

POSITIVE DENSITY FOR CONSECUTIVE RUNS OF SUMS OF TWO SQUARES

NOAM KIMMEL¹ AND VIVIAN KUPERBERG²

¹*Raymond and Beverly Sackler School of Mathematical Sciences, Tel Aviv University,
Tel Aviv 69978, Israel*
(noamkimmel@mail.tau.ac.il)

²*Département Mathematik, ETH Zürich, Rämistrasse 101, 8092 Zürich, Switzerland*
(vivian.kuperberg@math.ethz.ch)

(Received 26 June 2024; revised 22 March 2025; accepted 24 March 2025)

Abstract We study the distribution of consecutive sums of two squares in arithmetic progressions. We show that for any odd squarefree modulus q , any two reduced congruence classes a_1 and $a_2 \pmod q$, and any $r_1, r_2 \geq 1$, a positive density of sums of two squares begin a chain of r_1 consecutive sums of two squares, all of which are $a_1 \pmod q$, followed immediately by a chain of r_2 consecutive sums of two squares, all of which are $a_2 \pmod q$. This is an analog of the result of Maynard for the sequence of primes, showing that for any reduced congruence class $a \pmod q$ and for any $r \geq 1$, a positive density of primes begin a sequence of r consecutive primes, all of which are $a \pmod q$.

Contents

1	Introduction	2
2	Statement of sieve results and proof of the main theorem	5
2.1.	GPY sieve setup	5
2.2.	Conventions and notation	8
2.3.	Statement of the main sieve theorem	9
2.4.	Proof of Theorem 1	11

N.K. was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant agreement No. 786758). V.K. was supported by the NSF Mathematical Sciences Research Program through the grant DMS-2202128. The authors would like to thank the anonymous referee for helpful comments.

Keywords: Sums of two squares; sieve methods; arithmetic progressions

2020 Mathematics subject classification: Primary 11N69; 11N36

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<https://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

3	Proofs of sieve results	16
3.1.	Sieve notation and setup	16
3.2.	Auxiliary lemmas for sieve weights	18
3.3.	Auxiliary lemmas for sums of two squares	23
3.4.	Estimating S_1	26
3.5.	Estimating $S_2^{(m)}$	27
3.6.	Estimating $S_3^{(m_1, m_2)}$	29
3.7.	Estimating $S_4^{(m)}$	34
3.8.	Estimating $S_5^{(m)}$	36
3.9.	Estimating $S_6^{(b)}$	41
4	Singular series estimates	46
4.1.	Averaging over \mathcal{B} and ν_0	46
4.2.	Averaging over $S_6^{(b)}(\nu_1)$	47
	References	51
	Index of Notation	52

1. Introduction

The sequence of primes is known, by the prime number theorem in arithmetic progressions, to be equidistributed among reduced congruence classes to any modulus q . To be precise, for any modulus q and for any reduced congruence class $a \bmod q$, let $\pi(x; q, a)$ denote the number of primes $p \leq x$ with $p \equiv a \bmod q$ and let $\pi(x)$ denote the number of primes $p \leq x$. Then

$$\pi(x; q, a) = \frac{\pi(x)}{\phi(q)}(1 + o(1)).$$

Much less is known about analogous questions for strings of consecutive primes. Let p_n denote the sequence of primes in increasing order. For any $M \geq 1$, for a fixed modulus q and any M -tuple $\mathbf{a} = [a_1, \dots, a_M]$ of reduced residue classes mod q , let $\pi(x; q, \mathbf{a})$ denote the number of strings of consecutive primes matching the residue classes of \mathbf{a} . That is, define

$$\pi(x; q, \mathbf{a}) := \#\{p_n \leq x : p_{n+i-1} \equiv a_i \pmod{q} \quad \forall 1 \leq i \leq M\}.$$

Any randomness-based model of the primes would suggest that M -tuples of consecutive primes equidistribute among the possibilities for \mathbf{a} , as is the case when $M = 1$. That is, one would expect that $\pi(x; q, \mathbf{a}) \sim \frac{\pi(x)}{\phi(q)^M}$ as $x \rightarrow \infty$. Lemke Oliver and Soundararajan [12] provide a heuristic argument based on the Hardy–Littlewood k -tuples conjectures for estimating $\pi(x; q, \mathbf{a})$ which agrees with this expectation (although it also predicts large second-order terms creating biases among the patterns).

However, little is known about $\pi(x; q, \mathbf{a})$ when $M \geq 2$. In most cases, it is not even known that $\pi(x; q, \mathbf{a})$ tends to infinity as $x \rightarrow \infty$ (i.e., it is not known that \mathbf{a} occurs infinitely often as a consecutive pattern in the sequence of primes mod q). If $\phi(q) = 2$ and $a_1 \not\equiv a_2 \pmod q$ are distinct reduced congruence classes, then $\pi(x; q, [a_1, a_2])$ and $\pi(x; q, [a_2, a_1])$ must each tend to infinity as an immediate consequence of Dirichlet's theorem for primes in arithmetic progressions; Knapowski and Turán [11] observed that if $\phi(q) = 2$, all four patterns of length 2 occur infinitely often.

As for arbitrary q , Shiu [19] used the Maier matrix method to prove that for any constant tuple \mathbf{a} of any length, $\pi(x; q, \mathbf{a})$ tends to infinity as $x \rightarrow \infty$. That is, for any fixed reduced residue class $a \pmod q$, there are infinitely many arbitrarily long strings of consecutive primes, all of which are congruent to $a \pmod q$. This result was rederived by Banks, Freiberg and Turnage-Butterbaugh [2] using new developments in sieve theory. Maynard [14] showed further that a positive density of primes begin strings of M consecutive primes, all of which are congruent to $a \pmod q$ – that is, that $\pi(x; q, \mathbf{a}) \gg \pi(x)$ whenever \mathbf{a} is a constant pattern.

It is not currently known that $\pi(x; q, \mathbf{a})$ tends to infinity for any other case, leading to the question of what more can be proven for other arithmetic sequences. In previous work [10], the authors considered the sequence of integer sums of two squares. Let \mathbf{E} denote the set of sums of two squares and let E_n denote the increasing sequence of sums of two squares, so that

$$\mathbf{E} = \{a^2 + b^2 : a, b \in \mathbb{Z}\} = \{E_n : n \in \mathbb{N}\}.$$

Let $N(x)$ denote the number of sums of two squares less than x . A number n is in \mathbf{E} if and only if every prime congruent to 3 mod 4 divides n to an even power; that is, if n factors as $n = \prod_p p^{e_p}$, then e_p is even whenever $p \equiv 3 \pmod 4$. For a modulus $q = \prod_p p^{e_p}$ and a congruence class $a \pmod q$, write $(a, q) = \prod_p p^{f_p}$, where $f_p \leq e_p$ for all p . There are infinitely many $n \in \mathbf{E}$ congruent to $a \pmod q$ if and only if the following two conditions hold:

- for any prime $p \equiv 3 \pmod 4$, f_p is either even or $f_p = e_p$, and
- if $e_2 - f_2 \geq 2$, then $\frac{a}{2^{f_2}} \not\equiv 3 \pmod 4$.

We will call a congruence class $a \pmod q$ **E-admissible** if it satisfies these conditions (i.e., if there exists a solution to $x^2 + y^2 \equiv a \pmod q$). For a modulus q , an integer $M \geq 1$, and an M -tuple $\mathbf{a} = [a_1, \dots, a_M]$ of **E-admissible** residue classes mod q , let

$$N(x; q, \mathbf{a}) := \#\{E_n \leq x : E_{n+i-1} \equiv a_i \pmod q \quad \forall 1 \leq i \leq M\}.$$

Just as in the prime case, one expects $N(x; q, \mathbf{a})$ to tend to infinity for any tuple of **E-admissible** residue classes, and in fact, one expects $N(x; q, \mathbf{a}) \gg N(x)$. In other words, one expects $N(x; q, \mathbf{a})$ to represent a positive proportion of sums of two squares. When the modulus $q \equiv 1 \pmod 4$ is a prime, David, Devin, Nam and Schlitt [3] develop heuristics for second-order terms in the asymptotics of $N(x; q, \mathbf{a})$ analogously to [12]. Their heuristics are based on the analog of the Hardy–Littlewood k -tuples conjecture in the setting of sums of two squares, which was developed in [4]. For \mathbf{a} of length 1, these second-order terms are reminiscent of Chebyshev's bias, and were considered by Gorodetsky in [7].

The authors [10] proved that for any modulus q , for any 3-tuple of \mathbf{E} -admissible residue classes $[a_1, a_2, a_3]$,

$$\lim_{x \rightarrow \infty} N(x; q, [a_1, a_2, a_3]) \rightarrow \infty.$$

They also showed that for any odd, squarefree modulus q , for any residues a_1 and a_2 with $(a_i, q) = 1$, for any tuple of the form $[a_1, \dots, a_1, a_2, \dots, a_2]$ (i.e., the concatenation of two constant tuples with values a_1 and a_2),

$$\lim_{x \rightarrow \infty} N(x; q, [a_1, \dots, a_1, a_2, \dots, a_2]) \rightarrow \infty. \quad (1)$$

Note that this result does not extend to all \mathbf{E} -admissible residue classes a_1 and a_2 .

In this paper, we strengthen (1) by proving the following theorem.

Theorem 1. *Let $q \geq 1$ be a squarefree odd modulus and let \tilde{a}_1 and \tilde{a}_2 be reduced residue classes modulo q . Let $M \geq 1$, and let $\mathbf{a} = [a_1, \dots, a_M]$ be a tuple of residue classes such that for some $1 \leq M_1 \leq M$, $a_i = \tilde{a}_1$ whenever $i \leq M_1$ and $a_i = \tilde{a}_2$ whenever $i > M_1$. Then $N(x; q, \mathbf{a}) \gg N(x)$.*

That is, any concatenation of two constant tuples appears with positive density among consecutive increasing sums of two squares modulo q .

Remark. Again, this result does not extend to all \mathbf{E} -admissible residue classes; \tilde{a}_1 and \tilde{a}_2 must be relatively prime to q . For squarefree odd q , in fact, all residue classes modulo q are \mathbf{E} -admissible. For fixed squarefree odd q , and for \tilde{a}_1, \tilde{a}_2 modulo q such that if $p | (\tilde{a}_i, q)$, then $p \equiv 1 \pmod{4}$, we expect our proof to apply with only minor adjustments in the computations of the technical results. We also expect that Theorem 1 extends with essentially no new ideas to the case where q is not squarefree, if substantially more care is taken on the background lemmas on evaluating sums of two squares in Section 3.3. Finally, our proof may apply essentially as written to the case where (\tilde{a}_i, q) is divisible by primes that are $3 \pmod{4}$. However, these should appear with a smaller (yet still positive) density (for example, there are more sums of two squares that are $1 \pmod{3}$ than that are $0 \pmod{3}$), and it may be that understanding the case when q is not squarefree is necessary for understanding this case.

The proof of Theorem 1 follows along the same basic idea as Maynard's result [14] that constant tuples appear with positive density among consecutive increasing primes. This work in turn expands on the work of Maynard [13], in which he shows that for any m , for any large enough k , and for any \mathcal{P} -admissible (that is, admissible in a precise sense with respect to the sequence of prime numbers) k -tuple of linear forms $\{L_1(n), \dots, L_k(n)\}$, there exist infinitely many n such that at least m of the $L_i(n)$ are simultaneously prime. In [14], for a tuple $\{L_1(n) = qn + a_1, \dots, L_k(n) = qn + a_k\}$ where each $L_i(n)$ is chosen such that $L_i(n) \equiv a \pmod{q}$ for all i , Maynard shows that for infinitely many n , at least m of the $L_i(n)$ are simultaneously prime *and* the numbers in between the outputs of the $L_i(n)$ have small prime factors (and thus are not themselves prime). He then averages over many such tuples of $L_i(n)$ in order to obtain a lower bound of positive density.

In the setting of sums of two squares, stronger sieving results are available than those that are available in the prime case. McGrath [15] showed that for any m , for large enough k , for any k -tuple $\{h_1, \dots, h_k\}$ which is \mathcal{P} -admissible, and for any partition of $\{h_1, \dots, h_k\}$ into m sub-tuples or ‘bins’, for infinitely many n , there exists an h_i in each bin such that $n + h_i \in \mathbf{E}$. Banks, Freiberg and Maynard [1] use a similar, but weaker, result in the case of primes to show that a positive proportion of real numbers are limit points of the sequence of normalized prime gaps, work which was refined in [17] and [16].

In order to prove Theorem 1, we strengthen the sieve result of McGrath [15] in the same way that Maynard [14] had expanded his previous work [13]. Our paper is organized as follows. In Section 2, we will state our sieve theoretic results and use them to prove Theorem 1. In Section 3, we will prove the sieve theoretic results. Our notation and setup is explained in Section 2.1, with an additional explanation of more technical sieve notation in Section 3.1. Finally, in Section 4, we evaluate certain averages of ‘singular series’ constants that appear in the proof of Theorem 1.

2. Statement of sieve results and proof of the main theorem

2.1. GPY sieve setup

Our argument will follow the Goldston–Pintz–Yıldırım method for detecting primes in \mathcal{P} -admissible k -tuples, building off of work of Maynard [14], which uses a rather sophisticated version of this method, and of McGrath [15], which develops a second-moment version of this method for sums of two squares.

An \mathcal{P} -admissible k -tuple of linear forms $(\ell_1(n), \dots, \ell_k(n))$ is one such that, for every prime p , there exists some $a \bmod p$ with $\ell_i(a) \not\equiv 0 \bmod p$ for all $1 \leq i \leq k$. Using the GPY method, Maynard [13] showed that for all integers $m \geq 2$, there exists large enough k such that for any \mathcal{P} -admissible k -tuple of linear forms $(\ell_1(n), \dots, \ell_k(n))$, there are many integers $n \geq 1$ for which at least m of the values $\ell_1(n), \dots, \ell_k(n)$ are simultaneously prime.

This statement follows from the construction of positive weights $w(n)$ such that for all x ,

$$\sum_{x \leq n < 2x} \left(\sum_{i=1}^k \mathbf{1}_{\mathcal{P}}(\ell_i(n)) - m + 1 \right) w(n) > 0, \quad (2)$$

where $\mathbf{1}_{\mathcal{P}}$ denotes the indicator function of the set \mathcal{P} of prime numbers. The inequality (2) implies that there exists a strictly positive summand, so that for some n with $x \leq n < 2x$,

$$\sum_{i=1}^k \mathbf{1}_{\mathcal{P}}(\ell_i(n)) > m - 1,$$

and thus, there are at least m primes among the values of $\ell_i(n)$.

We will require a version of this technique that is adapted in three different ways: first, we will detect sums of two squares instead of primes; second, we will need a ‘second moment’ adaptation to detect slightly more delicate patterns among the sequence of sums of two squares; and third, we will exclude certain values of n so that we will be able to average over many different k -tuples.

We begin by defining a certain weighted indicator function of sums of two squares. For any function f (say, the indicator functions $\mathbf{1}_{\mathcal{P}}$ or $\mathbf{1}_{\mathbf{E}}$), in practice, applying the ‘second moment’ adaptation requires an understanding of two-point correlations of the form

$$\sum_{x \leq n < 2x} f(\ell_i(n))f(\ell_j(n)).$$

Estimates for two-point correlations of the standard indicator function of sums of two squares are not known, so we will instead make use of Hooley’s ρ -function, which was first introduced in [8] and also used in this context by McGrath [15].

The ρ function is defined by

$$\rho(n) = r_2(n)t(n), \quad (3)$$

where $r_2(n)$ is the representation function of n , given by

$$\begin{aligned} r_2(n) &:= \#\{(x, y) \in \mathbb{Z}^2 : x^2 + y^2 = n\} \\ &= 4 \sum_{\substack{d|n \\ d \text{ odd}}} (-1)^{\frac{d-1}{2}} \end{aligned} \quad (4)$$

and

$$t(n) = t_{x, \theta_1}(n) := \sum_{\substack{a|n \\ a \leq v \\ p|a \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)}{g_2(a)} \left(1 - \frac{\log a}{\log v}\right), \quad (v = x^{\theta_1}). \quad (5)$$

Here, θ_1 is a fixed small constant with $\theta_1 < 1/18$; for example, Hooley takes $\theta_1 = 1/20$. Moreover, g_2 is the multiplicative function defined on primes via

$$g_2(p) = \begin{cases} 2 - \frac{1}{p} & \text{if } p \equiv 1 \pmod{4} \\ \frac{1}{p} & \text{if } p \equiv 3 \pmod{4}. \end{cases} \quad (6)$$

Using the indicator function ρ , McGrath [15] uses a second-moment bound to prove the existence of sums of two squares in different ‘bins’ of the same tuple. To state this precisely, fix $M, k \geq 1$, and let K denote the product $K = Mk$. Let $q \geq 1$ be a fixed odd integer, and fix a tuple \mathcal{H}^* of size K such that $4|h_i$, $(h_i, q) = 1$, and for $\ell_i(n) = qn + h_i$, the tuple of linear forms $\{\ell_1(n), \dots, \ell_K(n)\}$ is \mathcal{P} -admissible (indeed, McGrath’s result is phrased as requiring the tuple to be \mathcal{P} -admissible, *not* \mathbf{E} -admissible). Suppose further that we have a fixed partition $\mathcal{H} = B_1 \sqcup \dots \sqcup B_M$ where $|B_i| = k$ for all i . McGrath showed that there exists a real number $u \geq 1$ and a nonnegative weight function $w(n)$ such that for all sufficiently large x ,

$$\sum_{x < n \leq 2x} \left[u^2 - \sum_{i=1}^M \left(\sum_{\ell \in B_i} \rho(\ell(n)) - u \right)^2 \right] w(n) > 0. \quad (7)$$

The positivity of the left-hand side of (7) implies that for all sufficiently large x , there exists some n with $x < n \leq 2x$ such that

$$\sum_{i=1}^M \left(\sum_{\ell \in B_i} \rho(\ell(n)) - u \right)^2 < u^2.$$

If for any bin B_i , there is no $\ell \in B_i$ with $\ell(n) \in \mathbf{E}$, then $\sum_{\ell \in B_i} \rho(\ell(n)) = 0$, and thus,

$$u^2 \leq \left(\sum_{\ell \in B_i} \rho(\ell(n)) - u \right)^2 \leq \sum_{i=1}^M \left(\sum_{\ell \in B_i} \rho(\ell(n)) - u \right)^2 < u^2,$$

a contradiction. Thus, in particular, the inequality (7) implies that for all sufficiently large x , there exists an n with $x < n \leq 2x$ and such that for every bin B_i , there exists an $\ell \in B_i$ with $\ell(n) = qn + h \in \mathbf{E}$.

Our aim is to combine this second-moment version of the GPY sieve setup with the goal of excluding certain values of n for each tuple \mathcal{H}^* in order to be able to average over many different tuples. In particular, we will choose weights $w(n)$ such that for any n making a positive contribution to the left-hand side of (7), $\ell_i(n)$ does not have any ‘small’ prime factors $p \equiv 3 \pmod{4}$ for any of the ℓ_i , and for any $b \leq \eta\sqrt{\log x}$ which is not in \mathcal{H} (i.e., $b \neq h_i$), the integer $qn + b$ is divisible exactly once by some ‘small’ prime $p \equiv 3 \pmod{4}$. These may seem like artificial constraints to place on the values n , but in fact, n that do not satisfy these constraints are exceptionally rare; intuitively, although it cannot be proven explicitly, the weights $w(n)$ place emphasis on those n where all $\ell_i(n) \in \mathbf{E}$ (or close to it), and values $qn + b$ that are outside of the tuple are unlikely to be sums of two squares. In [14], Maynard takes advantage of a similar device to average over different subsets \mathcal{H}^* , which allows him to prove a lower bound of positive density on the tuples he is counting.

Our precise setup is as follows. As in the setup of Theorem 1, we let q be a fixed odd squarefree modulus, and we also fix the parameters M and two congruence classes \tilde{a}_1 and \tilde{a}_2 modulo q , as well as M_1 with $1 \leq M_1 \leq M$. We will consider tuples of length K , where $K = kM$, split into bins of size k . We define integers a_1, \dots, a_K as follows. For i with $1 \leq i \leq M_1 k$, we let a_i be the smallest positive integer with $a_i \equiv \tilde{a}_1 \pmod{q}$ and $a_i \equiv 1 \pmod{4}$, whereas for i with $M_1 k + 1 \leq i \leq K$, we let a_i be the second-smallest positive integer with $a_i \equiv \tilde{a}_2 \pmod{q}$ and $a_i \equiv 1 \pmod{4}$ (that is, $a_i - 4q$ is the smallest such positive integer). The values of a_i for $M_1 k + 1 \leq i \leq K$ are shifted by q to ensure that $a_{i_1} < a_{i_2}$ whenever $1 \leq i_1 \leq M_1 k < i_2 \leq K$. Note that there are only two distinct values for the a_i , but for ease of notation, we define K values a_i , even though these values are repetitive.

Then, for any tuple of integers $\mathbf{b} = (b_1, \dots, b_K)$ with $b_i \equiv 3 \pmod{4}$ and $3 \leq b_i \leq \frac{\eta}{q}\sqrt{\log x}$ for all i , we will define the K -tuple $\mathcal{L} = \mathcal{L}(\mathbf{b}) = \{\ell_i(n)\}_{i=1}^K$ of linear forms given by

$$\ell_i(n) := qn + a_i + qb_i. \quad (8)$$

Here, η is a positive constant to be set later. Note that the constraints on a_i and b_i modulo 4 imply that whenever $n \equiv 1 \pmod{4}$, we also have $\ell_i(n) \equiv 1 \pmod{4}$.

We will ultimately average over many different choices of \mathbf{b} . Our average will be taken over \mathbf{b} lying in a slightly restricted set of tuples \mathcal{B} , where we define

$$\mathcal{B} := \left\{ \mathbf{b} = (b_1, \dots, b_K) \left| \begin{array}{l} b_1 = 3 \\ b_i \equiv 3 \pmod{4}, \quad 3 \leq b_i \leq \frac{\eta}{2q} \sqrt{\log x} \quad \forall 2 \leq i \leq M_1 k \\ \frac{\eta}{2q} \sqrt{\log x} < b_i \leq \frac{\eta}{q} \sqrt{\log x} \quad \forall M_1 k < i \leq K \end{array} \right. \right\}. \quad (9)$$

The key consequence of this definition (along with the definition of the a_i 's) is that for any n , $\ell_{i_1}(n) < \ell_{i_2}(n)$ whenever $1 \leq i_1 \leq M_1 k$ and $M_1 k + 1 \leq i_2 \leq K$.

As described above, we will write $\mathcal{L} = B_1 \sqcup \dots \sqcup B_M$, where

$$B_i := \{\ell_{(i-1)k+1}(n), \dots, \ell_{ik}(n)\}. \quad (10)$$

The B_i , which we refer to as *bins*, partition the tuple \mathcal{L} into M bins, each of size k .

