



# A variation of the prime $k$ -tuples conjecture with applications to quantum limits

Oliver McGrath<sup>1</sup>

Received: 3 December 2020 / Revised: 15 November 2021 / Accepted: 15 November 2021 /

Published online: 3 December 2021

© The Author(s) 2021

## Abstract

Let  $\mathcal{H}^* = \{h_1, h_2, \dots\}$  be an ordered set of integers. We give sufficient conditions for the existence of increasing sequences of natural numbers  $a_j$  and  $n_k$  such that  $n_k + h_{a_j}$  is a sum of two squares for every  $k \geq 1$  and  $1 \leq j \leq k$ . Our method uses a novel modification of the Maynard–Tao sieve together with a second moment estimate. As a special case of our result, we deduce a conjecture due to D. Jakobson which has several implications for quantum limits on flat tori.

## 1 Introduction

We say that a set  $\mathcal{H} = \{h_1, \dots, h_k\}$  of distinct integers is *admissible* if  $\#\{\mathcal{H} \pmod{p}\} < p$  for every prime  $p$ . An outstanding problem in analytic number theory is the prime  $k$ -tuples conjecture, which asserts the following.

**Conjecture 1.1** *Let  $\mathcal{H} = \{h_1, \dots, h_k\}$  be admissible. Then there exists infinitely many integers  $n$  such that the translates  $n + h_1, \dots, n + h_k$  are prime.*

A proof of this conjecture is far out of reach of current techniques. However, we have been successful in establishing various weak versions of this result using sieve methods. For example, the Maynard–Tao sieve can be used to show that  $\gg \log k$  of the translates are simultaneously prime infinitely often, when  $k$  is sufficiently large (cf. [9,11]).

We extend the definition of admissibility to infinite ordered sets and say  $\mathcal{H}^* = \{h_1, h_2, \dots\}$  is admissible if the finite truncation  $\{h_1, \dots, h_k\} \subseteq \mathcal{H}^*$  is admissible for every  $k \geq 1$ . In this paper we are interested in the following variation of this conjecture, for numbers representable as a sum of two squares.

---

Communicated by Wei Zhang.

✉ Oliver McGrath  
oliver.mcgrath@maths.ox.ac.uk

<sup>1</sup> Mathematical Institute, University of Oxford, Oxford OX2 6GG, UK

**Conjecture 1.2** Let  $\mathcal{H}^* = \{h_1, h_2, \dots\}$  be admissible. Then there exists an increasing sequence of integers  $n_k$  such that, for every  $k \geq 1$ , the translates  $n_k + h_1, \dots, n_k + h_k$  are sums of two squares.

We remark that if we replaced “sums of two squares” with “prime” here, then this would simply be a reformulation of Conjecture 1.1. (It is easy to show that any finite admissible set can be extended to an infinite admissible set.)

Our interest in this version of the conjecture stems from a problem which appears towards the end of D. Jakobson’s “*Quantum limits on flat tori*” paper [8]. In this paper Jakobson is concerned with characterising the possible quantum limits that can arise on the standard flat  $d$ -dimensional torus  $\mathbb{T}^d = \mathbb{R}^d / \mathbb{Z}^d$ . A complete classification of such objects is established in two dimensions, with possible behaviours in higher dimensions described unconditionally for  $d \geq 4$ , and conditionally for  $d = 3$  on a weak version of Conjecture 1.2 (cf. [8, Conjecture 8.2]).

In this paper we establish Jakobson’s conjecture.

**Theorem 1.3** There exists increasing sequences of natural numbers  $a_j$  and  $M_k$  such that  $M_k - a_j^2$  is a sum of two squares for  $1 \leq j \leq k$ . Moreover, the sequence  $a_j$  is such that:

- (1)  $r_2(a_j) < r_2(a_{j+1})$  for all  $j \geq 1$ .
- (2) The even parts are uniformly bounded; that is to say, if we write  $a_j = 2^{b_j} m_j$  where  $(m_j, 2) = 1$ , then  $b_j = O(1)$  uniformly for  $j \geq 1$ .

Here,  $r_2(n)$  denotes the number of representations of  $n$  as a sum of two squares. We deduce Theorem 1.3 from the following general result.

**Theorem 1.4** Let  $\mathcal{H}^* = \{h_1, h_2, \dots\}$  be admissible such that each  $h_i$  is divisible by 4. Then there exists increasing sequences of natural numbers  $a_j$  and  $n_k$  such that  $n_k + h_{a_j}$  is a sum of two squares for every  $k \geq 1$  and  $1 \leq j \leq k$ .

For example, Theorem 1.4 applied to the admissible set  $\mathcal{H}^* = \{h_1, h_2, \dots\}$  with elements  $h_i = -(2 \cdot 5^i)^2$ , yields a solution to Theorem 1.3.

As mentioned above, Theorem 1.3 allows us conclude results about quantum limits on flat tori. Let  $(\lambda_j)_{j \geq 1}$  be a sequence of eigenvalues of the Laplace operator  $\nabla$  on  $\mathbb{T}^d$  such that  $\lambda_j \rightarrow \infty$ , and let  $\varphi_j$  be corresponding eigenfunctions with  $\|\varphi_j\|_2 = 1$ . If the sequence of probability measures  $d\mu_j = |\varphi_j|^2 dx$  has a weak- $*$  limit  $d\nu$ , then we call  $d\nu$  a *quantum limit*. (Here  $dx$  is the normalised Riemannian volume.)

It can be shown that all limits of such sequences  $d\mu_j$  are absolutely continuous with respect to the Lebesgue measure on  $\mathbb{T}^d$  (cf. [8, Theorem 1.3]), and so one can consider the Fourier expansion

$$d\nu = \sum_{\tau \in \mathbb{Z}^d} c_\tau e^{2\pi i \langle \tau, x \rangle} dx. \quad (1.1)$$

Among other things, Jakobson shows that in two dimensions all quantum limits are necessarily trigonometric polynomials (cf. [8, Theorem 1.2]). The same result isn’t true for  $d \geq 4$ , and conjecturally not true for  $d = 3$  either (cf. [8, Conjecture 8.2]) and

the following discussion). With Theorem 1.3, we can now complete this aspect of the classification of quantum limits on flat tori.

**Theorem 1.5** *There exists quantum limits on  $\mathbb{T}^3$  that are not trigonometric polynomials.*

As further consequences to Theorem 1.3 we are able to show the following results for quantum limits whose Fourier expansions are described as in (1.1).

**Theorem 1.6** *Let  $\epsilon > 0$ . We have the following.*

- (i) *For  $d \geq 4$  there exists quantum limits  $\nu$  on  $\mathbb{T}^d$  with densities that are not in  $l^{2-\epsilon}$  (i.e. for which  $\sum_{\tau} |c_{\tau}|^{2-\epsilon}$  diverges).*
- (ii) *For  $d \geq 5$  there exists quantum limits  $\nu$  on  $\mathbb{T}^d$  for which*

$$\limsup_{\rho \rightarrow \infty} \frac{\Sigma(\rho)}{\rho^{d-4-\epsilon}} = +\infty,$$

where  $\Sigma(\rho)$  is defined as

$$\Sigma(\rho) = \sum_{\substack{\tau \in \mathbb{Z}^d \\ |\tau| < \rho}} |c_{\tau}|. \tag{1.2}$$

The results contained in Theorem 1.6 improve upon various results found in [8]. Part (i) was previously shown for  $d \geq 5$ , and has now been extended to the case  $d = 4$  where it is now optimal (cf. [8, Theorem 1.4]). Part (ii) improves on the weaker lower bound

$$\limsup_{\rho \rightarrow \infty} \frac{\Sigma(\rho)}{\rho^{d-5-\epsilon}} = +\infty$$

which was shown for  $d \geq 6$ . The lower bound we prove is believed to be optimal for all  $d \geq 5$  (cf. [8, Proposition 1.2] and comments shortly after).

**Remark 1.7** It is well-known that the eigenvalues of  $\nabla$  on  $\mathbb{T}^d$  are the numbers  $4\pi^2 k$  for non-negative integers  $k$ , and they occur with multiplicity  $r_d(k)$  (the number of representations of  $k$  as the sum of  $d$  squares). This means various constructions associated to quantum limits on flat tori can be translated to problems in number theory involving sums of squares.

**Remark 1.8** Jakobson shows how Theorem 1.3 follows from weak form of the prime  $k$ -tuples conjecture, essentially by using the fact primes  $p \equiv 1 \pmod{4}$  are the sum of two squares (cf. discussion at the end of [8, Section 8]). We note that the weak form of the conjecture Jakobson uses is still far out of reach of current methods.

## 2 Outline of new sieve ideas

In this section, let  $\mathcal{A} \subseteq \mathbb{N}$  denote a set of arithmetic interest, which for our purposes is the set of numbers representable as a sum of two squares (but the following discussion

holds more generally). We will denote random variables by boldfaced letters, for example  $\mathbf{X}$ . We will let  $\mathbb{P}(\cdot)$  denote a probability measure and by  $\mathbb{E}[\cdot]$  the expectation operator.

