second big sum on the right and where $\ell_1^*, \dots, \ell_r^*$ depend on t. Here $\rho_t^*(n)$ counts n which are divisible by some $p_j \sim x^{\ell_j^*}$ and have one additional factor in $[x^{\ell_r^*}, 2x^{\ell_r^*}]$, so that $\sum_{n \in \mathcal{B}(y)} \rho_t^*(n) = O(\frac{y}{x^b (\log x)^{r+2}})$. Since there are only $O((\log x)^r)$ choices for $\ell_1^*, \dots, \ell_r^*$, this error can easily be discarded. On the other hand, just like in the treatment of good* functions, the requirements placed on $\mathcal{G}$ ensure that $\sum_t \rho_t(n)$ can be written as a sum of functions with property (β) and an error term $\Theta(n)$ with $\sum_{n \in \mathcal{B}(y)} \Theta(n) = O(\frac{y}{x^b (\log x)^2})$. In short, the above discussion of decompositions of $1_{\mathbb{P}}(n)$ also applies to functions of the form described in (2.211). These will only be relevant on rare occasions and for simplicity's sake we also call them good*.)

2.6.2 Suitable ranges

To construct good* and good** functions, we need to know elements of $\mathcal{R}^*(a)$ and $\mathcal{R}^{**}(a)$. Recall here that $a \in [0.475 - \varepsilon, 0.77 - \varepsilon]$, with $\mathcal{A}(y) = [y, y + y/y^a]$. We remind ourselves of several definitions from earlier sections:

Let $\Xi^\star(\ell_1^*, \dots, \ell_r^*, \beta)$ and $\Xi^{\star\star}(\ell_1^*, \dots, \ell_r^*)$ be as given in Definition 2.4. Recall that $\{\ell_i\}_{i=1}^j \in \Xi^\star(\ell_1^*, \dots, \ell_r^*, \beta) \cup \Xi^{\star\star}(\ell_1^*, \dots, \ell_r^*)$ implies that $\ell_1, \dots, \ell_j \in [0,1]$, $\sum_{i=1}^j \ell_i = 1$ and $j \leq 10^{20}$ and that there exist disjoint subsets $X_1, \dots, X_r$ of $\{1, \dots, j\}$ with

$$\sum_{i \in X_s} \ell_i \in \left[\ell_s^* - \frac{\varepsilon}{10^{100}}, \ell_s^* + \frac{\varepsilon}{10^{100}}\right] \quad \text{for} \quad s \leq r.$$

If $\{\ell_i\}_{i=1}^j \in \Xi^\star(\ell_1^*, \dots, \ell_r^*, \beta)$, additionally $\ell_i \leq \beta + 10^{-100}\varepsilon$ for all but at most one $i \in \{1, \dots, j\} \setminus \bigcup_{i=1}^r X_r$.

Let $\chi_0(a)$, $\chi_1(a)$, $\chi_2(a)$ and $\chi_3(a)$ be as described in (2.9), (2.10), (2.11) and (2.12), respectively. Recall that we say that $\{\ell_i\}_{i=1}^j \in \Xi^\star(\ell_1^*, \dots, \ell_r^*, \beta) \cup \Xi^{\star\star}(\ell_1^*, \dots, \ell_r^*)$ satisfies one of the options (1), (2) or (3) for a given a if one of the following three conditions holds:

(1) There exists $k \in \{1, \dots, j\}$ with $\ell_k \geq \chi_0(a)$.

(2) There exists $I_1 \subseteq \{1, \dots, j\}$ with $\sum_{i \in I_1} \ell_i \in \chi_1(a)$.

(3) There exist $I_2, I_3 \subseteq \{1, \dots, j\}$ with $\sum_{i \in I_2} \ell_i \in \chi_2(a)$ and $\sum_{i \in I_3} \ell_i \in \chi_3(a)$.

Then recall that $\mathcal{R}^*(a)$ denotes the set of tuples $(\ell_1^*, \dots, \ell_r^*, \beta)$ which have $\beta \in [0.01, 0.15]$, $r \in \{0, \dots, 5\}$, $\ell_1^*, \dots, \ell_r^* \in [\beta, 0.5 + \varepsilon]$, $\ell_i^* \geq \ell_{i+1}^*$ and $\ell_1^* + \dots + \ell_r^* \leq 0.75$ and for which every $\{\ell_i\}_{i=1}^j \in \Xi^\star(\ell_1^*, \dots, \ell_r^*, \beta)$ satisfies one of the options (1), (2) or

(3). Further, $\mathcal{R}^{**}(a)$ denotes the set of tuples $(\ell_1^*, \ldots, \ell_r^*)$ which have $r \in \{2,4,6\}$, $\ell_1^*, \ldots, \ell_r^* \in [0.01-\varepsilon, 0.5+\varepsilon]$, $\ell_i^* \geq \ell_{i+1}^*$ and $\ell_1^* + \cdots + \ell_r^* \leq 0.99$ and for which every $\{\ell_i\}_{i=1}^j \in \Xi^{\star\star}(\ell_1^*, \ldots, \ell_r^*)$ satisfies one of the options (1), (2) or (3).

To find many good* and good** functions, we now wish to compute elements of $\mathcal{R}^*(a)$ and $\mathcal{R}^{**}(a)$. We begin by listing a few simple criteria which ensure that $\{\ell_i\}_{i=1}^j$ has some $I \subseteq \{1, \ldots, j\}$ such that $\sum_{i \in I} \ell_i$ is contained in a fixed interval $[a_1 - \frac{\varepsilon_1}{2}, a_2 + \frac{\varepsilon_1}{2}]$.

Lemma 2.29. *Let $r \in \{0, \ldots, 5\}$ and suppose $\ell_1^*, \ldots, \ell_r^*$ satisfy $0.01 \leq \ell_s^* \leq 0.5 + \varepsilon_1$ and $\ell_s^* \geq \ell_{s+1}^*$. Let $\Xi^\star(\ell_1^*, \ldots, \ell_r^*, \beta)$ and $\Xi^{\star\star}(\ell_1^*, \ldots, \ell_r^*)$ be as given in Definition 2.4.*

Let $\chi = [a_1 - \frac{\varepsilon_1}{2}, a_2 + \frac{\varepsilon_1}{2}] \subseteq [0,1]$. Let $\rho > 0$. Let $b_1 \in [0,1]$ and $b_2 \in [0,1]$.

Assume that one of the following conditions holds:

(A) *$\sum_{s=1}^r \ell_s^* \leq a_2$ and $1 - \rho \geq a_1$.*

(B) *For some $k \in \{2, \ldots, r\}$, $\sum_{s=1}^{k-1} \ell_s^* \leq a_2$ and $\ell_k^* \leq b_1$ and $1 - \rho - (r-k+1) b_1 \geq a_1$.*

(C) *$\sum_{s=2}^r \ell_s^* \leq a_2$ and $\ell_1^* \leq b_2$ and $1 - \rho - b_2 \geq a_1$.*

(D) *$a_1 \geq 0.5 + \varepsilon_1$, $\sum_{s=1}^r \ell_s^* \leq 1 - \rho - (2a_1 - 1)$ and ($\sum_{s=1}^k \ell_s^* \in [1 - a_2 - (2a_1 - 1), a_2]$ for some k).*

(E) *There exists some $S \subseteq \{1, \ldots, r\}$ with $\sum_{s \in S} \ell_s^* \in [a_1, a_2]$.*

For conditions (A), (B), (C) and (D), additionally assume that $a_2 - a_1 \geq \beta$. (This assumption is not required when using condition (E).)

Then if $\{\ell_i\}_{i=1}^j \in \Xi^\star(\ell_1^*, \ldots, \ell_r^*, \beta)$ and $\ell_i \leq \rho \;\forall\; i \in \{1, \ldots, j\}$, $\exists\; I_\chi \subseteq \{1, \ldots, j\}$ with

$$\sum_{i \in I_\chi} \ell_i \in \chi. \tag{2.212}$$

Alternatively, assume some $S \subseteq \{1, \ldots, r\}$ satisfies $\sum_{s \in S} \ell_s^* \in [a_1, a_2]$ and suppose $\{\ell_i\}_{i=1}^j \in \Xi^{\star\star}(\ell_1^*, \ldots, \ell_r^*)$. Then there exists some $I_\chi \subseteq \{1, \ldots, j\}$ with $\sum_{i \in I_\chi} \ell_i \in \chi$.

Proof. Let $\{\ell_i\}_{i=1}^j \in \Xi^\star(\ell_1^*, \ldots, \ell_r^*, \beta)$ and let $X_1, \ldots, X_r \subseteq \{1, \ldots, j\}$ be disjoint with

$$\sum_{i \in X_s} \ell_i \in \left[\ell_s^* - \frac{\varepsilon}{10^{100}}, \ell_s^* + \frac{\varepsilon}{10^{100}}\right]$$

for $s \leq r$ and $\ell_i \leq \beta + 10^{-100}\varepsilon$ for all but at most one $i \in \{1, \ldots, j\} \setminus \bigcup_{i=1}^r X_r$. Assume that $\ell_i \leq \rho$ for every $i \in \{1, \ldots, j\}$.

Then there exists $Y_1 \subseteq \{1, \ldots, j\} \setminus \bigcup_{i=1}^{r} X_r$ with

$$\sum_{s=1}^{r} \sum_{i \in X_s} \ell_i + \sum_{i \in Y_1} \ell_i \geq 1 - \rho \tag{2.213}$$

and $\ell_i \leq \beta + 10^{-100}\varepsilon$ for every $i \in Y_1$. With condition (A) we have $\sum_{s=1}^{r} \ell_s^* \leq a_2$, while $\sum_{s=1}^{r} \sum_{i \in X_s} \ell_i + \sum_{i \in Y_1} \ell_i \geq 1 - \rho \geq a_1$. Recall also that $a_2 - a_1 \geq \beta$. Thus, starting with $\sum_{s=1}^{r} \sum_{i \in X_s} \ell_i \leq a_2 + \varepsilon_1/2$ and adding elements of Y_1, we must eventually hit interval $[a_1 - \varepsilon_1/2, a_2 + \varepsilon_1/2]$. There exists $Y_2 \subseteq Y_1$ with

$$\sum_{s=1}^{r} \sum_{i \in X_s} \ell_i + \sum_{i \in Y_2} \ell_i \in \left[a_1 - \frac{\varepsilon_1}{2}, a_2 + \frac{\varepsilon_1}{2}\right].$$

Taking $I_\chi = \bigcup_{s=1}^{r} X_s \cup Y_2$, (2.212) is satisfied for condition (A).

With condition (B) we have $\ell_k^* \leq b_1$ and since $\ell_s^* \geq \ell_{s+1}^*$, (2.213) gives

$$\sum_{s=1}^{k-1} \sum_{i \in X_s} \ell_i + \sum_{i \in Y_1} \ell_i \geq 1 - \rho - (r - k + 1)b_1 - \varepsilon_1/2 \geq a_1 - \varepsilon_1/2.$$

On the other hand, we also have $\sum_{s=1}^{k-1} \sum_{i \in X_s} \ell_i \leq \sum_{s=1}^{k-1} \ell_s^* + \varepsilon_1/2 \leq a_2 + \varepsilon_1/2$. Since $\ell_i \leq \beta + 10^{-100}\varepsilon \leq a_2 - a_1 + 10^{-100}\varepsilon$ for $i \in Y_1$, there exists some $Y_2 \subseteq Y_1$ such that $I_\chi = \bigcup_{s=1}^{k-1} X_s \cup Y_2$ satisfies (2.212).

With condition (C) we have $\ell_1^* \leq b_2$, so that (2.213) gives

$$\sum_{s=2}^{r} \sum_{i \in X_s} \ell_i + \sum_{i \in Y_1} \ell_i \geq 1 - \rho - b_2 - \varepsilon_1/2 \geq a_1 - \varepsilon_1/2,$$

while $\sum_{s=2}^{r} \ell_s^* \leq a_2$, and hence $\exists Y_2 \subseteq Y_1$ such that $I_\chi = \bigcup_{s=2}^{r} X_s \cup Y_2$ satisfies (2.212).

With condition (D) we have $\sum_{s=1}^{k} \ell_s^* \in [1 - a_2 - (2a_1 - 1), a_2]$. If $\sum_{s=1}^{k} \ell_s^* \in [a_1, a_2]$, we simply take $I_\chi = \bigcup_{s=1}^{k} X_s$ to satisfy (2.212). If $\sum_{s=1}^{k} \ell_s^* \in [1 - a_2, 1 - a_1]$, we recall that $\{\ell_i\}_{i=1}^{j} \in \Xi^\star(\ell_1^*, \ldots, \ell_r^*, \beta)$ implies $\sum_{i=1}^{j} \ell_i = 1$ and take $I_\chi = \{1, \ldots, j\} \setminus (\bigcup_{s=1}^{k} X_s)$ to satisfy (2.212). This leaves the cases $\sum_{s=1}^{k} \ell_s^* \in [1 - a_1, a_1]$ and $\sum_{s=1}^{k} \ell_s^* \in [1 - a_2 - (2a_1 - 1), 1 - a_2]$. But condition (D) and (2.213) give

$$\sum_{i \in Y_1} \ell_i \geq 1 - \rho - \sum_{s=1}^{r} \sum_{i \in X_s} \ell_i \geq 2a_1 - 1 - \varepsilon_1/20.$$

So if $\sum_{s=1}^{k} \ell_s^* \in [1 - a_1, a_1]$, we have $\sum_{s=1}^{k} \sum_{i \in X_s} \ell_i + \sum_{i \in Y_1} \ell_i \geq (1 - a_1) + (2a_1 - 1) - \varepsilon_1/10 = a_1 - \varepsilon_1/10$ and if $\sum_{s=1}^{k} \ell_s^* \in [1 - a_2 - (2a_1 - 1), 1 - a_2]$, then $\sum_{s=1}^{k} \sum_{i \in X_s} \ell_i +$

$\sum_{i\in Y_1} \ell_i \geq 1 - a_2 - \varepsilon_1/10$. Since also $\ell_i - 10^{-100}\varepsilon \leq \beta \leq a_2 - a_1 = (1-a_1) - (1-a_2)$ for $i \in Y_1$, there exists $Y_2 \subseteq Y_1$ with

$$\sum_{s=1}^{k} \sum_{i\in X_s} \ell_i + \sum_{i\in Y_2} \ell_i \in [1 - a_2 - \varepsilon_1/10, 1 - a_1] \cup [a_1 - \varepsilon_1/10, a_2]$$

and one of $I_\chi = \bigcup_{s=1}^{k} X_s \cup Y_2$ or $I_\chi = \{1, \ldots, j\} \setminus (\bigcup_{s=1}^{k} X_s \cup Y_2)$ satisfies (2.212).

On the other hand, if $\{\ell_i\}_{i=1}^{j} \in \Xi^\star(\ell_1^*, \ldots, \ell_r^*, \beta)$ or $\{\ell_i\}_{i=1}^{j} \in \Xi^{\star\star}(\ell_1^*, \ldots, \ell_r^*)$ and there exists some $S \subseteq \{1, \ldots, r\}$ with $\sum_{s\in S} \ell_s^* \in [a_1, a_2]$, the definitions of $\Xi^\star$ and $\Xi^{\star\star}$ tell us that there exist some disjoint subsets $X_1, \ldots, X_r$ of $\{1, \ldots, j\}$ with

$$\sum_{i\in X_s} \ell_i \in \left[\ell_s^* - \frac{\varepsilon}{10^{100}}, \, \ell_s^* + \frac{\varepsilon}{10^{100}}\right]$$

and (2.212) is satisfied by taking $I_\chi = \bigcup_{s\in S} X_s$. This covers the final case of the proof, the only case for which assumption $\beta \leq a_2 - a_1$ is not needed. □

We are now ready to list elements of $\mathcal{R}^\star(a)$ and $\mathcal{R}^{\star\star}(a)$. We split the argument up according to the size of $\tau = x^a$.

Corollary 2.30. *Let $\varepsilon > 0$ and $a \in [0.475 - \varepsilon, 0.53]$.*

Let $r \in \{0, \ldots, 5\}$. Suppose $\ell_1^*, \ldots, \ell_r^*$ satisfy $\beta \leq \ell_i^* \leq 0.5 + \varepsilon$, $\ell_i^* \geq \ell_{i+1}^*$ and $\ell_1^* + \cdots + \ell_r^* \leq 0.75$. Assume that one of the following conditions (i) – (iii) *holds:*

(i) $(r = 0$ *or* $r = 1)$ *and* $\beta = 0.07$.

(ii) $(r = 2$ *or* $r = 3)$ *and* $\ell_2^* \leq \frac{(0.71 - \ell_1^*)}{2} + \frac{\varepsilon_1}{4}$ *and* $\beta = 0.07$.

(iii) $(r = 4$ *or* $r = 5)$ *and* $\ell_4^* \leq \frac{(0.71 - \ell_1^* - \ell_2^* - \ell_3^*)}{2} + \frac{\varepsilon_1}{4}$ *and* $\beta = 0.07$.

Then $(\ell_1^*, \ldots, \ell_r^*, \beta)$ is contained in $\mathcal{R}^\star(a)$.

Alternatively, let $r \in \{2, 4, 6\}$, $0.07 - \frac{\varepsilon_1}{2} \leq \ell_i^* \leq 0.53$, $\ell_i^\star \geq \ell_{i+1}^\star$ and $\ell_1^\star + \cdots + \ell_r^\star \leq 0.99$. Assume that one of the following conditions (a) – (f) *holds:*

(a) $r = 2$ *and* $\ell_1^* \in [0.29 - \frac{\varepsilon_1}{2}, 0.36 + \frac{\varepsilon_1}{2}]$.

(b) $r = 2$ *and* $\ell_2^* \in [0.64 - \ell_1^* - \frac{\varepsilon_1}{2}, 0.71 - \ell_1^* + \frac{\varepsilon_1}{2}]$.

(c) $r = 4$ *and* $\ell_2^* \leq \frac{(0.71 - \ell_1^*)}{2} + \frac{\varepsilon_1}{4}$ *and* $\ell_3^* \geq 0.64 - \ell_1^* - \ell_2^* - \frac{\varepsilon_1}{2}$.

(d) $r = 4$ *and* $\ell_1^* \in [0.22 - \frac{\varepsilon_1}{2}, 0.29 + \frac{\varepsilon_1}{2}]$ *and* $\ell_4^* \leq 0.36 - \ell_1^* + \frac{\varepsilon_1}{2}$.

(e) $r = 6$ *and* $\ell_i^* + \ell_j^* \in [0.29 - \frac{\varepsilon_1}{2}, 0.36 + \frac{\varepsilon_1}{2}]$ *for some* $i \neq j$.

(f) $r = 6$ *and* $\ell_1^* + \cdots + \ell_5^* \le 0.71 + \frac{\varepsilon_1}{2}$, $\ell_1^* \ge 0.18 - \frac{\varepsilon_1}{2}$ *and* $\ell_3^* \in [0.145 - \frac{\varepsilon_1}{2}, 0.29 + \frac{\varepsilon_1}{2}]$.

Then $(\ell_1^*, \ldots, \ell_r^*)$ *is contained in* $\mathcal{R}^{\star\star}(a)$.

Proof. Suppose $a \le 0.53$. Recall $\chi_0(a) = 0.29 - \varepsilon_1$ and $\chi_1(a) = [0.29 - \varepsilon_1, 0.36 + \varepsilon_1] \cup [0.64 - \varepsilon_1, 0.71 + \varepsilon_1]$.

Case 1: Treatment of $\Xi^\star$. Let $\{\ell_i\}_{i=1}^j \in \Xi^\star(\ell_1^*, \ldots, \ell_r^*, 0.07)$. If there exists $k \in \{1, \ldots, j\}$ with $\ell_k \ge \chi_0(a)$, then $\{\ell_i\}_{i=1}^j$ satisfies option (1). So we may assume that $\ell_i < \chi_0(a)$ for each i and take $\rho = \chi_0(a)$ in Lemma 2.29.

Note that $\sum_{s=1}^r \ell_s^* \le 0.71 + \frac{\varepsilon_1}{2}$ in each of the conditions (i), (ii) and (iii). We also have $1 - (0.29 - \varepsilon_1) = 0.71 + \varepsilon_1 \ge 0.64$ and $(0.71 + \frac{\varepsilon_1}{2}) - (0.64 + \frac{\varepsilon_1}{2}) = 0.07 + \varepsilon_1$. Hence we may apply Lemma 2.29 with $\beta = 0.07$, $a_1 = 0.64 - \frac{\varepsilon_1}{2}$ and $a_2 = 0.71 + \frac{\varepsilon_1}{2}$ (so that $\chi = [0.64 - \varepsilon_1, 0.71 + \varepsilon_1]$) and $\rho = 0.29 - \varepsilon_1$. Condition (A) is satisfied and there is some $I_2 \subseteq \{1, \ldots, j\}$ with $\sum_{i \in I_2} \ell_i \in [0.64 - \varepsilon_1, 0.71 + \varepsilon_1]$. But then $\{\ell_i\}_{i=1}^j$ satisfies option (2).

Case 2: Treatment of $\Xi^{\star\star}$. Now let $\{\ell_i\}_{i=1}^j \in \Xi^{\star\star}(\ell_1^*, \ldots, \ell_r^*)$. Option (a) gives $\ell_1^* \in [0.29 - \frac{\varepsilon_1}{2}, 0.36 + \frac{\varepsilon_1}{2}]$, (b) gives $\ell_1^* + \ell_2^* \in [0.64 - \frac{\varepsilon_1}{2}, 0.71 + \frac{\varepsilon_1}{2}]$, (c) gives $\ell_1^* + \ell_2^* + \ell_3^* \in [0.64 - \frac{\varepsilon_1}{2}, 0.71 + \frac{\varepsilon_1}{2}]$, (d) gives $\ell_1^* + \ell_4^* \in [0.29 - \frac{\varepsilon_1}{2}, 0.36 + \frac{\varepsilon_1}{2}]$, (e) gives $\ell_i^* + \ell_j^* \in [0.29 - \frac{\varepsilon_1}{2}, 0.36 + \frac{\varepsilon_1}{2}]$ and (f) gives $\ell_1^* + \cdots + \ell_5^* \in [0.64 - \frac{\varepsilon_1}{2}, 0.71 + \frac{\varepsilon_1}{2}]$. In every case there is some $S \subseteq \{1, \ldots, r\}$ with $\sum_{s \in S} \ell_s^* \in [0.29 - \frac{\varepsilon_1}{2}, 0.36 + \frac{\varepsilon_1}{2}] \cup [0.64 - \frac{\varepsilon_1}{2}, 0.71 + \frac{\varepsilon_1}{2}]$. Applying Lemma 2.29 with $\chi = [0.29 - \varepsilon_1, 0.36 + \varepsilon_1]$ or $\chi = [0.64 - \varepsilon_1, 0.71 + \varepsilon_1]$, we find that $\{\ell_i\}_{i=1}^j$ satisfies option (2). □

Corollary 2.31. *Let* $\varepsilon > 0$ *and* $a \in (0.53, 0.545]$.

Let $r \in \{0, \ldots, 5\}$. *Suppose* $\ell_1^*, \ldots, \ell_r^*$ *satisfy* $\beta \le \ell_i^* \le 0.5 + \varepsilon$, $\ell_i^* \ge \ell_{i+1}^*$ *and* $\ell_1^* + \cdots + \ell_r^* \le 0.75$. *Assume that one of the following conditions* (i) – (iv) *holds:*

(i) $r \in \{0, 1\}$ *and* $\beta = 0.08$.

(ii) $r \in \{2, 3\}$, $\ell_1^* \ge 0.474 + \frac{\varepsilon_1}{2}$, $\ell_2^* \le \min\{0.595 - \ell_1^* + \frac{\varepsilon_1}{2}, \frac{(0.715 - \ell_1^*)}{2} + \frac{\varepsilon_1}{4}\}$ *and* $\beta \le 0.09$.

(iii) $r \in \{2, 3\}$ *and* $\ell_1^* \le 0.427 + \frac{\varepsilon_1}{2}$, $\ell_2^* \le \frac{(0.655 - \ell_1^*)}{2} + \frac{\varepsilon_1}{4}$ *and* $\beta = 0.08$.

(iv) $r \in \{4, 5\}$, $\ell_1^* \le 0.285 + \frac{\varepsilon_1}{2}$, $\ell_2^* \le \frac{(0.655 - \ell_1^*)}{2} + \frac{\varepsilon_1}{4}$, $\ell_4^* \le \frac{(0.655 - \sum_{i=1}^3 \ell_3^*)}{2} + \frac{\varepsilon_1}{4}$ *and* $\beta = 0.08$.

Then $(\ell_1^*, \ldots, \ell_r^*, \beta)$ *is contained in* $\mathcal{R}^\star(a)$.

Alternatively, let $r = 2$, $0.08 - \frac{\varepsilon_1}{2} \le \ell_i^* \le 0.53$, $\ell_i^* \ge \ell_{i+1}^*$ *and* $\ell_1^\star + \cdots + \ell_r^\star \le 0.99$. *Assume that one of the following conditions* (a) – (c) *holds:*

(a) $r = 2$ *and* $\ell_1^* \in [0.315 - \frac{\varepsilon_1}{2}, 0.345 + \frac{\varepsilon_1}{2}] \cup [0.427 - \frac{\varepsilon_1}{2}, 0.474 + \frac{\varepsilon_1}{2}]$.

(b) $r = 2$ *and* $\ell_1^* + \ell_2^* \in [0.427 - \frac{\varepsilon_1}{2}, 0.474 + \frac{\varepsilon_1}{2}] \cup [0.526 - \frac{\varepsilon_1}{2}, 0.573 + \frac{\varepsilon_1}{2}] \cup [0.655 - \frac{\varepsilon_1}{2}, 0.685 + \frac{\varepsilon_1}{2}]$.

(c) $r = 2$ *and* $\ell_1^* \in [0.285 - \frac{\varepsilon_1}{2}, 0.375 + \frac{\varepsilon_1}{2}]$, $\ell_1^* + \ell_2^* \in [0.405 - \frac{\varepsilon_1}{2}, 0.485 + \frac{\varepsilon_1}{2}] \cup [0.515 - \frac{\varepsilon_1}{2}, 0.595 + \frac{\varepsilon_1}{2}]$.

Then $(\ell_1^*, \ldots, \ell_r^*)$ *is contained in* $\mathcal{R}^{\star\star}(a)$.

Proof. Suppose $a \in (0.53, 0.545]$. Recall $\chi_0(a) = 0.315 - \varepsilon_1$, $\chi_1(a) = [0.315 - \varepsilon_1, 0.345 + \varepsilon_1] \cup [0.427 - \varepsilon_1, 0.474 + \varepsilon_1] \cup [0.526 - \varepsilon_1, 0.573 + \varepsilon_1] \cup [0.655 - \varepsilon_1, 0.685 + \varepsilon_1]$, $\chi_2(a) = [0.405 - \varepsilon_1, 0.485 + \varepsilon_1] \cup [0.515 - \varepsilon_1, 0.595 + \varepsilon_1]$ and $\chi_3(a) = [0.285 - \varepsilon_1, 0.375 + \varepsilon_1] \cup [0.625 - \varepsilon_1, 0.715 + \varepsilon_1]$.

Case 1: Treatment of $\Xi^\star$. Let $\{\ell_i\}_{i=1}^j \in \Xi^\star(\ell_1^*, \ldots, \ell_r^*, \beta)$. If there exists $k \in \{1, \ldots, j\}$ with $\ell_k \geq \chi_0(a)$, then $\{\ell_i\}_{i=1}^j$ satisfies option (1). So we may assume that $\ell_i < \chi_0(a)$ for each i and take $\rho = \chi_0(a)$ in Lemma 2.29.

With each of the conditions (i), (ii), (iii) and (iv) we have $\sum_{s=1}^r \ell_s^* \leq 0.715 + \frac{\varepsilon_1}{2}$. Also, $1 - 0.315 = 0.685 \geq 0.625$, while $0.715 - 0.625 = 0.09$. Hence we may apply Lemma 2.29 with $\beta \leq 0.09$, $\chi = [0.625 - \varepsilon_1, 0.715 + \varepsilon_1]$ and $\rho = 0.315 - \varepsilon_1$. Condition (A) is satisfied and there is some $I_3 \subseteq \{1, \ldots, j\}$ with $\sum_{i \in I_3} \ell_i \in \chi_3(a)$.

On the other hand, with condition (i) we also have $\sum_{s=1}^r \ell_s^* \leq 0.595$, while $1 - 0.315 = 0.685 \geq 0.515$ and $0.595 - 0.525 = 0.08$. We may apply Lemma 2.29 with $\beta = 0.08$, $\chi = [0.515 - \varepsilon_1, 0.595 + \varepsilon_1]$ and $\rho = 0.315$. Condition (A) is satisfied and there is some $I_2 \subseteq \{1, \ldots, j\}$ with $\sum_{i \in I_2} \ell_i \in \chi_2(a)$. With condition (ii) we instead have $\ell_1^* + \ell_2^* \in [0.554 - \frac{\varepsilon_1}{2}, 0.595 + \frac{\varepsilon_1}{2}]$ and by condition (E) of Lemma 2.29 there exists $I_2 \subseteq \{1, \ldots, j\}$ with $\sum_{i \in I_2} \ell_i \in [0.515 - \varepsilon_1, 0.595 + \varepsilon_1] \subseteq \chi_2(a)$. (Note that $\beta \leq 0.09$ is allowed in this case.) Finally, with condition (iii) or (iv) we have $\sum_{s=1}^r \ell_s^* \leq 0.655 + \frac{\varepsilon_1}{2} = 1 - 0.315 - (2 \cdot 0.515 - 1) + \frac{\varepsilon_1}{2}$, while $\ell_1^* + \ell_2^* \leq \ell_1^* + (0.655 - \ell_1^*)/2 + \frac{\varepsilon_1}{2} \leq 0.595 + \frac{\varepsilon_1}{2}$. If $\ell_1^* + \ell_2^* < 0.375$, we must have $\ell_i^* < 0.19$ for $i \geq 2$ and so either $\ell_1^* + \cdots + \ell_r^* \leq 0.375$ or there exists k with $\ell_1^* + \cdots + \ell_k^* \in [0.405 - 0.03, 0.595]$. In the first case we use condition (A) and in the second case we use condition (D) of Lemma 2.29, taking $\beta = 0.08$, $\chi = [0.515 - \varepsilon_1, 0.595 + \varepsilon_1]$ and $\rho = 0.315 - \varepsilon_1$, to deduce that there exists $I_2 \subseteq \{1, \ldots, j\}$ with $\sum_{i \in I_2} \ell_i \in \chi_2(a)$. But then $\{\ell_i\}_{i=1}^j$ satisfies option (3).

Case 2: Treatment of $\Xi^{\star\star}$. Now let $\{\ell_i\}_{i=1}^j \in \Xi^{\star\star}(\ell_1^*, \ldots, \ell_r^*)$. With (a) we have $\ell_1^* \in [0.315 - \frac{\varepsilon_1}{2}, 0.345 + \frac{\varepsilon_1}{2}] \cup [0.427 - \frac{\varepsilon_1}{2}, 0.474 + \frac{\varepsilon_1}{2}] \subseteq \chi_1(a)$, with (b) we have $\ell_1^* + \ell_2^* \in [0.427 - \frac{\varepsilon_1}{2}, 0.474 + \frac{\varepsilon_1}{2}] \cup [0.526 - \frac{\varepsilon_1}{2}, 0.573 + \frac{\varepsilon_1}{2}] \cup [0.655 - \frac{\varepsilon_1}{2}, 0.685 + \frac{\varepsilon_1}{2}] \subseteq \chi_1(a)$ and with (c) we have $\ell_1^* \in [0.285 - \frac{\varepsilon_1}{2}, 0.375 + \frac{\varepsilon_1}{2}] \subseteq \chi_3(a)$ and $\ell_1^* + \ell_2^* \in [0.405 -$

$\frac{\varepsilon_1}{2}, 0.485 + \frac{\varepsilon_1}{2}] \cup [0.515 - \frac{\varepsilon_1}{2}, 0.595 + \frac{\varepsilon_1}{2}] \subseteq \chi_2(a)$. Lemma 2.29 tells us that $\{\ell_i\}_{i=1}^j$ satisfies option (2) or (3). $\square$

Corollary 2.32. *Let $\varepsilon > 0$ and $a \in (0.545, 0.57]$.*

Let $r \in \{0, \ldots, 5\}$. Suppose $\ell_1^*, \ldots, \ell_r^*$ satisfy $\beta \leq \ell_i^* \leq 0.5 + \varepsilon$, $\ell_i^* \geq \ell_{i+1}^*$ and $\ell_1^* + \cdots + \ell_r^* \leq 0.75$. Assume that one of the following conditions (i) – (iv) *holds:*

(i) *$(r = 0$ or $r = 1)$ and $\beta = 0.075$.*

(ii) *$(r = 2$ or $r = 3)$ and $\ell_2^* \in [0.475 - \ell_1^* - \frac{\varepsilon_1}{2}, \min\{0.525 - \ell_1^*, 0.14\} + \frac{\varepsilon_1}{2}]$ and $\beta = 0.075$.*

(iii) *$(r = 2$ or $r = 3)$ and $\ell_2^* \leq \frac{(0.6 - \ell_1^*)}{2} + \frac{\varepsilon_1}{4}$ and $\beta = 0.075$.*

(iv) *$(r = 4$ or $r = 5)$ and $\ell_2^* \leq 0.4 - \ell_1^* + \frac{\varepsilon_1}{2}$ and $\ell_4^* \leq \frac{(0.615 - \sum_{i=1}^3 \ell_3^*)}{2} + \frac{\varepsilon_1}{4}$ and $\beta = 0.075$.*

Then $(\ell_1^*, \ldots, \ell_r^*, \beta)$ is contained in $\mathcal{R}^\star(a)$.

Alternatively, let $r \in \{2, 4\}$, $0.075 - \frac{\varepsilon_1}{2} \leq \ell_i^\star \leq 0.53$, $\ell_i^\star \geq \ell_{i+1}^\star$ and $\ell_1^\star + \cdots + \ell_r^\star \leq 0.99$. Assume that one of the following conditions (a) – (c) *holds:*

(a) *$r = 2$ and $\ell_1^* \in [0.4 - \frac{\varepsilon_1}{2}, 0.475 + \frac{\varepsilon_1}{2}]$.*

(b) *$r = 2$ and $\ell_2^* \in [0.4 - \ell_1^* - \frac{\varepsilon_1}{2}, 0.475 - \ell_1^* + \frac{\varepsilon_1}{2}] \cup [0.525 - \ell_1^* - \frac{\varepsilon_1}{2}, 0.6 - \ell_1^* + \frac{\varepsilon_1}{2}]$.*

(c) *$r = 4$ and $\ell_4^* \in [0.4 - \ell_1^* - \ell_2^* - \frac{\varepsilon_1}{2}, 0.475 - \ell_1^* - \ell_2^* + \frac{\varepsilon_1}{2}]$.*

Then $(\ell_1^*, \ldots, \ell_r^*)$ is contained in $\mathcal{R}^{\star\star}(a)$.

Proof. Suppose $a \in (0.545, 0.57]$. Recall $\chi_0(a) = 0.335 - \varepsilon_1$ and $\chi_1(a) = [0.4 - \varepsilon_1, 0.475 + \varepsilon_1] \cup [0.525 - \varepsilon_1, 0.6 + \varepsilon_1]$.

Case 1: Treatment of $\Xi^\star$. Let $\{\ell_i\}_{i=1}^j \in \Xi^\star(\ell_1^*, \ldots, \ell_r^*, 0.075)$. If there exists $k \in \{1, \ldots, j\}$ with $\ell_k \geq \chi_0(a)$, then $\{\ell_i\}_{i=1}^j$ satisfies option (1). So we may assume that $\ell_i < \chi_0(a)$ for each i and take $\rho = \chi_0(a)$ in Lemma 2.29.

With conditions (i) and (iii) we have $\ell_1^* + \cdots + \ell_r^* \leq 0.6 + \frac{\varepsilon_1}{2}$. We apply Lemma 2.29 with $\beta = 0.075$, $\chi = [0.525 - \varepsilon_1, 0.6 + \varepsilon_1] \subseteq \chi_1(a)$, $\rho = 0.335 - \varepsilon_1$ and condition (A) to deduce that option (2) of Proposition 2.3 is satisfied. With condition (ii) we have $\ell_1^* + \ell_2^* \leq 0.525 + \frac{\varepsilon_1}{2}$, while $\ell_3^* \leq \ell_2^* \leq 0.14 + \frac{\varepsilon_1}{2}$ and $r \in \{2, 3\}$. Note $1 - 0.335 - (3 - 3 + 1)0.14 = 0.525$. For $r = 2$ we again use condition (A). For $r = 3$ we apply Lemma 2.29 with $\beta = 0.075 - \varepsilon_1$, $\chi = [0.525 - \varepsilon_1, 0.6 + \varepsilon_1]$, $\rho = 0.335 - \varepsilon_1$ and $k = 3$ in condition (B) to deduce that option (2) is satisfied. Finally, in condition (iv) we have $\ell_1^* + \cdots + \ell_r^* \leq 0.615 + \frac{\varepsilon_1}{2} = 1 - 0.335 - (2 \cdot 0.525 - 1) + \frac{\varepsilon_1}{2}$. We also have $\ell_1^* + \ell_2^* \leq 0.4 + \frac{\varepsilon_1}{2}$ and $\ell_2^* \leq 0.2 + \frac{\varepsilon_1}{4}$. If $\ell_1^* + \cdots + \ell_r^* \leq 0.4$, we again use condition (A) of Lemma 2.29. Otherwise there exists k with $\ell_1^* + \cdots + \ell_k^* \in [0.4 - \frac{\varepsilon_1}{2}, 0.6 + \frac{\varepsilon_1}{2}]$. So we may apply Lemma 2.29 with $\beta = 0.075$, $\chi = [0.525 - \varepsilon_1, 0.6 + \varepsilon_1]$, $\rho = 0.335 - \varepsilon_1$ and condition (D) to obtain that $\{\ell_i\}_{i=1}^j$ satisfies option (2).

Case 2: Treatment of $\Xi^{\star\star}$. The cases (a) – (c) can be treated just like the corresponding cases of Corollary 2.30 and Corollary 2.31. The values of ℓ_i^* have been carefully restricted to ensure that there exists a set $S_1 \subseteq \{1, \ldots, r\}$ with $\sum_{i \in S_1} \ell_i^* + \delta \in \chi_1(a)$ for all $\delta \in [-\frac{\varepsilon_1}{2}, \frac{\varepsilon_1}{2}]$ or two sets $S_2, S_3 \subseteq \{1, \ldots, r\}$ with $\sum_{i \in S_j} \ell_i^* + \delta \in \chi_j(a)$ for $\delta \in [-\frac{\varepsilon_1}{2}, \frac{\varepsilon_1}{2}]$ and $j \in \{2, 3\}$. For a given $\{\ell_i\}_{i=1}^j \in \Xi^{\star\star}(\ell_1^*, \ldots, \ell_r^*)$, Lemma 2.29 again tells us that $\{\ell_i\}_{i=1}^j$ satisfies option (2) or (3). $\square$

Corollary 2.33. *Let $\varepsilon > 0$ and $a \in (0.57, 0.59]$.*

Let $r \in \{0, \ldots, 5\}$. Suppose $\ell_1^*, \ldots, \ell_r^*$ satisfy $\beta \leq \ell_i^* \leq 0.5 + \varepsilon$, $\ell_i^* \geq \ell_{i+1}^*$ and $\ell_1^* + \cdots + \ell_r^* \leq 0.75$. Assume that one of the following conditions (i) – (viii) *holds:*

(i) $r \in \{0, 1\})$ *and* $\beta = 0.075$.

(ii) $r \in \{2, 3\}$ *and* $\ell_1^* \geq 0.455 - \frac{\varepsilon_1}{2}$, $\ell_2^* \leq \min\{0.58 - \ell_1^* + \frac{\varepsilon_1}{2}, \frac{(0.685 - \ell_1^*)}{2} + \frac{\varepsilon_1}{4}\}$ *and* $\beta = 0.075$.

(iii) $r \in \{2, 3\}$ *and* $\ell_1^* \in [0.42 - \frac{\varepsilon_1}{2}, 0.455 + \frac{\varepsilon_1}{2}]$, $\ell_2^* \leq \frac{(0.685 - \ell_1^*)}{2} + \frac{\varepsilon_1}{4}$, $\beta \leq 0.105$.

(iv) $r \in \{2, 3\}$, $\ell_1^* \in [0.315 - \frac{\varepsilon_1}{2}, 0.38 + \frac{\varepsilon_1}{2}]$, $\ell_2^* \leq \max\{0.145 + \frac{\varepsilon_1}{2}, \frac{(0.62 - \ell_1^*)}{2} + \frac{\varepsilon_1}{4}\}$ *and* $\beta = 0.075$.

(v) $r \in \{2, 3\}$ *and* $\ell_1^* \in [0.29 - \frac{\varepsilon_1}{2}, 0.315 + \frac{\varepsilon_1}{2}]$, $\ell_2^* \leq \frac{(0.62 - \ell_1^*)}{2} + \frac{\varepsilon_1}{4}$ *and* $\beta = 0.075$.

(vi) $r \in \{2, 3\}$ *and* $\ell_1^* \leq 0.29 + \frac{\varepsilon_1}{2}$, $\ell_2^* \leq \min\{0.2275 + \frac{\varepsilon_1}{2}, \frac{(0.685 - \ell_1^*)}{2} + \frac{\varepsilon_1}{4}\}]$ *and* $\beta = 0.075$.

(vii) $r \in \{2, 3\}$ *and* $\ell_1^* \leq 0.29 + \frac{\varepsilon_1}{2}$, $\ell_2^* \in [0.42 - \ell_1^* - \frac{\varepsilon_1}{2}, 0.455 - \ell_1^* + \frac{\varepsilon_1}{2}]$ *and* $\beta \leq 0.105$.

(viii) $r \in \{4, 5\}$ *and* $\ell_1^* \leq 0.29 + \frac{\varepsilon_1}{2}$, $\ell_4^* \leq \frac{(0.62 - \sum_{i=1}^3 \ell_3^*)}{2} + \frac{\varepsilon_1}{4}$ *and* $\beta = 0.075$.

Then $(\ell_1^*, \ldots, \ell_r^*, \beta)$ is contained in $\mathcal{R}^\star(a)$.

Alternatively, let $r \in \{2, 4\}$, $0.075 - \frac{\varepsilon_1}{2} \leq \ell_i^* \leq 0.53$, $\ell_i^* \geq \ell_{i+1}^*$ and $\ell_1^* + \cdots + \ell_r^* \leq 0.99$. Assume that one of the following conditions (a) – (g) *holds:*

(a) $r = 2$ *and* $\ell_1^* \in [0.38 - \frac{\varepsilon_1}{2}, 0.42 + \frac{\varepsilon_1}{2}]$.

(b) $r = 2$ *and* $\ell_1^* \in [0.42 - \frac{\varepsilon_1}{2}, 0.455 + \frac{\varepsilon_1}{2}]$ *and* $\ell_2^* \in [0.58 - \ell_1^* - \frac{\varepsilon_1}{2}, 0.685 - \ell_1^* + \frac{\varepsilon_1}{2}]$.

(c) $r \in \{2, 4\}$ *and* $\ell_1^* \in [0.315 - \frac{\varepsilon_1}{2}, 0.38 + \frac{\varepsilon_1}{2}]$, $\ell_1^* + \ell_i^* \in [0.38 - \frac{\varepsilon_1}{2}, 0.455 + \frac{\varepsilon_1}{2}] \cup [0.545 - \frac{\varepsilon_1}{2}, 0.62 + \frac{\varepsilon_1}{2}]$ *for some* $i \neq 1$.

(d) $r \in \{2, 4\}$ *and* $\ell_i^* \in [0.38 - \ell_1^* - \frac{\varepsilon_1}{2}, 0.42 - \ell_1^* + \frac{\varepsilon_1}{2}] \cup [0.58 - \ell_1^* - \frac{\varepsilon_1}{2}, 0.62 - \ell_1^* + \frac{\varepsilon_1}{2}]$ *for some* $i \neq 1$.

(e) $r = 4$ *and* $\ell_2^* \in [0.42 - \ell_1^* - \frac{\varepsilon_1}{2}, 0.455 - \ell_1^* + \frac{\varepsilon_1}{2}]$, $\ell_4^* \in [0.315 - \ell_1^* - \frac{\varepsilon_1}{2}, 0.42 - \ell_1^* + \frac{\varepsilon_1}{2}]$.

(f) $r = 4$ *and* $\ell_1^* \leq 0.29 + \frac{\varepsilon_1}{2}$, $\ell_2^* \in [0.315 - \ell_1^* - \frac{\varepsilon_1}{2}, 0.38 - \ell_1^* + \frac{\varepsilon_1}{2}]$, $\ell_1^* + \ell_2^* + \ell_4^* \in [0.38 - \frac{\varepsilon_1}{2}, 0.455 + \frac{\varepsilon_1}{2}]$.

(g) $r = 4$ *and* $\ell_1^* \leq 0.29 + \frac{\varepsilon_1}{2}$, $\ell_2^* \in [0.315 - \ell_1^* - \frac{\varepsilon_1}{2}, 0.38 - \ell_1^* + \frac{\varepsilon_1}{2}]$, $\ell_2^* + \ell_3^* + \ell_4^* \in [0.38 - \frac{\varepsilon_1}{2}, 0.455 + \frac{\varepsilon_1}{2}]$.

Then $(\ell_1^*, \ldots, \ell_r^*)$ *is contained in* $\mathcal{R}^{\star\star}(a)$.

Proof. Suppose $a \in (0.57, 0.59]$. Recall $\chi_0(a) = 0.33 - \varepsilon_1$, $\chi_1(a) = [0.38 - \varepsilon_1, 0.42 + \varepsilon_1] \cup [0.58 - \varepsilon_1, 0.62 + \varepsilon_1]$, $\chi_2(a) = [0.38 - \varepsilon_1, 0.455 + \varepsilon_1] \cup [0.545 - \varepsilon_1, 0.62 + \varepsilon_1]$ and $\chi_3(a) = [0.315 - \varepsilon_1, 0.42 + \varepsilon_1] \cup [0.58 - \varepsilon_1, 0.685 + \varepsilon_1]$.

<u>Case 1</u>: Treatment of $\Xi^\star$. Let $\{\ell_i\}_{i=1}^j \in \Xi^\star(\ell_1^*, \ldots, \ell_r^*, \beta)$. If there exists $k \in \{1, \ldots, j\}$ with $\ell_k \geq \chi_0(a)$, then $\{\ell_i\}_{i=1}^j$ satisfies option (1). So we may assume that $\ell_i < \chi_0(a)$ for each i and take $\rho = \chi_0(a)$ in Lemma 2.29.

With each of the conditions (i), (ii), (iii), (iv), (v), (vi), (vii) and (viii) we have $\ell_1^* + \cdots + \ell_r^* \leq 0.685 + \frac{\varepsilon_1}{2}$. We apply Lemma 2.29 with $\beta \leq 0.105$, $\chi = [0.58 - \varepsilon_1, 0.685 + \varepsilon_1] \subseteq \chi_3(a)$, $\rho = 0.33 - \varepsilon_1$ and condition (A) to deduce that there exists $I_3 \subseteq \{1, \ldots, j\}$ with $\sum_{i \in I_3} \ell_i \in \chi_3(a)$.

In condition (iii) we also have $\ell_1^* \in [0.42 - \frac{\varepsilon_1}{2}, 0.455 + \frac{\varepsilon_1}{2}]$ and in condition (vii) we have $\ell_1^* + \ell_2^* \in [0.42 - \frac{\varepsilon_1}{2}, 0.455 + \frac{\varepsilon_1}{2}]$. By condition (E) of Lemma 2.29, there exists $I_2 \subseteq \{1, \ldots, j\}$ with $\sum_{i \in I_2} \ell_i \in \chi_2(a)$. With condition (i), (v) and (viii) (as well as (iv), if $\max\{0.145 + \frac{\varepsilon_1}{2}, (0.62 - \ell_1^*)/2 + \frac{\varepsilon_1}{4}\} = (0.62 - \ell_1^*)/2 + \frac{\varepsilon_1}{4})$, we have $\ell_1^* + \cdots + \ell_r^* \leq 0.62 + \frac{\varepsilon_1}{2}$ and $\beta = 0.075$. Applying Lemma 2.29 with $\beta = 0.075$, $\chi = [0.545 - \varepsilon_1, 0.62 + \varepsilon_1]$, $\rho = 0.33 - \varepsilon_1$ and condition (A), we again find $I_2 \subseteq \{1, \ldots, j\}$ with $\sum_{i \in I_2} \ell_i \in \chi_2(a)$. With condition (ii) we observe that $\ell_2^* \leq (0.685 - \ell_1^*)/2 + \frac{\varepsilon_1}{4} \leq 0.125$ for $\ell_1^* \geq 0.455 - \frac{\varepsilon_1}{2}$, while $\ell_1^* + \ell_2^* \leq 0.62 + \frac{\varepsilon_1}{2}$. We note that $1 - 0.33 - 0.125 = 0.545$ and apply Lemma 2.29 with $\beta = 0.075$, $\chi = [0.545 - \varepsilon_1, 0.62 + \varepsilon_1]$, $\rho = 0.33 - \varepsilon_1$ and condition (A) or (B) (where $k = 3$ and $b_1 = 0.125$), depending on whether $r = 2$ or $r = 3$. Similarly, if $\max\{0.145 + \frac{\varepsilon_1}{2}, (0.62 - \ell_1^*)/2 + \frac{\varepsilon_1}{4}\} = 0.145 + \frac{\varepsilon_1}{2}$, then condition (iv) is treated by observing that $1 - 0.33 - 2 \cdot 0.145 = 0.38$ and taking $\beta = 0.075$, $\chi = [0.38 - \varepsilon_1, 0.455 + \varepsilon_1]$, $\rho = 0.33 - \varepsilon_1$ and $k = 2$ and $b_1 = 0.145 + \frac{\varepsilon_1}{2}$ in condition (B) of Lemma 2.29. Finally, in condition (vi) we have $\ell_1^* \leq 0.29 + \frac{\varepsilon_1}{2}$ and $\sum_{i=2}^3 \ell_i^* \leq 0.455 + \frac{\varepsilon_1}{2}$. Since $1 - 0.33 - 0.29 = 0.38$, we can apply Lemma 2.29 with $\beta = 0.075$, $\chi = [0.38 - \varepsilon_1, 0.455 + \varepsilon_1]$, $\rho = 0.33 - \varepsilon_1$ and $b_2 = 0.29 + \frac{\varepsilon_1}{2}$ in condition (C) to deduce that there exists $I_2 \subseteq \{1, \ldots, j\}$ with $\sum_{i \in I_2} \ell_i \in \chi_2(a)$. Hence $\{\ell_i\}_{i=1}^j$ satisfies option (3).

Case 2: Treatment of $\Xi^{\star\star}$. The cases (a) – (g) can be treated just like the corresponding cases of Corollary 2.30 and Corollary 2.31. The values of ℓ_i^* have been carefully restricted to ensure that there exists a set $S_1 \subseteq \{1,\ldots,r\}$ with $\sum_{i\in S_1} \ell_i^* + \delta \in \chi_1(a)$ for all $\delta \in [-\frac{\varepsilon_1}{2}, \frac{\varepsilon_1}{2}]$ or two sets $S_2, S_3 \subseteq \{1,\ldots,r\}$ with $\sum_{i\in S_j} \ell_i^* + \delta \in \chi_j(a)$ for $\delta \in [-\frac{\varepsilon_1}{2}, \frac{\varepsilon_1}{2}]$ and $j \in \{2,3\}$. For a given $\{\ell_i\}_{i=1}^j \in \Xi^{\star\star}(\ell_1^*,\ldots,\ell_r^*)$, Lemma 2.29 again tells us that $\{\ell_i\}_{i=1}^j$ satisfies option (2) or (3). $\square$

Corollary 2.34. *Let $\varepsilon > 0$ and $a \in (0.59, 0.61]$.*

Let $r \in \{0,\ldots,5\}$. Suppose $\ell_1^*,\ldots,\ell_r^*$ satisfy $\beta \le \ell_i^* \le 0.5+\varepsilon$, $\ell_i^* \ge \ell_{i+1}^*$ and $\ell_1^* + \cdots + \ell_r^* \le 0.75$. Assume that one of the following conditions (i) – (vii) *holds:*

(i) *$(r = 0$ or $r = 1)$ and $\beta = 0.07$.*

(ii) *$(r = 2$ or $r = 3)$ and $\ell_1^* \ge 0.435 - \frac{\varepsilon_1}{2}$, $\ell_2^* \le \min\{0.105 + \frac{\varepsilon_1}{2}, \frac{(0.67-\ell_1^*)}{2} + \frac{\varepsilon_1}{4}\}$, $\beta = 0.07$.*

(iii) *$(r = 2$ or $r = 3)$ and $\ell_1^* \in [0.42 - \frac{\varepsilon_1}{2}, 0.435 + \frac{\varepsilon_1}{2}]$, $\ell_2^* \le \frac{(0.67-\ell_1^*)}{2} + \frac{\varepsilon_1}{4}$ and $\beta \le 0.09$.*

(iv) *$(r = 2$ or $r = 3)$ and $\ell_1^* \in [0.33 - \frac{\varepsilon_1}{2}, 0.365 + \frac{\varepsilon_1}{2}]$, $\ell_2^* \le 0.1524 + \frac{\varepsilon_1}{2}$, $\beta = 0.07$.*

(v) *$(r = 2$ or $r = 3)$ and $\ell_1^* \in [0.305 - \frac{\varepsilon_1}{2}, 0.33 + \frac{\varepsilon_1}{2}]$, $\ell_2^* \le \frac{(0.635-\ell_1^*)}{2} + \frac{\varepsilon_1}{4}$, $\beta = 0.07$.*

(vi) *$(r = 2$ or $r = 3)$ and $\ell_1^* \le 0.305 + \frac{\varepsilon_1}{2}$, $\ell_2^* \le 0.2099 + \frac{\varepsilon_1}{2}$, $\beta = 0.07$.*

(vii) *$(r = 4$ or $r = 5)$ and $\ell_1^* \le 0.305 + \frac{\varepsilon_1}{2}$, $\ell_4^* \le \frac{(0.635-\sum_{i=1}^3 \ell_3^*)}{2} + \frac{\varepsilon_1}{4}$, $\beta = 0.07$.*

Then $(\ell_1^*,\ldots,\ell_r^*,\beta)$ is contained in $\mathcal{R}^\star(a)$.

Alternatively, let $r \in \{2,4\}$, $0.07 - \frac{\varepsilon_1}{2} \le \ell_i^* \le 0.53$, $\ell_i^\star \ge \ell_{i+1}^\star$ and $\ell_1^\star + \cdots + \ell_r^\star \le 0.99$. Assume that one of the following conditions (a) – (c) *holds:*

(a) *$r \in \{2,4\}$ and $\exists\ S \subseteq \{1,\ldots,r\}$ with $\sum_{i\in S} \ell_i^* \in [0.365 - \frac{\varepsilon_1}{2}, 0.42 + \frac{\varepsilon_1}{2}] \cup [0.58 - \frac{\varepsilon_1}{2}, 0.635 + \frac{\varepsilon_1}{2}]$.*

(b) *$r = 2$ and $\ell_1^* \in [0.42 - \frac{\varepsilon_1}{2}, 0.435 + \frac{\varepsilon_1}{2}]$ and $\ell_2^* \in [0.58 - \ell_1^* - \frac{\varepsilon_1}{2}, 0.67 - \ell_1^* + \frac{\varepsilon_1}{2}]$.*

(c) *$r = 2$ and $\ell_1^* \in [0.33 - \frac{\varepsilon_1}{2}, 0.365 + \frac{\varepsilon_1}{2}]$, $\ell_1^* + \ell_2^* \in [0.365 - \frac{\varepsilon_1}{2}, 0.435 + \frac{\varepsilon_1}{2}] \cup [0.565 - \frac{\varepsilon_1}{2}, 0.635 + \frac{\varepsilon_1}{2}]$.*

Then $(\ell_1^*,\ldots,\ell_r^*)$ is contained in $\mathcal{R}^{\star\star}(a)$.

Proof. Suppose $a \in (0.59, 0.61]$. Recall $\chi_0(a) = 0.33 - \varepsilon_1$, $\chi_1(a) = [0.365 - \varepsilon_1, 0.42 + \varepsilon_1] \cup [0.58 - \varepsilon_1, 0.635 + \varepsilon_1]$, $\chi_2(a) = [0.365 - \varepsilon_1, 0.435 + \varepsilon_1] \cup [0.565 - \varepsilon_1, 0.635 + \varepsilon_1]$ and $\chi_3(a) = [0.33 - \varepsilon_1, 0.42 + \varepsilon_1] \cup [0.58 - \varepsilon_1, 0.67 + \varepsilon_1]$.

Case 1: Treatment of $\Xi^\star$. Let $\{\ell_i\}_{i=1}^j \in \Xi^\star(\ell_1^*,\ldots,\ell_r^*,\beta)$. If there exists $k \in \{1,\ldots,j\}$ with $\ell_k \ge \chi_0(a)$, then $\{\ell_i\}_{i=1}^j$ satisfies option (1). So we may assume that $\ell_i < \chi_0(a)$ for each i and take $\rho = \chi_0(a)$ in Lemma 2.29.

With each of the conditions (i), (ii), (iii), (iv), (v) and (vii) we have $\ell_1^* + \cdots + \ell_r^* \leq 0.67 + \frac{\varepsilon_1}{2}$. We apply Lemma 2.29 with $\beta \leq 0.09$, $\chi = [0.58 - \varepsilon_1, 0.67 + \varepsilon_1] \subseteq \chi_3(a)$, $\rho = 0.33 - \varepsilon_1$ and condition (A) to deduce that there exists $I_3 \subseteq \{1, \ldots, j\}$ with $\sum_{i \in I_3} \ell_i \in \chi_3(a)$. On the other hand, with condition (vi) we have $\ell_1^* \leq 0.305 + \frac{\varepsilon_1}{2}$ and $\ell_2^* \leq 0.2099 + \frac{\varepsilon_1}{2}$, so that $\sum_{i=2}^{3} \ell_i^* \leq 0.42 + \frac{\varepsilon_1}{2}$. Note $1 - 0.33 - 0.305 = 0.365$. We apply Lemma 2.29 with $\beta = 0.07$, $\chi = [0.33 - \varepsilon_1, 0.42 + \varepsilon_1]$, $\rho = 0.33 - \varepsilon_1$ and $b_2 = 0.305 + \frac{\varepsilon_1}{2}$ in condition (C) to deduce that there exists $I_3 \subseteq \{1, \ldots, j\}$ with $\sum_{i \in I_3} \ell_i \in \chi_3(a)$.

In condition (iii) we have $\ell_1^* \in [0.42 - \frac{\varepsilon_1}{2}, 0.435 + \frac{\varepsilon_1}{2}]$. By condition (E) of Lemma 2.29, there exists $I_2 \subseteq \{1, \ldots, j\}$ with $\sum_{i \in I_2} \ell_i \in \chi_2(a)$. With condition (i), (v) and (vii), we have $\ell_1^* + \cdots + \ell_r^* \leq 0.635 + \frac{\varepsilon_1}{2}$ and $\beta = 0.07$. Applying Lemma 2.29 with $\beta = 0.07$, $\chi = [0.565 - \varepsilon_1, 0.635 + \varepsilon_1]$, $\rho = 0.33 - \varepsilon_1$ and condition (A), we again find $I_2 \subseteq \{1, \ldots, j\}$ with $\sum_{i \in I_2} \ell_i \in \chi_2(a)$. With condition (ii) we have $\ell_1^* + \ell_2^* < 0.635$, while $\ell_3^* \leq 0.105 + \frac{\varepsilon_1}{2}$. We note that $1 - 0.33 - 0.105 = 0.565$ and apply Lemma 2.29 with $\beta = 0.07$, $\chi = [0.565 - \varepsilon_1, 0.635 + \varepsilon_1]$, $\rho = 0.33 - \varepsilon_1$ and condition (A) or (B) (where $k = 3$ and $b_1 = 0.105 + \frac{\varepsilon_1}{2}$), depending on whether $r = 2$ or $r = 3$. Condition (iv) is treated by observing that $1 - 0.33 - 2 \cdot 0.1525 = 0.365$ and taking $\beta = 0.07$, $\chi = [0.365 - \varepsilon_1, 0.435 + \varepsilon_1]$, $\rho = 0.33 - \varepsilon_1$ and $k = 2$ and $b_1 = 0.1525 + \frac{\varepsilon_1}{2}$ in condition (B) of Lemma 2.29. Finally, in condition (vi) we have $\ell_1^* \leq 0.305 + \frac{\varepsilon_1}{2}$ and $\sum_{i=2}^{3} \ell_i^* \leq 0.42 + \frac{\varepsilon_1}{2}$. Since $1 - 0.33 - 0.305 = 0.365$, we can apply Lemma 2.29 with $\beta = 0.07$, $\chi = [0.365 - \varepsilon_1, 0.435 + \varepsilon_1]$, $\rho = 0.33 - \varepsilon_1$ and $b_2 = 0.305 + \frac{\varepsilon_1}{2}$ in condition (C) to deduce that there exists $I_2 \subseteq \{1, \ldots, j\}$ with $\sum_{i \in I_2} \ell_i \in \chi_2(a)$. Hence $\{\ell_i\}_{i=1}^{j}$ satisfies option (3).

Case 2: Treatment of $\Xi^{\star\star}$. The cases (a) – (c) can be treated just like the corresponding cases of Corollary 2.30 and Corollary 2.31. The values of ℓ_i^* have been carefully restricted to ensure that there exists a set $S_1 \subseteq \{1, \ldots, r\}$ with $\sum_{i \in S_1} \ell_i^* + \delta \in \chi_1(a)$ for all $\delta \in [-\frac{\varepsilon_1}{2}, \frac{\varepsilon_1}{2}]$ or two sets $S_2, S_3 \subseteq \{1, \ldots, r\}$ with $\sum_{i \in S_j} \ell_i^* + \delta \in \chi_j(a)$ for $\delta \in [-\frac{\varepsilon_1}{2}, \frac{\varepsilon_1}{2}]$ and $j \in \{2, 3\}$. For a given $\{\ell_i\}_{i=1}^{j} \in \Xi^{\star\star}(\ell_1^*, \ldots, \ell_r^*)$, Lemma 2.29 again tells us that $\{\ell_i\}_{i=1}^{j}$ satisfies option (2) or (3). □

Corollary 2.35. *Let $\varepsilon > 0$ and $a \in (0.61, 0.77 - \varepsilon]$.*

Let $r \in \{0, \ldots, 5\}$. Suppose $\ell_1^*, \ldots, \ell_r^*$ satisfy $\beta \leq \ell_i^* \leq 0.5 + \varepsilon$, $\ell_i^* \geq \ell_{i+1}^*$ and $\ell_1^* + \cdots + \ell_r^* \leq 0.75$. Assume that one of the following conditions (i) – (vi) *holds:*

(i) *($r = 0$ or $r = 1$) and $\beta = 0.065$.*

(ii) *($r = 2$ or $r = 3$) and $\ell_1^* \geq 0.42 - \frac{\varepsilon_1}{2}$, $\ell_2^* \leq \min\{0.58 - \ell_1^*, 0.1\} + \frac{\varepsilon_1}{2}$ and $\beta = 0.065$.*

(iii) ($r = 2$ *or* $r = 3$) *and* $\ell_1^* \in [0.325 - \frac{\varepsilon_1}{2}, 0.355 + \frac{\varepsilon_1}{2}]$, $\ell_2^* \le \frac{(0.645-\ell_1^*)}{2} + \frac{\varepsilon_1}{4}$ *and* $\beta = 0.065$.