For certain real numbers $\xi, \eta > 0$ (to be fixed later), a certain real number u , and a nonnegative weight function $w_n(\mathcal{L})$, we consider a sum of the shape

$$\sum_{x < n \leq 2x} \left[u^2 - \sum_{i=1}^M \left(\sum_{\ell \in B_i} \rho(\ell(n)) - u \right)^2 - \sum_{j=1}^K \sum_{\substack{p < x^\xi \\ p \equiv 3 \pmod{4} \\ p | \ell_j(n)}} u^2 - \sum_{\substack{b \leq \eta \sqrt{\log x} \\ \ell^{(b)} \notin \mathcal{L}}} \mathbf{1}_{S(\xi)}(\ell^{(b)}(n)) u^2 \right] w_n(\mathcal{L}), \quad (11)$$

where $S(\xi)$ is the set of integers such that for all primes $p < x^\xi$ which satisfy $p \equiv 3 \pmod{4}$, either $p \nmid n$ or $p^2 | n$. We write $\ell^{(b)}(n) := qn + b$, so that the final sum in (11) is a sum over $b \leq \eta \sqrt{\log x}$ such that $\ell^{(b)} \notin \mathcal{L}$. A choice of weights $w_n(\mathcal{L})$ such that (11) is positive implies that for some n with $x < n \leq 2x$,

$$u^2 - \sum_{i=1}^M \left(\sum_{\ell \in B_i} \rho(\ell(n)) - u \right)^2 - \sum_{j=1}^K \sum_{\substack{p < x^\xi \\ p \equiv 3 \pmod{4} \\ p | \ell_j(n)}} u^2 - \sum_{\substack{b \leq \eta \sqrt{\log x} \\ \ell^{(b)} \notin \mathcal{L}}} \mathbf{1}_{S(\xi)}(\ell^{(b)}(n)) u^2 > 0,$$

which in turn implies that

- for each i , there exists a linear form $\ell \in B_i$ with $\rho(\ell(n)) \neq 0$, and thus, $\ell(n) \in \mathbf{E}$;
- for each j , with $1 \leq j \leq K$, $\ell_j(n)$ is not divisible by any prime $p < x^\xi$ with $p \equiv 3 \pmod{4}$; and
- for each $b \leq \eta \sqrt{\log x}$ with $\ell^{(b)}$ not in \mathcal{L} , we have $\ell^{(b)}(n) \notin S(\xi)$, so there exists some prime $p < x^\xi$ with $p \equiv 3 \pmod{4}$ such that $p | \ell^{(b)}(n)$.

In order to take advantage of this positivity argument, we will need to evaluate the sums over n appearing in (11). These evaluations are accomplished in Theorem 2, which we state in the next section before completing the proof of Theorem 1.

2.2. Conventions and notation

Before stating our main sieve theorem and presenting the proof of Theorem 1, we first fix some notation and conventions that we will use throughout the paper. An index for key quantities appears after the references.

All asymptotic notation, such as $O(\cdot)$, $o(\cdot)$, \ll , and \gg , should be interpreted as referring to the limit $x \rightarrow \infty$. We will use Vinogradov $f \ll g$ to mean $f = O(g)$; that is, $|f| \leq Cg$ for some absolute constant C . Any constants are absolute unless otherwise noted. For all sums or products over a variable p (or p'), the variable p will be assumed to lie in the prime numbers; all other sums and products will be assumed to be taken over variables lying in the natural numbers $\mathbb{N}_{\geq 1}$ unless otherwise specified.

Recall that the squarefree odd modulus q is fixed throughout. We denote $q = q_1 q_3$, where q_1 is a product of primes that are $1 \pmod{4}$ and q_3 is a product of primes that are $3 \pmod{4}$.

Let $\theta_2 > 0$ be a fixed positive real number such that $0 < \theta_1 + \theta_2 < 1/18$, and let $R = x^{\theta_2/2}$. Letting $D_0 = \eta\sqrt{\log x}$ for a constant $\eta > 0$ to be fixed later, we define

$$W = \prod_{\substack{p \leq D_0 \\ p \equiv 3 \pmod{4} \\ p \nmid q}} p. \quad (12)$$

Note that $q_3 W$ is the product of all primes $p \leq D_0$ which are $3 \pmod{4}$. This definition of W differs from that of McGrath [15] because, while the value of D_0 is much larger than that used by McGrath, it is not divisible by any primes $p \equiv 1 \pmod{4}$.

We denote by A the *Landau–Ramanujan constant*, given by

$$A = \frac{1}{\sqrt{2}} \prod_{p \equiv 3 \pmod{4}} \left(1 - \frac{1}{p^2}\right)^{-\frac{1}{2}} = \frac{\pi}{4} \prod_{p \equiv 1 \pmod{4}} \left(1 - \frac{1}{p^2}\right)^{1/2}. \quad (13)$$

We also make use of a normalization constant B , defined as

$$B = \frac{A}{\Gamma(1/2)\sqrt{L(1, \chi_4)}} \cdot \frac{\phi(q_3 W)(\log R)^{1/2}}{q_3 W} = \frac{2A}{\pi} \frac{\phi(q_3 W)(\log R)^{1/2}}{q_3 W}. \quad (14)$$

Here, χ_4 denotes the nontrivial Dirichlet character modulo 4. Finally, we will denote by V the constant given by

$$V = \prod_{p \equiv 1 \pmod{4}} \left(1 + \frac{1}{(2p-1)^2}\right) \approx 1.016. \quad (15)$$

For K -tuples in \mathbb{N}^K , we will use the notation that a boldface letter such as \mathbf{d} represents a tuple $\mathbf{d} = (d_1, \dots, d_K)$, whereas a non-boldface d represents the product of the entries $\prod_{i=1}^K d_i$. Given tuples \mathbf{d} and \mathbf{e} , we will let $[\mathbf{d}, \mathbf{e}]$ denote the product of the least common multiples $\prod_{i=1}^K [d_i, e_i]$, let (\mathbf{d}, \mathbf{e}) denote the product of the greatest common divisors $\prod_{i=1}^K (d_i, e_i)$, and let $\mathbf{d}|\mathbf{e}$ denote the K conditions that $d_i | e_i$ for $1 \leq i \leq K$.

2.3. Statement of the main sieve theorem

We are now ready to state our main sieving theorem, which we will use in the next section to deduce Theorem 1.

Theorem 2. Fix $\mathbf{b} \in \mathcal{B}$ and let $\mathcal{L}(\mathbf{b})$ be the fixed K -tuple of linear forms $\{\ell_i(n)\}_{i=1}^K$ given by (8). Let ν_0 be a fixed residue class modulo W such that for all $\ell \in \mathcal{L}$, $(\ell(\nu_0), W) = 1$.

Then there exists a choice of nonnegative weights $w_n(\mathcal{L}) \geq 0$, as well as a constant $L_K(F)$, such that

$$w_n(\mathcal{L}) \ll \left(\frac{\log R}{\log D_0} \right)^K \prod_{i=1}^K \prod_{\substack{p|\ell_i(n) \\ p \equiv 3 \pmod{4}}} 4 \quad (16)$$

and the following estimates hold:

a) Let $S_1(\nu_0)$ be the sum defined by

$$S_1(\nu_0) := \sum_{\substack{x < n \leq 2x \\ n \equiv 1 \pmod{4} \\ n \equiv \nu_0 \pmod{W}}} w_n(\mathcal{L}).$$

Then

$$S_1(\nu_0) = (1 + o(1)) \frac{B^K x}{4W} L_K(F). \quad (17)$$

b) Let $S_2^{(m)}(\nu_0)$ be the sum defined by

$$S_2^{(m)}(\nu_0) := \sum_{\substack{x < n \leq 2x \\ n \equiv 1 \pmod{4} \\ n \equiv \nu_0 \pmod{W}}} \rho(\ell_m(n)) w_n(\mathcal{L}).$$

Then

$$S_2^{(m)}(\nu_0) = (1 + o(1)) \frac{4\pi \sqrt{\frac{\log R}{\log v}} B^K x}{(\pi + 2)\sqrt{KW}} L_K(F). \quad (18)$$

c) Let $S_3^{(m_1, m_2)}(\nu_0)$ be the sum defined by

$$S_3^{(m_1, m_2)}(\nu_0) := \sum_{\substack{x < n \leq 2x \\ n \equiv 1 \pmod{4} \\ n \equiv \nu_0 \pmod{W}}} \rho(\ell_{m_1}(n)) \rho(\ell_{m_2}(n)) w_n(\mathcal{L}).$$

Then

$$S_3^{(m_1, m_2)}(\nu_0) \leq (1 + o(1)) \frac{64\pi^2 \frac{\log R}{\log v} B^K x}{(\pi + 2)^2 KW} V L_K(F), \quad (19)$$

where V is the constant defined in (15).

d) Let $S_4^{(m)}(\nu_0)$ be the sum defined by

$$S_4^{(m)}(\nu_0) := \sum_{\substack{x < n \leq 2x \\ n \equiv 1 \pmod{4} \\ n \equiv \nu_0 \pmod{W}}} \rho(\ell_m(n))^2 w_n(\mathcal{L}).$$

Then

$$S_4^{(m)}(\nu_0) = (1 + o(1)) \frac{8\pi \sqrt{\frac{\log R}{\log v}} \left(\frac{\log x}{\log v} + 1 \right) B^K x}{(\pi + 2)\sqrt{KW}} \prod_{\substack{p \equiv 1 \pmod{4} \\ p \nmid q_1}} \left(1 + \frac{1}{(2p-1)^2} \right) L_K(F). \quad (20)$$

e) Assume that $\xi > 0$ satisfies $\xi < \frac{1}{K}$, and let $S_5^{(m)}(\nu_0)$ be the sum defined by

$$S_5^{(m)}(\nu_0) := \sum_{\substack{x < n \leq 2x \\ n \equiv 1 \pmod{4} \\ n \equiv \nu_0 \pmod{W}}} \sum_{\substack{p < x^\xi \\ p \equiv 3 \pmod{4} \\ p \mid \ell_m(n)}} w_n(\mathcal{L}).$$

Then $S_5^{(m)}(\nu_0)$ satisfies

$$S_5^{(m)}(\nu_0) \ll \frac{K^2 \xi^2}{\theta_2^2} \frac{B^K x}{W} L_K(F). \quad (21)$$

f) Let ν_1 be a congruence class modulo $q_3^2 W^2$ such that $(\ell(\nu_1), q_3^2 W^2)$ is a square for all $\ell \in \mathcal{L}$. Fix $3 < b \leq \eta \sqrt{\log x}$ and a constant ξ with $0 < \xi < 1/4$. Let $S_6^{(b)}(\nu_1)$ be defined by

$$S_6^{(b)}(\nu_1) := \sum_{\substack{x < n \leq 2x \\ n \equiv 1 \pmod{4} \\ n \equiv \nu_1 \pmod{q_3^2 W^2}}} \mathbf{1}_{S(\xi)}(\ell^{(b)}(n)) w_n(\mathcal{L}).$$

Then

$$S_6^{(b)}(\nu_1) \ll_K \frac{x}{4q_3^2 W^2} \xi^{-1/2} \left(\frac{\theta_2}{2} \right)^{-1/2} \left(\frac{\log R}{\log D_0} \right)^{\frac{K-1}{2}} L_K(F). \quad (22)$$

This theorem is key in all of our computations and will be proven in Section 3. In the remainder of this section, we derive our main result as a consequence of Theorem 2.

2.4. Proof of Theorem 1

The goal of this subsection is to prove Theorem 1 as a consequence of Theorem 2 and the evaluations of the linear functionals therein.

We will consider an average of \mathbf{E} -admissible tuples $\mathcal{L} = \mathcal{L}(\mathbf{b}) = \{\ell_i(n)\}_{i=1}^K$, given by (8), over the set \mathcal{B} (defined in (9)) of K -tuples \mathbf{b} . We consider the sum

$$S = \sum_{\substack{\mathbf{b} \in \mathcal{B} \\ \mathcal{L} = \mathcal{L}(\mathbf{b}) \text{ } \mathbf{E}\text{-admissible}}} \sum_{\substack{\nu_1 \bmod q_3^2 W^2 \\ (\ell(\nu_1), W) = 1 \forall \ell \in \mathcal{L}}} \sum_{\substack{x < n \leq 2x \\ n \equiv \nu_1 \bmod q_3^2 W^2}} \left[u^2 - \sum_{i=1}^M \left(\sum_{\ell \in B_i} \rho(\ell(n)) - u \right)^2 \right. \\ \left. - \sum_{i=1}^K \sum_{\substack{p < x^\xi \\ p \equiv 3 \bmod 4 \\ p | \ell_i(n)}} u^2 - \sum_{\substack{b \leq \eta \sqrt{\log x} \\ \ell^{(b)} \notin \mathcal{L} \\ (\ell^{(b)}(\nu_1), q_3^2 W^2) = \square}} \mathbf{1}_{S(\xi)}(\ell^{(b)}(n)) u^2 \right] w_n(\mathcal{L}). \quad (23)$$

For technical reasons involving the final sum, we will initially sum over congruence classes modulo $q_3^2 W^2$ instead of modulo W . However, note that the condition that $(\ell(\nu_1), W) = 1$ is determined only by the congruence class of $\nu_1 \bmod W$, so this is in some sense really a sum over congruence classes modulo W .

Here, $w_n(\mathcal{L})$ are the weights given by Theorem 2 for the \mathbf{E} -admissible set $\mathcal{L} = \mathcal{L}(\mathbf{b})$. For fixed \mathcal{L} , the term in the square parentheses in (23) is positive only if the following conditions all hold:

- (i) for each i with $1 \leq i \leq M$, there exists some $\ell \in B_i$ with $\rho(\ell(n)) \neq 0$, or equivalently with $\ell(n) \in \mathbf{E}$;
- (ii) for each $\ell \in \mathcal{L}$, $\ell(n)$ has no prime factors p with $p < x^\xi$ and $p \equiv 3 \bmod 4$; and
- (iii) for all other $\ell^{(b)} \notin \mathcal{L}$ with $b \leq \eta \sqrt{\log x}$, and $(\ell^{(b)}(n), q_3^2 W^2)$ a square, $\ell^{(b)}(n)$ has a prime factor p with $p < x^\xi$, $p \equiv 3 \bmod 4$, and $p | \ell^{(b)}(n)$.

This has two crucial implications. One is that no n can make a positive contribution from two different tuples \mathcal{L} , since if n makes a positive contribution for any \mathcal{L} , then the values $\ell(n)$ are uniquely determined as the integers in $[qn, qn + \eta \sqrt{\log x}]$ which are

- (i) congruent to 1 mod 4,
- (ii) congruent to $\tilde{a}_1 \bmod q$ if they lie in $[qn + a_1, qn + a_1 + (\eta/2)\sqrt{\log x}]$, or congruent to $\tilde{a}_2 \bmod q$ if they lie in $[qn + a_K + (\eta/2)\sqrt{\log x}, qn + a_K + \sqrt{\log x}]$, and
- (iii) not divisible to an odd power by any primes $p < x^\xi$ with $p \equiv 3 \bmod 4$.

The second observation is that if n makes a positive contribution for a tuple \mathcal{L} , then since for all $\ell^{(b)} \notin \mathcal{L}$ with $b \leq \eta \sqrt{\log x}$, $\ell^{(b)}(n) = qn + b \notin \mathbf{E}$, we have that the sums of two squares appearing in \mathcal{L} (of which there is at least one in each bin) must be *consecutive* sums of two squares.

Also, if n makes a positive contribution, then none of the $\ell_i(n)$ can have any prime factors $p \equiv 3 \bmod 4$ which are less than x^ξ , so each $\ell_i(n)$ can have at most $O(1/\xi)$ prime

factors $p \equiv 3 \pmod{4}$. In particular, this implies by (16) that

$$w_n(\mathcal{L}) \ll \left(\frac{\log R}{\log D_0} \right)^K \prod_{i=1}^K \prod_{\substack{p|\ell_i(n) \\ p \equiv 3 \pmod{4}}} 4 \ll \left(\frac{\log R}{\log D_0} \right)^K \exp(O(K/\xi)), \quad (24)$$

for any pair n and \mathcal{L} making a positive contribution to (23).

We now evaluate the sum in (23). To begin with, we can swap the order of summation for the various different terms to get

$$\begin{aligned} S = & \sum_{\substack{\mathbf{b} \in \mathcal{B} \\ \mathcal{L} = \mathcal{L}(\mathbf{b}) \text{ adm.}}} \left[\sum_{\substack{\nu_0 \bmod W \\ (\ell(\nu_0), W) = 1 \forall \ell \in \mathcal{L}}} \left(u^2(1-M)S_1(\nu_0) + 2u \sum_{m=1}^K S_2^{(m)}(\nu_0) \right. \right. \\ & - \sum_{i=1}^M \sum_{\substack{\ell_{m_1}, \ell_{m_2} \in B_i \\ m_1 \neq m_2}} S_3^{(m_1, m_2)}(\nu_0) - \sum_{m=1}^K S_4^{(m)}(\nu_0) - u^2 \sum_{i=1}^K S_5^{(i)}(\nu_0) \Big) \\ & \left. - u^2 \sum_{\substack{\nu_1 \bmod q_3^2 W^2 \\ (\ell(\nu_1), W) = 1 \forall \ell \in \mathcal{L}}} \sum_{\substack{b \leq \eta \sqrt{\log x} \\ \ell^{(b)} \notin \mathcal{L} \\ (\ell^{(b)}(\nu_1), q_3^2 W^2) = \square}} S_6^{(b)}(\nu_1) \right], \end{aligned} \quad (25)$$

where the sums $S_1(\nu_0), S_2^{(m)}(\nu_0), S_3^{(m_1, m_2)}(\nu_0), S_4^{(m)}(\nu_0), S_5(\nu_0)$ and $S_6^{(b)}(\nu_1)$ are in the notation of Theorem 2.

We now wish to use our estimates from Theorem 2. For the sum $S_6^{(b)}(\nu_1)$, we will require a more careful analysis that takes the averaging over \mathbf{b}, ν_1, b into account. Specifically, we require the following lemma, which is proven in Section 4.2.

Lemma 3. *With the notation above,*

$$\begin{aligned} & \sum_{\substack{\mathbf{b} \in \mathcal{B} \\ \mathcal{L} = \mathcal{L}(\mathbf{b}) \text{ adm.}}} \sum_{\substack{\nu_1 \bmod q_3^2 W^2 \\ (\ell(\nu_1), W) = 1 \forall \ell \in \mathcal{L}}} \sum_{\substack{b \leq \eta \sqrt{\log x} \\ \ell^{(b)} \notin \mathcal{L} \\ (\ell^{(b)}(\nu_1), q_3^2 W^2) = \square}} S_6^{(b)}(\nu_1) \\ & \ll_K \xi^{-1/2} \theta_2^{\frac{K}{2}-1} L_K(F) \left(\frac{q_3}{\phi(q_3)} \right)^K \frac{\eta^K}{q_3^2 q^{K-1}} \frac{(\log x)^{K-\frac{1}{2}}}{(\log D_0)^K} x, \end{aligned}$$

where the implied constant depends only on K .

Applying the estimates (17), (18), (19), (20) and (21) from Theorem 2 to (25), we get

$$\begin{aligned}
 S \geq & (1+o(1)) \frac{B^K x}{W} L_K(F) \sum_{\substack{\mathbf{b} \in \mathcal{B} \\ \mathcal{L}=\mathcal{L}(\mathbf{b}) \text{ adm.}}} \sum_{\substack{\nu_0 \bmod W \\ (\ell(\nu_0), W)=1 \\ \forall \ell \in \mathcal{L}}} \left[u^2(1-M) \frac{1}{4} + 2u \sum_{m=1}^K \frac{4\pi \sqrt{\frac{\log R}{\log v}}}{(\pi+2)\sqrt{K}} \right. \\
 & - \sum_{i=1}^M \sum_{\substack{\ell_{m_1}, \ell_{m_2} \in B_i \\ m_1 \neq m_2}} \frac{64\pi^2 \left(\frac{\log R}{\log v}\right)}{(\pi+2)^2 K} V - \sum_{m=1}^K \frac{8\pi \sqrt{\frac{\log R}{\log v}} \left(\frac{\log x}{\log v} + 1\right)}{(\pi+2)\sqrt{K}} \prod_{\substack{p \equiv 1 \bmod 4 \\ p \nmid q_1}} \left(1 + \frac{1}{(2p-1)^2}\right) \\
 & \left. - u^2 \sum_{i=1}^K O\left(\frac{K^2 \xi^2}{4} L_K(F)\right) \right] - u^2 \sum_{\substack{\mathbf{b} \in \mathcal{B} \\ \mathcal{L}=\mathcal{L}(\mathbf{b}) \text{ adm.}}} \sum_{\substack{\nu_1 \bmod q_3^2 W^2 \\ (\ell(\nu_1), W)=1 \\ \forall \ell \in \mathcal{L}}} \sum_{\substack{b \leq \eta \sqrt{\log x} \\ \ell^{(b)} \notin \mathcal{L} \\ (\ell^{(b)}(\nu_1), q_3^2 W^2) = \square}} S_6^{(b)}(\nu_1).
 \end{aligned} \tag{26}$$

We now use Lemma 3 to evaluate the last triple sum, and simplify using the facts that $\log R = \frac{\theta_2}{2} \log x$ and $\log v = \theta_1 \log x$, which gives

$$\begin{aligned}
 S \geq & (1+o(1)) \left(\sum_{\substack{\mathbf{b} \in \mathcal{B} \\ \mathcal{L}=\mathcal{L}(\mathbf{b}) \text{ adm.}}} \sum_{\substack{\nu_0 \bmod W \\ (\ell(\nu_0), W)=1 \forall \ell \in \mathcal{L}}} 1 \right) \frac{B^K x}{W} L_K(F) \left[\frac{u^2(1-M)}{4} + \frac{8u\pi\sqrt{K}\sqrt{\theta_2/2\theta_1}}{\pi+2} \right. \\
 & - M \frac{k(k+1)}{2} \frac{64\pi^2(\theta_2/2\theta_1)}{(\pi+2)^2 K} V - \frac{8\pi\sqrt{K}\sqrt{\theta_2/2\theta_1} \left(\frac{1}{\theta_1} + 1\right)}{(\pi+2)} \prod_{\substack{p \equiv 1 \bmod 4 \\ p \nmid q_1}} \left(1 + \frac{1}{(2p-1)^2}\right) \\
 & \left. - O(u^2 K^3 \xi^2) \right] - O\left(u^2 \xi^{-1/2} \theta_2^{\frac{K}{2}-1} L_K(F) \left(\frac{q_3}{\phi(q_3)}\right)^K \frac{\eta^K}{q_3^2 q^{K-1}} \frac{(\log x)^{K-\frac{1}{2}}}{(\log D_0)^K} x\right).
 \end{aligned}$$

We will make the change of variables

$$u = \frac{\pi}{\pi+2} \sqrt{\frac{\theta_2}{2\theta_1}} \tilde{u},$$

so that the sum above simplifies to

$$S \geq (1+o(1)) \left(\sum_{\substack{\mathbf{b} \in \mathcal{B} \\ \mathcal{L}=\mathcal{L}(\mathbf{b})}} \sum_{\substack{\nu_0 \bmod W \\ \text{adm. } (\ell(\nu_0), W)=1 \forall \ell \in \mathcal{L}}} 1 \right) \frac{B^K x}{W} L_K(F) \left[\left(\frac{\pi}{\pi+2} \right)^2 \frac{\theta_2}{2\theta_1} \right. \\ \left. \left(\frac{\tilde{u}^2(1-M)}{4} + 8\sqrt{K}\tilde{u} - 32V \left(\frac{K}{M} + 1 \right) \right) - O_{\theta_1, \theta_2}(\sqrt{K}) - O(u^2 K^3 \xi^2) \right] - \\ O \left(u^2 \xi^{-1/2} \theta_2^{\frac{K}{2}-1} L_K(F) \left(\frac{q_3}{\phi(q_3)} \right)^K \frac{\eta^K}{q_3^2 q^{K-1}} \frac{(\log x)^{K-\frac{1}{2}}}{(\log D_0)^K} x \right).$$

We then set $\tilde{u} = \frac{16\sqrt{K}}{M-1}$ to maximize the expression above, so that (recalling that $K = Mk$)

$$S \geq (1+o(1)) \left(\sum_{\substack{\mathbf{b} \in \mathcal{B} \\ \mathcal{L}=\mathcal{L}(\mathbf{b})}} \sum_{\substack{\nu_0 \bmod W \\ \text{adm. } (\ell(\nu_0), W)=1 \forall \ell \in \mathcal{L}}} 1 \right) \frac{B^K x}{W} L_K(F) \\ \times \left[\left(\frac{\pi}{\pi+2} \right)^2 \left(\frac{\theta_2}{2\theta_1} \right) \cdot 32 \left(k \frac{(2-V)M+V}{M-1} - V \right) - O_{\theta_1, \theta_2}(\sqrt{K}) - O(u^2 K^3 \xi^2) \right] \\ - O \left(u^2 \xi^{-1/2} \theta_2^{\frac{K}{2}-1} L_K(F) \left(\frac{q_3}{\phi(q_3)} \right)^K \frac{\eta^K}{q_3^2 q^{K-1}} \frac{(\log x)^{K-\frac{1}{2}}}{(\log D_0)^K} x \right).$$

Recall that $V \approx 1.016 < 2$, so for a given M , we can pick k large enough in terms of M , θ_1 and θ_2 so that the quantity

$$\Delta = \left(\frac{\pi}{\pi+2} \right)^2 \left(\frac{\theta_2}{2\theta_1} \right) \cdot 32 \left(K \frac{(2-V)M+V}{M(M-1)} - V \right) - O_{\theta_1, \theta_2}(\sqrt{K})$$

will be positive. We can then pick the constant ξ to be a small enough multiple of K^{-4} so that the term $O(u^2 K^3 \xi^2)$ will be negligible (for example, smaller than $\frac{\Delta}{100}$). Note that this is consistent with the constraint from the evaluation of $S_5^{(m)}$ that $\xi < \frac{1}{K}$.