### 2.1 A model problem

Our aim is to prove Theorem 1.4. By a pigeonhole argument (see Proposition 5.1), it suffices to consider the following model problem.

**Model Problem.** Fix an admissible set  $\mathcal{H}^* = \{h_1, h_2, \dots\}$  of integers and a partition  $\mathcal{H}^* = B_1 \cup B_2 \cup \dots$  where each bin  $B_i$  is a fixed, finite size  $k_i$ . Is it the case that for every  $M \geq 1$  there exists elements  $h_{a_1}, \dots, h_{a_M}$  and infinitely many integers  $n$  such that  $h_{a_j} \in B_j$  and  $n + h_{a_j} \in \mathcal{A}$  for  $1 \leq j \leq M$ ?

We realise the above set-up as the output of a sieving process. For notational purposes we order  $B_i = \{h_{k_0+\dots+k_{i-1}+1}, \dots, h_{k_0+\dots+k_i}\}$  for  $i \geq 1$ , with the convention that  $k_0 = 0$ . Let  $k = k_0 + \dots + k_M$  for some large  $M$ . Given  $n \in [N, 2N)$  for some large  $N$ , let  $\mathbf{X}_i$  denote the random variable that counts the number of  $h \in B_i$  such that  $n + h \in \mathcal{A}$ , and let  $\mathbf{X} = \mathbf{X}_1 + \dots + \mathbf{X}_M$ .

The current method we use to detect primes in  $k$ -tuples is the GPY method. For general sets  $\mathcal{A}$ , the aim is to show the first moment inequality

$$S_{\mathcal{A}} = \sum_{N \leq n < 2N} \left( \sum_{i=1}^k \mathbb{1}_{\mathcal{A}}(n + h_i) - m \right) w(n) > 0 \tag{2.1}$$

holds for some integer  $m \geq 1$ , where  $\mathbb{1}_{\mathcal{A}}$  denotes the indicator function of the set  $\mathcal{A}$  and  $w(n)$  are non-negative weights (cf. [4,9,11]). If we normalise the weights to sum to 1, then this is saying “if we choose  $n$  randomly from the interval  $[N, 2N)$  with probability  $w(n)$ , then  $\mathbb{E}[\mathbf{X}] > m$ .” From this we can deduce the existence of an  $n \in [N, 2N)$  for which  $m + 1$  of the translates  $n + h_i \in \mathcal{A}$ . We say such a translate has been “accepted.” Exactly which translates are accepted is unknown. This is a limitation of the first moment method.

It is clear that for our model problem, we require more information about which translates  $n + h$  appear. Namely, we need to be obtaining an accepted translate from each of the  $M$  bins  $B_1, \dots, B_M$  (recall  $k = k_0 + \dots + k_M$ ). This presents two obvious difficulties.

- (1) For any  $1 \leq i \leq k$  the probability of the event  $n + h_i \in \mathcal{A}$  depends on  $k$ , and tends to 0 as  $k \rightarrow \infty$ . This would mean any bin of *fixed* size expects to get fewer and fewer elements as  $k$  gets large. In particular, we cannot hope the hypotheses hold for every  $M \geq 1$ .
- (2) Even in the situation where  $\mathbb{E}[\mathbf{X}_i] > 1$  holds for each  $1 \leq i \leq M$ , we cannot conclude anything about  $\mathbb{P}(\mathbf{X}_1 > 0) \cap \dots \cap (\mathbf{X}_M > 0)$  unless we input some information about the joint distribution of the bins.

We are able to overcome these issues by modifying the sieve weights and using a second moment estimate.

### 2.2 Choice of sieve weights

We solve the first problem by modifying the sieve weights to put more emphasis on the earlier bins. This way, we can guarantee that  $\mathbb{P}(n + h \in \mathcal{A} | h \in B_i) = c_i$  where the constant  $c_i$  depends solely on the bin. This also means that we can guarantee  $\mathbb{E}[\mathbf{X}_i]$  is large for each  $i$  (provided  $k_i$  is large enough in terms of  $c_i$ ). We will consider Maynard–Tao sieve weights with a fixed factorisation

$$w(n) = \left( \sum_{\substack{d_1, \dots, d_k \\ d_i | n+h_i}} \prod_{i=1}^M \lambda_{d_{k_0+\dots+k_{i-1}+1}, \dots, d_{k_0+\dots+k_i}}^{(i)} \right)^2, \tag{2.2}$$

where

$$\lambda_{d_{k_0+\dots+k_{i-1}+1}, \dots, d_{k_0+\dots+k_i}}^{(i)} \approx \left( \prod_{j=k_0+\dots+k_{i-1}+1}^{k_0+\dots+k_i} \mu(d_j) \right) f_i(d_{k_0+\dots+k_{i-1}+1}, \dots, d_{k_0+\dots+k_i}), \tag{2.3}$$

and  $f_i$  is a suitable smooth function supported on the simplex

$$R_{B_i, \beta_i} = \{(x_{k_0+\dots+k_{i-1}+1}, \dots, x_{k_0+\dots+k_i}) \in [0, 1]^{k_i} : 0 \leq \sum_{j=k_0+\dots+k_{i-1}+1}^{k_0+\dots+k_i} x_j \leq \beta_i\}. \tag{2.4}$$

Here  $(\beta_i)_{i \geq 1}$  is a sequence of real numbers such that  $\sum_{i=1}^\infty \beta_i \leq 1$  (cf. the sieve weights defined in [9, Proposition 4.1]). We will take  $\beta_i = 2^{-i}$ , and in this instance one might say “we have allocated 50% of the sieve power to  $B_1$ .”

### 2.3 Concentration of measure

We can deal with the second problem by showing the random variables  $\mathbf{X}_i$  exhibit “enough” independence. This is precisely what concentration of measure arguments are used for. For example, an application of the union bound and Chebychev’s inequality tells us that

$$\mathbb{P}(|\mathbf{X}_i - \mathbb{E}[\mathbf{X}_i]| < t_i \text{ for all } i) \geq 1 - \sum_{i=1}^M \frac{\mathbb{E}[\mathbf{X}_i - \mathbb{E}[\mathbf{X}_i]]^2}{t_i^2} \tag{2.5}$$

where  $t_i \geq 1$  are concentration parameters. Thus, if we can show the variances  $\mathbb{E}[\mathbf{X}_i - \mathbb{E}[\mathbf{X}_i]]^2$  are small, then we should be able to show each random variable concentrates in a (small) interval about its mean with high probability. In particular, we should be able to show that we get (at least) one accepted translate coming from each bin after the sieving process. We implement this analytically by using a second moment estimate (see Proposition 5.2).

**Remark 2.1** A similar second moment estimate was considered by Banks, Freiberg and Maynard in their paper [1]. They showed a partition result (for primes), where the bins are allowed to grow with  $k$ . A key aspect of our work is that the bin sizes are fixed. Moreover they only needed upper bounds of the correct order of magnitude for the sieve sums, whereas we require precise asymptotics.

### 2.4 Hooley’s $\rho$ function

In practice, utilising a second moment estimate requires an understanding of the two-point correlations

$$\sum_{N \leq n < 2N} \rho_{\mathcal{A}}(n+h)\rho_{\mathcal{A}}(n+h') \tag{2.6}$$

where  $\rho_{\mathcal{A}}$  is a non-negative function supported on  $\mathcal{A}$ . This means our methods are limited to cases in which estimates of the above type are known. In particular we cannot deal with the case of primes, as evaluating the above sum asymptotically with  $\mathbb{1}_{\mathbb{P}}$  or the von Mangoldt function  $\Lambda$  (say) is equivalent to the twin prime conjecture.<sup>1</sup>

One can do much better when working with sums of two squares. We cannot evaluate (2.6) asymptotically using the indicator function, but we can if we work with the representation function  $r_2(n)$  instead. Unfortunately  $r_2(n)$  is too large for our purposes, and it proves necessary to consider a weighted version instead. In Hooley’s work [7] on the distribution of numbers representable as the sum of two squares, he considers a weighted representation function  $\rho(n) = t(n)r_2(n)$  where

$$t(n) = t_{N,\theta_1}(n) = \sum_{\substack{a|n, a \leq v \\ p|a \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)}{g_2(a)} \left(1 - \frac{\log a}{\log v}\right), \quad (v = N^{\theta_1}) \tag{2.7}$$

and  $\theta_1$  is a suitably small, fixed constant (for example Hooley takes  $\theta_1 = 1/20$ ). Here  $g_2(p)$  is the multiplicative function defined on primes by

$$g_2(p) = \begin{cases} 2 - \frac{1}{p} & \text{if } p \equiv 1 \pmod{4}, \\ \frac{1}{p} & \text{if } p \equiv 3 \pmod{4}. \end{cases} \tag{2.8}$$

The  $t(n)$  factor acts to dampen down the oscillations due to  $r_2(n)$ . Thus  $\rho(n)$  acts a proxy for the indicator function  $\mathbb{1}_{n=\square+\square}$  and moreover asymptotics for (2.6) are available for  $\rho(n)$ . This is the function we will be working with.