(iv) ($r = 2$ *or* $r = 3$) *and* $\ell_1^* \le 0.325 + \frac{\varepsilon_1}{2}$, $\ell_2^* \le 0.2099 + \frac{\varepsilon_1}{2}$ *and* $\beta = 0.065$.

(v) ($r = 4$ *or* $r = 5$) *and* $\ell_1^* \le 0.325 + \frac{\varepsilon_1}{2}$, $\ell_4^* \le \frac{(0.645-\sum_{i=1}^{3}\ell_i^*)}{2} + \frac{\varepsilon_1}{4}$ *and* $\beta = 0.065$.

(vi) ($r = 4$ *or* $r = 5$) *and* $\ell_1^* \le 0.325 + \frac{\varepsilon_1}{2}$, $\ell_4^* \le \frac{(0.42-\ell_2^*-\ell_3^*)}{2} + \frac{\varepsilon_1}{4}$ *and* $\beta = 0.065$.

Then $(\ell_1^*, \dots, \ell_r^*, \beta)$ *is contained in* $\mathcal{R}^\star(a)$.

Alternatively, let $r \in \{2, 4\}$, $0.065 - \frac{\varepsilon_1}{2} \le \ell_i^* \le 0.53$, $\ell_i^\star \ge \ell_{i+1}^\star$ *and* $\ell_1^\star + \dots + \ell_r^\star \le 0.99$. *Assume that one of the following conditions* (a) – (b) *holds:*

(a) $r = 2$ *and* $\ell_1^* \in [0.355 - \frac{\varepsilon_1}{2}, 0.42 + \frac{\varepsilon_1}{2}]$.

(b) $r \in \{2, 4\}$ *and* $\ell_1^* + \ell_i^* \in [0.355 - \frac{\varepsilon_1}{2}, 0.42 + \frac{\varepsilon_1}{2}] \cup [0.58 - \frac{\varepsilon_1}{2}, 0.645 + \frac{\varepsilon_1}{2}]$ *for some* $i \ne 1$.

Then $(\ell_1^*, \dots, \ell_r^*)$ *is contained in* $\mathcal{R}^{\star\star}(a)$.

Proof. Suppose $a \in (0.61, 0.77 - \varepsilon]$. Recall $\chi_0(a) = 0.32 - \varepsilon_1$ and $\chi_1(a) = [0.355 - \varepsilon_1, 0.42 + \varepsilon_1] \cup [0.58 - \varepsilon_1, 0.645 + \varepsilon_1]$.

Case 1: Treatment of $\Xi^\star$. Let $\{\ell_i\}_{i=1}^j \in \Xi^\star(\ell_1^*, \dots, \ell_r^*, 0.065)$. If there exists $k \in \{1, \dots, j\}$ with $\ell_k \ge \chi_0(a)$, then $\{\ell_i\}_{i=1}^j$ satisfies option (1). So we may assume that $\ell_i < \chi_0(a)$ for each i and take $\rho = \chi_0(a)$ in Lemma 2.29.

In conditions (i), (iii) and (v) we have $\ell_1^* + \dots + \ell_r^* \le 0.645 + \frac{\varepsilon_1}{2}$. We apply Lemma 2.29 with $\beta = 0.065$, $\chi = [0.58 - \varepsilon_1, 0.645 + \varepsilon_1]$, $\rho = 0.32 - \varepsilon_1$ and condition (A) to deduce that option (2) holds. With condition (ii) we have $\ell_1^* + \ell_2^* \in [0.42 - \frac{\varepsilon_1}{2}, 0.58 + \frac{\varepsilon_1}{2}]$, while $\ell_2^* \le 0.1 + \frac{\varepsilon_1}{2}$ and $r \in \{2, 3\}$. For $r = 2$ we again use condition (A), while for $r = 3$ we note $1 - 0.32 - (3 - 3 + 1)0.1 = 0.58$ and then apply Lemma 2.29 with $\beta = 0.065$, $\chi = [0.58 - \varepsilon_1, 0.645 + \varepsilon_1]$, $\rho = 0.32 - \varepsilon_1$ and $k = 3$ and $b_1 = 0.1 + \frac{\varepsilon_1}{2}$ in condition (B). Finally, with (iv) or (vi) we have $\sum_{i=2}^r \ell_i^* \le 0.42 + \frac{\varepsilon_1}{2}$ and $\ell_1^* \le 0.325 + \frac{\varepsilon_1}{2}$ and note $1 - 0.32 - 0.325 = 0.355$. We apply Lemma 2.29 with $\beta = 0.065$, $\chi = [0.355 - \varepsilon_1, 0.42 + \varepsilon_1] \subseteq \chi_1(a)$, $\rho = 0.32 - \varepsilon_1$ and $b_2 = 0.325 + \frac{\varepsilon_1}{2}$ in condition (C) to deduce that $\{\ell_i\}_{i=1}^j$ satisfies option (2) of Proposition 2.3.

Case 2: Treatment of $\Xi^{\star\star}$. The cases (a) and (b) can be treated just like the corresponding cases of Corollary 2.30 and Corollary 2.31. The values of ℓ_i^* have been carefully restricted to ensure that there exists a set $S_1 \subseteq \{1, \dots, r\}$ with $\sum_{i \in S_1} \ell_i^* + \delta \in \chi_1(a)$ for all $\delta \in [-\frac{\varepsilon_1}{2}, \frac{\varepsilon_1}{2}]$ or two sets $S_2, S_3 \subseteq \{1, \dots, r\}$ with $\sum_{i \in S_j} \ell_i^* + \delta \in \chi_j(a)$ for $\delta \in [-\frac{\varepsilon_1}{2}, \frac{\varepsilon_1}{2}]$ and $j \in \{2, 3\}$. For a given $\{\ell_i\}_{i=1}^j \in \Xi^{\star\star}(\ell_1^*, \dots, \ell_r^*)$, Lemma 2.29 again tells us that $\{\ell_i\}_{i=1}^j$ satisfies option (2) or (3). □

2.6.3 Computations and the Buchstab identity

To complete the proof of Proposition 2.3, we just need to use the elements of $\mathcal{R}^\star(a)$ and $\mathcal{R}^{\star\star}(a)$ given in Corollaries 2.30, 2.31, 2.32, 2.33, 2.34 and 2.35 to decompose $1_{\mathbb{P}}(n)$ into a sum which is as described in (2.210): $1_{\mathbb{P}}(n) = \sum_{i=1}^{s_1} \Psi_i(n) + \sum_{i=1}^{s_2} \Theta_i(n)$ with $s_1, s_2 \leq 100$ and each $\Psi_i : [x, 6x] \cap \mathbb{N} \to \mathbb{R}$ good* or good** and Θ_i non-negative with $\sum_{i=1}^{s} \sum_{n \in \mathcal{B}(y)} \Theta_i(n) \leq \frac{0.9999y}{x^b \log x}$.

2.6.3.1 Useful identities and notation

Recall the Buchstab identity: For $z_1 < z_2$,

$$\psi(n, z_2) = \psi(n, z_1) - \sum_{\substack{n = p_1 m \\ z_1 \leq p_1 < z_2}} \psi(m, p_1).$$

Recall $X = (7x)^{1/2}$. We only need to decompose $1_{\mathbb{P}}(n)$ when $n \in [x, 6x]$, where $1_{\mathbb{P}}(n) = \psi(n, X)$. The desired decomposition will be obtained via repeated use of the Buchstab identity, starting with $\psi(n, X)$.

To denote the expressions which appear in repeated applications of the Buchstab identity, we define the following shorthand notation: Let $\mathcal{R} \subseteq [0,1]^r$ and let $\beta \in [0,1]$. Using ℓ_i^* as a shorthand for $\log_x(p_i)$, we then set

$$\Omega^\star(n; \mathcal{R}, x^\beta) = \sum_{\substack{n = p_1 \dots p_r m \\ x^\beta \leq p_r < \dots < p_1 < X \\ (\ell_1^*, \dots, \ell_r^*) \in \mathcal{R}}} \psi(m, x^\beta),$$

$$\Omega^{\star\star}(n; \mathcal{R}, x^\beta) = \sum_{\substack{n = p_1 \dots p_r m \\ x^\beta \leq p_r < \dots < p_1 < X \\ (\ell_1^*, \dots, \ell_r^*) \in \mathcal{R}}} \psi(m, p_r).$$

In our computations, $\mathcal{R}$ will typically be defined via some inequalities of the form $k_1 \ell_1^* + \dots + k_r \ell_r^* \leq c$. To decide if $\Omega^\star(n; \mathcal{R}, x^\beta)$ is a good* function and $\Omega^{\star\star}(n; \mathcal{R}, x^\beta)$ is a good** function, we will compare these inequalities with the conditions of Corollaries 2.30, 2.31, 2.32, 2.33, 2.34 and 2.35.

When $\Omega^{\star\star}(n; \mathcal{R}, x^\beta)$ is not good**, we will have to compute $\sum_{n \in \mathcal{B}(y)} \Omega^{\star\star}(n; \mathcal{R}, x^\beta)$. For reasonably nice sets $\mathcal{R}$, a generalised prime number theorem can be used to bound $\sum_{n \in \mathcal{B}(y)} \Omega^{\star\star}(n; \mathcal{R}, x^\beta)$ from above. We define $I(\mathcal{R}, x^\beta)$ as follows:

$$I(\mathcal{R}, x^\beta) = \int_\beta^{0.5+\epsilon_1} \int_\beta^{\min\{\alpha_1, \frac{(1-\alpha_1)}{2}\}} \cdots \int_\beta^{\min\{\alpha_{r-1}, \frac{(1-\sum_{i=1}^{r-1} \alpha_i)}{2}\}} \omega\left(\frac{1 - \sum_{i=1}^r \alpha_i}{\alpha_r}\right) \frac{1_{\mathcal{R}}(\alpha_1, \dots, \alpha_r)}{\alpha_1 \dots \alpha_{r-1} \alpha_r^2} \, \mathrm{d}\alpha_r \dots \mathrm{d}\alpha_1.$$

Here $\omega(u)$ is the Buchstab function, which satisfies $\omega(u) \leq \max\{0.6, \frac{1}{u}\}$ for $u \geq 1$. (See pages 16 and 17 of [14] for a detailed discussion of $\omega(u)$.) We have also adopted the convention that $\int_c^d = 0$ whenever $c > d$.

For nice sets $\mathcal{R}$ (such as polytopes) then

$$\sum_{n \in \mathcal{B}(y)} \Omega^{\star\star}(n; \mathcal{R}, x^{\beta}) \leq \frac{y(1+o(1))}{x^b \log x} I(\mathcal{R}, x^{\beta}). \tag{2.214}$$

In particular, (2.214) will be applicable to all sums $\sum_{n \in \mathcal{B}(y)} \Omega^{\star\star}(n; \mathcal{R}, x^{\beta})$ which appear in our Harman's sieve computations on the following pages.

2.6.3.2 Sieve computations

Having defined all this useful notation, we are now finally ready to prove Proposition 2.3. The case $a \in (0.53, 0.545]$ corresponds to one of the more complicated situations and its treatment also demonstrates the full range of arguments required to deal with all other cases. For illustrative purposes, we thus begin with a very detailed proof of (2.210) for $a \in (0.53, 0.545]$, and then discuss the hardest case, $a \in (0.57, 0.59]$. The remaining four cases are postponed to Appendix A.

<u>Case 1</u>: $a \in (0.53, 0.545]$.

After two applications of the Buchstab identity with $z_1 = x^{0.08}$, and after using ℓ_i^* as a shorthand for $\log_x(p_i)$, we have

$$\begin{aligned}
\psi(n, X) &= \psi(n, x^{0.08}) - \sum_{\substack{n = p_1 m \\ x^{0.08} \leq p_1 < X}} \psi(m, p_1) \\
&= \psi(n, x^{0.08}) - \sum_{\substack{n = p_1 m \\ x^{0.08} \leq p_1 < X}} \psi(m, x^{0.08}) + \sum_{\substack{n = p_1 p_2 m \\ x^{0.08} \leq p_2 < p_1 < X}} \psi(m, p_2) \\
&= \psi(n, x^{0.08}) - \sum_{\substack{n = p_1 m \\ x^{0.08} \leq p_1 < X}} \psi(m, x^{0.08}) \qquad (2.215) \\
&+ \sum_{\substack{n = p_1 p_2 m \\ x^{0.08} \leq p_2 < p_1 < X \\ \ell_1^* > 0.474}} \psi(m, p_2) + \sum_{\substack{n = p_1 p_2 m \\ x^{0.08} \leq p_2 < p_1 < X \\ 0.427 \leq \ell_1^* \leq 0.474}} \psi(m, p_2) + \sum_{\substack{n = p_1 p_2 m \\ x^{0.08} \leq p_2 < p_1 < X \\ 0.375 < \ell_1^* < 0.427}} \psi(m, p_2) \qquad (2.216) \\
&+ \sum_{\substack{n = p_1 p_2 m \\ x^{0.08} \leq p_2 < p_1 < X \\ \ell_1^* \in [0.285, 0.315) \cup (0.345, 0.375]}} \psi(m, p_2) + \sum_{\substack{n = p_1 p_2 m \\ x^{0.08} \leq p_2 < p_1 < X \\ 0.315 \leq \ell_1^* \leq 0.345}} \psi(m, p_2) + \sum_{\substack{n = p_1 p_2 m \\ x^{0.08} \leq p_2 < p_1 < X \\ \ell_1^* < 0.285}} \psi(m, p_2). \qquad (2.217)
\end{aligned}$$

We thus define the following regions:

$$\mathcal{R}_{2,1} = \{(\ell_1^*, \ell_2^*) \in [0,1]^2 : \ell_1^* > 0.474\},$$

$$\mathcal{R}_{2,2} = \{(\ell_1^*, \ell_2^*) \in [0,1]^2 : 0.427 \leq \ell_1^* < 0.474\},$$
$$\mathcal{R}_{2,3} = \{(\ell_1^*, \ell_2^*) \in [0,1]^2 : 0.375 \leq \ell_1^* < 0.427\},$$
$$\mathcal{R}_{2,4} = \{(\ell_1^*, \ell_2^*) \in [0,1]^2 : \ell_1^* \in [0.285, 0.315) \cup (0.345, 0.375]\},$$
$$\mathcal{R}_{2,5} = \{(\ell_1^*, \ell_2^*) \in [0,1]^2 : 0.315 \leq \ell_1^* \leq 0.345\},$$
$$\mathcal{R}_{2,6} = \{(\ell_1^*, \ell_2^*) \in [0,1]^2 : \ell_1^* < 0.285\}.$$

Using the notation of Section 2.6.3.1, we then have

$$\psi(n, X) = \psi(n, x^{0.08}) - \Omega^\star(n; [0,1], x^{0.08}) + \sum_{i=1}^{6} \Omega^{\star\star}(n; \mathcal{R}_{2,i}, x^{0.08}).$$

By condition (i) of Corollary 2.31, $\psi(n, x^{0.08})$ and $-\Omega^\star(n; [0,1], x^{0.08})$ are good* functions. By condition (a) of Corollary 2.31, $\Omega^{\star\star}(n; \mathcal{R}_{2,2}, x^{0.08})$ and $\Omega^{\star\star}(n; \mathcal{R}_{2,5}, x^{0.08})$ are good** functions. It remains to treat $\Omega^{\star\star}(n; \mathcal{R}_{2,i}, x^{0.08})$ for $i \in \{1, 3, 4, 6\}$.

Step 1: Treatment of $\Omega^{\star\star}(n; \mathcal{R}_{2,1}, x^{0.08})$.

Recall $\mathcal{R}_{2,1} = \{(\ell_1^*, \ell_2^*) \in [0,1]^2 : \ell_1^* > 0.474\}$. We partition $\mathcal{R}_{2,1}$ as follows:

$$\mathcal{R}_{2,1,1} = \{(\ell_1^*, \ell_2^*) \in \mathcal{R}_{2,1} : \ell_2^* \geq \min\{0.595 - \ell_1^*, (0.715 - \ell_1^*)/2\},$$
$$\mathcal{R}_{2,1,2} = \{(\ell_1^*, \ell_2^*) \in \mathcal{R}_{2,1} : \ell_2^* < \min\{0.595 - \ell_1^*, (0.715 - \ell_1^*)/2\}.$$

We then have

$$\Omega^{\star\star}(n; \mathcal{R}_{2,1}, x^{0.08}) = \Omega^{\star\star}(n; \mathcal{R}_{2,1,1}, x^{0.08}) + \Omega^{\star\star}(n; \mathcal{R}_{2,1,2}, x^{0.08}).$$

Write $\Theta_1(n) = \Omega^{\star\star}(n; \mathcal{R}_{2,1,1}, x^{0.08})$. Using a generalised prime number theorem, we can bound $\sum_{n \in \mathcal{B}(y)} \Theta_1(n)$ from above by $\frac{y(1+o(1))}{x^b \log x} I(\mathcal{R}_{2,1,1}, x^{0.08})$. Note that $\mathcal{R}_{2,1,1} = \{(\ell_1^*, \ell_2^*) \in [0,1]^2 : \ell_1^* > 0.474, \ell_2^* \geq \min\{0.595 - \ell_1^*, (0.715 - \ell_1^*)/2\}$. Hence

$$\sum_{n \in \mathcal{B}(y)} \Theta_1(n) \leq \frac{y(1+o(1))}{x^b \log x} \int_{0.474}^{0.5+\varepsilon_1} \int_{\min\{0.595-\alpha_1, (0.715-\alpha_1)/2\}}^{(1-\alpha_1)/2} \frac{\omega\left(\frac{1-\alpha_1-\alpha_2}{\alpha_2}\right)}{\alpha_1 \alpha_2^2} \mathrm{d}\alpha_2 \mathrm{d}\alpha_1 < \frac{0.185y}{x^b \log x}.$$

Next we apply two Buchstab iterations to $\Omega^{\star\star}(n; \mathcal{R}_{2,1,2}, x^{0.08})$, the first with $z_1 = x^{0.08}$ and the second with $z_1 = x^{0.09}$.

$$\Omega^{\star\star}(n; \mathcal{R}_{2,1,2}, x^{0.08}) = \sum_{\substack{n = p_1 p_2 m \\ x^{0.08} \leq p_2 < p_1 < X \\ \ell_1^* > 0.474 \\ \ell_2^* < \min\{0.595 - \ell_1^*, (0.715 - \ell_1^*)/2\}}} \psi(m, x^{0.08}) - \sum_{\substack{n = p_1 p_2 p_3 m \\ x^{0.08} \leq p_3 < p_2 < p_1 < X \\ \ell_1^* > 0.474,\, \ell_3^* \leq 0.09 \\ \ell_2^* < \min\{0.595 - \ell_1^*, (0.715 - \ell_1^*)/2\}}} \psi(m, p_3) \tag{2.218}$$

$$- \sum_{\substack{n = p_1 p_2 p_3 m \\ x^{0.08} \leq p_3 < p_2 < p_1 < X \\ \ell_1^* > 0.474,\, \ell_3^* > 0.09 \\ \ell_2^* < \min\{0.595 - \ell_1^*, (0.715 - \ell_1^*)/2\}}} \psi(m, x^{0.09}) + \sum_{\substack{n = p_1 p_2 p_3 p_4 m \\ x^{0.09} \leq p_4 < p_3 < p_2 < p_1 < X \\ \ell_1^* > 0.474 \\ \ell_2^* < \min\{0.595 - \ell_1^*, (0.715 - \ell_1^*)/2\}}} \psi(m, p_4). \tag{2.219}$$

By condition (ii) of Corollary 2.31, both sums in (2.218) are good* functions. (The second sum is of the special type of good* described at the end of Section 2.6.1.) The first sum in (2.219) is also good* by condition (ii) of Corollary 2.31. Denote the second sum in (2.219) by $\Theta_2(n)$. Then $\sum_{n\in\mathcal{B}(y)}\Theta_2(n)$ is bounded above by:

$$\frac{y(1+o(1))}{x^b\log x}\int_{0.474}^{0.5+\varepsilon_1}\int_{0.09}^{\min\{0.595-\alpha_1,\frac{(0.715-\alpha_1)}{2}\}}\int_{0.09}^{\alpha_2}\int_{0.09}^{\min\{\alpha_3,\frac{(1-\sum_{i=1}^{3}\alpha_i)}{2}\}}\frac{\omega\left(\frac{1-\sum_{i=1}^{4}\alpha_i}{\alpha_4}\right)}{\alpha_1\alpha_2\alpha_3\alpha_4^2}d\underline{\alpha}<\frac{0.001y}{x^b\log x}.$$

Step 2: Treatment of $\Omega^{\star\star}(n;\mathcal{R}_{2,3},x^{0.08})$.

Recall $\mathcal{R}_{2,3}=\{(\ell_1^*,\ell_2^*)\in[0,1]^2:0.375\le\ell_1^*<0.427\}$. We partition $\mathcal{R}_{2,3}$ as follows:

$$\begin{aligned}
\mathcal{R}_{2,3,1}&=\{(\ell_1^*,\ell_2^*)\in\mathcal{R}_{2,3}:\ell_2^*\in(0.573-\ell_1^*,0.655-\ell_1^*)\cup(0.685-\ell_1^*,\ell_1^*)\},\\
\mathcal{R}_{2,3,2}&=\{(\ell_1^*,\ell_2^*)\in\mathcal{R}_{2,3}:\ell_1^*+\ell_2^*\in[0.427,0.474]\cup[0.526,0.573]\cup[0.655,0.685]\},\\
\mathcal{R}_{2,3,3}&=\{(\ell_1^*,\ell_2^*)\in\mathcal{R}_{2,3}:\ell_2^*\in(0.474-\ell_1^*,0.526-\ell_1^*),\ell_2^*>(0.655-\ell_1^*)/2\},\\
\mathcal{R}_{2,3,4}&=\{(\ell_1^*,\ell_2^*)\in\mathcal{R}_{2,3}:\ell_2^*\in(0.474-\ell_1^*,0.526-\ell_1^*),\ell_2^*\le(0.655-\ell_1^*)/2\}.
\end{aligned}$$

We then have

$$\Omega^{\star\star}(n;\mathcal{R}_{2,3},x^{0.08})=\sum_{i=1}^{4}\Omega^{\star\star}(n;\mathcal{R}_{2,3,i},x^{0.08}).$$

By condition (b) of Corollary 2.31, $\Omega^{\star\star}(n;\mathcal{R}_{2,3,2},x^{0.08})$ is a sum of three good** functions. Write $\Theta_3(n)=\Omega^{\star\star}(n;\mathcal{R}_{2,3,1},x^{0.08})$ and $\Theta_4(n)=\Omega^{\star\star}(n;\mathcal{R}_{2,3,3},x^{0.08})$. Then

$$\sum_{n\in\mathcal{B}(y)}\Theta_3(n)\le\frac{y(1+o(1))}{x^b\log x}\int_{0.375}^{0.427}\int_{\substack{[0.573-\alpha_1,0.655-\alpha_1]\\ \cup[0.685-\alpha_1,(1-\alpha_1)/2]}}\frac{\omega\left(\frac{1-\alpha_1-\alpha_2}{\alpha_2}\right)}{\alpha_1\alpha_2^2}d\alpha_2 d\alpha_1<\frac{0.175y}{x^b\log x},$$

$$\sum_{n\in\mathcal{B}(y)}\Theta_4(n)\le\frac{y(1+o(1))}{x^b\log x}\int_{0.375}^{0.427}\int_{\max\{0.08,0.474-\alpha_1,(0.655-\alpha_1)/2\}}^{0.526-\alpha_1}\frac{\omega\left(\frac{1-\alpha_1-\alpha_2}{\alpha_2}\right)}{\alpha_1\alpha_2^2}d\alpha_2 d\alpha_1<\frac{0.01y}{x^b\log x}.$$

Now we apply two more Buchstab iterations to $\Omega^{\star\star}(n;\mathcal{R}_{2,3,4},x^{0.08})$:

$$\begin{aligned}
\Omega^{\star\star}(n;\mathcal{R}_{2,3,4},x^{0.08})=&\sum_{\substack{n=p_1p_2m\\ x^{0.08}\le p_2<p_1<X\\ 0.375<\ell_1^*<0.427\\ \ell_2^*\in(0.474-\ell_1^*,0.526-\ell_1^*)\\ \ell_2^*\le(0.655-\ell_1^*)/2}}\psi(m,x^{0.08})-\sum_{\substack{n=p_1p_2p_3m\\ x^{0.08}<p_3<p_2<p_1<X\\ 0.375<\ell_1^*<0.427\\ \ell_2^*\in(0.474-\ell_1^*,0.526-\ell_1^*)\\ \ell_2^*\le(0.655-\ell_1^*)/2}}\psi(m,x^{0.08}) &(2.220)\\
&+\sum_{\substack{n=p_1p_2p_3p_4m\\ x^{0.08}\le p_4<p_3<p_2<p_1<X\\ 0.375<\ell_1^*<0.427\\ \ell_2^*\in(0.474-\ell_1^*,0.526-\ell_1^*)\\ \ell_2^*\le(0.655-\ell_1^*)/2}}\psi(m,p_4). &(2.221)
\end{aligned}$$

Writing $\mathcal{R}_{3,3,4} = \{(\ell_1^*, \ell_2^*, \ell_3^*) \in [0,1]^3 : (\ell_1^*, \ell_2^*) \in \mathcal{R}_{2,3,4}\}$ and $\mathcal{R}_{4,3,4} = \{(\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*) \in [0,1]^4 : (\ell_1^*, \ell_2^*) \in \mathcal{R}_{2,3,4}\}$, this equation can also be expressed as

$$\Omega^{\star\star}(n; \mathcal{R}_{2,3,4}, x^{0.08}) = \Omega^{\star}(n; \mathcal{R}_{2,3,4}, x^{0.08}) - \Omega^{\star}(n; \mathcal{R}_{3,3,4}, x^{0.08}) + \Omega^{\star\star}(n; \mathcal{R}_{4,3,4}, x^{0.08}).$$

By condition (iii) of Corollary 2.31, both $\Omega^{\star}(n; \mathcal{R}_{2,3,4}, x^{0.08})$ and $-\Omega^{\star}(n; \mathcal{R}_{3,3,4}, x^{0.08})$ are good* functions. Write $\Theta_5(n) = \Omega^{\star\star}(n; \mathcal{R}_{4,3,4}, x^{0.08})$. Then $\sum_{n \in \mathcal{B}(y)} \Theta_5(n)$ is bounded above by:

$$\frac{y(1+o(1))}{x^b \log x} \int_{0.375}^{0.427} \int_{\max\{0.08, 0.474-\alpha_1\}}^{\min\{0.526-\alpha_1, (0.655-\alpha_1)/2\}} \int_{0.08}^{\alpha_2} \int_{0.08}^{\min\{\alpha_3, \frac{(1-\sum_{i=1}^{3}\alpha_i)}{2}\}} \frac{\omega\left(\frac{1-\sum_{i=1}^{4}\alpha_i}{\alpha_4}\right)}{\alpha_1\alpha_2\alpha_3\alpha_4^2} \mathrm{d}\underline{\alpha} < \frac{0.013y}{x^b \log x}.$$

Step 3: Treatment of $\Omega^{\star\star}(n; \mathcal{R}_{2,4}, x^{0.08})$.

Recall $\mathcal{R}_{2,4} = \{(\ell_1^*, \ell_2^*) \in [0,1]^2 : \ell_1^* \in [0.285, 0.315) \cup (0.345, 0.375]\}$. We define the following partition of $\mathcal{R}_{2,4}$:

$$\begin{aligned}
\mathcal{R}_{2,4,1} &= \{(\ell_1^*, \ell_2^*) \in \mathcal{R}_{2,4} : \ell_2^* > 0.595 - \ell_1^*\},\\
\mathcal{R}_{2,4,2} &= \{(\ell_1^*, \ell_2^*) \in \mathcal{R}_{2,4} : \ell_2^* \in [0.515 - \ell_1^*, 0.595 - \ell_1^*]\},\\
\mathcal{R}_{2,4,3} &= \{(\ell_1^*, \ell_2^*) \in \mathcal{R}_{2,4} : \ell_2^* \in (0.485 - \ell_1^*, 0.515 - \ell_1^*), \ell_2^* > (0.655 - \ell_1^*)/2\},\\
\mathcal{R}_{2,4,4} &= \{(\ell_1^*, \ell_2^*) \in \mathcal{R}_{2,4} : \ell_2^* \in (0.485 - \ell_1^*, 0.515 - \ell_1^*), \ell_2^* \leq (0.655 - \ell_1^*)/2\},\\
\mathcal{R}_{2,4,5} &= \{(\ell_1^*, \ell_2^*) \in \mathcal{R}_{2,4} : \ell_2^* \in [0.405 - \ell_1^*, 0.485 - \ell_1^*]\},\\
\mathcal{R}_{2,4,6} &= \{(\ell_1^*, \ell_2^*) \in \mathcal{R}_{2,4} : \ell_2^* < 0.405 - \ell_1^*\}.
\end{aligned}$$

We then have

$$\Omega^{\star\star}(n; \mathcal{R}_{2,4}, x^{0.08}) = \sum_{i=1}^{6} \Omega^{\star\star}(n; \mathcal{R}_{2,4,i}, x^{0.08}).$$

By condition (c) of Corollary 2.31, both $\Omega^{\star\star}(n; \mathcal{R}_{2,4,2}, x^{0.08})$ and $\Omega^{\star\star}(n; \mathcal{R}_{2,4,5}, x^{0.08})$ are the sum of two good** functions. Write $\Theta_6(n) = \Omega^{\star\star}(n; \mathcal{R}_{2,4,1}, x^{0.08})$ and $\Theta_7(n) = \Omega^{\star\star}(n; \mathcal{R}_{2,4,3}, x^{0.08})$. We bound $\sum_{n \in \mathcal{B}(y)} \Theta_i(n)$, $i \in \{6, 7\}$, by $\frac{y(1+o(1))}{x^b \log x} I(\mathcal{R}_{2,4,1}, x^{0.08})$ and $\frac{y(1+o(1))}{x^b \log x} I(\mathcal{R}_{2,4,3}, x^{0.08})$, respectively. More precisely:

$$\sum_{n \in \mathcal{B}(y)} \Theta_6(n) \leq \frac{y(1+o(1))}{x^b \log x} \int_{\substack{[0.285, 0.315] \\ \cup [0.345, 0.375]}} \int_{0.595-\alpha_1}^{\min\{\alpha_1, (1-\alpha_1)/2\}} \frac{\omega\left(\frac{1-\alpha_1-\alpha_2}{\alpha_2}\right)}{\alpha_1\alpha_2^2} \mathrm{d}\alpha_2 \mathrm{d}\alpha_1 < \frac{0.08y}{x^b \log x},$$

$$\sum_{n \in \mathcal{B}(y)} \Theta_7(n) \leq \frac{y(1+o(1))}{x^b \log x} \int_{\substack{[0.285, 0.315] \\ \cup [0.345, 0.375]}} \int_{\max\{0.485-\alpha_1, \frac{(0.655-\alpha_1)}{2}\}}^{0.515-\alpha_1} \frac{\omega\left(\frac{1-\alpha_1-\alpha_2}{\alpha_2}\right)}{\alpha_1\alpha_2^2} \mathrm{d}\alpha_2 \mathrm{d}\alpha_1 < \frac{0.062y}{x^b \log x}.$$

This leaves $\Omega^{\star\star}(n;\mathcal{R}_{2,4,4},x^{0.08})$ and $\Omega^{\star\star}(n;\mathcal{R}_{2,4,6},x^{0.08})$. We apply two more Buchstab iterations to each. Defining the regions

$$\mathcal{R}_{j,4,i}=\{(\ell_1^*,\ldots,\ell_j^*)\in[0,1]^j:(\ell_1^*,\ell_2^*)\in\mathcal{R}_{2,4,i}\},$$

the resulting equality can be expressed as follows: For $i\in\{4,6\}$,

$$\Omega^{\star\star}(n;\mathcal{R}_{2,4,i},x^{0.08})=\Omega^{\star}(n;\mathcal{R}_{2,4,i},x^{0.08})-\Omega^{\star}(n;\mathcal{R}_{3,4,i},x^{0.08})+\Omega^{\star\star}(n;\mathcal{R}_{4,4,i},x^{0.08}).$$

For $i\in\{4,6\}$, $\Omega^{\star}(n;\mathcal{R}_{2,4,i},x^{0.08})$ and $\Omega^{\star}(n;\mathcal{R}_{3,4,i},x^{0.08})$ are good* by condition (iii) of Corollary 2.31. Write $\Theta_8(n)=\Omega^{\star\star}(n;\mathcal{R}_{4,4,4},x^{0.08})$ and $\Theta_9(n)=\Omega^{\star\star}(n;\mathcal{R}_{4,4,6},x^{0.08})$. Note that $\ell_1^*<0.315$ implies $0.485-\ell_1^*>(0.655-\ell_1^*)/2$. Hence $\sum_{n\in\mathcal{B}(y)}\Theta_8(n)$ and $\sum_{n\in\mathcal{B}(y)}\Theta_9(n)$ are bounded above by the following quantities, respectively:

$$\frac{y(1+o(1))}{x^b\log x}\int_{0.345}^{0.375}\int_{0.485-\alpha_1}^{(0.655-\alpha_1)/2}\int_{0.08}^{\alpha_2}\int_{0.08}^{\min\{\alpha_3,\frac{(1-\sum_{i=1}^{3}\alpha_i)}{2}\}}\frac{\omega\left(\frac{1-\alpha_1-\alpha_2-\alpha_3-\alpha_4}{\alpha_4}\right)}{\alpha_1\alpha_2\alpha_3\alpha_4^2}\mathrm{d}\underline{\alpha}<\frac{0.012y}{x^b\log x},$$
$$\frac{y(1+o(1))}{x^b\log x}\int_{0.285}^{0.315}\int_{0.08}^{0.405-\alpha_1}\int_{0.08}^{\alpha_2}\int_{0.08}^{\min\{\alpha_3,\frac{(1-\sum_{i=1}^{3}\alpha_i)}{2}\}}\frac{\omega\left(\frac{1-\alpha_1-\alpha_2-\alpha_3-\alpha_4}{\alpha_4}\right)}{\alpha_1\alpha_2\alpha_3\alpha_4^2}\mathrm{d}\underline{\alpha}<\frac{0.003y}{x^b\log x}.$$

Step 4: Treatment of $\Omega^{\star\star}(n;\mathcal{R}_{2,6},x^{0.08})$.

Recall $\mathcal{R}_{2,6}=\{(\ell_1^*,\ell_2^*)\in[0,1]^2:\ell_1^*<0.285\}$. We partition $\mathcal{R}_{2,6}$ as follows:

$$\begin{aligned}\mathcal{R}_{2,6,1}&=\{(\ell_1^*,\ell_2^*)\in\mathcal{R}_{2,6}:\ell_2^*\geq 0.526-\ell_1^*\},\\ \mathcal{R}_{2,6,2}&=\{(\ell_1^*,\ell_2^*)\in\mathcal{R}_{2,6}:\ell_2^*<0.526-\ell_1^*,\ell_2^*\geq(0.655-\ell_1^*)/2\},\\ \mathcal{R}_{2,6,3}&=\{(\ell_1^*,\ell_2^*)\in\mathcal{R}_{2,6}:\ell_2^*<0.526-\ell_1^*,\ell_2^*<(0.655-\ell_1^*)/2\}.\end{aligned}$$

We then have

$$\Omega^{\star\star}(n;\mathcal{R}_{2,6},x^{0.08})=\sum_{i=1}^{3}\Omega^{\star\star}(n;\mathcal{R}_{2,6,i},x^{0.08}).$$

By condition (b) of Corollary 2.31, $\Omega^{\star\star}(n;\mathcal{R}_{2,6,1},x^{0.08})$ is a good** function. Write $\Theta_{10}(n)=\Omega^{\star\star}(n;\mathcal{R}_{2,6,2},x^{0.08})$. Then $\sum_{n\in\mathcal{B}(y)}\Theta_{10}(n)\leq\frac{y(1+o(1))}{x^b\log x}I(\mathcal{R}_{2,6,2},x^{0.08})$. So

$$\sum_{n\in\mathcal{B}(y)}\Theta_{10}(n)\leq\frac{y(1+o(1))}{x^b\log x}\int_{0.655/3}^{0.285}\int_{(0.655-\alpha_1)/2}^{\min\{\alpha_1,0.526-\alpha_1\}}\frac{\omega\left(\frac{1-\alpha_1-\alpha_2}{\alpha_2}\right)}{\alpha_1\alpha_2^2}\mathrm{d}\alpha_2\mathrm{d}\alpha_1<\frac{0.01y}{x^b\log x}.$$

Finally, we apply up to four Buchstab iterations to $\Omega^{\star\star}(n;\mathcal{R}_{2,6,3},x^{0.08})$. We define:

$$\mathcal{R}_{j,6,3}=\{(\ell_1^*,\ldots,\ell_j^*)\in[0,1]^j:(\ell_1^*,\ell_2^*)\in\mathcal{R}_{2,6,3}\}\quad\text{for}\quad j\in\{3,4\},$$

$$
\begin{aligned}
\mathcal{R}_{4,6,3,1} &= \{(\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*) \in \mathcal{R}_{4,6,3} : \ell_4^* > (0.655 - \ell_1^* - \ell_2^* - \ell_3^*)/2\},\\
\mathcal{R}_{4,6,3,2} &= \{(\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*) \in \mathcal{R}_{4,6,3} : \ell_4^* \leq (0.655 - \ell_1^* - \ell_2^* - \ell_3^*)/2\},\\
\mathcal{R}_{j,6,3,2} &= \{(\ell_1^*, \dots, \ell_j^*) \in [0,1]^j : (\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*) \in \mathcal{R}_{4,6,3,2}\} \quad \text{for} \quad j \in \{5,6\}.
\end{aligned}
$$

Since $\mathcal{R}_{4,6,3,1} \cup \mathcal{R}_{4,6,3,2} = \mathcal{R}_{4,6,3}$, the Buchstab identity gives

$$
\begin{aligned}
\Omega^{\star\star}(n; \mathcal{R}_{2,6,3}, x^{0.08}) = \Omega^{\star}(n; \mathcal{R}_{2,6,3}, x^{0.08}) &- \Omega^{\star}(n; \mathcal{R}_{3,6,3}, x^{0.08})\\
&+ \Omega^{\star\star}(n; \mathcal{R}_{4,6,3,1}, x^{0.08}) + \Omega^{\star}(n; \mathcal{R}_{4,6,3,2}, x^{0.08})\\
&- \Omega^{\star}(n; \mathcal{R}_{5,6,3,2}, x^{0.08}) + \Omega^{\star\star}(n; \mathcal{R}_{6,6,3,2}, x^{0.08}).
\end{aligned}
$$

$\Omega^{\star}(n; \mathcal{R}_{2,6,3}, x^{0.08})$ and $-\Omega^{\star}(n; \mathcal{R}_{3,6,3}, x^{0.08})$ are good* functions by condition (iii) of Corollary 2.31, while $\Omega^{\star\star}(n; \mathcal{R}_{6,6,3,2}, x^{0.08})$ and $\Omega^{\star}(n; \mathcal{R}_{2,6,3}, x^{0.08})$ are also good* by condition (iv) of Corollary 2.31. Set $\Theta_{11}(n) = \Omega^{\star\star}(n; \mathcal{R}_{4,6,3,1}, x^{0.08})$ and $\Theta_{12}(n) = \Omega^{\star\star}(n; \mathcal{R}_{6,6,3,2}, x^{0.08})$. It follows that

$$
\Theta_{11}(n) = \sum_{\substack{n = p_1 p_2 p_3 p_4 m \\ x^{0.08} \leq p_4 < p_3 < p_2 < p_1 < X \\ \ell_1^* < 0.285 \\ \ell_2^* < \min\{(0.655 - \ell_1^*)/2, 0.526 - \ell_1^*\} \\ \ell_4^* > (0.655 - \ell_1^* - \ell_2^* - \ell_3^*)/2}} \psi(m, p_4) \qquad \text{and} \qquad \Theta_{12}(n) = \sum_{\substack{n = p_1 p_2 p_3 p_4 p_5 p_6 m \\ x^{0.08} \leq p_6 < p_5 < p_4 < p_3 < p_2 < p_1 < X \\ \ell_1^* < 0.285 \\ \ell_2^* < \min\{(0.655 - \ell_1^*)/2, 0.526 - \ell_1^*\} \\ \ell_4^* \leq (0.655 - \ell_1^* - \ell_2^* - \ell_3^*)/2}} \psi(m, p_6).
$$

Quantity $\sum_{n \in \mathcal{B}(y)} \Theta_{11}(n)$ is bounded above by

$$
\frac{y(1+o(1))}{x^b \log x} \int_{0.08}^{0.285} \int_{0.08}^{\min\{\alpha_1, \frac{(0.655-\alpha_1)}{2}\}} \int_{0.08}^{\alpha_2} \int_{\max\{0.08, \frac{(0.655 - \sum_{i=1}^3 \alpha_i)}{2}\}}^{\min\{\alpha_3, \frac{(1 - \sum_{i=1}^3 \alpha_i)}{2}\}} \frac{\omega\left(\frac{1 - \sum_{i=1}^4 \alpha_i}{\alpha_4}\right)}{\alpha_1 \alpha_2 \alpha_3 \alpha_4^2} \mathrm{d}\underline{\alpha} < \frac{0.296y}{x^b \log x}.
$$

Finally, to bound $\sum_{n \in \mathcal{B}(y)} \Theta_{12}(n)$, note that $\ell_1^* + \cdots + \ell_5^* \leq 0.655$ implies $\ell_3^* < 0.17$. Hence it is sufficient to count the number of integers $n = p_1 \dots p_6 m$ in $\mathcal{B}(y)$ which have $(m, P(x^{0.08})) = 1$ and ($p_1, p_2 \in [x^{0.17}, x^{0.285}]$ or $p_1 \in [x^{0.17}, x^{0.285}]$, $p_2 \in [x^{0.08}, x^{0.17}]$ or $p_1, p_2 \in [x^{0.08}, x^{0.17}]$) and $p_i \in [x^{0.08}, x^{0.17}]$ for $i \geq 3$. The double counting introduced by ignoring the conditions $p_i > p_{i+1}$ can be removed by dividing by 6! or 5! or 4!2!, respectively. We obtain the following upper bound:

$$
\sum_{n \in \mathcal{B}(y)} \Theta_{12}(n) \leq \frac{y(1+o(1))}{x^b \log x} \left(\frac{\log\left(\frac{0.17}{0.08}\right)^6}{0.08 \cdot 6!} + \frac{\log\left(\frac{0.17}{0.08}\right)^5 \log\left(\frac{0.285}{0.17}\right)}{0.08 \cdot 5!} + \frac{\log\left(\frac{0.17}{0.08}\right)^4 \log\left(\frac{0.285}{0.17}\right)^2}{0.08 \cdot 4!2!} \right) < \frac{0.04y}{x^b \log x}.
$$

In summary, we have shown that $1_{\mathbb{P}}(n)$ can be written as $1_{\mathbb{P}}(n) = \Psi(n) + \sum_{i=1}^{12} \Theta_i(n)$, where $\Psi(n)$ is a sum of no more than 100 good* and good** functions and where each Θ_i is non-negative and

$$
\sum_{i=1}^{12} \sum_{n \in \mathcal{B}(y)} \Theta_i(n) < \frac{0.887y}{x^b \log x}.
$$

This concludes the proof of Proposition 2.3 for the case $a \in (0.53, 0.545]$.

<u>Case 2</u>: $a \in (0.57, 0.59]$.

Now we treat the hardest case, $a \in (0.57, 0.59]$. We define the regions

$$\begin{aligned}
\mathcal{R}_{2,1} &= \{(\ell_1^*, \ell_2^*) \in [0,1]^2 : \ell_1^* > 0.455\},\\
\mathcal{R}_{2,2} &= \{(\ell_1^*, \ell_2^*) \in [0,1]^2 : 0.42 < \ell_1^* \leq 0.455\},\\
\mathcal{R}_{2,3} &= \{(\ell_1^*, \ell_2^*) \in [0,1]^2 : 0.38 < \ell_1^* \leq 0.42\},\\
\mathcal{R}_{2,4} &= \{(\ell_1^*, \ell_2^*) \in [0,1]^2 : 0.315 < \ell_1^* \leq 0.38\},\\
\mathcal{R}_{2,5} &= \{(\ell_1^*, \ell_2^*) \in [0,1]^2 : 0.29 < \ell_1^* \leq 0.315\},\\
\mathcal{R}_{2,6} &= \{(\ell_1^*, \ell_2^*) \in [0,1]^2 : \ell_1^* \leq 0.29\}.
\end{aligned}$$

After two applications of the Buchstab identity with $z_1 = x^{0.075}$, we have

$$\psi(n, X) = \psi(n, x^{0.075}) - \Omega^\star(n; [0,1], x^{0.075}) + \sum_{i=1}^{6} \Omega^{\star\star}(n; \mathcal{R}_{2,i}, x^{0.075}).$$

By condition (i) of Corollary 2.33, $\psi(n, x^{0.075})$ and $-\Omega^\star(n; [0,1], x^{0.075})$ are good* functions. By condition (a) of Corollary 2.33, $\Omega^{\star\star}(n; \mathcal{R}_{2,3}, x^{0.075})$ is a good** function. It remains to treat $\Omega^{\star\star}(n; \mathcal{R}_{2,i}, x^{0.075})$ for $i \in \{1, 2, 4, 5, 6\}$.

<u>Step 1</u>: Treatment of $\Omega^{\star\star}(n; \mathcal{R}_{2,1}, x^{0.075})$.

Recall $\mathcal{R}_{2,1} = \{(\ell_1^*, \ell_2^*) \in [0,1]^2 : \ell_1^* > 0.455\}$. We now define the following regions:

$$\begin{aligned}
\mathcal{R}_{2,1,1} &= \{(\ell_1^*, \ell_2^*) \in \mathcal{R}_{2,1} : \ell_2^* > 0.62 - \ell_1^*\},\\
\mathcal{R}_{2,1,2} &= \{(\ell_1^*, \ell_2^*) \in \mathcal{R}_{2,1} : \ell_2^* \in [0.58 - \ell_1^*, 0.62 - \ell_1^*]\},\\
\mathcal{R}_{2,1,3} &= \{(\ell_1^*, \ell_2^*) \in \mathcal{R}_{2,1} : \ell_2^* < 0.58 - \ell_1^*, \ell_2^* \geq (0.685 - \ell_1^*)/2\},\\
\mathcal{R}_{2,1,4} &= \{(\ell_1^*, \ell_2^*) \in \mathcal{R}_{2,1} : \ell_2^* < 0.58 - \ell_1^*, \ell_2^* < (0.685 - \ell_1^*)/2\}.
\end{aligned}$$

Additionally, following the same notational conventions as in the treatment of the case $a \in (0.53, 0.545]$, we extend tuples by writing

$$\mathcal{R}_{j,1,4} = \{(\ell_1^*, \ldots, \ell_j^*) \in [0,1]^j : (\ell_1^*, \ell_2^*) \in \mathcal{R}_{2,1,4}\}.$$

Applying two more Buchstab iterations to the region $\mathcal{R}_{2,1,4}$, then

$$\begin{aligned}
\Omega^{\star\star}(n; \mathcal{R}_{2,1}, x^{0.075}) &= \Omega^{\star\star}(n; \mathcal{R}_{2,1,1}, x^{0.075}) + \Omega^{\star\star}(n; \mathcal{R}_{2,1,2}, x^{0.075})\\
&\quad + \Omega^{\star\star}(n; \mathcal{R}_{2,1,3}, x^{0.075}) + \Omega^\star(n; \mathcal{R}_{2,1,4}, x^{0.075})
\end{aligned}$$

$$- \Omega^\star(n;\mathcal{R}_{3,1,4},x^{0.075}) + \Omega^\star(n;\mathcal{R}_{4,1,4},x^{0.075}).$$

$\Omega^\star(n;\mathcal{R}_{2,1,4},x^{0.075})$ and $-\Omega^\star(n;\mathcal{R}_{3,1,4},x^{0.075})$ are good* functions by condition (ii) of Corollary 2.33. Further, $\Omega^{\star\star}(n;\mathcal{R}_{2,1,2},x^{0.075})$ is good** by condition (d) of Corollary 2.33. Now we set $\Theta_1(n) = \Omega^{\star\star}(n;\mathcal{R}_{2,1,1},x^{0.075})$, $\Theta_2(n) = \Omega^{\star\star}(n;\mathcal{R}_{2,1,3},x^{0.075})$ and $\Theta_3(n) = \Omega^\star(n;\mathcal{R}_{4,1,4},x^{0.075})$. We have computed the following upper bounds:

$$\sum_{n\in\mathcal{B}(y)} \Theta_1(n) \le \frac{y(1+o(1))}{x^b\log x} I(\mathcal{R}_{2,1,1},x^{0.075}) < \frac{0.2029y}{x^b\log x},$$

$$\sum_{n\in\mathcal{B}(y)} \Theta_2(n) \le \frac{y(1+o(1))}{x^b\log x} I(\mathcal{R}_{2,1,3},x^{0.075}) < \frac{0.0099y}{x^b\log x},$$

$$\sum_{n\in\mathcal{B}(y)} \Theta_3(n) \le \frac{y(1+o(1))}{x^b\log x} I(\mathcal{R}_{4,1,4},x^{0.075}) < \frac{0.0038y}{x^b\log x}.$$

Step 2: Treatment of $\Omega^{\star\star}(n;\mathcal{R}_{2,2},x^{0.075})$.

Recall $\mathcal{R}_{2,2} = \{(\ell_1^*,\ell_2^*) \in [0,1]^2 : 0.42 < \ell_1^* \le 0.455\}$. We define the following regions:

$$\begin{aligned}
\mathcal{R}_{2,2,1} &= \{(\ell_1^*,\ell_2^*) \in \mathcal{R}_{2,2} : \ell_2^* > 0.685 - \ell_1^*\},\\
\mathcal{R}_{2,2,2} &= \{(\ell_1^*,\ell_2^*) \in \mathcal{R}_{2,2} : \ell_2^* \in [0.58 - \ell_1^*, 0.685 - \ell_1^*]\},\\
\mathcal{R}_{2,2,3} &= \{(\ell_1^*,\ell_2^*) \in \mathcal{R}_{2,2} : \ell_2^* \in ((0.685 - \ell_1^*)/2, 0.58 - \ell_1^*)\},\\
\mathcal{R}_{2,2,4} &= \{(\ell_1^*,\ell_2^*) \in \mathcal{R}_{2,2} : \ell_2^* \le (0.685 - \ell_1^*)/2, \ell_2^* \le 0.105\},\\
\mathcal{R}_{2,2,5} &= \{(\ell_1^*,\ell_2^*) \in \mathcal{R}_{2,2} : \ell_2^* \le (0.685 - \ell_1^*)/2, \ell_2^* > 0.105\},\\
\mathcal{R}_{j,2,5} &= \{(\ell_1^*,\ldots,\ell_j^*) \in [0,1]^j : (\ell_1^*,\ell_2^*) \in \mathcal{R}_{2,2,5}\} \quad \text{for} \quad j \in \{3,4\}.
\end{aligned}$$

Next we apply two Buchstab iterations with $z_1 = x^{0.105}$ to $\mathcal{R}_{2,2,5}$.

$$\begin{aligned}
\Omega^{\star\star}(n;\mathcal{R}_{2,2},x^{0.075}) = &\sum_{i=1}^{4} \Omega^{\star\star}(n;\mathcal{R}_{2,2,i},x^{0.075}) + \Omega^\star(n;\mathcal{R}_{2,2,5},x^{0.075})\\
&- \Omega^\star(n;\mathcal{R}_{3,2,5},x^{0.105}) + \Omega^{\star\star}(n;\mathcal{R}_{4,2,5},x^{0.105}).
\end{aligned}$$

Since condition (iii) of Corollary 2.33 allows $\beta \le 0.105$, $\Omega^{\star\star}(n;\mathcal{R}_{2,2,4},x^{0.075})$ and $\Omega^\star(n;\mathcal{R}_{2,2,5},x^{0.075})$ and $-\Omega^\star(n;\mathcal{R}_{3,2,5},x^{0.105})$ are good*. ($\Omega^{\star\star}(n;\mathcal{R}_{2,2,4},x^{0.075})$ is of the special type of good* described at the end of Section 2.6.1.) By condition (b) of Corollary 2.33, $\Omega^{\star\star}(n;\mathcal{R}_{2,2,2},x^{0.075})$ is good**. We now set $\Theta_4(n) = \Omega^{\star\star}(n;\mathcal{R}_{2,2,1},x^{0.075})$, $\Theta_5(n) = \Omega^{\star\star}(n;\mathcal{R}_{2,2,3},x^{0.075})$ and $\Theta_6(n) = \Omega^{\star\star}(n;\mathcal{R}_{4,2,5},x^{0.105})$. We have computed the following upper bounds:

$$\sum_{n\in\mathcal{B}(y)} \Theta_4(n) \le \frac{y(1+o(1))}{x^b\log x} I(\mathcal{R}_{2,2,1},x^{0.075}) < \frac{0.0345y}{x^b\log x},$$

$$\sum_{n\in\mathcal{B}(y)} \Theta_5(n) \le \frac{y(1+o(1))}{x^b \log x} I(\mathcal{R}_{2,2,3}, x^{0.075}) < \frac{0.0502y}{x^b \log x},$$
$$\sum_{n\in\mathcal{B}(y)} \Theta_6(n) \le \frac{y(1+o(1))}{x^b \log x} I(\mathcal{R}_{4,2,5}, x^{0.105}) < \frac{0.0004y}{x^b \log x}.$$

Step 3: Treatment of $\Omega^{\star\star}(n; \mathcal{R}_{2,4}, x^{0.075})$.

Recall $\mathcal{R}_{2,4} = \{(\ell_1^*, \ell_2^*) \in [0,1]^2 : 0.315 < \ell_1^* \le 0.38\}$. We define the following regions:

$$\begin{aligned}
&\mathcal{R}_{2,4,1} = \{(\ell_1^*, \ell_2^*) \in \mathcal{R}_{2,4} : \ell_2^* > 0.62 - \ell_1^*\},\\
&\mathcal{R}_{2,4,2} = \{(\ell_1^*, \ell_2^*) \in \mathcal{R}_{2,4} : \ell_2^* \in [0.38 - \ell_1^*, 0.455 - \ell_1^*) \cup (0.545 - \ell_1^*, 0.62 - \ell_1^*]\},\\
&\mathcal{R}_{2,4,3} = \{(\ell_1^*, \ell_2^*) \in \mathcal{R}_{2,4} : \ell_2^* \in [0.455 - \ell_1^*, 0.545 - \ell_1^*], \ell_2^* > \max\{0.145, \tfrac{(0.62-\ell_1^*)}{2}\}\},\\
&\mathcal{R}_{2,4,4} = \{(\ell_1^*, \ell_2^*) \in \mathcal{R}_{2,4} : \ell_2^* \in [0.455 - \ell_1^*, 0.545 - \ell_1^*], \ell_2^* \le \max\{0.145, \tfrac{(0.62-\ell_1^*)}{2}\}\},\\
&\mathcal{R}_{j,4,4} = \{(\ell_1^*, \ldots, \ell_j^*) \in [0,1]^j : (\ell_1^*, \ell_2^*) \in \mathcal{R}_{2,4,4}\} \quad \text{for} \quad j \in \{3,4\},\\
&\mathcal{R}_{4,4,4,1} = \{(\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*) \in \mathcal{R}_{4,4,4} : \ell_4^* > 0.455 - \ell_1^*\},\\
&\mathcal{R}_{4,4,4,2} = \{(\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*) \in \mathcal{R}_{4,4,4} : \ell_4^* \le 0.455 - \ell_1^*\}.
\end{aligned}$$

By condition (iv) of Corollary 2.33, $\Omega^{\star\star}(n; \mathcal{R}_{2,4}, x^{0.075})$ can be written as

$$\Omega^{\star\star}(n; \mathcal{R}_{2,4}, x^{0.075}) = \Psi_4(n) + \sum_{i=1}^{3} \Omega^{\star\star}(n; \mathcal{R}_{2,4,i}, x^{0.075}) + \sum_{i=1}^{2} \Omega^{\star\star}(n; \mathcal{R}_{4,4,4,i}, x^{0.075})$$

where $\Psi_4(n)$ is a sum of four good* functions, resulting from Buchstab iterations applied to region $\mathcal{R}_{2,4,4}$. By condition (c) of Corollary 2.33, $\Omega^{\star\star}(n; \mathcal{R}_{2,4,2}, x^{0.075})$ and $\Omega^{\star\star}(n; \mathcal{R}_{4,4,4,2}, x^{0.075})$ are good**. We write $\Theta_7(n) = \Omega^{\star\star}(n; \mathcal{R}_{2,4,1}, x^{0.075})$, $\Theta_8(n) = \Omega^{\star\star}(n; \mathcal{R}_{2,4,3}, x^{0.075})$ and $\Theta_9(n) = \Omega^{\star\star}(n; \mathcal{R}_{4,4,4,1}, x^{0.075})$. Then

$$\sum_{n\in\mathcal{B}(y)} \Theta_7(n) \le \frac{y(1+o(1))}{x^b \log x} I(\mathcal{R}_{2,4,1}, x^{0.075}) < \frac{0.0889y}{x^b \log x},$$
$$\sum_{n\in\mathcal{B}(y)} \Theta_8(n) \le \frac{y(1+o(1))}{x^b \log x} I(\mathcal{R}_{2,4,3}, x^{0.075}) < \frac{0.1993y}{x^b \log x},$$
$$\sum_{n\in\mathcal{B}(y)} \Theta_9(n) \le \frac{y(1+o(1))}{x^b \log x} I(\mathcal{R}_{4,4,4,1}, x^{0.075}) < \frac{0.0114y}{x^b \log x}.$$

Step 4: Treatment of $\Omega^{\star\star}(n; \mathcal{R}_{2,5}, x^{0.075})$.

Recall $\mathcal{R}_{2,5} = \{(\ell_1^*, \ell_2^*) \in [0,1]^2 : 0.29 < \ell_1^* \le 0.315\}$. We define the following regions:

$$\mathcal{R}_{2,5,1} = \{(\ell_1^*, \ell_2^*) \in \mathcal{R}_{2,5} : \ell_2^* > 0.62 - \ell_1^*\},$$

$$
\begin{aligned}
\mathcal{R}_{2,5,2} &= \{(\ell_1^*, \ell_2^*) \in \mathcal{R}_{2,5} : [0.58 - \ell_1^*, 0.62 - \ell_1^*]\},\\
\mathcal{R}_{2,5,3} &= \{(\ell_1^*, \ell_2^*) \in \mathcal{R}_{2,5} : \ell_2^* \in [(0.62 - \ell_1^*)/2, 0.58 - \ell_1^*)\},\\
\mathcal{R}_{2,5,4} &= \{(\ell_1^*, \ell_2^*) \in \mathcal{R}_{2,4} : \ell_2^* < (0.62 - \ell_1^*)/2\},\\
\mathcal{R}_{j,5,4} &= \{(\ell_1^*, \dots, \ell_j^*) \in [0,1]^j : (\ell_1^*, \ell_2^*) \in \mathcal{R}_{2,5,4}\} \quad \text{for} \quad j \in \{3,4\},\\
\mathcal{R}_{4,5,4,1} &= \{(\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*) \in \mathcal{R}_{4,5,4} : \ell_4^* > 0.42 - \ell_1^*\},\\
\mathcal{R}_{4,5,4,2} &= \{(\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*) \in \mathcal{R}_{4,5,4} : \ell_4^* \in [0.38 - \ell_1^*, 0.42 - \ell_1^*]\},\\
\mathcal{R}_{4,5,4,3} &= \{(\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*) \in \mathcal{R}_{4,5,4} : \ell_4^* < 0.38 - \ell_1^*\}.
\end{aligned}
$$

By condition (v) of Corollary 2.33, $\Omega^{\star\star}(n; \mathcal{R}_{2,5}, x^{0.075})$ can be written as

$$
\Omega^{\star\star}(n; \mathcal{R}_{2,5}, x^{0.075}) = \Psi_5(n) + \sum_{i=1}^{3} \Omega^{\star\star}(n; \mathcal{R}_{2,5,i}, x^{0.075}) + \sum_{i=1}^{3} \Omega^{\star\star}(n; \mathcal{R}_{4,5,4,i}, x^{0.075})
$$

where $\Psi_5(n)$ is a sum of two good* functions, resulting from Buchstab iterations applied to region $\mathcal{R}_{2,5,4}$. By condition (d) of Corollary 2.33, $\Omega^{\star\star}(n; \mathcal{R}_{2,5,2}, x^{0.075})$ and $\Omega^{\star\star}(n; \mathcal{R}_{4,5,4,2}, x^{0.075})$ are good**. We write $\Theta_{10}(n) = \Omega^{\star\star}(n; \mathcal{R}_{2,5,1}, x^{0.075})$, $\Theta_{11}(n) = \Omega^{\star\star}(n; \mathcal{R}_{2,5,3}, x^{0.075})$, $\Theta_{12}(n) = \Omega^{\star\star}(n; \mathcal{R}_{4,5,4,1}, x^{0.075})$, $\Theta_{13}(n) = \Omega^{\star\star}(n; \mathcal{R}_{4,5,4,3}, x^{0.075})$. We then compute

$$
\begin{aligned}
\sum_{n \in \mathcal{B}(y)} \Theta_{10}(n) &\le \frac{y(1+o(1))}{x^b \log x} I(\mathcal{R}_{2,5,1}, x^{0.075}) < \frac{0.0007y}{x^b \log x},\\
\sum_{n \in \mathcal{B}(y)} \Theta_{11}(n) &\le \frac{y(1+o(1))}{x^b \log x} I(\mathcal{R}_{2,5,3}, x^{0.075}) < \frac{0.1343y}{x^b \log x},\\
\sum_{n \in \mathcal{B}(y)} \Theta_{12}(n) &\le \frac{y(1+o(1))}{x^b \log x} I(\mathcal{R}_{4,5,4,1}, x^{0.075}) < \frac{0.0020y}{x^b \log x},\\
\sum_{n \in \mathcal{B}(y)} \Theta_{13}(n) &\le \frac{y(1+o(1))}{x^b \log x} I(\mathcal{R}_{4,5,4,3}, x^{0.075}) < \frac{0.0092y}{x^b \log x}.
\end{aligned}
$$

Step 5: Treatment of $\Omega^{\star\star}(n; \mathcal{R}_{2,6}, x^{0.075})$.