By Lemma 21, the sums over \mathbf{b} and ν_0 are bounded below by

$$\sum_{\substack{\mathbf{b} \in \mathcal{B} \\ \mathcal{L}=\mathcal{L}(\mathbf{b})}} \sum_{\substack{\nu_0 \bmod W \\ \text{adm. } (\ell(\nu_0), W)=1 \forall \ell \in \mathcal{L}}} 1 \gg_K \left(\frac{\eta}{q} \right)^{K-1} (\log x)^{\frac{K-1}{2}} W \left(\frac{\phi(W)}{W} \right)^K.$$

Thus, by definition of B ,

$$\frac{B^K x}{W} \sum_{\substack{\mathbf{b} \in \mathcal{B} \\ \mathcal{L} = \mathcal{L}(\mathbf{b})}} \sum_{\substack{\nu_0 \bmod W \\ \text{adm. } (\ell(\nu_0), W) = 1 \forall \ell \in \mathcal{L}}} 1 \gg_K \left(\frac{\eta}{q}\right)^{K-1} \left(\frac{q_3}{\phi(q_3)}\right)^K \frac{(\log x)^{K-\frac{1}{2}}}{(\log D_0)^K} x.$$

Returning to S , we have that

$$\begin{aligned} S &\gg_K \left(\frac{\eta}{q}\right)^{K-1} \left(\frac{q_3}{\phi(q_3)}\right)^K L_K(F) \left(\Delta - \frac{1}{100}\Delta\right) \frac{(\log x)^{K-\frac{1}{2}}}{(\log D_0)^K} x \\ &\quad + O\left(u^2 \xi^{-1/2} \theta_2^{\frac{K}{2}-1} L_K(F) \left(\frac{q_3}{\phi(q_3)}\right)^K \frac{\eta^K}{q_3^2 q^{K-1}} \frac{(\log x)^{K-\frac{1}{2}}}{(\log D_0)^K} x\right). \end{aligned}$$

We can now set the parameter η to be sufficiently small (in terms of $K, M, \theta_1, \theta_2, \xi$) such that the big-O term will be negligible, which implies that

$$S \gg_{K, M, \theta_1, \theta_2, \xi, \eta} \left(\frac{1}{q}\right)^{K-1} \left(\frac{q_3}{\phi(q_3)}\right)^K \frac{(\log x)^{K-\frac{1}{2}}}{(\log D_0)^K} x. \quad (27)$$

Equation (24) implies that

$$S \ll \#\{E_n \leq x : E_{n+i-1} \equiv a_i \bmod q \forall 1 \leq i \leq M\} \times \exp(O(K/\xi)) \left(\frac{\log R}{\log D_0}\right)^K,$$

which along with equation (27) and the fact that $\log R = \frac{\theta_2}{2} \log x$, implies that

$$\#\{E_n \leq x : E_{n+i-1} \equiv a_i \bmod q \forall 1 \leq i \leq M\} \gg_{K, M, \theta_1, \theta_2, \xi, \eta} \left(\frac{1}{q}\right)^{K-1} \left(\frac{q_3}{\phi(q_3)}\right)^K \frac{x}{\sqrt{\log x}}.$$

This completes the proof.

3. Proofs of sieve results

The goal of this section is to prove Theorem 2. Throughout, fix $\eta > 0$ and let $\mathcal{L} = \{\ell_i(n)\}_{1 \leq i \leq K}$ be a fixed tuple of linear forms $\ell_i(n) = qn + a_i + qb_i$, where $qb_i \leq \eta\sqrt{\log x}$ for all i . Let ν_0 be a congruence class modulo W such that $(\ell(\nu_0), W) = 1$ for all $\ell \in \mathcal{L}$.

This section will be organized as follows. In Section 3.1, we introduce notation that will be used throughout, and define the sieve weights $w_n(\mathcal{L})$. Sections 3.2 and 3.3 contain lemmas and computations that will be used throughout the proof of Theorem 2; the estimate (16) is proven in Lemma 8. Finally, equations (17), (18), (19), (20), (21) and (22) are proven (respectively) in Sections 3.4, 3.5, 3.6, 3.7, 3.8 and 3.9, which completes the proof of Theorem 2.

3.1. Sieve notation and setup

We begin by fixing some notation in preparation for defining the weights $w_n(\mathcal{L})$. Recall that W is the product of primes $p \equiv 3 \bmod 4$ satisfying $p \leq D_0$ and $(p, q) = 1$. In particular, this means that if a prime $p \equiv 3 \bmod 4$ satisfies $p|\ell_i(n)$ and $p|\ell_j(n)$ for $x < n \leq 2x$ and for two distinct linear forms $\ell_i, \ell_j \in \mathcal{L}$, then $p|q_3 W$. Let $\mathcal{D}_K \subset \mathbb{Z}^K$ denote the set of K -tuples

$\mathbf{d} = (d_i)$ such that for all i , $(d_i, q_3 W) = 1$, such that $(d_i, d_j) = 1$ for all $i \neq j$, and such that each d_i is divisible only by primes congruent to 3 mod 4.

Let $F: [0, 1]^K \rightarrow \mathbb{R}$ be a smooth function defined as follows. Let $R_K = \{(x_1, \dots, x_K) \in [0, 1]^K : \sum_{i=1}^K x_i \leq 1\}$. Define $F(t_1, \dots, t_K)$ as

$$F(t_1, \dots, t_K) := \prod_{i=1}^K g(Kt_i), \quad (28)$$

where

$$g(t) = \begin{cases} \frac{1}{1+t}, & \text{if } t \leq 1, \\ 0, & \text{otherwise.} \end{cases} \quad (29)$$

Note that F is supported on the set R_K .

We are now ready to define the sieve weights $w_n(\mathcal{L})$, which are nearly identical in structure to the multi-dimensional Selberg sieve weights used in, among other papers, [13] and [15]. We define

$$w_n(\mathcal{L}) = \left(\sum_{\substack{\mathbf{d} \in \mathcal{D}_K \\ d_i | \ell_i(n)}} \lambda_{\mathbf{d}} \right)^2, \quad (30)$$

where

$$\lambda_{\mathbf{d}} = \left(\prod_{i=1}^K \mu(d_i) d_i \right) \sum_{\substack{\mathbf{r} \in \mathcal{D}_K \\ d_i | r_i \forall i}} \frac{\mu(r)^2}{\phi(r)} F\left(\frac{\log r_1}{\log R}, \dots, \frac{\log r_K}{\log R}\right). \quad (31)$$

We will write

$$y_{\mathbf{r}} := F\left(\frac{\log r_1}{\log R}, \dots, \frac{\log r_K}{\log R}\right) \quad (32)$$

where F is a function defined in (28).

The results of our sieve evaluations will depend on the following functionals on F :

$$\begin{aligned} L_K(F) &:= \int_0^1 \cdots \int_0^1 [F(x_1, \dots, x_K)]^2 \prod_{i=1}^K \frac{dx_i}{\sqrt{x_i}}, \\ L_{K;m}(F) &:= \int_0^1 \cdots \int_0^1 \left[\int_0^1 F(x_1, \dots, x_K) \frac{dx_m}{\sqrt{x_m}} \right]^2 \prod_{\substack{i=1 \\ i \neq m}}^K \frac{dx_i}{\sqrt{x_i}}, \\ L_{K;m_1, m_2}(F) &:= \int_0^1 \cdots \int_0^1 \left[\int_0^1 \left(\int_0^1 F(x_1, \dots, x_K) \frac{dx_{m_1}}{\sqrt{x_{m_1}}} \right) \frac{dx_{m_2}}{\sqrt{x_{m_2}}} \right]^2 \prod_{\substack{i=1 \\ i \neq m_1, m_2}}^K \frac{dx_i}{\sqrt{x_i}}. \end{aligned} \quad (33)$$

Using the function F that is explicitly given by (28) and (29), we can evaluate each of $L_K(F)$, $L_{K;m}(F)$, and $L_{K;m_1, m_2}(F)$; each of these will be a constant depending only on K . More convenient, however, is using the following lemma, which relates each of these values to $L_K(F)$.

Lemma 4 ([15], Lemma 6.4: Evaluation of sieve functionals). *Let $F(t_1, \dots, t_K)$ be given by equation (28) and let $L_K(F)$, $L_{K;m_1}(F)$, and $L_{K;m_1,m_2}(F)$ denote the functionals defined in (33). Then for any m_1, m_2 ,*

$$\frac{L_{K;m_1}(F)}{L_K(F)} = \frac{\pi^2}{\pi+2} \sqrt{\frac{1}{K}} \quad \text{and} \quad \frac{L_{K;m_1,m_2}(F)}{L_K(F)} = \left(\frac{\pi^2}{\pi+2} \right)^2 \frac{1}{K}.$$

3.2. Auxiliary lemmas for sieve weights

This subsection and the next collect various lemmas that will be used throughout our estimates. To begin with, we present several lemmas concerning the sieve weights defined in Section 3.1.

Lemma 5.

- (i) *Let $\mathbf{r}, \mathbf{s} \in \mathcal{D}_K$ with $s_i = r_i$ for all $i \neq j$ and $s_j = Ar_j$ for some $A \in \mathbb{N}$. Then for $y_{\mathbf{r}}$ and $y_{\mathbf{s}}$ defined by (32), we have*

$$y_{\mathbf{s}} = y_{\mathbf{r}} + O\left(K \frac{\log A}{\log R} y_{\mathbf{r}}\right).$$

- (ii) *Let $\mathbf{r}, \mathbf{s} \in \mathcal{D}_K$ with $r = s$ and let A be the product of primes dividing r but not (\mathbf{r}, \mathbf{s}) . Then for $y_{\mathbf{r}}$ and $y_{\mathbf{s}}$ defined by (32), we have*

$$y_{\mathbf{s}} = y_{\mathbf{r}} + O\left(K \frac{\log A}{\log R} (y_{\mathbf{r}} + y_{\mathbf{s}})\right).$$

Proof. Recall that $y_{\mathbf{r}} = F\left(\frac{\log r_1}{\log R}, \dots, \frac{\log r_K}{\log R}\right)$, where $F(t_1, \dots, t_K) := \prod_{i=1}^K g(Kt_i)$ and $g(t) = \frac{1}{1+t}$ for $t \leq 1$ and $g(t) = 0$ otherwise. Given $u, v \geq 0$ with $|u - v| < \varepsilon$, we have

$$\frac{1}{1+Ku} = \frac{1+O(K\varepsilon)}{1+Kv}.$$

Let $u_i = \log r_i / \log R$, $v_i = \log s_i / \log R$ and $\varepsilon_i = v_i - u_i$. In part (i), $\varepsilon_i = 0$ for $i \neq j$ and $\varepsilon_j = \log A / \log R$. Thus,

$$\frac{1}{1+Kv_j} = \frac{1+O\left(\frac{K \log A}{\log R}\right)}{1+Ku_j}.$$

Multiplying by $\prod_{i \neq j} 1/(1+Ku_i)$ gives the result for (i).

Now consider part (ii), and let \mathbf{t} be the vector with $t_i = [r_i, s_i]$. Applying part (i) to each component in turn implies that

$$y_{\mathbf{s}} = y_{\mathbf{t}} + O\left(K y_{\mathbf{s}} \sum_{i=1}^K \frac{\log([r_i, s_i]) / s_i}{\log R}\right) = y_{\mathbf{t}} + O\left(K y_{\mathbf{s}} \frac{\log A}{\log R}\right).$$

The same holds for \mathbf{r} and \mathbf{t} , which implies (ii). □

The following lemma is a standard evaluation of sums of multiplicative functions that appear frequently in sieve computations.

Lemma 6. Let $A_1, A_2, L > 0$ and let γ be a multiplicative function satisfying

$$0 \leq \frac{\gamma(p)}{p} \leq 1 - \frac{1}{A_1}$$

and

$$-L \leq \sum_{w \leq p < z} \frac{\gamma(p) \log p}{p} - \frac{1}{2} \log \left(\frac{z}{w} \right) \leq A_2$$

for $2 \leq w \leq z$. Let g be the multiplicative function defined by $g(p) = \frac{\gamma(p)}{p - \gamma(p)}$, and let $G : [0, 1] \rightarrow \mathbb{R}$ be a piece-wise differentiable function. Then

$$\sum_{d < z} \mu^2(d) g(d) G \left(\frac{\log d}{\log z} \right) = c_\gamma \frac{(\log z)^{\frac{1}{2}}}{\Gamma(1/2)} \int_0^1 G(t) \frac{dt}{\sqrt{t}} + O \left(c_\gamma L G_{\max} (\log z)^{-\frac{1}{2}} \right),$$

where

$$c_\gamma = \prod_p \left(1 - \frac{\gamma(p)}{p} \right)^{-1} \left(1 - \frac{1}{p} \right)^{\frac{1}{2}}$$

and

$$G_{\max} = \sup_{t \in [0, 1]} (|G(t)| + |G'(t)|).$$

Proof. This is [6, Lemma 4] with $\kappa = \frac{1}{2}$. □

Using Lemma 6, we show the following lemma, which is similar to [14, Lemma 8.4].

Lemma 7. Let Q be a squarefree modulus of the form $Q = q_3 W \alpha$ with $\alpha = R^{O(K)}$. Let f be a multiplicative function with $f(p) = p + O(K)$, and let $G : \mathbb{R} \rightarrow \mathbb{R}$ be a smooth decreasing function supported on $[0, 1]$.

Then for K sufficiently large, we have

$$\begin{aligned} & \sum_{\substack{\mathbf{e} \in \mathcal{D}_K \\ (e, Q)=1}} \frac{\mu^2(e)}{f(e)} \prod_{i=1}^K G \left(\frac{\log e_i}{\log R} \right) \\ &= (1 + o(1)) \left(\frac{\phi(\alpha)}{\alpha} \right)^{\frac{K}{2}} \left(\frac{e^{-\gamma/2}}{\Gamma(1/2)} \right)^K \left(\frac{\log R}{\log D_0} \right)^{K/2} \left(\int_{t_1, \dots, t_K \geq 0} \prod_{i=1}^K G(t_i) \frac{dt_i}{\sqrt{t_i}} \right). \end{aligned}$$

Proof. We would like to apply Lemma 6 for each variable. However, the variables e_i are not independent, since $\mathbf{e} \in \mathcal{D}_K$ implies $(e_i, e_j) = 1$ for $i \neq j$. Our first task is to decouple these variables with a negligible penalty.

Denote the sum evaluated in the lemma by

$$S = \sum_{\substack{\mathbf{e} \in \mathcal{D}_K \\ (e, Q)=1}} \frac{\mu^2(e)}{f(e)} \prod_{i=1}^K G \left(\frac{\log e_i}{\log R} \right).$$

If $p|(e_r, e_j)$ for some $1 \leq r < j \leq K$, then we must have $p > D_0$ since $(e, Q) = 1$. Thus, using also the fact that G is decreasing, we have that

$$\begin{aligned} & \sum_{\substack{\mathbf{e} \in \mathbb{N}_K \\ (e, Q) = 1 \\ p|(e_i, e_j)}} \prod_{j=1}^K \frac{\mu^2(e_j)}{f(e_j)} \prod_{i=1}^K G\left(\frac{\log e_i}{\log R}\right) \\ & \ll \frac{1}{f(p)^2} \sum_{\substack{\mathbf{e} \in \mathbb{N}_K \\ (e, Q) = 1}} \prod_{j=1}^K \frac{\mu^2(e_j)}{f(e_j)} G\left(\frac{\log e_r + \log p}{\log R}\right) G\left(\frac{\log e_j + \log p}{\log R}\right) \prod_{\substack{i=1 \\ i \neq r, j}}^K G\left(\frac{\log e_i}{\log R}\right) \\ & \ll (p + O(K))^{-2} S. \end{aligned}$$

Writing

$$S' = \sum_{\substack{\mathbf{e} \in \mathbb{N}_K \\ (e, Q) = 1}} \prod_{j=1}^K \frac{\mu^2(e_j)}{f(e_j)} \prod_{i=1}^K G\left(\frac{\log e_i}{\log R}\right) = \left(\sum_{\substack{n \in \mathbb{N} \\ (n, Q) = 1}} \frac{\mu^2(n)}{f(n)} G\left(\frac{\log n}{\log R}\right) \right)^K,$$

it follows that

$$S - S' \ll S \binom{K}{2} \sum_{p > D_0} \frac{1}{(p + O(K))^2} \ll_K \frac{S}{D_0}.$$

We now consider S' in place of S and apply Lemma 6, where we take $g(p) = \frac{1}{f(p)}$ and $\gamma(p) = \frac{p}{f(p)+1}$ for $p \nmid Q$ and $g(p) = \gamma(p) = 0$ for $p|Q$. For this, we need a bound on the constant L from Lemma 6. If we did not have the restriction $g(p) = \gamma(p) = 0$ for $p|Q$, then using the prime number theorem in arithmetic progressions, we could take L to be a constant, since

$$\sum_{\substack{w \leq p < z \\ p \equiv 3 \pmod{4}}} \frac{\log p}{p + O(K)} - \frac{1}{2} \log\left(\frac{z}{w}\right) \ll_K 1.$$

It follows that in our case, we can choose L satisfying

$$L \ll_K 1 + \sum_{\substack{p \leq D_0 \\ p \equiv 3 \pmod{4}}} \frac{\log p}{p} + \sum_{p|\alpha} \frac{\log p}{p}.$$

The first sum is $(1 + o(1)) \frac{1}{2} \log D_0$. For the second sum, we have the bound

$$\sum_{p|\alpha} \frac{\log p}{p} \ll_K \log \log R$$

since $\alpha = R^{O(K)}$ and the sum is dominated by taking the smallest possible primes. This implies that

$$L \ll_K \frac{1}{2} \log D_0 + \log \log R \ll_K \log \log R.$$

We will write

$$\Omega_G = G_{\max} \cdot \left(\int_{t \geq 0} G(t) \frac{dt}{\sqrt{t}} \right)^{-1},$$

where G_{\max} is defined as in Lemma 6. Applying Lemma 6 successively K times, we get

$$\begin{aligned} S' &= \frac{c_\gamma^K}{\Gamma(1/2)^K} (\log R)^{K/2} \left(\int_{t_1, \dots, t_K \geq 0} \prod_{i=1}^K G(t_i) \frac{dt_i}{\sqrt{t_i}} \right) \left(1 + \sum_{\ell=1}^K \binom{K}{\ell} O_K \left(\frac{\Omega_G \log \log R}{\log R} \right)^\ell \right) \\ &= \frac{c_\gamma^K}{\Gamma(1/2)^K} (\log R)^{K/2} \left(\int_{t_1, \dots, t_K \geq 0} \prod_{i=1}^K G(t_i) \frac{dt_i}{\sqrt{t_i}} \right) \left(1 + O_{K,G} \left(\frac{\log \log R}{\log R} \right) \right), \end{aligned}$$

where c_γ is given in Lemma 6 and satisfies

$$c_\gamma = (1 + o(1)) \prod_{p|\alpha} \left(1 - \frac{1}{p} \right)^{\frac{1}{2}} \prod_{p \leq D_0} \left(1 - \frac{1}{p} \right)^{\frac{1}{2}} = (1 + o(1)) \left(\frac{\phi(\alpha)}{\alpha} \right)^{\frac{1}{2}} \frac{e^{-\gamma/2}}{(\log D_0)^{\frac{1}{2}}}$$

Combining our estimates completes the proof. \square

Lemma 7 will be useful for many of our computations to follow. For now, we use it to verify that the weights $w_n(\mathcal{L})$ given in (30) are bounded above as in (16), which we show in the following lemma.

Lemma 8. *We have the bounds*

(i)

$$|\lambda_{\mathbf{d}}| \ll_K \left(\frac{\log R}{\log D_0} \right)^{K/2}, \quad (34)$$

(ii)

$$w_n(\mathcal{L}) \ll_K \left(\frac{\log R}{\log D_0} \right)^K \prod_{i=1}^K \prod_{\substack{p|\ell_i(n) \\ p \equiv 3 \pmod{4}}} 4,$$

(iii)

$$w_n(\mathcal{L}) \ll_K R^{2+o(1)}. \quad (35)$$

Lemma 8, (ii) is precisely (16), whereas (i) and (iii) will be used in our sieve estimates.

Proof. By our choice of $y_{\mathbf{r}}$, for any $\mathbf{d} \in \mathcal{D}_K$, we have

$$|\lambda_{\mathbf{d}}| = \prod_{i=1}^K d_i \sum_{\substack{\mathbf{r} \in \mathcal{D}_K \\ \mathbf{d}|\mathbf{r}}} \frac{y_{\mathbf{r}}}{\phi(r)} = \frac{d}{\phi(d)} \sum_{\mathbf{d}|\mathbf{r} \in \mathcal{D}_K} \frac{F\left(\frac{\log r_1}{\log R}, \dots, \frac{\log r_K}{\log R}\right)}{\phi(r/d)}. \quad (36)$$

Since F is decreasing in each argument, we can bound $|\lambda_{\mathbf{d}}|$ above by replacing each $\frac{\log r_i}{\log R}$ in the argument of F with $\sigma_i = \frac{(\log r_i/d_i)}{\log R}$.

We now apply Lemma 7 with $Q = dq_3W$, which gives

$$\begin{aligned} \sum_{\mathbf{d}|\mathbf{r} \in \mathcal{D}_K} \frac{F(\sigma_1, \dots, \sigma_K)}{\phi(r/d)} &= \sum_{\mathbf{d}|\mathbf{r} \in \mathcal{D}_K} \frac{1}{\phi(r/d)} \prod_{i=1}^K g(K\sigma_i) \\ &\sim \left(\frac{\phi(d)}{d}\right)^{\frac{K}{2}} \left(\frac{e^{-\gamma/2}}{\Gamma(1/2)}\right)^K \left(\frac{\log R}{\log D_0}\right)^{K/2} \left(\int_{t_1, \dots, t_K \geq 0} \prod_{i=1}^K g(Kt_i) \frac{dt_i}{\sqrt{t_i}}\right). \end{aligned}$$

Evaluating the integral gives

$$\int_{t_1, \dots, t_K \geq 0} \prod_{i=1}^K g(Kt_i) \frac{dt_i}{\sqrt{t_i}} = \left(\frac{\pi}{2\sqrt{K}}\right)^K.$$

Substituting this expression back into (36), we get that

$$|\lambda_{\mathbf{d}}| \ll_K \left(\frac{\phi(d)}{d}\right)^{\frac{K}{2}-1} \left(\frac{\log R}{\log D_0}\right)^{K/2} \ll \left(\frac{\log R}{\log D_0}\right)^{K/2}$$

for $K \geq 4$. This completes the proof of (34).

For the second claim, recall that

$$\begin{aligned} w_n(\mathcal{L}) &= \left(\sum_{\substack{\mathbf{d} \in \mathcal{D}_K \\ d_i | \ell_i(n)}} \lambda_{\mathbf{d}}\right)^2 \ll_K \left(\frac{\log R}{\log D_0}\right)^K \#\{\mathbf{d} \in \mathcal{D}_K : d_i | \ell_i(n)\}^2 \\ &\ll_K \left(\frac{\log R}{\log D_0}\right)^K \prod_{i=1}^K \prod_{\substack{p | \ell_i(n) \\ p \equiv 3 \pmod{4}}} 4, \end{aligned}$$

as desired.