### 2.5 Outline of the paper

In Sect. 3 we deduce the results about quantum limits contained in Theorem’s 1.5 and 1.6 from Theorem 1.3. In Sect. 5 we state a few preliminary lemmas that will be needed in the sieve calculations. We defer the proofs of these results to the Appendix.

<sup>1</sup> This is precisely why the authors were limited to using upper bounds and not asymptotics in [1] (cf. Remark 2.1).

In Sect. 6 we state our main sieve results, and from them we deduce Theorem 1.4. We isolate a key lemma (see Lemma 6.6) from which all of our sieve estimates follow. Sections 7 and 8 are dedicated to proving this lemma.

### 3 Proofs of quantum limit results

In this section we deduce the results of Theorem 1.5 and Theorem 1.6 from Theorem 1.3. Following [8], we note that  $\varphi_k$  is an eigenfunction of the Laplacian on  $\mathbb{T}^d$  with eigenvalue  $\lambda_k = 4\pi^2 n_k$  for some  $n_k \in \mathbb{N}$  if and only if its Fourier expansion is of the form

$$\varphi_k(x) = \sum_{\substack{\xi \in \mathbb{Z}^d \\ |\xi|^2 = n_k}} a_\xi e^{2\pi i \langle \xi, x \rangle}, \tag{3.1}$$

for  $a_\xi \in \mathbb{C}$ . Moreover  $\|\varphi_k\|_2 = 1$  if and only if  $\sum_\xi |a_\xi|^2 = 1$ . It follows that

$$\begin{aligned} |\varphi_k(x)|^2 &= \sum_{\tau \in \mathbb{Z}^d} b_\tau(k) e^{2\pi i \langle \tau, x \rangle}, \\ b_\tau(k) &= \sum_{\substack{\xi - \eta = \tau \\ |\xi|^2 = |\eta|^2 = n_k}} a_\xi \overline{a_\eta}. \end{aligned} \tag{3.2}$$

Let  $d\nu$  be a quantum limit on  $\mathbb{T}^d$  with Fourier expansion as in (1.1). By  $|\varphi_k|^2 dx \rightarrow d\nu$  weak- $*$  as  $k \rightarrow \infty$  we mean that for every  $\tau \in \mathbb{Z}^d$  we have  $c_\tau = \lim_{k \rightarrow \infty} b_\tau(k)$ .

Fix  $a_1 < a_2 < \dots$  and  $M_1 < M_2 < \dots$  as in the statement of Theorem 1.3, and let  $b_j^{(k)}, c_j^{(k)} \in \mathbb{Z}$  be such

$$M_k = a_j^2 + (b_j^{(k)})^2 + (c_j^{(k)})^2, \quad (1 \leq j \leq k). \tag{3.3}$$

Let  $0 < \epsilon < 2$  and let  $F = F_\epsilon : \mathbb{N} \rightarrow \mathbb{N}$  be a rapidly increasing function whose rate of growth will be specified later. As we are assuming both  $a_i \rightarrow \infty$  and  $r(a_i) \rightarrow \infty$ , by passing to a subsequence if necessary (and relabelling the indices of the sequence  $M_k$ ), we may suppose  $a_i, r(a_i) \gg F(i)$ .

We will require information about the number of integer points on the surface of the  $d$ -dimensional sphere. For this we recall the following results: writing  $n = 2^k m$  and letting  $\sigma(n) = \sum_{d|n} d$  denote the sum-of-divisors function, we have the identities

$$\begin{aligned} r_3(n^2) &= 6 \prod_{p^a || m} (\sigma(p^a) - (-1)^{\frac{p-1}{2}} \sigma(p^{a-1})), \\ r_4(n^2) &= 24\sigma(m^2), \\ r_d(n^2) &= C_d(n^2)n^{d-2} \quad \text{for } d \geq 5. \end{aligned}$$

Here  $C_d(n^2)$  is a singular series which satisfies  $C_d(n^2) \asymp_d 1$ .

We prove each statement similarly - in each case we consider a suitable sequence of  $L^2$ -normalised eigenfunctions with eigenvalues  $\lambda_k = 4\pi^2 M_k$  and show that the limit has the desired property.

(Proof of (Theorem 1.3  $\Rightarrow$  Theorem 1.5)) Consider the sequence of  $L^2$ -normalised eigenfunctions on  $\mathbb{T}^3$  that arise by choosing coefficients

$$a_\xi = \begin{cases} \sqrt{\frac{2^k}{2^k-1}} \cdot \frac{1}{2^{(j+1)/2}} & \text{if } \xi = (\pm a_j, b_j^{(k)}, c_j^{(k)}) \text{ for some } j, \\ 0 & \text{otherwise.} \end{cases}$$

Fix  $i \geq 1$ . With this choice, for any  $k \geq i$  we obtain

$$b_{(2a_i, 0, 0)}(k) = \sum_{\substack{\xi - \eta = (2a_i, 0, 0) \\ |\xi|^2 = |\eta|^2 = M_k}} a_\xi \overline{a_\eta} = \frac{2^k}{2^k - 1} \cdot \frac{1}{2^{i+1}},$$

because the  $a_i$  are distinct and so the only contribution to the sum comes from  $\xi = (a_i, b_i^{(k)}, c_i^{(k)})$  and  $\eta = (-a_i, b_i^{(k)}, c_i^{(k)})$ . Hence

$$c_{(2a_i, 0, 0)} = \lim_{k \rightarrow \infty} b_{(2a_i, 0, 0)}(k) = \frac{1}{2^{i+1}} > 0,$$

which proves the theorem. □

(Proof of (Theorem 1.3  $\Rightarrow$  Theorem 1.6)) It suffices to prove part (i) for  $d = 4$ , by identifying the eigenfunctions on  $\mathbb{T}^d$  with the eigenfunctions on  $\mathbb{T}^{d+l}$  all of whose non-zero frequencies lie in the subspace  $\{(x_1, \dots, x_{d+l}) : x_{d+1} = \dots = x_{d+l} = 0\} \subseteq \mathbb{Z}^{d+l}$ .

Consider the sequence of  $L^2$ -normalised eigenvectors on  $\mathbb{T}^4$  that arise by choosing

$$a_\xi = \begin{cases} \sqrt{\frac{2^k}{2^k-1}} \cdot \frac{1}{(2^j r(a_j^2))^{1/2}} & \text{if } \xi = (X, Y, b_j^{(k)}, c_j^{(k)}) \text{ for some } j \text{ and } X^2 + Y^2 = a_j^2, \\ 0 & \text{otherwise.} \end{cases}$$

Fix  $i$  and suppose  $k \geq i$ . Given two non-zero coefficients  $a_\xi, a_{\xi'}$ , corresponding to vectors of the form  $\xi = (X, Y, b_i^{(k)}, c_i^{(k)})$  and  $\xi' = (X', Y', b_i^{(k)}, c_i^{(k)})$ , we see the difference vector is

$$\xi - \xi' = (X - X', Y - Y', 0, 0),$$

and the norm of this vector is  $\leq 2a_i$  by the triangle inequality. From (3.2), it follows that if we sum  $b_\tau(k)$  over all  $|\tau| \leq 2a_i$  then we pick up all such differences. There are  $r(a_i^2)^2$  of them, leading to

$$\sum_{\substack{\tau \in \mathbb{Z}^d \\ |\tau| \leq 2a_i}} |b_\tau(k)|^{2-\epsilon} \geq \left(\frac{2^k}{2^k - 1}\right)^{2-\epsilon} \cdot \frac{(2^i r(a_i^2))^\epsilon}{4^i}.$$



Taking the limit as  $k \rightarrow \infty$  we conclude that

$$\sum_{\substack{\tau \in \mathbb{Z}^d \\ |\tau| \leq 2a_i}} |c_\tau|^{2-\epsilon} \geq \frac{(2^i r(a_i^2))^\epsilon}{4^i} \geq \frac{(2^i r(a_i))^\epsilon}{4^i} \gg \frac{(2^i F(i))^\epsilon}{4^i}.$$

Now we can choose  $F = F_\epsilon$  so that the expression on the right hand side is unbounded as  $i \rightarrow \infty$ . It follows that  $\sum_\tau |c_\tau|^{2-\epsilon}$  doesn't converge, proving part (i).

For part (ii), fix  $d \geq 5$ . We proceed as in part (i), except this time because  $d \geq 5$  we have the lower bound  $r_{d-2}(a_i^2) \gg_d a_i^{d-4}$ . We remark that to obtain this bound for  $d \in \{5, 6\}$  we are using property (2) given by Theorem 1.3. For  $d \geq 7$  the bound holds without this extra assumption on our sequence.