Recall $\mathcal{R}_{2,6} = \{(\ell_1^*, \ell_2^*) \in [0,1]^2 : \ell_1^* \le 0.29\}$. We define the following regions:

$$
\begin{aligned}
\mathcal{R}_{2,6,1} &= \{(\ell_1^*, \ell_2^*) \in \mathcal{R}_{2,6} : \ell_2^* > \min\{0.2275, (0.685 - \ell_1^*)/2\},\\
\mathcal{R}_{2,6,2} &= \{(\ell_1^*, \ell_2^*) \in \mathcal{R}_{2,6} : \ell_2^* \le \min\{0.2275, (0.685 - \ell_1^*)/2\},\\
\mathcal{R}_{4,6,2} &= \{(\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*) \in [0,1]^4 : (\ell_1^*, \ell_2^*) \in \mathcal{R}_{2,6,2}\},\\
\mathcal{R}_{4,6,2,a} &= \{(\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*) \in \mathcal{R}_{4,6,2} : \ell_4^* > \tfrac{(0.62 - \ell_1^* - \ell_2^* - \ell_3^*)}{2}\},\\
\mathcal{R}_{4,6,2,a,1} &= \{(\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*) \in \mathcal{R}_{4,6,2,a} : \ell_2^* > 0.455 - \ell_1^*, \ell_3^* < 0.38 - \ell_1^*\},
\end{aligned}
$$

$$\mathcal{R}_{4,6,2,a,2} = \{(\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*) \in \mathcal{R}_{4,6,2,a} : \ell_2^* > 0.455 - \ell_1^*, \ell_3^* \in [0.38 - \ell_1^*, 0.42 - \ell_1^*]\},$$
$$\mathcal{R}_{4,6,2,a,3} = \{(\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*) \in \mathcal{R}_{4,6,2,a} : \ell_2^* > 0.455 - \ell_1^*, \ell_3^* > 0.42 - \ell_1^*\},$$
$$\mathcal{R}_{4,6,2,a,4} = \{(\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*) \in \mathcal{R}_{4,6,2,a} : \ell_2^* \in [0.42 - \ell_1^*, 0.455 - \ell_1^*], \ell_4^* \geq 0.105\},$$
$$\mathcal{R}_{4,6,2,a,5} = \{(\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*) \in \mathcal{R}_{4,6,2,a} : \ell_2^* \in (0.38 - \ell_1^*, 0.42 - \ell_1^*)\},$$
$$\mathcal{R}_{4,6,2,a,6} = \{(\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*) \in \mathcal{R}_{4,6,2,a} : \ell_2^* \in [0.315 - \ell_1^*, 0.38 - \ell_1^*]\},$$
$$\mathcal{R}_{4,6,2,a,7} = \{(\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*) \in \mathcal{R}_{4,6,2,a} : \ell_2^* < 0.315 - \ell_1^*\},$$
$$\mathcal{R}_{4,6,2,b} = \{(\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*) \in \mathcal{R}_{4,6,2} : \ell_4^* \leq \tfrac{(0.62 - \ell_1^* - \ell_2^* - \ell_3^*)}{2}\},$$
$$\mathcal{R}_{6,6,2,b} = \{(\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*, \ell_5^*, \ell_6^*) \in [0,1]^6 : (\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*) \in \mathcal{R}_{4,6,2,b}\}.$$

Using conditions (vi), (vii) and (viii) of Corollary 2.33 and applying the Buchstab identity, we find that $\Omega^{\star\star}(n; \mathcal{R}_{2,6}, x^{0.075})$ equals

$$\begin{aligned}\Omega^{\star\star}(n; \mathcal{R}_{2,6}, x^{0.075}) = {} & \Psi_6(n) + \Omega^{\star\star}(n; \mathcal{R}_{2,6,1}, x^{0.075}) \\ & + \sum_{i=1}^{7} \Omega^{\star\star}(n; \mathcal{R}_{4,6,2,a,i}, x^{0.075}) + \Omega^{\star\star}(n; \mathcal{R}_{6,6,2,b}, x^{0.075}),\end{aligned}$$

where $\Psi_6(n)$ is a sum of nine good* functions. Both $\Omega^{\star\star}(n; \mathcal{R}_{4,6,2,a,2}, x^{0.075})$ and $\Omega^{\star\star}(n; \mathcal{R}_{4,6,2,a,5}, x^{0.075})$ are good** by condition (d) of Corollary 2.33. Now we define $\Theta_{14}(n) = \Omega^{\star\star}(n; \mathcal{R}_{2,6,1}, x^{0.075})$ and $\Theta_{15}(n) = \Omega^{\star\star}(n; \mathcal{R}_{4,6,2,a,1}, x^{0.075})$ and $\Theta_{16}(n) = \Omega^{\star\star}(n; \mathcal{R}_{4,6,2,a,7}, x^{0.075})$ and $\Theta_{17}(n) = \Omega^{\star\star}(n; \mathcal{R}_{6,6,2,b}, x^{0.075})$. We then compute

$$\sum_{n \in \mathcal{B}(y)} \Theta_{14}(n) \leq \frac{y(1+o(1))}{x^b \log x} I(\mathcal{R}_{2,6,1}, x^{0.075}) < \frac{0.1145y}{x^b \log x},$$
$$\sum_{n \in \mathcal{B}(y)} \Theta_{15}(n) \leq \frac{y(1+o(1))}{x^b \log x} I(\mathcal{R}_{4,6,2,a,1}, x^{0.075}) < \frac{0.0116y}{x^b \log x},$$
$$\sum_{n \in \mathcal{B}(y)} \Theta_{16}(n) \leq \frac{y(1+o(1))}{x^b \log x} I(\mathcal{R}_{4,6,2,a,7}, x^{0.075}) < \frac{0.0127y}{x^b \log x},$$
$$\sum_{n \in \mathcal{B}(y)} \Theta_{17}(n) \leq \frac{y(1+o(1))}{x^b \log x} I(\mathcal{R}_{6,6,2,b}, x^{0.075}) < \frac{0.0524y}{x^b \log x}.$$

It remains to treat $\Omega^{\star\star}(n; \mathcal{R}_{4,6,2,a,i}, x^{0.075})$ for $i \in \{3, 4, 6\}$. Note

$$\Omega^{\star\star}(n; \mathcal{R}_{2,6}, x^{0.075}) = \Psi_{6,1}(n) + \sum_{i=14}^{17} \Theta_i(n) + \sum_{i \in \{3,4,6\}} \Omega^{\star\star}(n; \mathcal{R}_{4,6,2,a,i}, x^{0.075}),$$

where $\Psi_{6,1}(n)$ is a sum of eleven functions that are good* or good**.

We split $\mathcal{R}_{4,6,2,a,3}$, $\mathcal{R}_{4,6,2,a,4}$ and $\mathcal{R}_{4,6,2,a,6}$ up as follows:

$$\mathcal{R}_{4,6,2,a,3,1} = \{(\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*) \in \mathcal{R}_{4,6,2,a,3} : \ell_4^* < 0.38 - \ell_1^*\},$$

$$\mathcal{R}_{4,6,2,a,3,2} = \{(\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*) \in \mathcal{R}_{4,6,2,a,3} : \ell_4^* \in [0.38 - \ell_1^*, 0.42 - \ell_1^*]\},$$
$$\mathcal{R}_{4,6,2,a,3,3} = \{(\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*) \in \mathcal{R}_{4,6,2,a,3} : \ell_4^* > 0.42 - \ell_1^*\},$$
$$\mathcal{R}_{4,6,2,a,4,1} = \{(\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*) \in \mathcal{R}_{4,6,2,a,4} : \ell_4^* > 0.42 - \ell_1^*\},$$
$$\mathcal{R}_{4,6,2,a,4,2} = \{(\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*) \in \mathcal{R}_{4,6,2,a,4} : \ell_4^* \in [0.315 - \ell_1^*, 0.42 - \ell_1^*]\},$$
$$\mathcal{R}_{4,6,2,a,6,1} = \{(\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*) \in \mathcal{R}_{4,6,2,a,6} : \ell_4^* > 0.455 - \ell_2^* - \ell_3^*\},$$
$$\mathcal{R}_{4,6,2,a,6,2} = \{(\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*) \in \mathcal{R}_{4,6,2,a,6} : \ell_4^* \in [0.38 - \ell_2^* - \ell_3^*, 0.455 - \ell_2^* - \ell_3^*]\},$$
$$\mathcal{R}_{4,6,2,a,6,3} = \{(\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*) \in \mathcal{R}_{4,6,2,a,6} : \ell_4^* \in (0.455 - \ell_1^* - \ell_2^*, 0.38 - \ell_2^* - \ell_3^*)\},$$
$$\mathcal{R}_{4,6,2,a,6,4} = \{(\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*) \in \mathcal{R}_{4,6,2,a,6} : \ell_4^* \leq \min\{0.455 - \ell_1^* - \ell_2^*, 0.38 - \ell_2^* - \ell_3^*\}\}.$$

This gives the following equality:

$$\Omega^{\star\star}(n; \mathcal{R}_{2,6}, x^{0.075}) = \Psi_{6,1}(n) + \sum_{i=14}^{17} \Theta_i(n) + \sum_{j=1}^{3} \Omega^{\star\star}(n; \mathcal{R}_{4,6,2,a,3,j}, x^{0.075})$$
$$+ \sum_{j=1}^{2} \Omega^{\star\star}(n; \mathcal{R}_{4,6,2,a,4,j}, x^{0.075}) + \sum_{j=1}^{4} \Omega^{\star\star}(n; \mathcal{R}_{4,6,2,a,6,j}, x^{0.075}).$$

The function $\Omega^{\star\star}(n; \mathcal{R}_{4,6,2,a,3,2}, x^{0.075})$ is good** by condition (d) of Corollary 2.33 and $\Omega^{\star\star}(n; \mathcal{R}_{4,6,2,a,4,2}, x^{0.075})$ is good** by condition (e) of Corollary 2.33. The function $\Omega^{\star\star}(n; \mathcal{R}_{4,6,2,a,6,4}, x^{0.075})$ is good** by condition (f) of Corollary 2.33 and the function $\Omega^{\star\star}(n; \mathcal{R}_{4,6,2,a,6,2}, x^{0.075})$ is good** by condition (g) of Corollary 2.33. Now write

$$\Theta_{18}(n) = \Omega^{\star\star}(n; \mathcal{R}_{4,6,2,a,3,1}, x^{0.075}),$$
$$\Theta_{19}(n) = \Omega^{\star\star}(n; \mathcal{R}_{4,6,2,a,3,3}, x^{0.075}),$$
$$\Theta_{20}(n) = \Omega^{\star\star}(n; \mathcal{R}_{4,6,2,a,4,1}, x^{0.075}),$$
$$\Theta_{21}(n) = \Omega^{\star\star}(n; \mathcal{R}_{4,6,2,a,6,1}, x^{0.075}),$$
$$\Theta_{22}(n) = \Omega^{\star\star}(n; \mathcal{R}_{4,6,2,a,6,3}, x^{0.075}).$$

We then compute

$$\sum_{n \in \mathcal{B}(y)} \Theta_{18}(n) \leq \frac{y(1 + o(1))}{x^b \log x} I(\mathcal{R}_{4,6,2,a,3,1}, x^{0.075}) < \frac{0.0129y}{x^b \log x},$$
$$\sum_{n \in \mathcal{B}(y)} \Theta_{19}(n) \leq \frac{y(1 + o(1))}{x^b \log x} I(\mathcal{R}_{4,6,2,a,3,3}, x^{0.075}) < \frac{0.0027y}{x^b \log x},$$
$$\sum_{n \in \mathcal{B}(y)} \Theta_{20}(n) \leq \frac{y(1 + o(1))}{x^b \log x} I(\mathcal{R}_{4,6,2,a,4,1}, x^{0.075}) < \frac{0.0012y}{x^b \log x},$$
$$\sum_{n \in \mathcal{B}(y)} \Theta_{21}(n) \leq \frac{y(1 + o(1))}{x^b \log x} I(\mathcal{R}_{4,6,2,a,6,1}, x^{0.075}) < \frac{0.0048y}{x^b \log x},$$

$$\sum_{n\in\mathcal{B}(y)} \Theta_{22}(n) \leq \frac{y(1+o(1))}{x^b \log x} I(\mathcal{R}_{4,6,2,a,6,3}, x^{0.075}) < \frac{0.0252y}{x^b \log x},$$

In summary, we have shown that $1_{\mathbb{P}}(n)$ can be written as $1_{\mathbb{P}}(n) = \Psi(n) + \sum_{i=1}^{22} \Theta_i(n)$, where $\Psi(n)$ is a sum of no more than 100 good* and good** functions and where each Θ_i is non-negative and

$$\sum_{i=1}^{22} \sum_{n\in\mathcal{B}(y)} \Theta_i(n) < \frac{0.9855y}{x^b \log x}.$$

This concludes the proof of Proposition 2.3 for the case $a \in (0.57, 0.59]$.

The remaining cases of Proposition 2.3, $a \in (0.59, 0.61]$, $a \leq 0.53$, $a \in (0.545, 0.57]$ and $a > 0.61$, are covered in Appendix A.

2.6.4 Final comments

Section 2.6 completed the proof of Theorem 2.1 and gave

$$\sum_{p_n \leq x} (p_{n+1} - p_n)^2 \ll_\varepsilon x^{1+\nu+\varepsilon}. \tag{2.222}$$

with $\nu = 0.23$. More careful computations with Harman's sieve could improve this upper bound a bit further, but the necessary effort increases rapidly as ν decreases.

For $\nu = 0.23$, the case $a \approx 0.59$ was particularly tricky: For $a \in [0.57, 0.59]$ we constructed a minorant $\rho(n)$ with $\sum_{n\in\mathcal{B}(y)}(1_{\mathbb{P}}(n) - \rho(n)) \approx \frac{0.99y}{x^b \log x}$. This is already dangerously close to the required strict upper bound of $\frac{y}{x^b \log x}$. As we decrease ν, new bad terms appear, but there is not much room left to discard them. For instance, if we were to apply the arguments of Section 2.4 with $\nu = 0.22$ and $a \in [0.57, 0.59]$, we would have to take $\chi_2(a) \subseteq [0.39, 0.45] \cup [0.55, 0.61]$, $\chi_3(a) \subseteq [0.32, 0.415] \cup [0.585, 0.68]$ and $\chi_1(a) \subseteq [0.39, 0.415] \cup [0.585, 0.61]$ and this is likely insufficient input for Harman's sieve. However, we could split $[0.57, 0.59]$ into many shorter intervals and this extends the suitable ranges $\chi_1(a)$, $\chi_2(a)$ and $\chi_3(a)$ by about 0.01 for each smaller interval. A proof of (2.13), using our techniques, with 1.22 in the place of 1.23 and $\tau \in [x^{0.57}, x^{0.59}]$, thus still appears somewhat feasible, although it would require much greater computational effort.

For $\nu = 0.21$, our suitable ranges are already very short, even for fixed a. For example, $a = 0.59$ would require $\chi_2(a) \subseteq [0.385, 0.44] \cup [0.56, 0.615]$, $\chi_3(a) \subseteq [0.33, 0.41] \cup$

$[0.59, 0.67]$ and $\chi_1(a) \subseteq [0.385, 0.41] \cup [0.59, 0.615]$. Thus bound (2.13) with $\nu = 0.21$ appears almost certainly out of reach of our techniques.

We conclude this chapter with the remark that Guth and Maynard [12] very recently proved stronger Dirichlet polynomial bounds in the critical range $\sigma \in [0.7, 0.8]$. These very new results appeared after the submission of this thesis and have therefore not been implemented in our work. However, it appears that Guth and Maynard's new results allow one to break the barrier at $\nu = 0.25$ without the use of Harman's sieve. Further, if used in conjunction with our other methods and bounds, they should enable us to give a significantly improved upper bound on $\sum_{p_n \leq x}(p_{n+1} - p_n)^2$. A bound of $O_\varepsilon(x^{1.2+\varepsilon})$ now certainly appears well within reach.

Chapter 3

Primes in arithmetic progressions to smooth moduli

3.1 Introduction

In this chapter we study equidistribution estimates for primes in arithmetic progressions. For $\theta > 0$ and $\mathcal{Q} \subseteq \mathbb{N}$, consider the following statement: For any $a \in \mathbb{N}$,

$$\sum_{\substack{q \le x^{\theta-\varepsilon} \\ q \in \mathcal{Q} \\ (q,a)=1}} \left| \sum_{\substack{p \le x \\ p \equiv a(q)}} 1 - \frac{1}{\phi(q)} \sum_{\substack{p \le x \\ (p,q)=1}} 1 \right| \ll_{A,\varepsilon} \frac{x}{(\log x)^A}. \tag{3.1}$$

Bombieri [4] and Vinogradov [43] showed in 1965 that (3.1) holds with the exponent $\theta = \frac{1}{2}$ and $\mathcal{Q} = \mathbb{N}$. The Elliott–Halberstam conjecture suggests that (3.1) also holds with $\theta = 1$, but to date (3.1) remains unproven for any $\theta > \frac{1}{2}$ when $\mathcal{Q} = \mathbb{N}$.

Better exponents have been obtained for smaller sets $\mathcal{Q}$: In the 1980s, Bombieri, Fouvry, Friedlander and Iwaniec proved variants of (3.1) in which the absolute value signs were replaced by suitable coefficients. For instance, Fouvry and Iwaniec [8] obtained the exponent of distribution $\frac{9}{17}$ for well-factorable coefficients and Bombieri, Friedlander and Iwaniec [5] improved this exponent to $\frac{4}{7}$. Bombieri, Friedlander and Iwaniec [5] also obtained the exponent $\frac{29}{56}$ for a less restrictive set of coefficients, provided that $\mathcal{Q}$ only contains integers with a divisor of a certain suitable size. More recently, Maynard [32] proved that (3.1) holds with $\theta = \frac{11}{21}$ and $\mathcal{Q} = \{q_1 q_2 : q_1 \le x^{1/21}, q_2 \le x^{10/21-\varepsilon}\}$, provided that the implied constant is allowed to depend on a. Further, Maynard [33] proved that (3.1) holds with $\theta = \frac{1}{2} + \delta$ for $\delta \in (0, \frac{1}{1000})$ when $\mathcal{Q} = \{q_1 q_2 q_3 : q_1 \le Q_1, q_2 \le Q_2, q_3 \le Q_3\}$ with $x^{40\delta} < Q_2 < x^{1/20-7\delta}$ and $x^{1/10+12\delta} Q_2^{-1} < Q_3 < x^{1/10-4\delta} Q_2^{-3/5}$.

In this work, we are interested in the case $\mathcal{Q} = \{q \in \mathbb{N} : q \mid P(x^\delta)\}$ where $P(x^\delta) = \prod_{p<x^\delta} p$. This is the set of squarefree, x^δ-smooth integers. Equidistribution estimates for primes in APs to smooth moduli played a key role in Zhang's proof of bounded gaps between primes [45]. Zhang obtained the exponent of distribution $\frac{1}{2}+\frac{1}{584}$ for a variant of (3.1). Polymath [37] improved on these equidistribution estimates, showing that (3.1) holds with $\theta = \frac{1}{2} + \frac{7}{300}$ when $\mathcal{Q} = \{q \in \mathbb{N} : q \mid P(x^\delta)\}$. We prove the following:

Theorem 3.1. *Let $P(z) = \prod_{p<z} p$. Let $\varepsilon > 0$. There exists $\delta > 0$, dependent on ε, such that the following is true: For every $A > 0$ and every $a \in \mathbb{Z}$, we have*

$$\sum_{\substack{q \leq x^{1/2+1/40-\varepsilon} \\ q \mid P(x^\delta) \\ (q,a)=1}} \left| \sum_{\substack{p\leq x \\ p\equiv a(q)}} 1 - \frac{1}{\phi(q)} \sum_{\substack{p \leq x \\ (p,q)=1}} 1 \right| \ll_{A,\varepsilon} \frac{x}{(\log x)^A}.$$

The key ingredient of our proof is an improved exponential sum bound for what was the limiting case in previous works. By taking advantage of the precise shape of the exponential phases appearing in this problem, we are able to exploit the summation over a variable which was previously treated trivially. A sketch of this argument can be found in Section 3.1.1.

Since equidistribution estimates of the form described in Theorem 3.1 are closely linked to bounds on the infimum limit of gaps between primes, we also study

$$H_m = \liminf_{n\to\infty}(p_{n+m} - p_n).$$

For large m, a first upper bound on this quantity was given by Maynard [30], who showed that $H_m \ll m^3 \exp(4m)$. A similar bound was also obtained by Tao around the same time (unpublished). The methods of Maynard and Tao were further optimized in the Polymath paper [38], leading to the bound $H_m \ll \exp((4 - \frac{28}{157})m)$. Note here that $4 - \frac{28}{157} \approx 3.822$. Baker and Irving [2] combined the techniques of the previous papers with Harman's sieve, which gave $H_m \ll \exp(3.815m)$. Using the method of Polymath together with Theorem 3.1, we are immediately able to deduce that $H_m \ll \exp(3.81m)$, improving on the result of Baker and Irving. However, if we additionally combine our new equidistribution estimates with Harman's sieve, like in [2], we arrive at the following bound:

Theorem 3.2. *The quantity $H_m = \liminf_{n\to\infty}(p_{n+m} - p_n)$ satisfies*

$$H_m \ll \exp(3.8075m).$$

Our new equidistribution estimates can also be used to improve bounds on H_m for fixed small $m \geq 2$. Denoting by $H(k)$ the diameter of the narrowest admissible k-tuple, Polymath [38] showed that

$$\begin{aligned} H_2 &\leq H(35410) \leq 398130, \\ H_3 &\leq H(1649821) \leq 24797814, \\ H_4 &\leq H(75845707) \leq 1431556072, \\ H_5 &\leq H(3473955908) \leq 80550202480. \end{aligned}$$

For given $m \in \{2, 3, 4, 5\}$, we are now able to decrease the size of k for which it is known that $H_m \leq H(k)$. Sutherland computed corresponding bounds on $H(k)$, giving us the following improvement:

Corollary 3.1. *Write $H_m = \liminf_{n\to\infty}(p_{n+m} - p_n)$. Denote by $H(k)$ the diameter of the narrowest admissible k-tuple. Then the following is true:*

$$\begin{aligned} H_2 &\leq H(35265) \leq 396504, \\ H_3 &\leq H(1624545) \leq 24407016, \\ H_4 &\leq H(73807570) \leq 1391051532, \\ H_5 &\leq H(3340375663) \leq 77510685234. \end{aligned}$$

3.1.1 A sketch of the proof of Theorem 3.1

We now outline the main arguments of the proof of Theorem 3.1. Many details are suppressed for clarity.

By Heath-Brown's identity (see Section 1.1.2.1), in order to prove that the primes have exponent of distribution $1/2 + 2\omega$ to smooth moduli, it suffices to show the following: If $\alpha_i : \mathbb{N} \to \mathbb{C}$ with α_i supported on $[N_i, 2N_i]$, $\prod_{i=1}^{j} N_i \asymp x$ and ($\alpha_i(n) = 1$ when $N_i \gg x^{\delta}$), then

$$\sum_{\substack{q \leq x^{1/2+2\omega-\varepsilon} \\ q|P(x^\delta) \\ (q,a)=1}} \left| \sum_{\substack{n\equiv a(q) \\ n\in[x,2x]}} (\alpha_1 \star \cdots \star \alpha_j)(n) - \frac{1}{\phi(q)} \sum_{\substack{(n,q)=1 \\ n\in[x,2x]}} (\alpha_1 \star \cdots \star \alpha_j)(n) \right| \ll_{A,\delta} \frac{x}{(\log x)^A}. \tag{3.2}$$

The current techniques for establishing estimates of the form (3.2) are highly dependent on the values of $N_1, \ldots, N_j$. For instance, if we use the Type I/II/III estimates of Polymath [37] and take $\omega = 7/600$, then (3.2) holds if there are distinct

$i_1, i_2, i_3 \in \{1, \dots, j\}$ with $N_{i_1} N_{i_2} N_{i_3} \geq x^{0.9}$ and $N_{i_j} \gg x^{\delta}$ and (3.2) also holds if there exists $I \subseteq \{1, \dots, j\}$ with $\prod_{i \in I} N_i \in [x^{0.4}, x^{0.6}]$. Every choice of $\alpha_1 \star \cdots \star \alpha_j$ under consideration satisfies one of these two properties, and so Polymath [37] obtained the exponent of distribution $\frac{1}{2} + \frac{7}{300}$. However, if we increase ω just a little further, then the results of Polymath do not provide sufficiently good equidistribution estimates whenever $N_1, \dots, N_5$ are all close to $x^{0.2}$.

To establish a larger exponent of distribution, our goal is to prove that (3.2) still holds for $N_1, \dots, N_j$ with $\prod_{i \in I} N_i \approx x^{0.4}$ if ω is a bit larger than $7/600$. (We thus improve the Type I equidistribution estimates of Polymath [37].) In [37] it was shown that in order to obtain such a result, it is enough to prove that

$$\sum_{d \asymp \Delta} \left| \sum_{n \asymp N} \sum_{\substack{y \asymp Y \\ (y,m)=1}} e_{md}\left(\frac{By}{n + Cd} \right) \right| \ll \frac{\Delta N}{x^{\epsilon}} \tag{3.3}$$

for some choice of Δ with $1 \ll \Delta \ll N$ and some $\epsilon > 0$ and for $N \approx x^{0.4}$, $Y \approx x^{8\omega}$, $m \approx x^{1+8\omega}/(N\Delta)$ and B with $(B, md) = 1$. (This is primarily a consequence of Linnik's dispersion method and completion of sums.) The sum over d will be incorporated into the q-van der Corput process to reduce the modulus to m. The optimal choice of Δ balances the resulting diagonal and off-diagonal terms: larger choices of Δ correspond to better bounds for the off-diagonal terms, but worse bounds for the diagonal terms.

3.1.1.1 The q-van der Corput method - Polymath

We first recap the arguments of Polymath [37], paying particular attention to the diagonal terms. Using the Cauchy-Schwarz inequality, we have

$$\begin{aligned}
\sum_{d \asymp \Delta} \left| \sum_{n \asymp N} \sum_{\substack{y \asymp Y \\ (y,m)=1}} e_{md}\left(\frac{By}{n + Cd} \right) \right| &\ll \sum_{d \asymp \Delta} \sum_{c(d)} \sum_{\substack{y \asymp Y \\ (y,m)=1}} \left| \sum_{\substack{n \asymp N \\ n \equiv c(d)}} e_m\left(\frac{By}{d(n + Cd)} \right) \right| \\
&\ll (\Delta^2 Y)^{1/2} \left(\sum_{d \asymp \Delta} \sum_{c(d)} \sum_{\substack{y \asymp Y \\ (y,m)=1}} \sum_{\substack{n, \widetilde{n} \asymp N \\ n, \widetilde{n} \equiv c(d)}} e_m\left(\frac{By(\widetilde{n} - n)}{d(n + Cd)(\widetilde{n} + Cd)} \right) \right)^{1/2} \\
&\ll (\Delta^2 Y)^{1/2} \left(\sum_{\substack{y \asymp Y \\ (y,m)=1}} \sum_{k \ll \frac{N}{\Delta}} \sum_{d \asymp \Delta} \sum_{n \asymp N} e_m\left(\frac{Byk}{n(n + kd)} \right) \right)^{1/2} \\
&\ll \Delta^{3/2} N^{1/2} Y + \Delta Y \sup_{\substack{y \asymp Y \\ (y,m)=1}} \left(\sum_{1 \leq k \ll \frac{N}{\Delta}} \sum_{d \asymp \Delta} \sum_{n \asymp N} e_m\left(\frac{Byk}{n(n + kd)} \right) \right)^{1/2}.
\end{aligned}$$

Here $\Delta^{3/2}N^{1/2}Y$ is the contribution of the diagonal terms (which correspond to $n = \widetilde{n}$). We seek an upper bound of $\Delta N/x^\epsilon$, so this forces $\Delta < N/Y^2$. For the off-diagonal terms, Polymath [37] used the q-van der Corput process and Deligne's work on the Riemann Hypothesis over finite fields to prove that

$$\sum_{1\le k\ll \frac{N}{\Delta}} \sum_{d\asymp\Delta} \sum_{n\asymp N} e_m\left(\frac{Byk}{n(n+kd)}\right) \ll \frac{Nm^{2/3}}{\Delta^{1/2}}. \tag{3.4}$$

Recall that $m \approx x^{1+8\omega}/(N\Delta)$, $Y \approx x^{8\omega}$ and $\Delta \approx N/Y^2$. To get an upper bound of $\Delta N/x^\epsilon$ on the LHS of (3.3), we require $(Nm^{2/3}/\Delta^{1/2})^{1/2} \ll N/(x^\epsilon Y)$ and hence $N \gg x^{4/17+240\omega/17+12\epsilon/17}$. But recall that we are interested in the case $N \approx x^{0.4}$. This now forces $\omega \le 7/600$.

Larger Δ gave better bounds on the LHS of (3.4), but the maximal possible size of Δ was determined by the diagonal terms, which contributed $\Delta^{3/2}N^{1/2}Y$. So to get an improvement, the key idea is to change the shape of the diagonal terms, reducing their contribution. Observe that the argument above made no use of the summation over y. This summation can be exploited to get smaller diagonal terms.

3.1.1.2 The q-van der Corput method - Variant

Now we introduce our new argument. Notice that if we set $y/n \equiv c \bmod d$, then

$$e_{md}\left(\frac{By}{n+Cd}\right) = e_d\left(\frac{By}{m(n+Cd)}\right) e_m\left(\frac{By}{d(n+Cd)}\right) = e_d\left(\frac{Bc}{m}\right) e_m\left(\frac{By}{d(n+Cd)}\right)$$

and the d-factor no longer depends on n and y, allowing us to apply the q-van der Corput process with both n and y on the inside. More precisely, we have

$$\sum_{d\asymp\Delta}\left|\sum_{n\asymp N}\sum_{y\asymp Y} e_{md}\left(\frac{By}{n+Cd}\right)\right| \ll \sum_{d\asymp\Delta}\sum_{c(d)}\left|\sum_{n\asymp N}\sum_{\substack{y\asymp Y\\ y/n\equiv c(d)}} e_m\left(\frac{By}{d(n+Cd)}\right)\right|$$

$$\ll \Delta\left(\sum_{d\asymp\Delta}\sum_{c(d)}\sum_{\substack{n,\widetilde{n}\asymp N\\ y,\widetilde{y}\asymp Y\\ y/n\equiv\widetilde{y}/\widetilde{n}\equiv c(d)}} e_m\left(\frac{By}{d(n+Cd)} - \frac{B\widetilde{y}}{d(\widetilde{n}+Cd)}\right)\right)^{1/2}$$

$$\ll \Delta\left(\sum_{y,\widetilde{y}\asymp Y}\sum_{d\asymp\Delta}\sum_{\substack{n,\widetilde{n}\asymp N\\ y\widetilde{n}\equiv\widetilde{y}n(d)}} e_m\left(\frac{B(y\widetilde{n}-\widetilde{y}n+C(y-\widetilde{y})d)}{d(n+Cd)(\widetilde{n}+Cd)}\right)\right)^{1/2}.$$

To prove that (3.3) holds, it thus suffices to show that for all $y, \widetilde{y} \asymp Y$,

$$\sum_{d\asymp\Delta}\sum_{\substack{n,\widetilde{n}\asymp N\\ y\widetilde{n}\equiv\widetilde{y}n(d)}} e_m\left(\frac{B(y\widetilde{n}-\widetilde{y}n+C(y-\widetilde{y})d)}{d(n+Cd)(\widetilde{n}+Cd)}\right) \ll \frac{N^2}{x^{2\epsilon}Y^2}. \tag{3.5}$$

Observe that $y\widetilde{n} \equiv \widetilde{y}n(d)$ if and only if $\widetilde{n} = (n\widetilde{y} + kd)/y$ for some $k \ll YN/\Delta$ with $y \mid n\widetilde{y} + kd$. Then

$$e_m\left(\frac{B(y\widetilde{n} - \widetilde{y}n + C(y - \widetilde{y})d)}{d(n + Cd)(\widetilde{n} + Cd)}\right) = e_m\left(\frac{B(k + C(y - \widetilde{y}))}{(n + Cd)(\frac{\widetilde{y}}{y}n + \frac{kd}{y} + Cd)}\right).$$

The diagonal terms of (3.5) now correspond to choices of n, k and d for which $k + C(y-\widetilde{y}) = 0$ and $y \mid n\widetilde{y}+kd$. Fix some $d \asymp \Delta$. The diagonal terms have $k = -C(y-\widetilde{y})$ with $y \mid n\widetilde{y} - dC(y - \widetilde{y})$. Assuming $(y, \widetilde{y}) = 1$, only $O(N/Y)$ choices of n satisfy this divisibility condition. Overall, the diagonal terms thus contribute $O(\Delta N/Y)$ to (3.5). This forces $\Delta < N/Y$. (The value of Δ is thus allowed to be larger than in the work of Polymath [37], where $\Delta < N/Y^2$.)

For the off-diagonal terms, we first notice that the condition $y \mid n\widetilde{y} + kd$ gives $n = yn_1 + F_k d$ for some $n_1 \in \mathbb{N}$ and some constant F_k (dependent on k). We further write $\frac{\widetilde{y}}{y}n + \frac{kd}{y} \equiv \widetilde{y}n_1 + G_k d \bmod m$ and have the relationship $\frac{F_k}{y} - \frac{G_k}{\widetilde{y}} \equiv -\frac{kd}{y\widetilde{y}} \bmod m$. Substituting and using Deligne's bounds for the summation over d and n, we obtain:

$$\begin{aligned}
&\sum_{\substack{k \ll YN/\Delta \\ k+C(y-\widetilde{y}) \neq 0}} \sum_{d \asymp \Delta} \sum_{\substack{n \asymp N \\ y \mid n\widetilde{y}+kd}} e_m\left(\frac{B(k + C(y - \widetilde{y}))}{(n + Cd)(\frac{\widetilde{y}}{y}n + \frac{kd}{y} + Cd)}\right) \\
&= \sum_{\substack{k \ll YN/\Delta \\ k+C(y-\widetilde{y}) \neq 0}} \sum_{d \asymp \Delta} \sum_{n_1 \asymp N} e_m\left(\frac{B(\frac{k}{y\widetilde{y}} + C(\frac{1}{\widetilde{y}} - \frac{1}{y}))}{(n_1 + (\frac{F_k}{y} + \frac{C}{y})d)(n_1 + (\frac{G_k}{\widetilde{y}} + \frac{C}{\widetilde{y}})d)}\right) \ll \frac{YNm}{\Delta}.
\end{aligned}$$

Then (3.5) and hence (3.3) hold provided that $m \ll N\Delta/(x^{2\epsilon}Y^3)$. Recalling that $\Delta \approx N/Y$ and $m \approx x^{1+8\omega}/(N\Delta)$, we require $N \gg x^{1/4+12\omega}$. At $\omega = 1/80$, this condition is satisfied by $N \approx x^{0.4}$, concluding the sketch of our proof.

In this argument, we exploited the simple shape of the quotient $(\frac{By}{n+Cd})$ inside the exponential sums of (3.3). This is what gave us simple expressions for the diagonal terms and what allowed us to easily deal with arising modularity conditions via substitution of linear expressions.

3.1.2 Notation

Below is a list of notational conventions and important definitions.

3.1.2.1 Basics

In this chapter, lower case Roman letters denote integers and upper case Roman letters denote real numbers (unless specified otherwise). In particular, all sums and products are over the integers, with the exception of sums and products over p or p_i, which are over the primes.

We let $h \sim H$ denote the condition $h \in (H, 2H] \cup [2H, H)$. Here H may be positive or negative. (If $a > b$, then $(a, b]$ denotes the empty set.) We also use Vinogradov notation ($\ll$ and $\gg$) and write $X \asymp Y$ if $Y \ll X \ll Y$. If we write $X \ll_A Y$, then the implied constant may depend on A.

For $\delta > 0$, we set $P(x^\delta) = \prod_{p<x^\delta} p$. Divisors of $P(x^\delta)$ are squarefree and x^δ-smooth.

As usually, Dirichlet convolutions are defined as follows: If $\alpha, \beta : \mathbb{N} \to \mathbb{C}$, then

$$(\alpha \star \beta)(n) = \sum_{d|n} \alpha(d)\beta\left(\frac{n}{d}\right).$$

The function $\phi(n)$ is the Euler totient function, $\Lambda(n)$ is the von Mangoldt function and $\mu(n)$ is the Möbius function. If $Q(n)$ is a statement about integers, then $1_{Q(n)}$ is the indicator function of that property.

In all our proofs, we assume without loss of generality that x is very large.

3.1.2.2 Exponentials

For $q \in \mathbb{N}$ and $c \in \mathbb{Z}$, we define $e_q(c) = \exp(2\pi i c/q)$. For $a, b \in \mathbb{Z}/q\mathbb{Z}$,

$$e_q\left(\frac{a}{b}\right) = \begin{cases} e_q(\overline{a/b}) & \text{if } a/b \text{ is well-defined mod } q, \\ 0 & \text{otherwise.} \end{cases}$$

Here $\overline{a/b}$ represents an integer c which satisfies $c \equiv a/b \bmod q$.

By the Chinese Remainder Theorem, if $(q_1, q_2) = 1$, then

$$e_{q_1 q_2}(c) = e_{q_1}\left(\frac{c}{q_2}\right) e_{q_2}\left(\frac{c}{q_1}\right).$$

More generally, for $c \in \mathbb{Z}$ and pairwise coprime natural numbers $q_1, \ldots, q_k$,

$$e_{q_1 \ldots q_k}(c) = \prod_{i=1}^{k} e_{q_i}\left(\frac{c}{\prod_{j\neq i} q_j}\right). \tag{3.6}$$

3.1.2.3 Sequences

For a given sequence, we now formally define the exponent of distribution to smooth moduli, using essentially the same definition as in [2]:

Definition 3.1 (Exponent of distribution). *We say that a function $f : \mathbb{N} \to \mathbb{C}$, supported on $[x, 2x]$, has exponent of distribution θ to smooth moduli if the following holds: For any $\varepsilon > 0$ there exists $\delta > 0$ such that for any $a \in \mathbb{Z}$ and $A > 0$, we have*

$$\sum_{\substack{q \le x^{\theta-\varepsilon} \\ q \mid P(x^\delta) \\ (q,a)=1}} \left| \sum_{n \equiv a(q)} f(n) - \frac{1}{\phi(q)} \sum_{(n,q)=1} f(n) \right| \ll_{A,\varepsilon} \frac{x}{(\log x)^A}.$$

We work with coefficient sequences, as defined in Definition 2.5 of Polymath [37]. For convenience, we recall their definition here, and correct a typo in (iii):

Definition 3.2 (Coefficient sequences). *Given x, a coefficient sequence with respect to x is a finitely supported sequence $\alpha : \mathbb{N} \to \mathbb{R}$ which satisfies $|\alpha(n)| \ll (\tau(n) \log x)^{O(1)}$.*

(i) *α is located at scale N if it is supported on $[cN, CN]$ for some $1 \ll c \ll C \ll 1$.*

(ii) *If α is located at scale N, it is said to have the Siegel-Walfisz property if*

$$\left| \sum_{\substack{n = a(q) \\ (n,r)=1}} \alpha(n) - \frac{1}{\phi(q)} \sum_{(n,qr)=1} \alpha(n) \right| \ll_A \tau(qr)^{O(1)} \frac{N}{(\log x)^A}$$

for any $q, r \ge 1$, any $A > 1$, and any residue class a mod q with $(a,q) = 1$.

(iii) *α is said to be shifted smooth at scale N if there exists a constant $x_0 \in \mathbb{R}$ and a smooth function $\psi : \mathbb{R} \to \mathbb{C}$ supported on $[c, C]$, with $|\psi^{(j)}(t)| \ll_j (\log x)^{O_j(1)}$ for all $t \in \mathbb{R}$, such that*

$$\alpha(n) = \psi\left(\frac{n - x_0}{N}\right).$$

If we can take $x_0 = 0$, α is said to be smooth at scale N.

3.2 Key propositions

We now state our two main propositions and show how Theorem 3.1, Theorem 3.2 and Corollary 3.1 follow from these.

3.2.1 Equidistribution estimates

The following result improves on part (iii) of Theorem 2.8 of [37].

Proposition 3.1. *Let $\omega, \delta, \sigma > 0$. Suppose the following 3 inequalities are satisfied:*

$$\begin{cases} 72\omega + 24\delta < 1, \\ 48\omega + 16\delta + 4\sigma < 1, \\ 64\omega + 20\delta + 2\sigma < 1. \end{cases}$$

Let β be a coefficient sequence at scale N with $N \in [x^{1/2-\sigma}, x^{1/2}]$. Additionally, suppose that β has the Siegel–Walfisz property. Let α be a coefficient sequence at scale M with $M \asymp \frac{x}{N}$.

Then for any $A > 0$ and any $a \in \mathbb{Z}$,

$$\sum_{\substack{q \le x^{1/2+2\omega} \\ q | P(x^\delta) \\ (q,a)=1}} \left| \sum_{n \equiv a(q)} (\alpha \star \beta)(n) - \frac{1}{\phi(q)} \sum_{(n,q)=1} (\alpha \star \beta)(n) \right| \ll_{A,\delta} \frac{x}{(\log x)^A}. \tag{3.7}$$

3.2.1.1 Comment 1: Comparison with Polymath

The bound (3.7) is of a form similar to Polymath's $\text{Type}_I^{(i)}[\omega, \delta, \sigma]$ and $\text{Type}_{II}^{(i)}[\omega, \delta, \sigma]$ estimates (see [37], Definition 2.6). We only made the following modification: In part (iii) of Theorem 2.8 of [37], the sum was taken over $q \in \mathcal{D}_I^{(4)}(x^\delta)$, the set of 4-tuply x^δ-densely divisible integers, whereas we merely sum over x^δ-smooth q. We made this change purely for convenience's sake and our proof could also be adapted to the case $q \in \mathcal{D}_I^{(3)}(x^\delta)$, the set of 3-tuply x^δ-densely divisible integers.

Polymath obtained the equidistribution estimate (3.7) when

$$\begin{cases} 68\omega + 14\delta < 1, \\ \frac{160}{3}\omega + 16\delta + \frac{34}{9}\sigma < 1, \\ 64\omega + 18\delta + 2\sigma < 1. \end{cases} \tag{3.8}$$

In order to prove that the primes have exponent of distribution θ to smooth moduli, one needs to show that (3.7) holds for $\delta = 0$, $\sigma = \frac{1}{10}$ and $\omega < \frac{\theta}{2} - \frac{1}{4}$. Substituting these values into (3.8), the inequality $\frac{160}{3}\omega + 16\delta + \frac{34}{9}\sigma < 1$ forces $\theta \le \frac{1}{2} + \frac{7}{300}$. In order to improve on Polymath's exponent of distribution $\frac{1}{2} + \frac{7}{300}$, we therefore have to weaken the requirement $\frac{160}{3}\omega + 16\delta + \frac{34}{9}\sigma < 1$.

The three conditions in (3.8) are essentially the consequence of various exponential sum bounds and a rough outline of the proof (with $\sigma = 0.1$ and $\delta = 0$) can be found in Section 3.1.1.1. To improve on this result, we use an alternative version of the

q-van der Corput method. A rough sketch of the argument (with $\sigma = 0.1$ and $\delta = 0$) can be found in Section 3.1.1.2. Using this method, we obtain (3.7) if:

$$\begin{cases} 72\omega + 24\delta < 1, \\ 48\omega + 16\delta + 4\sigma < 1, \\ 64\omega + 20\delta + 2\sigma < 1. \end{cases} \tag{3.9}$$

Note in particular that $\frac{160}{3}\omega + 16\delta + \frac{34}{9}\sigma < 1$ has been replaced by $48\omega + 16\delta + 4\sigma < 1$. Substituting in $\delta = 0$, $\sigma = \frac{1}{10}$ and $\omega < \frac{\theta}{2} - \frac{1}{4}$, we obtain the condition $\theta \leq \frac{1}{2} + \frac{1}{40}$, which improves upon the prior requirement $\theta \leq \frac{1}{2} + \frac{7}{300}$.

<u>Caution</u>: So far we have only looked at the second inequality in (3.9). However, the third inequality is now equally important. Substituting $\delta = 0$, $\sigma = \frac{1}{10}$ and $\omega < \frac{\theta}{2} - \frac{1}{4}$ into $64\omega + 20\delta + 2\sigma < 1$, we again get $\theta \leq \frac{1}{2} + \frac{1}{40}$. Therefore the third condition in (3.9) now also prevents further improvements of the exponent of distribution θ. Note that condition $64\omega + 20\delta + 2\sigma < 1$ in (3.9) corresponds to condition $64\omega + 18\delta + 2\sigma < 1$ in (3.8). Ignoring δ, both (3.8) and (3.9) require $64\omega + 2\sigma < 1$ – our new approach did not weaken this requirement. This is not a coincidence: we improved on Polymath by reducing the size of certain diagonal terms in the q-van der Corput method, but the condition $64\omega + 2\sigma < 1$ did not depend on the size of these diagonal terms. This is a roadblock to further improvements of the exponent of distribution to smooth moduli.

As a final remark, note that while we improved on (3.8) when $\delta = 0$, the coefficients of δ in (3.9) are larger than those in (3.8). As such, Polymath's results are better than our results when δ is somewhat larger than ω. This is a rather unfortunate consequence of our amendments to the q-van der Corput method. The size of the coefficients of δ is important when proving explicit bounds on H_2, H_3, H_4 and H_5, and for that reason the larger coefficients of δ have limited our improvements of bounds on H_2, H_3, H_4 and H_5

Using the new conditions (3.9), we now derive Theorem 3.1 from Proposition 3.1.

3.2.1.2 Proof of Theorem 3.1 (assuming Proposition 3.1)

We want to deduce that for any $\varepsilon > 0$ there exists $\delta > 0$ such that for every $A > 0$ and for every $a \in \mathbb{Z}$,

$$\sum_{\substack{q \leq x^{1/2+1/40-\varepsilon} \\ q \mid P(x^\delta) \\ (q,a)=1}} \left| \sum_{\substack{p \leq x \\ p \equiv a(q)}} 1 - \frac{1}{\phi(q)} \sum_{\substack{p \leq x \\ (p,q)=1}} 1 \right| \ll_{A,\varepsilon} \frac{x}{(\log x)^A}.$$

By partial summation and dyadic decomposition, it suffices to show that for given $\varepsilon > 0$, sufficiently small $\delta > 0$ and $a \in \mathbb{Z}$,

$$\sum_{\substack{q \le x^{1/2+1/40-\varepsilon} \\ q \mid P(x^\delta) \\ (q,a)=1}} \left| \sum_{\substack{n \in [x,2x] \\ n \equiv a(q)}} \Lambda(n) - \frac{1}{\phi(q)} \sum_{\substack{n \in [x,2x] \\ (n,q)=1}} \Lambda(n) \right| \ll_{A,\delta} \frac{x}{(\log x)^A}. \tag{3.10}$$

In the terminology of Polymath [38], (3.10) is referred to as a Motohashi-Pintz-Zhang estimate: We ask for MPZ$[\omega, \delta]$ to hold when $\omega = \frac{1}{80} - \frac{\varepsilon}{2}$. Polymath [37] gave the following sufficient condition for when MPZ$[\omega, \delta]$ is true:

Lemma 3.2 (Combinatorial Lemma)**.** *Consider $\omega \in (0, \frac{1}{12})$ and $\delta \in (0, \frac{1}{4} + \omega)$ with $68\omega + 14\delta < 1$. Suppose there exists $\sigma \in (\frac{1}{10}, \frac{1}{2})$ with $\frac{\sigma}{2} > \omega$ and $\sigma > \frac{1}{18} + \frac{28}{9}\omega + \frac{2}{9}\delta$, such that* (3.7) (*the conclusion of Proposition* 3.1) *holds for the given ω, δ and σ.*

Then MPZ$[\omega, \delta]$ *is true.*

Proof. This follows from the proof of Lemma 2.7 of [37] by applying part (iv) and (v) of Theorem 2.8 of [37] and additionally replacing $\mathcal{Q} = \{q \le Q : q \in \mathcal{D}_I^{(i)}(x^\delta)\}$ with $\mathcal{Q} = \{q \le Q : q \mid P(x^\delta)\}$ everywhere. □

Using Proposition 3.1 and Lemma 3.2, we deduce:

Corollary 3.3. *Let $\omega > 0$ and $\delta > 0$ with $80\omega + \frac{80}{3}\delta < 1$. Then* MPZ$[\omega, \delta]$ *holds.*

Proof. By Proposition 3.1, (3.7) holds if $72\omega + 24\delta < 1$ and $48\omega + 16\delta + 4\sigma < 1$ and $64\omega + 20\delta + 2\sigma < 1$. The second and third inequality rearrange to $\sigma < \frac{1}{4} - 12\omega - 4\delta$ and $\sigma < \frac{1}{2} - 32\omega - 10\delta$. By Lemma 3.2, MPZ$[\omega, \delta]$ thus holds provided that $72\omega + 24\delta < 1$ and provided the following inequality is satisfied:

$$\max\left\{\frac{1}{18} + \frac{28\omega}{9} + \frac{2\delta}{9}, \frac{1}{10}\right\} < \min\left\{\frac{1}{4} - 12\omega - 4\delta, \frac{1}{2} - 32\omega - 10\delta\right\}.$$

After some careful computations, this condition rearranges to $80\omega + \frac{80}{3}\delta < 1$. □

Proof of Theorem 3.1 (*assuming Prop.* 3.1) – *Final Step:* Recall that (3.10) corresponds to MPZ$[\frac{1}{80} - \frac{\varepsilon}{2}, \delta]$. Hence we have the desired bound (3.10) whenever ε and δ satisfy $80(\frac{1}{80} - \frac{\varepsilon}{2}) + \frac{80}{3}\delta = 1 - 40\varepsilon + \frac{80}{3}\delta < 1$. Choosing $\delta = \varepsilon$, this inequality holds. Theorem 3.1 follows. □

In [37] and [38] it was shown and used that MPZ$[\omega, \delta]$ holds when $\frac{600\omega}{7} + \frac{180\delta}{7} < 1$. Our alternative criterion $80\omega + \frac{80}{3}\delta < 1$, stated in Corollary 3.3, is an improvement over Polymath's result when $\omega > \frac{\delta}{6}$.

3.2.1.3 A first improvement of bounds on H_m

We recall now from [38] some of the main steps of Maynard's and Tao's method for bounding H_m: First, if the weak Dickson-Hardy-Littlewood prime tuples conjecture DHL$[k, m+1]$ holds, then $H_m \le H(k)$, where $H(k)$ is the diameter of the narrowest admissible k-tuple. By Theorem 3.10 of [38], DHL$[k, m+1]$ holds if MPZ$[\omega, \delta]$ holds and $M_k^{[\frac{\delta}{1/4+\omega}]} > \frac{m}{1/4+\omega}$, where $M_k^{[\alpha]}$ is as defined in equation (34) of [38]. By the proof of Theorem 6.7 of [38], $M_k^{[\alpha]} > \log k - O_\alpha(1)$. Finally, by Theorem 3.3 of [38], $H(k) \le k \log k + k \log\log k + o(k)$. Combining all this with Corollary 3.3, we have

$$H_m \ll_\epsilon \exp\left(\frac{(1+\epsilon)m}{(1/4+1/80)}\right).$$

In particular, we have $H_m \ll \exp(3.81m)$, as claimed in the introduction.

3.2.1.4 Proof of Corollary 3.1 (assuming Proposition 3.1)

To get bounds on $H_2, \ldots, H_5$, we again follow the arguments of Polymath [38].

In particular, we use Theorem 6.7 of [38] with

$$(k, c\log k, T\log k) = \begin{cases} (35265, 0.99479, 0.85213) & \text{for } m = 2, \\ (1624545, 1.00422, 0.80148) & \text{for } m = 3, \\ (73807570, 1.00712, 0.77003) & \text{for } m = 4, \\ (3340375663, 1.0079318, 0.7490925) & \text{for } m = 5, \end{cases}$$

to deduce that $M_k^{[\frac{\delta}{1/4+\omega}]} > \frac{m}{1/4+\omega}$ holds for the following choices of ω and δ:

$$(\omega, \delta) = \begin{cases} (0.00556625, 0.0207987) & \text{for } (m,k) = (2, 35265), \\ (0.00768602, 0.0144419) & \text{for } (m,k) = (3, 1624545), \\ (0.00883292, 0.0110012) & \text{for } (m,k) = (4, 73807570), \\ (0.0095447064335, 0.0088658806979) & \text{for } (m,k) = (5, 3340375663). \end{cases}$$

Each of these choices of ω and δ satisfies $80\omega + \frac{80}{3}\delta < 1$, so that MPZ$[\omega, \delta]$ holds. Thus DHL$[35265, 3]$, DHL$[1624545, 4]$, DHL$[73807570, 5]$ and DHL$[3340375663, 6]$ are true. In particular, $H_2 \le H(35265)$, $H_3 \le H(1624545)$, $H_4 \le H(73807570)$ and $H_5 \le H(3340375663)$. Sutherland computed the following:

$$H(35265) \le 396504,$$
$$H(1624545) \le 24407016,$$
$$H(73807570) \le 1391051532,$$

$$H(3340375663) \leq 77510685234.$$

For $k \in \{35265, 1624545, 73807570\}$, a link to admissible k-tuples of lengths 396504, 24407016 and 1391051532 can be found on page 7 of the arXiv preprint [42] corresponding to this thesis chapter. To compute bound $H(3340375663) \leq 77510685234$, the shifted Schinzel sieve was used as described in Section 10.2.2 of Polymath [38], with $[s, s+x] = [3343896484, 80854581718]$ and $m = 13090609$. This concludes the proof of Corollary 3.1. $\square$

3.2.1.5 Comment 2: Proof outline for Proposition 3.1

As highlighted above, Proposition 3.1 is really the key result of Chapter 3. In Section 3.1.1 we already gave a sketch of the most important ideas of Chapter 3, and hence of the proof of Proposition 3.1. However, the proof outline glossed over many important technical details. A rigorous account of the key steps of the proof of Proposition 3.1 can be found in the following places:

Polymath ([37], Section 5) used Linnik's dispersion method and completion of sums to reduce the proof of equidistribution estimates of the form given in (3.7) to bounding exponential sums of the form given in (3.3). The precise shape of these exponential sums can be found in our Lemma 3.5, where we summarize the results of the first steps of the Polymath paper.

After handling some small technical details, the q-van der Corput method is then applied in Lemma 3.8. In Lemma 3.9 we remove the resulting diagonal terms. To bound the off-diagonal terms, we wish to apply the exponential sum bounds of ([37], Section 8). Initially, our expressions do not have the correct form and we spend several lemmas removing various divisibility conditions and linear congruence conditions. In Lemma 3.19, we are finally ready to apply the exponential sum bounds of Polymath and complete the proof of Proposition 3.1.

3.2.2 Harman's sieve

We now use Harman's sieve to further improve our bounds on H_m. Section 3.4 is focused on the construction of a suitable minorant for $1_{\mathbb{P}}(n)$, using as input the Type I/II equidistribution estimate given in Proposition 3.1 and the Type III equidistribution estimate given in part (v) of Theorem 2.8 of [37]. Write

$$\psi(n, y) = \begin{cases} 1 & \text{if } p \mid n \rightarrow p \geq y, \\ 0 & \text{otherwise.} \end{cases}$$

We will prove the following:

Proposition 3.2. *Fix some large x, write $p_j = x^{\alpha_j}$ and set*

$$\rho(n) = 1_{\mathbb{P}}(n) \quad - \sum_{\substack{n=p_1p_2p_3p_4n_5 \\ 0.19038\leq\alpha_4<\alpha_3<\alpha_2<\alpha_1<0.40481 \\ \alpha_1+\alpha_2<0.40481 \\ \alpha_2+\alpha_3+\alpha_4>0.59519}} \psi(n_5, p_4) \quad - \sum_{\substack{n=p_2p_3p_4p_5p_6 \\ 0.19038\leq\alpha_2,\alpha_3,\alpha_4,\alpha_5,\alpha_6\leq 0.23848 \\ \alpha_2,\alpha_4>\alpha_3 \\ \alpha_2+\alpha_4<0.40481 \\ \alpha_2+\alpha_3+\alpha_5>0.59519 \\ \alpha_6\geq\alpha_5}} 1.$$

Then $\rho(n)$ is a minorant for $1_{\mathbb{P}}(n)$ on $[x, 2x]$, has exponent of distribution 0.5253 to smooth moduli, and

$$\sum_{n\in[x,2x]} \rho(n) \geq (1 - 2\cdot 10^{-5} + o(1))\frac{x}{\log x}.$$

3.2.2.1 Proof of Theorem 3.2 (assuming Proposition 3.2)

Baker and Irving [2] combined results of [38] with Harman's sieve as follows:

Lemma 3.4 (Baker, Irving). *Fix a small $\xi > 0$ and $c_1 \in (0,1)$. Suppose that for every large x, there exists a function $\rho : [x, 2x] \to \mathbb{R}$ with the following properties:*

(i) *$\rho(n)$ is a minorant for the indicator function of primes, that is, $\rho(n) \leq 1_{\mathbb{P}}(n)$.*

(ii) *If $\rho(n) \neq 0$, then all prime factors of n exceed x^{ξ}.*

(iii) *The function $\rho(n)$ has exponent of distribution θ to smooth moduli.*

(iv) *The sum of $\rho(n)$ over $[x, 2x]$ satisfies*

$$\sum_{n\in[x,2x]} \rho(n) = (1 - c_1 + o(1))\frac{x}{\log x}.$$

Then we have, for every $\epsilon > 0$,

$$H_m = \liminf_{n\to\infty}(p_{n+m} - p_n) \ll_{\epsilon} \exp\left(\left(\frac{2(1+\epsilon)}{\theta(1-c_1)}\right) m\right).$$

Proof. This is Lemma 1 of [2]. □

Lemma 3.4 and Proposition 3.2 give $H_m \ll \exp(3.8075m)$, concluding the proof of Theorem 3.2.

3.3 Equidistribution estimates

In this section we consider a coefficient sequence α at scale M and a coefficient sequence β at scale N with $x \ll MN \ll x$. We assume that β has the Siegel–Walfisz property. We write $N = x^\gamma$. Our aim is to prove the following equidistribution estimate for certain $\omega > 0$, $\delta > 0$ and $\gamma \in (12\omega + 6\delta, \frac{1}{2} - 2\omega - 8\varepsilon)$: For every $A \geq 1$ and every integer $a \in \mathbb{Z}$,

$$\sum_{\substack{q \leq x^{1/2+2\omega} \\ q \mid P(x^\delta) \\ (q,a)=1}} \left| \sum_{n \equiv a(q)} (\alpha \star \beta)(n) - \frac{1}{\phi(q)} \sum_{(n,q)=1} (\alpha \star \beta)(n) \right| \ll_{A,\delta} \frac{x}{(\log x)^A}. \tag{3.11}$$

3.3.1 A summary of Polymath's results

We begin by summarizing some of the results of Section 5 and Section 8.1 of Polymath [37]. These results provide a starting point for our proof.

Lemma 3.5 (Reduction of Type I bounds to exponential sums [37]). *Let $\omega, \delta > 0$. Let $\varepsilon \in (0, 10^{-100}\delta)$. Let $N = x^\gamma$, where $\gamma \in (12\omega + 6\delta, \frac{1}{2} - 2\omega - 8\varepsilon)$.*
Let α and β be coefficient sequences at scales M and N with $x \ll MN \ll x$. Assume β has the Siegel-Walfisz property.

Denote by Z_1 the set of tuples $(Q, R, U, V, H^*, q_0, \ell, a, b_1, b_2, \psi_M, \psi_N)$ which satisfy the following properties: Q, R, U and V are positive. $q_0 \in \mathbb{N}$ with $q_0 \ll Q$ and $(q_0, ab_1b_2) = 1$. ℓ satisfies $0 \neq |\ell| \ll \frac{N}{R}$. ψ_M and ψ_N are real, non-negative coefficient sequences, smooth at scales M and N. Further, for $H = x^\varepsilon RQ^2(q_0M)^{-1}$, the quantities Q, R, U, V, H^* satisfy $H \geq 1$, $1 \ll |H^*| \ll H$ and

$$x^{-4\varepsilon-\delta}N \ll R \ll x^{-2\varepsilon}N, \tag{3.12}$$

$$x^{1/2-\varepsilon} \ll QR \ll x^{1/2+2\omega+\varepsilon}, \tag{3.13}$$

$$q_0^{-1}x^{-\delta-5\varepsilon}Q/H \ll U \ll q_0^{-1}x^{-5\varepsilon}Q/H, \tag{3.14}$$

$$x^{5\varepsilon}H \ll V \ll x^{\delta+5\varepsilon}H, \tag{3.15}$$

$$UV \asymp Q/q_0. \tag{3.16}$$

For a given $z \in Z_1$, we define the indicator function $C_0(n) = 1_{\frac{b_1}{n} \equiv \frac{b_2}{n+\ell r}(q_0)}$. For $h \ll H$, we set $\varphi^\star_H(h) = \frac{1}{M}\sum_m \psi_M(m) e(\frac{-mh}{rq_0u_1v_1q_2})$ and $\varphi^{\star\star}_H(h) = \frac{1}{M}\sum_m \psi_M(m) e(\frac{-mh}{rq_0u_1v_2q_2})$.

$$\Psi(n, h_1, h_2, r, u_1, v_1, v_2, q_2) = e_r\left(\frac{ah_1}{nq_0u_1v_1q_2}\right) e_{q_0u_1v_1}\left(\frac{b_1h_1}{nrq_2}\right) e_{q_2}\left(\frac{b_2h_1}{(n+\ell r)rq_0u_1v_1}\right)$$

$$e_r\bigg(-\frac{ah_2}{nq_0u_1v_2q_2}\bigg)e_{q_0u_1v_2}\bigg(-\frac{b_1h_2}{nrq_2}\bigg)e_{q_2}\bigg(-\frac{b_2h_2}{(n+\ell r)rq_0u_1v_2}\bigg).$$

For $z \in Z_1$ and $C_0(n)$, $\varphi_H^\star(h)$, $\varphi_H^{\star\star}(h)$ and $\Psi(n,h_1,h_2,r,u_1,v_1,v_2,q_2)$ as above, we set

$$\Sigma_1(z) = \sum_{\substack{r,u_1,v_1,v_2,q_2\\ r\asymp R\\ u_1\asymp U\\ v_1,v_2\asymp V\\ q_2\asymp Q/q_0\\ rq_0u_1[v_1,v_2]q_2|P(x^\delta)\\ (rq_0u_1v_1v_2q_2,ab_1b_2)=1}}\bigg|\sum_{\substack{h_1,h_2\\ h_1\sim H^*\\ h_2\sim H^*}}\sum_{\substack{n\\ (n,rq_0u_1v_1v_2)=1\\ (n+\ell r,q_0q_2)=1}} C_0(n)\varphi_H^\star(h_1)\overline{\varphi_H^{\star\star}(h_2)}\psi_N(n)\Psi(n,h_1,h_2,r,u_1,v_1,v_2,q_2)\bigg|.$$

Suppose that $\Sigma_1(z) = O_\varepsilon((q_0,\ell)RQNUV^2x^{-4\varepsilon})$ for every $z \in Z_1$. Then for $A > 0$,

$$\sum_{\substack{q\le x^{1/2+2\omega}\\ q|P(x^\delta)\\ (q,a)=1}}\bigg|\sum_{n\equiv a(q)}(\alpha\star\beta)(n) - \frac{1}{\phi(q)}\sum_{(n,q)=1}(\alpha\star\beta)(n)\bigg| \ll_{A,\varepsilon} \frac{x}{(\log x)^A}.$$

This lemma follows almost directly from the arguments of Section 5 and Section 8.1 of [37]. Some small changes have to be made, so for completeness' sake a full proof of Lemma 3.5, starting from the point at which we first diverge from [37], can be found in the appendix of this thesis.