For the third claim, note that since $\lambda_{\mathbf{d}}$ is supported on $d = d_1 \cdots d_K < R$, we have

$$w_n(\mathcal{L}) \ll_K \left(\frac{\log R}{\log D_0}\right)^K \left(\sum_{d_1 \cdots d_K < R} 1\right)^2 \ll_K R^{2+o(1)} \left(\sum_{d_1 \cdots d_K < R} \frac{1}{d_1 \cdots d_K}\right)^2 \ll_K R^{2+o(1)}.$$

□

Finally, throughout our sieve arguments, we will make crucial use of [15, Lemma 6.6], which we restate below for convenience.

Lemma 9. *Let $J \subseteq \{1, \dots, K\}$ (possibly empty) and $p_1, p_2 \in \mathbb{P} \cup \{1\}$ be fixed. Write $I = \{1, \dots, K\} \setminus J$. Define the sieve sum $S_{J, p_1, p_2, m} = S_{J, p_1, p_2, m, f, g}$ by*

$$S_{J, p_1, p_2, m} = \sum_{\substack{\mathbf{d}, \mathbf{e} \in \mathcal{D}_K \\ q_3W, [d_1, e_1], \dots, [d_K, e_K] \text{ coprime} \\ p_1 | d_m, p_2 | e_m}} \lambda_{\mathbf{d}} \lambda_{\mathbf{e}} \prod_{i \in I} f([d_i, e_i]) \prod_{j \in J} g([d_j, e_j]),$$

with weights $\lambda_{\mathbf{d}}$ defined as in (30). If $J = \emptyset$, we define $f(p) = 1/p$ (and in this case, there is no dependence on g in the sum). Otherwise, f and g are nonzero multiplicative

functions defined on primes by

$$f(p) = \frac{1}{p} + O\left(\frac{1}{p^2}\right), \quad g(p) = \frac{1}{p^2} + O\left(\frac{1}{p^3}\right),$$

and moreover, we assume that $f(p) \neq 1/p$. We write S_J for $S_{J,1,1,m}$. Then for $|J| \in \{0, 1, 2\}$, we have the following:

(i) If $m \in J$, then

$$S_{J,p_1,p_2,m} \ll \frac{F_{\max}^2 B^{K+|J|} (\log \log R)^2}{(p_1 p_2 / (p_1, p_2))^2}.$$

(ii) If $m \notin J$, then

$$S_{J,p_1,p_2,m} \ll \frac{F_{\max}^2 B^{K+|J|} (\log \log R)^2}{p_1 p_2 / (p_1, p_2)}.$$

(iii) We have

$$S_J = (1 + o(1)) B^{K+|J|} L_J(F),$$

where we write $L_J(F)$ as shorthand for $L_{K;j \in J}(F)$, defined analogously to (33).

Remark. This is [15, Lemma 6.6] with $q_3 W$ in place of W . Note also that the restriction $\mathbf{d}, \mathbf{e} \in \mathcal{D}_K$ in the definition of $S_{J,p_1,p_2,m}$ does not affect the statement since $\lambda_{\mathbf{d}} = 0$ for $\mathbf{d} \notin \mathcal{D}_K$.

3.3. Auxiliary lemmas for sums of two squares

In this subsection, we record several useful results on estimates of the functions ρ and r_2 . To begin with, we have the following lemma, which is [15, Lemma 5.3], and will be used in the proof of (18).

Lemma 10. Let $(\alpha, r) = (d, r) = 1$, where d and r are squarefree, odd and $\ll x^{O(1)}$. Then

$$\sum_{\substack{n \leq x \\ n \equiv \alpha \pmod{r} \\ n \equiv 1 \pmod{4} \\ d|n}} r_2(n) = \frac{g_1(r) g_2(d)}{2rd} \pi x + R_1(x; d, r),$$

where g_2 is defined as in (6), g_1 is the multiplicative function defined on primes by

$$g_1(p) = 1 - \chi(p)/p,$$

and

$$R_1(x; d, r) \ll_{\varepsilon} ((rd)^{\frac{1}{2}} + x^{\frac{1}{3}}) d^{\frac{1}{2}} x^{\varepsilon}.$$

The following lemma is nearly identical to [15, Lemma A.3] and will be used to prove (19).

Lemma 11. Let r be a modulus and suppose that $(\alpha, r) = (\alpha + h, r) = (c_1, r) = (c_2, r) = 1$ and $4|h$, where c_1, c_2, r are squarefree and odd, and of size $\ll x^{O(1)}$. Then we have

$$\sum_{\substack{n \leq x \\ n \equiv \alpha \pmod{r} \\ n \equiv 1 \pmod{4} \\ c_1 | n \\ c_2 | n+h}} r_2(n)r_2(n+h) = \frac{g_1(r)^2 \Gamma(h, c_1, c_2, r)}{r} \pi^2 x + R_2(x; c_1, c_2, r),$$

where

$$\Gamma(h, c_1, c_2, r) = \frac{g_2(c_1)g_2(c_2)}{c_1 c_2} \sum_{(t, 2r)=1} \frac{c_t(h)(c_1, t)(c_2, t)\chi[(c_1^2, t)]\chi[(c_2^2, t)]}{t^2 \Psi(c_1, t)\Psi(c_2, t)},$$

where $\Psi(u, t) := g_2((u, t/(u, t)))$, $c_t(h)$ is the Ramanujan sum $c_t(h) = \sum_{d|(t, h)} \mu\left(\frac{t}{d}\right) d$, and

$$R_2(x; c_1, c_2, r) \ll_{\varepsilon} r^{\frac{1}{2}} c_1 c_2 x^{\frac{3}{4}+\varepsilon} + c_1^{\frac{1}{2}} c_2^{\frac{1}{2}} x^{\frac{5}{6}+\varepsilon}.$$

The only difference between Lemma 11 and [15, Lemma A.3] is that the latter requires the additional constraint that $(c_1, c_2) = 1$. However, this constraint is not used in the proof, which refers heavily to the proof of [18, Lemma 4]. Note that if $(c_1, c_2) \nmid h$, then the sum over $r_2(n)r_2(n+h)$ is empty; in this case, $\Gamma(h, c_1, c_2, r) = 0$ and the equation still holds. One can see that $\Gamma(h, c_1, c_2, r) = 0$ when $(c_1, c_2) \nmid h$ by noting that Γ is multiplicative over primes $p|[h, c_1, c_2, r]$ and that $c_t(h) = \mu(t)$ whenever $(t, h) = 1$. Then if any prime p divides (c_1, c_2) but $p \nmid h$, the p -component of $\Gamma(h, c_1, c_2, r)$ is 0.

The following lemma is [15, Lemma 5.5] and will be used in the proof of (20).

Lemma 12. Let $(a, r) = (d, r) = 1$, where d and r are squarefree, odd, and $\ll x^{O(1)}$. Then

$$\sum_{\substack{n \leq x \\ n \equiv a \pmod{r} \\ n \equiv 1 \pmod{4} \\ d | n}} r_2^2(n) = \frac{g_3(r)g_4(d)}{rd} \left(\log x + A_2 + 2 \sum_{p|r} g_5(p) - 2 \sum_{p|d} g_6(p) \right) x + O_{\varepsilon} \left(r x^{\frac{3}{4}+\varepsilon} \right),$$

where

$$\begin{aligned} g_3(p) &:= \begin{cases} \frac{(p-1)^2}{p(p+1)} & \text{if } p \equiv 1 \pmod{4} \\ g_1(p) & \text{if } p \equiv 3 \pmod{4} \end{cases} & g_4(p) &:= \begin{cases} \frac{4p^2-3p+1}{p(p+1)} & \text{if } p \equiv 1 \pmod{4} \\ g_2(p) & \text{if } p \equiv 3 \pmod{4} \end{cases} \\ g_5(p) &:= \begin{cases} \frac{(2p+1)\log p}{p^2-1} & \text{if } p \equiv 1 \pmod{4} \\ \frac{\log p}{p^2-1} & \text{if } p \equiv 3 \pmod{4} \end{cases} & g_6(p) &:= \begin{cases} \frac{(p-1)^2(2p+1)\log p}{(p+1)(4p^2-3p+1)} & \text{if } p \equiv 1 \pmod{4} \\ \log p & \text{if } p \equiv 3 \pmod{4} \end{cases} \end{aligned}$$

and A_2 is a constant given explicitly in [15, Lemma 5.5].

The following lemma is [15, Lemma 5.6], and will be used to prove (18), (19) and (20).

Lemma 13. Let Q be a squarefree integer such that $Q \mid \prod_{p \leq (\log \log x)^3} p$ and any prime p dividing Q is $1 \pmod{4}$. Define $X_{x,Q}$, $Z_{x,Q}^{(1)}$, and $Z_{x,Q}^{(2)}$ as follows:

$$\begin{aligned} X_{x,Q} &= \sum_{\substack{a \leq v \\ (a,Q)=1 \\ p|a \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)}{a} \log \frac{v}{a}, \\ Z_{x,Q}^{(1)} &= \sum_{\substack{a,b \leq v \\ (a,Q)=(b,Q)=1 \\ p|ab \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)\mu(b)g_4([a,b])}{g_2(a)g_2(b)[a,b]} \log \frac{v}{a} \log \frac{v}{b} \\ Z_{x,Q}^{(2)} &= \sum_{\substack{a,b \leq v \\ (a,Q)=(b,Q)=1 \\ p|ab \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)\mu(b)g_4([a,b])}{g_2(a)g_2(b)[a,b]} \log \frac{v}{a} \log \frac{v}{b} \sum_{p|ab} g_6(p). \end{aligned}$$

Then

$$\begin{aligned} X_{x,Q} &= (1+o(1)) \frac{8A \log^{\frac{1}{2}} v}{\pi g_1(Q)}, \\ Z_{x,Q}^{(1)} &= (1+o(1)) 8A \frac{g_7(Q) \log^{\frac{1}{2}} v}{\phi(Q) g_1(Q)} \prod_{\substack{p \equiv 1 \pmod{4} \\ p \nmid Q}} \left(1 + \frac{1}{(2p-1)^2}\right), \text{ and} \\ Z_{x,Q}^{(2)} &= -(1+o(1)) 4A \frac{g_7(Q) \log^{\frac{3}{2}} v}{\phi(Q) g_1(Q)} \prod_{\substack{p \equiv 1 \pmod{4} \\ p \nmid Q}} \left(1 + \frac{1}{(2p-1)^2}\right), \end{aligned}$$

where g_7 is the multiplicative function defined on primes by $g_7(p) = p+1$.

Proof. The proof of this lemma is for the most part identical to the proof in Appendix B of [15], so here we will restrict ourselves to highlighting the differences in the argument.

In general, the application of the Selberg–Delange method is identical to that described in [15], with the same arguments applying for bounding, for example, relevant analytic functions; the only change that need be made to McGrath’s arguments is replacing W by an arbitrary Q (which must divide the W that McGrath uses) everywhere. Thus, by following McGrath’s proof, we get that, for

$$K_1(s)^2 = \left(1 - \frac{1}{2^s}\right)^{-1} \prod_{p \equiv 3 \pmod{4}} \left(1 - \frac{1}{p^{2s}}\right)^{-1}$$

and

$$G_1(Q, s) = \prod_{\substack{p|Q \\ p \equiv 1 \pmod{4}}} \left(1 - \frac{1}{p^s}\right)^{-1},$$

we have

$$X_{x,Q} = \frac{K_1(1)G_1(Q)}{\Gamma(3/2)\sqrt{L(1,\chi_4)}}(\log v)^{\frac{1}{2}} + O_\varepsilon\left(\frac{(\log x)^\varepsilon}{(\log v)^{\frac{1}{2}}}\right),$$

which simplifies to the desired expression.

In the same manner, expressions for $Z_{x,Q}^{(1)}$ and $Z_{x,Q}^{(2)}$ can be derived. \square

3.4. Estimating S_1

In this section, we will prove Theorem 2, equation (17), which we restate in the following proposition.

Proposition 14. *Consider the sum $S_1(\nu_0)$ defined by*

$$S_1(\nu_0) := \sum_{\substack{x \leq n \leq 2x \\ n \equiv 1 \pmod{4} \\ n \equiv \nu_0 \pmod{W}}} w_n(\mathcal{L}).$$

Then

$$S_1(\nu_0) = (1 + o(1)) \frac{B^K x}{4W} L_K(F).$$

Proof. We have

$$S_1(\nu_0) = \sum_{\substack{x \leq n \leq 2x \\ n \equiv 1 \pmod{4} \\ n \equiv \nu_0 \pmod{W}}} w_n(\mathcal{L}) = \sum_{\substack{x \leq n \leq 2x \\ n \equiv 1 \pmod{4} \\ n \equiv \nu_0 \pmod{W}}} \left(\sum_{\mathbf{d} \in \mathcal{D}_K} \lambda_{\mathbf{d}} \right)^2.$$

By expanding the square and swapping the order of summation, we get that

$$\begin{aligned} S_1(\nu_0) &= \sum_{\mathbf{d}, \mathbf{e} \in \mathcal{D}_K} \lambda_{\mathbf{d}} \lambda_{\mathbf{e}} \sum_{\substack{x \leq n \leq 2x \\ qn + a_i + qb_i \equiv 0 \pmod{[d_i, e_i]} \\ n \equiv 1 \pmod{4} \\ n \equiv \nu_0 \pmod{W}}} 1 \\ &= \sum_{\mathbf{d}, \mathbf{e} \in \mathcal{D}_K} \lambda_{\mathbf{d}} \lambda_{\mathbf{e}} \left(\frac{x}{4W} \prod_{i=1}^K \frac{1}{[d_i, e_i]} + O(1) \right) \\ &= \frac{x}{4W} \sum_{\mathbf{d}, \mathbf{e} \in \mathcal{D}_K} \frac{\lambda_{\mathbf{d}} \lambda_{\mathbf{e}}}{[\mathbf{d}, \mathbf{e}]} + O\left(\sum_{\mathbf{d}, \mathbf{e} \in \mathcal{D}_K} \lambda_{\mathbf{d}} \lambda_{\mathbf{e}} \right), \end{aligned}$$

where the second line follows from the first by the Chinese remainder theorem and the observation that 4, W and the $[d_i, e_i]$ are all pairwise relatively prime. In the notation of Lemma 9, the sum in the main term is precisely the sum $S_{\emptyset} = S_{\emptyset, 1, 1, m}$, which by Lemma 9 is equal to $B^K L_K(F)$.

It remains only to bound the error term; We have

$$\left| \sum_{\mathbf{d}, \mathbf{e} \in \mathcal{D}_K} \lambda_{\mathbf{d}} \lambda_{\mathbf{e}} \right| \ll \lambda_{\max}^2 |\mathcal{D}_K|^2.$$

We use (34) to bound λ_{\max} , and the bound

$$|\mathcal{D}_K| \leq \sum_{n \leq R} \tau_K(n) \ll_{\epsilon} R^{1+\epsilon}. \quad (37)$$

Since $R = x^{\theta_2/2}$ with $\theta_2 < \frac{1}{18}$, we conclude

$$\left| \sum_{\mathbf{d}, \mathbf{e} \in \mathcal{D}_K} \lambda_{\mathbf{d}} \lambda_{\mathbf{e}} \right| \ll_{K, \epsilon} R^{2+\epsilon} = x^{\theta_2+\epsilon} \ll x^{1/18+\epsilon},$$

which is negligible. This completes the proof. \square

3.5. Estimating $S_2^{(m)}$

In this section, we will prove Theorem 2, equation (18), which we restate in the following proposition.

Proposition 15. For fixed $1 \leq m \leq K$, consider the sum $S_2^{(m)}(\nu_0)$ defined by

$$S_2^{(m)}(\nu_0) := \sum_{\substack{x < n \leq 2x \\ n \equiv 1 \pmod{4} \\ n \equiv \nu_0 \pmod{W}}} \rho(\ell_m(n)) w_n(\mathcal{L}).$$

Then

$$S_2^{(m)}(\nu_0) = (1 + o(1)) \frac{4\pi \sqrt{\frac{\log R}{\log v}} B^K x}{(\pi + 2)\sqrt{KW}} L_K(F).$$

Proof. By definition of ρ , we have that

$$\rho(\ell_m(n)) = \frac{r_2(\ell_m(n))}{\log v} \sum_{\substack{a | \ell_m(n) \\ p | a \Rightarrow p \equiv 1 \pmod{4} \\ a \leq v}} \frac{\mu(a)}{g_2(a)} \log\left(\frac{v}{a}\right).$$

Note that since $a | \ell_m(n)$, which is relatively prime to g , we must have $(a, q) = 1$ for any a in the definition of $\rho(\ell_m(n))$. Since a is only divisible by primes that are $1 \pmod{4}$, the only nontrivial constraint on a is that $(a, q_1) = 1$. By expanding the definitions of $w_n(\mathcal{L})$ and ρ in the expression for $S_2^{(m)}(\nu_0)$ and changing the order of summation, we get

$$S_2^{(m)}(\nu_0) = \frac{1}{\log v} \sum_{\mathbf{d}, \mathbf{e} \in \mathcal{D}_K} \lambda_{\mathbf{d}} \lambda_{\mathbf{e}} \sum_{\substack{a \leq v \\ (a, q_1) = 1 \\ p | a \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)}{g_2(a)} \log \frac{v}{a} \sum_{\substack{x < n \leq 2x \\ n \equiv 1 \pmod{4} \\ n \equiv \nu_0 \pmod{W} \\ [d_i, e_i] | \ell_i(n) \forall i \\ a | \ell_m(n)}} r_2(\ell_m(n)). \quad (38)$$

The parameter a is supported on integers whose prime factors are all 1 mod 4, whereas by the definition of \mathcal{D}_K , each $[d_i, e_i]$ is only divisible by primes that are 3 mod 4. Thus, the $[d_i, e_i]$'s and a are pairwise coprime. Also, all prime factors of each $[d_i, e_i]$ are larger than D_0 , whereas each $qb_i < D_0$, so each $[d_i, e_i]$ is coprime to each qb_i .

Thus, by Lemma 10, (applied with $r = qW \prod_{i \neq m} [d_i, e_i]$ and with $d = a[d_m, e_m]$), we get

$$\sum_{\substack{x < n \leq 2x \\ n \equiv 1 \pmod{4} \\ n \equiv \nu_0 \pmod{W} \\ [d_i, e_i] | \ell_i(n) \forall i \\ a | \ell_m(n)}} r_2(\ell_m(n)) = \frac{g_1\left(qW \prod_{i \neq m} [d_i, e_i]\right) g_2(a[d_m, e_m])}{2qW a \prod_i [d_i, e_i]} \pi q x + O_\epsilon\left(\left((WR^3 v)^{1/2} + x^{1/3}\right) v^{1/2} x^\epsilon\right). \quad (39)$$

Recall that $v = x^{\theta_1}$, $R^2 = x^{\theta_2}$, $\theta_1 + \theta_2 < 1/18$, and $W \ll_\varepsilon x^\varepsilon$ for all $\varepsilon > 0$. Thus,

$$S_2^{(m)}(\nu_0) = \frac{\pi x}{\log v} \sum_{\mathbf{d}, \mathbf{e} \in \mathcal{D}_K} \lambda_{\mathbf{d}} \lambda_{\mathbf{e}} \sum_{\substack{a \leq v \\ (a, q_1) = 1 \\ p | a \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)}{g_2(a)} \log \frac{v}{a} \frac{g_1\left(qW \prod_{i \neq m} [d_i, e_i]\right) g_2(a[d_m, e_m])}{2W a \prod_i [d_i, e_i]} + E,$$

where

$$|E| \ll_\varepsilon \frac{1}{\log v} \left| \sum_{\mathbf{d}, \mathbf{e} \in \mathcal{D}_K} \lambda_{\mathbf{d}} \lambda_{\mathbf{e}} \right| \sum_{\substack{a \leq v \\ p | a \Rightarrow p \equiv 1 \pmod{4}}} \frac{1}{g_2(a)} \log \frac{v}{a} x^{\frac{1}{3} + (\theta_1 + \theta_2)/2 + \varepsilon} \\ \ll \lambda_{\max}^2 |\mathcal{D}_K|^2 v x^{\frac{1}{3} + (\theta_1 + \theta_2)/2 + \varepsilon} \ll_{\varepsilon'} x^{\frac{1}{3} + (\theta_1 + \theta_2)\frac{3}{2} + \varepsilon'} \ll x^{\frac{5}{12} + \varepsilon'},$$

where we used (34) to bound the λ_{\max} and (37) for $|\mathcal{D}_K|$.

It remains to consider the main term, which is given by

$$\begin{aligned} & \frac{\pi x}{\log v} \sum_{\mathbf{d}, \mathbf{e} \in \mathcal{D}_K} \lambda_{\mathbf{d}} \lambda_{\mathbf{e}} \sum_{\substack{a \leq v \\ (a, q_1) = 1 \\ p | a \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)}{g_2(a)} \log \left(\frac{v}{a}\right) \frac{g_1\left(qW \prod_{i \neq m} [d_i, e_i]\right) g_2(a[d_m, e_m])}{2W a \prod_i [d_i, e_i]} \\ &= \frac{\pi x g_1(qW)}{2W \log v} \sum_{\mathbf{d}, \mathbf{e} \in \mathcal{D}_K} \lambda_{\mathbf{d}} \lambda_{\mathbf{e}} \frac{g_1\left(\prod_{i \neq m} [d_i, e_i]\right) g_2([d_m, e_m])}{\prod_i [d_i, e_i]} \sum_{\substack{a \leq v \\ (a, q_1) = 1 \\ p | a \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)}{a} \log \frac{v}{a}. \end{aligned}$$

The inside sum is exactly X_{x, q_1} , as defined in Lemma 13. Also, all primes dividing $[d_m, e_m]$ are congruent to 3 mod 4, so $g_2([d_m, e_m]) = \frac{1}{[d_m, e_m]}$. We can thus rewrite the

main term as

$$\frac{X_{x,q_1} \pi x g_1(qW)}{2W \log v} \sum_{\mathbf{d}, \mathbf{e} \in \mathcal{D}_K} \lambda_{\mathbf{d}} \lambda_{\mathbf{e}} \prod_{i \neq m} \left(\frac{g_1([d_i, e_i])}{[d_i, e_i]} \right) \frac{1}{[d_m, e_m]^2}. \quad (40)$$

By Lemma 13, $X_{x,q_1} = (1 + o(1)) \frac{8A\sqrt{\log v}}{g_1(q_1)\pi}$, and by Lemma 9 applied with $f(p) = \frac{g_1(p)}{p}$ and $g(p) = 1/p^2$,

$$\sum_{\mathbf{d}, \mathbf{e} \in \mathcal{D}_K} \lambda_{\mathbf{d}} \lambda_{\mathbf{e}} \prod_{i \neq m} \left(\frac{g_1([d_i, e_i])}{[d_i, e_i]} \right) \frac{1}{[d_m, e_m]^2} = S_{\{m\}} = B^{K+1} L_{K;m}(F) (1 + o(1)).$$

Then (40) can be rewritten as

$$\frac{\pi x g_1(qW)}{2W \log v} \cdot \frac{8A\sqrt{\log v}}{g_1(q_1)\pi} B^{K+1} L_{K;m}(F) (1 + o(1)).$$

By definition of g_1 ,

$$\frac{g_1(q_3 W) \phi(q_3 W)}{q_3 W} = \frac{1}{2A^2} + O\left(\frac{1}{D_0}\right). \quad (41)$$

Recalling also that $B = \frac{2A}{\pi} \frac{\phi(q_3 W)}{q_3 W} (\log R)^{\frac{1}{2}}$, we get

$$S_2^{(m)}(\nu_0) = (1 + o(1)) \frac{4 \sqrt{\frac{\log R}{\log v}} B^K x}{\pi W} L_{K;m}(F),$$

which, along with the fact from Lemma 4 that $L_{K;m}(F) = \frac{\pi^2}{\pi+2} \frac{L_K(F)}{\sqrt{K}}$, completes the proof. \square

3.6. Estimating $S_3^{(m_1, m_2)}$

In this section, we will prove Theorem 2, equation (19), which we restate in the following proposition.