Now consider the sequence of eigenvectors on  $\mathbb{T}^d$  with densities

$$a_\xi = \begin{cases} \sqrt{\frac{2^k}{2^k-1}} \cdot \frac{1}{(2^j r_{d-2}(a_j^2))^{1/2}} & \text{if } \xi = (X_1, \dots, X_{d-2}, b_j^{(k)}, c_j^{(k)}) \text{ for some } j \text{ and } X_1^2 + \dots + X_{d-2}^2 = a_j^2, \\ 0 & \text{otherwise.} \end{cases}$$

Fix  $i$  and suppose  $k \geq i$ . As above we can conclude

$$\sum_{\substack{\tau \in \mathbb{Z}^d \\ |\tau| \leq 2a_i}} |b_\tau(k)| \geq \left( \frac{2^k}{2^k-1} \right) \cdot \frac{r_{d-2}(a_i^2)}{2^i}.$$

Taking the limit as  $k \rightarrow \infty$  we conclude that

$$\sum_{\substack{\tau \in \mathbb{Z}^d \\ |\tau| \leq 2a_i}} |c_\tau| \geq \frac{r_{d-2}(a_i^2)}{2^i} \gg_d \frac{a_i^{d-4}}{2^i}.$$

It follows that

$$\frac{\sum_{|\tau| \leq 2a_i} |c_\tau|}{(2a_i)^{d-4-\epsilon}} \gg_d \frac{(2a_i)^\epsilon}{2^i} \gg_d \frac{(2F(i))^\epsilon}{2^i}.$$

Choosing  $F = F_\epsilon$  appropriately and letting  $i \rightarrow \infty$ , we see that for this choice of quantum limit we have

$$\limsup_{\rho \rightarrow \infty} \frac{\Sigma(\rho)}{\rho^{d-4-\epsilon}} = +\infty,$$

where  $\Sigma(\rho)$  is defined as in (1.2). This proves part (ii). □

### 4 Notation

We will use both Landau and Vinogradov asymptotic notation throughout the paper.  $N$  will denote a large integer, and all asymptotic notation is to be understood as referring to the limit as  $N \rightarrow \infty$ . Any dependencies of the implied constants on other parameters

$A$  will be denoted by a subscript, for example  $X \ll_A Y$  or  $X = O_A(Y)$ , unless stated otherwise. We let  $\epsilon$  denote a small positive constant, and we adopt the convention it is allowed to change at each occurrence, and even within a line.

We will denote the non-trivial Dirichlet character (mod 4) by  $\chi_4$ , and we may omit the subscript and simply write  $\chi$ . As usual, we let  $\varphi(n)$  denote the Euler-Totient function,  $\tau_r(n)$  denote the number of ways of writing  $n$  as the product of  $r$  natural numbers,  $\mu(n)$  denote the Möbius function, and  $r_d(n)$  denote the number of representations of  $n$  as the sum of  $d$  squares. For the rest of the paper we will write  $r(n)$  when  $d = 2$ . For integers  $a, b$  we let  $(a, b)$  denote their highest common factor, and  $[a, b]$  denote their lowest common multiple.

We define the Ramanujan–Landau constant

$$A = \frac{1}{\sqrt{2}} \prod_{p \equiv 3 \pmod{4}} \left(1 - \frac{1}{p^2}\right)^{-\frac{1}{2}} \quad (4.1)$$

which will appear in many of our results.

## 5 Preliminaries

In this section, we formalise some of the notions discussed in Sect. 2, and state a few key estimates that will be required later in the sieve calculations.

### 5.1 A pigeonhole argument

The following proposition allows us to go from the set-up in Theorem 1.4 to the model problem discussed in Sect. 2.

**Proposition 5.1** (Pigeonhole argument for infinite bin set-up). *Fix  $\mathcal{A} \subseteq \mathbb{N}$  and a set  $\mathcal{H}^* = \{h_1, h_2, \dots\}$  of integers. Suppose that there exists a partition  $\mathcal{H}^* = B_1 \cup B_2 \cup \dots$  where each bin  $B_i$  is a fixed, finite size, such that for every  $M \geq 1$ , there exists infinitely many  $n$  and  $M$  translates  $n + h_{i,M} \in \mathcal{A}$  with  $h_{i,M} \in B_i$  for  $1 \leq i \leq M$ . Then there exists increasing sequences  $a_j$  and  $n_k$  such that for every  $k \geq 1$  we have  $n_k + h_{a_j} \in \mathcal{A}$  for  $1 \leq j \leq k$  and moreover  $h_{a_j} \in B_j$  for all  $j$ .*

**Proof** With the above set-up, obtain translates  $n + h_{i,M} \in \mathcal{A}$  with  $h_{i,M} \in B_i$  for  $1 \leq i \leq M$  for each  $M \geq 1$ . Record this process in the following infinite table:

$B_1$	$B_2$	$B_3$	...	$B_M$	$B_{M+1}$	...
$h_{1,1}$			...			
$h_{1,2}$	$h_{2,2}$		...			
$h_{1,3}$	$h_{2,3}$	$h_{3,3}$	...			
$\vdots$	$\vdots$	$\vdots$	$\vdots$			
$h_{1,M}$	$h_{2,M}$	$h_{3,M}$	...	$h_{M,M}$		
$h_{1,M+1}$	$h_{2,M+1}$	$h_{3,M+1}$	...	$h_{M,M+1}$	$h_{M+1,M+1}$	
$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	$\vdots$	$\vdots$

Look at the first column. By the pigeonhole principle, since  $B_1$  is finite, there must exist an element  $h_{a_1} \in B_1$  which appears infinitely many times. Choose the smallest such  $h_{a_1}$ , and choose any  $n_1 \in \mathbb{N}$  for which  $n_1 + h_{a_1} \in \mathcal{A}$ . Now erase all the rows that do not start with  $h_{a_1}$ , and look at the remaining (infinite) table. Again, since  $B_2$  is finite, some element  $h_{a_2} \in B_2$  must occur infinitely many times in the second column. Choose the smallest such  $h_{a_2}$ , and choose any  $n_2 > n_1$  such that  $n_2 + h_{a_2} \in \mathcal{A}$  (which we can do because there are infinitely many such  $n_2$ ). By construction this  $n_2$  will be such that  $n_2 + h_{a_1} \in \mathcal{A}$ . Now erase all rows that don't start with  $h_{a_1}, h_{a_2}$ , and repeat this process for  $B_3$ , and so on. We will end up with increasing sequences  $a_j$  and  $n_k$  which by construction satisfy the required conditions.  $\square$

### 5.2 A second moment estimate

As discussed in Sect. 2, our work will require input about the joint distribution of the bins, which we will achieve via concentration of measure arguments. The following second moment estimate will suffice for our purposes.

**Proposition 5.2** (Second moment estimate). *Fix  $\mathcal{A} \subset \mathbb{N}$  and  $\mathcal{H} = \{h_1, \dots, h_k\}$  a set of integers. Suppose we have a partition  $\mathcal{H} = B_1 \cup \dots \cup B_M$ . Let  $\mu_i, t_i \geq 1$  be real numbers for  $1 \leq i \leq M$ . Let  $\rho_{\mathcal{A}}$  be a non-negative function supported on  $\mathcal{A}$  and  $w(n)$  be non-negative weights for each integer  $n$ . If*

$$\sum_{N \leq n < 2N} \left[ \min_{j=1, \dots, M} \frac{\mu_j^2}{t_j^2} - \sum_{i=1}^M \left( \frac{\sum_{h \in B_i} \rho_{\mathcal{A}}(n+h) - \mu_i}{t_i} \right)^2 \right] w(n) > 0, \tag{5.1}$$

then there exists an  $n \in [N, 2N)$  and elements  $h_{a_i} \in B_i$  such that  $n + h_{a_i} \in \mathcal{A}$  for  $1 \leq i \leq M$ .

**Proof** By positivity we deduce the existence of an  $n \in [N, 2N)$  such that

$$\sum_{i=1}^M \left( \frac{\sum_{h \in B_i} \rho_{\mathcal{A}}(n+h) - \mu_i}{t_i} \right)^2 < \min_{j=1, \dots, M} \frac{\mu_j^2}{t_j^2}.$$

If  $n + h \notin \mathcal{A}$  for all  $h \in B_i$ , then by assumption on the support of  $\rho_{\mathcal{A}}$  the left hand side of the above expression is  $\geq \mu_i^2/t_i^2$ , a contradiction.  $\square$

Thus, if the second moment estimate (5.1) holds for all  $M \geq 1$  and sufficiently large  $N$ , then we are in a situation where the hypotheses of Proposition 5.1 are satisfied.

### 5.3 Estimates in arithmetic progressions

We require an understanding of how  $\rho(n)$  and  $\rho(n)\rho(n + h)$  behaves in arithmetic progressions for our sieve calculations. Essentially this reduces down to understanding the corresponding sums for  $r(n)$  and  $r(n)r(n + h)$ , where the estimates we need are known with power-saving error terms. This means that the error terms in the sieve calculations can be bounded trivially (cf. with the case of primes [4,9], where we have to use equi-distribution results such as the Bombieri-Vinogradov theorem to bound the error terms that arise).

We have the following lemmas. We note that the functions  $g_1, \dots, g_7$  defined in this section will be used frequently throughout the rest of the paper.