3.3.2 The q-van der Corput method

We now prepare to carry out the variant of the q-van der Corput method which was described in Section 3.1.1.2.

3.3.2.1 Diagonal terms + Splitting r

Our first lemma has two main aims: On the one hand, we remove certain inconvenient terms from our sums. (To be precise, we remove combinations of h_1, h_2, v_1, v_2 with $h_1v_2 = h_2v_1$.) On the other hand, we split up r, writing $r = dr_1$ for d and r_1 of a certain good size. The factor d will later be incorporated into the q-van der Corput method to reduce the modulus.

Lemma 3.6 (Extraction of a suitable factor d). *Let Z_1 and Σ_1 be as described in Lemma 3.5. For a given $z_1 \in Z_1$, we set*

$$D = \frac{N}{x^{50\varepsilon}H^2}. \tag{3.17}$$

Denote by $Z_2(z_1)$ the set of tuples

$$(\Delta, r_1, u_1, v_1, v_2, q_2, d_0, \psi_{\Delta_1})$$

which satisfy the following properties: Δ *is a positive real number with* $x^{-\delta}D \ll \Delta \ll D$ *and* $\psi_{\Delta_1}(d_1)$ *is a real, non-negative coefficient sequence, smooth at scale* $\Delta_1 = \Delta/(x^{5\varepsilon})$. *Further,* $r_1, u_1, v_1, v_2, q_2, d_0 \in \mathbb{N}$ *with* $u_1 \asymp U$, $v_1, v_2 \asymp V$, $d_0 \asymp \Delta$, $r_1 \asymp \frac{R}{\Delta}$, $q_2 \asymp \frac{Q}{q_0}$, $r_1 q_0 u_1 [v_1, v_2] q_2 \mid P(x^\delta)$ *and* $(r_1 q_0 u_1 v_1 v_2 q_2, a b_1 b_2) = 1$.

Write $C(n) = 1_{\frac{b_1}{n} \equiv \frac{b_2}{n+\ell d r_1}(q_0)}$. *For given* $z_1 \in Z_1$, $z_2 \in Z_2(z_1)$, *we set* $\Sigma_2(z_1, z_2)$ *equal*

$$\sum_{\substack{d_1 \\ (d, r_1 q_0 u_1)=1 \\ (d, [v_1, v_2] q_2)=1 \\ d \text{ squarefree} \\ \text{where } d = d_0 + d_1}} \psi_{\Delta_1}(d_1) \Bigg| \sum_{\substack{h_1, h_2 \\ h_1 \sim H^* \\ h_2 \sim H^* \\ h_1 v_2 \neq h_2 v_1}} \sum_{\substack{n \\ (n, d r_1 q_0)=1 \\ (n, u_1 v_1 v_2)=1 \\ (n+\ell d r_1, q_0)=1 \\ (n+\ell d r_1, q_2)=1}} C(n) \varphi_H^\star(h_1) \overline{\varphi_H^{\star\star}(h_2)} \psi_N(n) \Psi(n, h_1, h_2, d r_1, u_1, v_1, v_2, q_2) \Bigg|.$$

If $\Sigma_2(z_1, z_2) = O_\varepsilon\left(\dfrac{q_0 (q_0, \ell) \Delta N (v_1, v_2)}{x^{10\varepsilon}}\right)$ *for all* $z_2 \in Z_2(z_1)$, *then*

$$\Sigma_1(z_1) = O_\varepsilon\left(\frac{(q_0, \ell) R Q N U V^2}{x^{4\varepsilon}}\right).$$

Proof. Our aim is to bound $\Sigma_1(z_1)$, which is given by

$$\sum_{\substack{r, u_1, v_1, v_2, q_2 \\ r \asymp R, u_1 \asymp U \\ v_1, v_2 \asymp V, q_2 \asymp Q/q_0 \\ r q_0 u_1 [v_1, v_2] q_2 \mid P(x^\delta) \\ (r q_0 u_1 v_1 v_2 q_2, a b_1 b_2)=1}} \Bigg| \sum_{\substack{h_1, h_2 \\ h_1 \sim H^* \\ h_2 \sim H^*}} \sum_{\substack{n \\ (n, r q_0 u_1 v_1 v_2)=1 \\ (n+\ell r, q_0 q_2)=1}} C_0(n) \varphi_H^\star(h_1) \overline{\varphi_H^{\star\star}(h_2)} \psi_N(n) \Psi(n, h_1, h_2, r, u_1, v_1, v_2, q_2) \Bigg|.$$

We now factorise r. Recall that $R \gg x^{-4\varepsilon-\delta} N$. Hence $D \leq x^{-50\varepsilon} N \ll x^{-46\varepsilon+\delta} R$. On the other hand,

$$H = \frac{x^\varepsilon R Q^2}{q_0 M} = \frac{x^\varepsilon (RQ)^2}{q_0 M R} \ll \frac{(x^{1+4\omega+3\varepsilon})(x^{4\varepsilon+\delta})}{q_0 M N} \ll \frac{x^{4\omega+\delta+7\varepsilon}}{q_0} \tag{3.18}$$

and so $D = x^{-50\varepsilon} N / H^2 \gg x^{-8\omega-2\delta-64\varepsilon} N$. Using that $\gamma > 12\omega + 6\delta$, then $D \gg 1$. For $r \in \mathbb{N}$ with $r \mid P(x^\delta)$ and $r \asymp R$, we can thus find d and r_1 such that $r = d r_1$ with $x^{-\delta} D \leq d \leq D$. We define a quantity $\Upsilon_0(\Delta, r_1, u_1, q_2)$ as follows:

$$\Upsilon_0 = \sum_{\substack{d, v_1, v_2 \\ d \asymp \Delta \\ v_1, v_2 \asymp V}}^{*} \Bigg| \sum_{\substack{h_1, h_2 \\ h_1 \sim H^* \\ h_2 \sim H^*}} \sum_{\substack{n \\ (n, d r_1 q_0 u_1 v_1 v_2)=1 \\ (n+\ell d r_1, q_0 q_2)=1}} C(n) \varphi_H^\star(h_1) \overline{\varphi_H^{\star\star}(h_2)} \psi_N(n) \Psi(n, h_1, h_2, d r_1, u_1, v_1, v_2, q_2) \Bigg|.$$

(Here $\sum^*$ indicates that the outer sum is restricted to d, v_1, v_2 with $d r_1 q_0 u_1 [v_1, v_2] q_2 \mid P(x^\delta)$ and $(d r_1 q_0 u_1 v_1 v_2 q_2, a b_1 b_2) = 1$.) Then by dyadic decomposition,

$$\Sigma_1(z_1) \ll_\varepsilon \left(\frac{x^{0.5\varepsilon} R U Q}{q_0 \Delta}\right) \sup \Upsilon_0(\Delta, r_1, u_1, q_2),$$

where the supremum is taken over all choices of Δ with $x^{-\delta}D \ll \Delta \ll D$ and all choices of r_1, u_1, q_2 with $r_1 \asymp \frac{R}{\Delta}$, $u_1 \asymp U$, $q_2 \asymp \frac{Q}{q_0}$ and $(r_1u_1q_2, ab_1b_2) = 1$. To show that $\Sigma_1(z_1) \ll_\varepsilon x^{-4\varepsilon}(q_0, \ell)RQNUV^2$, it thus suffices to show that $\sup \Upsilon_0(\Delta, r_1, u_1, q_2) \ll_\varepsilon x^{-4.5\varepsilon}q_0(q_0, \ell)\Delta NV^2$.

Next we wish to exclude choices of h_i and v_i for which $h_1v_2 - h_2v_1 = 0$. By the Triangle inequality, $\Upsilon_0 \le \Upsilon_1 + \Upsilon_2$, where Υ_1 and Υ_2 are given by

$$\Upsilon_1 = \sideset{}{^*}\sum_{\substack{d,v_1,v_2\\ d\asymp\Delta\\ v_1,v_2\asymp V}} \Bigg| \sum_{\substack{h_1,h_2\\ h_1,h_2\sim H^*\\ h_1v_2-h_2v_1=0}} \sum_{\substack{n\\ (n,dr_1q_0u_1v_1v_2)=1\\ (n+\ell dr_1,q_0q_2)=1}} C(n)\varphi_H^\star(h_1)\overline{\varphi_H^{\star\star}(h_2)}\psi_N(n)\Psi(n,h_1,h_2,dr_1,u_1,v_1,v_2,q_2)\Bigg|,$$

$$\Upsilon_2 = \sideset{}{^*}\sum_{\substack{d,v_1,v_2\\ d\asymp\Delta\\ v_1,v_2\asymp V}} \Bigg| \sum_{\substack{h_1,h_2\\ h_1,h_2\sim H^*\\ h_1v_2-h_2v_1\neq0}} \sum_{\substack{n\\ (n,dr_1q_0u_1v_1v_2)=1\\ (n+\ell dr_1,q_0q_2)=1}} C(n)\varphi_H^\star(h_1)\overline{\varphi_H^{\star\star}(h_2)}\psi_N(n)\Psi(n,h_1,h_2,dr_1,u_1,v_1,v_2,q_2)\Bigg|.$$

Recall that $H^* \ll |H| \ll x^{-5\varepsilon}V$. A trivial bound on Υ_1 is thus given by

$$\Upsilon_1 \ll x^{o(1)}\Delta H^*VN \ll x^{-5\varepsilon+o(1)}\Delta NV^2.$$

To get the desired bound on $\Sigma_1(z_1)$, it thus suffices to show that $\Upsilon_2 \ll_\varepsilon \frac{q_0(q_0,\ell)\Delta NV^2}{x^{4.5\varepsilon}}$. Consider the quantity $\Upsilon_2'(v_1, v_2)$, defined to equal

$$\sum_{\substack{d\asymp\Delta\\ (d,r_1q_0u_1)=1\\ (d,[v_1,v_2]q_2)=1\\ d \text{ squarefree}}} \Bigg| \sum_{\substack{h_1,h_2\\ h_1,h_2\sim H^*\\ h_1v_2-h_2v_1\neq0}} \sum_{\substack{n\\ (n,dr_1q_0u_1v_1v_2)=1\\ (n+\ell dr_1,q_0q_2)=1}} C(n)\varphi_H^\star(h_1)\overline{\varphi_H^{\star\star}(h_2)}\psi_N(n)\Psi(n,h_1,h_2,dr_1,u_1,v_1,v_2,q_2)\Bigg|$$

for v_1, v_2 with $r_1q_0u_1[v_1,v_2]q_2 \mid P(x^\delta)$ and $(v_1v_2, ab_1b_2) = 1$. (Note here that we removed the conditions $d \mid P(x^\delta)$ and $(d, ab_1b_2) = 1$. This increases the sum.) Now

$$\Upsilon_2 \ll \sum_{v_1\asymp V}\sum_{v_2\asymp V} \Upsilon_2'(v_1, v_2).$$

For technical reasons, it will later be important to have a shorter summation range for d. Currently we are looking at $d \asymp \Delta$. Recall that $\Delta_1 = \Delta/(x^{5\varepsilon})$. We partition the set of d with $d \asymp \Delta$ into sets of the form $\{d_0 + d_1 : \Delta_1 < d_1 \le 2\Delta_1\}$. We only need $O(x^{5\varepsilon})$ choices of d_0 to cover the range $d \asymp \Delta$. Define $\Upsilon_3(v_1, v_2, d_0)$ as follows:

$$\Upsilon_3 = \sum_{\substack{d_1\\ (d,r_1q_0u_1)=1\\ (d,[v_1,v_2]q_2)=1\\ d \text{ squarefree}\\ \text{where } d=d_0+d_1}} \psi_{\Delta_1}(d_1)\Bigg| \sum_{\substack{h_1,h_2\\ h_1,h_2\sim H^*\\ h_1v_2\neq h_2v_1}} \sum_{\substack{n\\ (n,dr_1q_0)=1\\ (n,u_1v_1v_2)=1\\ (n+\ell dr_1,q_0)=1\\ (n+\ell dr_1,q_2)=1\\ \text{where } d=d_0+d_1}} C(n)\varphi_H^\star(h_1)\overline{\varphi_H^{\star\star}(h_2)}\psi_N(n)\Psi(n,h_1,h_2,dr_1,u_1,v_1,v_2,q_2)\Bigg|.$$

Here $d_0 \asymp \Delta$ and $\psi_{\Delta_1}(d_1)$ is a coefficient sequence which is smooth at scale Δ_1 and has $\psi_{\Delta_1}(d_1) \geq 0$ for all $d_1 \in \mathbb{R}$ and $\psi_{\Delta_1}(d_1) \geq 1$ for all $d_1 \in [\Delta_1, 2\Delta_1]$. We then have

$$\Upsilon_2'(v_1, v_2) \ll x^{5\varepsilon} \sup \Upsilon_3(v_1, v_2, d_0).$$

Now observe that $\Upsilon_3(v_1, v_2, d_0)$ is of the form $\Sigma_2(z_1, z_2)$ for some $z_2 \in Z_2(z_1)$. Hence, by the assumptions of Lemma 3.6, we have $\Upsilon_3(v_1, v_2, d_0) \ll_\varepsilon x^{-10\varepsilon} q_0(q_0, \ell)\Delta N(v_1, v_2)$. This in turn implies $\Upsilon_2'(v_1, v_2) \ll_\varepsilon x^{-5\varepsilon} q_0(q_0, \ell)\Delta N(v_1, v_2)$. Summing over $v_1, v_2 \asymp V$, we then obtain the desired bound

$$\Upsilon_2 \ll_\varepsilon x^{-5\varepsilon} q_0(q_0, \ell)\Delta N \sum_{v_1 \asymp V} \sum_{v_2 \asymp V} (v_1, v_2) \ll x^{-4.5\varepsilon} q_0(q_0, \ell)\Delta N V^2.$$

Hence $\Sigma_2(z_1, z_2) = O_\varepsilon(\frac{q_0(q_0,\ell)\Delta N(v_1,v_2)}{x^{10\varepsilon}})$ indeed implies $\Sigma_1(z_1) = O_\varepsilon(\frac{(q_0,\ell)RQNUV^2}{x^{4\varepsilon}})$, as proposed. $\square$

3.3.2.2 Simplifying exponentials

We now inspect $\Psi(n, h_1, h_2, dr_1, u_1, v_1, v_2, q_2)$ more closely.

Lemma 3.7 (Factorising Ψ). *Consider $n, d, h_1, h_2, \ell, r_1, q_0, u_1, v_1, v_2, q_2, a, b_1, b_2 \in \mathbb{Z}$. Assume that $r_1 q_0 u_1 [v_1, v_2] q_2 \mid P(x)$ and $(r_1 q_0 u_1 v_1 v_2 q_2, dab_1 b_2) = (n, dr_1 q_0 u_1 v_1 v_2) = (n + \ell dr_1, q_0 q_2) = 1$. Let $\Psi(n, h_1, h_2, dr_1, u_1, v_1, v_2, q_2)$ be as defined in Lemma* 3.5.

Write $v_1^* = v_1/(v_1, v_2)$ and $v_2^* = v_2/(v_1, v_2)$. Write $v^* = (v_1, v_2)$. Then

$$\Psi(n, h_1, h_2, dr_1, u_1, v_1, v_2, q_2) = e_d\left(\frac{a(h_1 v_2^* - h_2 v_1^*)}{n r_1 q_0 u_1 v^* v_1^* v_2^* q_2}\right) e_m\left(\frac{A(h_1 v_2^* - h_2 v_1^*)}{d(n + Bd)}\right),$$

where $m = r_1 q_0 u_1 v^* v_1^* v_2^* q_2 = r_1 q_0 u_1 [v_1, v_2] q_2$ and where A and B are residue classes $\bmod m$ which are independent of n, d, h_1 and h_2. Furthermore, A satisfies $(A, m) = 1$.

Proof. Using the Chinese Remainder Theorem as stated in (3.6), we find the following:

$$\begin{aligned}
&\Psi(n, h_1, h_2, dr_1, \dots) \\
&= e_{dr_1}\left(\frac{ah_1}{nq_0 u_1 v_1 q_2} - \frac{ah_2}{nq_0 u_1 v_2 q_2}\right) e_{q_0 u_1 v_1}\left(\frac{b_1 h_1}{ndr_1 q_2}\right) e_{q_0 u_1 v_2}\left(-\frac{b_1 h_2}{ndr_1 q_2}\right) \\
&e_{q_2}\left(\frac{b_2 h_1}{(n + \ell dr_1)dr_1 q_0 u_1 v_1} - \frac{b_2 h_2}{(n + \ell dr_1)dr_1 q_0 u_1 v_2}\right) \\
&= e_d\left(\frac{a(h_1 v_2 - h_2 v_1)}{nr_1 q_0 u_1 v_1 v_2 q_2}\right) e_{r_1}\left(\frac{a(h_1 v_2 - h_2 v_1)}{ndq_0 u_1 v_1 v_2 q_2}\right) e_{v^*}\left(\frac{b_1(h_1 v_2^* - h_2 v_1^*)}{ndr_1 q_0 u_1 v_1^* v_2^* q_2}\right) \\
&e_{q_0 u_1 v_1^*}\left(\frac{b_1 h_1}{ndr_1 v^* q_2}\right) e_{q_0 u_1 v_2^*}\left(-\frac{b_1 h_2}{ndr_1 v^* q_2}\right) e_{q_2}\left(\frac{b_2(h_1 v_2 - h_2 v_1)}{(n + \ell dr_1)dr_1 q_0 u_1 v_1 v_2}\right)
\end{aligned}$$

$$
\begin{aligned}
&= e_d\left(\frac{a(h_1v_2 - h_2v_1)}{nr_1q_0u_1v_1v_2q_2}\right) e_{r_1}\left(\frac{a(h_1v_2^* - h_2v_1^*)}{ndq_0u_1v^*v_1^*v_2^*q_2}\right) e_{v^*}\left(\frac{b_1(h_1v_2^* - h_2v_1^*)}{ndr_1q_0u_1v_1^*v_2^*q_2}\right) \\
&e_{q_0u_1}\left(\frac{b_1(h_1v_2^* - h_2v_1^*)}{ndr_1v^*v_1^*v_2^*q_2}\right) e_{v_1^*}\left(\frac{b_1(h_1v_2^* - h_2v_1^*)}{ndr_1q_0u_1v^*v_2^*q_2}\right) e_{v_2^*}\left(\frac{b_1(h_1v_2^* - h_2v_1^*)}{ndr_1q_0u_1v^*v_1^*q_2}\right) \\
&e_{q_2}\left(\frac{b_2(h_1v_2^* - h_2v_1^*)}{(n+\ell dr_1)dr_1q_0u_1v^*v_1^*v_2^*}\right) \\
&= e_d\left(\frac{a(h_1v_2^* - h_2v_1^*)}{nr_1q_0u_1v^*v_1^*v_2^*q_2}\right) e_m\left(\frac{A(h_1v_2^* - h_2v_1^*)}{d(n+Bd)}\right)
\end{aligned}
$$

where $m = r_1q_0u_1v^*v_1^*v_2^*q_2 = r_1q_0u_1[v_1, v_2]q_2$ and where A and B are residue classes mod m which are independent of n, d, h_1 and h_2. Since $(ab_1b_2, r_1q_0u_1v_1v_2q_2) = 1$, we have that a, b_1 and b_2 are coprime to m and A satisfies $(A, m) = 1$. □

3.3.2.3 q-van der Corput process

We now apply the q-van der Corput method. This is where our arguments diverge from the work of Polymath [37]: On page 97 of [37], sums were partitioned according to the value of n mod d. Instead, we now partition according to the value of $(h_1v_2^* - h_2v_1^*)/n$ mod d. By including both h and d in the innermost sum of the q-van der Corput method, the contribution of diagonal terms is decreased. This allows for a larger choice of D: In [37], the proof used $D \approx N/H^4$, but now we may work with $D \approx N/H^2$, which leads to a smaller modulus m in the exponential sums. This change is what allows for an improvement of the equidistribution estimates.

Lemma 3.8 (Alternative q-van der Corput method)**.** *Let Z_1, $Z_2(z_1)$ and $\Sigma_2(z_1, z_2)$ be as described in Lemma 3.5 and Lemma 3.6. For given $z_1 \in Z_1$ and $z_2 \in Z_2(z_1)$, denote by $Z_3(z_1, z_2)$ the set of tuples*

$$(W_1, Y, w_0, w_1, w_2, A, B, C(n))$$

which satisfy the following: $Y \in \mathbb{R}$ with $1 \ll |Y| \ll \frac{H^*V}{(v_1,v_2)}$ and $W_1 \in (0, \infty)$ with $1 \ll W_1 \ll \Delta$. A and B are residue classes mod m with $(A, m) = 1$, where $m = r_1q_0u_1[v_1, v_2]q_2$. Further, $w_0, w_1, w_2 \in \mathbb{N}$ with $w_0 \mid w_1$, w_1 squarefree, $(w_1, m) = 1$, $w_1 \asymp W_1$ and $w_2 \mid m^\star = m^{\lfloor 2\log x\rfloor}$. Finally, $C(n)$ is the indicator function of a set E_d mod q_0, dependent on d mod q_0, which has at most (q_0, ℓ) elements.

Set $\mathcal{L} = \{h_1v_2^* - h_2v_1^* : h_1, h_2 \sim H^*\}$, where $v_1^* = v_1/(v_1, v_2)$ and $v_2^* = v_2/(v_1, v_2)$.

For given $z_1 \in Z_1$, $z_2 \in Z_2(z_1)$ and $z_3 \in Z_3(z_1, z_2)$, define $\Sigma_3(z_1, z_2, z_3)$ to equal

$$\sum_{\substack{y,\widetilde{y}\in\mathcal{L} \\ y,\widetilde{y}\sim Y \\ w_1|(y,\widetilde{y}) \\ (y,m^*)=w_2 \\ (\widetilde{y},m^*)=w_2}} \Bigg| \sum_{\substack{d_1 \\ w_0w_1|d \\ (\frac{d}{w_1},\frac{my\widetilde{y}}{w_1^2})=1 \\ where \\ d=d_0+d_1}} \psi_{\Delta_1}(d_1) \sum_{\substack{n,\widetilde{n} \\ (n,w_1r_1q_0u_1v_1v_2)=1 \\ (\widetilde{n},w_1r_1q_0u_1v_1v_2)=1 \\ (\widetilde{n}+\ell dr_1,q_0q_2)=1 \\ (n+\ell dr_1,q_0q_2)=1 \\ y\widetilde{n}\equiv\widetilde{y}n \mod d}} C(n)C(\widetilde{n})\psi_N(n)\psi_N(\widetilde{n})e_m\left(\frac{A(y(\widetilde{n}+Bd)-\widetilde{y}(n+Bd))}{d(n+Bd)(\widetilde{n}+Bd)}\right)\Bigg|.$$

We then have the bound

$$\Sigma_2(z_1, z_2) \ll_\varepsilon \sup_{z_3\in Z_3(z_1,z_2)} (x^{3.5\varepsilon}(v_1, v_2)W_1^{1/2}\Delta)\, \Sigma_3(z_1, z_2, z_3)^{1/2}.$$

In particular, if $\Sigma_3(z_1, z_2, z_3) = O_\varepsilon(\frac{q_0^2(q_0,\ell)^2N^2}{x^{27\varepsilon}W_1})$, then $\Sigma_2(z_1, z_2) = O_\varepsilon(\frac{q_0(q_0,\ell)\Delta N(v_1,v_2)}{x^{10\varepsilon}})$.

Proof. Our goal is to bound $\Sigma_2(z_1, z_2)$, which is given by

$$\sum_{\substack{d_1 \\ (d,r_1q_0u_1)=1 \\ (d,[v_1,v_2]q_2)=1 \\ d \text{ squarefree} \\ \text{where } d=d_0+d_1}} \psi_{\Delta_1}(d_1) \Bigg| \sum_{\substack{h_1,h_2 \\ h_1,h_2\sim H^* \\ h_1v_2\neq h_2v_1}} \sum_{\substack{n \\ (n,dr_1q_0)=1 \\ (n,u_1v_1v_2)=1 \\ (n+\ell dr_1,q_0)=1 \\ (n+\ell dr_1,q_2)=1}} C(n)\varphi_H^\star(h_1)\overline{\varphi_H^{\star\star}(h_2)}\psi_N(n)\Psi(n,h_1,h_2,dr_1,u_1,v_1,v_2,q_2)\Bigg|.$$

Here $C(n) = 1_{\frac{b_1}{n}\equiv\frac{b_2}{n+\ell dr_1}(q_0)}$, and this is the indicator function of a set E_d mod q_0, which depends on the value of d mod q_0.

We begin our proof by splitting up the sums inside the absolute value signs of $\Sigma_2(z_1, z_2)$ according to the value of $(h_1v_2^* - h_2v_1^*)/n$ mod d, where $v_1^* = v_1/(v_1, v_2)$ and $v_2^* = v_2/(v_1, v_2)$. We recall Lemma 3.7. If $(h_1v_2^* - h_2v_1^*)/n \equiv \gamma$ mod d, then

$$\begin{aligned}\Psi(n, h_1, h_2, dr_1, u_1, v_1, v_2, q_2) &= e_d\left(\frac{a(h_1v_2^* - h_2v_1^*)}{nr_1q_0u_1v^*v_1^*v_2^*q_2}\right) e_m\left(\frac{A(h_1v_2^* - h_2v_1^*)}{d(n + Bd)}\right) \\ &= e_d\left(\frac{a\gamma}{r_1q_0u_1v^*v_1^*v_2^*q_2}\right) e_m\left(\frac{A(h_1v_2^* - h_2v_1^*)}{d(n + Bd)}\right),\end{aligned}$$

where $m = r_1q_0u_1[v_1, v_2]q_2$ and A and B are residue classes mod m, independent of n, d, h_1 and h_2, with $(A, m) = 1$. The d-factor no longer depends on h_1, h_2 and n. So we find that $\Sigma_2(z_1, z_2) \ll E_1(z_1, z_2)$, where E_1 is defined to equal

$$\sum_{\substack{d_1 \\ (d,m)=1 \\ d \text{ squarefree} \\ \text{where } d=d_0+d_1}} \psi_{\Delta_1}(d_1)\sum_{\gamma(d)} \Bigg|\sum_{\substack{h_1,h_2 \\ h_1\sim H^* \\ h_2\sim H^* \\ h_1v_2\neq h_2v_1}} \sum_{\substack{n \\ (n,dr_1q_0u_1v_1v_2)=1 \\ (n+\ell dr_1,q_0q_2)=1 \\ \frac{(h_1v_2^*-h_2v_1^*)}{n}\equiv\gamma(d)}} C(n)\varphi_H^\star(h_1)\overline{\varphi_H^{\star\star}(h_2)}\psi_N(n)e_m\left(\frac{A(h_1v_2^* - h_2v_1^*)}{d(n + Bd)}\right)\Bigg|.$$

The condition $(n, d) = 1$ could become bothersome at a later stage. To ease its removal, we now partition according to the value of $w_1 = (h_1v_2^* - h_2v_1^*, d)$ and will later only insist that $(n,w_1)=1$. Similarly, choices of $h_1v_2^*-h_2v_1^*$ which share prime factors with m may cause issues, and so we partition according to the value of $w_2=(h_1v_2^*-h_2v_1^*, m^{\lfloor 2\log x\rfloor})$. (Since $2^{\lfloor 2\log x\rfloor}>x$, this means that we partition according to the largest factor of $h_1v_2^*-h_2v_1^*$ which consists solely of powers of primes which divide m.) Writing $m^*=m^{\lfloor 2\log x\rfloor}$, we get $\Sigma_2(z_1,z_2)\ll E_2(z_1,z_2)$, where E_2 is

$$\sum_{\substack{d_1,w_1,w_2\\(d,m)=1\\d\text{ squarefree}\\w_1|d\\w_2|m^*\\\text{where } d=d_0+d_1}}\psi_{\Delta_1}(d_1)\sum_{\gamma(d)}\Bigg|\sum_{\substack{h_1,h_2\\h_1\sim H^*\\h_2\sim H^*\\h_1v_2\neq h_2v_1\\(h_1v_2^*-h_2v_1^*,d)=w_1\\(h_1v_2^*-h_2v_1^*,m^*)=w_2}}\sum_{\substack{n\\(n,dr_1q_0u_1v_1v_2)=1\\(n+\ell dr_1,q_0q_2)=1\\\frac{(h_1v_2^*-h_2v_1^*)}{n}\equiv\gamma(d)}}C(n)\varphi_H^{\star}(h_1)\overline{\varphi_H^{\star\star}(h_2)}\psi_N(n)e_m\bigg(\frac{A(h_1v_2^*-h_2v_1^*)}{d(n+Bd)}\bigg)\Bigg|.$$

Now we count the number of $w_2\mid m^*$ which have $w_2\ll x$. Since $|h_1v_2^*-h_2v_1^*|\ll x$, these are the only w_2 which contribute a non-zero amount to $\Sigma_2(z_1,z_2)$. Recall that m is squarefree and denote the number of prime factors of m by k. Here $k\le\lfloor 2\log x\rfloor$. There is an injection from the set $\{w_2\in\mathbb{N}:w_2\mid m^*, w_2\ll x\}$ to the set $I_k=\{n\in\mathbb{N}:n=p_1^{\alpha_1}\dots p_k^{\alpha_k}, n\ll x\}$, where $p_1,\dots,p_k$ are the k smallest primes. (Denoting the prime factors of m in ascending order by $q_1,\dots,q_k$, we send $w_2=q_1^{\alpha_1}\dots q_k^{\alpha_k}$ to $p_1^{\alpha_1}\dots p_k^{\alpha_k}$, which is less than or equal to w_2.) But I_k is the set of p_k-smooth numbers bounded by x. Here $p_k=O(k\log k)=O_\varepsilon((\log x)^{1+\varepsilon/2})$ since $k\le\lfloor 2\log x\rfloor$. It is known that the number of $(\log x)^{1+\varepsilon/2}$-smooth numbers less than x is bounded by $O_\varepsilon(x^{1-1/(1+\varepsilon/2)+o(1)})=O_\varepsilon(x^\varepsilon)$. (See Hildebrand and Tenenbaum [22].) Hence we only need to consider $O_\varepsilon(x^\varepsilon)$ different values of w_2.

Next we partition according to the size of w_1, restricting to $w_1\asymp W_1$ for some W_1 with $1\ll W_1\ll\Delta$. We also fix w_2. Finally, we restrict the value of $h_1v_2^*-h_2v_1^*$ to a dyadic interval, writing $h_1v_2^*-h_2v_1^*\sim Y$ for some Y with $1\ll|Y|\ll\frac{H^*V}{(v_1,v_2)}$. Here we use that $h_1v_2^*-h_2v_1^*\neq 0$. (Note that $h_1v_2^*-h_2v_1^*$ and Y may be negative, but not zero.) Using dyadic decomposition, we define $\Upsilon_0(W_1,Y,w_1,w_2)$ to equal

$$\Upsilon_0=\sum_{\substack{d_1\\(d,m)=1\\d\text{ squarefree}\\w_1|d\\\text{where } d=d_0+d_1}}\psi_{\Delta_1}(d_1)\sum_{\gamma(d)}\Bigg|\sideset{}{^*}\sum_{h_1,h_2}\sum_{\substack{n\\(n,dr_1q_0)=1\\(n,u_1v_1v_2)=1\\(n+\ell dr_1,q_0q_2)=1\\\frac{(h_1v_2^*-h_2v_1^*)}{n}\equiv\gamma(d)}}C(n)\varphi_H^{\star}(h_1)\overline{\varphi_H^{\star\star}(h_2)}\psi_N(n)e_m\bigg(\frac{A(h_1v_2^*-h_2v_1^*)}{d(n+Bd)}\bigg)\Bigg|,$$

where $\sideset{}{^*}\sum$ denotes the sum over h_1 and h_2 with $h_1,h_2\sim H^*$, $h_1v_2^*-h_2v_1^*\sim Y$, $(h_1v_2^*-h_2v_1^*,d)=w_1$ and $(h_1v_2^*-h_2v_1^*,m^*)=w_2$. We now have

$$\Sigma_2(z_1,z_2)\ll_\varepsilon\sup_{W_1,Y,w_1,w_2}(x^{2\varepsilon}W_1)\Upsilon_0(W_1,Y,w_1,w_2).$$

Here the supremum is taken over choices of W_1 and Y with $1 \ll W_1 \ll \Delta$ and $1 \ll |Y| \ll \frac{H^*V}{(v_1,v_2)}$ and over all squarefree $w_1 \asymp W_1$ with $(w_1, m) = 1$ and all $w_2 \mid m^*$.

We retained the condition that d is squarefree up to this point to ensure that we only need to consider squarefree w_1. Now we remove the condition, replacing it by $(d/w_1, w_1) = 1$. Observe that $\Upsilon_0 \ll \Upsilon_1$ for

$$\Upsilon_1 = \sum_{\substack{d_1 \\ (d,m)=1 \\ (d/w_1,w_1)=1 \\ w_1 | d \\ \text{where } d=d_0+d_1}} \psi_{\Delta_1}(d_1) \sum_{\gamma(d)} \Big| \sideset{}{^*}\sum_{h_1,h_2} \sum_{\substack{n \\ (n,dr_1q_0)=1 \\ (n,u_1v_1v_2)=1 \\ (n+\ell dr_1,q_0q_2)=1 \\ \frac{(h_1v_2^*-h_2v_1^*)}{n} \equiv \gamma(d)}} C(n)\varphi_H^\star(h_1)\overline{\varphi_H^{\star\star}(h_2)}\psi_N(n)e_m\left(\frac{A(h_1v_2^*-h_2v_1^*)}{d(n+Bd)}\right)\Big|.$$

(Here we used that $\psi_{\Delta_1}(d_1) \geq 0$ and that $w_1 \mid d$ and d squarefree imply $(\frac{d}{w_1}, w_1) = 1$.)

Since we only sum over h_1 and h_2 with $(h_1v_2^* - h_2v_1^*, d) = w_1$ and n with $(n, d) = 1$, residue classes $\gamma(d)$ with $(\gamma, w_1) \neq w_1$ do not contribute to Υ_1. Hence, for each residue class mod (d/w_1), only one lifted residue class mod d contributes to Υ_1. Furthermore, $\frac{h_1v_2^*-h_2v_1^*}{n}$ is invertible mod (d/w_1). We may replace summation over $\gamma(d)$ by summation over $\gamma_1(d/w_1)$ and invert $\frac{h_1v_2^*-h_2v_1^*}{n}$, as follows:

$$\Upsilon_1 = \sum_{\substack{d_1 \\ (d,m)=1 \\ (d/w_1,w_1)=1 \\ w_1 | d \\ \text{where } d=d_0+d_1}} \psi_{\Delta_1}(d_1) \sum_{\gamma_1(d/w_1)} \Big| \sideset{}{^*}\sum_{h_1,h_2} \sum_{\substack{n \\ (n,dr_1q_0)=1 \\ (n,u_1v_1v_2)=1 \\ (n+\ell dr_1,q_0q_2)=1 \\ \frac{n}{(h_1v_2^*-h_2v_1^*)} \equiv \gamma_1(\frac{d}{w_1})}} C(n)\varphi_H^\star(h_1)\overline{\varphi_H^{\star\star}(h_2)}\psi_N(n)e_m\left(\frac{A(h_1v_2^*-h_2v_1^*)}{d(n+Bd)}\right)\Big|.$$

The condition $(n, d) = 1$ ensures that only congruence classes $\gamma_1(\frac{d}{w_1})$ with $(\gamma_1, \frac{d}{w_1}) = 1$ contribute to Υ_1. If we replace $(n, d) = 1$ by $(n, w_1) = 1$, we get new summands corresponding to $(\gamma_1, \frac{d}{w_1}) \neq 1$ and this increases the sum. Hence

$$\Upsilon_1 \ll \sum_{\substack{d_1 \\ (d,m)=1 \\ (d/w_1,w_1)=1 \\ w_1 | d \\ \text{where } d=d_0+d_1}} \psi_{\Delta_1}(d_1) \sum_{\gamma_1(d/w_1)} \Big| \sideset{}{^*}\sum_{h_1,h_2} \sum_{\substack{n \\ (n,w_1r_1q_0)=1 \\ (n,u_1v_1v_2)=1 \\ (n+\ell dr_1,q_0q_2)=1 \\ \frac{n}{(h_1v_2^*-h_2v_1^*)} \equiv \gamma_1(\frac{d}{w_1})}} C(n)\varphi_H^\star(h_1)\overline{\varphi_H^{\star\star}(h_2)}\psi_N(n)e_m\left(\frac{A(h_1v_2^*-h_2v_1^*)}{d(n+Bd)}\right)\Big|.$$

Next we apply the Cauchy-Schwarz inequality. We have $\Upsilon_0^2 \ll \Upsilon_1^2 \ll_\varepsilon \Upsilon_2\Upsilon_3$, where

$$\Upsilon_2 = x^\varepsilon \Delta \max\left\{\Delta_1 W_1^{-1}, 1\right\} \ll x^\varepsilon \Delta^2 W_1^{-1},$$

$$\Upsilon_3 = \sum_{\substack{d_1 \\ (d,m)=1 \\ (d/w_1,w_1)=1 \\ w_1 | d \\ \text{where } d=d_0+d_1}} \psi_{\Delta_1}(d_1) \sum_{\gamma_1(\frac{d}{w_1})} \sideset{}{^*}\sum_{h_1,h_2} \sideset{}{^*}\sum_{\widetilde{h}_1,\widetilde{h}_2} \sum_{\substack{n,\widetilde{n} \\ (n,w_1r_1q_0u_1v_1v_2)=1 \\ (\widetilde{n},w_1r_1q_0u_1v_1v_2)=1 \\ (n+\ell dr_1,q_0q_2)=(\widetilde{n}+\ell dr_1,q_0q_2)=1 \\ \frac{n}{(h_1v_2^*-h_2v_1^*)} \equiv \frac{\widetilde{n}}{(\widetilde{h}_1v_2^*-\widetilde{h}_2v_1^*)} \equiv \gamma_1(\frac{d}{w_1})}} \varphi_H^\star(h_1)\overline{\varphi_H^{\star\star}(h_2)}\overline{\varphi_H^\star(\widetilde{h}_1)}\varphi_H^{\star\star}(\widetilde{h}_2) \ (\star)$$

$$\text{with } (\star) = C(n)C(\widetilde{n})\psi_N(n)\psi_N(\widetilde{n})e_m\left(\frac{A(h_1v_2^*-h_2v_1^*)}{d(n+Bd)}\right)e_m\left(-\frac{A(\widetilde{h}_1v_2^*-\widetilde{h}_2v_1^*)}{d(\widetilde{n}+Bd)}\right).$$

In particular, we now have the upper bound

$$\Sigma_2(z_1,z_2) \ll_\varepsilon \sup_{W_1,Y,w_1,w_2}(x^{2.5\varepsilon}W_1^{1/2}\Delta)\Upsilon_3(W_1,Y,w_1,w_2)^{1/2}.$$

Since $(h_1v_2^*-h_2v_1^*,d)=w_1$ and $(\widetilde{h}_1v_2^*-\widetilde{h}_2v_1^*,d)=w_1$ and $(w_1,\frac{d}{w_1})=1$, the condition $\frac{n}{(h_1v_2^*-h_2v_1^*)} \equiv \frac{\widetilde{n}}{(\widetilde{h}_1v_2^*-\widetilde{h}_2v_1^*)} \equiv \gamma_1(\frac{d}{w_1})$ together with the sum over $\gamma_1(\frac{d}{w_1})$ reduce to

$$(h_1v_2^*-h_2v_1^*)\widetilde{n} \equiv (\widetilde{h}_1v_2^*-\widetilde{h}_2v_1^*)n \bmod d.$$

Note also that $|\varphi_H^\star(h)|,|\varphi_H^{\star\star}(h)| \ll (\log M)^{O(1)}$ for $|h| \ll H$. Hence we have the following upper bound:

$$\Upsilon_3 \ll x^{o(1)} \sum_{\substack{h_1,\widetilde{h}_1,h_2,\widetilde{h}_2 \\ h_1,\widetilde{h}_1,h_2,\widetilde{h}_2\sim H^* \\ h_1v_2^*-h_2v_1^*,\widetilde{h}_1v_2^*-\widetilde{h}_2v_1^*\sim Y \\ w_1|h_1v_2^*-h_2v_1^*,\, w_1|\widetilde{h}_1v_2^*-\widetilde{h}_2v_1^* \\ (h_1v_2^*-h_2v_1^*,m^*)=w_2 \\ (\widetilde{h}_1v_2^*-\widetilde{h}_2v_1^*,m^*)=w_2}} \left| \sum_{\substack{d \\ w_1|d \\ (d,m)=1 \\ (d/w_1,(h_1v_2^*-h_2v_1^*)/w_1)=1 \\ (d/w_1,(\widetilde{h}_1v_2^*-\widetilde{h}_2v_1^*)/w_1)=1 \\ (d/w_1,w_1)=1}} \psi_{\Delta_1}(d-d_0) \sum_{\substack{n,\widetilde{n} \\ (n,w_1r_1q_0u_1v_1v_2)=1 \\ (\widetilde{n},w_1r_1q_0u_1v_1v_2)=1 \\ (n+\ell dr_1,q_0q_2)=1 \\ (\widetilde{n}+\ell dr_1,q_0q_2)=1 \\ (h_1v_2^*-h_2v_1^*)\widetilde{n}\equiv(\widetilde{h}_1v_2^*-\widetilde{h}_2v_1^*)n \bmod d}} (\star) \right|$$

$$\text{with } (\star) = C(n)C(\widetilde{n})\psi_N(n)\psi_N(\widetilde{n})e_m\left(\frac{A(h_1v_2^*-h_2v_1^*)}{d(n+Bd)}\right)e_m\left(-\frac{A(\widetilde{h}_1v_2^*-\widetilde{h}_2v_1^*)}{d(\widetilde{n}+Bd)}\right).$$

Now consider $y \sim Y$ and count the number of choices of (h_1,h_2) with $h_1 \sim H^*$ and $h_2 \sim H^*$ such that $y = h_1v_2^*-h_2v_1^*$. (Recall here that v_1 and v_2 are already fixed.) Suppose that (h_1^*,h_2^*) is a solution. Since $(v_1^*,v_2^*)=1$, the other possible solutions to $h_1v_2^*-h_2v_1^*=y$ are of the form $(h_1,h_2)=(h_1^*+kv_1^*,h_2^*+kv_2^*)$. However, since $v_1^* \asymp V/(v_1,v_2)$ and since we also require $h_1 \sim H^*$, the number of possible k is bounded by $O(\max\{1,(v_1,v_2)H^*/V\}) = O((v_1,v_2))$. Now write $\mathcal{L}(v_1^*,v_2^*,H^*) = \{h_1v_2^*-h_2v_1^* : h_1,h_2\sim H^*\}$. Set Υ_4 equal to

$$\sum_{\substack{y,\widetilde{y}\in\mathcal{L} \\ y,\widetilde{y}\sim Y \\ w_1|(y,\widetilde{y}) \\ (y,m^*)=w_2 \\ (\widetilde{y},m^*)=w_2}} \left| \sum_{\substack{d_1 \\ w_1|d \\ (\frac{d}{w_1},\frac{my\widetilde{y}}{w_1^2})=1 \\ (d/w_1,w_1)=1 \\ \text{where } d=d_0+d_1}} \psi_{\Delta_1}(d_1) \sum_{\substack{n,\widetilde{n} \\ (n\widetilde{n},w_1r_1q_0)=1 \\ (n\widetilde{n},u_1v_1v_2)=1 \\ ((n+\ell dr_1),q_0q_2)=1 \\ ((\widetilde{n}+\ell dr_1),q_0q_2)=1 \\ y\widetilde{n}\equiv\widetilde{y}n \bmod d}} C(n)C(\widetilde{n})\psi_N(n)\psi_N(\widetilde{n})e_m\left(\frac{A(y(\widetilde{n}+Bd)-\widetilde{y}(n+Bd))}{d(n+Bd)(\widetilde{n}+Bd)}\right)\right|.$$

We then have $\Upsilon_3 \ll_\varepsilon (x^\varepsilon(v_1,v_2)^2)\Upsilon_4$. In particular,

$$\Sigma_2(z_1,z_2) \ll \sup_{W_1,Y,w_1,w_2}(x^{3\varepsilon}(v_1,v_2)W_1^{1/2}\Delta)\,\Upsilon_4(W_1,Y,w_1,w_2)^{1/2}.$$

Finally, we remove the condition $(\frac{d}{w_1}, w_1) = 1$: By Möbius inversion,

$$\Upsilon_4 \ll \sum_{\substack{w_0 \\ w_0 | w_1}} |\mu(w_0)| \sum_{\substack{y,\widetilde{y} \in \mathcal{L} \\ y,\widetilde{y} \sim Y \\ w_1 | (y,\widetilde{y}) \\ (y,m^*)=w_2 \\ (\widetilde{y},m^*)=w_2}} \left| \sum_{\substack{d \\ w_0 w_1 | d \\ (\frac{d}{w_1}, \frac{my\widetilde{y}}{w_1^2})=1}} \psi_{\Delta_1}(d-d_0) \sum_{\substack{n,\widetilde{n} \\ (n\widetilde{n}, w_1 r_1 q_0 u_1 v_1 v_2)=1 \\ ((n+\ell d r_1)(\widetilde{n}+\ell d r_1), q_0 q_2)=1 \\ y\widetilde{n} \equiv \widetilde{y} n \bmod d}} (\star) \right|$$

$$\text{with } (\star) = C(n)C(\widetilde{n})\psi_N(n)\psi_N(\widetilde{n}) e_m\left(\frac{A(y(\widetilde{n}+Bd) - \widetilde{y}(n+Bd))}{d(n+Bd)(\widetilde{n}+Bd)}\right).$$

But w_1 has only $O_\varepsilon(x^\varepsilon)$ divisors. For a given $w_0 \mid w_1$, define $\Upsilon_5(w_0)$ to equal

$$\sum_{\substack{y,\widetilde{y} \in \mathcal{L} \\ y,\widetilde{y} \sim Y \\ w_1 | (y,\widetilde{y}) \\ (y,m^*)=w_2 \\ (\widetilde{y},m^*)=w_2}} \left| \sum_{\substack{d_1 \\ w_0 w_1 | d \\ (\frac{d}{w_1}, \frac{my\widetilde{y}}{w_1^2})=1 \\ \text{where} \\ d=d_0+d_1}} \psi_{\Delta_1}(d_1) \sum_{\substack{n,\widetilde{n} \\ (n\widetilde{n}, w_1 r_1 q_0)=1 \\ (n\widetilde{n}, u_1 v_1 v_2)=1 \\ ((n+\ell d r_1), q_0 q_2)=1 \\ ((\widetilde{n}+\ell d r_1), q_0 q_2)=1 \\ y\widetilde{n} \equiv \widetilde{y} n \bmod d}} C(n)C(\widetilde{n})\psi_N(n)\psi_N(\widetilde{n}) e_m\left(\frac{A(y(\widetilde{n}+Bd) - \widetilde{y}(n+Bd))}{d(n+Bd)(\widetilde{n}+Bd)}\right) \right|.$$

Observe that we now have the bound

$$\Sigma_2(z_1, z_2) \ll \sup_{W_1, Y, w_1, w_2, w_0} (x^{3.5\varepsilon}(v_1, v_2) W_1^{1/2} \Delta)\, \Upsilon_5(W_1, Y, w_1, w_2, w_0)^{1/2},$$

where the supremum is taken over w_0 with $w_0 \mid w_1$. But $\Upsilon_5(W_1, Y, w_1, w_2, w_0)$ is of the same form as $\Sigma_3(z_1, z_2, z_3)$ for some $z_3 \in Z_3(z_1, z_2)$. This concludes the proof. $\square$

3.3.2.4 Large divisors

We have reduced the problem of obtaining equidistribution estimates to estimating exponential sums of the form Σ_3 given in Lemma 3.8. Tuples $(n, \widetilde{n}, y, \widetilde{y})$ with $y(\widetilde{n} + Bd) - \widetilde{y}(n + Bd) = 0$ contribute a large amount to Σ_3. We also have no non-trivial estimates for the exponential sums when $(y(\widetilde{n}+Bd)-\widetilde{y}(n+Bd))/d$ shares a very large divisor with m. In the next lemma we thus remove such bad choices of $(n, \widetilde{n}, y, \widetilde{y})$. Here it is of great importance that D has been chosen sufficiently small. (Due to the involvement of y, the diagonal/bad terms here are different from the ones in [37], and this is what allows us to choose a larger D.)

Lemma 3.9 (Removing diagonal terms)**.** *Let Z_1, $Z_2(z_1)$, $Z_3(z_1, z_2)$ and $\Sigma_3(z_1, z_2, z_3)$ be as described in Lemma 3.5, Lemma 3.6 and Lemma 3.8.*

For given $z_1 \in Z_1$, $z_2 \in Z_2(z_1)$ and $z_3 \in Z_3(z_1, z_2)$, denote by $Z_4(z_1, z_2, z_3)$ the set of tuples $(y, \widetilde{y})$ which satisfy the following: $y, \widetilde{y} \in \mathbb{Z} \setminus \{0\}$ with $y, \widetilde{y} \sim Y$, $w_1 \mid (y, \widetilde{y})$ and $(y, m^*) = (\widetilde{y}, m^*) = w_2$, where $m = r_1 q_0 u_1 [v_1, v_2] q_2$ and $m^* = m^{\lfloor 2 \log x \rfloor}$.

Set $T(z_1,z_2,z_3)=\max\{\frac{1}{(w_2,m)},\frac{1}{H}\}\frac{x^{\delta+100\varepsilon}H^2N}{(v_1,v_2)\Delta_1}$ *and define* $\Sigma_4(z_1,z_2,z_3,z_4)$ *to equal*

$$\left|\sum_{\substack{d\\ w_0w_1|d\\ (\frac{d}{w_1},\frac{my\widetilde{y}}{w_1^2})=1}}\psi_{\Delta_1}(d-d_0)\sum_{\substack{n,\widetilde{n}\\ (n\widetilde{n},w_1r_1q_0u_1v_1v_2)=1\\ ((n+\ell dr_1)(\widetilde{n}+\ell dr_1),q_0q_2)=1\\ y\widetilde{n}\equiv\widetilde{y}n \mod d\\ ((y(\widetilde{n}+Bd)-\widetilde{y}(n+Bd))/d,m)\le T(z_1,z_2,z_3)}}C(n)C(\widetilde{n})\psi_N(n)\psi_N(\widetilde{n})e_m\left(\frac{A(y(\widetilde{n}+Bd)-\widetilde{y}(n+Bd))}{d(n+Bd)(\widetilde{n}+Bd)}\right)\right|.$$

We then have the bound

$$\Sigma_3(z_1,z_2,z_3)\ll_\varepsilon\sup_{z_4\in Z_4(z_1,z_2,z_3)}\left(\max\left\{\frac{x^{\delta+10\varepsilon}H^3}{W_1(w_2,m)},\frac{H^4}{W_1(w_2,m)}\right\}\right)\Sigma_4(z_1,z_2,z_3,z_4)+\left(\frac{N^2}{x^{47\varepsilon}W_1}\right).$$

Note here that if $\Sigma_4(z_1,z_2,z_3,z_4)=O_\varepsilon\left(\min\left\{\frac{(w_2,m)}{x^{\delta+10\varepsilon}H^3},\frac{(w_2,m)}{H^4}\right\}\frac{q_0^2(q_0,\ell)^2N^2}{x^{27\varepsilon}}\right)$, then

$$\Sigma_3(z_1,z_2,z_3)=O_\varepsilon\left(\frac{q_0^2(q_0,\ell)^2W_1^{-1}N^2}{x^{27\varepsilon}}\right).$$

Proof. Our goal is to bound $\Sigma_3(z_1,z_2,z_3)$, which is given by

$$\sum_{\substack{y,\widetilde{y}\in\mathcal{L}\\ y,\widetilde{y}\sim Y\\ w_1|(y,\widetilde{y})\\ (y,m^*)=w_2\\ (\widetilde{y},m^*)=w_2}}\left|\sum_{\substack{d_1\\ w_0w_1|d\\ (\frac{d}{w_1},\frac{my\widetilde{y}}{w_1^2})=1\\ \text{where}\\ d=d_0+d_1}}\psi_{\Delta_1}(d_1)\sum_{\substack{n,\widetilde{n}\\ (n,w_1r_1q_0u_1v_1v_2)=1\\ (\widetilde{n},w_1r_1q_0u_1v_1v_2)=1\\ (\widetilde{n}+\ell dr_1,q_0q_2)=1\\ (n+\ell dr_1,q_0q_2)=1\\ y\widetilde{n}\equiv\widetilde{y}n \bmod d}}C(n)C(\widetilde{n})\psi_N(n)\psi_N(\widetilde{n})e_m\left(\frac{A(y(\widetilde{n}+Bd)-\widetilde{y}(n+Bd))}{d(n+Bd)(\widetilde{n}+Bd)}\right)\right|.$$

Notice that d divides $(y(\widetilde{n}+Bd)-\widetilde{y}(n+Bd))$ since $y\widetilde{n}\equiv\widetilde{y}n \bmod d$. Write

$$S(q;y,\widetilde{y},d)=\left\{(n,\widetilde{n}):\gcd\left(\frac{y(\widetilde{n}+Bd)-\widetilde{y}(n+Bd)}{d},m\right)=q\right\},$$
$$T^\star(y,\widetilde{y},d)=\bigcup_{q\le\max\left\{\frac{1}{(w_2,m)},\frac{1}{H}\right\}\frac{x^{\delta+100\varepsilon}H^2N}{(v_1,v_2)\Delta_1}}S(q;y,\widetilde{y},d).$$

We partition the set of $(n,\widetilde{n})$, using $S(q;y,\widetilde{y},d)$ with $q>\max\{\frac{1}{(w_2,m)},\frac{1}{H}\}\frac{x^{\delta+100\varepsilon}H^2N}{(v_1,v_2)\Delta_1}$ and using $T^\star(y,\widetilde{y},d)$. We define

$$\Upsilon_T=\sum_{\substack{y,\widetilde{y}\in\mathcal{L}\\ y,\widetilde{y}\sim Y\\ w_1|(y,\widetilde{y})\\ (y,m^*)=w_2\\ (\widetilde{y},m^*)=w_2}}\left|\sum_{\substack{d\\ w_0w_1|d\\ (\frac{d}{w_1},\frac{my\widetilde{y}}{w_1^2})=1}}\psi_{\Delta_1}(d-d_0)\sum_{\substack{n,\widetilde{n}\\ (n\widetilde{n},w_1r_1q_0u_1v_1v_2)=1\\ ((n+\ell dr_1)(\widetilde{n}+\ell dr_1),q_0q_2)=1\\ y\widetilde{n}\equiv\widetilde{y}n \bmod d\\ (n,\widetilde{n})\in T^\star(y,\widetilde{y},d)}}(\star)\right|,$$

$$\Upsilon_S(q)=\sum_{\substack{y,\widetilde{y}\in\mathcal{L}\\ y,\widetilde{y}\sim Y\\ w_1|(y,\widetilde{y})\\ (y,m^*)=w_2\\ (\widetilde{y},m^*)=w_2}}\left|\sum_{\substack{d\\ w_0w_1|d\\ (\frac{d}{w_1},\frac{my\widetilde{y}}{w_1^2})=1}}\psi_{\Delta_1}(d-d_0)\sum_{\substack{n,\widetilde{n}\\ (n\widetilde{n},w_1r_1q_0u_1v_1v_2)=1\\ ((n+\ell dr_1)(\widetilde{n}+\ell dr_1),q_0q_2)=1\\ y\widetilde{n}\equiv\widetilde{y}n \bmod d\\ (n,\widetilde{n})\in S(q;y,\widetilde{y},d)}}(\star)\right|,$$

$$\text{with } (\star) = C(n)C(\widetilde{n})\psi_N(n)\psi_N(\widetilde{n})e_m\left(\frac{A(y(\widetilde{n}+Bd)-\widetilde{y}(n+Bd))}{d(n+Bd)(\widetilde{n}+Bd)}\right).$$

The number of divisors of m is bounded by $O_\varepsilon(x^\varepsilon)$, so $S(q;y,\widetilde{y},d)$ is empty for all but $O_\varepsilon(x^\varepsilon)$ choices of q. Therefore we have the upper bound

$$\Sigma_3(z_1,z_2,z_3) \ll_\varepsilon \Upsilon_T + \sup_{q>T(z_1,z_2,z_3)} x^\varepsilon\,\Upsilon_S(q).$$

Initially, we focus on $\Upsilon_S(q)$. Fix $q > T(z_1,z_2,z_3) = \max\{\frac{1}{(w_2,m)},\frac{1}{H}\}\frac{x^{\delta+100\varepsilon}H^2N}{(v_1,v_2)\Delta_1}$. We wish to count the number of choices of $(y,\widetilde{y},n,\widetilde{n},d)$ which appear in $\Upsilon_S(q)$ with $(n,\widetilde{n}) \in S(q;y,\widetilde{y},d)$. For use in the computations below, recall that $H \ll x^{4\omega+\delta+7\varepsilon}$, $\Delta \ll D = \frac{N}{x^{50\varepsilon}H^2}$ and $Y \ll \frac{H^*V}{(v_1,v_2)} \ll \frac{x^{\delta+5\varepsilon}H^2}{(v_1,v_2)}$. Hence

$$\frac{Y}{q} \ll \frac{\min\{(w_2,m),H\}}{x^{150\varepsilon}H^2}.$$

We first count the number of $y \in \mathcal{L}$ with $(w_2,m) \mid y$. Note that (w_2,m) is squarefree since m is squarefree. Write $w_2^\circ = ((w_2,m),v_1^*)$. Suppose $y = h_1v_2^* - h_2v_1^*$ with $h_1,h_2 \sim H^*$. Since $(v_1^*,v_2^*) = 1$, we have $(w_2,m) \mid y$ if and only if ($h_1 \equiv 0 \bmod w_2^\circ$ and $h_2 \equiv (v_2^*/v_1^*)h_1 \bmod (w_2,m)/w_2^\circ$). Thus there are $O(H^*/w_2^\circ)$ choices for h_1, and for each h_1, there are $O(\max\{\frac{w_2^\circ H^*}{(w_2,m)},1\})$ choices for h_2. In total, the number of $y \in \mathcal{L}$ with $w_2 \mid y$ is thus bounded by $O(\max\{\frac{H^2}{(w_2,m)},H\})$. (By the exact same kind of argument, and using that $w_1 \asymp W_1$ is squarefree and coprime to w_2, we also have that the number of $y \in \mathcal{L}$ with $w_1 \mid y$ and $w_2 \mid y$ is bounded by $O(\max\{\frac{H^2}{W_1(w_2,m)},H\})$.)

To count the number of choices of $(y,\widetilde{y},n,\widetilde{n},d)$ which appear in $\Upsilon_S(q)$ with $(n,\widetilde{n}) \in S(q;y,\widetilde{y},d)$, we first consider the case ($y=\widetilde{y}$ and $n=\widetilde{n}$). There are $O(\Delta/(w_0W_1))$ choices for d with $d \asymp \Delta$ and $w_0w_1 \mid d$, there are $O((H^*)^2)$ choices for $y \sim Y$ with $y \in \mathcal{L} = \{h_1v_2^* - h_2v_1^* : h_1,h_2 \sim H^*\}$ and there are $O(N)$ choices for n. The total contribution of case ($y=\widetilde{y}$ and $n=\widetilde{n}$) is thus $O(W_1^{-1}\Delta H^2N) = O(x^{-50\varepsilon}W_1^{-1}N^2)$.

Next we consider the case ($y=\widetilde{y}$ and $n \neq \widetilde{n}$). Then $(n,\widetilde{n}) \in S(q;y,\widetilde{y},d)$ implies $q \mid \frac{1}{d}(y(\widetilde{n}-n))$. Given $d \asymp \Delta$ with $w_0w_1 \mid d$, $y \sim Y$ with $y \in \mathcal{L}$ and $w_2 \mid y$, and $n \asymp N$, there are only $O(NY/(\Delta q))$ suitable choices for $\widetilde{n}-n$, each of which gives only one choice for $\widetilde{n}$. There are $O(\max\{H^2/(w_2,m),H\})$ choices for $y \in \mathcal{L}$ with $w_2 \mid y$, there are $O(\Delta/(w_0W_1))$ choices for $d \asymp \Delta$ with $w_0w_1 \mid d$ and there are $O(N)$ choices for $n \asymp N$. Overall, this bounds the contribution of the case $y=\widetilde{y}$ by $O(\max\{\frac{H^2}{(w_2,m)},H\}\frac{\Delta N^2Y}{W_1\Delta q}) = O(x^{-50\varepsilon}W_1^{-1}N^2)$. We thus assume $y \neq \widetilde{y}$.

Next we treat the case $n = \widetilde{n}$. Fix a value of $d \asymp \Delta$ with $w_1 \mid d$. We count the number of $(n, y, \widetilde{y})$ with $q \mid \frac{1}{d}(y - \widetilde{y})(n + Bd)$. First we note that dq has $O_\varepsilon(x^\varepsilon)$ divisors and fix some $\gamma \mid dq$. We now assume that $(y - \widetilde{y}, dq) = \gamma$. Then we must have $\frac{dq}{\gamma} \mid (n + Bd)$. There are $O(\frac{Y}{\gamma})$ choices for $y - \widetilde{y}$, $O(\frac{Y}{w_2})$ choices for $y \sim Y$ with $w_2 \mid y$, and $O(\frac{\gamma N}{\Delta q} + 1)$ choices for n. Note that $(\frac{Y}{\gamma})(\frac{\gamma N}{\Delta q} + 1) = O(\frac{YN}{\Delta q} + Y)$. Overall, this bounds the contribution of the case $n = \widetilde{n}$ by $O_\varepsilon(x^\varepsilon W_1^{-1} \Delta Y^2 \frac{1}{w_2}(\frac{N}{\Delta q} + 1)) = O_\varepsilon(x^{\delta - 50\varepsilon} W_1^{-1} N + x^{2\delta} W_1^{-1} H^2 N) = O(x^{-50\varepsilon} W_1^{-1} N^2)$ since $H^2 \ll x^{8\omega + 3\delta} \ll x^{-3\delta} N$. We thus now also assume $n \neq \widetilde{n}$.