Proposition 16. *For fixed $1 \leq m_1, m_2 \leq K$ lying in the same bin of the tuple \mathcal{L} , consider the sum $S_3^{(m_1, m_2)}(\nu_0)$ defined by*

$$S_3^{(m_1, m_2)}(\nu_0) := \sum_{\substack{x < n \leq 2x \\ n \equiv 1 \pmod{4} \\ n \equiv \nu_0 \pmod{W}}} \rho(\ell_{m_1}(n)) \rho(\ell_{m_2}(n)) w_n(\mathcal{L}).$$

Then

$$S_3^{(m_1, m_2)}(\nu_0) \leq (1 + o(1)) \frac{64\pi^2 \left(\frac{\log R}{\log v}\right) B^K x}{(\pi + 2)^2 K W} V L_K(F).$$

Proof. We begin by expanding the definition of $w_n(\mathcal{L})$ to rewrite $S_3^{(m_1, m_2)}(\nu_0)$ as

$$S_3^{(m_1, m_2)}(\nu_0) = \sum_{\mathbf{d}, \mathbf{e} \in \mathcal{D}_K} \lambda_{\mathbf{d}} \lambda_{\mathbf{e}} \sum_{\substack{x < n \leq 2x \\ n \equiv 1 \pmod{4} \\ n \equiv \nu_0 \pmod{W} \\ [d_i, e_i] | \ell_i(n) \forall i}} \rho(\ell_{m_1}(n)) \rho(\ell_{m_2}(n)).$$

Upon expanding the definition of ρ and changing the order of summation, the inside sum over n is equal to

$$\begin{aligned} & \frac{1}{\log^2 v} \sum_{\substack{a, b \leq v \\ (a, q_1) = (b, q_1) = 1 \\ p | a, b \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)\mu(b)}{g_2(a)g_2(b)} \log \frac{v}{a} \log \frac{v}{b} \\ & \times \sum_{\substack{x < n \leq 2x \\ n \equiv 1 \pmod{4} \\ n \equiv \nu_0 \pmod{W} \\ [d_i, e_i] | \ell_i(n) \forall i \\ a | (qn + a_{m_1} + qb_{m_1}) \\ b | (qn + a_{m_2} + qb_{m_2})}} r_2(qn + a_{m_1} + qb_{m_1}) r_2(qn + a_{m_2} + qb_{m_2}). \end{aligned}$$

Since $\ell_{m_1}(n)$ and $\ell_{m_2}(n)$ are always relatively prime to q , we must also always have $(a, q) = (b, q) = 1$; since a and b are only divisible by primes congruent to 1 mod 4, this is equivalent to the constraint that $(a, q_1) = (b, q_1) = 1$. For the inside sum to be nonzero, $[d_1, e_1], \dots, [d_K, e_K]$, and W must all be pairwise coprime, and each of these must be coprime to both a and b . Moreover, if any prime p divides (a, b) , then $p | q(b_{m_2} - b_{m_1})$. We thus have

$$\begin{aligned} S_3^{(m_1, m_2)}(\nu_0) &= \sum_{\mathbf{d}, \mathbf{e} \in \mathcal{D}_K} \frac{\lambda_{\mathbf{d}} \lambda_{\mathbf{e}}}{\log^2 v} \sum_{\substack{a, b \leq v \\ (a, q_1) = (b, q_1) = 1 \\ p | a, b \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)\mu(b)}{g_2(a)g_2(b)} \log \frac{v}{a} \log \frac{v}{b} \\ &\times \sum_{\substack{\ell_{m_1}(x) < n \leq \ell_{m_1}(2x) \\ n \equiv 1 \pmod{4} \\ n \equiv \alpha \pmod{qW \prod_{i \neq m_1, m_2} [d_i, e_i]} \\ a[d_{m_1}, e_{m_1}] | n \\ b[d_{m_2}, e_{m_2}] | n+h}} r_2(n) r_2(n+h), \end{aligned}$$

where α is relatively prime to $qW \prod_{i \neq m_1, m_2} [d_i, e_i]$ and $h = q(b_{m_2} - b_{m_1})$.

We now apply Lemma 11 to estimate the inner sum, taking $r = qW \prod_{i \neq m_1, m_2} [d_i, e_i]$, $c_1 = a[d_{m_1}, e_{m_1}]$ and $c_2 = b[d_{m_2}, e_{m_2}]$. Note that Lemma 11 does not require that c_1 and c_2 be relatively prime, but that the main term is 0 unless $(c_1, c_2) | h$. Thus,

$$\begin{aligned} S_3^{(m_1, m_2)}(\nu_0) &= \sum_{\mathbf{d}, \mathbf{e} \in \mathcal{D}_K} \frac{\lambda_{\mathbf{d}} \lambda_{\mathbf{e}}}{\log^2 v} \sum_{\substack{a, b \leq v \\ (a, q_1) = (b, q_1) = 1 \\ p | a, b \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)\mu(b)}{g_2(a)g_2(b)} \log \frac{v}{a} \log \frac{v}{b} \\ &\times \left[\pi^2 q x \frac{g_1^2 \left(qW \prod_{i \neq m_1, m_2} [d_i, e_i] \right)}{qW \prod_{i \neq m_1, m_2} [d_i, e_i]} \frac{g_2(a[d_{m_1}, e_{m_1}]) g_2(b[d_{m_2}, e_{m_2}])}{a[d_{m_1}, e_{m_1}] b[d_{m_2}, e_{m_2}]} \right] \end{aligned}$$

$$\begin{aligned}
& \times \sum_{(t, 2qW \prod_{i \neq m_1, m_2} [d_i, e_i])=1} \frac{c_t(h)}{t^2} \frac{(a, t)(b, t)}{\Psi(a, t)\Psi(b, t)} \prod_{i=1,2} \frac{([d_{m_i}, e_{m_i}], t) \chi([d_{m_i}, e_{m_i}]^2, t)}{\Psi([d_{m_i}, e_{m_i}], t)} \\
& + O\left(x^{\frac{5}{6}+\epsilon} a^{\frac{1}{2}} b^{\frac{1}{2}} [d_{m_1}, e_{m_1}]^{\frac{1}{2}} [d_{m_2}, e_{m_2}]^{\frac{1}{2}}\right) \\
& + O\left(x^{\frac{3}{4}+\epsilon} \left(qW \prod_{i \neq m_1, m_2} [d_i, e_i]\right)^{\frac{1}{2}} ab[d_{m_1}, e_{m_1}][d_{m_2}, e_{m_2}]\right).
\end{aligned}$$

Taking absolute values and noting that $d_{m_i}, e_{m_i} \ll R^{\frac{1}{K}}$, the first error term is bounded by

$$\lambda_{\max}^2 |\mathcal{D}_K|^2 \left(x^{\frac{5}{6}+\epsilon} v^{\frac{3}{2}} R^{\frac{1}{K}} + x^{\frac{3}{4}+\epsilon} v^2 R^2 \right) \ll \lambda_{\max}^2 |\mathcal{D}_K|^2 x^{\frac{5}{6}} v^2 R^2.$$

Using (34) to bound λ_{\max} , (37) for $|\mathcal{D}_K|$, and the fact that $\theta_1 + \theta_2 < \frac{1}{18}$, we get that the error term is bounded by

$$\ll_{\epsilon'} x^{\frac{5}{6}+2(\theta_1+\theta_2)+\epsilon'} \ll x^{\frac{17}{18}+\epsilon'},$$

which is negligible.

We return to the main term which, after some simplification, and recalling that $g_2(p) = \frac{1}{p}$ for $p \equiv 3 \pmod{4}$, becomes

$$\begin{aligned}
S_3^{(m_1, m_2)}(\nu_0) & \sim \frac{\pi^2 x g_1^2(qW)}{W \log^2 v} \sum_{\mathbf{d}, \mathbf{e} \in \mathcal{D}_K} \lambda_{\mathbf{d}} \lambda_{\mathbf{e}} \frac{\prod_{i \neq m_1, m_2} g_1^2([d_i, e_i])}{[d_{m_1}, e_{m_1}]^2 [d_{m_2}, e_{m_2}]^2 \prod_{i \neq m_1, m_2} [d_i, e_i]} \\
& \times \sum_{\substack{a, b \leq v \\ (a, q_1) = (b, q_1) = 1 \\ p|a, b \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)\mu(b)}{ab} \log \frac{v}{a} \log \frac{v}{b} \sum_{(t, 2qW \prod_{i \neq m_1, m_2} [d_i, e_i])=1} \frac{c_t(h)}{t^2} \\
& \times \prod_{i=1,2} \frac{([d_{m_i}, e_{m_i}], t) \chi([d_{m_i}, e_{m_i}]^2, t)}{\Psi([d_{m_i}, e_{m_i}], t)} \frac{(a, t)(b, t)}{\Psi(a, t)\Psi(b, t)}.
\end{aligned} \tag{42}$$

The sum over t is multiplicative, and can thus be written as a product $\Sigma_1 \times \Sigma_3$, where Σ_1 ranges over integers divisible only by primes congruent to 1 mod 4 and Σ_3 ranges over integers divisible only by primes congruent to 3 mod 4. That is,

$$\Sigma_3 = \sum_{\substack{t \\ p|t \Rightarrow p \equiv 3 \pmod{4} \\ (t, q_3 W \prod_{i \neq m_1, m_2} [d_i, e_i])=1}} \frac{c_t(h)}{t^2} \prod_{i=1,2} \frac{([d_{m_i}, e_{m_i}], t) \chi([d_{m_i}, e_{m_i}]^2, t)}{\Psi([d_{m_i}, e_{m_i}], t)},$$

and

$$\Sigma_1 = \sum_{\substack{t \\ (t, q_1)=1 \\ p|t \Rightarrow p \equiv 1 \pmod{4}}} \frac{c_t(h)}{t^2} \frac{(a, t)(b, t)}{\Psi(a, t)\Psi(b, t)}.$$

We have used the fact that in Σ_3 we have $(a, t) = (b, t) = \Psi(a, t) = \Psi(b, t) = 1$, and in Σ_1 we have $([d_{m_i}, e_{m_i}], t) = \chi([d_{m_i}, e_{m_i}]^2, t) = \Psi([d_{m_i}, e_{m_i}], t) = 1$.

We begin by considering Σ_3 . Using the definition $c_t(h) = \sum_{t_1|(t,h)} \mu(t/t_1)t_1$, and swapping sums and relabeling via $t_2 = t/t_1$, we get

$$\Sigma_3 = \sum_{\substack{t_1, t_2 \mid p|t_1 t_2 \Rightarrow p \equiv 3 \pmod{4} \\ (t_1 t_2, q_3 W \prod_{i \neq m_1, m_2} [d_i, e_i]) = 1} \frac{t_1 \mu(t_2)}{(t_1 t_2)^2} \prod_{i=1,2} \frac{([d_{m_i}, e_{m_i}], t_1 t_2) \chi([d_{m_i}, e_{m_i}]^2, t_1 t_2)}{\Psi([d_{m_i}, e_{m_i}], t_1 t_2)}.$$

Any prime $p|h$ with $p \equiv 3 \pmod{4}$ divides $q_3 W$, and $t_1|h$ must be co-prime with $q_3 W$, so $t_1 = 1$. Furthermore, t_2 is squarefree due to the term $\mu(t_2)$, which implies that $\Psi([d_{m_i}, e_{m_i}], t_2) = 1$. Thus,

$$\begin{aligned} \Sigma_3 &= \sum_{\substack{t_2 \mid p|t_2 \Rightarrow p \equiv 3 \pmod{4} \\ (t_2, q_3 W \prod_{i \neq m_1, m_2} [d_i, e_i]) = 1}} \frac{\mu(t_2)}{t_2^2} \prod_{i=1,2} ([d_{m_i}, e_{m_i}], t_2) \chi([d_{m_i}, e_{m_i}]^2, t_2)) \\ &= \prod_{\substack{p \nmid q_3 W \prod_{i \neq m_1, m_2} [d_i, e_i] \\ p \equiv 3 \pmod{4}}} \left(1 - \frac{1}{p^2}\right) \prod_{\substack{p \nmid q_3 W \prod_{i \neq m_1, m_2} [d_i, e_i] \\ p \mid \prod_{i=1,2} [d_{m_i}, e_{m_i}] \\ p \equiv 3 \pmod{4}}} \left(1 + \frac{1}{p}\right). \end{aligned}$$

Since $q_3 W$ is the product of all primes congruent to 3 mod 4 smaller than D_0 , we have

$$\begin{aligned} \Sigma_3 &= \prod_{\substack{p > D_0 \\ p \equiv 1 \pmod{4}}} \left(1 - \frac{1}{p^2}\right) \prod_{i \neq m_1, m_2} \left(\prod_{p \mid [d_i, e_i]} \left(1 - \frac{1}{p^2}\right)^{-1} \right) g_1([d_{m_1}, e_{m_1}]) g_1([d_{m_2}, e_{m_2}]) \\ &= (1 + o(1)) g_1([d_{m_1}, e_{m_1}]) g_1([d_{m_2}, e_{m_2}]) \prod_{i \neq m_1, m_2} f([d_i, e_i]), \end{aligned}$$

where $f(n) = \prod_{p|n} \left(1 - \frac{1}{p^2}\right)^{-1}$.

Plugging this back into (42), we get

$$\begin{aligned} S_3^{(m_1, m_2)}(\nu_0) &\sim \frac{\pi^2 x g_1^2(qW)}{W \log^2 v} \sum_{\mathbf{d}, \mathbf{e} \in \mathcal{D}_K} \lambda_{\mathbf{d}} \lambda_{\mathbf{e}} \frac{\prod_{i \neq m_1, m_2} g_1^2([d_i, e_i]) f([d_i, e_i]) \prod_{i=1,2} g_1([d_{m_i}, e_{m_i}])}{[d_{m_1}, e_{m_1}]^2 [d_{m_2}, e_{m_2}]^2 \prod_{i \neq m_1, m_2} [d_i, e_i]} \\ &\times \sum_{\substack{a, b \leq v \\ (a, q_1) = (b, q_1) = 1 \\ p|a, b \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)\mu(b)}{ab} \log \frac{v}{a} \log \frac{v}{b} \sum_{\substack{t \\ (t, q_1) = 1 \\ p|t \Rightarrow p \equiv 1 \pmod{4}}} \frac{c_t(h)}{t^2} \frac{(a, t)(b, t)}{\Psi(a, t)\Psi(b, t)}. \end{aligned}$$

In the notation of Lemma 9, the sum over \mathbf{d} and \mathbf{e} (which is independent of a, b, t) is of the form S_J for $J = \{m_1, m_2\}$, so that by Lemma 9, we get

$$S_3^{(m_1, m_2)}(\nu_0) \sim \frac{\pi^2 x g_1^2(qW)}{W \log^2 v} B^{K+2} L_{K; m_1, m_2}(F) \\ \times \sum_{\substack{a, b \leq v \\ (a, q_1) = (b, q_1) = 1 \\ p|a, b \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)\mu(b)}{ab} \log \frac{v}{a} \log \frac{v}{b} \sum_{\substack{t \\ (t, q_1) = 1 \\ p|t \Rightarrow p \equiv 1 \pmod{4}}} \frac{c_t(h)}{t^2} \frac{(a, t)(b, t)}{\Psi(a, t)\Psi(b, t)}.$$

Switching the order of summation gives

$$S_3^{(m_1, m_2)}(\nu_0) \sim \frac{\pi^2 x g_1^2(qW)}{W \log^2 v} B^{K+2} L_{K; m_1, m_2}(F) \quad (43) \\ \times \sum_{\substack{t \\ (t, q_1) = 1 \\ p|t \Rightarrow p \equiv 1 \pmod{4}}} \frac{c_t(h)}{t^2} \sum_{\substack{a, b \leq v \\ (a, q_1) = (b, q_1) = 1 \\ p|a, b \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)\mu(b)}{ab} \frac{(a, t)(b, t)}{\Psi(a, t)\Psi(b, t)} \log \frac{v}{a} \log \frac{v}{b}.$$

Denoting the inner sum as $\Sigma_{a, b}(t)$, we can write

$$\Sigma_{a, b}(t) = \left(\sum_{\substack{a \leq v \\ (a, q_1) = 1 \\ p|a \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)(a, t)}{a \Psi(a, t)} \log \frac{v}{a} \right)^2.$$

The calculations from [9, Lemma 6] (along much the same lines as Lemma 13) imply that

$$\Sigma_{a, b}(t) = \left(\frac{8A}{\pi g_1(q_1)} C(t) \log^{\frac{1}{2}} v + O\left(\frac{t^{\frac{1}{4}}}{\log^{\frac{1}{2}} v}\right) \right)^2,$$

where $C(t)$ is the constant in [9, Lemma 6], given by

$$C(t) = \begin{cases} \prod_{p|t, p \equiv 1 \pmod{4}} \left(2 - \frac{1}{p}\right)^{-1} & \text{if } p|t \text{ and } p \equiv 1 \pmod{4} \text{ implies } p^2|t \\ 0 & \text{otherwise.} \end{cases}$$

Note that in [9, Lemma 6], the Landau–Ramanujan constant is normalized as $\sqrt{2}A$, and that the statements of [9, Lemma 5] and [9, Lemma 6] are missing another factor of $\sqrt{2}$.

Plugging this estimate back into (43), we get

$$S_3^{(m_1, m_2)}(\nu_0) \sim \frac{64A^2 x g_1^2(q_3 W) B^{K+2}}{W \log v} L_{K; m_1, m_2}(F) \left(\sum_{\substack{t \\ (t, q_1) = 1 \\ p|t \Rightarrow p \equiv 1 \pmod{4}}} \frac{c_t(h) C^2(t)}{t^2} + O(E) \right),$$

where

$$\begin{aligned} E &\ll \frac{1}{\log v} \sum_{\substack{t \\ (t, q_1)=1 \\ p|t \Rightarrow p \equiv 1 \pmod{4}}} \frac{|c_t(h)|C(t)}{t^{\frac{7}{4}}} + \frac{1}{\log^2 v} \sum_{\substack{t \\ (t, q_1)=1 \\ p|t \Rightarrow p \equiv 1 \pmod{4}}} \frac{|c_t(h)|}{t^{\frac{3}{2}}} \\ &\ll \frac{1}{\log v} \sum_{\substack{t \\ (t, q_1)=1 \\ p|t \Rightarrow p \equiv 1 \pmod{4}}} \frac{|c_t(h)|}{t^{\frac{3}{2}}}. \end{aligned}$$

From [9, Eq 15], we have that $E \ll \frac{1}{\log v} \sigma_{-\frac{1}{2}}(h)$, which implies $E \ll \frac{\log \log x}{\log x}$ since $h \leq \eta(\log x)^{\frac{1}{2}}$. As for the main term, from [9, Eq 18], we have

$$\sum_{\substack{t \\ (t, q_1)=1 \\ p|t \Rightarrow p \equiv 1 \pmod{4}}} \frac{c_t(h)C^2(t)}{t^2} = \prod_{\substack{p^\beta \| h \\ p \equiv 1 \pmod{4} \\ p \nmid q_1}} \left(1 + \frac{1}{(2p-1)^2} \left(1 - \frac{1}{p^{\beta-1}} - \frac{1}{p^\beta} \right) \right).$$

Thus,

$$\begin{aligned} &S_3^{(m_1, m_2)} \\ &\sim \frac{64A^2 x g_1^2(q_3 W) B^{K+2}}{W \log v} \prod_{\substack{p^\beta \| b_{m_2} - b_{m_1} \\ p \equiv 1 \pmod{4} \\ p \nmid q_1}} \left(1 + \frac{1}{(2p-1)^2} \left(1 - \frac{1}{p^{\beta-1}} - \frac{1}{p^\beta} \right) \right) L_{K; m_1, m_2}(F). \end{aligned}$$

The product over $p^\beta \| (b_{m_2} - b_{m_1})$ is bounded above by V (defined in (15)), so that

$$S_3^{(m_1, m_2)}(\nu_0) \leq (1 + o(1)) \frac{64A^2 x g_1^2(q_3 W) B^{K+2}}{W \log v} V L_{K; m_1, m_2}(F).$$

Finally, using the identity that $B = \frac{2A}{\pi} \frac{\phi(q_3 W)(\log R)^{1/2}}{q_3 W}$ as well as applying (41) and Lemma 4 completes the proof. \square

3.7. Estimating $S_4^{(m)}$

In this section, we will prove Theorem 2, equation (20), which we restate in the following proposition.

Proposition 17. For fixed $1 \leq m \leq K$, consider the sum $S_4^{(m)}(\nu_0)$ defined by

$$S_4^{(m)}(\nu_0) = \sum_{\substack{x < n \leq 2x \\ n \equiv 1 \pmod{4} \\ n \equiv \nu_0 \pmod{W}}} \rho^2(\ell_m(n)) w_n(\mathcal{L}).$$

Then

$$S_4^{(m)}(\nu_0) = (1 + o(1)) \frac{8\pi \sqrt{\frac{\log R}{\log v}} \left(\frac{\log x}{\log v} + 1 \right) B^K x}{(\pi + 2)\sqrt{KW}} \prod_{\substack{p \equiv 1 \pmod{4} \\ p \nmid q_1}} \left(1 + \frac{1}{(2p-1)^2} \right) L_K(F).$$

Proof. We first expand the definitions of $w_n(\mathcal{L})$ and $\rho^2(\ell_m(n))$ and swap the order of summation to write

$$S_4^{(m)}(\nu_0) = \frac{1}{\log^2 v} \sum_{\mathbf{d}, \mathbf{e} \in \mathcal{D}_K} \lambda_{\mathbf{d}} \lambda_{\mathbf{e}} \sum_{\substack{a, b \leq v \\ (a, q_1) = (b, q_1) = 1 \\ p|ab \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)\mu(b)}{g_2(a)g_2(b)} \log \frac{v}{a} \log \frac{v}{b} \sum_{\substack{x < n \leq 2x \\ n \equiv 1 \pmod{4} \\ n \equiv \nu_0 \pmod{W} \\ [d_i, e_i] | \ell_i(n) \forall i \\ [a, b] | \ell_m(n)}} r_2^2(\ell_m(n)). \quad (44)$$

The quantities $W, [d_1, e_1], \dots, [d_k, e_k], q$, and $[a, b]$ must be pairwise coprime because of the support of \mathcal{D}_K . We use Lemma 12 in order to evaluate the inner sum; we will apply Lemma 12 with $r = qW \prod_{i \neq m} [d_i, e_i]$ and $d = [a, b][d_m, e_m]$. The sum (44) can then be written as

$$\frac{1}{\log^2 v} \sum_{\mathbf{d}, \mathbf{e} \in \mathcal{D}_K} \lambda_{\mathbf{d}} \lambda_{\mathbf{e}} \sum_{\substack{a, b \leq v \\ p|ab \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)\mu(b)}{g_2(a)g_2(b)} \log \frac{v}{a} \log \frac{v}{b} \times \quad (45)$$

$$\left(\frac{g_3(r)g_4(d)}{rd} \left(\log qx + A_2 + 2 \sum_{p|r} g_5(p) - 2 \sum_{p|d} g_6(p) \right) qx + O_\epsilon \left((qx)^{\frac{3}{4} + \theta_2 + \epsilon} \right) \right).$$

Taking absolute values, the error term from (45) is bounded by

$$\ll_\epsilon \lambda_{\max}^2 |\mathcal{D}_K|^2 v^2 (qx)^{\frac{3}{4} + \theta_2 + \epsilon} \ll_{\epsilon'} x^{\frac{3}{4} + 2(\theta_2 + \theta_1) + \epsilon'},$$

where we used (34) to bound the λ_{\max} and (37) for $|\mathcal{D}_K|$. This error term is negligible, since $\theta_1 + \theta_2 < \frac{1}{18}$.