**Lemma 5.3** *Suppose  $(a, q) = (d, q) = 1$  where  $d, q$  are square-free and odd, of size  $\ll N^{O(1)}$ . Then we have*

$$\sum_{\substack{n \leq N \\ n \equiv a \pmod{q} \\ n \equiv 1 \pmod{4} \\ d|n}} r(n) = \frac{g_1(q)g_2(d)}{2qd} \pi N + R_1(N; d, q),$$

where  $g_2$  is defined as in (2.8),  $g_1$  is the multiplicative function defined on primes by  $g_1(p) = 1 - \chi(p)/p$ , and

$$R_1(N; d, q) \ll_{\epsilon} ((qd)^{\frac{1}{2}} + N^{\frac{1}{3}})d^{\frac{1}{2}}N^{\epsilon}$$

**Lemma 5.4** *Suppose that  $(a, q) = (a + h, q) = (d_1, q) = (d_2, q) = (d_1, d_2) = 1$  and  $4|h$ , where  $d_1, d_2, q$  are square-free and odd, of size  $\ll N^{O(1)}$ . Moreover suppose  $h > 0$  is fixed such that  $p|h \Rightarrow p|2q$ . Then we have*

$$\sum_{\substack{n \leq N \\ n \equiv a \pmod{q} \\ n \equiv 1 \pmod{4} \\ d_1|n \\ d_2|n+h}} r(n)r(n + h) = \frac{g_1(q)^2 \Gamma(d_1, d_2, q)}{q} \pi^2 N + R_2(N; d_1, d_2, q),$$

where

$$\Gamma(d_1, d_2, q) = \frac{g_2(d_1)g_2(d_2)}{d_1d_2} \sum_{(r, 2q)=1} \frac{\mu(r)(d_1, r)(d_2, r)\chi[(d_1^2, r)]\chi[(d_2^2, r)]}{r^2},$$

and

$$R_2(N; d_1, d_2, q) \ll_{\epsilon} q^{\frac{1}{2}}d_1d_2N^{\frac{3}{4}+\epsilon} + d_1^{\frac{1}{2}}d_2^{\frac{1}{2}}N^{\frac{5}{6}+\epsilon}.$$

**Lemma 5.5** *Let  $(a, q) = (d, q) = 1$  where  $d, q$  are square-free and odd, of size  $\ll N^{O(1)}$ . Then we have*

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

where

$$A_2 = 2\gamma - 1 + 2 \frac{L'(1, \chi_4)}{L(1, \chi_4)} - 2 \frac{\zeta'(2)}{\zeta(2)} + \frac{4}{3} \log 2.$$

Here  $g_3, g_4$  are the multiplicative functions defined on primes by

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

and  $g_5(p), g_6(p)$  are defined by

$$g_5(p) = \begin{cases} \frac{(2p+1) \log p}{p^2-1} & \text{if } p \equiv 1 \pmod{4}, \\ \frac{\log p}{p^2-1} & \text{if } p \equiv 3 \pmod{4}, \end{cases}$$

$$g_6(p) = \begin{cases} \frac{(p-1)^2(2p+1) \log p}{(p+1)(4p^2-3p+1)} & \text{if } p \equiv 1 \pmod{4}, \\ \log p & \text{if } p \equiv 3 \pmod{4}. \end{cases}$$

We will prove each of these results in Appendix A. Lemma 5.3 follows from two known results. Lemma 5.4 follows by adapting the method used by Plaksin in [10], where a similar sum is considered. Finally Lemma 5.5 can be shown using standard Perron’s formula arguments, together with a fourth moment estimate for Dirichlet L-functions.

When finding the corresponding estimates for  $\rho(n)$  the following sums naturally appear (for the definitions of  $W, W_1$  and  $D_0$  see (6.1) below the fold):

$$X_{N,W} = \sum_{\substack{a \leq v \\ (a,W)=1 \\ p|a \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)}{a} \log \frac{v}{a}, \quad (5.2)$$

$$Y_{N,W} = \sum_{\substack{a,b \leq v \\ (a,W)=(b,W)=1 \\ (a,b)=1 \\ p|a,b \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)\mu(b)}{g_7(a)g_7(b)} \log \frac{v}{a} \log \frac{v}{b}, \quad (5.3)$$

$$Z_{N,W}^{(1)} = \sum_{\substack{a,b \leq v \\ (a,W)=(b,W)=1 \\ p|a,b \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)\mu(b)g_4([a,b])}{g_2(a)g_2(b)[a,b]} \log \frac{v}{a} \log \frac{v}{b}, \quad (5.4)$$

$$Z_{N,W}^{(2)} = \sum_{\substack{a,b \leq v \\ (a,W)=(b,W)=1 \\ p|a,b \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)\mu(b)g_4([a,b])}{g_2(a)g_2(b)[a,b]} \log \frac{v}{a} \log \frac{v}{b} \sum_{p|[a,b]} g_6(p). \quad (5.5)$$

Here  $g_7$  is the multiplicative function defined on primes by  $g_7(p) = p + 1$ . The following lemma evaluates the auxiliary sums above.

**Lemma 5.6** (Auxiliary estimates for  $\rho(n)$ ). *We have*

$$\begin{aligned} X_{N,W} &= (1 + o(1)) \frac{8A \log^{\frac{1}{2}} v}{\pi g_1(W_1)}, \\ Y_{N,W} &= (1 + o(1)) \frac{64A^2 \log v}{\pi^2 g_1(W_1)^2}, \\ Z_{N,W}^{(1)} &= (1 + o(1)) \frac{32A^3 \log^{\frac{1}{2}} v}{\pi^2 g_1(W_1)^3}, \\ Z_{N,W}^{(2)} &= -(1 + o(1)) \frac{16A^3 \log^{\frac{3}{2}} v}{\pi^2 g_1(W_1)^3}, \end{aligned}$$

where  $A$  is defined as in (4.1). In each case one may take the  $o(1)$  term to be  $O(D_0^{-1})$ .

We remark that the estimate for  $X_{N,1}$  appears in Hooley's work (after correcting a misprint—cf. [7, Lemma 5] and note his slightly different definition of  $A$ ). We prove Lemma 5.6 in Appendix B. Each sum can be evaluated by the Selberg–Delange method.

## 6 The sieve set-up

We now state our sieve results and use them to deduce Theorem 1.4. For the rest of the paper  $k$  is fixed,  $\mathcal{H} = \{h_1, \dots, h_k\}$  is a fixed admissible set such that  $4|h_i$  for each  $i$ , and  $N$  is sufficiently large in terms of any fixed quantity. We allow any of the constants hidden in the Landau notation to depend on  $k$ , without explicitly specifying so.

We will employ a  $4W$ -trick in our sieve calculations. Let

$$W = \prod_{2 < p \leq D_0} p, \tag{6.1}$$

where  $D_0 = (\log \log N)^3$ , so that  $W \ll (\log N)^{2(\log \log N)^2} \ll_\epsilon N^\epsilon$  for any fixed  $\epsilon > 0$  by the prime number theorem. It will prove useful to define

$$W_1 = \prod_{\substack{p \leq D_0 \\ p \equiv 1 \pmod{4}}} p, \quad W_3 = \prod_{\substack{p \leq D_0 \\ p \equiv 3 \pmod{4}}} p, \tag{6.2}$$

so that  $W = W_1 W_3$ .

By admissibility of  $\mathcal{H}$  there exists a fixed residue class  $v_0 \pmod{W}$  such that  $(v_0 + h_i, W) = 1$  for each  $i$ . Fix  $1 \leq m, l \leq k$  with  $m \neq l$ . We consider four types of sums:

$$S_1 = \sum_{\substack{N \leq n < 2N \\ n \equiv v_0 \pmod{W} \\ n \equiv 1 \pmod{4}}} \left( \sum_{\substack{d_1, \dots, d_k \\ d_i | n + h_i \ \forall i}} \lambda_{d_1, \dots, d_k} \right)^2, \tag{6.3}$$

$$S_2^{(m)} = \sum_{\substack{N \leq n < 2N \\ n \equiv v_0 \pmod{W} \\ n \equiv 1 \pmod{4}}} \rho(n + h_m) \left( \sum_{\substack{d_1, \dots, d_k \\ d_i | n + h_i \ \forall i}} \lambda_{d_1, \dots, d_k} \right)^2, \tag{6.4}$$

$$S_3^{(m,l)} = \sum_{\substack{N \leq n < 2N \\ n \equiv v_0 \pmod{W} \\ n \equiv 1 \pmod{4}}} \rho(n + h_m) \rho(n + h_l) \left( \sum_{\substack{d_1, \dots, d_k \\ d_i | n + h_i \ \forall i}} \lambda_{d_1, \dots, d_k} \right)^2, \tag{6.5}$$

$$S_4^{(m)} = \sum_{\substack{N \leq n < 2N \\ n \equiv v_0 \pmod{W} \\ n \equiv 1 \pmod{4}}} \rho^2(n + h_m) \left( \sum_{\substack{d_1, \dots, d_k \\ d_i | n + h_i \ \forall i}} \lambda_{d_1, \dots, d_k} \right)^2. \tag{6.6}$$

Because  $k$  is fixed, we may assume that  $D_0$  is sufficiently large so that

$$p|h_i - h_j \Rightarrow p|2W. \tag{6.7}$$

**Remark 6.1** For the second moment estimate, it proves important to control the residue classes of the translates  $n + h \pmod{4}$ , hence the condition  $n \equiv 1 \pmod{4}$  in our sieve sums and also the assumption  $4|h$  for our admissible set. This is because of the inherent bias numbers representable as a sum of two squares have modulo 4.