Fix some d with $d - d_0 \asymp \Delta_1$ and $w_0 w_1 \mid d$, some $\widetilde{n} \asymp N$ and some $y_d \in \mathcal{L} - \mathcal{L}$ with $|y_d| \ll Y$, $y_d \neq 0$ and $w_2 \mid y_d$. Since $\mathcal{L} - \mathcal{L} \subseteq \{h_1 v_2^* - h_2 v_1^* : h_1, h_2 \ll H^*\}$, there are $O(\max\{\frac{H^2}{(w_2,m)}, H\})$ choices for y_d and so in total there are $O(\max\{\frac{H^2}{(w_2,m)}, H\} W_1^{-1} \Delta N)$ choices for $(d, \widetilde{n}, y_d)$. Consider y, $\widetilde{y}$ and n with $y - \widetilde{y} = y_d$ and $(n, \widetilde{n}) \in S(q; y, \widetilde{y}, d)$. Then q divides $\frac{1}{d}(y(\widetilde{n} + Bd) - \widetilde{y}(n + Bd))$ and $(y\widetilde{n} - \widetilde{y}n) + Bdy_d = Cdq$ for some $C \in \mathbb{Z}$. Observe $|y\widetilde{n} - \widetilde{y}n| \ll YN \ll \frac{1}{(v_1,v_2)} x^{\delta + 5\varepsilon} H^2 N$ and $|dq| \gg \max\{\frac{1}{(w_2,m)}, \frac{1}{H}\} \frac{x^{\delta + 100\varepsilon} H^2 N}{(v_1,v_2)}$. Hence $C = By_d/q + O\left(\max\{\frac{1}{(w_2,m)}, \frac{1}{H}\}^{-1}\right)$ and this bounds the number of possible choices for C. Choose one such C. We have $\widetilde{y}(\widetilde{n} - n) = Cdq - (Bd + \widetilde{n})y_d$. All values on the RHS have been fixed, whereas the LHS is assumed to be non-zero. This gives $O_\varepsilon(x^\varepsilon)$ choices for the factors on the LHS. The values of $\widetilde{n}$ and $\widetilde{n} - n$ determine the value of n. Overall, this bounds the number of $(y, \widetilde{y}, n, \widetilde{n}, d)$ which appear in $\Upsilon_S(q)$ with $(n, \widetilde{n}) \in S(q; y, \widetilde{y}, d)$ by $O(\frac{N^2}{x^{49\varepsilon} W_1})$. So

$$\Upsilon_S(q) \ll_\varepsilon x^{-48\varepsilon} W_1^{-1} N^2.$$

Now we look at Υ_T. Consider y and $\widetilde{y}$ with $y, \widetilde{y} \in \mathcal{L}$, $y, \widetilde{y} \sim Y$ and $w_1 \mid (y, \widetilde{y})$ and $w_2 \mid (y, \widetilde{y})$. As observed earlier, there are $O(\max\{\frac{H^4}{W_1^2 (w_2,m)^2}, H^2\})$ pairs $y, \widetilde{y}$ with $y, \widetilde{y} \in \mathcal{L}$ and $w_1 \mid (y, \widetilde{y})$ and $w_2 \mid (y, \widetilde{y})$. On the other hand, the number of pairs $y, \widetilde{y} \sim Y$ with $w_1 w_2 \mid (y, \widetilde{y})$ is clearly also bounded by $O(\frac{Y^2}{W_1 (w_2,m)^2})$. Now observe that $\frac{Y^2}{W_1 (w_2,m)^2} = O(\frac{x^{\delta + 10\varepsilon} H^3}{W_1 (w_2,m)})$ when $W_1 (w_2, m) > x^\delta H$. But when $W_1 (w_2, m) \leq x^\delta H$, we also have $H^2 = O(\frac{x^\delta H^3}{W_1 (w_2,m)})$. Overall, the number of relevant y and $\widetilde{y}$ is bounded by $O(\max\{\frac{x^{\delta + 10\varepsilon} H^3}{W_1 (w_2,m)}, \frac{H^4}{W_1 (w_2,m)}\})$. Recall that $\Sigma_4(z_1, z_2, z_3, (y, \widetilde{y}))$ is defined to equal

$$\left| \sum_{\substack{d \\ w_0 w_1 \mid d \\ (\frac{d}{w_1}, \frac{my\widetilde{y}}{w_1^2}) = 1}} \psi_{\Delta_1}(d - d_0) \sum_{\substack{n, \widetilde{n} \\ (n\widetilde{n}, w_1 r_1 q_0 u_1 v_1 v_2) = 1 \\ ((n + \ell d r_1)(\widetilde{n} + \ell d r_1), q_0 q_2) = 1 \\ y\widetilde{n} \equiv \widetilde{y} n \bmod d \\ (n, \widetilde{n}) \in T^\star(y, \widetilde{y}, d)}} C(n) C(\widetilde{n}) \psi_N(n) \psi_N(\widetilde{n}) e_m\left(\frac{A(y(\widetilde{n} + Bd) - \widetilde{y}(n + Bd))}{d(n + Bd)(\widetilde{n} + Bd)}\right) \right|.$$

We then have $\Upsilon_T \ll (\max\{\frac{x^{\delta+10\varepsilon}H^3}{W_1(w_2,m)}, \frac{H^4}{W_1(w_2,m)}\}) \sup_{(y,\widetilde{y})} \Sigma_4(z_1, z_2, z_3, (y, \widetilde{y}))$, where the supremum is taken over all choices of y and $\widetilde{y}$ with $y, \widetilde{y} \in \mathcal{L}$, $y, \widetilde{y} \sim Y$, $w_1 \mid (y, \widetilde{y})$ and $(y, m^*) = (\widetilde{y}, m^*) = w_2$. Thus

$$\Sigma_3(z_1, z_2, z_3) \ll_\varepsilon \sup_{z_4 \in Z_4(z_1,z_2,z_3)} \left(\max\left\{\frac{x^{\delta+10\varepsilon}H^3}{W_1(w_2,m)}, \frac{H^4}{W_1(w_2,m)}\right\}\right) \Sigma_4(z_1, z_2, z_3, z_4) + \left(\frac{N^2}{x^{47\varepsilon}W_1}\right).$$

This completes the proof. □

3.3.3 Summary - Results so far

By now, we have accumulated a lot of notation. For this reason we now summarize our results up to this point and simplify their presentation.

Lemma 3.10 (Summary). *Let $\omega, \delta > 0$. Let $\varepsilon \in (0, 10^{-100}\delta)$. Let $N = x^\gamma$, where $\gamma \in (12\omega + 6\delta, \frac{1}{2} - 2\omega - 8\varepsilon)$. Let α and β be coefficient sequences at scales M and N with $x \ll MN \ll x$. Assume β has the Siegel-Walfisz property.*

Denote by Z_5 the set of tuples

$$(Q, R, \Delta_1, \Lambda, m, q_0, \ell, z_1, d_0, v_1, v_2, w_1, w_2, c_1, c_2, l, \lambda, \widetilde{\lambda}, A, B, \psi_N, \psi_{\Delta_1}, C(n; d))$$

with the following properties: $m, q_0, \ell, z_1, d_0, v_1, v_2, w_1, w_2, c_1, c_2 \in \mathbb{N}$ with $m \mid P(x^\delta)$, $q_0 \mid m$ and $q_0 \ll Q$, $\ell \ll \frac{N}{R}$, $z_1 \ll x^{5\varepsilon}\Delta_1$ and $(z_1, m) = 1$, $d_0 \asymp x^{5\varepsilon}\Delta_1$, w_1 square-free and $w_1 \mid z_1$, $w_2 \mid m^*$ for $m^* = m^{\lfloor 2\log x \rfloor}$, and $c_1 \mid m$ and $c_2 \mid m$. Further, $l, \lambda, \widetilde{\lambda} \in \mathbb{Z} \setminus \{0\}$ with $|l| \ll N$ and $\lambda, \widetilde{\lambda} \sim \Lambda$ and $(\lambda, m^*) = (\widetilde{\lambda}, m^*) = w_2$. A and B are residue classes mod m with $(A, m) = 1$, ψ_N and ψ_{Δ_1} are non-negative, real coefficient sequences, smooth at scales N and Δ_1, and $C(n; d) = 1_{n \in E_d(q_0)}$ is the indicator function of a set E_d mod q_0 (dependent on d mod q_0) with $|E_d| \leq (q_0, \ell)$. Finally,

$$\begin{aligned}
&1 \leq H = \frac{x^\varepsilon RQ^2}{q_0 M} \ll \frac{x^{4\omega+\delta+7\varepsilon}}{q_0},\\
&x^{-4\varepsilon-\delta}N \ll R \ll x^{-2\varepsilon}N,\\
&x^{1/2-\varepsilon} \ll QR \ll x^{1/2+2\omega+\varepsilon},\\
&x^{5\varepsilon}H \ll v_1, v_2 \ll x^{\delta+5\varepsilon}H,\\
&\frac{N}{x^{\delta+55\varepsilon}H^2} \ll \Delta_1 \ll \frac{N}{x^{55\varepsilon}H^2},\\
&\frac{RQ^2H}{q_0(v_1,v_2)\Delta_1} \ll m \ll \frac{x^\delta RQ^2H}{q_0(v_1,v_2)\Delta_1},\\
&1 \ll \Lambda \ll \frac{1}{w_1(v_1,v_2)}x^{\delta+5\varepsilon}H^2.
\end{aligned}$$

For $z \in Z_5$, we set $T(z) = \max\{\frac{1}{(w_2,m)}, \frac{1}{H}\}\frac{x^{\delta+100\varepsilon}H^2N}{(v_1,v_2)\Delta_1}$ and define $\Sigma_5(z)$ to equal

$$\left|\sum_{\substack{d\\(d,m)=1\\(d,\lambda)=1\\(d,\widetilde{\lambda})=1}}\psi_{\Delta_1}(z_1d-d_0)\sum_{\substack{n,\widetilde{n}\\(n\widetilde{n},w_1c_1)=1\\((n+ld)(\widetilde{n}+ld),c_2)=1\\\lambda\widetilde{n}\equiv\widetilde{\lambda}n \mod (z_1d/w_1)\\((\lambda(\widetilde{n}+Bd)-\widetilde{\lambda}(n+Bd))/d,m)\leq T(z)}}C(n;d)C(\widetilde{n};d)\psi_N(n)\psi_N(\widetilde{n})e_m\left(\frac{A(\lambda(\widetilde{n}+Bd)-\widetilde{\lambda}(n+Bd))}{d(n+Bd)(\widetilde{n}+Bd)}\right)\right|.$$

If $\sup_{z\in Z_5}\Sigma_5(z) = O_\varepsilon\left(\min\left\{\frac{(w_2,m)}{x^{\delta+10\varepsilon}H^3}, \frac{(w_2,m)}{H^4}\right\}\frac{q_0^2(q_0,\ell)^2N^2}{x^{27\varepsilon}}\right)$, *then for $C>0$ and $a\in\mathbb{Z}$,*

$$\sum_{\substack{d\leq x^{1/2+2\omega}\\q|P(x^\delta)\\(q,a)=1}}\left|\sum_{n\equiv a(q)}(\alpha\star\beta)(n)-\frac{1}{\phi(q)}\sum_{(n,q)=1}(\alpha\star\beta)(n)\right| \ll_{C,\varepsilon}\frac{x}{(\log x)^C}.$$

Proof. This is simply a summary of Lemma 3.5, 3.6, 3.8 and 3.9: Recall that in Lemma 3.9, for given $\alpha_1\in Z_1$, $\alpha_2\in Z_2(\alpha_1)$, $\alpha_3\in Z_3(\alpha_1,\alpha_2)$ and $\alpha_4\in Z_4(\alpha_1,\alpha_2,\alpha_3)$, we defined $\Sigma(\alpha_1,\alpha_2,\alpha_3,\alpha_4)$ to equal

$$\left|\sum_{\substack{d\\w_0w_1|d\\(\frac{d}{w_1},\frac{my\widetilde{y}}{w_1^2})=1}}\psi_{\Delta_1}(d-d_0)\sum_{\substack{n,\widetilde{n}\\(n\widetilde{n},w_1r_1q_0u_1v_1v_2)=1\\((n+\ell dr_1)(\widetilde{n}+\ell dr_1),q_0q_2)=1\\y\widetilde{n}\equiv\widetilde{y}n \bmod d\\((y(\widetilde{n}+Bd)-\widetilde{y}(n+Bd))/d,m)\leq T(\alpha_1,\alpha_2,\alpha_3)}}C(n)C(\widetilde{n})\psi_N(n)\psi_N(\widetilde{n})e_m\left(\frac{A(y(\widetilde{n}+Bd)-\widetilde{y}(n+Bd))}{d(n+Bd)(\widetilde{n}+Bd)}\right)\right|.$$

Here $m = r_1q_0u_1[v_1,v_2]q_2 = r_1q_0u_1v_1v_2q_2/(v_1,v_2)$ with $r_1\asymp R/\Delta$, $u_1v_1,u_1v_2\asymp Q/q_0$, $x^{5\varepsilon}H\ll v_1,v_2\ll x^{\delta+5\varepsilon}H$ and $q_2\asymp Q/q_0$. Hence $\frac{1}{q_0(v_1,v_2)}x^{5\varepsilon}RQ^2H/\Delta\ll m\ll \frac{1}{q_0(v_1,v_2)}x^{\delta+5\varepsilon}RQ^2H/\Delta$.

We now write $z_1 = w_0w_1$. We observe that $w_1\mid z_1$, $(z_1,m)=1$ and $z_1\ll x^{5\varepsilon}\Delta_1$ when $\Sigma_4(\alpha_1,\alpha_2,\alpha_3,\alpha_4)$ is non-zero. Recall that $w_1\mid y$ and $w_1\mid\widetilde{y}$ and write $y=w_1\lambda$ and $\widetilde{y}=w_1\widetilde{\lambda}$. Then $\lambda,\widetilde{\lambda}\sim\Lambda = Y/w_1$, where $|\Lambda|\ll\frac{1}{w_1(v_1,v_2)}x^{\delta+5\varepsilon}H^2$. Write $d=z_1e$. If $(z_1/w_1,m\lambda\widetilde{\lambda})\neq 1$, then $\Sigma_4(\alpha_1,\alpha_2,\alpha_3,\alpha_4)$ is zero. Hence we assume $(z_1/w_1,m\lambda\widetilde{\lambda})=1$ and replace $(z_1e/w_1,m\lambda\widetilde{\lambda})=1$ by $(e,m\lambda\widetilde{\lambda})=1$ in the first sum in $\Sigma_4(\alpha_1,\alpha_2,\alpha_3,\alpha_4)$. Further, $w_1\lambda\widetilde{n}\equiv w_1\widetilde{\lambda}n \bmod z_1e$ is equivalent to $\lambda\widetilde{n}\equiv\widetilde{\lambda}n \bmod (z_1/w_1)e$ since $w_1\mid z_1$. Finally, since $(z_1,m)=1$ and $w_1\mid z_1$, we have $(\frac{w_1}{z_1e}(\lambda(\widetilde{n}+Bz_1e)-\widetilde{\lambda}(n+Bz_1e)),m) = (\frac{1}{e}(\lambda(\widetilde{n}+Bz_1e)-\widetilde{\lambda}(n+Bz_1e)),m)$.

So whenever $\Sigma_4(\alpha_1,\alpha_2,\alpha_3,\alpha_4)\neq 0$, there exists some $z\in Z_5$ with $\Sigma_4(\alpha_1,\alpha_2,\alpha_3,\alpha_4) = \Sigma_5(z)$. This choice of $\Sigma_5(z)$ has ℓz_1r_1 in the place of l, $r_1q_0u_1[v_1,v_2]$ in the place of c_1, q_0q_2 in the place of c_2, $Aw_1/z_1 \bmod m$ in the place of $A \bmod m$, Bz_1 in the place of B and Y/w_1 in the place of Λ.

We now assume $\sup_{z\in Z_5}\Sigma_5(z) = O_\varepsilon(\min\{\frac{(w_2,m)}{x^{\delta+10\varepsilon}H^3}, \frac{(w_2,m)}{H^4}\}\frac{q_0^2(q_0,\ell)^2N^2}{x^{27\varepsilon}})$. Then

$$\Sigma_4(\alpha_1,\alpha_2,\alpha_3,\alpha_4) \ll \sup_{z\in Z_5}\Sigma_5(z) = O_\varepsilon(\min\{\tfrac{(w_2,m)}{x^{\delta+10\varepsilon}H^3}, \tfrac{(w_2,m)}{H^4}\}\tfrac{q_0^2(q_0,\ell)^2N^2}{x^{27\varepsilon}})$$

for $\alpha_1 \in Z_1$, $\alpha_2 \in Z_2(\alpha_1)$, $\alpha_3 \in Z_3(\alpha_1,\alpha_2)$ and $\alpha_4 \in Z_4(\alpha_1,\alpha_2,\alpha_3)$. By Lemma 3.9, this implies $\Sigma_3(\alpha_1,\alpha_2,\alpha_3) = O_\varepsilon(\frac{q_0^2(q_0,\ell)^2W_1^{-1}N^2}{x^{27\varepsilon}})$. By Lemma 3.8, then $\Sigma_2(\alpha_1,\alpha_2) = O_\varepsilon(\frac{q_0(q_0,\ell)\Delta N(v_1,v_2)}{x^{10\varepsilon}})$. By Lemma 3.6, $\Sigma_1(\alpha_1) = O_\varepsilon(\frac{(q_0,\ell)RQNUV^2}{x^{4\varepsilon}})$. So by Lemma 3.5,

$$\sum_{\substack{d\le x^{1/2+2\omega}\\ q\mid P(x^\delta)\\ (q,a)=1}} \left| \sum_{n\equiv a(q)} (\alpha\star\beta)(n) - \frac{1}{\phi(q)}\sum_{(n,q)=1}(\alpha\star\beta)(n)\right| \ll_{C,\varepsilon} \frac{x}{(\log x)^C}. \qquad \square$$

3.3.4 Further adjustments

In Section 3.3.4 we simplify Σ_5 further, aiming to replace it by an expression to which the exponential sum bounds of [37], and in particular a version of Polymath's Proposition 8.4, can be applied.

3.3.4.1 Removal of $\widetilde{n}$

As a first step, we remove $\widetilde{n}$. We use $\lambda\widetilde{n} \equiv \widetilde{\lambda}n \bmod (\frac{z_1 d}{w_1})$ to replace $\widetilde{n}$ by a linear combination of d and n, and Taylor expand to separate variables n and d in $\psi_N(\widetilde{n})$.

Lemma 3.11 (Removing a linear congruence condition). *Let Z_5 and Σ_5 be as given in Lemma 3.10. For $z \in Z_5$, denote by $W_6(z)$ the set of tuples $(\psi_N^*, \psi_{\Delta_1}^*, d_\star, n_\star)$ which satisfy the following conditions: $\psi_N^*(n)$ and $\psi_{\Delta_1}^*(d_1)$ are coefficient sequences, smooth at scales N and Δ_1. $d_\star$ and $n_\star$ are residue classes mod q_0.*

For given $z \in Z_5$, $w \in W_6(z)$ and $k \in \mathbb{Z}$ with $|k| \ll \frac{w_1\Lambda N}{x^{5\varepsilon}\Delta_1}$, we set $\Sigma_6(z,w,k)$ equal to

$$\left| \sum_{\substack{d\\ d\equiv d_\star(q_0)\\ (d,m\lambda\widetilde{\lambda})=1}} \psi_{\Delta_1}^*(z_1 d - d_0) \sideset{}{^\bigstar}\sum_{\substack{n\\ n\equiv n_\star(q_0)}} \psi_N^*(n)\, e_m\!\left(\frac{A((\frac{z_1}{w_1})k + (\lambda-\widetilde{\lambda})B)}{(n+Bd)(\frac{1}{\lambda}(\widetilde{\lambda}n + (\frac{z_1}{w_1})kd) + Bd)}\right)\right|,$$

where $\sum^\bigstar$ denotes summation over integers n with $\lambda \mid \widetilde{\lambda}n + k(\frac{z_1}{w_1})d$, $(n, w_1c_1) = 1$, $(\frac{1}{\lambda}(\widetilde{\lambda}n + k(\frac{z_1}{w_1})d), w_1c_1) = 1$, $(n+ld, c_2) = 1$ and $(\frac{1}{\lambda}(\widetilde{\lambda}n + k(\frac{z_1}{w_1})d) + ld, c_2) = 1$.

We denote by Ξ the set of sequences (ξ_k) with $\xi_k \in [0,\infty)$ and $\displaystyle\sideset{}{^*}\sum_{|k|\ll\frac{w_1\Lambda N}{x^{5\varepsilon}\Delta_1}} \xi_k \ll_\varepsilon 1$.

($\sum^*$ indicates that we sum over k with $w_2 \mid k$ and $((\frac{z_1}{w_1})k + (\lambda-\widetilde{\lambda})B, m) \le T(z)$.)

Suppose that for every $z \in Z_5$, there exists $(\xi_k) \in \Xi$ such that for all k with $|k| \ll \frac{w_1 \Lambda N}{x^{5\varepsilon}\Delta_1}$, $w_2 \mid k$ and $((\frac{z_1}{w_1})k + (\lambda - \widetilde{\lambda})B, m) \le T(z)$,

$$\sup_{w \in W_6(z)} \Sigma_6(z,w,k) \ll_\varepsilon \xi_k \left(\min\left\{ \frac{(w_2,m)}{x^{\delta+10\varepsilon}H^3}, \frac{(w_2,m)}{H^4} \right\} \frac{q_0(q_0,\ell)N^2}{x^{27\varepsilon}} \right).$$

Then also $\sup_{z \in Z_5} \Sigma_5(z) \ll_\varepsilon \min\left\{ \frac{(w_2,m)}{x^{\delta+10\varepsilon}H^3}, \frac{(w_2,m)}{H^4} \right\} \frac{q_0^2(q_0,\ell)^2N^2}{x^{27\varepsilon}}$.

Proof. Our goal is to show $\Sigma_5(z) \ll_\varepsilon \min\{\frac{(w_2,m)}{x^{\delta+10\varepsilon}H^3}, \frac{(w_2,m)}{H^4}\} \frac{q_0^2(q_0,\ell)^2N^2}{x^{27\varepsilon}}$. Recall $\Sigma_5(z)$ is

$$\left| \sum_{\substack{d \\ (d,m)=1 \\ (d,\lambda)=1 \\ (d,\widetilde{\lambda})=1}} \psi_{\Delta_1}(z_1 d - d_0) \sum_{\substack{n,\widetilde{n} \\ (n\widetilde{n}, w_1 c_1)=1 \\ ((n+ld)(\widetilde{n}+ld), c_2)=1 \\ \lambda\widetilde{n} \equiv \widetilde{\lambda} n \bmod (z_1 d/w_1) \\ ((\lambda(\widetilde{n}+Bd) - \widetilde{\lambda}(n+Bd))/d, m) \le T(z)}} C(n;d) C(\widetilde{n};d) \psi_N(n)\psi_N(\widetilde{n}) e_m\!\left(\frac{A(\lambda(\widetilde{n}+Bd) - \widetilde{\lambda}(n+Bd))}{d(n+Bd)(\widetilde{n}+Bd)} \right) \right|.$$

Consider the condition $\lambda\widetilde{n} \equiv \widetilde{\lambda} n \bmod (z_1 d/w_1)$. Rearranging, we get

$$\widetilde{n} = \frac{\widetilde{\lambda} n + k(\frac{z_1}{w_1})d}{\lambda}$$

for some $k \in \mathbb{Z}$ with $\lambda \mid \widetilde{\lambda} n + k(\frac{z_1}{w_1})d$. Since $\psi_N(n)$ is a coefficient sequence at scale N, the contribution of n and $\widetilde{n}$ to $\Sigma_5(z)$ is only non-zero when $1 \le n, \widetilde{n} \ll N$. Since $\lambda, \widetilde{\lambda} \sim \Lambda$, then $|\lambda\widetilde{n} - \widetilde{\lambda} n| \ll \Lambda N$. Further, $\psi_{\Delta_1}(d_1)$ is a coefficient sequence at scale Δ_1 and $d_0 \asymp x^{5\varepsilon}\Delta_1$. Hence the contribution of d to $\Sigma_5(z)$ is only non-zero when $z_1 d \asymp x^{5\varepsilon}\Delta_1$. However, $k = (\frac{w_1}{z_1 d})(\lambda\widetilde{n} - \widetilde{\lambda} n)$ and so $|k| \ll \frac{w_1 \Lambda N}{x^{5\varepsilon}\Delta_1}$ whenever the contribution of k to $\Sigma_5(z)$ is non-zero. Observe that now $(\lambda(\widetilde{n}+Bd) - \widetilde{\lambda}(n+Bd)) = ((\frac{z_1}{w_1})kd + (\lambda - \widetilde{\lambda})Bd)$ and

$$e_m\left(\frac{A(\lambda(\widetilde{n}+Bd) - \widetilde{\lambda}(n+Bd))}{d(n+Bd)(\widetilde{n}+Bd)} \right) = e_m\left(\frac{A((\frac{z_1}{w_1})k + (\lambda - \widetilde{\lambda})B)}{(n+Bd)(\frac{1}{\lambda}(\widetilde{\lambda} n + (\frac{z_1}{w_1})kd) + Bd)} \right).$$

Hence we may write

$$\Sigma_5(z) = \left| \sum_{\substack{k \\ |k| \ll \frac{w_1 \Lambda N}{x^{5\varepsilon}\Delta_1} \\ ((\frac{z_1}{w_1})k + (\lambda-\widetilde{\lambda})B, m) \le T(z)}} \sum_{\substack{d \\ (d, m\lambda\widetilde{\lambda})=1}} \psi_{\Delta_1}(z_1 d - d_0) \sideset{}{^\star}\sum_n C(n;d) C\left(\frac{\widetilde{\lambda} n + k(\frac{z_1}{w_1})d}{\lambda}; d \right) (\star) \right|$$

$$\text{with } (\star) = \psi_N(n)\, \psi_N\left(\frac{\widetilde{\lambda} n + k(\frac{z_1}{w_1})d}{\lambda} \right) e_m\left(\frac{A((\frac{z_1}{w_1})k + (\lambda - \widetilde{\lambda})B)}{(n+Bd)(\frac{1}{\lambda}(\widetilde{\lambda} n + (\frac{z_1}{w_1})kd) + Bd)} \right).$$

Since $\psi_N(n)$ is a smooth coefficient sequence, there exists a smooth $\psi : \mathbb{R} \to \mathbb{C}$, supported on $[c, C]$, with $|\psi^{(j)}(t)| \ll_j (\log x)^{O_j(1)}$ and $\psi_N(n) = \psi(\frac{n}{N})$. We focus on

$$\psi_N\left(\frac{\widetilde{\lambda}n + k(\frac{z_1}{w_1})d}{\lambda}\right) = \psi\left(\frac{\widetilde{\lambda}n + (\frac{k}{w_1})d_0}{\lambda N} + \frac{(\frac{k}{w_1})(z_1 d - d_0)}{\lambda N}\right).$$

Since $\psi_{\Delta_1}(d_1)$ is smooth at scale Δ_1, the contribution of d to $\Sigma_5(z)$ is zero unless $|z_1 d - d_0| \ll \Delta_1$. But then

$$\left|\frac{(\frac{k}{w_1})(z_1 d - d_0)}{\lambda N}\right| \ll \frac{\left(\frac{w_1 \Lambda N}{x^{5\varepsilon}\Delta_1}\right)\Delta_1}{w_1 \Lambda N} \ll \frac{1}{x^{5\varepsilon}}.$$

Choosing $J = \lceil \frac{20}{\varepsilon} \rceil$ and using Taylor expansions, we may write

$$\psi_N\left(\frac{\widetilde{\lambda}n + k(\frac{z_1}{w_1})d}{\lambda}\right) = \sum_{j=0}^{J} \frac{1}{j!}\left(\frac{(\frac{k}{w_1})(z_1 d - d_0)}{\lambda N}\right)^j \psi^{(j)}\left(\frac{\widetilde{\lambda}n + (\frac{k}{w_1})d_0}{\lambda N}\right) + O_\varepsilon(x^{-100}). \tag{3.19}$$

Consider the function $\psi_\star(t) = \psi(t)\psi^{(j)}((\frac{\widetilde{\lambda}}{\lambda})t + (\frac{kd_0}{w_1 \lambda N}))$. Since $\lambda, \widetilde{\lambda} \sim \Lambda$, we have $|\frac{\widetilde{\lambda}}{\lambda}| \ll 1$. So the Leibniz product rule and $|\psi^{(i)}(t)| \ll_i (\log x)^{O_i(1)}$ tell us that $\psi_\star(t)$ is also smooth with $|\psi_\star^{(i)}(t)| \ll_i (\log x)^{O_i(1)}$. Further, the support of $\psi_\star(t)$ is contained in the support of $\psi(t)$. Overall, we see that $\psi_\star(n/N)$ is a coefficient sequence which is smooth at scale N.

Since $\psi_{\Delta_1}(d_1)$ is a smooth coefficient sequence, there also exists a smooth function $\phi : \mathbb{R} \to \mathbb{C}$, supported on $[c, C]$, with $|\phi^{(j)}(t)| \ll_j (\log x)^{O_j(1)}$ and $\psi_{\Delta_1}(d_1) = \phi(d_1/\Delta_1)$. Now set $\phi_\star(t) = \frac{1}{j!}((\frac{k}{w_1})(\frac{\Delta_1}{\lambda N})t)^j \phi(t)$. We already know that $|(\frac{k}{w_1})(\frac{\Delta_1}{\lambda N})| \ll \frac{1}{x^{5\varepsilon}} \ll 1$ and so by the product rule we find that $|\phi_\star^{(i)}(t)| \ll_i (\log x)^{O_i(1)}$. Hence the coefficient sequence $\phi_\star(d_1/\Delta_1)$ is smooth at scale Δ_1.

We substitute the RHS of (3.19) for $\psi_N((\widetilde{\lambda}n + \frac{z_1}{w_1}kd)/\lambda)$ in $\Sigma_5(z)$. The contribution of $1_{[cN,CN]}(n)O(x^{-100})$ is trivially bounded by $O_\varepsilon(x^{-90})$, while the substitution of $\frac{1}{j!}((\frac{k}{w_1 \lambda N})(z_1 d - d_0))^j \psi^{(j)}(\frac{1}{\lambda N}(\widetilde{\lambda}n + (\frac{k}{w_1})d_0))$ creates a new smooth coefficient sequence of the form $\psi_\star(\frac{n}{N})$ and a shifted smooth coefficient sequence of the form $\phi_\star(\frac{1}{\Delta_1}(z_1 d - d_0))$. Note that $J = \lceil \frac{20}{\varepsilon} \rceil = O_\varepsilon(1)$. We define

$$\Upsilon_1(\psi_N^\star, \psi_{\Delta_1}^\star) = \left| \sum_{\substack{k \\ |k| \ll \frac{w_1 \Lambda N}{x^{5\varepsilon}\Delta_1} \\ ((\frac{z_1}{w_1})k + (\lambda - \widetilde{\lambda})B, m) \le T(z)}} \sum_{\substack{d \\ (d, m\lambda\widetilde{\lambda}) = 1}} \psi_{\Delta_1}^*(z_1 d - d_0) \sideset{}{^\star}\sum_{n} \psi_N^*(n)\ (*)_1 \right|$$

with $(*)_1 = C(n;d)C\left(\frac{\widetilde{\lambda}n + k(\frac{z_1}{w_1})d}{\lambda}; d\right) e_m\left(\frac{A((\frac{z_1}{w_1})k + (\lambda - \widetilde{\lambda})B)}{(n+Bd)(\frac{1}{\lambda}(\widetilde{\lambda}n + (\frac{z_1}{w_1})kd) + Bd)}\right)$.

Taking the supremum over all coefficient sequences $\psi_N^*(n)$ and $\psi_{\Delta_1}^*(d_1)$ which are smooth at scales N and Δ_1, we get

$$\Sigma_5(z) \ll_\varepsilon \sup_{(\psi_N^*, \psi_{\Delta_1}^*)} \Upsilon_1(\psi_N^\star, \psi_{\Delta_1}^\star) + O_\varepsilon(x^{-90}).$$

Hence we have the desired bound on $\Sigma_5(z)$ if we can show that

$$\Upsilon_1 \ll_\varepsilon \min\left\{\tfrac{(w_2,m)}{x^{\delta+10\varepsilon}H^3}, \tfrac{(w_2,m)}{H^4}\right\} \tfrac{q_0^2(q_0,\ell)^2N^2}{x^{27\varepsilon}}.$$

Next we partition according to the value of k and write

$$\Upsilon_2^{(k)} = \left| \sum_{\substack{d \\ (d, m\lambda\widetilde{\lambda})=1}} \psi_{\Delta_1}^*(z_1 d - d_0) \sideset{}{^\star}\sum_n \psi_N^*(n)\ (*)_1 \right|.$$

Suppose that (ξ_k) is a sequence with $\sum_{|k| \ll \frac{w_1 \Lambda N}{x^{5\varepsilon}\Delta_1}} \xi_k 1_{((\frac{z_1}{w_1})k + (\lambda-\widetilde{\lambda})B, m) \le T(z)} = O_\varepsilon(1)$. If $\Upsilon_2^{(k)} = O_\varepsilon(\xi_k I)$ for each $k \in \mathbb{Z}$ with $|k| \ll \frac{w_1\Lambda N}{x^{5\varepsilon}\Delta_1}$ and $((\frac{z_1}{w_1})k + (\lambda - \widetilde{\lambda})B, m) \le T(z)$, then also $\Upsilon_1 = O_\varepsilon(I)$.

Finally, we also remove $C(n;d)$. Recall that $C(n;d)$ is the indicator function of a set E_d mod q_0, which depends on the value of d mod q_0. In particular, if we fix the values of n and d mod q_0, then $C(n;d)$ is constant. Since $\lambda, \widetilde{\lambda}$ and k are all already fixed in $\Upsilon_2^{(k)}$, the function $C((\widetilde{\lambda}n + k(\frac{w_1}{z_1})d)/\lambda; d)$ is also constant when n and d are fixed mod q_0. But E_d has at most (q_0, ℓ) elements for each d mod q_0 and thus $C(n;d)$ is only non-zero for at most $q_0(q_0,\ell)$ choices of n and d mod q_0. We recall that $\Sigma_6(z, (\psi_N^*, \psi_{\Delta_1}^*, d_\star, n_\star), k)$ is defined to equal

$$\left| \sum_{\substack{d \\ d \equiv d_\star (q_0) \\ (d, m\lambda\widetilde{\lambda})=1}} \psi_{\Delta_1}^*(z_1 d - d_0) \sideset{}{^\star}\sum_{\substack{n \\ n \equiv n_\star (q_0)}} \psi_N^*(n)\, e_m\left(\frac{A((\frac{z_1}{w_1})k + (\lambda - \widetilde{\lambda})B)}{(n+Bd)(\frac{1}{\lambda}(\widetilde{\lambda}n + (\frac{z_1}{w_1})kd) + Bd)}\right) \right|,$$

and observe that

$$\sup_{(\psi_N^*, \psi_{\Delta_1}^*)} \Upsilon_2^{(k)} \ll q_0(q_0,\ell) \sup_{(\psi_N^*, \psi_{\Delta_1}^*, d_\star, n_\star)} \Sigma_6(z, (\psi_N^*, \psi_{\Delta_1}^*, d_\star, n_\star), k).$$

The sum $\sideset{}{^\star}\sum$ is restricted to n which satisfy $\lambda \mid \widetilde{\lambda}n + k(\frac{z_1}{w_1})d$. But $w_2 \mid \lambda$ and $w_2 \mid \widetilde{\lambda}$, while $(z_1, m) = (d,m) = 1$, so that $((\frac{z_1}{w_1})d, w_2) = 1$. Hence if $w_2 \nmid k$, we immediately get $\Sigma_6(z,w,k)=0$. Additionally, we only need to consider k which satisfy $((\frac{z_1}{w_1})k+(\lambda-\widetilde{\lambda})B,m)\le T(z)$. But by the assumptions of Lemma 3.11, there exists $(\xi_k)\in\Xi$ so that for all $|k|\ll\frac{w_1\Lambda N}{x^{5\varepsilon}\Delta_1}$ with $w_2\mid k$ and $((\frac{z_1}{w_1})k+(\lambda-\widetilde{\lambda})B,m)\le T(z)$,

$$\sup_{w\in W_6(z)}\Sigma_6(z,w,k)\ll_\varepsilon \xi_k\left(\min\left\{\frac{(w_2,m)}{x^{\delta+10\varepsilon}H^3},\frac{(w_2,m)}{H^4}\right\}\frac{q_0(q_0,\ell)N^2}{x^{27\varepsilon}}\right).$$

Then also $\Upsilon_2^{(k)}\ll_\varepsilon \xi_k\min\{\frac{(w_2,m)}{x^{\delta+10\varepsilon}H^3},\frac{(w_2,m)}{H^4}\}\frac{q_0^2(q_0,\ell)^2N^2}{x^{27\varepsilon}}$ and so

$$\Upsilon_1\ll_\varepsilon \min\left\{\tfrac{(w_2,m)}{x^{\delta+10\varepsilon}H^3},\tfrac{(w_2,m)}{H^4}\right\}\tfrac{q_0^2(q_0,\ell)^2N^2}{x^{27\varepsilon}}.$$

It follows that $\Sigma_5(z)\ll_\varepsilon\min\{\frac{(w_2,m)}{x^{\delta+10\varepsilon}H^3},\frac{(w_2,m)}{H^4}\}\frac{q_0^2(q_0,\ell)^2N^2}{x^{27\varepsilon}}$, as proposed. □

3.3.4.2 Divisibility by λ

Next we remove the condition $\lambda\mid\widetilde{\lambda}n+(\frac{z_1}{w_1})kd$.

Lemma 3.12 (Removing a divisibility condition)**.** *Let Z_5 be as given in Lemma 3.10 and let $W_6(z)$ and $\Sigma_6(z,w,k)$ be as given in Lemma 3.11. For $z\in Z_5$ and $k\in\mathbb{Z}$, denote by $W_7(z,k)$ the set of tuples $(\psi_{\frac{\Delta^*}{z_1}},\psi_{(\frac{\lambda^*}{\lambda})N},F_k,G_k)$ which satisfy the following: $\psi_{\frac{\Delta^*}{z_1}}(d)$ and $\psi_{(\frac{\lambda^*}{\lambda})N}(n)$ are shifted smooth coefficient sequences at scales $(\frac{\Delta^*}{z_1})$ and $(\frac{\lambda^*}{\lambda})N$, with corresponding shifts $x_0\ll x$. Here $\Delta^*=\min\{\frac{N}{\Lambda x^{5\varepsilon}},\Delta_1\}$ and $\lambda^*=(\lambda,\widetilde{\lambda})$. F_k and G_k are residue classes mod m with*

$$\left(\tfrac{\lambda^*}{\widetilde{\lambda}}\right)(G_k+B)-\left(\tfrac{\lambda^*}{\lambda}\right)(F_k+B)\equiv\left\{\left(\tfrac{\lambda^*}{\lambda}\right)\left(\tfrac{w_2}{\widetilde{\lambda}}\right)\right\}\left\{\tfrac{1}{w_2}\left(\left(\tfrac{z_1}{w_1}\right)k+(\lambda-\widetilde{\lambda})B\right)\right\}\bmod m. \tag{3.20}$$

For $k\in\mathbb{Z}$, $z\in Z_5$, $w_6\in W_6(z)$ and $w_7\in W_7(z,k)$, we let $\Sigma_7(z,w_6,w_7,k)$ equal

$$\left|\sum_{\substack{d\\ d\equiv d_\star(q_0)\\ (d,m\lambda\widetilde{\lambda})=1}}\psi_{\frac{\Delta^*}{z_1}}(d)\sum_{n_1}^{\bigstar_1}\psi_{(\frac{\lambda^*}{\lambda})N}(n_1)\,e_m\left(\frac{(\frac{\lambda^*}{\lambda})(\frac{\lambda^*}{\widetilde{\lambda}})A((\frac{z_1}{w_1})k+(\lambda-\widetilde{\lambda})B)}{(n_1+(\frac{\lambda^*}{\lambda})(F_k+B)d)(n_1+(\frac{\lambda^*}{\widetilde{\lambda}})(G_k+B)d)}\right)\right|,$$

where $\bigstar_1$ denotes a sum over n_1 with $(\frac{\lambda}{\lambda^*})n_1+F_kd\equiv n_\star(q_0)$, $((\frac{\lambda}{\lambda^*})n_1+F_kd,w_1c_1)=((\frac{\widetilde{\lambda}}{\lambda^*})n_1+G_kd,w_1c_1)=1$, $((\frac{\lambda}{\lambda^*})n_1+F_kd+ld,c_2)=1$ and $((\frac{\widetilde{\lambda}}{\lambda^*})n_1+G_kd+ld,c_2)=1$.

Then for all $z\in Z_5$, $w_6\in W_6(z)$ and $k\in\mathbb{Z}$ which satisfy $|k|\ll\frac{w_1\Lambda N}{x^{5\varepsilon}\Delta_1}$, $w_2\mid k$ and $((\frac{z_1}{w_1})k+(\lambda-\widetilde{\lambda})B,m)\le T(z)$, we have

$$\Sigma_6(z,w_6,k)\ll_\varepsilon\sup_{w_7\in W_7(z,k)}\left(\frac{\Delta_1}{\Delta^*}\right)(\Sigma_7(z,w_6,w_7,k)+O_\varepsilon(x^{-90})).$$

In particular, to show that $\Sigma_6(z,w_6,k) \ll_\varepsilon \xi_k\left(\min\left\{\frac{(w_2,m)}{x^{\delta+10\varepsilon}H^3},\frac{(w_2,m)}{H^4}\right\}\frac{q_0(q_0,\ell)N^2}{x^{27\varepsilon}}\right)$, it suffices to prove that for all $w_7\in W_7(z,k)$

$$\Sigma_7(z,w_6,w_7,k)+O_\varepsilon(x^{-90}) \ll_\varepsilon \xi_k\left(\frac{\Delta^*}{\Delta_1}\right)\left(\min\left\{\frac{(w_2,m)}{x^{\delta+10\varepsilon}H^3},\frac{(w_2,m)}{H^4}\right\}\frac{q_0(q_0,\ell)N^2}{x^{27\varepsilon}}\right).$$

Proof. Our goal is to bound $\Sigma_6(z,w_6,k)$, which is given by

$$\Sigma_6(z,w_6,k)=\left|\sum_{\substack{d\\ d\equiv d_\star(q_0)\\ (d,m\lambda\widetilde{\lambda})=1}}\psi^*_{\Delta_1}(z_1d-d_0)\sum_{\substack{n\\ n\equiv n_\star(q_0)\\ \lambda|\widetilde{\lambda}n+k(\frac{z_1}{w_1})d\\ (n,w_1c_1)=(n+ld,c_2)=1\\ (\frac{1}{\lambda}(\widetilde{\lambda}n+k(\frac{z_1}{w_1})d),w_1c_1)=1\\ (\frac{1}{\lambda}(\widetilde{\lambda}n+k(\frac{z_1}{w_1})d)+ld,c_2)=1}}\psi^*_N(n)\,e_m\left(\frac{A((\frac{z_1}{w_1})k+(\lambda-\widetilde{\lambda})B)}{(n+Bd)(\frac{1}{\lambda}(\widetilde{\lambda}n+(\frac{z_1}{w_1})kd)+Bd)}\right)\right|.$$

Here we assume that $k\in\mathbb{Z}$ with $|k|\ll\frac{w_1\Lambda N}{x^{5\varepsilon}\Delta_1}$, $w_2\mid k$ and $((\frac{z_1}{w_1})k+(\lambda-\widetilde{\lambda})B,m)\le T(z)$.

We now consider the condition $\lambda\mid\widetilde{\lambda}n+(\frac{z_1}{w_1})kd$. Write $\lambda^*=(\lambda,\widetilde{\lambda})$. The divisibility condition implies that $\lambda^*\mid(\frac{z_1}{w_1})kd$. But $(d,m\lambda\widetilde{\lambda})=1$ in Σ_6. So if $\lambda^*\nmid(\frac{z_1}{w_1})k$, then $\Sigma_6(z,w_6,k)=0$. Hence we may assume that $\lambda^*\mid(\frac{z_1}{w_1})k$. Note that $\lambda\mid\widetilde{\lambda}n+(\frac{z_1}{w_1})kd$ is equivalent to $\frac{\lambda}{\lambda^*}\mid\frac{\widetilde{\lambda}}{\lambda^*}n+(\frac{z_1k}{w_1\lambda^*})d$. But since $(\lambda/\lambda^*,\widetilde{\lambda}/\lambda^*)=1$, this is further equivalent to $n\equiv-(\frac{\widetilde{\lambda}}{\lambda^*})^{-1}(\frac{z_1k}{w_1\lambda^*})d \bmod \frac{\lambda}{\lambda^*}$. Choose $F_k\in\{1,\dots,\frac{\lambda}{\lambda^*}\}$ with $F_k\equiv-(\frac{\widetilde{\lambda}}{\lambda^*})^{-1}(\frac{z_1k}{w_1\lambda^*})$ mod $\frac{\lambda}{\lambda^*}$. Then $n=F_kd+(\frac{\lambda}{\lambda^*})n_1$ for some $n_1\in\mathbb{Z}$. Further, $\frac{1}{\lambda}(\widetilde{\lambda}n+(\frac{z_1}{w_1})kd)=\frac{1}{\lambda}(\widetilde{\lambda}(F_kd+(\frac{\lambda}{\lambda^*})n_1)+(\frac{z_1}{w_1})kd)$. But there exists $G_k\in\mathbb{Z}$ such that $(\frac{\widetilde{\lambda}}{\lambda^*})F_k=-\frac{1}{\lambda^*}(\frac{z_1}{w_1})k+G_k(\frac{\lambda}{\lambda^*})$ and so $\frac{1}{\lambda}(\widetilde{\lambda}F_kd+(\frac{z_1}{w_1})kd)=G_kd$. Hence

$$n=\left(\tfrac{\lambda}{\lambda^*}\right)n_1+F_kd \qquad\text{and}\qquad \tfrac{1}{\lambda}\left(\widetilde{\lambda}n+\left(\tfrac{z_1}{w_1}\right)kd\right)=\left(\tfrac{\widetilde{\lambda}}{\lambda^*}\right)n_1+G_kd.$$

Therefore, $\Sigma_6(z,w_6,k)$ transforms as follows:

$$\Sigma_6=\left|\sum_{\substack{d\\ d\equiv d_\star(q_0)\\ (d,m\lambda\widetilde{\lambda})=1}}\psi^*_{\Delta_1}(z_1d-d_0)\sum_{\substack{n_1\\ (\frac{\lambda}{\lambda^*})n_1+F_kd\equiv n_\star(q_0)\\ ((\frac{\lambda}{\lambda^*})n_1+F_kd,w_1c_1)=1\\ ((\frac{\widetilde{\lambda}}{\lambda^*})n_1+G_kd,w_1c_1)=1\\ ((\frac{\lambda}{\lambda^*})n_1+F_kd+ld,c_2)=1\\ ((\frac{\widetilde{\lambda}}{\lambda^*})n_1+G_kd+ld,c_2)=1}}\psi^*_N((\tfrac{\lambda}{\lambda^*})n_1+F_kd)\,e_m\left(\frac{A((\frac{z_1}{w_1})k+(\lambda-\widetilde{\lambda})B)}{((\frac{\lambda}{\lambda^*})n_1+F_kd+Bd)((\frac{\widetilde{\lambda}}{\lambda^*})n_1+G_kd+Bd)}\right)\right|.$$

Recall that $(\lambda,m^*)=(\widetilde{\lambda},m^*)=w_2$. Here w_2 is the largest factor of λ (and $\widetilde{\lambda}$) which only consists of powers of primes that divide m. Since $w_2\mid\lambda^*$, this further implies that $(\frac{\lambda}{\lambda^*})$ and $(\frac{\widetilde{\lambda}}{\lambda^*})$ are coprime to m.

$$\Sigma_6=\left|\sum_{\substack{d\\ d\equiv d_\star(q_0)\\ (d,m\lambda\widetilde{\lambda})=1}}\psi^*_{\Delta_1}(z_1d-d_0)\sum_{n_1}^{\bigstar_1}\psi^*_N((\tfrac{\lambda}{\lambda^*})n_1+F_kd)e_m\left(\frac{(\frac{\lambda^*}{\lambda})(\frac{\lambda^*}{\widetilde{\lambda}})A((\frac{z_1}{w_1})k+(\lambda-\widetilde{\lambda})B)}{(n_1+(\frac{\lambda^*}{\lambda})(F_k+B)d)(n_1+(\frac{\lambda^*}{\widetilde{\lambda}})(G_k+B)d)}\right)\right|.$$

Observe now that $(\frac{\lambda^*}{\widetilde{\lambda}})(G_k + B) - (\frac{\lambda^*}{\lambda})(F_k + B) = (\frac{\lambda^*}{\lambda\widetilde{\lambda}})(\frac{z_1}{w_1})k + (\frac{\lambda^*}{\lambda\widetilde{\lambda}})(\lambda - \widetilde{\lambda})B$. We assumed that $w_2 \mid k$. Hence we may write

$$\left(\tfrac{\lambda^*}{\widetilde{\lambda}}\right)(G_k + B) - \left(\tfrac{\lambda^*}{\lambda}\right)(F_k + B) \equiv \left\{\left(\tfrac{\lambda^*}{\lambda}\right)\left(\tfrac{w_2}{\widetilde{\lambda}}\right)\right\}\left\{\tfrac{1}{w_2}\left(\left(\tfrac{z_1}{w_1}\right)k + (\lambda - \widetilde{\lambda})B\right)\right\} \mod m, \tag{3.21}$$

where both of the two factors surrounded by $\{\}$–brackets are well-defined residue classes mod m.

Next we use Taylor expansions to treat $\psi_N^*((\frac{\lambda}{\lambda^*})n_1 + F_k d)$, which depends on both n_1 and d. Currently, the summation range of d is still a bit too long, so we first use a partition of unity to shorten it: Choose a smooth function $\phi : \mathbb{R} \to [0, \infty)$ with $\mathrm{supp}(\phi) \subseteq [-2, 2]$ and $\phi(t) \geq 1$ on $[-1, 1]$. Define $\phi_m : \mathbb{R} \to [0, \infty)$ by setting $\phi_m(t) = (\sum_{j=-4}^{4} \phi(t - m + j))^{-1}\phi(t - m)$ on $[m - 3, m + 3]$ and $\phi_m(t) = 0$ elsewhere. The function ϕ_m is well-defined and smooth since $(\sum_{j=-4}^{4} \phi(t - m + j))^{-1} > 0$ on $[m - 3, m + 3]$. Further, for every $t \in \mathbb{R}$, $\sum_{m=-\infty}^{\infty} \phi_m(t) = 1$. Now recall that $\psi_{\Delta_1}^*(d_1)$ is a smooth coefficient sequence and there is a smooth function $\psi : \mathbb{R} \to \mathbb{C}$, supported on $[c, C]$, with $|\psi^{(j)}(t)| \ll_j (\log x)^{O_j(1)}$ and $\psi_{\Delta_1}^*(d_1) = \psi(\frac{d_1}{\Delta_1})$. Set

$$\Delta^* = \min\left\{\frac{N}{\Lambda x^{5\varepsilon}}, \Delta_1\right\}$$

and observe that

$$\begin{aligned}
\psi_{\Delta_1}^*(z_1 d - d_0) &= \psi\left(\left(\frac{\Delta^*}{\Delta_1}\right)\frac{(z_1 d - d_0)}{\Delta^*}\right)\sum_{m=-\infty}^{\infty}\phi_m\left(\frac{z_1 d - d_0}{\Delta^*}\right) \\
&= \sum_{m=-\infty}^{\infty}\psi\left(\left(\frac{\Delta^*}{\Delta_1}\right)\frac{(d - d_0/z_1)}{(\Delta^*/z_1)}\right)\phi_5\left(\frac{(d - d_0/z_1)}{(\Delta^*/z_1)} - (m - 5)\right) \\
&= \sum_{m \ll \Delta_1/\Delta^*}\tau_m\left(\frac{d - x_m}{(\Delta^*/z_1)}\right),
\end{aligned}$$

where $x_m = d_0/z_1 + (\Delta^*/z_1)(m - 5) \in \mathbb{R}$ and $\tau_m(t) = \psi((\frac{\Delta^*}{\Delta_1})(t + m - 5))\phi_5(t)$. In the last line we used that τ_m is identical to zero unless $m \ll \Delta_1/\Delta^*$. The support of τ_m is contained in $[3, 7]$ and since $\Delta^*/\Delta_1 \ll 1$, we also have $|\tau^{(j)}(t)| \ll_j (\log x)^{O_j(1)}$. Thus $\tau_m((d - x_m)/(\Delta^*/z_1))$ is a coefficient sequence which is shifted smooth at scale Δ^*/z_1. We then define

$$\Upsilon_0(\psi^{\circ}_{\frac{\Delta^*}{z_1}}) = \left|\sum_{\substack{d \\ d \equiv d_\star (q_0) \\ (d, m\lambda\widetilde{\lambda}) = 1}}\psi^{\circ}_{\frac{\Delta^*}{z_1}}(d)\sum_{n_1}^{\bigstar_1}\psi_N^*((\tfrac{\lambda}{\lambda^*})n_1 + F_k d)\, e_m\!\left(\frac{(\frac{\lambda^*}{\lambda})(\frac{\lambda^*}{\widetilde{\lambda}})A((\frac{z_1}{w_1})k + (\lambda - \widetilde{\lambda})B)}{(n_1 + (\frac{\lambda^*}{\lambda})(F_k + B)d)(n_1 + (\frac{\lambda^*}{\widetilde{\lambda}})(G_k + B)d)}\right)\right|.$$

We now substitute $\sum \tau_m(\frac{d-x_m}{(\Delta^*/z_1)})$ for $\psi^*_{\Delta_1}(z_1 d - d_0)$ in Σ_6. Taking a supremum over all coefficient sequences $\psi^\circ_{\frac{\Delta^*}{z_1}}(d)$ which are shifted smooth at scale $\frac{\Delta^*}{z_1}$ and have shift $x_0 \ll x^{5\varepsilon}\Delta_1$, we find that

$$\Sigma_6 \ll_\varepsilon \left(\frac{\Delta_1}{\Delta^*}\right) \sup_{\psi^\circ_{\frac{\Delta^*}{z_1}}} \Upsilon_0(\psi^\circ_{\frac{\Delta^*}{z_1}}).$$

We have thus shortened the summation range of d. In particular, for a given Υ_0, there exists $x_0 \in \mathbb{R}$ with $x_0 \ll x^{5\varepsilon}\Delta_1$ such that the contribution of d to Υ_0 is zero unless $|d - x_0| \ll \Delta^*/z_1$. Note that whenever d contributes to Υ_0,

$$\frac{|F_k(d-x_0)|}{N} \ll \left|\frac{\lambda}{\lambda^*}\right| \frac{\Delta^*}{z_1 N} \ll \frac{1}{x^{5\varepsilon}}$$

We set $\psi^*_N(n) = \phi(\frac{n}{N})$. Choosing $J = \lceil \frac{20}{\varepsilon} \rceil$ and using Taylor expansions,

$$\psi^*_N\big((\tfrac{\lambda}{\lambda^*})\, n_1 + F_k d\big) = \sum_{j=0}^{J} \frac{1}{j!}\left(\frac{F_k(d-x_0)}{N}\right)^j \phi^{(j)}\!\left(\frac{n_1 + (\frac{\lambda^*}{\lambda})F_k x_0}{(\frac{\lambda^*}{\lambda})N}\right) + O_\varepsilon(x^{-100} 1_{n_1 \ll x}). \tag{3.22}$$

The sequence $\phi^{(j)}(\frac{n_1 + (\frac{\lambda^*}{\lambda})F_k x_0}{(\frac{\lambda^*}{\lambda})N})$ is a shifted smooth coefficient sequence in n_1 at scale $(\frac{\lambda^*}{\lambda})N$, while the sequence $\frac{1}{j!}\left(\frac{F_k(\Delta^*/z_1)}{N}\frac{(d-x_0)}{(\Delta^*/z_1)}\right)^j \psi^\circ_{\frac{\Delta^*}{z_1}}(d)$ is a shifted smooth coefficient sequence in d at scale $\frac{\Delta^*}{z_1}$. We now substitute the RHS of (3.22) for $\psi^*_N((\frac{\lambda}{\lambda^*})n_1 + F_k d)$ in Υ_0. The contribution of $O_\varepsilon(x^{-100}1_{n_1 \ll x})$ is trivially bounded by $O_\varepsilon(x^{-90})$. For the treatment of other terms, set

$$\Upsilon_1(\psi_{(\frac{\lambda^*}{\lambda})N}, \psi_{\frac{\Delta^*}{z_1}}) = \left| \sum_{\substack{d \\ d \equiv d_\star(q_0) \\ (d, m\lambda\widetilde{\lambda})=1}} \psi_{\frac{\Delta^*}{z_1}}(d) \sum_{n_1}^{\bigstar_1} \psi_{(\frac{\lambda^*}{\lambda})N}(n_1) e_m\!\left(\frac{(\frac{\lambda^*}{\lambda})(\frac{\lambda^*}{\widetilde{\lambda}})A((\frac{z_1}{w_1})k + (\lambda - \widetilde{\lambda})B)}{(n_1 + (\frac{\lambda^*}{\lambda})(F_k + B)d)(n_1 + (\frac{\lambda^*}{\widetilde{\lambda}})(G_k + B)d)}\right)\right|.$$

Here we consider those coefficient sequences $\psi_{(\frac{\lambda^*}{\lambda})N}(n_1)$ and $\psi_{\frac{\Delta^*}{z_1}}(d)$ which are shifted smooth at scales $(\frac{\lambda^*}{\lambda})N$ and $\frac{\Delta^*}{z_1}$, and for which the corresponding shifts x_0 are bounded by $x_0 \ll x$. By the discussion above, we find

$$\Upsilon_0(\psi^\circ_{\frac{\Delta^*}{z_1}}) \ll_\varepsilon \sup_{(\psi_{(\frac{\lambda^*}{\lambda})N}, \psi_{\frac{\Delta^*}{z_1}})} \Upsilon_1(\psi_{(\frac{\lambda^*}{\lambda})N}, \psi_{\frac{\Delta^*}{z_1}}) + O_\varepsilon(x^{-90}).$$

Hence we have the bound

$$\Sigma_6(z, w_6, k) \ll_\varepsilon \sup_{(\psi_{(\frac{\lambda^*}{\lambda})N}, \psi_{\frac{\Delta^*}{z_1}})} \left(\frac{\Delta_1}{\Delta^*}\right) (\Upsilon_1(\psi_{(\frac{\lambda^*}{\lambda})N}, \psi_{\frac{\Delta^*}{z_1}}) + O_\varepsilon(x^{-90})).$$

But $\Upsilon_1(\psi_{(\frac{\lambda^*}{\lambda})N}, \psi_{\frac{\Delta^*}{z_1}}) = \Sigma_7(z, w_6, w_7, k)$ for some $w_7 \in W_7(z, k)$. $\square$

3.3.4.3 Coprimality conditions

For a final simplification, we remove the various coprimality conditions by using Möbius inversion.

Lemma 3.13 (Removal of coprimality conditions)**.** *Let Z_5 and Σ_7 be as given in Lemma* 3.10 *and Lemma* 3.12*.*

For $z \in Z_5$, denote by $W_8(z)$ the set of tuples

$$(\Delta_2, N_2, \psi_{\Delta_2/z_1}, \psi_{N_2}, q_3, d^\star, n^\star, A_1, A_2, B_1)$$

which satisfy the following: $\Delta_2, N_2 \in (0, \infty)$ with $\Delta_2 \le \Delta^* = \min\{\frac{N}{\Lambda x^{5\varepsilon}}, \Delta_1\}$ and $N_2 \le N$. $q_3 \in \mathbb{N}$ with $q_3 \mid (\frac{m}{q_0})$. $\psi_{\Delta_2/z_1}(d)$ and $\psi_{N_2}(n)$ are shifted smooth coefficient sequences at scales Δ_2/z_1 and N_2. $d^\star$ and $n^\star$ are residue classes mod $(q_0 q_3)$ and A_1, A_2, B_1 are residue classes mod m with $(A_1, m) = (A_2, m) = 1$.

For given $z \in Z_5$, $w_8 \in W_8(z)$ and $k \in \mathbb{Z}$, set

$$\Sigma_8(z, w_8, k) = \left| \sum_{\substack{d \\ d \equiv d^\star (q_0 q_3)}} \sum_{\substack{n \\ n \equiv n^\star (q_0 q_3)}} \psi_{\Delta_2/z_1}(d) \psi_{N_2}(n) e_m\left(\frac{w_2 A_1 L_1}{(n + B_1 d)(n + (B_1 + L_1)d)}\right) \right|$$

where $L_1 \equiv A_2 \left\{ \frac{1}{w_2} \left(\left(\frac{z_1}{w_1} \right) k + (\lambda - \widetilde{\lambda}) B \right) \right\} \bmod m.$

Then for all $k \in \mathbb{Z}$ with $|k| \ll \frac{w_1 \Lambda N}{x^{5\varepsilon} \Delta_1}$, $w_2 \mid k$ and $((\frac{z_1}{w_1})k + (\lambda - \widetilde{\lambda})B, m) \le T(z)$, and for all $z \in Z_5$, $w_6 \in W_6(z)$ and $w_7 \in W_7(z, k)$, we have

$$\Sigma_7(z, w_6, w_7, k) \ll_\varepsilon \sup_{w_8 \in W_8(z)} \left(x^{2\varepsilon} q_3\right) \Sigma_8(z, w_8, k) + O_\varepsilon(x^{-89}).$$

In particular, to show $\Sigma_7(z, w_6, w_7, k) \ll_\varepsilon \xi_k \left(\frac{\Delta^*}{\Delta_1}\right) \left(\min\left\{\frac{(w_2, m)}{x^{\delta + 10\varepsilon} H^3}, \frac{(w_2, m)}{H^4}\right\} \frac{q_0 (q_0, \ell) N^2}{x^{27\varepsilon}}\right)$, it suffices to prove that

$$\Sigma_8(z, w_8, k) + O_\varepsilon(x^{-89}) \ll_\varepsilon \xi_k \left(\frac{1}{x^{2\varepsilon} q_3}\right) \left(\min\left\{\frac{(w_2, m)}{x^{\delta + 10\varepsilon} H^3}, \frac{(w_2, m)}{H^4}\right\} \frac{q_0 (q_0, \ell) N^2}{x^{27\varepsilon}}\right).$$

Proof. Our goal is to bound $\Sigma_7(z, w_6, w_7, k)$, which is given by

$$\Sigma_7 = \left| \sum_{\substack{d \\ d \equiv d_\star (q_0) \\ (d, m\lambda\widetilde{\lambda}) = 1}} \psi_{\frac{\Delta^*}{z_1}}(d) \sum_{\substack{n_1 \\ (\frac{\lambda}{\lambda^*}) n_1 + F_k d \equiv n_\star (q_0) \\ ((\frac{\lambda}{\lambda^*}) n_1 + F_k d, w_1 c_1) = 1 \\ ((\frac{\widetilde{\lambda}}{\lambda^*}) n_1 + G_k d, w_1 c_1) = 1 \\ ((\frac{\lambda}{\lambda^*}) n_1 + F_k d + ld, c_2) = 1 \\ ((\frac{\widetilde{\lambda}}{\lambda^*}) n_1 + G_k d + ld, c_2) = 1}} \psi_{(\frac{\lambda^*}{\lambda}) N}(n_1)\, e_m\left(\frac{(\frac{\lambda^*}{\lambda})(\frac{\lambda^*}{\widetilde{\lambda}}) A((\frac{z_1}{w_1})k + (\lambda - \widetilde{\lambda})B)}{(n_1 + (\frac{\lambda^*}{\lambda})(F_k + B)d)(n_1 + (\frac{\lambda^*}{\widetilde{\lambda}})(G_k + B)d)} \right) \right|.$$

Here we assume that $k \in \mathbb{Z}$ with $|k| \ll \frac{w_1 \Lambda N}{x^{5\varepsilon}\Delta_1}$, $w_2 \mid k$ and $((\frac{z_1}{w_1})k + (\lambda - \widetilde{\lambda})B, m) \le T(z)$.