We now evaluate the main term via a process that is identical to the one in [15, Proposition 6.2, part (iv)]. Using the notation from Lemma 13, the main term of (45) is

$$(1 + o(1)) \frac{g_3(qW)B^{K+1}qx}{W \log^2 v} \left(Z_{x, q_1}^{(1)} \log x - 2Z_{x, q_1}^{(2)} \right) L_{K; m}(F).$$

By Lemma 13, this is equal to

$$(1 + o(1)) \frac{g_3(qW)B^{K+1}qx}{W \log^2 v} \frac{8Ag_7(q_1)}{\phi(q_1)g_1(q_1)} \sqrt{\log v} (\log x + \log v)$$

$$\times \prod_{\substack{p \equiv 1 \pmod{4} \\ p \nmid q_1}} \left(1 + \frac{1}{(2p-1)^2} \right) L_{K; m}(F)$$

$$= (1 + o(1)) \frac{8Ag_3(qW)g_7(q_1)B^{K+1} \log^{\frac{3}{2}} v \left(\frac{\log x}{\log v} + 1 \right) qx}{qW \phi(q_1)g_1(q_1) \log^2 v}$$

$$\times \prod_{\substack{p \equiv 1 \pmod{4} \\ p \nmid q_1}} \left(1 + \frac{1}{(2p-1)^2} \right) L_{K; m}(F).$$

Recalling the definition of B and equation (41), we get that

$$S_4^{(m)}(\nu_0) = (1 + o(1)) \frac{g_3(q_1)g_7(q_1)}{\phi(q_1)g_1(q_1)} \frac{8B^k \sqrt{\frac{\log R}{\log v}} \left(\frac{\log x}{\log v} + 1 \right) x}{W\pi} \prod_{\substack{p \equiv 1 \pmod{4} \\ p \nmid q_1}} \left(1 + \frac{1}{(2p-1)^2} \right) L_{K;m}(F).$$

Observing that the factors dividing q_1 cancel and applying the identity from Lemma 4 that $L_{K;m}(F) = \frac{\pi^2}{\pi+2} \frac{L_K(F)}{\sqrt{K}}$ completes the proof. \square

3.8. Estimating $S_5^{(m)}$

In this section, we will prove Theorem 2, equation (21), which we restate in the following proposition.

Proposition 18. *Let $\xi > 0$ be a constant with $\xi < \frac{1}{K}$. For fixed $1 \leq m \leq K$, define $S_5^{(m)}(\nu_0)$ to be the sum*

$$S_5^{(m)}(\nu_0) := \sum_{\substack{x < n \leq 2x \\ n \equiv 1 \pmod{4} \\ n \equiv \nu_0 \pmod{W}}} \sum_{\substack{p < x^\xi \\ p \equiv 3 \pmod{4} \\ p \mid \ell_m(n)}} w_n(\mathcal{L}).$$

Then

$$S_5^{(m)}(\nu_0) \ll \frac{K^2 \xi^2}{\theta_2^2} \frac{B^K x}{W} L_K(F).$$

The proof of the proposition relies on the following lemma, which we state and prove before turning to the main proof of Proposition 18.

Lemma 19. *Define*

$$T = \frac{x}{W} \sum_{\mathbf{u}, \mathbf{v} \in \mathcal{D}_K} \frac{y_{\mathbf{u}} y_{\mathbf{v}}}{\phi(u)\phi(v)} \prod_{p \mid uv} |\sigma_p(\mathbf{u}, \mathbf{v})|,$$

where

$$\sigma_p(\mathbf{u}, \mathbf{v}) = \sum_{\substack{\mathbf{d} \mid \mathbf{u}, \mathbf{e} \mid \mathbf{v} \\ d_i, e_i \mid p \ \forall i}} \frac{\mu(d)\mu(e)de}{[\mathbf{d}, \mathbf{e}]} = \begin{cases} p-1 & \text{if } p \mid (\mathbf{u}, \mathbf{v}) \\ 0 & \text{if } p \mid uv, p \nmid (\mathbf{u}, \mathbf{v}) \\ 1 & \text{if } p \nmid uv \end{cases}.$$

Then

$$T \ll \frac{x}{W} \left(\frac{e^{-\gamma/2}}{\Gamma(1/2)} \right)^K \left(\frac{\log R}{\log D_0} \right)^{K/2} L_K(F).$$

Proof. First note that if $u \neq v$, then for some prime p , $\sigma_p(\mathbf{u}, \mathbf{v}) = 0$, so these terms do not contribute. Thus, for a fixed $\mathbf{u} \in \mathcal{D}_K$,

$$\sum_{\mathbf{v} \in \mathcal{D}_K} \frac{\prod_{p|uv} |\sigma_p(\mathbf{u}, \mathbf{v})|}{\phi(v)} = \prod_{p|u} \left(\sum_{\substack{\mathbf{w} \in \mathcal{D}_K \\ w_i | p \forall i}} \frac{|\sigma_p(\mathbf{u}, \mathbf{w})|}{\phi(w)} \right) = 1.$$

This, along with the bound that $y_{\mathbf{u}} y_{\mathbf{v}} \ll y_{\mathbf{u}}^2 + y_{\mathbf{v}}^2$, implies that

$$\begin{aligned} T &\ll \frac{x}{W} \sum_{\mathbf{u}, \mathbf{v} \in \mathcal{D}_K} \frac{y_{\mathbf{u}}^2 + y_{\mathbf{v}}^2}{\phi(u)\phi(v)} \prod_{p|uv} |\sigma_p(\mathbf{u}, \mathbf{v})| \ll \frac{x}{W} \sum_{\mathbf{u} \in \mathcal{D}_K} \frac{y_{\mathbf{u}}^2}{\phi(u)} \left(\sum_{\mathbf{v} \in \mathcal{D}_K} \frac{\prod_{p|uv} |\sigma_p(\mathbf{u}, \mathbf{v})|}{\phi(v)} \right) \\ &\ll \frac{x}{W} \sum_{\mathbf{u} \in \mathcal{D}_K} \frac{y_{\mathbf{u}}^2}{\phi(u)}. \end{aligned}$$

By Lemma 7,

$$\sum_{\mathbf{r} \in \mathcal{D}_K} \frac{y_{\mathbf{r}}^2}{\phi(r)} = \sum_{\mathbf{r} \in \mathcal{D}_K} \frac{\mu(r)^2}{\phi(r)} \prod_{i=1}^K g \left(K \frac{\log r_1}{\log R} \right)^2 \ll \left(\frac{e^{-\gamma/2}}{\Gamma(1/2)} \right)^K \left(\frac{\log R}{\log D_0} \right)^{K/2} L_K(F),$$

as desired. \square

We are now ready to prove Proposition 18.

Proof of Proposition 18. Expanding the square and swapping the order of summation gives

$$S_5^{(m)}(\nu_0) = \sum_{\substack{p < x^\xi \\ p \equiv 3 \pmod{4}}} \sum_{\mathbf{d}, \mathbf{e} \in \mathcal{D}_K} \lambda_{\mathbf{d}} \lambda_{\mathbf{e}} \sum_{\substack{x < n \leq 2x \\ n \equiv 1 \pmod{4} \\ n \equiv \nu_0 \pmod{W} \\ [d_i, e_i] | \ell_i(n) \\ p | \ell_m(n)}} 1.$$

By choice of $\nu_0 \pmod{W}$, if $p | \ell_m(n)$, then $p > D_0$. Because of the support of \mathcal{D}_K , if $\lambda_{\mathbf{d}} \neq 0$ and $\lambda_{\mathbf{e}} \neq 0$, then any prime $p \equiv 3 \pmod{4}$ can divide at most one of the $\ell_i(n)$. Thus, if $p | \ell_m(n)$, then $p \nmid \ell_i(n)$ for all $i \neq m$, which implies that $(d_i e_i, p) = 1$ for all $i \neq m$. By the Chinese remainder theorem, the inner sum is of the form $\frac{x}{Q} + O(1)$, where $Q = 4W[d_m, e_m, p] \prod_{i \neq m} [d_i, e_i]$. Note that $Q < 4WR^2 x^\xi$, and for any fixed Q , there are $O(\tau_{3k+4}(Q))$ choices of $\mathbf{d}, \mathbf{e}, p$ giving rise to the modulus Q . Thus, the error term from the Chinese remainder theorem application and (34) makes a contribution that is

$$\ll \sum_{Q < 4WR^2 x^\xi} \tau_{3k+4}(Q) \lambda_{\max}^2 \ll_\varepsilon R^2 x^\xi x^\varepsilon = x^{\theta_2 + \xi + \varepsilon},$$

which is negligible for ξ small.

The remaining term is given by

$$\frac{x}{4W} \sum_{\substack{D_0 < p < x^\xi \\ p \equiv 3 \pmod{4}}} \frac{1}{p} \sum_{\substack{\mathbf{d}, \mathbf{e} \in \mathcal{D}_K \\ (d_i e_i, p) = 1 \forall i \neq m}} \frac{\lambda_{\mathbf{d}} \lambda_{\mathbf{e}} p}{[d_m, e_m, p] \prod_{i \neq m} [d_i, e_i]}.$$

Expanding the definitions of λ_d, λ_e and rearranging, this is

$$\begin{aligned} & \frac{x}{4W} \sum_{\substack{D_0 < p < x^\xi \\ p \equiv 3 \pmod{4}}} \frac{1}{p} \sum_{\substack{\mathbf{d}, \mathbf{e} \in \mathcal{D}_K \\ (d_i e_i, p) = 1 \forall i \neq m}} \frac{\mu(d)\mu(e)dep}{[d_m, e_m, p] \prod_{i \neq m} [d_i, e_i]} \sum_{\substack{\mathbf{r}, \mathbf{s} \in \mathcal{D}_K \\ \mathbf{d} | \mathbf{r}, \mathbf{e} | \mathbf{s}}} \frac{y_r y_s}{\phi(r)\phi(s)} \\ &= \frac{x}{4W} \sum_{\substack{D_0 < p < x^\xi \\ p \equiv 3 \pmod{4}}} \frac{1}{p} \sum_{\mathbf{r}, \mathbf{s} \in \mathcal{D}_K} \frac{y_r y_s}{\phi(r)\phi(s)} \sum_{\substack{\mathbf{d}, \mathbf{e} \in \mathcal{D}_K \\ (d_i e_i, p) = 1 \forall i \neq m \\ \mathbf{d} | \mathbf{r}, \mathbf{e} | \mathbf{s}}} \frac{\mu(d)\mu(e)dep}{[d_m, e_m, p] \prod_{i \neq m} [d_i, e_i]}. \end{aligned} \quad (46)$$

The inside sum is multiplicative over $p' | rs$; write $\sigma_{p'}(\mathbf{r}, \mathbf{s}, p)$ for the p' component. If $p' \neq p$, then

$$\sigma_{p'}(\mathbf{r}, \mathbf{s}, p) = \sum_{\substack{\mathbf{d}, \mathbf{e} \in \mathcal{D}_K \\ d_i, e_i | p' \forall i \\ \mathbf{d} | \mathbf{r}, \mathbf{e} | \mathbf{s}}} \frac{\mu(d)\mu(e)dep}{[d_m, e_m, p] \prod_{i \neq m} [d_i, e_i]} = \begin{cases} p' - 1 & \text{if } p' | (\mathbf{r}, \mathbf{s}) \\ -1 & \text{if } p' | r, p' | s, p' \nmid (\mathbf{r}, \mathbf{s}) \\ 0 & \text{otherwise,} \end{cases}$$

where we recall that $(\mathbf{r}, \mathbf{s}) = \prod_i (r_i, s_i)$. If $p' = p$, then

$$\sigma_p(\mathbf{r}, \mathbf{s}, p) = \sum_{\substack{\mathbf{d}, \mathbf{e} \in \mathcal{D}_K \\ d_m, e_m | p \\ d_i = e_i = 1 \forall i \neq m \\ \mathbf{d} | \mathbf{r}, \mathbf{e} | \mathbf{s}}} \frac{\mu(d)\mu(e)dep}{[d_m, e_m, p] \prod_{i \neq m} [d_i, e_i]} = \begin{cases} (p-1)^2 & \text{if } p | (r_m, s_m) \\ -(p-1) & \text{if } p | r_m s_m, p \nmid (r_m, s_m) \\ 1 & \text{otherwise.} \end{cases}$$

Let $f_{\mathbf{u}}(\mathbf{r}) = (r_1, \dots, r_m / (r_m, p), \dots, r_k)$ be the vector formed by removing a possible factor of p from r_m . Then our expression (46) can be written as

$$\frac{x}{4W} \sum_{\substack{D_0 < p < x^\xi \\ p \equiv 3 \pmod{4}}} \frac{1}{p} \sum_{\substack{\mathbf{u}, \mathbf{s} \in \mathcal{D}_K \\ (u_m, p) = 1}} \frac{y_s}{\phi(s)} \sum_{\substack{\mathbf{r} \\ f_{\mathbf{u}}(\mathbf{r}) = \mathbf{u}}} \frac{y_r}{\phi(r)} \prod_{p' | rs} \sigma_{p'}(\mathbf{r}, \mathbf{s}, p).$$

We split the sum above into several parts. Let Σ_1 be the summands where $p | u_j$ for some $j \neq m$, and let Σ_2 be the summands where $p \nmid u_i$ for all i . Define

$$T = \frac{x}{W} \sum_{\mathbf{u}, \mathbf{v} \in \mathcal{D}_K} \frac{y_u y_v}{\phi(u)\phi(v)} \prod_{p' | uv} |\sigma_{p'}(\mathbf{u}, \mathbf{v})|, \quad (47)$$

where

$$\sigma_{p'}(\mathbf{u}, \mathbf{v}) = \sum_{\substack{\mathbf{d} | \mathbf{u}, \mathbf{e} | \mathbf{v} \\ d_i, e_i | p' \forall i}} \frac{\mu(d)\mu(e)de}{[\mathbf{d}, \mathbf{e}]}.$$

We will bound both Σ_1 and Σ_2 in terms of T , showing first that $\Sigma_1 \ll \frac{1}{D_0}T$. We have

$$\Sigma_1 = \frac{x}{4W} \sum_{\substack{D_0 < p < x^\xi \\ p \equiv 3 \pmod{4}}} \frac{1}{p} \sum_{i \neq m} \sum_{\substack{\mathbf{u}, \mathbf{s} \in \mathcal{D}_K \\ (u_m, p) = 1 \\ p | u_i}} \frac{y_{\mathbf{s}}}{\phi(s)} \sum_{\substack{\mathbf{r} \\ f_{\mathbf{u}}(\mathbf{r}) = \mathbf{u}}} \frac{y_{\mathbf{r}}}{\phi(r)} \sigma_p(\mathbf{r}, \mathbf{s}, p) \prod_{\substack{p' | rs \\ p' \neq p}} \sigma_{p'}(\mathbf{r}, \mathbf{s}, p).$$

Given \mathbf{u} , there is only one vector \mathbf{r} such that $f_{\mathbf{u}}(\mathbf{r}) = \mathbf{u}$; namely, $\mathbf{r} = \mathbf{u}$. Thus,

$$\Sigma_1 = \frac{x}{4W} \sum_{\substack{D_0 < p < x^\xi \\ p \equiv 3 \pmod{4}}} \frac{1}{p} \sum_{i \neq m} \sum_{\substack{\mathbf{u}, \mathbf{s} \in \mathcal{D}_K \\ (u_m, p) = 1 \\ p | u_i}} \frac{y_{\mathbf{s}} y_{\mathbf{u}}}{\phi(s) \phi(u)} \sigma_p(\mathbf{u}, \mathbf{s}, p) \prod_{\substack{p' | us \\ p' \neq p}} \sigma_{p'}(\mathbf{u}, \mathbf{s}).$$

Denote by \mathbf{u}' the vector obtained from \mathbf{u} by removing all factors of p . Then $\phi(u) = (p-1)\phi(u')$ and $\sigma_p(\mathbf{u}, \mathbf{s}, p) = \sigma_p(\mathbf{u}', \mathbf{s}, p) = \mu((s_m, p))\phi((s_m, p))$, which is independent of \mathbf{u} because we already require $(u_m, p) = 1$. Thus,

$$\Sigma_1 = \frac{x}{4W} \sum_{\mathbf{s} \in \mathcal{D}_K} \frac{y_{\mathbf{s}}}{\phi(s)} \sum_{\substack{D_0 < p < x^\xi \\ p \equiv 3 \pmod{4}}} \frac{1}{p} \sum_{\substack{\mathbf{u}' \in \mathcal{D}_K \\ p \nmid u'_i}} \frac{1}{\phi(u')} \sigma_p(\mathbf{u}', \mathbf{s}, p) \prod_{\substack{p' | u's \\ p' \neq p}} \sigma_{p'}(\mathbf{u}', \mathbf{s}) \sum_{\substack{\mathbf{u} \in \mathcal{D}_K \\ \mathbf{u} \rightarrow \mathbf{u}'}} \frac{y_{\mathbf{u}}}{(p-1)}.$$

By Lemma 5, we have $y_{\mathbf{u}} = y_{\mathbf{u}'}(1 + O(K\xi))$. By assumption, $K\xi \ll 1$, so (recalling that the weights $y_{\mathbf{r}}$ are nonnegative), $y_{\mathbf{u}} \ll y_{\mathbf{u}'}$ and

$$\Sigma_1 \ll \frac{x}{W}(K-1) \sum_{\mathbf{s} \in \mathcal{D}_K} \frac{y_{\mathbf{s}}}{\phi(s)} \sum_{\substack{D_0 < p < x^\xi \\ p \equiv 3 \pmod{4}}} \frac{1}{p(p-1)} \sum_{\substack{\mathbf{u}' \in \mathcal{D}_K \\ p \nmid u'_i}} \frac{y_{\mathbf{u}'}}{\phi(u')} |\sigma_p(\mathbf{u}', \mathbf{s}, p)| \prod_{\substack{p' | u's \\ p' \neq p}} |\sigma_{p'}(\mathbf{u}', \mathbf{s})|.$$

To bound Σ_1 , we now further split it into subsums. First, let T_1 consist of all those terms with \mathbf{s} such that $p \nmid s_i$ for all i . In this case, $\sigma_p(\mathbf{u}', \mathbf{s}, p) = 1$, so

$$T_1 \ll \frac{x}{W}(K-1) \sum_{\substack{D_0 < p < x^\xi \\ p \equiv 3 \pmod{4}}} \frac{1}{p(p-1)} \sum_{\substack{\mathbf{s} \in \mathcal{D}_K \\ p \nmid s_i \forall i}} \frac{y_{\mathbf{s}}}{\phi(s)} \sum_{\substack{\mathbf{u}' \in \mathcal{D}_K \\ p \nmid u'_i \forall i}} \frac{y_{\mathbf{u}'}}{\phi(u')} \prod_{p' | u's} |\sigma_{p'}(\mathbf{u}', \mathbf{s})|.$$

Dropping the requirement that $p \nmid s_i$, $p \nmid u'_i$ only increases T_1 . The sum over p is then independent of the rest of the expression, and converges to a constant that is $\ll \frac{1}{D_0}$, which in turn implies that $T_1 \ll \frac{K}{D_0}T$, where T is defined in (47).

Now consider T_2 , the terms \mathbf{s} in Σ_1 such that $p | s_i$ for some $i \neq m$. In this case, $\sigma_p(\mathbf{u}', \mathbf{s}, p) = 1$, so

$$T_2 \ll \frac{x}{W}(K-1)^2 \sum_{D_0 < p < x^\xi} \frac{1}{p(p-1)} \sum_{\substack{\mathbf{s} \in \mathcal{D}_K \\ p | s_1}} \frac{y_{\mathbf{s}}}{\phi(s)} \sum_{\substack{\mathbf{u}' \in \mathcal{D}_K \\ p \nmid u'_i \forall i}} \frac{y_{\mathbf{u}'}}{\phi(u')} \prod_{\substack{p' | u's \\ p' \neq p}} |\sigma_{p'}(\mathbf{u}', \mathbf{s})|.$$

Let \mathbf{s}' be the vector obtained by removing the factor of p from \mathbf{s} . Once again, $y_{\mathbf{s}} \ll y_{\mathbf{s}'}$, so

$$T_2 \ll \frac{x}{W}(K-1)^2 \sum_{\substack{D_0 < p < x^\xi \\ p \equiv 3 \pmod{4}}} \frac{1}{p(p-1)^2} \sum_{\substack{\mathbf{s}' \in \mathcal{D}_K \\ p \nmid s'_i \forall i}} \frac{y_{\mathbf{s}'}}{\phi(s')} \sum_{\substack{\mathbf{u}' \in \mathcal{D}_K \\ p \nmid u'_i \forall i}} \frac{y_{\mathbf{u}'}}{\phi(u')} \prod_{p' | u's} |\sigma_{p'}(\mathbf{u}', \mathbf{s})|.$$

Once more, we can remove the constraints that $p \nmid s_i$ and $p \nmid u'_i$ and evaluate the sum over p to get that $T_2 \ll \frac{K^2}{D_0^2} T$.

Finally, consider T_3 , the subsum of Σ_1 with those \mathbf{s} such that $p \mid s_m$. In this case, $\sigma_p(\mathbf{u}', \mathbf{s}, p) = -(p-1)$. A similar computation gives

$$T_3 \ll \frac{x}{W} \sum_{\substack{D_0 < p < x^\xi \\ p \equiv 3 \pmod{4}}} \frac{1}{p(p-1)^2} \sum_{\substack{\mathbf{s}' \in \mathcal{D}_K \\ p \nmid s_i \forall i}} \frac{y_{\mathbf{s}'}}{\phi(s')} \sum_{\substack{\mathbf{u}' \in \mathcal{D}_K \\ p \nmid u'_i \forall i}} \frac{y_{\mathbf{u}'}}{\phi(u')} (p-1) \prod_{p' \mid \mathbf{u}' \mathbf{s}} |\sigma_{p'}(\mathbf{u}', \mathbf{s})| \ll \frac{1}{D_0} T.$$

Thus, $|\Sigma_1| \ll T_1 + T_2 + T_3 \ll \frac{K^2}{D_0} T$.