Our first Proposition evaluates these sums for general half-dimensional Maynard–Tao sieve weights. Fix  $0 < \theta_1 < 1/18$  in the definition of  $\rho(n)$  (see (2.7)). We also define the normalisation constant

$$B = \frac{A}{\Gamma(1/2)\sqrt{L(1, \chi_4)}} \cdot \frac{\varphi(W_3)(\log R)^{\frac{1}{2}}}{W_3} = \frac{2A}{\pi} \cdot \frac{\varphi(W_3)(\log R)^{\frac{1}{2}}}{W_3}. \tag{6.8}$$

**Proposition 6.2** (Half-dimensional Maynard–Tao sieve estimates). *Let  $R = N^{\theta_2/2}$  for some small fixed positive constant  $\theta_2$  such that  $0 < \theta_1 + \theta_2 < 1/18$ . Let  $\lambda_{d_1, \dots, d_k}$  be defined in terms of a fixed smooth function  $F$  by*

$$\lambda_{d_1, \dots, d_k} = \left( \prod_{i=1}^k \mu(d_i) d_i \right) \sum_{\substack{r_1, \dots, r_k \\ d_i | r_i \forall i \\ (r_i, W) = 1 \forall i \\ p | r_i \Rightarrow p \equiv 3 \pmod{4} \forall i}} \frac{\mu(\prod_{i=1}^k r_i)^2}{\prod_{i=1}^k \varphi(r_i)} F\left(\frac{\log r_1}{\log R}, \dots, \frac{\log r_k}{\log R}\right),$$

whenever  $\prod_{i=1}^k d_i \leq R$  is squarefree,  $(\prod_{i=1}^k d_i, W) = 1$  and  $p | \prod_{i=1}^k d_i \Rightarrow p \equiv 3 \pmod{4}$ , and let  $\lambda_{d_1, \dots, d_k} = 0$  otherwise. Moreover let  $F$  be supported on  $R_k = \{(x_1, \dots, x_k) \in [0, 1]^k : \sum_{i=1}^k x_i \leq 1\}$ . Then we have

$$\begin{aligned} S_1 &= (1 + o(1)) \frac{B^k N}{4W} L_k(F), \\ S_2^{(m)} &= (1 + o(1)) \frac{4\sqrt{\frac{\log R}{\log v}} B^k N}{\pi W} L_{k;m}(F), \\ S_3^{(m,l)} &= (1 + o(1)) \frac{64\left(\frac{\log R}{\log v}\right) B^k N}{\pi^2 W} L_{k;m,l}(F), \\ S_4^{(m)} &= (1 + o(1)) \frac{2\sqrt{\frac{\log R}{\log v}} \left(\frac{\log N}{\log v} + 1\right) B^k N}{\pi W} L_{k;m}(F) \end{aligned}$$

provided  $L_k(F)$ ,  $L_{k;m}(F)$  and  $L_{k;m,l}(F)$  are non-zero, where

$$\begin{aligned} L_k(F) &= \int_0^1 \dots \int_0^1 \left[ F(x_1, \dots, x_k) \right]^2 \prod_{i=1}^k \frac{dx_i}{\sqrt{x_i}} \\ L_{k;m}(F) &= \int_0^1 \dots \int_0^1 \left[ \int_0^1 F(x_1, \dots, x_k) \frac{dx_m}{\sqrt{x_m}} \right]^2 \prod_{\substack{i=1 \\ i \neq m}}^k \frac{dx_i}{\sqrt{x_i}}, \\ L_{k;m,l}(F) &= \int_0^1 \dots \int_0^1 \left[ \int_0^1 \left( \int_0^1 F(x_1, \dots, x_k) \frac{dx_m}{\sqrt{x_m}} \right) \frac{dx_l}{\sqrt{x_l}} \right]^2 \prod_{\substack{i=1 \\ i \neq m,l}}^k \frac{dx_i}{\sqrt{x_i}}. \end{aligned}$$



From this, one can deduce the corresponding results for the modification of the Maynard–Tao sieve described in Sect. 2.

**Proposition 6.3** (Modified Maynard–Tao sieve estimates). *Suppose in addition to the hypotheses of Proposition 6.2 we have a partition  $\mathcal{H} = \{h_1, \dots, h_k\} = B_1 \cup \dots \cup B_M$  into bins  $B_i$  of fixed and finite size  $k_i$ . Write  $B_i = \{h_{k_0+\dots+k_{i-1}+1}, \dots, h_{k_0+\dots+k_i}\}$  with the convention that  $k_0 = 0$ . Suppose further we have a corresponding factorisation*

$$F(x_1, \dots, x_k) = \prod_{i=1}^M F_i(x_{k_0+\dots+k_{i-1}+1}, \dots, x_{k_0+\dots+k_i}),$$

where each  $F_i$  is smooth and supported on the simplex

$$R_{B_i, \beta_i} = \{(x_{k_0+\dots+k_{i-1}+1}, \dots, x_{k_0+\dots+k_i}) \in [0, 1]^{k_i} : 0 \leq \sum_{j=k_0+\dots+k_{i-1}+1}^{k_0+\dots+k_i} x_j \leq \beta_i\}.$$

Here  $(\beta_i)_{i=1}^\infty$  is a sequence of real numbers such that  $\sum_{i=1}^\infty \beta_i \leq 1$ . Then for  $h_m, h_l \in B_j$  we have

$$\begin{aligned} S_1 &= (1 + o(1)) \frac{B^k N}{4W} \left( \prod_{i=1}^M L_{|B_i|}(F) \right), \\ S_2^{(m)} &= (1 + o(1)) \frac{4\sqrt{\frac{\log R}{\log v}} B^k N}{\pi W} \left( \prod_{i=1}^M L_{|B_i|}(F_i) \right) \frac{L_{|B_j|;m}(F_j)}{L_{|B_j|}(F_j)}, \\ S_3^{(m,l)} &= (1 + o(1)) \frac{64\left(\frac{\log R}{\log v}\right) B^k N}{\pi^2 W} \left( \prod_{i=1}^M L_{|B_i|}(F_i) \right) \frac{L_{|B_j|;m,l}(F_j)}{L_{|B_j|}(F_j)}, \\ S_4^{(m)} &= (1 + o(1)) \frac{2\sqrt{\frac{\log R}{\log v}} \left(\frac{\log N}{\log v} + 1\right) B^k N}{\pi W} \left( \prod_{i=1}^M L_{|B_i|}(F_i) \right) \frac{L_{|B_j|;m}(F_j)}{L_{|B_j|}(F_j)}. \end{aligned}$$

**Proof** The hypotheses imply  $F = \prod_{i=1}^M F_i$  is also smooth and supported on  $R_k$ , and hence the results of Proposition 6.2 apply. It suffices to show the functionals factorise in the forms stated. Because our set-up ensures that  $\text{supp}(F) = \text{supp}(F_1) \times \dots \times \text{supp}(F_M)$  one can easily check that if  $h_m, h_l \in B_j$  then

$$\begin{aligned} L_k(F) &= \prod_{i=1}^M L_{|B_i|}(F_i), \\ L_{k;m}(F) &= \left( \prod_{i=1}^M L_{|B_i|}^{(0)}(F_i) \right) \frac{L_{|B_j|;m}(F_j)}{L_{|B_j|}(F_j)}, \end{aligned}$$

$$L_{k;m,l}(F) = \left( \prod_{i=1}^M L_{|B_i|}^{(0)}(F_i) \right) \frac{L_{|B_j|;m,l}(F_j)}{L_{|B_j|}(F_j)}.$$

□

With the following lemma we will be in a position to prove Theorem 1.4.

**Lemma 6.4** (Evaluation of sieve functionals). *Let  $F(t_1, \dots, t_k) = \prod_{i=1}^k g(kt_i)$  where*

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

Then for any  $m, l$  we have

$$\begin{aligned} \frac{L_{k;m}(F)}{L_k(F)} &= \frac{\pi^2}{\pi + 2} \cdot \sqrt{\frac{\beta}{k}}, \\ \frac{L_{k;m,l}(F)}{L_k(F)} &= \left( \frac{\pi^2}{\pi + 2} \right)^2 \cdot \frac{\beta}{k}. \end{aligned}$$

**Proof** The definition implies  $F$  is supported on the cube  $[0, \frac{\beta}{k}]^k \subseteq R_{k,\beta}$ . In this case the functionals factorise completely and the lemma follows from the fact

$$\begin{aligned} \int_0^{\frac{\beta}{k}} \frac{dt}{\sqrt{t}(1 + \frac{kt}{\beta})^2} &= \frac{\pi + 2}{4} \cdot \sqrt{\frac{\beta}{k}}, \\ \int_0^{\frac{\beta}{k}} \frac{dt}{\sqrt{t}(1 + \frac{kt}{\beta})} &= \frac{\pi}{2} \cdot \sqrt{\frac{\beta}{k}}. \end{aligned}$$

□

**Remark 6.5** We have restricted the support of our functions to the cube  $[0, \frac{\beta}{k}]^k \subseteq R_{k,\beta}$  so that the integrals can be evaluated exactly. This essentially means we are using weights of similar strength to the original GPY weights (cf. [4]). For the half-dimensional case one can show that for large  $k$  these weights are essentially optimal. (In particular, following a similar optimisation process as in [9, Section 7] one arrives at the same results as above.)