Step 1: Removal of coprimality conditions (Part 1). We first remove coprimality conditions which do not involve factors of m. Recall that $(\lambda, m^*) = w_2$ and $(\widetilde{\lambda}, m^*) = w_2$. Hence we may replace $(d, m\lambda\widetilde{\lambda}) = 1$ by $(d, m(\lambda/w_2)(\widetilde{\lambda}/w_2)) = 1$, where both λ/w_2 and $\widetilde{\lambda}/w_2$ are coprime to m. Recall also that $(w_1, m) = 1$. For the moment, we wish to remove the requirements $(d, (\lambda/w_2)(\widetilde{\lambda}/w_2)) = 1$, $((\frac{\lambda}{\lambda^*})n_1 + F_k d, w_1) = 1$ and $((\frac{\widetilde{\lambda}}{\lambda^*})n_1 + G_k d, w_1) = 1$. Using Möbius inversion, we observe that Σ_7 equals

$$\left| \sum_{\substack{f_1, f_2, f_3 \\ f_1 | \frac{\lambda}{w_2}\frac{\widetilde{\lambda}}{w_2} \\ f_2 | w_1 \\ f_3 | w_1}} \mu(f_1)\mu(f_2)\mu(f_3) \sum_{\substack{d \\ d \equiv d_\star (q_0) \\ f_1 | d \\ (d,m)=1}} \psi_{\frac{\Delta^*}{z_1}}(d) \sum_{\substack{n_1 \\ n_1 + (\frac{\lambda^*}{\lambda})F_k d \equiv (\frac{\lambda^*}{\lambda})n_\star (q_0) \\ f_2 | (\frac{\lambda}{\lambda^*})n_1 + F_k d \\ f_3 | (\frac{\widetilde{\lambda}}{\lambda^*})n_1 + G_k d \\ ((\frac{\lambda}{\lambda^*})n_1 + F_k d, c_1) = 1 \\ ((\frac{\widetilde{\lambda}}{\lambda^*})n_1 + G_k d, c_1) = 1 \\ ((\frac{\lambda}{\lambda^*})n_1 + F_k d + ld, c_2) = 1 \\ ((\frac{\widetilde{\lambda}}{\lambda^*})n_1 + G_k d + ld, c_2) = 1}} \psi_{(\frac{\lambda^*}{\lambda})N}(n_1) e_m\left(\frac{(\frac{\lambda^*}{\lambda})(\frac{\lambda^*}{\widetilde{\lambda}})A((\frac{z_1}{w_1})k + (\lambda - \widetilde{\lambda})B)}{(n_1 + (\frac{\lambda^*}{\lambda})(F_k + B)d)(n_1 + (\frac{\lambda^*}{\widetilde{\lambda}})(G_k + B)d)} \right) \right|.$$

For $f_1, f_2, f_3 \in \mathbb{N}$ with $f_1 \mid \frac{\lambda}{w_2}\frac{\widetilde{\lambda}}{w_2}$, $f_2 \mid w_1$ and $f_3 \mid w_1$, we define $\Upsilon_1(f_1, f_2, f_3)$ to equal

$$\Upsilon_1 = \left| \sum_{\substack{d_1 \\ f_1 d_1 \equiv d_\star (q_0) \\ (d_1, m) = 1}} \psi_{\frac{\Delta^*}{z_1}}(f_1 d_1) \sum_{\substack{n_1 \\ n_1 + (\frac{\lambda^*}{\lambda})F_k f_1 d_1 \equiv (\frac{\lambda^*}{\lambda})n_\star (q_0) \\ f_2 | (\frac{\lambda}{\lambda^*})n_1 + F_k f_1 d_1 \\ f_3 | (\frac{\widetilde{\lambda}}{\lambda^*})n_1 + G_k f_1 d_1 \\ ((\frac{\lambda}{\lambda^*})n_1 + F_k f_1 d_1, c_1) = 1 \\ ((\frac{\widetilde{\lambda}}{\lambda^*})n_1 + G_k f_1 d_1, c_1) = 1 \\ ((\frac{\lambda}{\lambda^*})n_1 + F_k f_1 d_1 + l f_1 d_1, c_2) = 1 \\ ((\frac{\widetilde{\lambda}}{\lambda^*})n_1 + G_k f_1 d_1 + l f_1 d_1, c_2) = 1}} \psi_{(\frac{\lambda^*}{\lambda})N}(n_1) e_m\left(\frac{(\frac{\lambda^*}{\lambda})(\frac{\lambda^*}{\widetilde{\lambda}})A((\frac{z_1}{w_1})k + (\lambda - \widetilde{\lambda})B)}{(n_1 + (\frac{\lambda^*}{\lambda})(F_k + B) f_1 d_1)(n_1 + (\frac{\lambda^*}{\widetilde{\lambda}})(G_k + B) f_1 d_1)} \right) \right|.$$

Since $\frac{\lambda}{w_2}\frac{\widetilde{\lambda}}{w_2}$ and w_1 have only $O_\varepsilon(x^{\varepsilon/3})$ divisors, we then have

$$\Sigma_7(z, w_6, w_7, k) \ll_\varepsilon \sup_{f_1, f_2, f_3} (x^\varepsilon) \Upsilon_1(f_1, f_2, f_3), \tag{3.23}$$

where the supremum is taken over $f_1, f_2, f_3 \in \mathbb{N}$ with $f_1 \mid \frac{\lambda}{w_2}\frac{\widetilde{\lambda}}{w_2}$, $f_2 \mid w_1$ and $f_3 \mid w_1$.

Next we consider the conditions $(\frac{\lambda}{\lambda^*})n_1 \equiv -F_k f_1 d_1 \bmod (f_2)$ and $(\frac{\widetilde{\lambda}}{\lambda^*})n_1 \equiv -G_k f_1 d_1 \bmod (f_3)$. Write $f_2^* = (f_2, \lambda/\lambda^*)$ and $f_3^* = (f_3, \widetilde{\lambda}/\lambda^*)$. Write $f_2^\circ = f_2^*/(f_2^*, F_k f_1)$ and $f_3^\circ = f_3^*/(f_3^*, G_k f_1)$ and set $f_{23} = [f_2^\circ, f_3^\circ]$. For $(\frac{\lambda}{\lambda^*})n_1 \equiv -F_k f_1 d_1 \bmod (f_2)$ and $(\frac{\widetilde{\lambda}}{\lambda^*})n_1 \equiv -G_k f_1 d_1 \bmod (f_3)$ to simultaneously have solutions, we require that $f_{23} \mid d_1$. Write $d_1 = f_{23} d_2$. In that case, the conditions are satisfied if $n_1 \equiv -(\frac{\lambda}{\lambda^* f_2^*})^{-1}(\frac{F_k f_1 f_{23}}{f_2^*})d_2 \bmod (f_2/f_2^*)$ and $n_1 \equiv -(\frac{\widetilde{\lambda}}{\lambda^* f_3^*})^{-1}(\frac{G_k f_1 f_{23}}{f_3^*})d_2 \bmod (f_3/f_3^*)$. Now recall that w_1 is squarefree. Hence f_2/f_2^* and f_3/f_3^* are also squarefree. Suppose $p \mid (f_2/f_2^*, f_3/f_3^*)$. If $-(\frac{\lambda}{\lambda^* f_2^*})^{-1}(\frac{F_k f_1 f_{23}}{f_2^*}) \not\equiv -(\frac{\widetilde{\lambda}}{\lambda^* f_3^*})^{-1}(\frac{G_k f_1 f_{23}}{f_3^*}) \bmod p$, then $-(\frac{\lambda}{\lambda^* f_2^*})^{-1}(\frac{F_k f_1 f_{23}}{f_2^*})d_2 \equiv -(\frac{\widetilde{\lambda}}{\lambda^* f_3^*})^{-1}(\frac{G_k f_1 f_{23}}{f_3^*})d_2 \bmod p$ only holds if $p \mid d_2$. Denote the product of $p \mid (f_2/f_2^*, f_3/f_3^*)$ with $-(\frac{\lambda}{\lambda^* f_2^*})^{-1}(\frac{F_k f_1 f_{23}}{f_2^*}) \not\equiv -(\frac{\widetilde{\lambda}}{\lambda^* f_3^*})^{-1}(\frac{G_k f_1 f_{23}}{f_3^*}) \bmod p$ by f_4. Choose $H_k \in \{1, \ldots, \frac{1}{f_4}[f_2/f_2^*, f_3/f_3^*]\}$ with $H_k \equiv -(\frac{\lambda}{\lambda^* f_2^*})^{-1}(\frac{F_k f_1 f_{23}}{f_2^*}) \bmod p$ for every $p \mid f_2/(f_2^* f_4)$ and $H_k \equiv -(\frac{\widetilde{\lambda}}{\lambda^* f_3^*})^{-1}(\frac{G_k f_1 f_{23}}{f_3^*}) \bmod p$ for every $p \mid f_3/(f_3^* f_4)$. Overall, we then find that the conditions $(\frac{\lambda}{\lambda^*})n_1 \equiv -F_k f_1 d_1 \bmod (f_2)$ and $(\frac{\widetilde{\lambda}}{\lambda^*})n_1 \equiv -G_k f_1 d_1 \bmod (f_3)$ are satisfied if and only if $d_1 = f_{23} f_4 d_3$ for some $d_3 \in \mathbb{Z}$ and $n_1 = [f_2/f_2^*, f_3/f_3^*]n_3 + H_k f_4 d_3$ for some $n_3 \in \mathbb{Z}$. To simplify notation, we now set $f^* = f_1 f_{23} f_4$, where $f^* \mid \frac{\lambda}{w_2}\frac{\widetilde{\lambda}}{w_2} w_1$ and $(f^*, m) = 1$, and we set $w_3 = [f_2/f_2^*, f_3/f_3^*]$, where $w_3 \mid w_1$ and $(w_3, m) = 1$. Then $\Upsilon_1(f_1, f_2, f_3)$ transforms as follows:

$$\Upsilon_1(f_1,f_2,f_3) = \left| \sum_{\substack{d_3 \\ d_3 \equiv (f^*)^{-1} d_\star (q_0) \\ (d_3,m)=1}} \psi_{\frac{\Delta^*}{z_1}}(f^* d_3) \sum_{\substack{n_3 \\ n_3 \equiv -\frac{1}{w_3}(H_k f_4 + (\frac{\lambda^*}{\lambda}) F_k f^*)(f^*)^{-1} d_\star + \frac{1}{w_3}(\frac{\lambda^*}{\lambda}) n_\star (q_0) \\ ((\frac{\lambda}{\lambda^*}) w_3 n_3 + ((\frac{\lambda}{\lambda^*}) H_k f_4 + F_k f^*) d_3, c_1) = 1 \\ ((\frac{\widetilde{\lambda}}{\lambda^*}) w_3 n_3 + ((\frac{\widetilde{\lambda}}{\lambda^*}) H_k f_4 + G_k f^*) d_3, c_1) = 1 \\ ((\frac{\lambda}{\lambda^*}) w_3 n_3 + ((\frac{\lambda}{\lambda^*}) H_k f_4 + F_k f^* + l f^*) d_3, c_2) = 1 \\ ((\frac{\widetilde{\lambda}}{\lambda^*}) w_3 n_3 + ((\frac{\widetilde{\lambda}}{\lambda^*}) H_k f_4 + G_k f^* + l f^*) d_3, c_2) = 1}} \psi_{(\frac{\lambda^*}{\lambda})N}(w_3 n_3 + H_k f_4 d_3) \ (\star) \right|$$

where $(\star) = e_m\left(\frac{(\frac{1}{w_3})^2 (\frac{\lambda^*}{\lambda})(\frac{\lambda^*}{\widetilde{\lambda}}) A((\frac{z_1}{w_1})k + (\lambda - \widetilde{\lambda})B)}{(n_3 + \frac{1}{w_3}(H_k f_4 + (\frac{\lambda^*}{\lambda})(F_k + B) f^*) d_3)(n_3 + \frac{1}{w_3}(H_k f_4 + (\frac{\lambda^*}{\widetilde{\lambda}})(G_k + B) f^*) d_3)} \right)$.

Step 2: Taylor expansions. Next we use Taylor expansions once more, this time to treat $\psi_{(\frac{\lambda^*}{\lambda})N}(w_3 n_3 + H_k f_4 d_3)$. Since $\psi_{(\lambda^*/\lambda)N}(n)$ is a shifted smooth coefficient sequence, there exist a constant $x_0 \in \mathbb{R}$ and a smooth function $\phi : \mathbb{R} \to \mathbb{C}$, supported on $[c, C]$, with $|\phi^{(j)}(t)| \ll_j (\log x)^{O(1)}$ and $\psi_{(\lambda^*/\lambda)N}(n) = \phi(\frac{n - x_0}{(\lambda^*/\lambda)N})$. Since $\psi_{\frac{\Delta^*}{z_1}}(d)$ is a shifted smooth coefficient sequence at scale Δ^*/z_1, there also exist a constant $y_0 \in \mathbb{R}$ and a smooth function $\varphi : \mathbb{R} \to \mathbb{C}$, supported on $[c, C]$, with $|\varphi^{(j)}(t)| \ll_j (\log x)^{O(1)}$ and $\psi_{\frac{\Delta^*}{z_1}}(d) = \varphi(\frac{d - y_0}{\Delta^*/z_1})$. (Looking at the definition of $W_7(z, k)$, we also note the additional assumptions $x_0 \ll x$ and $y_0 \ll x$.) Recall that $\Delta^* \ll N/(\Lambda x^{5\varepsilon})$ and observe that

$$\left| \frac{H_k f_4 (\Delta^*/(z_1 f^*))}{(\lambda^*/\lambda)N} \frac{(d_3 - \frac{y_0}{f^*})}{\Delta^*/(z_1 f^*)} \right| \ll \left| \frac{[f_2/f_2^*, f_3/f_3^*]\Delta^*\Lambda}{z_1 N} \right| \ll \left| \frac{w_1 \Delta^* \Lambda}{z_1 N} \right| \ll \frac{\Delta^* \Lambda}{N} \ll \frac{1}{x^{5\varepsilon}}$$

whenever $|f^* d_3 - y_0| \ll \Delta^*/z_1$. Choosing $J = \lceil \frac{20}{\varepsilon} \rceil$, we may write

$$\psi_{(\frac{\lambda^*}{\lambda})N}(w_3 n_3 + H_k f_4 d_3) = \sum_{j=0}^{J} \frac{1}{j!} \left(\frac{H_k f_4 (d_3 - \frac{y_0}{f^*})}{(\lambda^*/\lambda)N} \right)^j \phi^{(j)}\left(\frac{w_3 n_3 - x_0 + (\frac{H_k f_4}{f^*}) y_0}{(\lambda^*/\lambda)N} \right) \quad (3.24)$$

$$+ O_\varepsilon(x^{-100} 1_{n_3 \ll x}).$$

The sequence $\phi^{(j)}((n_3 - \frac{x_0}{w_3} + (\frac{H_k f_4}{f^* w_3})y_0)/(\frac{\lambda^* N}{\lambda w_3}))$ is a shifted smooth coefficient sequence in n_3 at scale $(\frac{\lambda^*}{\lambda w_3})N$. The sequence $\frac{1}{j!}\left(\frac{H_k f_4(d_3 - \frac{y_0}{f^*})}{(\lambda^*/\lambda)N}\right)^j \varphi\left(\frac{(d_3 - \frac{y_0}{f^*})}{\Delta^*/(z_1 f^*)}\right)$ is a shifted smooth coefficient sequence in d_3 at scale $\frac{\Delta^*}{z_1 f^*}$. We now substitute the RHS of (3.24) for $\psi_{(\frac{\lambda^*}{\lambda})N}(w_3 n_3 + H_k f_4 d_3)$ in $\Upsilon_1(f_1, f_2, f_3)$. Set $\Upsilon_2(f_1, f_2, f_3, \psi_{(\frac{\lambda^*}{\lambda w_3})N}, \psi_{\frac{\Delta^*}{z_1 f^*}})$ equal to

$$\Upsilon_2 = \left| \sum_{\substack{d_3 \\ d_3 \equiv (f^*)^{-1} d_\star \bmod (q_0) \\ (d_3, m) = 1}} \psi_{\frac{\Delta^*}{z_1 f^*}}(d_3) \sum_{\substack{n_3 \\ n_3 \equiv -\frac{1}{w_3}(H_k f_4 + (\frac{\lambda^*}{\lambda})F_k f^*)(f^*)^{-1} d_\star + \frac{1}{w_3}(\frac{\lambda^*}{\lambda}) n_\star \bmod (q_0) \\ ((\frac{\lambda}{\lambda^*})w_3 n_3 + ((\frac{\lambda}{\lambda^*})H_k f_4 + F_k f^*)d_3, c_1) = 1 \\ ((\frac{\widetilde{\lambda}}{\lambda^*})w_3 n_3 + ((\frac{\widetilde{\lambda}}{\lambda^*})H_k f_4 + G_k f^*)d_3, c_1) = 1 \\ ((\frac{\lambda}{\lambda^*})w_3 n_3 + ((\frac{\lambda}{\lambda^*})H_k f_4 + F_k f^* + l f^*)d_3, c_2) = 1 \\ ((\frac{\widetilde{\lambda}}{\lambda^*})w_3 n_3 + ((\frac{\widetilde{\lambda}}{\lambda^*})H_k f_4 + G_k f^* + l f^*)d_3, c_2) = 1}} \psi_{(\frac{\lambda^*}{\lambda w_3})N}(n_3) \;\; (\star) \right|$$

$$\text{where } (\star) = e_m\left(\frac{(\frac{1}{w_3})^2(\frac{\lambda^*}{\lambda})(\frac{\lambda^*}{\widetilde{\lambda}})A((\frac{z_1}{w_1})k + (\lambda - \widetilde{\lambda})B)}{(n_3 + \frac{1}{w_3}(H_k f_4 + (\frac{\lambda^*}{\lambda})(F_k + B)f^*)d_3)(n_3 + \frac{1}{w_3}(H_k f_4 + (\frac{\lambda^*}{\widetilde{\lambda}})(G_k + B)f^*)d_3)}\right).$$

Taking the supremum over all coefficient sequences $\psi_{(\frac{\lambda^*}{\lambda w_3})N}(n_3)$ and $\psi_{\frac{\Delta^*}{z_1 f^*}}(d_3)$ which are shifted smooth at scales $(\frac{\lambda^*}{\lambda w_3})N$ and $\frac{\Delta^*}{z_1 f^*}$, we then have

$$\Upsilon_1(f_1, f_2, f_3) \ll_\varepsilon \sup_{(\psi_{(\frac{\lambda^*}{\lambda w_3})N}, \psi_{\frac{\Delta^*}{z_1 f^*}})} \Upsilon_2(f_1, f_2, f_3, \psi_{(\frac{\lambda^*}{\lambda w_3})N}, \psi_{\frac{\Delta^*}{z_1 f^*}}) + O_\varepsilon(x^{-90}). \tag{3.25}$$

Step 3: Removal of coprimality conditions (Part 2). Finally, we now remove the remaining coprimality conditions. Again we use Möbius inversion. For $e_0 \mid m$, $e_1, e_2 \mid c_1$ and $e_3, e_4 \mid c_2$, we define $\Upsilon_3(f_1, f_2, f_3, \psi_{(\frac{\lambda^*}{\lambda w_3})N}, \psi_{\frac{\Delta^*}{z_1 f^*}}, e_0, e_1, e_2, e_3, e_4)$ by

$$\Upsilon_3 = \left| \sum_{\substack{d_3 \\ d_3 \equiv (f^*)^{-1} d_\star \bmod (q_0) \\ d_3 \equiv 0 \bmod (e_0)}} \psi_{\frac{\Delta^*}{z_1 f^*}}(d_3) \sum_{\substack{n_3 \\ n_3 \equiv -\frac{1}{w_3}(H_k f_4 + (\frac{\lambda^*}{\lambda})F_k f^*)(f^*)^{-1} d_\star + \frac{1}{w_3}(\frac{\lambda^*}{\lambda}) n_\star \bmod (q_0) \\ (\frac{\lambda}{\lambda^*})w_3 n_3 + ((\frac{\lambda}{\lambda^*})H_k f_4 + F_k f^*)d_3 \equiv 0 \bmod (e_1) \\ (\frac{\widetilde{\lambda}}{\lambda^*})w_3 n_3 + ((\frac{\widetilde{\lambda}}{\lambda^*})H_k f_4 + G_k f^*)d_3 \equiv 0 \bmod (e_2) \\ (\frac{\lambda}{\lambda^*})w_3 n_3 + ((\frac{\lambda}{\lambda^*})H_k f_4 + F_k f^* + l f^*)d_3 \equiv 0 \bmod (e_3) \\ (\frac{\widetilde{\lambda}}{\lambda^*})w_3 n_3 + ((\frac{\widetilde{\lambda}}{\lambda^*})H_k f_4 + G_k f^* + l f^*)d_3 \equiv 0 \bmod (e_4)}} \psi_{(\frac{\lambda^*}{\lambda w_3})N}(n_3) \;\; (\star) \right|$$

$$\text{where } (\star) = e_m\left(\frac{(\frac{1}{w_3})^2(\frac{\lambda^*}{\lambda})(\frac{\lambda^*}{\widetilde{\lambda}})A((\frac{z_1}{w_1})k + (\lambda - \widetilde{\lambda})B)}{(n_3 + \frac{1}{w_3}(H_k f_4 + (\frac{\lambda^*}{\lambda})(F_k + B)f^*)d_3)(n_3 + \frac{1}{w_3}(H_k f_4 + (\frac{\lambda^*}{\widetilde{\lambda}})(G_k + B)f^*)d_3)}\right).$$

Using the same argument as in Step 1, we then have

$$\Upsilon_2(f_1, f_2, f_3, \psi_{(\frac{\lambda^*}{\lambda w_3})N}, \psi_{\frac{\Delta^*}{z_1 f^*}}) \ll_\varepsilon \sup_{e_0, e_1, e_2, e_3, e_4} (x^\varepsilon) \Upsilon_3(f_1, f_2, f_3, \psi_{(\frac{\lambda^*}{\lambda w_3})N}, \psi_{\frac{\Delta^*}{z_1 f^*}}, e_0, e_1, e_2, e_3, e_4), \tag{3.26}$$

where the supremum is taken over e_0, e_1, e_2, e_3, e_4 with $e_0 \mid m$, $e_1, e_2 \mid c_1$ and $e_3, e_4 \mid c_2$.

Now recall that $c_1 \mid m$ and $c_2 \mid m$ and $(\lambda, m^*) = (\widetilde{\lambda}, m^*) = w_2$. Since $(\lambda, \widetilde{\lambda}) = \lambda^*$, then $w_2 \mid \lambda^*$ and $((\frac{\lambda}{\lambda^*})w_3, mc_1c_2) = ((\frac{\widetilde{\lambda}}{\lambda^*})w_3, mc_1c_2) = 1$. So the summation conditions of Υ_3 can be rewritten as follows:

$$d_3 \equiv \frac{d_\star}{f^*} \bmod (q_0) \quad \text{and} \quad d_3 \equiv 0 \bmod (e_0), \tag{3.27}$$

$$n_3 \equiv -\frac{1}{w_3}\left(H_k f_4 + \left(\frac{\lambda^*}{\lambda}\right) F_k f^*\right)\frac{d_\star}{f^*} + \frac{1}{w_3}\left(\frac{\lambda^*}{\lambda}\right) n_\star \bmod (q_0), \tag{3.28}$$

$$n_3 \equiv -\frac{1}{w_3}\left(H_k f_4 + \left(\frac{\lambda^*}{\lambda}\right) F_k f^*\right) d_3 \bmod (e_1), \tag{3.29}$$

$$n_3 \equiv -\frac{1}{w_3}\left(H_k f_4 + \left(\frac{\lambda^*}{\widetilde{\lambda}}\right) G_k f^*\right) d_3 \bmod (e_2), \tag{3.30}$$

$$n_3 \equiv -\frac{1}{w_3}\left(H_k f_4 + \left(\frac{\lambda^*}{\lambda}\right) (F_k f^* + l f^*)\right) d_3 \bmod (e_3), \tag{3.31}$$

$$n_3 \equiv -\frac{1}{w_3}\left(H_k f_4 + \left(\frac{\lambda^*}{\widetilde{\lambda}}\right) (G_k f^* + l f^*)\right) d_3 \bmod (e_4). \tag{3.32}$$

Now denote by g_1 the least common multiple of the six integers q_0, e_0, e_1, e_2, e_3 and e_4. Denote by g_2 the least common multiple of q_0 and e_0. Denote by g_3 the least common multiple of q_0, e_1, e_2, e_3 and e_4. Note that $g_1 = [g_2, g_3]$ and $q_0 \mid (g_2, g_3)$.

Step 4: A final partition. We now partition Υ_3 according to two new conditions:

$$d_3 \equiv d^\circ \bmod \left(\frac{g_1}{g_2}\right) \quad \text{and} \quad n_3 \equiv n^\circ \bmod \left(\frac{g_1}{g_3}\right). \tag{3.33}$$

The conditions (3.27), (3.28), (3.29), (3.30), (3.31) and (3.32) may be incompatible. However, if they are compatible, then (3.27) and the LHS of (3.33) determine the congruence class of d_3 mod g_1. Replacing d_3 by this congruence class in (3.28), (3.29), (3.30), (3.31) and (3.32), these expressions, together with the RHS of (3.33), in turn determine the congruence class of n_3 mod g_1. We write $d_3 \equiv d_4$ mod g_1 and $n_3 \equiv n_4$ mod g_1. The only choices we made were the value of d° and the value of n°, so there are at most $(\frac{g_1}{g_2})(\frac{g_1}{g_3})$ choices of d_4 and n_4 mod g_1 which contribute to Υ_3. Set

$$\Upsilon_4(\dots, d_4, n_4) = \left| \sum_{\substack{d_3 \\ d_3 \equiv d_4 (g_1)}} \sum_{\substack{n_3 \\ n_3 \equiv n_4 (g_1)}} \psi_{\frac{\Delta^*}{z_1 f^*}}(d_3)\psi_{(\frac{\lambda^*}{\lambda w_3})N}(n_3) \ (\star) \right|$$

$$\text{where } (\star) = e_m\left(\frac{(\frac{1}{w_3})^2(\frac{\lambda^*}{\lambda})(\frac{\lambda^*}{\widetilde{\lambda}})A((\frac{z_1}{w_1})k+(\lambda-\widetilde{\lambda})B)}{(n_3+\frac{1}{w_3}(H_k f_4+(\frac{\lambda^*}{\lambda})(F_k+B)f^*)d_3)(n_3+\frac{1}{w_3}(H_k f_4+(\frac{\lambda^*}{\widetilde{\lambda}})(G_k+B)f^*)d_3)}\right).$$

For $\underline{w} = (f_1, f_2, f_3, \psi_{(\frac{\lambda^*}{\lambda w_3})N}, \psi_{\frac{\Delta^*}{z_1 f^*}}, e_0, e_1, e_2, e_3, e_4)$, we then have

$$\Upsilon_3(\underline{w}) \ll \left(\frac{g_1}{g_2}\right)\left(\frac{g_1}{g_3}\right) \sup_{(d_4,n_4)} \Upsilon_4(\underline{w}, d_4, n_4) \ll \left(\frac{g_1}{q_0}\right) \sup_{(d_4,n_4)} \Upsilon_4(\underline{w}, d_4, n_4), \tag{3.34}$$

where the supremum is taken over all residue classes d_4 and n_4 mod g_1. (In the second upper bound we also used that $[g_2, g_3/q_0] = g_1$ and $g_2g_3/q_0 \geq g_1$.)

To conclude the proof, we now simplify the expression for Υ_4 a bit further. Set

$$\begin{aligned} B_1 &= \frac{1}{w_3}\left(H_k f_4 + \left(\frac{\lambda^*}{\lambda}\right)(F_k + B)f^*\right), \\ L_1 &= \frac{1}{w_3}\left(\left(\frac{\lambda^*}{\widetilde{\lambda}}\right)(G_k + B)f^* - \left(\frac{\lambda^*}{\lambda}\right)(F_k + B)f^*\right). \end{aligned}$$

We recall (3.20) to observe that

$$L_1 \equiv \left\{\left(\frac{f^*}{w_3}\right)\left(\frac{\lambda^*}{\lambda}\right)\left(\frac{w_2}{\widetilde{\lambda}}\right)\right\}\left\{\frac{1}{w_2}\left(\left(\frac{z_1}{w_1}\right)k + (\lambda - \widetilde{\lambda})B\right)\right\} \bmod m.$$

The residue class in the first pair of {}-brackets is primitive. We choose a residue class A_1 mod m such that $A_1 \equiv (\frac{w_3}{f^*})(\frac{\widetilde{\lambda}}{w_2})(\frac{1}{w_3})^2(\frac{\lambda^*}{\widetilde{\lambda}})A$ mod m. Here $(A_1, m) = 1$. The enumerator of the fraction inside $e_m()$ in Υ_4 now reads $w_2A_1L_1$ and we have $((\frac{z_1}{w_1})k + (\lambda - \widetilde{\lambda})B, m) = (w_2A_1L_1, m)$.

Furthermore, we relabel our coefficient sequences, writing $\psi_{\Delta_2/z_1}(d_3) = \psi_{\frac{\Delta^*}{z_1 f^*}}(d_3)$ and $\psi_{N_2}(n_3) = \psi_{(\frac{\lambda^*}{\lambda w_3})N}(n_3)$. Here $\Delta_2 \leq \Delta^*$ and $N_2 \leq N$. We recall that g_1 is divisible by q_0 and write $g_1 = q_0q_3$. Then

$$\Upsilon_4(\ldots, d_4, n_4) = \left| \sum_{\substack{d_3 \\ d_3 \equiv d_4 (q_0q_3)}} \sum_{\substack{n_3 \\ n_3 \equiv n_4 (q_0q_3)}} \psi_{\Delta_2/z_1}(d_3)\psi_{N_2}(n_3) e_m\left(\tfrac{w_2A_1L_1}{(n_3 + B_1d_3)(n_3 + (B_1 + L_1)d_3)}\right)\right|.$$

We now see that $\Upsilon_4(f_1, f_2, f_3, \psi_{(\frac{\lambda^*}{\lambda w_3})N}, \psi_{\frac{\Delta^*}{z_1 f^*}}, e_0, e_1, e_2, e_3, e_4, d_4, n_4) = \Sigma_8(z, w_8, k)$ for some $w_8 \in W_8(z)$. Putting (3.23), (3.25), (3.26) and (3.34) together, we obtain the proposed bound

$$\Sigma_7(z, w_6, w_7, k) \ll_\varepsilon \sup_{w_8 \in W_8(z)} \left(x^{2\varepsilon}q_3\right) \Sigma_8(z, w_8, k) + O_\varepsilon(x^{-89}). \qquad \square$$

3.3.5 Exponential sums

We have finally arrived at an expression to which the exponential sum bounds of Section 8 of Polymath [37] can be applied. In particular, we use bound (8.21) of Proposition 8.4 of Polymath [37], a version of which is stated below:

Lemma 3.14 (Deligne's bounds [37]). *Let $\delta > 0$ and $\varepsilon \in (0, \frac{\delta}{10^{100}})$. Let $m \mid P(x^\delta)$ with $m \ll x^{O(1)}$ and let Δ and N be positive real numbers with $\Delta, N \ll x^{O(1)}$. Let A, B, L, γ_1 and γ_2 be residue classes mod m. Let ψ_Δ and ψ_N be shifted smooth coefficient sequences at scales Δ and N. Then*

$$\begin{aligned}&\sum_d \sum_n \psi_\Delta(d)\psi_N(n) e_m\left(\frac{AL}{(n+Bd+\gamma_1)(n+(B+L)d+\gamma_2)}\right)\\ &\ll_\varepsilon x^\varepsilon (AL, m)\left(\frac{N}{m^{1/2}} + m^{1/2}\right)\left(\frac{\Delta}{m^{1/2}} + m^{1/2}\right).\end{aligned}$$

This lemma is not quite applicable to $\Sigma_8(z, w_8, k)$ yet, since $\Sigma_8(z, w_8, k)$ also involves the congruence conditions $d \equiv d^\star \bmod (q_0q_3)$ and $n \equiv n^\star \bmod (q_0q_3)$. Proposition 8.4 of [37] actually allows for the presence of such congruence conditions, but its stated bounds are not sufficiently strong for our purposes, and it potentially also contains some small mistakes: some factors q_0 appear to be in the wrong place. Hence we give a different version of this proposition below.

Lemma 3.15 (Deligne's bounds with moduli). *Let $\delta > 0$ and $\varepsilon \in (0, \frac{\delta}{10^{100}})$. Let $m \mid P(x^\delta)$ with $m \ll x^{O(1)}$. Let $0 < \Delta, N \ll x^{O(1)}$. Let A, B and L be residue classes mod m. Let ψ_Δ and ψ_N be shifted smooth coefficient sequences at scales Δ and N. Let $q \mid m$ and let $d^\star$ and $n^\star$ be residue classes mod q. Then*

$$\sum_{\substack{d\\ d\equiv d^\star(q)}} \sum_{\substack{n\\ n\equiv n^\star(q)}} \psi_\Delta(d)\psi_N(n) e_m\Big(\tfrac{AL}{(n+Bd)(n+(B+L)d)}\Big) \ll_\varepsilon \tfrac{x^\varepsilon(AL,m)}{q}\big(\tfrac{N}{m^{1/2}} + m^{1/2}\big)\big(\tfrac{\Delta}{m^{1/2}} + m^{1/2}\big).$$

Proof. To remove the congruence conditions, we write $d = d_1q + d^\star$ and $n = n_1q + n^\star$.

$$\begin{aligned}\Upsilon &= \sum_{\substack{d\\ d\equiv d^\star(q)}} \sum_{\substack{n\\ n\equiv n^\star(q)}} \psi_\Delta(d)\psi_N(n) e_m\Big(\tfrac{AL}{(n+Bd)(n+(B+L)d)}\Big)\\ &= \sum_{d_1}\sum_{n_1} \psi_\Delta(qd_1 + d^\star)\psi_N(qn_1 + n^\star) e_m\Big(\tfrac{AL}{(qn_1+n^\star+B(qd_1+d^\star))(qn_1+n^\star+(B+L)(qd_1+d^\star))}\Big).\end{aligned}$$

Write $\psi_{\Delta/q}(d_1) = \psi_\Delta(qd_1 + d^\star)$ and $\psi_{N/q}(n_1) = \psi_N(qn_1 + n^\star)$. Observe that $\psi_{\Delta/q}(d_1)$ is a shifted smooth coefficient sequence at scale Δ/q and $\psi_{N/q}(n_1)$ is a shifted smooth coefficient sequence at scale N/q. Set $\gamma_1 = n^\star + Bd^\star$ and $\gamma_2 = n^\star + (B+L)d^\star$. Since m is squarefree, we may write $m = qm_1$, where $(q, m_1) = 1$. In particular, q has an inverse mod m_1. Then note that

$$\begin{aligned}&e_m\Big(\tfrac{AL}{(qn_1+n^\star+B(qd_1+d^\star))(qn_1+n^\star+(B+L)(qd_1+d^\star))}\Big) = e_{qm_1}\Big(\tfrac{AL}{(qn_1+Bqd_1+\gamma_1)(qn_1+(B+L)qd_1+\gamma_2)}\Big)\\ &= e_q\Big(\tfrac{AL}{m_1(qn_1+Bqd_1+\gamma_1)(qn_1+(B+L)qd_1+\gamma_2)}\Big)\, e_{m_1}\Big(\tfrac{AL}{q(qn_1+Bqd_1+\gamma_1)(qn_1+(B+L)qd_1+\gamma_2)}\Big)\end{aligned}$$

$$= e_q\left(\tfrac{AL}{m_1\gamma_1\gamma_2}\right) e_{m/q}\left(\tfrac{(AL/q^3)}{(n_1+Bd_1+\gamma_1/q)(n_1+(B+L)d_1+\gamma_2/q)}\right).$$

Here $e_q(\frac{AL}{(m_1\gamma_1\gamma_2)})$ is a constant which is either equal to zero or has modulus 1. So

$$|\Upsilon| \le \left|\sum_{d_1}\sum_{n_1}\psi_{\Delta/q}(d_1)\psi_{N/q}(n_1)e_{m/q}\left(\frac{(AL/q^3)}{(n_1+Bd_1+\gamma_1/q)(n_1+(B+L)d_1+\gamma_2/q)}\right)\right|.$$

Now Lemma 3.14 is applicable to the expression above. We apply the lemma with Δ/q in the place of Δ, N/q in the place of N, m/q in the place of m and $A_1 \equiv AL/q^3$ mod m/q in the place of AL mod m. We also note that $(A_1, m/q) = (AL, m/q) \le (AL, m)$ since $(q, m/q) = 1$. The result follows. $\square$

3.3.5.1 Upper bounds on Σ_6

We combine Lemma 3.12, 3.13 and 3.15 to obtain upper bounds on Σ_6.

Lemma 3.16 (Bounding Σ_6)**.** *Let Z_5, $W_6(z)$ and Σ_6 be as given in Lemma* 3.10 *and Lemma* 3.11. *Then for all $z \in Z_5$, $w_6 \in W_6(z)$ and $k \in \mathbb{Z}$ which satisfy $|k| \ll \frac{w_1\Lambda N}{x^{5\varepsilon}\Delta_1}$, $w_2 \mid k$ and $((\frac{z_1}{w_1})k + (\lambda - \widetilde{\lambda})B, m) \le T(z)$, we have*

$$\Sigma_6(z, w_6, k) \ll_\varepsilon ((\tfrac{z_1}{w_1})k + (\lambda - \widetilde{\lambda})B, m)\left(\tfrac{x^{3\varepsilon}}{q_0}\right)\left(\tfrac{\Delta_1}{\Delta^*}\right)\left(\tfrac{N}{m^{1/2}} + m^{1/2}\right)\left(\tfrac{\Delta^*}{m^{1/2}} + m^{1/2}\right).$$

Proof. Consider $z \in Z_5$, $w_6 \in W_6(z)$ and $k \in \mathbb{Z}$ which satisfy $|k| \ll \frac{w_1\Lambda N}{x^{5\varepsilon}\Delta_1}$, $w_2 \mid k$ and $((\frac{z_1}{w_1})k + (\lambda - \widetilde{\lambda})B, m) \le T(z)$. We first recall from Lemma 3.12 that

$$\Sigma_6(z, w_6, k) \ll_\varepsilon \sup_{w_7 \in W_7(z,k)} \left(\tfrac{\Delta_1}{\Delta^*}\right)(\Sigma_7(z, w_6, w_7, k) + O_\varepsilon(x^{-90})).$$

Lemma 3.13 then gives

$$\Sigma_7(z, w_6, w_7, k) \ll_\varepsilon \sup_{w_8 \in W_8(z)} \left(x^{2\varepsilon}q_3\right)\Sigma_8(z, w_8, k) + O_\varepsilon(x^{-89}).$$

In combination, we have the bound

$$\Sigma_6(z, w_6, k) \ll_\varepsilon \sup_{w_8 \in W_8(z)} \left(\tfrac{\Delta_1}{\Delta^*}\right)\left(x^{2\varepsilon}q_3\right)\Sigma_8(z, w_8, k) + O_\varepsilon(x^{-50}). \tag{3.35}$$

Now we recall that for a given $w_8 \in W_8(z)$,

$$\Sigma_8(z, w_8, k) = \left|\sum_{\substack{d \\ d \equiv d^\star (q_0q_3)}}\sum_{\substack{n \\ n \equiv n^\star (q_0q_3)}}\psi_{\Delta_2/z_1}(d)\psi_{N_2}(n)e_m\left(\frac{w_2A_1L_1}{(n+B_1d)(n+(B_1+L_1)d)}\right)\right|,$$

where $L_1 \equiv A_2\big\{\frac{1}{w_2}\big(\big(\frac{z_1}{w_1}\big)k + (\lambda - \widetilde{\lambda})B\big)\big\} \bmod m$. By Lemma 3.15, then

$$\Sigma_8(z, w_8, k) \ll_\varepsilon \frac{x^\varepsilon(w_2A_1L_1, m)}{q_0q_3}\left(\tfrac{N_2}{m^{1/2}} + m^{1/2}\right)\Big(\tfrac{\Delta_2}{z_1m^{1/2}} + m^{1/2}\Big).$$

Recall that $w_8 \in W_8(z)$ implies $(A_1A_2, m) = 1$, $\Delta_2 \le \Delta^*$ and $N_2 \le N$. Thus

$$\Sigma_8(z, w_8, k) \ll_\varepsilon \frac{x^\varepsilon\big(\big(\frac{z_1}{w_1}\big)k+(\lambda-\widetilde{\lambda})B, m\big)}{q_0q_3}\left(\tfrac{N}{m^{1/2}} + m^{1/2}\right)\left(\tfrac{\Delta^*}{m^{1/2}} + m^{1/2}\right). \tag{3.36}$$

Substituting (3.36) into (3.35), we then get

$$\Sigma_6(z, w_6, k) \ll_\delta \frac{x^{3\varepsilon}\big(\big(\frac{z_1}{w_1}\big)k+(\lambda-\widetilde{\lambda})B, m\big)}{q_0}\left(\tfrac{\Delta_1}{\Delta^*}\right)\left(\tfrac{N}{m^{1/2}} + m^{1/2}\right)\left(\tfrac{\Delta^*}{m^{1/2}} + m^{1/2}\right).$$

This is what we wanted to show. □

3.3.5.2 Sums of greatest common divisors

Next we will bound Σ_5. Our bounds on Σ_6 involve gcds and so we need to bound sums of common divisors. We use the lemma below.

Lemma 3.17 (gcd sums)**.** *Let $\varepsilon > 0$, $m \in \mathbb{N}$ and $K, T > 0$ with $T \ge K$. Let $A, B \in \mathbb{Z}$ with $A \ne 0$. Then the following bound holds:*

$$\sum_{\substack{|k|\le K\\(Ak+B,m)\le T}} (Ak+B, m) \ll_\varepsilon m^\varepsilon (A, m)T.$$

Proof. We have

$$\sum_{\substack{|k|\le K\\(Ak+B,m)\le T}} (Ak+B, m) \le \sum_{d|m} \sum_{\substack{|k|\le K\\(Ak+B,m)\le T\\d|Ak+B}} d \le \sum_{\substack{d|m\\d\le T}} \sum_{\substack{|k|\le K\\d|Ak+B}} d.$$

Now write $w(d) = (A, d)$. We have $Ak + B \equiv 0 \bmod d$ if and only if ($w(d) \mid B$ and $\frac{A}{w(d)}k + \frac{B}{w(d)} = \ell\frac{d}{w(d)}$ for some $\ell \in \mathbb{Z}$) if and only if ($w(d) \mid B$ and $k \equiv -(\frac{A}{w(d)})^{-1}(\frac{B}{w(d)})$ mod $(\frac{d}{w(d)})$). For d with $d \mid m$ and $d \le T$, this implies that

$$\sum_{\substack{|k|\le K\\d|Ak+B}} d \ll d\left(\frac{K}{(d/w(d))} + 1\right) \ll d\left(\frac{(A,m)K}{d} + \frac{T}{d}\right) \ll (A, m)T$$

and hence we get

$$\sum_{\substack{|k|\le K\\(Ak+B,m)\le T}} (Ak+B, m) \ll \sum_{\substack{d|m\\d\le T}} (A, m)T \ll_\varepsilon m^\varepsilon (A, m)T. \qquad \square$$

Lemma 3.18 (Finding a sequence $(\xi_k) \in \Xi$)**.** *Let Z_5, Σ_6 and Ξ be as given in Lemma* 3.10 *and Lemma* 3.11*. Let $z \in Z_5$. Then for all $w_6 \in W_6(z)$ and $k \in \mathbb{Z}$ with $|k| \ll \frac{w_1 \Lambda N}{x^{5\varepsilon}\Delta_1}$, $w_2 \mid k$ and $((\frac{z_1}{w_1})k + (\lambda - \widetilde{\lambda})B, m) \leq T(z)$,*

$$\Sigma_6(z, w_6, k) \ll_\varepsilon \xi_k \left(\frac{x^{4\varepsilon}(w_2, m)T(z)}{q_0}\right)\left(\frac{\Delta_1}{\Delta^*}\right)\left(\frac{N}{m^{1/2}} + m^{1/2}\right)\left(\frac{\Delta^*}{m^{1/2}} + m^{1/2}\right),$$

where $\xi_k = ((\frac{z_1}{w_1})k + (\lambda - \widetilde{\lambda})B, m)(x^\varepsilon(w_2, m)T(z))^{-1}$.

Furthermore, $(\xi_k) \in \Xi$.

Proof. Fix $z \in Z_5$. We recall from Lemma 3.16 that for $w_6 \in W_6(z)$ and $k \in \mathbb{Z}$ with $|k| \ll \frac{w_1 \Lambda N}{x^{5\varepsilon}\Delta_1}$, $w_2 \mid k$ and $((\frac{z_1}{w_1})k + (\lambda - \widetilde{\lambda})B, m) \leq T(z) = \max\{\frac{1}{(w_2,m)}, \frac{1}{H}\}\frac{x^{\delta+100\varepsilon}H^2N}{(v_1,v_2)\Delta_1}$,

$$\Sigma_6(z, w_6, k) \ll_\varepsilon ((\tfrac{z_1}{w_1})k + (\lambda - \widetilde{\lambda})B, m)\left(\frac{x^{3\varepsilon}}{q_0}\right)\left(\frac{\Delta_1}{\Delta^*}\right)\left(\frac{N}{m^{1/2}} + m^{1/2}\right)\left(\frac{\Delta^*}{m^{1/2}} + m^{1/2}\right).$$

Recall $\Lambda \ll \frac{x^{\delta+5\varepsilon}H^2}{w_1(v_1,v_2)}$. So $T(z) = \max\{\frac{1}{(w_2,m)}, \frac{1}{H}\}\frac{x^{\delta+100\varepsilon}H^2N}{(v_1,v_2)\Delta_1} \geq \frac{x^{\delta+100\varepsilon}H^2N}{w_2(v_1,v_2)\Delta_1} \geq \frac{w_1\Lambda N}{x^{5\varepsilon}w_2\Delta_1}$. Recall that $(z_1, m) = 1$ and so $((\frac{z_1}{w_1})w_2, m) = (w_2, m)$. By Lemma 3.17,

$$\sum_{\substack{|k| \ll \frac{w_1\Lambda N}{x^{5\varepsilon}\Delta_1} \\ w_2 \mid k \\ ((\frac{z_1}{w_1})k + (\lambda-\widetilde{\lambda})B, m) \leq T(z)}} ((\tfrac{z_1}{w_1})k + (\lambda - \widetilde{\lambda})B, m) = \sum_{\substack{|k_1| \ll \frac{w_1\Lambda N}{x^{5\varepsilon}w_2\Delta_1} \\ ((\frac{z_1}{w_1})w_2k_1 + (\lambda-\widetilde{\lambda})B, m) \leq T(z)}} ((\tfrac{z_1}{w_1})w_2k_1 + (\lambda - \widetilde{\lambda})B, m) \ll_\varepsilon x^\varepsilon(w_2, m)T(z).$$

Setting $\xi_k = ((\frac{z_1}{w_1})k + (\lambda - \widetilde{\lambda})B, m)(x^\varepsilon(w_2, m)T(z))^{-1}$, we thus get $\sum^{*} \xi_k \ll_\varepsilon 1$ and $(\xi_k) \in \Xi$. This is what we wanted to show. ☐

3.3.5.3 Proof of Proposition 3.1

To prove Proposition 3.1, it remains to combine Lemma 3.10, Lemma 3.11 and Lemma 3.18.

Lemma 3.19 (Small γ)**.** *Let $\omega, \delta > 0$. Let $\varepsilon \in (0, 10^{-100}\delta)$. Let α and β be coefficient sequences at scales M and N with $x \ll MN \ll x$. Assume β has the Siegel-Walfisz property. Write $N = x^\gamma$ and suppose that γ satisfies*

$$\max\left\{\frac{1}{4} + 12\omega + 4\delta + 100\varepsilon,\, 32\omega + 10\delta + 400\varepsilon\right\} \leq \gamma \leq \frac{1}{2} - 4\omega - 2\delta - 50\varepsilon.$$

Then for every $C > 0$ and every $a \in \mathbb{Z}$,

$$\sum_{\substack{d \leq x^{1/2+2\omega} \\ q \mid P(x^\delta) \\ (q,a)=1}} \left| \sum_{n \equiv a(q)} (\alpha \star \beta)(n) - \frac{1}{\phi(q)} \sum_{(n,q)=1} (\alpha \star \beta)(n) \right| \ll_{C,\delta} \frac{x}{(\log x)^C}. \tag{3.37}$$

Proof. Lemma 3.18 tells us that for every $z \in Z_5$, there exists $(\xi_k) \in \Xi$ so that for all $w_6 \in W_6(z)$ and $k \in \mathbb{Z}$ with $|k| \ll \frac{w_1 \Lambda N}{x^{5\varepsilon}\Delta_1}$, $w_2 \mid k$ and $((\frac{z_1}{w_1})k + (\lambda - \widetilde{\lambda})B, m) \le T(z)$,

$$\Sigma_6(z, w_6, k) \ll_\delta \xi_k \left(\frac{x^{4\varepsilon}(w_2,m)T(z)}{q_0}\right)\left(\frac{\Delta_1}{\Delta^*}\right)\left(\frac{N}{m^{1/2}} + m^{1/2}\right)\left(\frac{\Delta^*}{m^{1/2}} + m^{1/2}\right).$$

Suppose the following bound is satisfied for every $z \in Z_5$:

$$\left(\frac{x^{4\varepsilon}T(z)}{q_0}\right)\left(\frac{\Delta_1}{\Delta^*}\right)\left(\frac{N}{m^{1/2}} + m^{1/2}\right)\left(\frac{\Delta^*}{m^{1/2}} + m^{1/2}\right) \ll \min\left\{\frac{1}{x^{\delta+10\varepsilon}H^3}, \frac{1}{H^4}\right\}\frac{q_0(q_0,\ell)N^2}{x^{27\varepsilon}}. \tag{3.38}$$

Then by Lemma 3.11, $\sup_{z\in Z_5} \Sigma_5(z) \ll_\varepsilon \min\{\frac{(w_2,m)}{x^{\delta+10\varepsilon}H^3}, \frac{(w_2,m)}{H^4}\}\frac{q_0^2(q_0,\ell)^2N^2}{x^{27\varepsilon}}$. However, by Lemma 3.10, this upper bound implies (3.37).

Recall that $T(z) = \max\{\frac{1}{(w_2,m)}, \frac{1}{H}\}\frac{x^{\delta+100\varepsilon}H^2N}{(v_1,v_2)\Delta_1} \le \frac{x^{\delta+100\varepsilon}H^2N}{(v_1,v_2)\Delta_1}$ and $\Delta^* \le N$. To prove (3.38) and hence (3.37), it thus remains to verify the following three bounds:

$$\frac{N\Delta^*}{m} \ll \left(\frac{\Delta^*}{\Delta_1}\right)\min\left\{\frac{1}{x^{\delta+10\varepsilon}H^3}, \frac{1}{H^4}\right\}\frac{(v_1,v_2)q_0^2(q_0,\ell)N\Delta_1}{x^{\delta+131\varepsilon}H^2}, \tag{3.39}$$

$$N \ll \left(\frac{\Delta^*}{\Delta_1}\right)\min\left\{\frac{1}{x^{\delta+10\varepsilon}H^3}, \frac{1}{H^4}\right\}\frac{(v_1,v_2)q_0^2(q_0,\ell)N\Delta_1}{x^{\delta+131\varepsilon}H^2}, \tag{3.40}$$

$$m \ll \left(\frac{\Delta^*}{\Delta_1}\right)\min\left\{\frac{1}{x^{\delta+10\varepsilon}H^3}, \frac{1}{H^4}\right\}\frac{(v_1,v_2)q_0^2(q_0,\ell)N\Delta_1}{x^{\delta+131\varepsilon}H^2}. \tag{3.41}$$

Recall that $\frac{N}{x^{\delta+55\varepsilon}H^2} \ll \Delta_1 \ll \frac{N}{x^{55\varepsilon}H^2}$, $\frac{RQ^2H}{q_0(v_1,v_2)\Delta_1} \ll m \ll \frac{x^\delta RQ^2H}{q_0(v_1,v_2)\Delta_1}$, $\Delta^* = \min\{\frac{N}{x^{5\varepsilon}\Lambda}, \Delta_1\}$, $\Lambda \ll \frac{1}{w_1(v_1,v_2)}x^{\delta+5\varepsilon}H^2$, $RQ^2 = x^{-\varepsilon}q_0MH$, $H = x^\varepsilon RQ^2(q_0M)^{-1}$ and $H \ll \frac{x^{4\omega+\delta+7\varepsilon}}{q_0}$.

We begin with (3.39). This condition rearranges to

$$\left(\frac{x^{\delta+131\varepsilon}}{(v_1,v_2)q_0^2(q_0,\ell)}\right)\frac{\max\{x^{\delta+10\varepsilon}H^5, H^6\}}{m} \ll 1. \tag{3.42}$$

Note that $m \gg \frac{RQ^2H}{q_0(v_1,v_2)\Delta_1} \gg \frac{x^{55\varepsilon}RQ^2H^3}{q_0(v_1,v_2)N} \gg \frac{x^{54\varepsilon}MH^4}{(v_1,v_2)N}$. Substituting lower bounds on m and upper bounds on H, the LHS of (3.42) is bounded by

$$\begin{aligned}\left(\frac{x^{\delta+131\varepsilon}}{(v_1,v_2)q_0^2(q_0,\ell)}\right)\frac{\max\{x^{\delta+10\varepsilon}H^5, H^6\}}{m} &\ll x^{-1+\delta+77\varepsilon}N^2\max\{x^{\delta+10\varepsilon}H, H^2\}\\ &\ll x^{-1+8\omega+3\delta+100\varepsilon}N^2.\end{aligned}$$

But we assumed $\gamma \le \frac{1}{2} - 4\omega - 2\delta - 50\varepsilon$. Hence (3.42) and (3.39) hold.

Next we consider (3.40). Rearranging, we find that this condition holds if

$$\frac{x^{\delta+131\varepsilon}}{\Delta^*}\max\{x^{\delta+10\varepsilon}H^5, H^6\} \ll 1.$$

We have $\Delta^* = \min\{\frac{N}{x^{5\varepsilon}\Lambda}, \Delta_1\} \geq \frac{N}{x^{\delta+55\varepsilon}H^2}$ and $H \ll \frac{x^{4\omega+\delta+7\varepsilon}}{q_0}$. Hence (3.40) holds if:

$$x^{32\omega+10\delta+400\varepsilon} \ll N.$$

It thus suffices to assume that $\gamma \geq 32\omega + 10\delta + 400\varepsilon$.

Now we consider the final condition, (3.41). Recall that $m \ll \frac{x^\delta RQ^2 H}{(v_1,v_2)\Delta_1} \ll \frac{q_0 x^\delta MH^2}{(v_1,v_2)\Delta_1}$. So (3.41) is true if

$$\frac{x^{1+2\delta+131\varepsilon}}{N^2\Delta^*\Delta_1} \max\left\{x^{\delta+10\varepsilon}H^7, H^8\right\} \ll 1.$$

Using $H \ll \frac{x^{4\omega+\delta+7\varepsilon}}{q_0}$, we see that (3.41) holds if the following inequalities are satisfied:

$$\frac{x^{1+32\omega+10\delta+200\varepsilon}}{N^2}\left(\frac{x^{5\varepsilon}\Lambda}{\Delta_1 N}\right) \ll 1 \quad \text{and} \quad \frac{x^{1+32\omega+10\delta+200\varepsilon}}{N^2}\left(\frac{1}{\Delta_1^2}\right) \ll 1. \tag{3.43}$$

Using $\Lambda \ll x^{\delta+5\varepsilon}H^2$ and $\Delta_1 \gg \frac{N}{x^{\delta+55\varepsilon}H^2}$, we observe

$$\frac{x^{1+32\omega+10\delta+200\varepsilon}}{N^2}\max\left\{\left(\frac{x^{5\varepsilon}\Lambda}{\Delta_1 N}\right), \left(\frac{1}{\Delta_1^2}\right)\right\} \ll \frac{x^{1+48\omega+16\delta+400\varepsilon}}{N^4}.$$

Hence (3.43) and (3.41) hold when $\gamma \geq 1/4 + 12\omega + 4\delta + 100\varepsilon$ and (3.37) is true. □

Proof of Proposition 3.1 (*Final Steps*) : We consider a coefficient sequence α at scale N and a coefficient sequence β at scale M, where $NM \asymp x$ and $N = x^\gamma$ with $\gamma \in [1/2 - \sigma, 1/2]$. We also assume that β has the Siegel-Walfisz property. Finally, we assume that the following three inequalities are satisfied:

$$\begin{cases} 72\omega + 24\delta < 1, \\ 48\omega + 16\delta + 4\sigma < 1, \\ 64\omega + 20\delta + 2\sigma < 1. \end{cases} \tag{3.44}$$

The second and third inequality rearrange to $1/4+12\omega+4\delta < 1/2-\sigma$ and $32\omega+10\delta < 1/2 - \sigma$. But $\gamma \geq 1/2 - \sigma$. Choosing ε sufficiently small compared to δ, we then find

$$\max\{1/4 + 12\omega + 4\delta + 100\varepsilon, 32\omega + 10\delta + 400\varepsilon\} \leq \gamma.$$

By Lemma 3.19 we thus have the desired equidistribution estimate (3.37) when also $\gamma \leq 1/2 - 4\omega - 2\delta - 50\varepsilon$.

On the other hand, Theorem 2.8 of Polymath [37] tells us that (3.37) holds if $68\omega + 14\delta < 1$ and $\gamma > \frac{4}{17} + \frac{240}{17}\omega + \frac{72}{17}\delta$ and $\gamma > 32\omega + 9\delta$. Observe that $\max\{\frac{4}{17} + \frac{240}{17}\omega + \frac{72}{17}\delta, 32\omega + 9\delta\} < \frac{1}{2} - 4\omega - 2\delta - 50\varepsilon$ if $72\omega + 24\delta < 1$. Hence under the assumptions (3.44), the equidistribution estimate (3.37) holds for every $\gamma \in [1/2 - \sigma, 1/2]$. This concludes the proof of Proposition 3.1. □

3.4 Harman's sieve

In this section we apply Harman's sieve, constructing a function $\rho(n)$ which is a minorant of the prime indicator function $1_{\mathbb{P}}(n)$ and has exponent of distribution $\theta = 0.5253 = 1/2 + 2 \cdot 0.01265$.

3.4.1 Motivation

We first explain our choice of θ. In the arguments of Maynard [30] and Polymath [38], bounds on H_m were derived using the exponent of distribution of $\Lambda(n)$. This exponent of distribution was determined by decomposing $\Lambda(n)$ via the Heath-Brown identity and treating various convolutions of coefficient sequences separately. Certain bad coefficient sequences, corresponding to convolutions of 5 sequences of length about $x^{0.2}$, formed a barrier to improving the exponent of distribution of $\Lambda(n)$ further. The idea behind a use of Harman's sieve is to discard these bad convolutions: Baker and Irving [2] noticed that if we replace $\Lambda(n)$ by $\log(x)\rho_\omega(n)$, where $\rho_\omega(n)$ is a minorant for $1_{\mathbb{P}}(n)$ which satisfies

$$\sum_{\substack{q \le x^{1/2+2\omega} \\ q \mid P(x^\delta) \\ (q,a)=1}} \left| \sum_{\substack{n \in [x,2x] \\ n \equiv a(q)}} \rho_\omega(n) - \frac{1}{\phi(q)} \sum_{\substack{n \in [x,2x] \\ (n,q)=1}} \rho_\omega(n) \right| \ll_{A,\delta} \frac{x}{(\log x)^A} \tag{3.45}$$

and $\sum_{x \le n \le 2x} \rho_w(n) = (1 - c_1(\omega) + o(1))\frac{x}{\log x}$, then we still have a bound of the form

$$H_m \ll_\epsilon \exp\left(\frac{(1+\epsilon)m}{(1/4+\omega)(1-c_1(\omega))}\right).$$

The quantity $c_1(\omega)$ describes the loss incurred by discarding bad convolutions and increases as ω increases, since fewer and fewer equidistribution estimates are available. (For $\omega = 1/80$ thus $c_1(\omega) = 0$.) Harman's sieve then gives an improvement over a direct treatment of $\Lambda(n)$ whenever the loss $c_1(\omega)$, incurred by discarding bad cases, does not outweigh the gain of a larger ω, so that $(1/4+\omega)(1-c_1(\omega)) > 1/4 + 1/80$.

3.4.1.1 Limitations

This particular application of Harman's sieve has a discontinuity, present in the same form both in Baker and Irving's and in our work:

The minorant $\rho_\omega(n)$ is constructed using the Buchstab identity, and in the process one must consider whether or not $\alpha_1 \star \alpha_2 \star \alpha_3$ with $\alpha_i(n) = 1_{\mathbb{P}}(n)$ on $[N_i, 2N_i]$ and

$N_1N_2N_3 \asymp x$ has exponent of distribution $1/2 + 2\omega$. To do so, each α_i is first decomposed further via the Heath-Brown identity and we consider $\alpha_i = \psi_i \star \beta_i$, where ψ_i and β_i have support $[A_i, 2A_i]$ and $[B_i, 2B_i]$, $A_iB_i \approx N_i$, $B_i \le x^{0.1}$ and $\psi_i(n) = 1$ on $[A_i, 2A_i]$. Good equidistribution estimates are available if $A_1A_2A_3$ is large or if the product of some subset of $\{A_1, A_2, A_3, B_1, B_2, B_3\}$ is close to $x^{0.5}$. In particular, if $\omega = 1/79$ then we either need that $A_1A_2A_3 \ge x^{0.8925}$ or that a product of $A_1, A_2, A_3, B_1, B_2, B_3$ is contained in $[x^{0.405}, x^{0.595}]$, and one of these two cases is always satisfied by $\psi_1 \star \psi_2 \star \psi_3 \star \beta_1 \star \beta_2 \star \beta_3$.

However, as soon as ω is increased slightly beyond $1/79$, our equidistribution estimates do not cover the case where $A_1A_2A_3$ is very close to $x^{0.8925}$ – the three smooth factors are too short for Polymaths's Type III estimate, but too long for our Type I/II estimate. When constructing $\rho_\omega(n)$, this forces us to discard many products of three primes in Buchstab's identity, leading to a sudden jump where $c_1(\omega)$ becomes much larger as ω increases beyond $1/79$.

This discontinuity determines the choice $\omega = 0.01265 < 1/79$. Here we still have $c_1(\omega) < 2 \cdot 10^{-5}$ and thus obtain the proposed bound $H_m = O(\exp(3.8075m))$.

3.4.2 Type I, II and III information

We now record the available Type I, II and III information for the case $\omega = 0.01265 \approx \frac{1}{79}$. (Here we are using Type I/II in the sense of Harman rather than Polymath.)