Now consider Σ_2 , given by

$$\Sigma_2 = \frac{x}{4W} \sum_{\substack{D_0 < p < x^\xi \\ p \equiv 3 \pmod{4}}} \frac{1}{p} \sum_{\substack{\mathbf{u}, \mathbf{s} \in \mathcal{D}_K \\ (u_i, p) = 1 \forall i}} \frac{y_{\mathbf{s}}}{\phi(s)} \sum_{\substack{\mathbf{r} \in \mathcal{D}_K \\ f_{\mathbf{u}}(\mathbf{r}) = \mathbf{u}}} \frac{y_{\mathbf{r}}}{\phi(r)} \prod_{p' \mid \mathbf{r} \mathbf{s}} \sigma_{p'}(\mathbf{r}, \mathbf{s}, p).$$

Observe that for fixed $\mathbf{u}, \mathbf{s} \in \mathcal{D}_K$ with $p \nmid u_i$ for all i ,

$$\sum_{\substack{\mathbf{r} \in \mathcal{D}_K \\ f_{\mathbf{u}}(\mathbf{r}) = \mathbf{u}}} \frac{\sigma_p(\mathbf{r}, \mathbf{s}, p)}{\phi(r)} = \frac{\mu((s_m, p)) \phi((s_m, p))}{\phi(u)} \left(1 - \frac{p-1}{p-1}\right) = 0. \quad (48)$$

We substitute $y_{\mathbf{r}} = y_{\mathbf{u}} + (y_{\mathbf{r}} - y_{\mathbf{u}})$ into Σ_2 . By (48), the $y_{\mathbf{u}}$ do not contribute, leaving only the contribution from $(y_{\mathbf{r}} - y_{\mathbf{u}})$. The only terms remaining have $\mathbf{r} \neq \mathbf{u}$, so that $p \mid r_m$. Thus,

$$\Sigma_2 = \frac{x}{4W} \sum_{\substack{D_0 < p < x^\xi \\ p \equiv 3 \pmod{4}}} \frac{1}{p} \sum_{\substack{\mathbf{r}, \mathbf{s} \in \mathcal{D}_K \\ p \mid r_m}} \frac{y_{\mathbf{s}}(y_{\mathbf{r}} - y_{\mathbf{u}})}{\phi(r)\phi(s)} \prod_{p' \mid \mathbf{r} \mathbf{s}} \sigma_{p'}(\mathbf{r}, \mathbf{s}, p).$$

By running the same argument for \mathbf{s} and a tuple \mathbf{v} obtained from \mathbf{s} by removing a factor of p from s_m (including bounding the terms where $p \mid v_i$ for some $i \neq m$ by $\frac{K^2}{D_0} T$ using identical arguments to the bound on Σ_1), we can also replace $y_{\mathbf{s}}$ by $y_{\mathbf{s}} - y_{\mathbf{v}}$. By Lemma 5, we have

$$(y_{\mathbf{r}} - y_{\mathbf{u}})(y_{\mathbf{s}} - y_{\mathbf{v}}) \ll y_{\mathbf{u}} y_{\mathbf{v}} K^2 \frac{(\log p)^2}{(\log R)^2},$$

so Σ_2 is given by

$$\begin{aligned} \Sigma_2 &= \frac{x}{4W} \sum_{\substack{D_0 < p < x^\xi \\ p \equiv 3 \pmod{4}}} \frac{1}{p} \sum_{\substack{\mathbf{r}, \mathbf{s} \in \mathcal{D}_K \\ p \mid (r_m, s_m)}} \frac{(y_{\mathbf{r}} - y_{\mathbf{u}})(y_{\mathbf{s}} - y_{\mathbf{v}})}{\phi(r)\phi(s)} \prod_{p' \mid \mathbf{r} \mathbf{s}} \sigma_{p'}(\mathbf{r}, \mathbf{s}, p) + O\left(\frac{K^2 T}{D_0}\right) \\ &\ll \frac{x K^2}{W} \sum_{\substack{D_0 < p < x^\xi \\ p \equiv 3 \pmod{4}}} \frac{1}{p} \left(\frac{\log p}{\log R}\right)^2 \sum_{\substack{\mathbf{u}, \mathbf{v} \in \mathcal{D}_K \\ (uv, p) = 1}} y_{\mathbf{u}} y_{\mathbf{v}} \prod_{\substack{p' \mid uv \\ p' \neq p}} |\sigma_{p'}(\mathbf{u}, \mathbf{v})| \sum_{\substack{\mathbf{r}, \mathbf{s} \in \mathcal{D}_K \\ r_m = pu_m \\ s_m = pv_m \\ r_i = u_i \forall i \neq m \\ s_i = v_i \forall i \neq m}} \frac{|\sigma_p(\mathbf{r}, \mathbf{s}, p)|}{\phi(r)\phi(s)} + \frac{K^2 T}{D_0}. \end{aligned}$$

The sum over \mathbf{r} and \mathbf{s} is equal to $\phi(u)^{-1}\phi(v)^{-1}$, so

$$\begin{aligned}\Sigma_2 &\ll \frac{xK^2}{W} \sum_{\substack{D_0 < p < x^\xi \\ p \equiv 3 \pmod{4}}} \frac{1}{p} \left(\frac{\log p}{\log R} \right)^2 \sum_{\substack{\mathbf{u}, \mathbf{v} \in \mathcal{D}_K \\ (uv, p)=1}} \frac{y_{\mathbf{u}} y_{\mathbf{v}}}{\phi(u)\phi(v)} \prod_{\substack{p' | uv \\ p' \neq p}} |\sigma_{p'}(\mathbf{u}, \mathbf{v})| + \frac{K^2 T}{D_0} \\ &\ll \frac{xK^2 \xi^2}{W \theta_2^2} \sum_{\mathbf{u}, \mathbf{v} \in \mathcal{D}_K} \frac{y_{\mathbf{u}} y_{\mathbf{v}}}{\phi(u)\phi(v)} \prod_{p' | uv} |\sigma_{p'}(\mathbf{u}, \mathbf{v})| + \frac{KT}{D_0} \ll \left(K^2 \frac{\xi^2}{\theta_2^2} + \frac{K^2}{D_0} \right) T.\end{aligned}$$

Altogether, we get that $S_5^{(m)}(\nu_0) \ll \left(\frac{K^2 \xi^2}{\theta_2^2} + \frac{K^2}{D_0} \right) T$. The contribution from the $\frac{K^2}{D_0}$ term vanishes as x grows large. The quantity T is evaluated in Lemma 19, giving

$$S_5^{(m)}(\nu_0) \ll \frac{xK^2 \xi^2}{W \theta_2^2} \left(\frac{e^{-\gamma/2}}{\Gamma(1/2)} \right)^K \left(\frac{\log R}{\log D_0} \right)^{K/2} L_K(F).$$

From the definition of B and Mertens' theorem, we get

$$\left(\frac{e^{-\gamma/2}}{\Gamma(1/2)} \right)^K \left(\frac{\log R}{\log D_0} \right)^{K/2} \sim B^k,$$

which completes the argument. \square

3.9. Estimating $S_6^{(b)}$

In this section, we will prove Theorem 2, equation (22), which we restate in the following proposition.

Proposition 20. *Let ν_1 be a congruence class modulo $q_3^2 W^2$ such that $(\ell(\nu_1), q_3^2 W^2)$ is a square for all $\ell \in \mathcal{L}$. Fix $3 < b \leq \eta \sqrt{\log x}$ and consider the linear form $\ell^{(b)}(n) := qn + b$. Fix a constant ξ with $0 < \xi < 1/4$, and define*

$$S_6^{(b)}(\nu_1) := \sum_{\substack{x < n \leq 2x \\ n \equiv 1 \pmod{4} \\ n \equiv \nu_1 \pmod{q_3^2 W^2}}} \mathbf{1}_{S(\xi)}(\ell^{(b)}(n)) w_n(\mathcal{L}),$$

where $S(\xi)$ is the set described in (11). Then

$$S_6^{(b)}(\nu_1) \ll_K \frac{x}{4q_3^2 W^2} \xi^{-1/2} \left(\frac{\theta_2}{2} \right)^{-1/2} \left(\frac{\log R}{\log D_0} \right)^{\frac{K-1}{2}} L_K(F).$$

Proof. We will apply Selberg's sieve to bound the function $\mathbf{1}_{S(\xi)}(\ell(n))$, while also evaluating the sum over sieve weights $w_n(\mathcal{L})$. We begin by defining the additional sieve weights.

Recall that $S(\xi)$ denotes the set of integers such that for all primes $p < x^\xi$ with $p \equiv 3 \pmod{4}$, either $p \nmid n$ or $p^2 \mid n$. Thus, for each prime $p \equiv 3 \pmod{4}$, with $D_0 < p < x^\xi$, we sieve by the set \mathcal{A}_p of integers $x \leq n \leq 2x$ such that $p \mid \ell(n)$ but $p^2 \nmid \ell(n)$. The sieving set \mathcal{A}_p has density function

$$g(p) = \begin{cases} \frac{1}{p} - \frac{1}{p^2} & p > D_0 \text{ and } p \equiv 3 \pmod{4} \\ 0 & \text{otherwise.} \end{cases}$$

We extend both $g(p)$ and \mathcal{A}_p multiplicatively to squarefree d , so that

$$|r_d| := |\mathcal{A}_d - g(d)x| \ll \tau(d).$$

We will use the upper bound Selberg sieve

$$\mathbf{1}_{S(\xi)}(\ell(n)) \leq \sum_{f \mid \ell(n)} \mu^+(f) |\mathcal{A}_f|,$$

where

$$\mu^+(f) = \frac{1}{\tilde{\lambda}_1} \sum_{[d_0, e_0] = f} \tilde{\lambda}_{d_0} \tilde{\lambda}_{e_0},$$

and $\tilde{\lambda}_d$ is a sequence of weights defined as follows. Define ‘diagonalizing vectors’ y_{r_0} via

$$y_{r_0} := \begin{cases} 1 & \text{if } (r_0, q_3 W) = 1, r_0 < x^\xi, \text{ and } p \mid r_0 \Rightarrow p \equiv 3 \pmod{4} \\ 0 & \text{otherwise,} \end{cases}$$

and define $\tilde{\lambda}_{d_0}$ to be

$$\tilde{\lambda}_{d_0} := \mu(d_0) \frac{d_0^2}{\phi(d_0)} \sum_{d_0 \mid r_0} \frac{y_{r_0}}{\phi(r_0)}. \quad (49)$$

By Möbius inversion, we also have the relation that

$$y_{r_0} = \mu(r_0) \phi(r_0) \sum_{r_0 \mid d_0} \frac{\tilde{\lambda}_{d_0} \phi(d_0)}{d_0^2}.$$

Note that $\tilde{\lambda}_{d_0}$ is supported on squarefree d_0 with $(d_0, q_3 W) = 1$, $d_0 < x^\xi$, and d_0 only divisible by primes congruent to 3 mod 4. Also, with this choice,

$$\tilde{\lambda}_1 = \sum_{\substack{r_0 < x^\xi \\ (r_0, q_3 W) = 1 \\ p \mid r_0 \Rightarrow p \equiv 3 \pmod{4}}} \frac{1}{\phi(r_0)} \gg \sqrt{\frac{\xi \log x}{\log D_0}}. \quad (50)$$

For what follows, we will fix the notation that

$$\phi_{\omega^*}(n) = n \prod_{p|n} \left(1 - \frac{K+1}{p}\right),$$

and define further ‘cross’-diagonalizing vectors

$$y_{\mathbf{r}, r_0} := \mu(r_0 r) \phi_{\omega^*}(r_0 r) \sum_{\substack{\mathbf{r}|\mathbf{d} \\ r_0|d_0 \\ (d_0, d)=1}} \frac{\lambda_{\mathbf{d}} \tilde{\lambda}_{d_0} \phi(d_0)}{d d_0^2}, \quad (51)$$

which satisfy the inverse relation that

$$\lambda_{\mathbf{d}} \tilde{\lambda}_{d_0} = \frac{\mu(d_0 d) d d_0^2}{\phi(d_0)} \sum_{\substack{\mathbf{d}|\mathbf{r} \\ d_0|r_0 \\ (r_0, r)=1}} \frac{y_{\mathbf{r}, r_0}}{\phi_{\omega^*}(r_r)}. \quad (52)$$

We are now ready to apply Selberg’s sieve, which gives that

$$\begin{aligned} S_6^{(b)}(\nu_1) &\leq \frac{1}{\tilde{\lambda}_1^2} \sum_{d_0, e_0, \mathbf{d}, \mathbf{e}} \tilde{\lambda}_{d_0} \tilde{\lambda}_{e_0} \lambda_{\mathbf{d}} \lambda_{\mathbf{e}} \sum_{\substack{x < n \leq 2x \\ [d_i, e_i] | \ell_i(n) \\ p|d_0 e_0 \Rightarrow p|\ell(n) \\ n \equiv 1 \pmod{4} \\ n \equiv \nu_1 \pmod{q_3^2 W^2}}} |\mathcal{A}_{[d_0, e_0]}| \\ &= \frac{1}{\tilde{\lambda}_1^2} \sum_{\substack{d_0, e_0, \mathbf{d}, \mathbf{e} \\ (d_0 e_0, de)=1}} \tilde{\lambda}_{d_0} \tilde{\lambda}_{e_0} \lambda_{\mathbf{d}} \lambda_{\mathbf{e}} \left(\frac{x g([d_0, e_0])}{4 q_3^2 W^2 \prod_{i=1}^K [d_i, e_i]} + O(\tau([d_0, e_0])) \right). \end{aligned}$$

The contribution from the $O(\tau([d_0, e_0]))$ term satisfies

$$\frac{1}{\tilde{\lambda}_1^2} \sum_{\substack{d_0, e_0, \mathbf{d}, \mathbf{e} \\ (d_0 e_0, de)=1}} |\tilde{\lambda}_{d_0}| |\tilde{\lambda}_{e_0}| |\lambda_{\mathbf{d}}| |\lambda_{\mathbf{e}}| \tau([d_0, e_0]) \ll R^{2+o(1)} \frac{1}{\tilde{\lambda}_1^2} \sum_{d_0, e_0} |\tilde{\lambda}_{d_0}| |\tilde{\lambda}_{e_0}| \tau([d_0, e_0]).$$

By construction of $\tilde{\lambda}_1$ and $\tilde{\lambda}_{d_0}$, we always have that $|\tilde{\lambda}_{d_0}|/|\tilde{\lambda}_1| \leq \frac{d_0^2}{\phi(d_0)^2}$. Thus, this term contributes $\ll_{\varepsilon} R^{2+\varepsilon} x^{2\xi+\varepsilon} \ll x^{\theta_2+2\xi+2\varepsilon}$, which is negligible since $\xi < 1/4$.

The remaining ‘main’ term is given by

$$\frac{x}{4 q_3^2 W^2 \tilde{\lambda}_1^2} \sum_{\substack{d_0, e_0 \\ (d_0, e_0, q_3 W)=1}} \tilde{\lambda}_{d_0} \tilde{\lambda}_{e_0} g([d_0, e_0]) \sum_{\substack{\mathbf{d}, \mathbf{e} \in \mathcal{D}_K \\ (de, d_0 e_0)=1}} \frac{\lambda_{\mathbf{d}} \lambda_{\mathbf{e}}}{\prod_{i=1}^K [d_i, e_i]}. \quad (53)$$

Substituting the formula (52) for $\lambda_{\mathbf{d}}\tilde{\lambda}_{d_0}$ into (53), we get that (53) is equal to

$$\begin{aligned}
 &= \frac{x}{4q_3^2 W^2 \tilde{\lambda}_1^2} \sum_{\substack{\mathbf{d}, \mathbf{e} \in \mathcal{D}_K \\ d_0, e_0 \\ (d_i e_i, d_j e_j)=1 \forall 0 \leq i < j \leq k}} \frac{\mu(d_0 d) \mu(e_0 e) d_0^2 d e_0^2 e g([d_0, e_0])}{[\mathbf{d}, \mathbf{e}] \phi(d_0) \phi(e_0)} \\
 &\quad \times \sum_{\substack{\mathbf{d} | \mathbf{r}, \mathbf{e} | \mathbf{s} \\ d_0 | r_0, e_0 | s_0}} \frac{y_{\mathbf{r}, r_0} y_{\mathbf{s}, s_0}}{\phi_{\omega^*}(r_0 r) \phi_{\omega^*}(s_0 s)} \\
 &= \frac{x}{4q_3^2 W^2 \tilde{\lambda}_1^2} \sum_{\substack{\mathbf{r}, \mathbf{s} \in \mathcal{D}_K \\ r_0, s_0 \\ (r_0 s_0, r s)=1}} \frac{y_{\mathbf{r}, r_0} y_{\mathbf{s}, s_0}}{\phi_{\omega^*}(r_0 r) \phi_{\omega^*}(s_0 s)} \\
 &\quad \times \sum_{\substack{\mathbf{d} | \mathbf{r}, \mathbf{e} | \mathbf{s} \\ d_0 | r_0, e_0 | s_0 \\ (d_i e_i, d_j e_j)=1 \forall 0 \leq i < j \leq k}} \frac{\mu(d_0 d) \mu(e_0 e) d_0^2 d e_0^2 e g([d_0, e_0])}{[\mathbf{d}, \mathbf{e}] \phi(d_0) \phi(e_0)}.
 \end{aligned}$$

The inner sum is multiplicative over $p \mid r s r_0 s_0$, where the p th factor is given by

$$\sigma_p(\mathbf{r}, \mathbf{s}, r_0, s_0) = \begin{cases} p-1 & p \mid r_i, p \mid s_i, i \geq 1 \\ \frac{p^2}{p-1} - 1 & p \mid r_0, p \mid s_0 \\ -1 & p \mid r_i, p \mid s_j, i \neq j, i, j \geq 0 \\ 0 & p \text{ divides exactly one of } r r_0 \text{ and } s s_0. \end{cases}$$

The product $\prod_{p \mid r s r_0 s_0} \sigma_p(\mathbf{r}, \mathbf{s}, r_0, s_0)$ is 0 unless $r r_0 = s s_0$. Then using the bound $y_{\mathbf{r}, r_0} y_{\mathbf{s}, s_0} \leq y_{\mathbf{r}, r_0}^2 + y_{\mathbf{s}, s_0}^2$ we see (by symmetry) that (53) is

$$\begin{aligned}
 &\leq \frac{x}{4q_3^2 W^2 \tilde{\lambda}_1^2} \sum_{\mathbf{r}, r_0} \frac{y_{\mathbf{r}, r_0}^2}{\phi_{\omega^*}^2(r r_0)} \sum_{\substack{\mathbf{s}, s_0 \\ s s_0 = r r_0}} \prod_{p \mid r r_0} |\sigma_p(\mathbf{r}, \mathbf{s}, r_0, s_0)| \\
 &= \frac{x}{4q_3^2 W^2 \tilde{\lambda}_1^2} \sum_{\mathbf{r}, r_0} \frac{y_{\mathbf{r}, r_0}^2}{\phi_{\omega^*}^2(r r_0)} \prod_{p \mid r} (K + p - 1) \prod_{p \mid r_0} \left(K + \frac{p^2}{p-1} - 1 \right) \\
 &\leq \frac{x}{4q_3^2 W^2 \tilde{\lambda}_1^2} \sum_{\mathbf{r}, r_0} y_{\mathbf{r}, r_0}^2 \prod_{p \mid r r_0} \left(\frac{K + \frac{p^2}{p-1} - 1}{(p - K - 1)^2} \right) \\
 &\ll \frac{x}{4q_3^2 W^2 \tilde{\lambda}_1^2} \sum_{\mathbf{r}, r_0} \frac{y_{\mathbf{r}, r_0}^2}{\prod_{p \mid r r_0} (p + O(K))}. \tag{54}
 \end{aligned}$$

In order to estimate this sum, we wish to express $y_{\mathbf{r}, r_0}$ in terms of $y_{\mathbf{r}}$ and \tilde{y}_{r_0} . This is very similar to the computation done in [14, Proposition 9.4]. Writing $y_{\mathbf{r}, r_0}$ as in (51) and

using the definition of $\lambda_{\mathbf{d}}$ in (31) and $\tilde{\lambda}_{d_0}$ in (49), we get

$$\begin{aligned} y_{\mathbf{r}, r_0} &= \mu(r_0 r) \phi_{\omega^*}(r_0 r) \sum_{\substack{\mathbf{r}|\mathbf{d} \\ r_0|d_0 \\ (d_0, d)=1}} \frac{\lambda_{\mathbf{d}} \tilde{\lambda}_{d_0} \phi(d_0)}{d d_0^2} \\ &= \mu(r_0 r) \phi_{\omega^*}(r_0 r) \sum_{r_0|d_0} \mu(d_0) \sum_{d_0|f_0} \frac{y_{f_0}}{\phi(f_0)} \sum_{\substack{\mathbf{r}|\mathbf{d} \\ (d, d_0)=1}} \mu(d) \sum_{\mathbf{d}|\mathbf{f}} \frac{y_{\mathbf{f}}}{\phi(f)} \\ &= \mu(r_0 r) \phi_{\omega^*}(r_0 r) \sum_{\substack{f_0, \mathbf{f} \\ r_0|f_0, \mathbf{r}|\mathbf{f}}} \frac{y_{\mathbf{f}} y_{f_0}}{\phi(f_0) \varphi(f)} \sum_{\substack{d_0, \mathbf{d} \\ r_0|d_0, \mathbf{r}|\mathbf{d} \\ d_0|f_0, \mathbf{d}|\mathbf{f} \\ (d, d_0)=1}} \mu(d) \mu(d_0). \end{aligned}$$

The inner sum is 0 unless every prime dividing one of f and f_0 but not the other is a divisor of rr_0 ; in that case, the inner sum is ± 1 . Thus, using the fact that $y_{\mathbf{r}} \geq y_{\mathbf{f}}$ (since F is decreasing), as well as the fact that $\tilde{y}_{r_0} \geq \tilde{y}_{f_0}$, we get that

$$y_{\mathbf{r}, r_0} \leq \phi_{\omega^*}(r_0 r) y_{\mathbf{r}} \tilde{y}_{r_0} \sum_{\substack{f_0 \\ r_0|f_0 \\ (f_0, q_3 W)=1}} \sum_{\substack{\mathbf{f} \in \mathcal{D}_K \\ \mathbf{r}|\mathbf{f} \\ f f_0 / (f, f_0)^2 | r r_0}} \frac{\mu^2(f_0)}{\phi(f_0) \phi(f)}.$$

Let $f_0 = r_0 f'_0 g_0$ and $f_i = r_i f'_i g_i$ for $1 \leq i \leq K$, where $f'_i = f_i / (f_i, r r_0)$ is the largest divisor of f_i that is relatively prime to rr_0 . In particular, $g_0 | r$ and $g_i | r_0$ for $1 \leq i \leq K$. Since $f f_0 / (f, f_0)^2 | r r_0$, we must have $f'_0 = \prod_{i=1}^K f'_i$. Thus, $y_{\mathbf{r}, r_0}$ is bounded by

$$\begin{aligned} y_{\mathbf{r}, r_0} &\leq \phi_{\omega^*}(r_0 r) y_{\mathbf{r}} \tilde{y}_{r_0} \frac{1}{\phi(r_0 r)} \sum_{\mathbf{f}' \in \mathcal{D}_K} \frac{1}{\phi^2(f')} \sum_{\substack{\mathbf{g} \in \mathcal{D}_K \\ g_i | r_0 \forall 1 \leq i \leq K}} \frac{1}{\phi(g)} \sum_{g_0 | r} \frac{1}{\phi(g_0)} \\ &\leq \phi_{\omega^*}(r_0 r) y_{\mathbf{r}} \tilde{y}_{r_0} \prod_{\substack{p > D_0 \\ p \equiv 3 \pmod{4}}} \left(1 + \frac{K}{(p-1)^2}\right) \prod_{p | r_0} \left(1 + \frac{K}{(p-1)}\right) \prod_{p | r} \left(1 + \frac{1}{(p-1)}\right). \end{aligned}$$

The first product is $\ll O_K(1)$. By the definition of ϕ_{ω^*} , we then have

$$y_{\mathbf{r}, r_0} \ll y_{\mathbf{r}} \tilde{y}_{r_0} \prod_{p | r_0} \left(1 + \frac{K}{p-1}\right) \left(1 - \frac{K+1}{p}\right) \prod_{p | r} \left(1 + \frac{1}{p-1}\right) \left(1 - \frac{K+1}{p}\right),$$

which in turn implies that $y_{\mathbf{r}, r_0} \ll y_{\mathbf{r}} \tilde{y}_{r_0}$ because both products are ≤ 1 .

Plugging this into (54), we get that

$$S_6^{(b)}(\nu_1) \ll \frac{x}{q_3^2 W^2 \tilde{\lambda}_1^2} \left(\sum_{\substack{r_0 \leq x^\epsilon \\ (r_0, W)=1}} \frac{\tilde{y}_{r_0}^2}{\prod_{p | r_0} (p + O(K))} \right) \left(\sum_{\mathbf{r} \in \mathcal{D}_K} \frac{y_{\mathbf{r}}^2}{\prod_{p | r} (p + O(K))} \right). \quad (55)$$

Recalling that $y_{r_0} = 1$ for $r_0 \leq x^\xi$ and $(r_0, q_3 W) = 1$, we have

$$\sum_{\substack{r_0 \leq x^\xi \\ (r_0, q_3 W) = 1}} \frac{y_{r_0}^2}{\prod_{p|r_0} (p + O(K))} \ll \left(\frac{\xi \log x}{\log D_0} \right)^{1/2}.$$

We can bound the sum over \mathbf{r} using Lemma 7. From the definition of $L_K(F)$, we then get

$$\sum_{\mathbf{r} \in \mathcal{D}_K} \frac{y_{\mathbf{r}}^2}{\prod_{p|\mathbf{r}} (p + O(K))} \ll \left(\frac{e^{-\gamma/2}}{\Gamma(1/2)} \right)^K \left(\frac{\log R}{\log D_0} \right)^{K/2} L_K(F).$$

Using these estimates as well as the bound (50) on $\lambda_{\mathbf{r}}$, equation (55) becomes

$$\begin{aligned} S_6^{(b)}(\nu_1) &\ll_K \frac{x}{q_3^2 W^2} \frac{\log D_0}{\xi \log x} \left(\frac{\xi \log x}{\log D_0} \right)^{1/2} \left(\frac{\log R}{\log D_0} \right)^{K/2} L_K(F) \\ &\ll \frac{x}{q_3^2 W^2} \xi^{-1/2} \theta_2^{-1/2} \left(\frac{\log R}{\log D_0} \right)^{\frac{K-1}{2}} L_K(F), \end{aligned}$$

as desired. \square

4. Singular series estimates

We now prove several computational lemmas providing bounds on sums over \mathbf{E} -admissible tuples $\mathcal{L} = \mathcal{L}(\mathbf{b})$ for $\mathbf{b} \in \mathcal{B}$. We begin with an average that appears in the sums over the terms $S_1(\nu_0)$ through $S_5^{(m)}(\nu_0)$ in the proof of Theorem 1, before turning to bounding $S_6^{(b)}(\nu_1)$ on average over different values of b .