We are now in a position to prove Theorem 1.4.

**Proof of Theorem 1.4** Let  $\mathcal{H} = \{h_1, h_2, \dots\}$  be a fixed admissible set. Fix real numbers  $\theta_1, \theta_2$  subject to  $0 < \theta_1 + \theta_2 < 1/18$  and define the constant

$$\Delta = \Delta(\theta_1, \theta_2) = \frac{\sqrt{2}(\pi + 2)}{32\pi} \frac{1 + \theta_1}{\sqrt{\theta_1\theta_2}}.$$

With notation as above, consider a partition  $\mathcal{H} = B_1 \cup B_2 \cup \dots$  where  $k_1 > 2\Delta^3$  and for  $i \geq 2$  we choose  $B_i$  such that  $k_i > 2^{7i}$ . By Proposition 5.2 we will be done if we can show, for every  $M \geq 1$ , the inequality

$$\sum_{\substack{N \leq n < 2N \\ n \equiv v_0 \pmod{W} \\ n \equiv 1 \pmod{4}}} \left[ \min_{j=1, \dots, M} \frac{\mu_j^2}{t_j^2} - \sum_{i=1}^M \left( \frac{\sum_{h \in B_i} \rho(n+h) - \mu_i}{t_i} \right)^2 \right] \left( \sum_{\substack{d_1, \dots, d_k \\ d_i | n+h_i \forall i}} \lambda_{d_1, \dots, d_k} \right)^2 > 0 \tag{6.9}$$

for all sufficiently large  $N$  and some choice of real numbers  $\mu_i, t_i \geq 1$ . Choose weights  $\lambda_{d_1, \dots, d_k}$  as in Proposition 6.2, and let  $F(x_1, \dots, x_k) = \prod_{i=1}^M F_i(x_{k_0+\dots+k_{i-1}+1}, \dots, x_{k_0+\dots+k_i})$  where each  $F_i$  is supported on  $R_{B_i, 2^{-i}}$ . Let  $F_i(x_{k_0+\dots+k_{i-1}+1}, \dots, x_{k_0+\dots+k_i}) = \prod_{j=k_0+\dots+k_{i-1}+1}^{k_0+\dots+k_i} g(k_i x_j)$  where for  $j \in \{b_{2i-1}, \dots, b_{2i}\}$  we define

$$g(x_j) = \begin{cases} \frac{1}{1+2^i x_j}, & \text{if } x_j \leq 2^{-i} \\ 0, & \text{otherwise} \end{cases}$$

Expanding out (6.9), we have the evaluate the expression

$$\left( \min_{j=1, \dots, M} \frac{\mu_j^2}{t_j^2} \right) S_1 - \sum_{i=1}^M \frac{1}{t_i^2} \left[ \sum_{\substack{h, h' \in B_i \\ h \neq h'}} S_3^{(h, h')} + \sum_{h \in B_i} S_4^{(h)} - 2\mu_i \sum_{h \in B_i} S_2^{(h)} + \mu_i^2 S_1 \right]$$

(where by abuse of notation we have written  $S_2^{(h)}$  for  $S_2^{(i)}$  where  $h = h_i$  say). A convenient choice of  $\mu_i, \lambda_i$  is

$$\mu_i = c \left( \frac{k_i}{2^i} \right)^{\frac{1}{2}}, \quad t_i = c \left( \frac{k_i}{2^i} \right)^{\frac{1}{3}},$$

where

$$c = c(\theta_1, \theta_2) = \frac{16\sqrt{\theta_2/2\theta_1}}{\pi} \left( \frac{\pi^2}{\pi + 2} \right).$$

Evaluating these sums using Proposition 6.3 and Lemma 6.4 we see that this is asymptotically

$$\frac{B^k N}{4W} \left( \prod_{i=1}^M L_{|B_i|}^{(0)}(F) \right) \left\{ \left( \frac{k_1}{2} \right)^{\frac{1}{3}} - \sum_{i=1}^M \left[ \Delta \left( \frac{2^i}{k_i} \right)^{\frac{1}{6}} - \frac{1}{2^i} \left( \frac{2^i}{k_i} \right)^{\frac{2}{3}} \right] \right\}.$$

Hence (6.9) will be satisfied for all sufficiently large  $N$  provided

$$\Delta \sum_{i=1}^M \left(\frac{2^i}{k_i}\right)^{\frac{1}{6}} < \left(\frac{k_1}{2}\right)^{\frac{1}{3}}. \tag{6.10}$$

But now our choice of bins ensures that

$$\Delta \sum_{i=1}^M \left(\frac{2^i}{k_i}\right)^{\frac{1}{6}} \leq \Delta \sum_{i=1}^M \frac{1}{2^i} \leq \Delta < \left(\frac{k_1}{2}\right)^{\frac{1}{3}},$$

and so (6.10) is satisfied for all  $M \geq 1$ . □

It remains to prove Proposition 6.2. Each sum can be treated similarly. The following lemma handles all of them at once. First, given a function  $F$  satisfying the hypotheses of Proposition 6.2, we define

$$F_{\max} = \sup_{(t_1, \dots, t_k) \in [0, 1]^k} \left( |F(t_1, \dots, t_k)| + \sum_{i=1}^k \left| \frac{\partial F}{\partial t_i}(t_1, \dots, t_k) \right| \right). \tag{6.11}$$

The lemma can now be stated as follows.

**Lemma 6.6** (General sieve lemma). *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*

$$S_{J, p_1, p_2, m} = \sum_{\substack{d_1, \dots, d_k \\ e_1, \dots, e_k \\ W, [d_1, e_1], \dots, [d_k, e_k] \text{ coprime} \\ p_1 | d_m, p_2 | e_m}} \lambda_{d_1, \dots, d_k} \lambda_{e_1, \dots, e_k} \prod_{i \in I} f([d_i, e_i]) \prod_{j \in J} g([d_j, e_j]),$$

with weights  $\lambda_{d_1, \dots, d_k}$  defined as in Proposition 6.2. If  $J = \emptyset$  we define  $f(p) = 1/p$  (and there is no dependence on  $g$  in the sum). Otherwise,  $f$  and  $g$  are non-zero 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  $S_{J, 1, 1, m}$ . Suppose  $\lambda_{d_1, \dots, d_k}$  satisfy the same hypotheses as in Proposition 6.2. 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 the integral operators are defined by Proposition 6.2 above, and we write  $L_J(F)$  as shorthand for  $L_{k; j \in J}(F)$ .

We now show how this implies Proposition 6.2.

*Proof of (Lemma 6.6  $\Rightarrow$  Proposition 6.2)* We consider each sum in turn. First we note that using the definition of  $\lambda_{d_1, \dots, d_k}$ , the exact same calculation as in [9, p. 394] gives

$$\sup_{d_1, \dots, d_k} |\lambda_{d_1, \dots, d_k}| \ll F_{\max} \sum_{\substack{u \leq R \\ p|u \Rightarrow p \equiv 3 \pmod{4}}} \frac{\mu^2(u) \tau_k(u)}{\varphi(u)} \ll F_{\max} (\log R)^{\frac{k}{2}},$$

and so we have a trivial bound<sup>2</sup>

$$\begin{aligned} \sum_{\substack{d_1, \dots, d_k \\ e_1, \dots, e_k}} |\lambda_{d_1, \dots, d_k} \lambda_{e_1, \dots, e_k}| &\ll \left( \sup_{d_1, \dots, d_k} |\lambda_{d_1, \dots, d_k}| \right)^2 \left( \sum_{\substack{d \leq R \\ p|d \Rightarrow p \equiv 3 \pmod{4}}} \tau_k(d) \right)^2 \\ &\ll F_{\max}^2 R^2 (\log R)^{2k} \ll_{\epsilon} F_{\max}^2 N^{\theta_2 + \epsilon}. \end{aligned} \tag{6.12}$$

As mentioned in Sect. 4, because we can obtain power-saving in the error terms for the formulae stated there, this trivial bound will suffice for our purposes.