Lemma 3.20. *Let $\delta = \frac{1}{10^{10}}$. Suppose $f : \mathbb{N} \to \mathbb{C}$ satisfies one of the following conditions:*

(I) *$f = \alpha \star \beta$, where α is coefficient sequence at scale M and β is a **smooth** coefficient sequence at scale N with $MN \asymp x$. Additionally, $N = x^\gamma$ with $\gamma \ge 0.33856 - \delta$.*

(II) *$f = \alpha \star \beta$, where α is coefficient sequence at scale M, β is a coefficient sequence at scale N with $MN \asymp x$, and α and β both have the Siegel–Walfisz property. Additionally, $N = x^\gamma$ with $0.40481 - \delta \le \gamma \le 0.59519 + \delta$.*

(III) *$f = \alpha \star \psi_1 \star \psi_2 \star \psi_3$, where α is a coefficient sequence at scale M and ψ_1, ψ_2, ψ_3 are **smooth** coefficient sequences at scales N_1, N_2, N_3 with $MN_1N_2N_3 \asymp x$. Additionally, $N_1N_2, N_1N_3, N_2N_3 \ge x^{0.595-\delta}$ and $x^{0.19-\delta} \le N_1, N_2, N_3 \le x^{0.405+\delta}$.*

Then for any integer $a \in \mathbb{Z}$ and for any $A > 0$, we have

$$\sum_{\substack{q \le x^{0.5253} \\ q \mid P(x^\delta) \\ (q,a)=1}} \left| \sum_{n \equiv a(q)} f(n) - \frac{1}{\phi(q)} \sum_{(n,q)=1} f(n) \right| \ll_A \frac{x}{(\log x)^A}.$$

In particular, f has exponent of distribution 0.5253 *to smooth moduli.*

Proof. For option (I), we use Lemma 5 of Baker and Irving [2]. We take $\eta = 0.0138825$ to get $\theta = 0.5253$ and observe that $0.33856 > 199/600 + 119\eta/240$. For option (II), we use Proposition 3.1. We take $\omega = 0.01265$ and observe that $0.40481 > \max\left\{\frac{1}{4} + 12\omega, 32\omega\right\}$. Finally, for option (III) we use part (v) of Theorem 2.8 of Polymath [37]. Again we take $\omega = 0.01265$ and observe that $0.595 > 1/2 + 1/18 + 28\omega/9$, while $0.405 < 1/2 - 1/18 - 28\omega/9$ and $0.19 > 2(1/18 + 28\omega/9)$. □

3.4.3 Good sifted sets

From here on out, we follow the arguments of Baker and Irving [2] closely, effectively only updating their proofs with better Type II information. As a first step, we record a version of Baker and Irving's Lemma 7. Recall that

$$\psi(n, y) = \begin{cases} 1 & \text{if } p \mid n \to p \ge y, \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 3.21. *Write $p_i = x^{\alpha_i}$. Let $\zeta = 1 - 0.33856 - 0.40481 = 0.25663$ and $\lambda = 0.59519 - 0.40481 = 0.19038$. Each of the functions given below has exponent of distribution* 0.5253 *to smooth moduli.*

$$\theta_1(n) = \psi(n, x^\lambda). \tag{1}$$

$$\theta_2(n) = \sum_{\substack{n = p_1 n_2 \\ 0.19038 \le \alpha_1 < 0.40481}} \psi(n_2, x^\lambda). \tag{2}$$

$$\theta_3(n) = \sum_{\substack{n = p_1 p_2 n_3 \\ 0.19038 \le \alpha_2 < \alpha_1 < 0.40481 \\ \alpha_1 + \alpha_2 < 0.40481}} \psi(n_3, x^\lambda). \tag{3}$$

$$\theta_4(n) = \sum_{\substack{n = p_1 p_2 n_3 \\ 0.19038 \le \alpha_2 < \alpha_1 < 0.40481 \\ \alpha_1 + \alpha_2 > 0.59519,\, \alpha_2 < \zeta}} \psi(n_3, x^\lambda). \tag{4}$$

$$\theta_5(n) = \sum_{\substack{n=p_1p_2p_3n_4\\ 0.19038\le\alpha_3<\alpha_2<\alpha_1<0.40481\\ \alpha_1+\alpha_2<0.40481,\,\alpha_3<\zeta}} \psi(n_4, x^\lambda). \tag{5}$$

$$\theta_6(n) = \sum_{\substack{n=n_1p_2p_3p_4\\ 0.19038\le\alpha_4<\alpha_3<\alpha_2<0.40481\\ \alpha_3+\alpha_4<0.40481,\,\alpha_2<\zeta,\,\alpha_4\ge\alpha_3}} \psi(n_1, x^\lambda). \tag{6}$$

Proof. This follows from the exact same arguments as Lemma 7 of [2]. We have simply updated the available Type II information.

More precisely, we use Lemma 14 of Baker and Weingartner [3] with $\alpha = 0.40481$, $\beta = 0.59519 - 0.40481 = 0.19038 = \lambda$, so that $\alpha + \beta = 0.59519$, and $M = x^{1-0.33856} = x^{0.66144}$. In the notation of that lemma we must then take $R < x^{0.40481}$ and $S < Mx^{-\alpha} = x^{0.25663} = x^\zeta$. After verifying that the sequences which appear in the proof of Lemma 14 of [3] have the Siegel-Walfisz property, we obtain that $\theta_k(n)$ has exponent of distribution 0.5253 to smooth moduli if the following is true: We can partition $\{1, \dots, j\}$ into I and J so that $\sum_{i\in I} \alpha_i < 0.40481$ and $\sum_{i\in J} \alpha_i < \zeta = 0.25663$. (Here j is the number of p_i which appear in $\theta_k(n)$. In (6) we partition $\{2, 3, 4\}$ rather than $\{1, \dots, j\}$.)

For (1), we simply take $I = J = \emptyset$. In (2), we choose $I = \{1\}$ and $J = \emptyset$. In (3) we take $I = \{1\}$ and $J = \{2\}$. We use that $\alpha_2 \le (\alpha_1 + \alpha_2)/2 < 0.40481/2 < \zeta$. In (4) we also take $I = \{1\}$ and $J = \{2\}$. For (5), we take $I = \{1, 2\}$ and $J = \{3\}$. Finally, in (6) we use $I = \{3, 4\}$ and $J = \{2\}$. $\square$

3.4.4 Buchstab decompositions

We now decompose the prime indicator function using Buchstab iterations. Write $\lambda = 0.19038$ and $\zeta = 0.25663$ and set $p_i = x^{\alpha_i}$. For $n \in [x, 2x] \cap \mathbb{N}$, we have

$$\begin{aligned}
1_{\mathbb{P}}(n) &= \psi(n, (3x)^{1/2})\\
&= \psi(n, x^\lambda) - \sum_{\substack{n=p_1n_2\\ \lambda\le\alpha_1<0.40481}} \psi(n_2, x^\lambda) - \sum_{\substack{n=p_1n_2\\ 0.40481\le\alpha_1<0.5}} \psi(n_2, p_1) &(3.46)\\
&+ \sum_{\substack{n=p_1p_2n_3\\ \lambda\le\alpha_2<\alpha_1<0.40481\\ \alpha_1+\alpha_2<0.40481}} \psi(n_3, x^\lambda) + \sum_{\substack{n=p_1p_2n_3\\ \lambda\le\alpha_2<\alpha_1<0.40481\\ 0.40481\le\alpha_1+\alpha_2\le0.59519}} \psi(n_3, p_2) + \sum_{\substack{n=p_1p_2n_3\\ \lambda\le\alpha_2<\alpha_1<0.40481\\ \alpha_1+\alpha_2>0.59519\\ \alpha_2<\zeta}} \psi(n_3, x^\lambda) &(3.47)\\
&+ \sum_{\substack{n=p_1p_2n_3\\ \lambda\le\alpha_2<\alpha_1<0.40481\\ \alpha_1+\alpha_2>0.59519\\ \alpha_2\ge\zeta}} \psi(n_3, p_2) - \sum_{\substack{n=p_1p_2p_3n_4\\ \lambda\le\alpha_3<\alpha_2<\alpha_1<0.40481\\ \alpha_1+\alpha_2<0.40481\\ \alpha_3<\zeta}} \psi(n_4, x^\lambda) - \sum_{\substack{n=p_1p_2p_3n_4\\ \lambda\le\alpha_3<\alpha_2<\alpha_1<0.40481\\ \alpha_1+\alpha_2<0.40481\\ \alpha_3\ge\zeta}} \psi(n_4, p_3) &(3.48)
\end{aligned}$$

$$+\sum_{\substack{n=p_1p_2p_3p_4n_5\\ \lambda\le\alpha_4<\alpha_3<\alpha_2<\alpha_1<0.40481\\ \alpha_1+\alpha_2<0.40481\\ \alpha_3<\zeta}}\psi(n_5,p_4)-\sum_{\substack{n=p_1p_2p_3n_4\\ \lambda\le\alpha_3<\alpha_2<\alpha_1<0.40481\\ \alpha_1+\alpha_2>0.59519\\ \alpha_2<\zeta}}\psi(n_4,p_3). \tag{3.49}$$

We may apply part (1) of Lemma 3.21 to the first function in (3.46) and part (2) to the second function in (3.46). Further, the third function can be treated directly, using our Type II information, recorded in Lemma 3.20. The first and third function in (3.47) can be treated via part (3) and part (4) of Lemma 3.21, respectively. On the other hand, the second function is treated directly, using Type II information from Lemma 3.20. For (3.48), part (5) of Lemma 3.21 can be applied to the second expression. If $\alpha_3\ge\zeta$, then $\alpha_1+\alpha_2\ge 2\cdot 0.25663>0.40481$. Hence the third sum is zero. This leaves only the first function in (3.48). Finally, for the moment, we also leave the two functions in (3.49) unchanged.

Overall we have found that there exists a function $\theta_0(n)$ which has exponent of distribution 0.5253 to smooth moduli and which satisfies that for all $n\in[x,2x]\cap\mathbb{N}$,

$$1_{\mathbb{P}}(n)=\theta_0(n)+\sum_{\substack{n=p_1p_2n_3\\ 0.19038\le\alpha_2<\alpha_1<0.40481\\ \alpha_1+\alpha_2>0.59519\\ \alpha_2\ge0.25663}}\psi(n_3,p_2)+\sum_{\substack{n=p_1p_2p_3p_4n_5\\ 0.19038\le\alpha_4<\alpha_3<\alpha_2<\alpha_1<0.40481\\ \alpha_1+\alpha_2<0.40481\\ \alpha_3<0.25663}}\psi(n_5,p_4)-\sum_{\substack{n=p_1p_2p_3n_4\\ 0.19038\le\alpha_3<\alpha_2<\alpha_1<0.40481\\ \alpha_1+\alpha_2>0.59519\\ \alpha_2<0.25663}}\psi(n_4,p_3).$$

3.4.5 3 prime factors

Next we inspect the first sum in the above expression for $1_{\mathbb{P}}(n)$ more closely. This corresponds to Lemma 10 of [2].

Lemma 3.22. *For $n\in[x,2x]\cap\mathbb{N}$, define the function*

$$\Gamma(n)=\sum_{\substack{n=p_1p_2n_3\\ 0.19038\le\alpha_2<\alpha_1<0.40481\\ \alpha_1+\alpha_2>0.59519\\ \alpha_2\ge0.25663}}\psi(n_3,p_2).$$

Then Γ has exponent of distribution 0.5253 to smooth moduli.

Proof. Consider $n=p_1p_2n_3$ with $0.19038\le\alpha_2<\alpha_1<0.40481$ and $\alpha_2\ge0.25663$. Suppose that n_3 has no prime factor less than p_2. Assume for a contradiction that n_3 is not prime. Then n has at least 4 prime factors, all of which are of size at least $x^{0.25663}$. But then n is certainly larger than $2x$, assuming that x is large. We work with $n\le 2x$, so this gives a contradiction. Hence it suffices to consider $n=p_1p_2p_3$.

We now look at the sum

$$\Gamma_0(n) = \sum_{\substack{n=p_1p_2p_3 \\ 0.19038 \le \alpha_2 < \alpha_1 < 0.40481 \\ \alpha_1+\alpha_2 > 0.59519 \\ \alpha_3 \ge \alpha_2 \ge 0.25663}} 1.$$

Applying the Heath-Brown identity like in Section 3 of Polymath [37], we find that $\Gamma_0(n)$ has exponent of distribution 0.5253 to smooth moduli provided that each coefficient sequence $\beta \in \mathcal{B}$, with $\mathcal{B}$ as described below, has exponent of distribution 0.5253 to smooth moduli.

The set $\mathcal{B}$ denotes the set of coefficient sequences

$$\beta = \beta_1 \star \cdots \star \beta_k$$

with $k \le 40$ which satisfy the following properties:

1. Each β_i is a coefficient sequence, located at some scale N_i, with $\prod_{i=1}^k N_i \asymp x$. Further, if $S \subseteq \{1, \ldots, k\}$, then $\bigstar_{i\in S}\beta_i$ is a coefficient sequence at scale $\prod_{i\in S} N_i$.
2. If $N_i \ge x^{0.1}$, then β_i is smooth at scale N_i.
3. If $S \subseteq \{1, \ldots, k\}$ with $\prod_{i\in S} N_i \ge x^\delta = x^{10^{-10}}$, then $\bigstar_{i\in S}\beta_i$ satisfies the Siegel-Walfisz property.
4. There exits a partition I_1, I_2, I_3 of $\{1, \ldots, k\}$ such that $x^{\gamma_i} = \prod_{i\in I_j} N_i$ satisfy $0.19038 - \frac{\delta}{10} \le \gamma_1, \gamma_2 \le 0.40481 + \frac{\delta}{10}$, $\gamma_1 + \gamma_2 \ge 0.59519 - \frac{\delta}{10}$, $\gamma_2 \ge 0.25663 - \frac{\delta}{10}$ and $\gamma_1, \gamma_3 \ge \gamma_2 - \frac{\delta}{10}$.

It remains to show that the functions in $\mathcal{B}$ have the desired exponent of distribution. As a first step, partition $\{1, \ldots, k\}$ into sets $J_1, \ldots, J_j$ such that $x^{\mu_s} = \prod_{i\in J_s} N_i$ satisfies $\mu_s \ge 0.05$ for all but at most one $s \in \{1, \ldots j\}$ and such that J_s is a singleton whenever $\mu_s \ge 0.1$. In particular, $\bigstar_{i\in J_s}\beta_i$ is smooth when $\mu_s \ge 0.1$. We also assume $\mu_1 \ge \cdots \ge \mu_j$.

If there exits s with $\mu_s \ge 0.33856-\delta$, then we may use the Type I information available in Lemma 3.20 to deduce that β has exponent of distribution 0.5253 to smooth moduli. If there exists $K \subseteq \{1, \ldots, j\}$ with $\sum_{s\in K} \mu_s \in [0.40481 - \delta, 0.59519 + \delta]$, we use Type II information to deduce the same. Hence we may assume that neither of these two statements holds.

Now consider $\mu_2 + \mu_3$. Since $\sum_{s \in K} \mu_s \notin [0.40481 - \delta, 0.59519 + \delta]$, we either have $\mu_2 + \mu_3 \geq 0.59519 + \delta$ or $\mu_2 + \mu_3 \leq 0.40481 - \delta$. Assume first that $\mu_2 + \mu_3 \geq 0.59519 + \delta$. Then also $\mu_1 + \mu_2, \mu_1 + \mu_3 \geq 0.59519 + \delta$. We certainly must have $\mu_1 \leq 0.405 + \delta$, since otherwise $\mu_1 \geq 0.59519 + \delta$ and $\mu_1 + \mu_2 + \mu_3 > 1$. Hence if also $\mu_3 \geq 0.19$, then $\bigstar_{i \in J_s} \beta_i$ is smooth for $s \leq 3$, and by part (III) of Lemma 3.20, the sequence

$$(\bigstar_{i \in J_1} \beta_i) \star (\bigstar_{i \in J_2} \beta_i) \star (\bigstar_{i \in J_3} \beta_i) \star (\bigstar_{i \notin J_1 \cup J_2 \cup J_3} \beta_i).$$

has exponent of distribution 0.5253 to smooth moduli. (The set $\{1, \ldots, j\} \setminus (J_1 \cup J_2 \cup J_3)$ may be empty. In that case we replace the final (empty) convolution by the delta function.) Thus we may assume that $\mu_3 < 0.19$. But this implies $\mu_2 \geq 0.40519 + \delta$, contradicting our earlier assumptions. This concludes the treatment of the case $\mu_2 + \mu_3 \geq 0.59519 + \delta$.

Finally, we consider the case $\mu_2 + \mu_3 \leq 0.40481 - \delta$. We have $\mu_2 + \cdots + \mu_j \geq 1 - 0.33856 = 0.66144$, so if $\mu_4 \leq 0.59519 - 0.40481 = 0.19038$, then there exists K with $\sum_{s \in K} \mu_s \in [0.40481 - \delta, 0.59519 + \delta]$, contradicting our earlier assumption. Similarly, since $\mu_4 \leq \mu_3 \leq 0.21$, we have $\mu_2 + \mu_3 + \mu_5 + \cdots + \mu_j \geq 1 - 0.33856 - 0.21 = 0.45144 > 0.40481$, and thus may also assume that $\mu_5 \geq 0.19038$. Notice that $\mu_1, \ldots, \mu_5 > 0.1$ and thus $J_1, \ldots, J_5$ are singletons. In particular, there exist N_i with $N_i = x^{\mu_i} \geq x^{0.19038}$ for $i \leq 5$. We also have $\mu_i + (\mu_6 + \cdots + \mu_j) \leq \mu_1 + (\mu_6 + \cdots + \mu_j) \leq 1 - 4 \cdot 0.19038 = 0.23848$ for $1 \leq i \leq 5$.

Looking at condition (4) of $\mathcal{B}$, we have $\gamma_1, \gamma_2, \gamma_3 \geq 0.25663 - \delta$. Each of $\{N_i : i \in I_1\}$ and $\{N_i : i \in I_2\}$ and $\{N_i : i \in I_3\}$ must then contain at least two choices of x^{μ_i} with $1 \leq i \leq 5$. This is because $\mu_i + (\mu_6 + \cdots + \mu_j) < 0.25663 - \delta$ for $1 \leq i \leq 5$. But $2 \cdot 3 = 6 > 5$, so there are not enough x^{μ_i} available to satisfy this requirement. Hence condition (4) tells us that $\mu_2 + \mu_3 \leq 0.40481 - \delta$ is impossible unless there exists K with $\sum_{s \in K} \mu_s \in [0.40481 - \delta, 0.59519 + \delta]$. This concludes the proof. □

3.4.6 5 prime factors

Next we consider products of 5 primes. This corresponds to Lemma 11 of [2].

Lemma 3.23. *For $n \in [x, 2x] \cap \mathbb{N}$, define the function*

$$\Gamma(n) = \sum_{\substack{n = p_1 p_2 p_3 p_4 n_5 \\ 0.19038 \leq \alpha_4 < \alpha_3 < \alpha_2 < \alpha_1 < 0.40481 \\ \alpha_1 + \alpha_2 < 0.40481 \\ \alpha_3 < 0.25663}} \psi(n_5, p_4).$$

If x is sufficiently large, we can write $\Gamma(n) = \Gamma_1(n)+\Gamma_2(n)$, where $\Gamma_1(n)$ has exponent of distribution 0.5253 *to smooth moduli and $\Gamma_2(n)$ is non-negative with*

$$\sum_{n\in[x,2x]} \Gamma_2(n) = \sum_{n\in[x,2x]} \sum_{\substack{n=p_1p_2p_3p_4n_5\\ 0.19038\le\alpha_4<\alpha_3<\alpha_2<\alpha_1<0.40481\\ \alpha_1+\alpha_2<0.40481\\ \alpha_2+\alpha_3+\alpha_4>0.59519}} \psi(n_5,p_4) \le \frac{3x}{10^7\log x}.$$

Proof. The inequality $\alpha_3 < 0.25663$ automatically follows from $\alpha_1+\alpha_2 < 0.40481$ and $\alpha_3 < \alpha_2$, so we may remove it. We now split $\Gamma(n)$ up into

$$\Gamma(n) = \Gamma_1(n)+\Gamma_2(n) = \sum_{\substack{n=p_1p_2p_3p_4n_5\\ 0.19038\le\alpha_4<\alpha_3<\alpha_2<\alpha_1<0.40481\\ \alpha_1+\alpha_2<0.40481\\ \alpha_2+\alpha_3+\alpha_4\le0.59519}} \psi(n_5,p_4) + \sum_{\substack{n=p_1p_2p_3p_4n_5\\ 0.19038\le\alpha_4<\alpha_3<\alpha_2<\alpha_1<0.40481\\ \alpha_1+\alpha_2<0.40481\\ \alpha_2+\alpha_3+\alpha_4>0.59519}} \psi(n_5,p_4).$$

Notice here that $\alpha_2+\alpha_3+\alpha_4 \ge 3\cdot 0.19038 > 0.40481$. Hence our Type II information tells us that $\Gamma_1(n)$ has exponent of distribution 0.5253 to smooth moduli. On the other hand, if $\alpha_2+\alpha_3+\alpha_4 > 0.59519$, then $\alpha_1+\alpha_2+\alpha_3+3\alpha_4 > (4/3)\cdot 0.59519+2\cdot 0.19038 > 1$. Hence if n_5 contributes positively to $\Gamma_2(n)$, then n_5 is prime. Using standard techniques, and in particular the prime number theorem for short intervals, we find that $\sum_{n\in[x,2x]}\Gamma_2(n)$ equals

$$\frac{x(1+o(1))}{\log x}\int_{0.19038}^{0.40481}\int_{0.19038}^{\min\{\alpha_1,0.40481-\alpha_1\}}\int_{0.19038}^{\alpha_2}\int_{\max\{0.19038,0.59519-\alpha_2-\alpha_3\}}^{\min\{\alpha_3,(1-\alpha_1-\alpha_2-\alpha_3)/2\}} \frac{1}{\alpha_1\alpha_2\alpha_3\alpha_4(1-\sum_{i=1}^4\alpha_i)}d\underline{\alpha}.$$

In particular, we have $\sum_{n\in[x,2x]}\Gamma_2(n) \le \frac{3x}{10^7\log x}$ when x is large. □

3.4.7 4 prime factors

It remains to look at products of 4 primes. This corresponds to Lemma 13 of [2], and we closely follow the steps of that proof.

Lemma 3.24. *For $n \in [x,2x]\cap\mathbb{N}$, define the function*

$$\Gamma(n) = \sum_{\substack{n=p_1p_2p_3n_4\\ 0.19038\le\alpha_3<\alpha_2<\alpha_1<0.40481\\ \alpha_1+\alpha_2>0.59519\\ \alpha_2<0.25663}} \psi(n_4,p_3).$$

If x is sufficiently large, we can write $\Gamma(n) = \Gamma_1(n)-\Gamma_2(n)$, where $\Gamma_1(n)$ has exponent of distribution 0.5253 *to smooth moduli and $\Gamma_2(n)$ is non-negative with*

$$\sum_{n\in[x,2x]} \Gamma_2(n) = \sum_{n\in[x,2x]} \sum_{\substack{n=p_2p_3p_4p_5p_6\\ 0.19038\le\alpha_2,\alpha_3,\alpha_4,\alpha_5,\alpha_6\le0.23848\\ \alpha_2,\alpha_4>\alpha_3\\ \alpha_2+\alpha_4<0.40481\\ \alpha_3+\alpha_4+\alpha_5>0.59519\\ \alpha_6\ge\alpha_5}} 1 \le \frac{0.00001x}{\log x}.$$

Proof. Notice that $\alpha_1+\alpha_2+3\alpha_3 > 0.59519+3\cdot 0.19038 > 1$. Hence if n_4 contributes to $\Lambda(n)$, then n_4 is prime. Notice also that $\alpha_1+\alpha_2 > 0.59519$ and $\alpha_2 < 0.25663$ imply $\alpha_1 > \alpha_2$. Further, $\alpha_i \geq 0.19038$ implies $\alpha_1 < 1-3\cdot 0.19038 < 0.59519$. We may write

$$\Gamma(n) = \Upsilon_1(n)+\Upsilon_2(n) = \sum_{\substack{n=p_1p_2p_3p_4\\ 0.19038\leq\alpha_3<\alpha_2<0.40481\\ \alpha_1+\alpha_2>0.59519\\ \alpha_2<0.25663\\ \alpha_4\geq\alpha_3}} 1 \quad - \sum_{\substack{n=p_1p_2p_3p_4\\ 0.19038\leq\alpha_3<\alpha_2<0.40481\\ \alpha_1+\alpha_2>0.59519\\ \alpha_2<0.25663\\ \alpha_4\geq\alpha_3\\ \alpha_1\in[0.40481,0.59519]}} 1.$$

Using the available Type II information, we immediately see that the second sum has exponent of distribution 0.5253 to smooth moduli. $\Upsilon_1(n)$ may be rewritten as

$$\Upsilon_1(n) = \sum_{\substack{n=p_1p_2p_3p_4\\ 0.19038\leq\alpha_3<\alpha_2<0.40481\\ \alpha_3+\alpha_4<0.40481\\ \alpha_2<0.25663\\ \alpha_4\geq\alpha_3}} 1.$$

First applying reversal of roles to p_1, and then using the Buchstab identity, we get

$$\begin{aligned}\Upsilon_1(n) &= \sum_{\substack{n=n_1p_2p_3p_4\\ 0.19038\leq\alpha_3<\alpha_2<0.40481\\ \alpha_3+\alpha_4<0.40481\\ \alpha_2<0.25663\\ \alpha_4\geq\alpha_3}} \psi(n_1,(\tfrac{2x}{p_2p_3p_4})^{1/2})\\ &= \sum_{\substack{n=n_1p_2p_3p_4\\ 0.19038\leq\alpha_3<\alpha_2<0.40481\\ \alpha_3+\alpha_4<0.40481\\ \alpha_2<0.25663\\ \alpha_4\geq\alpha_3}} \psi(n_1,x^{\lambda}) \quad - \sum_{\substack{n=p_2p_3p_4p_5n_6\\ 0.19038\leq\alpha_3<\alpha_2<0.40481\\ \alpha_3+\alpha_4<0.40481\\ \alpha_2<0.25663\\ \alpha_4\geq\alpha_3\\ 0.19038\leq\alpha_5<(1-\alpha_2-\alpha_3-\alpha_4)/2}} \psi(n_6,p_5).\end{aligned}$$

Notice that part (6) of Lemma 3.21 applies to the first sum in the second line, which hence has exponent of distribution 0.5253 to smooth moduli. Hence we may focus on the second sum. Observe that $6\cdot 0.19038 > 1$. Hence if n_6 contributes a non-zero amount to $\Upsilon_1(n)$, then we must have n_6 prime. Let

$$\Upsilon_3(n) = \sum_{\substack{n=p_2p_3p_4p_5p_6\\ 0.19038\leq\alpha_3<\alpha_2<0.40481\\ \alpha_3+\alpha_4<0.40481\\ \alpha_2<0.25663\\ \alpha_4\geq\alpha_3\\ 0.19038\leq\alpha_5\leq\alpha_6}} 1.$$

Our proof is finished if we can show that $\Upsilon_3(n)$ is the sum of a function $\Upsilon_4(n)$ which has exponent of distribution 0.5253 to smooth moduli and a non-negative function $\Gamma_2(n)$ with $\sum_{n\in[x,2x]}\Gamma_2(n) \leq \frac{10^{-5}x}{\log x}$.

The contribution of $(\alpha_2, \ldots, \alpha_6)$ with $\sum_{i \in S} \alpha_i \in [0.40481, 0.59519]$ to $\Upsilon_3(n)$ can be treated via the available Type II information. Notice also that ($\alpha_i \geq 0.19038$ for all i) implies $\alpha_i \leq 1 - 4 \cdot 0.19038 = 0.23848 < 0.25663$ for $i \in \{2, 3, 4, 5\}$. So we consider

$$\Upsilon_3^*(n) = \sum^*_{\substack{n=p_2p_3p_4p_5p_6 \\ 0.19038 \leq \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6 \leq 0.23848 \\ \alpha_2, \alpha_4 > \alpha_3 \\ \alpha_3 + \alpha_4 < 0.40481 \\ \alpha_6 \geq \alpha_5}} 1,$$

where $\sum^*$ denotes the sum over $(\alpha_2, \ldots, \alpha_6)$ with $\sum_{i \in S} \alpha_i \notin [0.40481, 0.59519]$ for all $S \subseteq \{2, \ldots, 6\}$. Observe that $\alpha_2 + \alpha_4 < 2 \cdot 0.24 < 0.59519$. Hence we may assume $\alpha_2 + \alpha_4 < 0.40481$. This immediately also implies $\alpha_3 + \alpha_4 < 0.40481$. Further, $\alpha_3 + \alpha_4 + \alpha_5 \geq 3 \cdot 0.19038 > 0.40481$ and so $\alpha_3 + \alpha_4 + \alpha_5 \geq 0.59519$.

$$\Upsilon_3^*(n) \leq \Gamma_2(n) = \sum_{\substack{n=p_2p_3p_4p_5p_6 \\ 0.19038 \leq \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6 \leq 0.23848 \\ \alpha_2, \alpha_4 > \alpha_3 \\ \alpha_2 + \alpha_4 < 0.40481 \\ \alpha_3 + \alpha_4 + \alpha_5 > 0.59519 \\ \alpha_6 \geq \alpha_5}} 1.$$

Using standard techniques, we find that $\sum_{n \in [x, 2x]} \Gamma_2(n)$ is asymptotically equal to

$$\frac{x}{\log x} \int_{0.19038}^{0.23848} \int_{0.19038}^{\min\{0.23848, \alpha_2\}} \int_{\max\{\alpha_3, 0.19038\}}^{\min\{0.23848, 0.40481 - \alpha_2\}} \int_{\max\{0.19038, 0.59519 - \alpha_3 - \alpha_4\}}^{\min\{0.23848, (1 - \sum_{i=2}^4 \alpha_i)/2\}} \frac{d\underline{\alpha}}{\prod_{i=2}^5 \alpha_i (1 - \sum_{i=2}^5 \alpha_i)}.$$

In particular, we have $\sum_{n \in [x, 2x]} \Gamma_2(n) \leq \frac{0.00001x}{\log x}$ when x is large. □

3.4.8 Construction of a minorant

Now combining the Buchstab decomposition given in Section 3.4.4 with Lemma 3.22, Lemma 3.23 and Lemma 3.24, we have that for $n \in [x, 2x] \cap \mathbb{N}$,

$$1_{\mathbb{P}}(n) = \theta^*(n) \;+ \sum_{\substack{n=p_1p_2p_3p_4n_5 \\ 0.19038 \leq \alpha_4 < \alpha_3 < \alpha_2 < \alpha_1 < 0.40481 \\ \alpha_1 + \alpha_2 < 0.40481 \\ \alpha_2 + \alpha_3 + \alpha_4 > 0.59519}} \psi(n_5, p_4) \;+ \sum_{\substack{n=p_2p_3p_4p_5p_6 \\ 0.19038 \leq \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6 \leq 0.23848 \\ \alpha_2, \alpha_4 > \alpha_3 \\ \alpha_2 + \alpha_4 < 0.40481 \\ \alpha_2 + \alpha_3 + \alpha_5 > 0.59519 \\ \alpha_6 \geq \alpha_5}} 1,$$

where $\theta^*(n)$ has exponent of distribution 0.5253 to smooth moduli and where

$$\sum_{n \in [x, 2x]} (1_{\mathbb{P}}(n) - \theta^*(n)) \leq \frac{0.0000103x}{\log x}$$

when x is sufficiently large.

Setting $\rho(n) = \theta^*(n)$, this concludes the proof of Proposition 3.2. □

Appendix A

Harman's sieve: Computations

In Appendix A, we complete the proof of Proposition 2.3 for the cases $a \in (0.59, 0.61]$, $a \le 0.53$, $a \in (0.545, 0.57]$ and $a > 0.61$.

(Note: The detailed treatment of the cases $a \in (0.53, 0.545]$ and $a \in (0.57, 0.59]$ was aimed at readers who are not very familiar with Harman's sieve. As such, we used the convenient, but somewhat non-standard notation of Section 2.6.3.1. The purpose of this appendix is to supply the computations needed to treat the remaining cases, but with a much smaller focus on details. As the intended audience is a different one, we now return to standard Harman's sieve notation.)

<u>Case 3</u>: $a \in (0.59, 0.61]$.

The treatment of $a \in (0.59, 0.61]$ follows almost the exact same steps as the earlier case $a \in (0.57, 0.59]$. Applying the Buchstab identity twice with $z_1 = x^{0.07}$, we have

$$\psi(n, X) = \psi(n, x^{0.07}) \; - \sum_{\substack{n=p_1 m \\ x^{0.07} \le p_1 < X}} \psi(m, x^{0.07}) \tag{A.1}$$

$$+ \sum_{\substack{n=p_1 p_2 m \\ x^{0.07} \le p_2 < p_1 < X \\ \ell_1^* > 0.435}} \psi(m, p_2) + \sum_{\substack{n=p_1 p_2 m \\ x^{0.07} \le p_2 < p_1 < X \\ 0.42 < \ell_1^* \le 0.435}} \psi(m, p_2) + \sum_{\substack{n=p_1 p_2 m \\ x^{0.07} \le p_2 < p_1 < X \\ 0.365 < \ell_1^* \le 0.42}} \psi(m, p_2) \tag{A.2}$$

$$+ \sum_{\substack{n=p_1 p_2 m \\ x^{0.07} \le p_2 < p_1 < X \\ 0.33 < \ell_1^* \le 0.365}} \psi(m, p_2) + \sum_{\substack{n=p_1 p_2 m \\ x^{0.07} \le p_2 < p_1 < X \\ 0.305 < \ell_1^* \le 0.33}} \psi(m, p_2) + \sum_{\substack{n=p_1 p_2 m \\ x^{0.07} \le p_2 < p_1 < X \\ \ell_1^* \le 0.305}} \psi(m, p_2). \tag{A.3}$$

By condition (i) of Corollary 2.34, $\psi(n, x^{0.07})$ and $-\sum_{n=p_1 m,\, 0.07 \le \ell_1^* \le 0.5+\varepsilon_1} \psi(m, x^{0.07})$ in (A.1) are good* functions. Denote the first sum in (A.2) by $\Upsilon_1(n)$ and the second sum by $\Upsilon_2(n)$. The third sum in (A.2) is a good** function by condition (a) of Corollary 2.34. Denote the first sum in (A.3) by $\Upsilon_3(n)$, the second by $\Upsilon_4(n)$ and the third by $\Upsilon_5(n)$. We now treat $\Upsilon_1(n)$, $\Upsilon_2(n)$, $\Upsilon_3(n)$, $\Upsilon_4(n)$ and $\Upsilon_5(n)$ further, beginning with $\Upsilon_1(n)$.

Step 1: Treatment of $\Upsilon_1(n)$. Applying two more Buchstab iterations to parts of $\Upsilon_1(n)$, condition (ii) of Corollary 2.34 tells us that

$$\Upsilon_1(n) = \Psi_1(n) \;+ \sum_{\substack{n=p_1p_2m \\ x^{0.07}\le p_2<p_1<X \\ \ell_1^*>0.435 \\ \ell_2^*>0.635-\ell_1^*}} \psi(m,p_2) \;+ \sum_{\substack{n=p_1p_2m \\ x^{0.07}\le p_2<p_1<X \\ \ell_1^*>0.435 \\ \ell_2^*\in[0.58-\ell_1^*,0.635-\ell_1^*]}} \psi(m,p_2) \tag{A.4}$$

$$+ \sum_{\substack{n=p_1p_2m \\ x^{0.07}\le p_2<p_1<X \\ \ell_1^*>0.435 \\ \ell_2^*<0.58-\ell_1^* \\ \ell_2^*\ge\min\{0.105,(0.67-\ell_1^*)/2\}}} \psi(m,p_2) \;+ \sum_{\substack{n=p_1p_2p_3p_4m \\ x^{0.07}\le p_4<p_3<p_2<p_1<X \\ \ell_1^*>0.435 \\ \ell_2^*<\min\{0.58-\ell_1^*,0.105,(0.67-\ell_1^*)/2\}}} \psi(m,p_4), \tag{A.5}$$

where $\Psi_1(n)$ is a sum of two good* functions. Denote the first big sum in (A.4) by $\Theta_1(n)$. The second big sum in (A.4) is good** by condition (a) of Corollary 2.34. Denote the first sum in (A.5) by $\Theta_2(n)$ and the second by $\Theta_3(n)$. Upper bounds for $\sum_{n\in\mathcal{B}(y)}\Theta_i(n)$, $i\in\{1,2,3\}$, are given below:

$$\frac{y(1+o(1))}{x^b\log x}\int_{0.435}^{0.5+\varepsilon_1}\int_{0.635-\alpha_1}^{(1-\alpha_1)/2}\frac{\omega\left(\frac{1-\alpha_1-\alpha_2}{\alpha_2}\right)}{\alpha_1\alpha_2^2}\mathrm{d}\alpha_2\mathrm{d}\alpha_1 < \frac{0.2182y}{x^b\log x},$$

$$\frac{y(1+o(1))}{x^b\log x}\int_{0.435}^{0.5+\varepsilon_1}\int_{\min\{0.105,(0.67-\alpha_1)/2\}}^{0.58-\alpha_1}\frac{\omega\left(\frac{1-\alpha_1-\alpha_2}{\alpha_2}\right)}{\alpha_1\alpha_2^2}\mathrm{d}\alpha_2\mathrm{d}\alpha_1 < \frac{0.0921y}{x^b\log x},$$

$$\frac{y(1+o(1))}{x^b\log x}\int_{0.435}^{0.5+\varepsilon_1}\int_{0.07}^{\min\{0.58-\alpha_1,0.105,\frac{(0.67-\alpha_1)}{2}\}}\int_{0.07}^{\alpha_2}\int_{0.07}^{\min\{\alpha_3,\frac{(1-\sum_{i=1}^3\alpha_i)}{2}\}}\frac{\omega\left(\frac{1-\sum_{i=1}^4\alpha_i}{\alpha_4}\right)}{\alpha_1\alpha_2\alpha_3\alpha_4^2}\mathrm{d}\underline{\alpha} < \frac{0.0083y}{x^b\log x}.$$

Step 2: Treatment of $\Upsilon_2(n)$. We apply two more Buchstab iterations to parts of $\Upsilon_2(n)$, one with $z_1 = x^{0.07}$ and one with $z_1 = x^{0.09}$. Noting that (iii) of Corollary 2.34 allows $\beta \le 0.09$, we find

$$\Upsilon_2(n) = \Psi_2(n) \;+ \sum_{\substack{n=p_1p_2m \\ x^{0.07}\le p_2<p_1<X \\ 0.42<\ell_1^*\le 0.435 \\ \ell_2^*\ge 0.67-\ell_1^*}} \psi(m,p_2) \;+ \sum_{\substack{n=p_1p_2m \\ x^{0.07}\le p_2<p_1<X \\ 0.42<\ell_1^*\le 0.435 \\ \ell_2^*\in(0.58-\ell_1^*,0.67-\ell_1^*)}} \psi(m,p_2) \tag{A.6}$$

$$+ \sum_{\substack{n=p_1p_2m \\ x^{0.07}\le p_2<p_1<X \\ 0.42<\ell_1^*\le 0.435 \\ \ell_2^*\in[(0.67-\ell_1^*)/2,0.58-\ell_1^*]}} \psi(m,p_2) \;+ \sum_{\substack{n=p_1p_2p_3p_4m \\ x^{0.07}\le p_4<p_3<p_2<p_1<X \\ 0.42<\ell_1^*\le 0.435 \\ \ell_2^*<(0.67-\ell_1^*)/2 \\ \ell_4^*\ge 0.09}} \psi(m,p_4), \tag{A.7}$$

where $\Psi_2(n)$ is a sum of three good* functions. Denote the first big sum in (A.6) by $\Theta_4(n)$. The second big sum in (A.6) is good** by condition (b) of Corollary 2.34.

Denote the first sum in (A.7) by $\Theta_5(n)$ and the second by $\Theta_6(n)$. Upper bounds for $\sum_{n\in\mathcal{B}(y)}\Theta_i(n)$, $i\in\{4,5,6\}$, are given below:

$$\frac{y(1+o(1))}{x^b\log x}\int_{0.42}^{0.435}\int_{0.67-\alpha_1}^{(1-\alpha_1)/2}\frac{\omega\left(\frac{1-\alpha_1-\alpha_2}{\alpha_2}\right)}{\alpha_1\alpha_2^2}\mathrm{d}\alpha_2\mathrm{d}\alpha_1<\frac{0.0189y}{x^b\log x},$$
$$\frac{y(1+o(1))}{x^b\log x}\int_{0.42}^{0.435}\int_{(0.67-\alpha_1)/2}^{0.58-\alpha_1}\frac{\omega\left(\frac{1-\alpha_1-\alpha_2}{\alpha_2}\right)}{\alpha_1\alpha_2^2}\mathrm{d}\alpha_2\mathrm{d}\alpha_1<\frac{0.0356y}{x^b\log x},$$
$$\frac{y(1+o(1))}{x^b\log x}\int_{0.42}^{0.435}\int_{0.09}^{(0.67-\alpha_1)/2}\int_{0.09}^{\alpha_2}\int_{0.09}^{\min\{\alpha_3,\frac{(1-\sum_{i=1}^{3}\alpha_i)}{2}\}}\frac{\omega\left(\frac{1-\sum_{i=1}^{4}\alpha_i}{\alpha_4}\right)}{\alpha_1\alpha_2\alpha_3\alpha_4^2}\mathrm{d}\underline{\alpha}<\frac{0.001y}{x^b\log x}.$$

Step 3: Treatment of $\Upsilon_3(n)$. By condition (iv) of Corollary 2.34,

$$\Upsilon_3(n)=\Psi_3(n)+\sum_{\substack{n=p_1p_2m\\ x^{0.07}\le p_2<p_1<X\\ 0.33<\ell_1^*\le 0.365\\ \ell_2^*>0.635-\ell_1^*}}\psi(m,p_2)+\sum_{\substack{n=p_1p_2m\\ x^{0.07}\le p_2<p_1<X\\ 0.33<\ell_1^*\le 0.365\\ \ell_2^*\in[0.365-\ell_1^*,0.435-\ell_1^*]\cup[0.565-\ell_1^*,0.635-\ell_1^*]}}\psi(m,p_2) \tag{A.8}$$
$$+\sum_{\substack{n=p_1p_2m\\ x^{0.07}\le p_2<p_1<X\\ 0.33<\ell_1^*\le 0.365\\ \ell_2^*\in[0.1524,0.565-\ell_1^*)}}\psi(m,p_2)+\sum_{\substack{n=p_1p_2p_3p_4m\\ x^{0.07}\le p_4<p_3<p_2<p_1<X\\ 0.33<\ell_1^*\le 0.365\\ \ell_2^*\in(0.435-\ell_1^*,0.1524)\\ \ell_4^*>0.435-\ell_1^*}}\psi(m,p_4)+\sum_{\substack{n=p_1p_2p_3p_4m\\ x^{0.07}\le p_4<p_3<p_2<p_1<X\\ 0.33<\ell_1^*\le 0.365\\ \ell_2^*\in(0.435-\ell_1^*,0.1524)\\ \ell_4^*\le 0.435-\ell_1^*}}\psi(m,p_4), \tag{A.9}$$

where $\Psi_3(n)$ is a sum of two good* functions. Denote the first big sum in (A.8) by $\Theta_7(n)$. The second big sum in (A.8) is good** by condition (c) of Corollary 2.34. Denote the first sum in (A.9) by $\Theta_8(n)$ and the second by $\Theta_9(n)$. The third sum is again good** by condition (c). We then have:

$$\sum_{n\in\mathcal{B}(y)}\Theta_7(n)\le\frac{y(1+o(1))}{x^b\log x}\int_{0.33}^{0.365}\int_{0.635-\alpha_1}^{\min\{\alpha_1,(1-\alpha_1)/2\}}\frac{\omega\left(\frac{1-\alpha_1-\alpha_2}{\alpha_2}\right)}{\alpha_1\alpha_2^2}\mathrm{d}\alpha_2\mathrm{d}\alpha_1<\frac{0.0367y}{x^b\log x},$$
$$\sum_{n\in\mathcal{B}(y)}\Theta_8(n)\le\frac{y(1+o(1))}{x^b\log x}\int_{0.33}^{0.365}\int_{0.1524}^{0.565-\alpha_1}\frac{\omega\left(\frac{1-\alpha_1-\alpha_2}{\alpha_2}\right)}{\alpha_1\alpha_2^2}\mathrm{d}\alpha_2\mathrm{d}\alpha_1<\frac{0.1186y}{x^b\log x},$$
$$\sum_{n\in\mathcal{B}(y)}\Theta_9(n)\le\frac{y(1+o(1))}{x^b\log x}\int_{0.33}^{0.365}\int_{0.435-\alpha_1}^{0.1524}\int_{0.435-\alpha_1}^{\alpha_2}\int_{0.435-\alpha_1}^{\min\{\alpha_3,\frac{(1-\sum_{i=1}^{3}\alpha_i)}{2}\}}\frac{\omega\left(\frac{1-\sum_{i=1}^{4}\alpha_i}{\alpha_4}\right)}{\alpha_1\alpha_2\alpha_3\alpha_4^2}\mathrm{d}\underline{\alpha}<\frac{0.0211y}{x^b\log x}.$$

Step 4: Treatment of $\Upsilon_4(n)$. By condition (v) of Corollary 2.34,

$$\Upsilon_4(n)=\Psi_4(n)+\sum_{\substack{n=p_1p_2m\\ x^{0.07}\le p_2<p_1<X\\ 0.305<\ell_1^*\le 0.33\\ \ell_2^*>0.635-\ell_1^*}}\psi(m,p_2)+\sum_{\substack{n=p_1p_2m\\ x^{0.07}\le p_2<p_1<X\\ 0.305<\ell_1^*\le 0.33\\ \ell_2^*\in(0.58-\ell_1^*,0.635-\ell_1^*]}}\psi(m,p_2) \tag{A.10}$$

$$+\sum_{\substack{n=p_1p_2m\\ x^{0.07}\le p_2<p_1<X\\ 0.305<\ell_1^*\le 0.33\\ \ell_2^*\in[(0.635-\ell_1^*)/2,0.58-\ell_1^*]}}\psi(m,p_2)+\sum_{\substack{n=p_1p_2p_3p_4m\\ x^{0.07}\le p_4<p_3<p_2<p_1<X\\ 0.305<\ell_1^*\le 0.33\\ \ell_2^*<(0.635-\ell_1^*)/2\\ \ell_4^*>0.42-\ell_1^*}}\psi(m,p_4)+\sum_{\substack{n=p_1p_2p_3p_4m\\ x^{0.07}\le p_4<p_3<p_2<p_1<X\\ 0.305<\ell_1^*\le 0.33\\ \ell_2^*<(0.635-\ell_1^*)/2\\ \ell_4^*\le 0.42-\ell_1^*}}\psi(m,p_4), \quad (A.11)$$

where $\Psi_4(n)$ is a sum of two good* functions. Denote the first big sum in (A.10) by $\Theta_{10}(n)$. The second big sum in (A.10) is good** by condition (a) of Corollary 2.34. Denote the first sum in (A.11) by $\Theta_{11}(n)$ and the second by $\Theta_{12}(n)$. The third sum is again good** by condition (a). We then have:

$$\sum_{n\in\mathcal{B}(y)}\Theta_{10}(n)\le\frac{y(1+o(1))}{x^b\log x}\int_{0.305}^{0.33}\int_{0.635-\alpha_1}^{\alpha_1}\frac{\omega\left(\frac{1-\alpha_1-\alpha_2}{\alpha_2}\right)}{\alpha_1\alpha_2^2}\mathrm{d}\alpha_2\mathrm{d}\alpha_1<\frac{0.0043y}{x^b\log x},$$

$$\sum_{n\in\mathcal{B}(y)}\Theta_{11}(n)\le\frac{y(1+o(1))}{x^b\log x}\int_{0.305}^{0.33}\int_{(0.635-\alpha_1)/2}^{0.58-\alpha_1}\frac{\omega\left(\frac{1-\alpha_1-\alpha_2}{\alpha_2}\right)}{\alpha_1\alpha_2^2}\mathrm{d}\alpha_2\mathrm{d}\alpha_1<\frac{0.1178y}{x^b\log x},$$

$$\sum_{n\in\mathcal{B}(y)}\Theta_{12}(n)\le\frac{y(1+o(1))}{x^b\log x}\int_{0.305}^{0.33}\int_{0.42-\alpha_1}^{\frac{(0.635-\alpha_1)}{2}}\int_{0.42-\alpha_1}^{\alpha_2}\int_{0.42-\alpha_1}^{\min\{\alpha_3,\frac{(1-\sum_{i=1}^3\alpha_i)}{2}\}}\frac{\omega\left(\frac{1-\sum_{i=1}^4\alpha_i}{\alpha_4}\right)}{\alpha_1\alpha_2\alpha_3\alpha_4^2}\mathrm{d}\underline{\alpha}<\frac{0.0062y}{x^b\log x}.$$

Step 5: Treatment of $\Upsilon_5(n)$. Using conditions (vi) and (vii) of Corollary 2.34, we find that $\Upsilon_5(n)$ equals

$$\Psi_5(n)+\sum_{\substack{n=p_1p_2m\\ x^{0.07}\le p_2<p_1<X\\ \ell_1^*\le 0.305\\ \ell_2^*\in[0.365-\ell_1^*,0.42-\ell_1^*)\cup(0.58-\ell_1^*,0.635-\ell_1^*]}}\psi(m,p_2)+\sum_{\substack{n=p_1p_2m\\ x^{0.07}\le p_2<p_1<X\\ \ell_1^*\le 0.305\\ \ell_2^*\in(0.2099,0.58-\ell_1^*]}}\psi(m,p_2) \quad (A.12)$$

$$+\sum_{\substack{n=p_1p_2p_3p_4m\\ x^{0.07}\le p_4<p_3<p_2<p_1<X\\ \ell_1^*\le 0.305\\ \ell_2^*\in[0.42-\ell_1^*,0.2099]\\ \ell_3^*>0.42-\ell_1^*\\ \ell_4^*>0.42-\ell_1^*}}\psi(m,p_4)+\sum_{\substack{n=p_1p_2p_3p_4m\\ x^{0.07}\le p_4<p_3<p_2<p_1<X\\ \ell_1^*\le 0.305\\ \ell_2^*\in[0.42-\ell_1^*,0.2099]\\ \ell_3^*>0.42-\ell_1^*\\ \ell_4^*\in[0.365-\ell_1^*,0.42-\ell_1^*]}}\psi(m,p_4)+\sum_{\substack{n=p_1p_2p_3p_4m\\ x^{0.07}\le p_4<p_3<p_2<p_1<X\\ \ell_1^*\le 0.305\\ \ell_2^*\in[0.42-\ell_1^*,0.2099]\\ \ell_3^*>0.42-\ell_1^*\\ \ell_4^*<0.365-\ell_1^*}}\psi(m,p_4) \quad (A.13)$$

$$+\sum_{\substack{n=p_1p_2p_3p_4m\\ x^{0.07}\le p_4<p_3<p_2<p_1<X\\ \ell_1^*\le 0.305\\ \ell_2^*\in[0.42-\ell_1^*,0.2099]\\ \ell_3^*\in[0.365-\ell_1^*,0.42-\ell_1^*]}}\psi(m,p_4)+\sum_{\substack{n=p_1p_2p_3p_4m\\ x^{0.07}\le p_4<p_3<p_2<p_1<X\\ \ell_1^*\le 0.305\\ \ell_2^*\in[0.42-\ell_1^*,0.2099]\\ \ell_3^*<0.365-\ell_1^*}}\psi(m,p_4) \quad (A.14)$$

$$+\sum_{\substack{n=p_1p_2p_3p_4m\\ x^{0.07}\le p_4<p_3<p_2<p_1<X\\ \ell_1^*\le 0.305\\ \ell_2^*<0.365-\ell_1^*\\ \ell_4^*>0.42-\ell_2^*-\ell_3^*\\ \ell_4^*>(0.635-\ell_1^*-\ell_2^*-\ell_3^*)/2}}\psi(m,p_4)+\sum_{\substack{n=p_1p_2p_3p_4m\\ x^{0.07}\le p_4<p_3<p_2<p_1<X\\ \ell_1^*\le 0.305\\ \ell_2^*<0.365-\ell_1^*\\ \ell_4^*\in[0.365-\ell_2^*-\ell_3^*,0.42-\ell_2^*-\ell_3^*]\\ \ell_4^*>(0.635-\ell_1^*-\ell_2^*-\ell_3^*)/2}}\psi(m,p_4) \quad (A.15)$$

$$+ \sum_{\substack{n=p_1p_2p_3p_4m \\ x^{0.07}\le p_4<p_3<p_2<p_1<X \\ \ell_1^*\le 0.305 \\ \ell_2^*<0.365-\ell_1^* \\ \ell_4^*<0.365-\ell_2^*-\ell_3^* \\ \ell_4^*>(0.635-\ell_1^*-\ell_2^*-\ell_3^*)/2}} \psi(m,p_4) + \sum_{\substack{n=p_1p_2p_3p_4p_5p_6m \\ x^{0.07}\le p_6<p_5<p_4<p_3<p_2<p_1<X \\ \ell_1^*\le 0.305 \\ \ell_2^*<0.365-\ell_1^* \\ \ell_4^*\le(0.635-\ell_1^*-\ell_2^*-\ell_3^*)/2}} \psi(m,p_6), \tag{A.16}$$

where $\Psi_5(n)$ is a sum of six good* functions. The first sum in (A.12), the second sum in (A.13), the first sum in (A.14) and the second sum in (A.15) are good** by condition (a) of Corollary 2.34. Enumerate the remaining seven big sums as $\Theta_{13}(n), \ldots, \Theta_{19}(n)$. Upper bounds for $\sum_{n\in\mathcal{B}(y)} \Theta_i(n)$, $i \in \{13, \ldots, 19\}$, are given below:

$$\frac{y(1+o(1))}{x^b \log x} \int_{0.2099}^{0.305} \int_{0.2099}^{\min\{\alpha_1, 0.58-\alpha_1\}} \frac{\omega\left(\frac{1-\alpha_1-\alpha_2}{\alpha_2}\right)}{\alpha_1\alpha_2^2} \mathrm{d}\alpha_2 \mathrm{d}\alpha_1 < \frac{0.1723y}{x^b \log x},$$

$$\frac{y(1+o(1))}{x^b \log x} \int_{0.21}^{0.305} \int_{0.42-\alpha_1}^{0.2099} \int_{0.42-\alpha_1}^{\alpha_2} \int_{0.42-\alpha_1}^{\min\{\alpha_3, \frac{(1-\sum_{i=1}^3 \alpha_i)}{2}\}} \frac{\omega\left(\frac{1-\sum_{i=1}^4 \alpha_i}{\alpha_4}\right)}{\alpha_1\alpha_2\alpha_3\alpha_4^2} \mathrm{d}\underline{\alpha} < \frac{0.0104y}{x^b \log x},$$

$$\frac{y(1+o(1))}{x^b \log x} \int_{0.21}^{0.305} \int_{0.42-\alpha_1}^{0.2099} \int_{0.42-\alpha_1}^{\alpha_2} \int_{0.07}^{\min\{\alpha_3, \frac{(1-\sum_{i=1}^3 \alpha_i)}{2}, 0.365-\alpha_1\}} \frac{\omega\left(\frac{1-\sum_{i=1}^4 \alpha_i}{\alpha_4}\right)}{\alpha_1\alpha_2\alpha_3\alpha_4^2} \mathrm{d}\underline{\alpha} < \frac{0.0212y}{x^b \log x},$$

$$\frac{y(1+o(1))}{x^b \log x} \int_{0.21}^{0.305} \int_{0.42-\alpha_1}^{0.2099} \int_{0.07}^{\min\{\alpha_2, 0.365-\alpha_1\}} \int_{0.07}^{\min\{\alpha_3, 0.365-\alpha_1\}} \frac{\omega\left(\frac{1-\sum_{i=1}^4 \alpha_i}{\alpha_4}\right)}{\alpha_1\alpha_2\alpha_3\alpha_4^2} \mathrm{d}\underline{\alpha} < \frac{0.0397y}{x^b \log x},$$

$$\frac{y(1+o(1))}{x^b \log x} \int_{0.07}^{0.295} \int_{0.07}^{\min\{\alpha_1, 0.365-\alpha_1\}} \int_{0.07}^{\alpha_2} \int_{\max\{\frac{(0.635-\sum_{i=1}^3 \alpha_i)}{2}, 0.42-\alpha_2-\alpha_3\}}^{\min\{\alpha_3, \frac{(1-\sum_{i=1}^3 \alpha_i)}{2}\}} \frac{\omega\left(\frac{1-\sum_{i=1}^4 \alpha_i}{\alpha_4}\right)}{\alpha_1\alpha_2\alpha_3\alpha_4^2} \mathrm{d}\underline{\alpha} < \frac{0.0105y}{x^b \log x},$$

$$\frac{y(1+o(1))}{x^b \log x} \int_{0.07}^{0.295} \int_{0.07}^{\min\{\alpha_1, 0.365-\alpha_1\}} \int_{0.07}^{\alpha_2} \int_{\max\{0.07, \frac{(0.635-\sum_{i=1}^3 \alpha_i)}{2}\}}^{\min\{\alpha_3, 0.365-\alpha_2-\alpha_3\}} \frac{\omega\left(\frac{1-\sum_{i=1}^4 \alpha_i}{\alpha_4}\right)}{\alpha_1\alpha_2\alpha_3\alpha_4^2} \mathrm{d}\underline{\alpha} < \frac{0.023y}{x^b \log x},$$

$$\frac{y(1+o(1))}{x^b \log x} \left(\frac{\log\left(\frac{0.2}{0.07}\right)^6}{0.07 \cdot 6!} + \frac{\log\left(\frac{0.165}{0.07}\right)^5 \log\left(\frac{0.23}{0.2}\right)}{0.07 \cdot 5!} + \frac{\log\left(\frac{0.135}{0.07}\right)^5 \log\left(\frac{0.295}{0.23}\right)}{0.07 \cdot 5!} \right) < \frac{0.0379y}{x^b \log x}.$$

In summary, we have shown that $1_{\mathbb{P}}(n)$ can be written as $1_{\mathbb{P}}(n) = \Psi(n) + \sum_{i=1}^{19} \Theta_i(n)$, where $\Psi(n)$ is a sum of no more than 100 good* and good** functions and where each Θ_i is non-negative and

$$\sum_{i=1}^{19} \sum_{n\in\mathcal{B}(y)} \Theta_i(n) < \frac{0.9937y}{x^b \log x}.$$

This concludes the proof of Proposition 2.3 for the case $a \in (0.59, 0.61]$.

<u>Case 4</u>: $a \le 0.53$.

For the remaining cases, $a \le 0.53$, $a \in (0.545, 0.57]$ and $a > 0.61$, option (2) and (3) of Definition 2.4 were not used in the search for elements of $\mathcal{R}^\star(a)$ and $\mathcal{R}^{\star\star}(a)$. The corresponding Buchstab decompositions are also considerably simpler. Thus we now skip past the detailed step-by-step derivation given in previous cases and immediately write down the final decomposition. We begin with case $a \leq 0.53$. By conditions (i), (ii) and (iii) of Corollary 2.30, Buchstab iterations give

$$\psi(n,X) = \Psi_1(n) + \sum_{\substack{n=p_1p_2m \\ x^{0.07}\leq p_2<p_1<X \\ \ell_1^*>0.36 \\ \ell_2^*>(0.71-\ell_1^*)/2 \\ \ell_2^*\notin[0.64-\ell_1^*,0.71-\ell_1^*]}} \psi(m,p_2) + \sum_{\substack{n=p_1p_2m \\ x^{0.07}\leq p_2<p_1<X \\ \ell_1^*>0.36 \\ \ell_2^*>(0.71-\ell_1^*)/2 \\ \ell_2^*\in[0.64-\ell_1^*,0.71-\ell_1^*]}} \psi(m,p_2) \tag{A.17}$$

$$+ \sum_{\substack{n=p_1p_2p_3p_4m \\ x^{0.07}\leq p_4<p_3<p_2<p_1<X \\ \ell_1^*>0.36 \\ \ell_2^*\leq(0.71-\ell_1^*)/2 \\ \ell_3^*<0.64-\ell_1^*-\ell_2^*}} \psi(m,p_4) + \sum_{\substack{n=p_1p_2p_3p_4m \\ x^{0.07}\leq p_4<p_3<p_2<p_1<X \\ \ell_1^*>0.36 \\ \ell_2^*\leq(0.71-\ell_1^*)/2 \\ \ell_3^*\geq0.64-\ell_1^*-\ell_2^*}} \psi(m,p_4) \tag{A.18}$$

$$+ \sum_{\substack{n=p_1p_2m \\ x^{0.07}\leq p_2<p_1<X \\ \ell_1^*\in[0.29,0.36]}} \psi(m,p_2) + \sum_{\substack{n=p_1p_2m \\ x^{0.07}\leq p_2<p_1<X \\ \ell_1^*<0.29 \\ \ell_2^*>(0.71-\ell_1^*)/2}} \psi(m,p_2) + \sum_{\substack{n=p_1p_2p_3p_4m \\ x^{0.07}\leq p_4<p_3<p_2<p_1<X \\ 0.22<\ell_1^*<0.29 \\ \ell_2^*\leq(0.71-\ell_1^*)/2 \\ \ell_4^*>(0.71-\ell_1^*-\ell_2^*-\ell_3^*)/2 \\ \ell_4^*>0.36-\ell_1^*}} \psi(m,p_4) \tag{A.19}$$

$$+ \sum_{\substack{n=p_1p_2p_3p_4m \\ x^{0.07}\leq p_4<p_3<p_2<p_1<X \\ 0.22<\ell_1^*<0.29 \\ \ell_2^*\leq(0.71-\ell_1^*)/2 \\ \ell_4^*>(0.71-\ell_1^*-\ell_2^*-\ell_3^*)/2 \\ \ell_4^*\leq0.36-\ell_1^*}} \psi(m,p_4) + \sum_{\substack{n=p_1p_2p_3p_4m \\ x^{0.07}\leq p_4<p_3<p_2<p_1<X \\ \ell_1^*\leq0.22 \\ \ell_2^*\leq(0.71-\ell_1^*)/2 \\ \ell_4^*>(0.71-\ell_1^*-\ell_2^*-\ell_3^*)/2}} \psi(m,p_4) \tag{A.20}$$

$$+ \sum_{\substack{n=p_1p_2p_3p_4p_5p_6m \\ x^{0.07}\leq p_6<p_5<p_4<p_3<p_2<p_1<X \\ \ell_1^*<0.29 \\ \ell_2^*\leq(0.71-\ell_1^*)/2 \\ \ell_4^*\leq(0.71-\ell_1^*-\ell_2^*-\ell_3^*)/2 \\ \mathcal{C}_1}} \psi(m,p_6) + \sum_{\substack{n=p_1p_2p_3p_4p_5p_6m \\ x^{0.07}\leq p_6<p_5<p_4<p_3<p_2<p_1<X \\ \ell_1^*<0.29 \\ \ell_2^*\leq(0.71-\ell_1^*)/2 \\ \ell_4^*\leq(0.71-\ell_1^*-\ell_2^*-\ell_3^*)/2 \\ \mathcal{C}_2}} \psi(m,p_6) \tag{A.21}$$

$$+ \sum_{\substack{n=p_1p_2p_3p_4p_5p_6m \\ x^{0.07}\leq p_6<p_5<p_4<p_3<p_2<p_1<X \\ \ell_1^*<0.29 \\ \ell_2^*\leq(0.71-\ell_1^*)/2 \\ \ell_4^*\leq(0.71-\ell_1^*-\ell_2^*-\ell_3^*)/2 \\ \mathcal{C}_3}} \psi(m,p_6) + \sum_{\substack{n=p_1p_2p_3p_4p_5p_6m \\ x^{0.07}\leq p_6<p_5<p_4<p_3<p_2<p_1<X \\ \ell_1^*<0.29 \\ \ell_2^*\leq(0.71-\ell_1^*)/2 \\ \ell_4^*\leq(0.71-\ell_1^*-\ell_2^*-\ell_3^*)/2 \\ \mathcal{C}_4}} \psi(m,p_6), \tag{A.22}$$

where $\Psi_1(n)$ is a sum of eight good* functions and where $\mathcal{C}_1$ denotes the condition ($\ell_i^*, \ell_j^* \in [0.145, 0.18]$ for some $i \neq j$), $\mathcal{C}_2$ denotes the condition ($\ell_3^* \in [0.145, 0.29]$ and $\ell_i^* \in [0.145, 0.18]$ for at most one i), $\mathcal{C}_3$ denotes the condition ($\ell_1^* > 0.18$, $\ell_2^* \in [0.145, 0.29]$ and $\ell_3^* < 0.145$) and $\mathcal{C}_4$ denotes the condition ($\ell_2^* < 0.145$). Conditions $\mathcal{C}_1$, $\mathcal{C}_2$, $\mathcal{C}_3$ and $\mathcal{C}_4$ partition the set of $(\ell_1^*, \ell_2^*, \ell_3^*, \ell_4^*, \ell_5^*, \ell_6^*)$ which satisfy $0.07 \leq \ell_6^* < \ell_5^* < \ell_4^* < \ell_3^* < \ell_2^* < \ell_1^* \leq 0.29$, $\ell_2^* \leq (0.71 - \alpha_1)/2$ and $\ell_4^* \leq (0.71 - \alpha_1 - \alpha_2 - \alpha_3)/2$.