4.1. Averaging over \mathcal{B} and ν_0

Lemma 21. *In the notation of Section 2,*

$$\sum_{\substack{\mathbf{b} \in \mathcal{B} \\ \mathcal{L} = \mathcal{L}(\mathbf{b}) \text{ adm.}}} \sum_{\substack{\nu_0 \bmod W \\ (\ell(\nu_0), W) = 1 \forall \ell \in \mathcal{L}}} 1 \gg_K \left(\frac{\eta}{q} \right)^{K-1} (\log x)^{\frac{K-1}{2}} \left(\frac{\phi(W)}{W} \right)^K W.$$

Proof. We first consider the number of $\mathbf{b} \in \mathcal{B}$ that will produce an admissible tuple $\mathcal{L}(\mathbf{b})$. If $\mathcal{L}(\mathbf{b})$ is not admissible, then there is some prime $p \leq K$, $p \equiv 3 \pmod{4}$, $p \nmid q$ such that $\prod_{\ell \in \mathcal{L}(\mathbf{b})} \ell(n)$ is always divisible by p . In order to prevent this situation, we can consider only those \mathbf{b} for which each b_i , $i \geq 2$ satisfies $q + a_i + qb_i \not\equiv 0, 1 \pmod{p}$ for all $p \leq K$, $p \equiv 3 \pmod{4}$, $p \nmid 2q$. Having excluded two congruence classes for each prime p , together with the linear form corresponding to b_1 , the tuple $\mathcal{L}(\mathbf{b})$ cannot cover all of the congruence classes mod p .

Thus, for each $2 \leq i \leq K$, we can choose b_i from a set of size

$$\frac{\eta}{8q} \sqrt{\log x} \prod_{\substack{p|W \\ 2 < p \leq K}} \left(1 - \frac{2}{p} \right) \gg_K \frac{\eta}{q} \sqrt{\log x}$$

while ensuring that the resulting \mathbf{b} is admissible. It follows that there are at least $\gg_K \left(\frac{n}{q}\right)^{K-1} (\log x)^{(K-1)/2}$ choices of $\mathbf{b} \in \mathcal{B}$ with $\mathcal{L}(\mathbf{b})$ admissible.

For each \mathbf{b} with $\mathcal{L}(\mathbf{b})$, we now consider the sum over ν_0 . For fixed \mathbf{b} , this is bounded by

$$\sum_{\substack{\nu_0 \bmod W \\ (\ell(\nu_0), W)=1 \forall \ell \in \mathcal{L}}} 1 \gg W \prod_{\substack{p|W \\ 2 < p \leq K}} \left(\frac{1}{p}\right) \prod_{\substack{p|W \\ p > K}} \left(1 - \frac{K}{p}\right) \gg_K W \left(\frac{\phi(W)}{W}\right)^K,$$

which, along with the number of choices of \mathbf{b} yielding admissible tuples, completes the proof. \square

4.2. Averaging over $S_6^{(b)}(\nu_1)$

In this section, we will analyze the sum over $S_6^{(b)}(\nu_1)$ terms appearing in the proof of Theorem 1, and in particular provide the proof of Lemma 3. To begin with, the $S_6^{(b)}(\nu_1)$ sum can be bounded via Theorem 2 by

$$\begin{aligned} & \sum_{\substack{\mathbf{b} \in \mathcal{B} \\ \mathcal{L} = \mathcal{L}(\mathbf{b}) \text{ adm.}}} \sum_{\substack{\nu_1 \bmod W^2 \\ (\ell(\nu_1), W)=1 \forall \ell \in \mathcal{L}}} \sum_{\substack{b \leq \eta \sqrt{\log x} \\ qn+b \notin \mathcal{L} \\ (q\nu_1+b, W^2)=\square}} S_6^{(b)}(\nu_1) \\ & \ll_K \frac{x}{q_3^2 W^2} \xi^{-1/2} \theta_2^{-1/2} \left(\frac{\log R}{\log D_0}\right)^{\frac{K-1}{2}} L_K(F) \sum_{\substack{\mathbf{b} \in \mathcal{B} \\ \mathcal{L} = \mathcal{L}(\mathbf{b}) \text{ adm.}}} \sum_{\substack{b \leq \eta \sqrt{\log x} \\ qn+b \notin \mathcal{L}}} \sum_{\substack{\nu_1 \bmod W^2 \\ (\ell(\nu_1), W)=1 \forall \ell \in \mathcal{L} \\ (q\nu_1+b, W^2)=\square}} 1. \end{aligned} \quad (56)$$

Our next task is estimating the sums over \mathbf{b}, b and ν_1 . The constraints on $\nu_1 \bmod W^2$ are multiplicative, so we can understand them separately for each $p \mid W$. For a fixed $p \mid W$, let $\tilde{N}_{p^2}(\mathcal{L}, b)$ denote the number of congruence classes $\nu \bmod p^2$ such that $p \mid \ell(\nu)$ for some $\ell \in \mathcal{L}$ or such that $p \mid \ell^b(\nu)$ but $p^2 \nmid \ell^b(\nu)$. Then we have, for fixed $\mathcal{L}(\mathbf{b})$ and fixed b , that

$$\sum_{\substack{\nu_1 \bmod W^2 \\ (\ell(\nu_1), W)=1 \forall \ell \in \mathcal{L} \\ (q\nu_1+b, W^2)=\square}} 1 = \prod_{p|W} (p^2 - \tilde{N}_{p^2}(\mathcal{L}, b)) = W^2 \left(\frac{\phi(W)}{W}\right)^{K+1} \prod_{p|W} \frac{1 - \tilde{N}_{p^2}(\mathcal{L}, b)/p^2}{(1 - 1/p)^{K+1}}.$$

The remaining sum over \mathbf{b} and b is bounded in the following proposition.

Proposition 22. *We have*

$$\sum_{\substack{\mathbf{b} \in \mathcal{B} \\ \mathcal{L}(\mathbf{b}) \text{ adm.}}} \sum_{\substack{b \leq \eta \sqrt{\log x} \\ qn+b \notin \mathcal{L}}} \prod_{p|W} \frac{1 - \tilde{N}_{p^2}(\mathcal{L}, b)/p^2}{(1 - 1/p)^{K+1}} \ll_K \frac{(\eta \sqrt{\log x})^K}{(8q)^{K-1}},$$

where the implied constant depends only on K .

Plugging this estimate into (56), we have

$$\sum_{\substack{\mathbf{b} \in \mathcal{B} \\ \mathcal{L} = \mathcal{L}(\mathbf{b})}} \sum_{\substack{\nu_1 \bmod W^2 \\ \text{adm. } (\ell(\nu_1), W) = 1 \forall \ell \in \mathcal{L}}} \sum_{\substack{b \leq \eta \sqrt{\log x} \\ qn + b \notin \mathcal{L} \\ (q\nu_1 + b, W^2) = \square}} S_6^{(b)} \\ \ll_K \frac{x}{q_3^2 W^2} L_K(F) \xi^{-1/2} \theta_2^{-1/2} \left(\frac{\log R}{\log D_0} \right)^{\frac{K-1}{2}} W^2 \frac{\phi(W)^{K+1}}{W^{K+1}} \frac{(\eta \sqrt{\log x})^K}{(8q)^{K-1}}.$$

Using the fact that

$$\frac{\phi(W)}{W} \asymp \frac{q_3}{\phi(q_3)} (\log D_0)^{-1}$$

and that $\log R = \theta_2 \log x$, we get

$$\sum_{\substack{\mathbf{b} \in \mathcal{B} \\ \mathcal{L} = \mathcal{L}(\mathbf{b})}} \sum_{\substack{\nu_1 \bmod W^2 \\ \text{adm. } (\ell(\nu_1), W) = 1 \forall \ell \in \mathcal{L}}} \sum_{\substack{b \leq \eta \sqrt{\log x} \\ qn + b \notin \mathcal{L} \\ (q\nu_1 + b, W^2) = \square}} S_6^{(b)} \\ \ll_K \xi^{-1/2} \theta_2^{\frac{K}{2}-1} L_K(F) \left(\frac{q_3}{\phi(q_3)} \right)^{K+1} \frac{\eta^K}{q_3^2 q^{K-1}} \frac{(\log x)^{K-\frac{1}{2}}}{(\log D_0)^K} x,$$

which completes the proof of Lemma 3.

It remains to prove Proposition 22. To do so, we will make use of the following lemma.

Lemma 23. Let $\tilde{N}_{p^2}(\mathcal{L}, b)$ denote the number of congruence classes $\nu \bmod p^2$ such that $p \mid \ell(\nu)$ for some $\ell \in \mathcal{L}$ or such that $p \mid \ell^b(\nu)$ but $p^2 \nmid \ell^b(\nu)$. Let $N_p(\mathcal{L}, b)$ denote the number of congruence classes $\nu \bmod p$ such that $p \mid \ell(\nu)$ for some $\ell \in \mathcal{L}$ or such that $p \mid \ell^b(\nu)$. Then

$$\prod_{p|W} \frac{1 - \tilde{N}_{p^2}(\mathcal{L}, b)/p^2}{(1 - 1/p)^{K+1}} \ll_K \prod_{\substack{p|W \\ p > K+1}} \frac{1 - N_p(\mathcal{L}, b)/p}{(1 - 1/p)^{K+1}},$$

where the implied constant depends only on K .

Proof. By definition, $\tilde{N}_{p^2}(\mathcal{L}, b)$ almost consists of all elements of a certain set of congruence classes modulo p when lifted to $\mathbb{Z}/p^2\mathbb{Z}$, with the possible exception of one congruence class ν modulo p^2 such that $\ell^b(\nu) \equiv 0 \bmod p^2$. In particular, this implies that

$$\tilde{N}_{p^2}(\mathcal{L}, b) = pN_p(\mathcal{L}, b) - E,$$

where E is either 0 or 1. Thus,

$$\prod_{p|W} \frac{1 - \tilde{N}_{p^2}(\mathcal{L}, b)/p^2}{(1 - 1/p)^{K+1}} \leq \prod_{p|W} \frac{1 - N_p(\mathcal{L}, b)/p + 1/p^2}{(1 - 1/p)^{K+1}},$$

since the numerator of each factor in the product is either unchanged or has increased.

We can then rewrite the right-hand side as

$$\begin{aligned} \prod_{p|W} \frac{1 - N_p(\mathcal{L}, b)/p + 1/p^2}{(1 - 1/p)^{K+1}} &= \prod_{\substack{p|W \\ p \leq K+1}} \frac{1 - N_p(\mathcal{L}, b)/p + 1/p^2}{(1 - 1/p)^{K+1}} \prod_{\substack{p|W \\ p > K+1}} \frac{1 - N_p(\mathcal{L}, b)/p + 1/p^2}{(1 - 1/p)^{K+1}} \\ &\leq \prod_{p \leq K+1} \frac{1 - 1/p}{(1 - 1/p)^{K+1}} \prod_{\substack{p|W \\ p > K+1}} \frac{(1 - N_p(\mathcal{L}, b)/p)}{(1 - 1/p)^{K+1}} \left(1 + \frac{1}{p(p - N_p(\mathcal{L}, b))} \right) \\ &\ll_K \prod_{\substack{p|W \\ p > K+1}} \frac{(1 - N_p(\mathcal{L}, b)/p)}{(1 - 1/p)^{K+1}} \prod_{\substack{p|W \\ p > K+1}} \left(1 + \frac{1}{p(p - K - 1)} \right). \end{aligned}$$

The second Euler product converges to a constant dependent only on K when extended over all primes $p > K + 1$, which completes the proof. \square

We are now ready to prove Proposition 22. This estimate is an analog of Gallagher's result [5] that the average value of the singular series constants appearing in the Hardy–Littlewood k -tuples conjecture is 1. Our proof will closely follow Gallagher's argument.

Proof of Proposition 22. We begin by applying Lemma 23 to bound the left-hand side by

$$\ll_K \sum_{\substack{\mathbf{b} \in \mathcal{B} \\ \mathcal{L}(\mathbf{b}) \text{ adm.}}} \sum_{\substack{b \leq \eta \sqrt{\log x} \\ qn + b \notin \mathcal{L}}} \prod_{\substack{p|W \\ p > K+1}} \frac{1 - N_p(\mathcal{L}, b)/p}{(1 - 1/p)^{K+1}},$$

where $N_p(\mathcal{L}, b)$ is the number of congruence classes $\nu \bmod p$ such that $p|\ell(\nu)$ for some $\ell \in \mathcal{L}$ or such that $p|\ell^b(\nu)$.

Let $\Delta(\mathcal{L}, b)$ denote the product

$$\Delta(\mathcal{L}, b) := \prod_{1 \leq i_1 < i_2 \leq K} (q(b_{i_2} - b_{i_1}) + a_{i_2} - a_{i_1}) \prod_{1 \leq i \leq K} (b - qb_i - a_i).$$

Thus, $1 \leq N_p(\mathcal{L}, b) \leq K + 1$, with equality on the right if and only if $p \nmid \Delta(\mathcal{L}, b)$. Define $a(p, N_p)$ via

$$\frac{1 - N_p/p}{(1 - 1/p)^{K+1}} = 1 + a(p, N_p),$$

and for squarefree r , define $a_{\mathcal{L}, b}(r)$ multiplicatively via $a_{\mathcal{L}, b}(r) = \prod_{p|r} a(p, N_p(\mathcal{L}, b))$, so that

$$\prod_{p|W} \frac{1 - N_p(\mathcal{L}, b)/p}{(1 - 1/p)^{K+1}} = \sum_{r|W} a_{\mathcal{L}, b}(r).$$

By the same reasoning as in Gallagher's proof of equation (3) in [5], for a constant x to be fixed later and for all $\varepsilon > 0$, we have

$$\begin{aligned} & \sum_{\substack{\mathbf{b} \in \mathcal{B} \\ \mathcal{L}(\mathbf{b}) \text{ adm.}}} \sum_{\substack{b \leq \eta\sqrt{\log x} \\ qn+b \notin \mathcal{L}}} \prod_{\substack{p|W \\ p > K+1}} \frac{1 - N_p(\mathcal{L}, b)/p}{(1 - 1/p)^{K+1}} \\ &= \sum_{r \leq x} \sum_{\substack{\mathbf{b} \in \mathcal{B} \\ \mathcal{L}(\mathbf{b}) \text{ adm.}}} \sum_{\substack{b \leq \eta\sqrt{\log x} \\ qn+b \notin \mathcal{L}}} a_{\mathcal{L}, b}(r) + O_{K, \varepsilon}(\eta^K (\log x)^{K/2} (x\eta \log x)^\varepsilon / x). \end{aligned} \quad (57)$$

The inner sums over $\mathbf{b} \in \mathcal{B}$ and b are equal to

$$\sum_{(N_p)_p} \prod_{\substack{p|r \\ p > K+1}} a(p, N_p) \left\{ \sum' 1 + O((\eta\sqrt{\log x})^{K-1}) \right\},$$

where each $(N_p)_p$ is a vector with positive integer entries for each prime $p|r$ with $p > K+1$, and where \sum' denotes the number of ways to choose values b_2, \dots, b_K and b , not necessarily distinct, such that each $b_i \equiv 3 \pmod{4}$, such that $1 < qb_i \leq \frac{\eta}{2}\sqrt{\log x}$ for $2 \leq i \leq jM$, such that $\frac{\eta}{2} < qb_i \leq \eta\sqrt{\log x}$ for $jM+1 \leq i \leq K$, such that $b \leq \eta\sqrt{\log x}$, and most crucially, such that b_1, \dots, b_K, b occupy precisely N_p congruence classes modulo p for each $p|r$. Recall that $b_1 = 3$ is fixed for all $\mathbf{b} \in \mathcal{B}$.

By the Chinese remainder theorem, for $r \leq \eta\sqrt{\log x}$,

$$\sum' = \left\{ \left(\frac{\eta\sqrt{\log x}}{8qr} \right)^{K-1} \left(\frac{\eta\sqrt{\log x}}{r} \right) + O \left(\frac{\eta\sqrt{\log x}}{r} \right)^{K-1} \right\} \times \prod_{p|r} \binom{p-1}{N_p-1} \sigma(K, N_p),$$

where $\sigma(K, N_p)$ denotes the number of surjective maps from $\{1, \dots, K\}$ onto $\{1, \dots, N_p\}$. Thus, the inner sum is

$$\left(\frac{\eta\sqrt{\log x}}{8qr} \right)^{K-1} \left(\frac{\eta\sqrt{\log x}}{r} \right) A(r) + O \left(\left(\frac{\eta\sqrt{\log x}}{r} \right)^{K-1} B(r) \right) + O \left((\eta\sqrt{\log x})^{K-1} C(r) \right),$$

where

$$\begin{aligned} A(r) &= \sum_{(N_p)_p} \prod_{p|r} a(p, N_p) \binom{p}{N_p} \sigma(K, N_p), \\ B(r) &= \sum_{(N_p)_p} \prod_{p|r} |a(p, N_p)| \binom{p}{N_p} \sigma(K, N_p), \text{ and} \\ C(r) &= \sum_{(N_p)_p} \prod_{p|r} |a(p, N_p)|. \end{aligned}$$

One can show via a combinatorial argument (identical to the one performed in [5]) that $A(r) = 0$ whenever $r > 1$. Also by the same arguments as in [5], we have $B(r) \leq C^{\omega(r)} \frac{r^K}{\phi(r)}$ and $C(r) \leq C^{\omega(r)} \frac{r}{\phi(r)}$ for a suitable constant C .

Altogether, we get that

$$\begin{aligned}
 & \sum_{\substack{\mathbf{b} \in \mathcal{B} \\ \mathcal{L}(\mathbf{b}) \text{ adm.}}} \sum_{\substack{b \leq \eta\sqrt{\log x} \\ qn+b \notin \mathcal{L}}} \prod_{\substack{p|W \\ p > K+1}} \frac{1 - N_p(\mathcal{L}, b)/p}{(1 - 1/p)^{K+1}} \\
 &= \left(\frac{\eta\sqrt{\log x}}{8q} \right)^{K-1} (\eta\sqrt{\log x}) + O \left((\eta\sqrt{\log x})^{K-1} \sum_{r \leq x} \frac{C^{\omega(r)} r}{\phi(r)} \right) \\
 &\quad + O(\eta^K (\log x)^{K/2} (x\eta \log x)^\varepsilon / x) \\
 &= \frac{(\eta\sqrt{\log x})^K}{(8q)^{K-1}} + O_{K, \varepsilon, q}((\eta\sqrt{\log x})^{K-1/2+\varepsilon}),
 \end{aligned}$$

choosing $x = (\eta\sqrt{\log x})^{1/2}$. This completes the proof. \square

Competing interests. The authors have no competing interests to declare.

Data availability statement. Data availability is not applicable to this article as no data was created or analysed.

References

- [1] BANKS WD, FREIBERG T AND MAYNARD J (2016) On limit points of the sequence of normalized prime gaps. *Proc. Lond. Math. Soc. (3)* **113**(4), 515–539. MR 3556490
- [2] BANKS WD, FREIBERG T AND TURNAGE-BUTTERBAUGH CL (2015) Consecutive primes in tuples. *Acta Arith.* **167**(3), 261–266. MR 3316460
- [3] DAVID C, DEVIN L, NAM J AND SCHLITT J (2022) Lemke Oliver and Soundararajan bias for consecutive sums of two squares. *Math. Ann.* **384**(3–4), 1181–1242. MR 4498471
- [4] FREIBERG T, KURLBERG P AND ROSENZWEIG L (2017) Poisson distribution for gaps between sums of two squares and level spacings for toral point scatterers. *Commun. Number Theory Phys.* **11**(4), 837–877. MR 3733767
- [5] GALLAGHER PX (1976) On the distribution of primes in short intervals. *Mathematika* **23**(1), 4–9. MR 409385
- [6] GOLDSTON DA, GRAHAM SW, PINTZ J AND YILID RIM CY (2009) Small gaps between products of two primes. *Proc. Lond. Math. Soc. (3)* **98**(3), 741–774. MR 2500871
- [7] GORODETSKY O (2023) Sums of two squares are strongly biased towards quadratic residues. *Algebra Number Theory* **17**(3), 775–804. MR 4578006
- [8] HOOLEY C (1963) On the number of divisors of a quadratic polynomial. *Acta Math.* **110**, 97–114. MR 153648
- [9] HOOLEY C (1971) On the intervals between numbers that are sums of two squares. *Acta Math.* **127**, 279–297. MR 294281
- [10] KIMMEL N AND KUPERBERG V (2023) Consecutive runs of sums of two squares. arXiv preprint 2306.12855.
- [11] KNAPOWSKI S AND TURÁN P (1977) On prime numbers $\equiv 1$ resp. 3. In *Number Theory and Algebra*. New York : Academic Press, 157–165. MR 0466043
- [12] LEMKE OLIVER RJ AND SOUNDARARAJAN K (2016) Unexpected biases in the distribution of consecutive primes. *Proc. Natl. Acad. Sci. USA* **113**(31), E4446–E4454. MR 3624386
- [13] MAYNARD J (2015) Small gaps between primes. *Ann. of Math. (2)* **181**(1), 383–413. MR 3272929

- [14] MAYNARD J (2016) Dense clusters of primes in subsets. *Compos. Math.* **152**(7), 1517–1554. MR 3530450
- [15] MCGRATH O (2022) A variation of the prime k-tuples conjecture with applications to quantum limits. *Math. Ann.* **384**(3–4), 1343–1407. MR 4498475
- [16] MERIKOSKI J (2020) Limit points of normalized prime gaps. *J. Lond. Math. Soc. (2)* **102**(1), 99–124. MR 4143728
- [17] PINTZ J (2018) A note on the distribution of normalized prime gaps. *Acta Arith.* **184**(4), 413–418. MR 3855375
- [18] PŁAKSIN VA (1988) The distribution of numbers represented by the sum of two squares. *Dokl. Akad. Nauk SSSR* **299**(6), 1320–1323. MR 947418
- [19] SHIU DKL (2000) Strings of congruent primes. *J. London Math. Soc. (2)* **61**(2), 359–373. MR 1760689

Index of Notation

Averaging argument

$\ell^b(n)$, 7

$\ell_i(n)$, 6

η , 8

b , 6

\mathcal{B} , 6

$\mathcal{L}, \mathcal{L}(b)$, 6

ξ , 7

B_i , 7

b_i , 6

K, k , 6

$S(\xi)$, 7

u , 7

Main problem setup

\mathbf{E} , 2

\tilde{a}_1, \tilde{a}_2 , 3

a_i , 3

M, M_1 , 3

$N(x)$, 2

$N(x; q, a)$, 3

q , 3

q_1, q_3 , 7

Sieve parameters

λ_d , 15

D_K , 14

θ_2 , 8

B , 8

D_0 , 8

F , 14

$g(t)$, 15

$L_k(F), L_{k;m}(F), L_{k;m_1, m_2}(F)$, 15

R , 8

R_K , 14

W , 8

$w_n(\mathcal{L})$, 15

y_r , 15

Sums of two squares

X_{x_0} , 22

$Z_{x,Q}^{(1)}$, 22

$Z_{x,Q}^{(2)}$, 22

$\Psi(u, t)$, 21

$c_t(h)$, 21

χ_4 , 8

$\rho(n)$, 5

$y\theta_1$, 5

A , 8

$g_1(n)$, 21

$g_2(n)$, 5

$g_3(n), g_4(n), g_5(n), g_6(n)$, 21

$g_7(n)$, 22

$r_2(n)$, 5

$t(n)$, 8

V , 8

v , 5