Denote the first big sum in (A.17) by $\Theta_1(n)$. The second sum in (A.17) is good** by condition (b) of Corollary 2.30. Denote the first sum in (A.18) by $\Theta_2(n)$. The second sum in (A.18) is good** by condition (c) of Corollary 2.30. The first sum in (A.19) is good** by condition (a) of Corollary 2.30. Denote the second sum by $\Theta_3(n)$ and the third sum by $\Theta_4(n)$. The first sum in (A.20) is good** by condition (d) of Corollary 2.30. Denote the second sum by $\Theta_5(n)$. Since $\ell_i^*, \ell_j^* \in [0.145, 0.18]$ implies $\ell_i^* + \ell_j^* \in [0.29, 0.36]$, the first sum in (A.21) is good** by condition (e) of Corollary 2.30. By condition (f), the second sum is also good. Denote the first sum in (A.22) by $\Theta_6(n)$ and the second sum by $\Theta_7(n)$.

Upper bounds for $\sum_{n\in\mathcal{B}(y)} \Theta_i(n)$, $i \in \{1, \ldots, 7\}$, are given below:

$$\frac{y(1+o(1))}{x^b \log x} \int_{0.36}^{0.5+\varepsilon_1} \int_{[(0.71-\alpha_1)/2, 0.64-\alpha_1] \cup [0.71-\alpha_1, (1-\alpha_1)/2]} \frac{\omega\left(\frac{1-\alpha_1-\alpha_2}{\alpha_2}\right)}{\alpha_1\alpha_2^2} \mathrm{d}\alpha_2 \mathrm{d}\alpha_1 < \frac{0.513y}{x^b \log x},$$

$$\frac{y(1+o(1))}{x^b \log x} \int_{0.36}^{0.5+\varepsilon_1} \int_{0.07}^{\min\{\alpha_1, (0.71-\alpha_1)/2\}} \int_{0.07}^{\min\{\alpha_2, 0.64-\alpha_1-\alpha_2\}} \int_{0.07}^{\alpha_3} \frac{\omega\left(\frac{1-\sum_{i=1}^4 \alpha_i}{\alpha_4}\right)}{\alpha_1\alpha_2\alpha_3\alpha_4^2} \mathrm{d}\underline{\alpha} < \frac{0.079y}{x^b \log x},$$

$$\frac{y(1+o(1))}{x^b \log x} \int_{0.07}^{0.29} \int_{(0.71-\alpha_1)/2}^{\alpha_1} \frac{\omega\left(\frac{1-\alpha_1-\alpha_2}{\alpha_2}\right)}{\alpha_1\alpha_2^2} \mathrm{d}\alpha_2 \mathrm{d}\alpha_1 < \frac{0.08y}{x^b \log x},$$

$$\frac{y(1+o(1))}{x^b \log x} \int_{0.22}^{0.29} \int_{0.07}^{\min\{\alpha_1, \frac{(0.71-\alpha_1)}{2}\}} \int_{0.07}^{\alpha_2} \int_{\max\{0.36-\alpha_1, \frac{(0.71-\alpha_1-\alpha_2-\alpha_3)}{2}\}}^{\alpha_3} \frac{\omega\left(\frac{1-\sum_{i=1}^4 \alpha_i}{\alpha_4}\right)}{\alpha_1\alpha_2\alpha_3\alpha_4^2} \mathrm{d}\underline{\alpha} < \frac{0.112y}{x^b \log x},$$

$$\frac{y(1+o(1))}{x^b \log x} \int_{0.07}^{0.22} \int_{0.07}^{\min\{\alpha_1, \frac{(0.71-\alpha_1)}{2}\}} \int_{0.07}^{\alpha_2} \int_{\max\{0.07, \frac{(0.71-\alpha_1-\alpha_2-\alpha_3)}{2}\}}^{\alpha_3} \frac{\omega\left(\frac{1-\sum_{i=1}^4 \alpha_i}{\alpha_4}\right)}{\alpha_1\alpha_2\alpha_3\alpha_4^2} \mathrm{d}\underline{\alpha} < \frac{0.063y}{x^b \log x},$$

$$\frac{y(1+o(1))}{x^b \log x} \frac{1}{4!} \int_{0.18}^{0.29} \int_{0.145}^{0.29} \int_{0.07}^{0.145} \int_{0.07}^{0.145} \int_{0.07}^{0.145} \int_{0.07}^{0.145} \frac{\omega\left(\frac{1-\alpha_1-\alpha_2-\alpha_3-\alpha_4-\alpha_5-\alpha_6}{\alpha_6}\right)}{0.07\alpha_1\alpha_2\alpha_3\alpha_4\alpha_5\alpha_6} \mathrm{d}\underline{\alpha} < \frac{0.056y}{x^b \log x},$$

$$\frac{y(1+o(1))}{x^b \log x} \frac{1}{5!} \int_{0.07}^{0.29} \int_{0.07}^{0.145} \int_{0.07}^{0.145} \int_{0.07}^{0.145} \int_{0.07}^{0.145} \int_{0.07}^{0.145} \frac{\omega\left(\frac{1-\alpha_1-\alpha_2-\alpha_3-\alpha_4-\alpha_5-\alpha_6}{\alpha_6}\right)}{0.07\alpha_1\alpha_2\alpha_3\alpha_4\alpha_5\alpha_6} \mathrm{d}\underline{\alpha} < \frac{0.035y}{x^b \log x}.$$

In summary, we have shown that $1_{\mathbb{P}}(n)$ can be written as $1_{\mathbb{P}}(n) = \Psi(n) + \sum_{i=1}^7 \Theta_i(n)$, where $\Psi(n)$ is a sum of no more than 100 good* and good** functions and where each Θ_i is non-negative and

$$\sum_{i=1}^{7} \sum_{n\in\mathcal{B}(y)} \Theta_i(n) < \frac{0.938y}{x^b \log x}.$$

This concludes the proof of Proposition 2.3 for the case $a \leq 0.53$.

<u>Case 5</u>: $a \in (0.545, 0.57]$.

Next we cover the case $a \in (0.545, 0.57]$. By conditions (i), (ii), (iii) and (iv) of Corollary 2.32, Buchstab iterations give

$$\psi(n, X) = \Psi_1(n) + \sum_{\substack{n=p_1p_2m \\ x^{0.075}\leq p_2<p_1<X \\ \ell_1^*>0.475 \\ \ell_2^*>0.6-\ell_1^*}} \psi(m, p_2) + \sum_{\substack{n=p_1p_2m \\ x^{0.075}\leq p_2<p_1<X \\ \ell_1^*>0.475 \\ \ell_2^*\leq 0.6-\ell_1^*}} \psi(m, p_2) + \sum_{\substack{n=p_1p_2m \\ x^{0.075}\leq p_2<p_1<X \\ 0.4\leq\ell_1^*\leq 0.475}} \psi(m, p_2) \quad \text{(A.23)}$$

$$+\sum_{\substack{n=p_1p_2m\\ x^{0.075}\le p_2<p_1<X\\ \ell_1^*<0.4\\ \ell_2^*>0.6-\ell_1^*}}\psi(m,p_2)+\sum_{\substack{n=p_1p_2m\\ x^{0.075}\le p_2<p_1<X\\ \ell_1^*<0.4\\ \ell_2^*\in[0.4-\ell_1^*,0.475-\ell_1^*]\cup[0.525-\ell_1^*,0.6-\ell_1^*]}}\psi(m,p_2) \tag{A.24}$$

$$+\sum_{\substack{n=p_1p_2m\\ x^{0.075}\le p_2<p_1<X\\ \ell_1^*<0.4\\ \ell_2^*\in(0.475-\ell_1^*,0.525-\ell_1^*)\\ \ell_2^*>0.14}}\psi(m,p_2)+\sum_{\substack{n=p_1p_2p_3p_4m\\ x^{0.075}\le p_4<p_3<p_2<p_1<X\\ \ell_1^*<0.4\\ \ell_2^*\in(0.475-\ell_1^*,0.525-\ell_1^*)\\ \ell_2^*\le 0.14}}\psi(m,p_4) \tag{A.25}$$

$$+\sum_{\substack{n=p_1p_2p_3p_4m\\ x^{0.075}\le p_4<p_3<p_2<p_1<X\\ \ell_2^*<0.4-\ell_1^*\\ \ell_4^*>(0.615-\ell_1^*-\ell_2^*-\ell_3^*)/2\\ \ell_4^*>0.475-\ell_1^*-\ell_2^*}}\psi(m,p_4)+\sum_{\substack{n=p_1p_2p_3p_4m\\ x^{0.075}\le p_4<p_3<p_2<p_1<X\\ \ell_2^*<0.4-\ell_1^*\\ \ell_4^*>(0.615-\ell_1^*-\ell_2^*-\ell_3^*)/2\\ \ell_4^*\in[0.4-\ell_1^*-\ell_2^*,0.475-\ell_1^*-\ell_2^*]}}\psi(m,p_4) \tag{A.26}$$

$$+\sum_{\substack{n=p_1p_2p_3p_4m\\ x^{0.075}\le p_4<p_3<p_2<p_1<X\\ \ell_2^*<0.4-\ell_1^*\\ \ell_4^*>(0.615-\ell_1^*-\ell_2^*-\ell_3^*)/2\\ \ell_4^*<0.4-\ell_1^*-\ell_2^*}}\psi(m,p_4)+\sum_{\substack{n=p_1p_2p_3p_4p_5p_6m\\ x^{0.075}\le p_6<p_5<p_4<p_3<p_2<p_1<X\\ \ell_2^*<0.4-\ell_1^*\\ \ell_4^*\le(0.615-\ell_1^*-\ell_2^*-\ell_3^*)/2}}\psi(m,p_6), \tag{A.27}$$

where $\Psi_1(n)$ is a sum of eight good* functions. Denote the first big sum in (A.23) by $\Theta_1(n)$. The second big sum is good** by condition (b) of Corollary 2.32 and the third is good** by condition (a). Denote the first sum in (A.24) by $\Theta_2(n)$. The second sum is again good** by condition (b). Denote the first sum in (A.25) by $\Theta_3(n)$ and the second by $\Theta_4(n)$. Denote the first sum in (A.26) by $\Theta_5(n)$. The second sum is good** by condition (c). Denote the first sum in (A.27) by $\Theta_6(n)$ and the second by $\Theta_7(n)$.

Upper bounds for $\sum_{n\in\mathcal{B}(y)}\Theta_i(n)$, $i\in\{1,\dots,7\}$, are given below:

$$\frac{y(1+o(1))}{x^b\log x}\int_{0.475}^{0.5+\varepsilon_1}\int_{0.6-\alpha_1}^{(1-\alpha_1)/2}\frac{\omega\left(\frac{1-\alpha_1-\alpha_2}{\alpha_2}\right)}{\alpha_1\alpha_2^2}\mathrm{d}\alpha_2\mathrm{d}\alpha_1<\frac{0.166y}{x^b\log x},$$

$$\frac{y(1+o(1))}{x^b\log x}\int_{0.3}^{0.4}\int_{0.6-\alpha_1}^{\min\{\alpha_1,(1-\alpha_1)/2\}}\frac{\omega\left(\frac{1-\alpha_1-\alpha_2}{\alpha_2}\right)}{\alpha_1\alpha_2^2}\mathrm{d}\alpha_2\mathrm{d}\alpha_1<\frac{0.187y}{x^b\log x},$$

$$\frac{y(1+o(1))}{x^b\log x}\int_{0.475/2}^{0.385}\int_{\max\{0.14,0.475-\alpha_1\}}^{\min\{\alpha_1,0.525-\alpha_1\}}\frac{\omega\left(\frac{1-\alpha_1-\alpha_2}{\alpha_2}\right)}{\alpha_1\alpha_2^2}\mathrm{d}\alpha_2\mathrm{d}\alpha_1<\frac{0.302y}{x^b\log x},$$

$$\frac{y(1+o(1))}{x^b\log x}\int_{0.335}^{0.4}\int_{0.475-\alpha_1}^{\min\{0.14,0.525-\alpha_1\}}\int_{0.075}^{\alpha_2}\int_{0.075}^{\min\{\alpha_3,\frac{(1-\sum_{i=1}^3\alpha_i)}{2}\}}\frac{\omega\left(\frac{1-\sum_{i=1}^4\alpha_i}{\alpha_4}\right)}{\alpha_1\alpha_2\alpha_3\alpha_4^2}\mathrm{d}\underline{\alpha}<\frac{0.032y}{x^b\log x},$$

$$\frac{y(1+o(1))}{x^b\log x}\int_{0.075}^{0.325}\int_{0.075}^{\min\{\alpha_1,0.4-\alpha_1\}}\int_{0.075}^{\alpha_2}\int_{\max\{\frac{(0.615-\sum_{i=1}^3\alpha_i)}{2},0.475-\alpha_1-\alpha_2\}}^{\min\{\alpha_3,\frac{(1-\sum_{i=1}^3\alpha_i)}{2}\}}\frac{\omega\left(\frac{1-\sum_{i=1}^4\alpha_i}{\alpha_4}\right)}{\alpha_1\alpha_2\alpha_3\alpha_4^2}\mathrm{d}\underline{\alpha}<\frac{0.07y}{x^b\log x},$$

$$\frac{y(1+o(1))}{x^b\log x}\int_{0.075}^{0.325}\int_{0.075}^{\min\{\alpha_1,0.4-\alpha_1\}}\int_{0.075}^{\alpha_2}\int_{\max\{0.075,\frac{(0.615-\sum_{i=1}^3\alpha_i)}{2}\}}^{\min\{\alpha_3,\frac{(1-\sum_{i=1}^3\alpha_i)}{2},0.4-\alpha_1-\alpha_2\}}\frac{\omega\left(\frac{1-\sum_{i=1}^4\alpha_i}{\alpha_4}\right)}{\alpha_1\alpha_2\alpha_3\alpha_4^2}\mathrm{d}\underline{\alpha}<\frac{0.01y}{x^b\log x},$$

$$\frac{y(1+o(1))}{x^b\log x}\left(\frac{\log\left(\frac{0.155}{0.075}\right)^6}{0.075\cdot 6!}+\frac{\log\left(\frac{0.155}{0.075}\right)^5\log\left(\frac{0.4}{0.155}\right)}{0.075\cdot 5!}+\frac{\log\left(\frac{0.155}{0.075}\right)^4\log\left(\frac{0.4}{0.155}\right)^2}{0.075\cdot 4!2!}\right)<\frac{0.1y}{x^b\log x}.$$

In the final upper bound, which is the upper bound for $\sum_{n\in\mathcal{B}(y)}\Theta_7(n)$, we used that $\ell_1^*+\cdots+\ell_5^*\le 0.615$ implies $\ell_3^*\le 0.155$. In summary, $1_{\mathbb{P}}(n)$ can be written as $1_{\mathbb{P}}(n)=\Psi(n)+\sum_{i=1}^{7}\Theta_i(n)$, where $\Psi(n)$ is a sum of no more than 100 good* and good** functions and where each Θ_i is non-negative and

$$\sum_{i=1}^{7}\sum_{n\in\mathcal{B}(y)}\Theta_i(n)<\frac{0.867y}{x^b\log x}.$$

This concludes the proof of Proposition 2.3 for the case $a\in(0.545,0.57]$.

Case 6: $a>0.61$.

This is our final case. By conditions (i), (ii), (iii), (iv), (v) and (vi) of Corollary 2.35, Buchstab iterations give

$$\psi(n,X)=\Psi_1(n)+\sum_{\substack{n=p_1p_2m\\ x^{0.065}\le p_2<p_1<X\\ \ell_1^*>0.42\\ \ell_2^*>0.645-\ell_1^*}}\psi(m,p_2)+\sum_{\substack{n=p_1p_2m\\ x^{0.065}\le p_2<p_1<X\\ \ell_1^*>0.42\\ \ell_2^*\in[0.58-\ell_1^*,0.645-\ell_1^*]}}\psi(m,p_2)+\sum_{\substack{n=p_1p_2m\\ x^{0.065}\le p_2<p_1<X\\ \ell_1^*>0.42\\ \ell_2^*<0.58-\ell_1^*\\ \ell_2^*\ge 0.1}}\psi(m,p_2) \tag{A.28}$$

$$+\sum_{\substack{n=p_1p_2p_3p_4m\\ x^{0.065}\le p_4<p_3<p_2<p_1<X\\ \ell_1^*>0.42\\ \ell_2^*<\min\{0.58-\ell_1^*,0.1\}}}\psi(m,p_4)+\sum_{\substack{n=p_1p_2m\\ x^{0.065}\le p_2<p_1<X\\ 0.355\le\ell_1^*\le 0.42}}\psi(m,p_2)+\sum_{\substack{n=p_1p_2m\\ x^{0.065}\le p_2<p_1<X\\ \ell_1^*<0.355\\ \ell_2^*>0.645-\ell_1^*}}\psi(m,p_2) \tag{A.29}$$

$$+\sum_{\substack{n=p_1p_2m\\ x^{0.065}\le p_2<p_1<X\\ \ell_1^*<0.355\\ \ell_2^*\in[0.355-\ell_1^*,0.42-\ell_1^*)\cup(0.58-\ell_1^*,0.645-\ell_1^*]}}\psi(m,p_2)+\sum_{\substack{n=p_1p_2m\\ x^{0.065}\le p_2<p_1<X\\ 0.325<\ell_1^*<0.355\\ \ell_2^*\in[0.42-\ell_1^*,0.58-\ell_1^*]\\ \ell_2^*>(0.645-\ell_1^*)/2}}\psi(m,p_2) \tag{A.30}$$

$$+\sum_{\substack{n=p_1p_2p_3p_4m\\ x^{0.065}\le p_4<p_3<p_2<p_1<X\\ 0.325<\ell_1^*<0.355\\ \ell_2^*\in[0.42-\ell_1^*,0.58-\ell_1^*]\\ \ell_2^*\le(0.645-\ell_1^*)/2\\ \ell_4^*>0.42-\ell_1^*}}\psi(m,p_4)+\sum_{\substack{n=p_1p_2p_3p_4m\\ x^{0.065}\le p_4<p_3<p_2<p_1<X\\ 0.325<\ell_1^*<0.355\\ \ell_2^*\in[0.42-\ell_1^*,0.58-\ell_1^*]\\ \ell_2^*\le(0.645-\ell_1^*)/2\\ \ell_4^*\in[0.355-\ell_1^*,0.42-\ell_1^*]}}\psi(m,p_4) \tag{A.31}$$

$$+\sum_{\substack{n=p_1p_2m\\ x^{0.065}\le p_2<p_1<X\\ \ell_1^*\le 0.325\\ \ell_2^*\in(\max\{0.42-\ell_1^*,0.2099\},0.58-\ell_1^*]}}\psi(m,p_2)+\sum_{\substack{n=p_1p_2p_3p_4m\\ x^{0.065}\le p_4<p_3<p_2<p_1<X\\ \ell_1^*\le 0.325\\ \ell_2^*\in[0.42-\ell_1^*,0.2099)\\ \ell_4^*>(0.42-\ell_2^*-\ell_3^*)/2\\ \ell_3^*<0.355-\ell_1^*}}\psi(m,p_4) \tag{A.32}$$

$$+\sum_{\substack{n=p_1p_2p_3p_4m\\ x^{0.065}\le p_4<p_3<p_2<p_1<X\\ \ell_1^*\le 0.325\\ \ell_2^*\in[0.42-\ell_1^*,0.2099)\\ \ell_4^*>(0.42-\ell_2^*-\ell_3^*)/2\\ \ell_3^*\in[0.355-\ell_1^*,0.42-\ell_1^*]}}\psi(m,p_4)+\sum_{\substack{n=p_1p_2p_3p_4m\\ x^{0.065}\le p_4<p_3<p_2<p_1<X\\ \ell_1^*\le 0.325\\ \ell_2^*\in[0.42-\ell_1^*,0.2099)\\ \ell_4^*>(0.42-\ell_2^*-\ell_3^*)/2\\ \ell_3^*>0.42-\ell_1^*\\ \ell_4^*<0.355-\ell_1^*}}\psi(m,p_4) \tag{A.33}$$

$$+\sum_{\substack{n=p_1p_2p_3p_4m\\ x^{0.065}\leq p_4<p_3<p_2<p_1<X\\ \ell_1^*\leq 0.325\\ \ell_2^*\in[0.42-\ell_1^*,0.2099)\\ \ell_4^*>(0.42-\ell_2^*-\ell_3^*)/2\\ \ell_3^*>0.42-\ell_1^*\\ \ell_4^*\in[0.355-\ell_1^*,0.42-\ell_1^*]}}\psi(m,p_4)+\sum_{\substack{n=p_1p_2p_3p_4m\\ x^{0.065}\leq p_4<p_3<p_2<p_1<X\\ \ell_1^*\leq 0.325\\ \ell_2^*\in[0.42-\ell_1^*,0.2099)\\ \ell_4^*>(0.42-\ell_2^*-\ell_3^*)/2\\ \ell_3^*>0.42-\ell_1^*\\ \ell_4^*>0.42-\ell_1^*}}\psi(m,p_4) \tag{A.34}$$

$$+\sum_{\substack{n=p_1p_2p_3p_4p_5p_6m\\ x^{0.065}\leq p_5<p_6<p_4<p_3<p_2<p_1<X\\ \ell_1^*\leq 0.325\\ \ell_2^*\in[0.42-\ell_1^*,0.2099)\\ \ell_4^*\leq(0.42-\ell_2^*-\ell_3^*)/2}}\psi(m,p_6)+\sum_{\substack{n=p_1p_2p_3p_4m\\ x^{0.065}\leq p_4<p_3<p_2<p_1<X\\ \ell_1^*\leq 0.325\\ \ell_2^*<0.355-\ell_1^*\\ \ell_4^*>(0.645-\ell_1^*-\ell_2^*-\ell_3^*)/2}}\psi(m,p_4) \tag{A.35}$$

$$+\sum_{\substack{n=p_1p_2p_3p_4p_5p_6m\\ x^{0.065}\leq p_6<p_5<p_4<p_3<p_2<p_1<X\\ \ell_1^*\leq 0.325\\ \ell_2^*<0.355-\ell_1^*\\ \ell_4^*\leq(0.645-\ell_1^*-\ell_2^*-\ell_3^*)/2}}\psi(m,p_6), \tag{A.36}$$

where $\Psi_1(n)$ is a sum of twelve good* functions.

Denote the first big sum in (A.28) by $\Theta_1(n)$ and the third by $\Theta_2(n)$. The second sum in (A.28) is good** by condition (b) of Corollary 2.35. Denote the first sum in (A.29) by $\Theta_3(n)$ and the third by $\Theta_4(n)$. The second sum in (A.29) is good** by condition (a) of Corollary 2.35. The first sum in (A.30) is good** by condition (b). Denote the second sum in (A.30) by $\Theta_5(n)$. Denote the first sum in (A.31) by $\Theta_6(n)$. The second sum in (A.31) is good** by condition (b) of Corollary 2.35. Denote the first sum in (A.32) by $\Theta_7(n)$ and the second by $\Theta_8(n)$. The first sum in (A.33) is good** by condition (b). Denote the second sum in (A.33) by $\Theta_9(n)$. The first sum in (A.34) is also good** by condition (b). Denote the second sum in (A.34) by $\Theta_{10}(n)$. Denote the first sum in (A.35) by $\Theta_{11}(n)$ and the second by $\Theta_{12}(n)$. Denote the sum in (A.36) by $\Theta_{13}(n)$.

Upper bounds for $\sum_{n\in\mathcal{B}(y)}\Theta_i(n)$, $i\in\{1,\ldots,13\}$, are given below:

$$\frac{y(1+o(1))}{x^b\log x}\int_{0.42}^{0.5+\varepsilon_1}\int_{0.645-\alpha_1}^{(1-\alpha_1)/2}\frac{\omega\left(\frac{1-\alpha_1-\alpha_2}{\alpha_2}\right)}{\alpha_1\alpha_2^2}\mathrm{d}\alpha_2\mathrm{d}\alpha_1<\frac{0.2194y}{x^b\log x},$$

$$\frac{y(1+o(1))}{x^b\log x}\int_{0.42}^{0.48}\int_{0.1}^{0.58-\alpha_1}\frac{\omega\left(\frac{1-\alpha_1-\alpha_2}{\alpha_2}\right)}{\alpha_1\alpha_2^2}\mathrm{d}\alpha_2\mathrm{d}\alpha_1<\frac{0.1769y}{x^b\log x},$$

$$\frac{y(1+o(1))}{x^b\log x}\int_{0.42}^{0.5}\int_{0.065}^{\min\{0.1,0.58-\alpha_1\}}\int_{0.065}^{\alpha_2}\int_{0.065}^{\min\{\alpha_3,\frac{(1-\sum_{i=1}^{3}\alpha_i)}{2}\}}\frac{\omega\left(\frac{1-\sum_{i=1}^{4}\alpha_i}{\alpha_4}\right)}{\alpha_1\alpha_2\alpha_3\alpha_4^2}\mathrm{d}\underline{\alpha}<\frac{0.0170y}{x^b\log x},$$

$$\frac{y(1+o(1))}{x^b\log x}\int_{0.3225}^{0.355}\int_{0.645-\alpha_1}^{\min\{\alpha_1,(1-\alpha_1)/2\}}\frac{\omega\left(\frac{1-\alpha_1-\alpha_2}{\alpha_2}\right)}{\alpha_1\alpha_2^2}\mathrm{d}\alpha_2\mathrm{d}\alpha_1<\frac{0.0191y}{x^b\log x},$$

$$\frac{y(1+o(1))}{x^b\log x}\int_{0.325}^{0.355}\int_{(0.645-\alpha_1)/2}^{0.58-\alpha_1}\frac{\omega\left(\frac{1-\alpha_1-\alpha_2}{\alpha_2}\right)}{\alpha_1\alpha_2^2}\mathrm{d}\alpha_2\mathrm{d}\alpha_1<\frac{0.1266y}{x^b\log x},$$

$$\frac{y(1+o(1))}{x^b\log x}\int_{0.325}^{0.355}\int_{0.42-\alpha_1}^{(0.645-\alpha_1)/2}\int_{0.42-\alpha_1}^{\alpha_2}\int_{0.42-\alpha_1}^{\min\{\alpha_3,\frac{(1-\sum_{i=1}^3\alpha_i)}{2}\}}\frac{\omega\left(\frac{1-\sum_{i=1}^4\alpha_i}{\alpha_4}\right)}{\alpha_1\alpha_2\alpha_3\alpha_4^2}\mathrm{d}\underline{\alpha}<\frac{0.0282y}{x^b\log x},$$

$$\frac{y(1+o(1))}{x^b\log x}\int_{0.2099}^{0.325}\int_{0.2099}^{\min\{\alpha_1,0.58-\alpha_1\}}\frac{\omega\left(\frac{1-\alpha_1-\alpha_2}{\alpha_2}\right)}{\alpha_1\alpha_2^2}\mathrm{d}\alpha_2\mathrm{d}\alpha_1<\frac{0.2102y}{x^b\log x},$$

$$\frac{y(1+o(1))}{x^b\log x}\int_{0.21}^{0.325}\int_{0.42-\alpha_1}^{0.21}\int_{0.065}^{\min\{\alpha_2,0.355-\alpha_1\}}\int_{\max\{0.065,\frac{(0.42-\alpha_2-\alpha_3)}{2}\}}^{\min\{\alpha_3,\frac{(1-\sum_{i=1}^3\alpha_i)}{2}\}}\frac{\omega\left(\frac{1-\sum_{i=1}^4\alpha_i}{\alpha_4}\right)}{\alpha_1\alpha_2\alpha_3\alpha_4^2}\mathrm{d}\underline{\alpha}<\frac{0.0249y}{x^b\log x},$$

$$\frac{y(1+o(1))}{x^b\log x}\int_{0.21}^{0.325}\int_{0.42-\alpha_1}^{0.21}\int_{0.42-\alpha_1}^{\alpha_2}\int_{\max\{0.065,\frac{(0.42-\alpha_2-\alpha_3)}{2}\}}^{\min\{\alpha_3,0.355-\alpha_1,\frac{(1-\sum_{i=1}^3\alpha_i)}{2}\}}\frac{\omega\left(\frac{1-\sum_{i=1}^4\alpha_i}{\alpha_4}\right)}{\alpha_1\alpha_2\alpha_3\alpha_4^2}\mathrm{d}\underline{\alpha}<\frac{0.0191y}{x^b\log x},$$

$$\frac{y(1+o(1))}{x^b\log x}\int_{0.21}^{0.325}\int_{0.42-\alpha_1}^{0.21}\int_{0.42-\alpha_1}^{\alpha_2}\int_{\max\{0.42-\alpha_1,\frac{(0.42-\alpha_2-\alpha_3)}{2}\}}^{\min\{\alpha_3,\frac{(1-\sum_{i=1}^3\alpha_i)}{2}\}}\frac{\omega\left(\frac{1-\sum_{i=1}^4\alpha_i}{\alpha_4}\right)}{\alpha_1\alpha_2\alpha_3\alpha_4^2}\mathrm{d}\underline{\alpha}<\frac{0.0280y}{x^b\log x},$$

$$\frac{y(1+o(1))}{x^b\log x}\frac{\log\left(\frac{0.325}{0.21}\right)\left(\frac{4!}{5!}\log\left(\frac{0.145}{0.065}\right)^5+\log\left(\frac{0.18}{0.145}\right)\log\left(\frac{0.145}{0.065}\right)^4+\log\left(\frac{0.21}{0.18}\right)\log\left(\frac{0.11}{0.065}\right)^4\right)}{0.065\cdot 4!}<\frac{0.0471y}{x^b\log x},$$

$$\frac{y(1+o(1))}{x^b\log x}\int_{0.065}^{0.325}\int_{0.065}^{\min\{\alpha_1,0.355-\alpha_1\}}\int_{0.065}^{\alpha_2}\int_{\max\{0.065,\frac{(0.645-\alpha_1-\alpha_2-\alpha_3)}{2}\}}^{\min\{\alpha_3,\frac{(1-\sum_{i=1}^3\alpha_i)}{2}\}}\frac{\omega\left(\frac{1-\sum_{i=1}^4\alpha_i}{\alpha_4}\right)}{\alpha_1\alpha_2\alpha_3\alpha_4^2}\mathrm{d}\underline{\alpha}<\frac{0.0180y}{x^b\log x},$$

$$\frac{y(1+o(1))}{x^b\log x}\left(\frac{\log\left(\frac{0.1775}{0.065}\right)^6}{0.065\cdot 6!}+\frac{\log\left(\frac{0.22}{0.1775}\right)\log\left(\frac{0.1775}{0.065}\right)^5}{0.065\cdot 5!}+\frac{\log\left(\frac{0.29}{0.22}\right)\log\left(\frac{0.135}{0.065}\right)^5}{0.065\cdot 5!}\right)<\frac{0.0576y}{x^b\log x}.$$

In summary, we have shown that $1_{\mathbb{P}}(n)$ can be written as $1_{\mathbb{P}}(n)=\Psi(n)+\sum_{i=1}^{13}\Theta_i(n)$, where $\Psi(n)$ is a sum of no more than 100 good* and good** functions and where each Θ_i is non-negative and

$$\sum_{i=1}^{13}\sum_{n\in\mathcal{B}(y)}\Theta_i(n)<\frac{0.9921y}{x^b\log x}.$$

This concludes the proof of Proposition 2.3 for the case $a>0.61$. $\square$

Appendix B

Proof of Lemma 3.5

In Appendix B, we give a proof of Lemma 3.5. It follows the arguments of Section 5 and Section 8.1 of Polymath [37] and only small changes need to be made. As a first step, we summarize the first pages of Section 5 of [37], which can be used largely unchanged.

Lemma B.1. *Let $\omega, \delta > 0$. Let $\varepsilon \in (0, 10^{-100}\delta)$. Let $N = x^\gamma$ with $\gamma \in (4\omega + 2\delta, 1/2]$. Let α and β be coefficient sequences at scales M and N with $x \ll MN \ll x$. Assume β has the Siegel-Walfisz property.*

Let $\psi_M(m)$ be a real, non-negative coefficient sequence which is smooth at scale M. Let $c_{q,r} \in \mathbb{C}$ with $|c_{q,r}| = 1$, $a, b_1, b_2 \in \mathbb{Z}$ with $(b_1 b_2, P(x^\delta)) = 1$ and $Q, R \in (0, \infty)$ with

$$x^{-4\varepsilon-\delta} N \ll R \ll x^{-2\varepsilon} N, \tag{B.1}$$

$$x^{1/2-\varepsilon} \ll QR \ll x^{1/2+2\omega+\varepsilon}. \tag{B.2}$$

Let $C^*(n) = 1_{\frac{b_1}{n} \equiv \frac{b_2}{n+\ell r}((q_1,q_2))}$ for n with $(n, q_1 r) = (n + \ell r, q_2) = 1$.
Write $D_0 = \exp((\log x)^{1/3})$ and set $\Sigma(b_1, b_2)$ equal to

$$\sum_{\substack{r \asymp R \\ (r,a)=1}} \sum_{\substack{|\ell| \ll \frac{N}{R} \\ \ell \neq 0}} \sum_{\substack{q_1 \asymp Q \\ q_1 r | P(x^\delta) \\ (q_1,a)=1 \\ (q_1, P(D_0))=1}} \sum_{\substack{q_2 \asymp Q \\ q_2 r | P(x^\delta) \\ (q_2,a)=1 \\ (q_2, P(D_0))=1}} c_{q_1,r}\overline{c_{q_2,r}} \sum_{\substack{n \\ (n,q_1 r)=1 \\ (n+\ell r, q_2)=1}} C^*(n)\beta(n)\overline{\beta(n+\ell r)} \sum_m \psi_M(m) 1_{m \equiv y\,([q_1,q_2]r)}$$

where y is a function of n, ℓ, a, b_i, r, q_i with $y \equiv \frac{b_1}{n}(q_1), y \equiv \frac{b_2}{n+\ell r}(q_2)$ and $y \equiv \frac{a}{n}(r)$.

Then suppose that $\Sigma(b_1, b_2) = X + O_{A,\varepsilon}(N^2 M R^{-1} (\log x)^{-A})$, where X does not depend on b_1 and b_2 and the implied constant does not depend on Q, R, ψ_M, $(c_{q,r})$, a, b_1 and b_2. In that case we have

$$\sum_{\substack{q \le x^{1/2+2\omega} \\ q | P(x^\delta) \\ (q,a)=1}} \left| \sum_{n \equiv a(q)} (\alpha \star \beta)(n) - \frac{1}{\phi(q)} \sum_{(n,q)=1} (\alpha \star \beta)(n) \right| \ll_{A,\varepsilon} \frac{x}{(\log x)^A}. \tag{B.3}$$

Proof. This is essentially a summary of pages 42 to 48 of Section 5.3 of [37], which concluded with an application of the dispersion method of Linnik. $\square$

The next step is to complete sums in $\Sigma(b_1, b_2)$, with the goal of reducing the proof of (B.3) to bounding certain exponential sums. For this purpose we record the following minor amendment of Lemma 4.9 of [37], which (unlike Lemma 4.9) preserves the smooth coefficients of h:

Lemma B.2 (Completion of sums)**.** *Let $M \geq 1$ be a real number and let ψ_M be a shifted smooth coefficient sequence at scale M. Let I be a finite set, let $c_i \in \mathbb{C}$, $q \in \mathbb{N}$ and let $a_i(q)$ be a residue class. Set*

$$\varphi_H(h) = \frac{1}{M} \sum_m \psi_M(m) e\left(\frac{-mh}{q}\right). \tag{B.4}$$

Then for $A > 0$, $\varepsilon > 0$ and $B \geq qM^{-1+\varepsilon}$,

$$\begin{aligned}\sum_{i \in I} c_i \sum_m \psi_M(m) 1_{m = a_i(q)} = &\frac{1}{q} \sum_{m \geq 1} \psi_M(m) \sum_{i \in I} c_i + O_{A,\varepsilon}\Big(M^{-A} \sum_{i \in I} |c_i|\Big) \\ &+ \frac{M}{q} \sum_{0 < |h| \leq B} \varphi_H(h) \sum_{i \in I} c_i e\left(\frac{a_i h}{q}\right).\end{aligned}$$

Proof. This is a direct consequence of the proof of Lemma 4.9 of [37]. $\square$

Before we apply Lemma B.2 to $\Sigma(b_1, b_2)$, it is convenient to record one more result of Polymath [37], which will be used to bound some of the error terms which soon arise:

Lemma B.3. *Let $\omega, \delta > 0$. Let M, N, Q and R be positive real numbers such that $MN \asymp x$ and such that conditions* (B.1) *and* (B.2) *of Lemma* B.1 *hold. Let $a, b_1, b_2 \in \mathbb{Z}$. Let $C^*(n)$ be as given in Lemma* B.1. *Write $D_0 = \exp((\log x)^{1/3})$. Then*

$$\sum_{r \asymp R} \sum_{\substack{|\ell| \ll \frac{N}{R} \\ \ell \neq 0}} \sum_{\substack{1 \neq q_0 \ll Q \\ (q_0, P(D_0)) = 1}} \sum_{\substack{q_1, q_2 \asymp Q \\ q_1 r, q_2 r | P(x^\delta) \\ (q_1, q_2) = q_0}} \frac{1}{[q_1, q_2]} \sum_{\substack{n \\ (n, q_1 r) = 1 \\ (n + \ell r, q_2) = 1}} (\tau(n)\tau(n + \ell r))^{O(1)} C^*(n) \ll_A \frac{N^2}{(\log x)^A}.$$

Proof. This is shown on page 49 of [37]. $\square$

Applying Lemma B.2 and Lemma B.3 to Lemma B.1, we obtain Lemma B.4, stated below. This is an amended version of Theorem 5.8 of [37], in which the smooth coefficients of variable h have been preserved.

Lemma B.4 (Exponential sum estimates)**.** *Let $\omega > 0$ and $\delta > 0$. Let $\varepsilon \in (0, 10^{-100}\delta)$. Let $N = x^\gamma$ with $\gamma \in (4\omega + 2\delta, 1/2]$. Let α and β be coefficient sequences at scales M and N with $x \ll MN \ll x$. Assume that β has the Siegel-Walfisz property.*

Write $D_0 = \exp((\log x)^{1/3})$. Consider $Q, R \in (0, \infty)$ such that conditions (B.1) *and* (B.2) *of Lemma* B.1 *are satisfied. Let $\psi_M(m)$ be a real, non-negative coefficient sequence, smooth at scale M. Let q_0 be a positive integer with $q_0 \ll Q$, $q_0 \mid P(x^\delta)$ and $(q_0, P(D_0)) = 1$, let $\ell \in \mathbb{Z}$ with $0 \neq |\ell| \ll \frac{N}{R}$ and let $a, b_1, b_2 \in \mathbb{Z}$ with $(q_0, ab_1b_2) = 1$. Write $H(q_0) = \frac{x^\varepsilon RQ^2}{q_0 M}$ and set Σ^* equal to*

$$\sum_{\substack{r \asymp R \\ (r,a)=1 \\ (r,b_1)=1 \\ (r,b_2)=1}} \sum\sum_{\substack{q_1,q_2 \asymp Q/q_0 \\ (q_1,q_2)=1 \\ (q_1q_2, ab_1b_2)=1 \\ q_0q_1r, q_0q_2r \mid P(x^\delta)}} \left| \sum_{\substack{|h| \le H(q_0) \\ h \neq 0}} \varphi_H(h) \sum_{\substack{n \\ (n, rq_0q_1)=1 \\ (n+\ell r, q_0q_2)=1}} C_0(n)\beta(n)\overline{\beta(n+\ell r)}\Phi_\ell(h, n, r, q_0, q_1, q_2) \right|,$$

where $\Phi_\ell(h, n, r, q_0, q_1, q_2) = e_r\left(\frac{ah}{nq_0q_1q_2}\right) e_{q_0q_1}\left(\frac{b_1h}{nrq_2}\right) e_{q_2}\left(\frac{b_2h}{(n+\ell r)rq_0q_1}\right).$

Here $C_0(n) = 1_{\frac{b_1}{n} \equiv \frac{b_2}{n+\ell r}(q_0)}$, while $\varphi_H(m)$ is as defined in Lemma B.2*, with $q = rq_0q_1q_2$.*

Suppose that $\Sigma^* \ll_\varepsilon (q_0, \ell)RQ^2N(x^\varepsilon q_0^2)^{-1}$, with the implied constant independent of the choice of Q, R, ψ_M, q_0, ℓ, a, b_1 and b_2. Then (B.3) *is satisfied.*

Proof. We consider $\Sigma(b_1, b_2)$, described in Lemma B.1, for b_1, b_2 with $(b_1b_2, P(x^\delta)) = 1$. For fixed r, ℓ, q_1 and q_2, we apply Lemma B.2 to the remaining variables m and n and find that $\Sigma(b_1, b_2) = \Sigma_0^*(b_1, b_2) + \Sigma_1^*(b_1, b_2) + O_A(N^2MR^{-1}(\log x)^{-A})$, where

$$\Sigma_0^*(b_1, b_2) = \left(\sum_m \psi_M(m)\right) \sum_{\substack{r \asymp R \\ (r,a)=1}} \sum_{\substack{|\ell| \ll \frac{N}{R} \\ \ell \neq 0}} \sum\sum_{\substack{q_1,q_2 \asymp Q \\ q_1r, q_2r \mid P(x^\delta) \\ (q_1q_2, a)=1 \\ (q_1q_2, P(D_0))=1}} \frac{c_{q_1,r}\overline{c_{q_2,r}}}{r[q_1, q_2]} \sum_{\substack{n \\ (n,q_1r)=1 \\ (n+\ell r, q_2)=1}} C^*(n)\beta(n)\overline{\beta(n+\ell r)},$$

$$\Sigma_1^*(b_1, b_2) = \sum_{\substack{r \asymp R \\ (r,a)=1}} \sum_{\substack{|\ell| \ll \frac{N}{R} \\ \ell \neq 0}} \sum\sum_{\substack{q_1,q_2 \asymp Q \\ q_1r, q_2r \mid P(x^\delta) \\ (q_1q_2, a)=1 \\ (q_1q_2, P(D_0))=1}} \frac{c_{q_1,r}\overline{c_{q_2,r}}}{r[q_1, q_2]} \sum_{\substack{|h| \le \frac{x^\varepsilon RQ^2}{(q_1,q_2)M} \\ h \neq 0}} M\varphi_H(h) \sum_{\substack{n \\ (n,q_1r)=1 \\ (n+\ell r, q_2)=1}} \frac{C^*(n)\beta(n)\overline{\beta(n+\ell r)}}{e_{[q_1,q_2]r}(-yh)}.$$

(Here the modulus of φ_H is $[q_1, q_2]r$.) Observe now that the component of $\Sigma_0^*(b_1, b_2)$ which has $(q_1, q_2) = 1$ is given by

$$X = \left(\sum_m \psi_M(m)\right) \sum_{\substack{r \asymp R \\ (r,a)=1}} \sum_{\substack{|\ell| \ll \frac{N}{R} \\ \ell \neq 0}} \sum\sum_{\substack{q_1,q_2 \\ q_1r, q_2r \mid P(x^\delta) \\ (q_1q_2, a)=1 \\ (q_1q_2, P(D_0))=1 \\ (q_1,q_2)=1}} \frac{c_{q_1,r}\overline{c_{q_2,r}}}{r[q_1, q_2]} \sum_{\substack{n \\ (n,q_1r)=1 \\ (n+\ell r, q_2)=1}} \beta(n)\overline{\beta(n+\ell r)}.$$

(Here we used that $C^*(n) = 1$ on this component.) Observe that X does not depend on b_1 and b_2. Writing $(q_1, q_2) = q_0$ and using Lemma B.3, the remaining terms of $\Sigma_0^*(b_1, b_2)$, for which $q_0 \neq 1$, are bounded in absolute value by

$$\frac{M(\log x)^{O(1)}}{R} \sum_{r \asymp R} \sum_{\substack{|\ell| \ll \frac{N}{R} \\ \ell \neq 0}} \sum_{\substack{1 \neq q_0 \ll Q \\ (q_0, P(D_0))=1}} \sum\sum_{\substack{q_1, q_2 \asymp Q \\ q_1 r | P(x^\delta) \\ q_2 r | P(x^\delta) \\ (q_1, q_2) = q_0}} \sum_{\substack{n \\ (n, q_1 r)=1 \\ (n+\ell r, q_2)=1}} \frac{(\tau(n)\tau(n+\ell r))^{O(1)}}{[q_1, q_2](C^*(n))^{-1}} \ll_A \frac{N^2 M}{R(\log x)^A}.$$

Now we look at $\Sigma_1^*(b_1, b_2)$. We split the sum up according to the value of (q_1, q_2) and deduce that $|\Sigma_1^*(b_1, b_2)| \ll E$, where E is given by

$$\sum_{\substack{|\ell| \ll \frac{N}{R} \\ \ell \neq 0}} \sum_{q_0 \ll Q} \frac{q_0 M}{RQ^2} \sum_{r \asymp R} \sum\sum_{\substack{q_1, q_2 \asymp Q/q_0 \\ (q_1, q_2)=1 \\ (rq_0 q_1 q_2, ab_1 b_2)=1 \\ q_0 q_1 r, q_0 q_2 r | P(x^\delta)}} \left| \sum_{\substack{|h| \leq H(q_0) \\ h \neq 0}} \varphi_H(h) \sum_{\substack{n \\ (n, rq_0 q_1)=1 \\ (n+\ell r, q_2)=1}} C_0(n)\beta(n)\overline{\beta(n+\ell r)}\Phi_\ell(h, n, r, q_0, q_1, q_2) \right|. \tag{B.5}$$

(Notice that we introduced the condition $(rq_0q_1q_2, b_1b_2) = 1$. Since $(b_1b_2, P(x^\delta)) = 1$, while $rq_0[q_1, q_2] \mid P(x^\delta)$, this condition is trivially satisfied and does not change the summation range. This is merely a technical restriction, which now allows us to drop the requirement $(b_1b_2, P(x^\delta)) = 1$.)

Note that the inner sums of (B.5) are of the form Σ^* described in the statement of Lemma B.4, and recall that we assumed that $\Sigma^* \ll_\varepsilon (q_0, \ell)RQ^2N(x^\varepsilon q_0^2)^{-1}$ for any choice of Q, R, ψ_M, q_0, ℓ, a, b_1 and b_2. Hence,

$$|\Sigma_1^*(b_1, b_2)| \ll_\varepsilon \sum_{q_0 \ll Q} \frac{q_0 M}{RQ^2} \sum_{\substack{|\ell| \ll \frac{N}{R} \\ \ell \neq 0}} \frac{(q_0, \ell)RQ^2N}{x^\varepsilon q_0^2} \ll_{A,\varepsilon} \frac{N^2 M}{R(\log x)^A}.$$

Overall we have shown that $\Sigma(b_1, b_2) = X + O_{A,\varepsilon}(N^2MR^{-1}(\log x)^{-A})$, with X not dependent on b_1 and b_2. By Lemma B.1 this implies that (B.3) is satisfied. □

For the final step of our proof of Lemma 3.5, we use the Cauchy-Schwarz inequality to remove β. We proceed like in Section 8.1 of [37].

Lemma B.5. *Let $\omega > 0$ and $\delta > 0$ and $\varepsilon \in (0, 10^{-100}\delta)$.*
Let $N = x^\gamma$, where $\gamma \in (4\omega + 2\delta, \frac{1}{2} - 2\omega - 8\varepsilon)$.
Let α and β be coefficient sequences at scales M and N with $x \ll MN \ll x$. Assume β has the Siegel-Walfisz property.

Let Z_1 and Σ_1 be as described in Lemma 3.5 *and let Σ^* be as described in Lemma* B.4. *Suppose that $\Sigma_1(z) = O_\varepsilon((q_0,\ell)RQNUV^2x^{-4\varepsilon})$ for every $z \in Z_1$. Then also*

$$\Sigma^* \ll_\varepsilon (q_0,\ell)RQ^2N(x^\varepsilon q_0^2)^{-1}.$$

Proof. We consider Σ^*, described in Lemma B.4. We wish to show that $\Sigma^* \ll_\varepsilon \frac{(q_0,\ell)RQ^2N}{x^\varepsilon q_0^2}$. Recall that Σ^* is equal to

$$\sum_{\substack{r\asymp R\\(r,ab_1b_2)=1}} \sum\sum_{\substack{q_1,q_2\asymp Q/q_0\\(q_1,q_2)=1\\(q_1q_2,ab_1b_2)=1\\q_0q_1r,q_0q_2r\mid P(x^\delta)}} \left| \sum_{0<|h|\le H} \varphi_H(h) \sum_{\substack{n\\(n,rq_0q_1)=1\\(n+\ell r,q_0q_2)=1}} C_0(n)\beta(n)\overline{\beta(n+\ell r)}\Phi_\ell(h,n,r,q_0,q_1,q_2)\right|,$$

where $H = \frac{x^\varepsilon RQ^2}{q_0M}$. (Since the value of q_0 has been fixed in Σ^*, it is no longer necessary to indicate the dependence of H on q_0.) Observe first that if $H < 1$, then the sum over h is empty and $\Sigma^* = 0$. Hence we may assume that $H \ge 1$. But then $q_0^{-1}x^{-5\varepsilon}Q/H \le x^{-5\varepsilon}(Q/q_0)$. Further, since $RQ \ll x^{1/2+2\omega+\varepsilon}$, $MN \asymp x$ and $N = x^\gamma$ with $\gamma \le 1/2 - 2\omega - 8\varepsilon$, we also have $q_0^{-1}x^{-\delta-5\varepsilon}Q/H \gg x^{-\delta+\varepsilon}$. Hence for any $q_1 \asymp Q/q_0$ with $q_1 \mid P(x^\delta)$ we can find u_1 and v_1 with $q_1 = u_1v_1$ and $q_0^{-1}x^{-\delta-5\varepsilon}Q/H \le u_1 \le q_0^{-1}x^{-5\varepsilon}Q/H$.

So, using dyadic decomposition, we deduce that the bound $\Sigma^* \ll_\varepsilon \frac{(q_0,\ell)RQ^2N}{x^\varepsilon q_0^2}$ holds if the following is true for every $H^* \in \mathbb{R}\setminus\{0\}$ with $1 \ll |H^*| \ll H$ and all U and V which satisfy (3.14), (3.15) and (3.16):

$$\sum_{r\asymp R} \sum_{\substack{u_1,v_1\\u_1\asymp U,v_1\asymp V\\q_0u_1v_1r\mid P(x^\delta)}} \sum_{\substack{q_2\asymp Q/q_0\\(u_1v_1,q_2)=1\\q_0q_2r\mid P(x^\delta)\\(rq_0u_1v_1q_2,ab_1b_2)=1}} \left|\sum_{h\sim H^*}\varphi_H^\star(h) \sum_{\substack{n\\(n,rq_0u_1v_1)=1\\(n+\ell r,q_0q_2)=1}} C_0(n)\beta(n)\overline{\beta(n+\ell r)}\Phi_\ell(h,n,r,q_0,u_1v_1,q_2)\right|$$
$$\ll_\varepsilon \frac{(q_0,\ell)RQ^2N}{x^{2\varepsilon}q_0^2}.$$

(Recall here that $\varphi_H(h)$ has modulus $rq_0q_1q_2$, $\varphi_H^\star(h)$ has modulus $rq_0u_1v_1q_2$ and $\varphi_H^{\star\star}(h)$ has modulus $rq_0u_1v_2q_2$.) Denote the LHS of the above expression by Σ_1^* and observe that Σ_1^* is equal to

$$\sum_{r\asymp R} \sum_{\substack{u_1,q_2\\u_1\asymp U\\q_2\asymp Q/q_0}} \sum_{\substack{n\\(n,rq_0u_1)=1\\(n+\ell r,q_0q_2)=1}} C_0(n)\beta(n)\overline{\beta(n+\ell r)} \sum_{h\sim H^*} \sum_{\substack{v_1\asymp V\\(v_1,n)=1}} c_{r,u_1,v_1,q_2}\varphi_H^\star(h)\Phi_\ell(h,n,r,q_0,u_1v_1,q_2),$$

where c_{r,u_1,v_1,q_2} are 1-bounded constants, supported on r, u_1, v_1 and q_2 with $q_0u_1v_1q_2r \mid P(x^\delta)$ and $(rq_0u_1v_1q_2, ab_1b_2) = 1$. Next we apply the Cauchy-Schwarz inequality. We find $(\Sigma_1^*)^2 \le \Upsilon_1\Upsilon_2$, where

$$\Upsilon_1 = \sum_{r\asymp R}\sum_{u_1\asymp U}\sum_{q_2\asymp Q/q_0}\sum_{\substack{n\\(n,rq_0u_1)=1\\(n+\ell r,q_0q_2)=1}} C_0(n)|\beta(n)\overline{\beta(n+\ell r)}|^2 \ll \frac{(q_0,\ell)RQUN}{q_0^2},$$

$$\begin{aligned}\Upsilon_2 =&\sum_{r\asymp R}\sum_{u_1\asymp U}\sum_{q_2\asymp \frac{Q}{q_0}}\sum_{\substack{n\\(n,rq_0u_1)=1\\(n+\ell r,q_0)=1\\(n+\ell r,q_2)=1}} C_0(n)\psi_N(n)\Bigg|\sum_{h\sim H^*}\sum_{\substack{v_1\asymp V\\(v_1,n)=1}} c_{r,u_1,v_1,q_2}\varphi_H^\star(h)\Phi_\ell(h,n,r,q_0,u_1v_1,q_2)\Bigg|^2\\ \le& \sum_{\substack{r,u_1,v_1,v_2,q_2\\r\asymp R\\u_1\asymp U\\v_1,v_2\asymp V\\q_2\asymp Q/q_0\\(rq_0u_1v_1v_2q_2,ab_1b_2)=1\\q_0u_1v_1q_2r\mid P(x^\delta)\\q_0u_1v_2q_2r\mid P(x^\delta)}}\Bigg|\sum_{\substack{h_1,h_2\\h_1\sim H^*\\h_2\sim H^*}}\sum_{\substack{n\\(n,rq_0)=1\\(n,u_1v_1v_2)=1\\(n+\ell r,q_0q_2)=1}} C_0(n)\varphi_H^\star(h_1)\overline{\varphi_H^{\star\star}(h_2)}\psi_N(n)\widetilde{\Phi}_\ell(h_1,h_2,n,r,q_0,u_1,v_1,v_2,q_2)\Bigg|\end{aligned}$$

for some smooth coefficient sequence ψ_N, located at scale N, with $\psi_N(n) \ge 0$ for all n and $\psi_N(n) \ge 1$ on the support of β. Here we used the notation

$$\widetilde{\Phi}_\ell(h_1,h_2,n,r,q_0,u_1,v_1,v_2,q_2) = \Phi_\ell(h_1,n,r,q_0,u_1v_1,q_2)\overline{\Phi_\ell(h_2,n,r,q_0,u_1v_2,q_2)}.$$

Then $\Sigma_1^* \ll_\varepsilon (q_0,\ell)RQ^2N(x^{2\varepsilon}q_0^2)^{-1}$ provided that $\Upsilon_2 \ll_\varepsilon (q_0,\ell)RQNUV^2x^{-4\varepsilon}$. However, looking at the notation of Lemma 3.5, we see that $\Upsilon_2 \ll \Sigma_1(z)$ for a suitable choice of $z \in Z_1$. Recall the assumption $\Sigma_1(z) = O_\varepsilon((q_0,\ell)RQNUV^2x^{-4\varepsilon})$ to conclude the proof. $\square$

Lemma 3.5 is simply a summary of Lemma B.1, Lemma B.4 and Lemma B.5.

Bibliography

[1] Roger Baker, Glyn Harman, and János Pintz. The difference between consecutive primes, II. *Proceedings of the London Mathematical Society*, 83(3):532–562, 2001.

[2] Roger Baker and Alastair Irving. Bounded intervals containing many primes. *Mathematische Zeitschrift*, 286:821–841, 2017.

[3] Roger Baker and Andreas Weingartner. A ternary diophantine inequality over primes. *Acta Arithmetica*, 162(2):159–196, 2014.

[4] Enrico Bombieri. On the large sieve. *Mathematika*, 12(2):201–225, 1965.

[5] Enrico Bombieri, John Friedlander, and Henryk Iwaniec. Primes in arithmetic progressions to large moduli. *Acta Mathematica*, 156:203–251, 1986.

[6] Jingrun Chen. On the representation of a large even integer as the sum of a prime and the product of at most two primes. *Scientia Sinica*, 16:157–176, 1973.

[7] Tomás Oliveira e Silva, Siegfried Herzog, and Silvio Pardi. Empirical verification of the even Goldbach conjecture and computation of prime gaps up to $4 \cdot 1018$. *Math. Comput.*, 83(288):2033–2060, 2013.

[8] Etienne Fouvry and Henryk Iwaniec. Primes in arithmetic progressions. *Acta Arithmetica*, 42(2):197–218, 1983.

[9] John Friedlander and Henryk Iwaniec. *Opera de Cribro*. American Mathematical Soc., 2010.

[10] P. X. Gallagher. On the distribution of primes in short intervals. *Mathematika*, 23(1):4–9, 1976.

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

[12] Larry Guth and James Maynard. New large value estimates for Dirichlet polynomials. 2024. arXiv:2405.20552.

[13] H. Halberstam and H. E. Richert. *Sieve methods.* Dover Publications, 2011.

[14] Glyn Harman. *Prime-Detecting Sieves.* Princeton University Press, 2012.

[15] D. R. Heath-Brown. The differences between consecutive primes. *Journal of the London Mathematical Society*, s2–18(1):7–13, 1978.

[16] D. R. Heath-Brown. The differences between consecutive primes, III. *Journal of the London Mathematical Society*, s2–20(2):177–178, 1979.

[17] D. R. Heath-Brown. Zero density estimates for the Riemann zeta-function and Dirichlet L-functions. *Journal of the London Mathematical Society*, s2–19(2):221–232, 1979.

[18] D. R. Heath-Brown. Prime numbers in short intervals and a generalized Vaughan identity. *Canadian Journal of Mathematics*, 34(6):1365–1377, 1982.

[19] D. R. Heath-Brown. The number of primes in a short interval. *Journal für die reine und angewandte Mathematik*, (389):22–63, 1988.

[20] D. R. Heath-Brown. The differences between consecutive smooth numbers. *Acta Arithmetica*, 184:267–285, 2018.

[21] D. R. Heath-Brown. The differences between consecutive primes, V. *International Mathematics Research Notices*, 2021(22):17514–17562, 2019.

[22] Adolf Hildebrand and Gerald Tenenbaum. Integers without large prime factors. *Journal de théorie des nombres de Bordeaux*, 5(2):411–484, 1993.

[23] Martin Huxley. On the difference between consecutive primes. *Inventiones mathematicae*, 15(2):164–170, 1971.

[24] Asim Islam. *Multiplicative properties of integers in short intervals.* PhD thesis, Royal Holloway, University of London, 2015.

[25] Aleksandar Ivić. *The Riemann zeta-function: The theory of the Riemann zeta-function with applications.* Wiley, New York, 1985.

[26] Olli Järviniemi. On large differences between consecutive primes. 2022. arXiv:2212.10965.

[27] Kaisa Matomäki. Large differences between consecutive primes. *The Quarterly Journal of Mathematics*, 58(4):489–518, 2007.

[28] Kaisa Matomäki and Joni Teräväinen. Almost primes in almost all short intervals II. 2022. arXiv:2207.05038.

[29] James Maynard. On the difference between consecutive primes. 2012. arXiv:1201.1787.

[30] James Maynard. Small gaps between primes. *Annals of mathematics*, 181:383–413, 2015.

[31] James Maynard. Primes represented by incomplete norm forms. *Forum of Mathematics, Pi*, 8(e3), 2020.

[32] James Maynard. Primes in arithmetic progressions to large moduli I: Fixed residue classes. *Memoirs of the American Mathematical Society*, to appear.

[33] James Maynard. Primes in arithmetic progressions to large moduli III: Uniform residue classes. *Memoirs of the American Mathematical Society*, to appear.

[34] Hugh Montgomery. *Topics in multiplicative number theory*. Springer, 1971.

[35] Yoichi Motohashi. *Lectures on sieve methods and prime number theory*. Springer, 1983.

[36] Adrian Peck. *On the differences between consecutive primes*. PhD thesis, University of Oxford, 1996.

[37] D. H. J. Polymath. New equidistribution estimates of Zhang type. *Algebra & Number Theory*, 8(9):2067–2199, 2014.

[38] D. H. J. Polymath. Variants of the Selberg sieve, and bounded intervals containing many primes. *Research in the Mathematical sciences*, 1:1–83, 2014.

[39] Atle Selberg. On the normal density of primes in small intervals and the difference between consecutive primes. *Arch. Math. Naturvid.*, 47:87–105, 1943.

[40] Peter Shiu. A Brun-Titchmarsh theorem for multiplicative functions. *Journal für die reine und angewandte Mathematik*, 313:161–170, 1980.

[41] Julia Stadlmann. On the mean square gap between primes. 2022. arXiv:2212.10867.

[42] Julia Stadlmann. On primes in arithmetic progressions and bounded gaps between many primes. 2023. arXiv:2309.00425.

[43] A. I. Vinogradov. The density hypothesis for Dirichet L-series. *Izv. Akad. Nauk SSSR Ser. Mat.*, 29:903–934, 1965.

[44] Gang Yu. The differences between consecutive primes. *Bulletin of the London Mathematical Society*, 28(3):242–248, 1996.

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