(i) Rewrite  $S_1$  in the form

$$S_1 = \sum_{\substack{d_1, \dots, d_k \\ e_1, \dots, e_k}} \lambda_{d_1, \dots, d_k} \lambda_{e_1, \dots, e_k} \sum_{\substack{N \leq n < 2N \\ n \equiv v_0 \pmod{W} \\ n \equiv 1 \pmod{4} \\ n \equiv -h_i \pmod{[d_i, e_i]}}} 1.$$

We may assume  $W, [d_1, e_1], \dots, [d_k, e_k]$  are pairwise coprime, as otherwise the inner sum is empty. In this case, by the Chinese Remainder Theorem, these congruences are equivalent to a single congruence  $(\text{mod } q)$  where  $q = 4W \prod_{i=1}^k [d_i, e_i]$ . The inner sum evaluates to

$$\frac{N}{q} + O(1).$$

<sup>2</sup> By Rankin's trick we have

$$\sum_{\substack{d \leq R \\ p|d \Rightarrow p \equiv 3 \pmod{4}}} \tau_k(d) \leq R \sum_{\substack{d \leq R \\ p|d \Rightarrow p \equiv 3 \pmod{4}}} \frac{\tau_k(d)}{d} \ll R (\log R)^{k/2}.$$

The error term contributes  $O_\epsilon(F_{\max}^2 N^{\theta_2+\epsilon})$  which is negligible. The main term is

$$\frac{N}{4W} \sum_{\substack{d_1, \dots, d_k \\ e_1, \dots, e_k \\ W, [d_1, e_1], \dots, [d_k, e_k] \text{ coprime}}} \lambda_{d_1, \dots, d_k} \lambda_{e_1, \dots, e_k} \prod_{i=1}^k \frac{1}{[d_i, e_i]}.$$

This is of the form  $S_J$  where  $|J| = 0$ . Evaluating it according to Lemma 6.6 we obtain

$$S_1 = (1 + o(1)) \frac{B^k N}{4W} L_k^{(0)}(F).$$

(ii) Rewrite  $S_2^{(m)}$  in the form

$$S_2^{(m)} = \sum_{\substack{d_1, \dots, d_k \\ e_1, \dots, e_k}} \lambda_{d_1, \dots, d_k} \lambda_{e_1, \dots, e_k} \sum_{\substack{N \leq n < 2N \\ n \equiv v_0 \pmod{W} \\ n \equiv 1 \pmod{4} \\ n \equiv -h_i \pmod{[d_i, e_i]} \forall i}} \rho(n + h_m).$$

By definition of  $\rho(n + h_m)$  this is equal to

$$\frac{1}{\log v} \sum_{\substack{d_1, \dots, d_k \\ e_1, \dots, e_k}} \lambda_{d_1, \dots, d_k} \lambda_{e_1, \dots, e_k} \sum_{\substack{a \leq v \\ p|a \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)}{g_2(a)} \log \frac{v}{a} \sum_{\substack{N \leq n < 2N \\ n \equiv v_0 \pmod{W} \\ n \equiv 1 \pmod{4} \\ n \equiv -h_i \pmod{[d_i, e_i]} \forall i \\ n \equiv -h_m \pmod{a}}} r(n + h_m).$$

From considering the support of  $\lambda_{d_1, \dots, d_k}$  we see that for non-zero contribution we may assume  $W, [d_1, e_1], \dots, [d_k, e_k]$  and  $a$  are pairwise coprime. In this case the inner sum can be evaluated according to Lemma 5.3, taking  $q = W \prod_{i \neq m} [d_i, e_i]$  and  $d = a[d_m, e_m]$ . As  $q \ll WR^2 \ll_\epsilon N^{\theta_2+\epsilon}$  and  $d \ll vR^2 \ll N^{\theta_1+\theta_2}$  we see that the inner sum evaluates to

$$\frac{g_1(q)g_2(d)}{2qd} \pi N + O_\epsilon(N^{\frac{1}{3} + \frac{1}{2}(\theta_1+\theta_2)+\epsilon}).$$

Bounding the sum over  $a$  trivially by  $v \log v$  and using (6.12), we see the error term contributes  $O_\epsilon(N^{\frac{1}{3} + \frac{3}{2}(\theta_1+\theta_2)+\epsilon})$  which is negligible. We obtain a main term

$$\frac{X_{N,W} g_1(W) \pi N}{2W \log v} \sum_{\substack{d_1, \dots, d_k \\ e_1, \dots, e_k \\ W, [d_1, e_1], \dots, [d_k, e_k] \text{ coprime}}} \lambda_{d_1, \dots, d_k} \lambda_{e_1, \dots, e_k} \prod_{i \neq m} \frac{g_1([d_i, e_i])}{[d_i, e_i]} \frac{1}{[d_m, e_m]^2},$$

where we have defined

$$X_{N,W} = \sum_{\substack{a \leq v \\ (a,W)=1 \\ p|a \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)}{a} \log \frac{v}{a}$$

as in (5.2). The sieve sum above is of the form  $S_J$  with  $|J| = 1$ . We can evaluate this by Lemma 6.6 to obtain

$$S_2^{(m)} = (1 + o(1)) \frac{X_{N,W} g_1(W) \pi B^{k+1} N}{2W \log v} L_{k;m}^{(1)}(F).$$

Recalling the definition of  $B$  in (6.8), evaluating  $X_{N,W}$  according to Lemma 5.6, and using the fact

$$\frac{g_1(W_3)\varphi(W_3)}{W_3} = \prod_{p|W_3} \left(1 - \frac{1}{p^2}\right) = \frac{1}{2A^2} + O(D_0^{-1}),$$

we obtain

$$S_2^{(m)} = (1 + o(1)) \frac{4\sqrt{\frac{\log R}{\log v}} B^k N}{\pi W} L_{k;m}^{(1)}(F).$$

(iii) Rewrite  $S_3^{(m,l)}$  in the form

$$S_3^{(m,l)} = \sum_{\substack{d_1, \dots, d_k \\ e_1, \dots, e_k}} \lambda_{d_1, \dots, d_k} \lambda_{e_1, \dots, e_k} \sum_{\substack{N \leq n < 2N \\ n \equiv v_0 \pmod{W} \\ n \equiv 1 \pmod{4} \\ n \equiv -h_i \pmod{[d_i, e_i]} \forall i}} \rho(n + h_m) \rho(n + h_l).$$

Expanding out the definition of  $\rho$  this is

$$\frac{1}{\log^2 v} \sum_{\substack{d_1, \dots, d_k \\ e_1, \dots, e_k}} \lambda_{d_1, \dots, d_k} \lambda_{e_1, \dots, e_k} \sum_{\substack{a, b \leq v \\ p|a, b \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)\mu(b)}{g_2(a)g_2(b)} \log \frac{v}{a} \log \frac{v}{b} \cdot \sum_{\substack{N \leq n < 2N \\ n \equiv v_0 \pmod{W} \\ n \equiv 1 \pmod{4} \\ n \equiv -h_i \pmod{[d_i, e_i]} \forall i \\ n \equiv -h_m \pmod{a} \\ n \equiv -h_l \pmod{b}}} r(n + h_m) r(n + h_l).$$

Similarly to the above, for non-zero contribution we may restrict to the case  $W, [d_1, e_1], \dots, [d_k, e_k], a, b$  are pairwise coprime (note that the last two congruences are solvable if and only if  $(a, b)|h_l - h_m$ , and in the case  $(a, 2W) = (b, 2W) = 1$  this is true if and only if  $(a, b) = 1$ ). We evaluate the inner sum

according to Lemma 5.4, taking  $q = W \prod_{i \neq m,l} [d_i, e_i]$ ,  $d_1 = a[d_m, e_m]$  and  $d_2 = b[d_l, e_l]$ . We note that  $q \ll_\epsilon N^{\theta_2 + \epsilon}$  and  $d_1, d_2 \ll N^{\theta_1 + \theta_2}$ . Using the fact  $\theta_1 + \theta_2 < 1/18$ , we see the second error term in the definition of  $R_2(N; d_1, d_2, q)$  dominates, and so the inner sum evaluates to

$$\frac{g_1(q)^2 \Gamma(d_1, d_2, q)}{q} \pi^2 N + O_\epsilon(N^{\frac{5}{6} + \theta_1 + \theta_2 + \epsilon}).$$

Bounding the rest of the sum trivially, we obtain a total error of size  $O_\epsilon(N^{\frac{5}{6} + 3\theta_1 + 2\theta_2 + \epsilon})$  which, again, is negligible in the range  $\theta_1 + \theta_2 < 1/18$ . We obtain a main term

$$\frac{g_1(W)^2 \pi^2 N}{W \log^2 v} \sum_{\substack{a, b \leq v \\ p|a, b \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)\mu(b)}{g_2(a)g_2(b)} \log \frac{v}{a} \log \frac{v}{b} \cdot \sum_{\substack{d_1, \dots, d_k \\ e_1, \dots, e_k \\ W, [d_1, e_1], \dots, [d_k, e_k] \text{ coprime}}} \lambda_{d_1, \dots, d_k} \lambda_{e_1, \dots, e_k} \prod_{i \neq m, l} \frac{g_1([d_i, e_i])^2}{[d_i, e_i]} \Gamma([d_m, e_m]a, [d_l, e_l]b, W \prod_{i \neq m, l} [d_i, e_i]).$$

For arbitrary (square-free) moduli  $d_1, d_l$  and  $q$ , we can write  $\Gamma(d_1, d_2, q)$  as a product over primes (cf. the definition of  $\Gamma(d_1, d_2, q)$  given in Lemma 5.4 and note that we are summing a multiplicative function). By considering the Euler-product and the various support restrictions on the variables  $d_i, e_i, a, b$ , one can write  $\Gamma([d_m, e_m]a, [d_l, e_l]b, W \prod_{i \neq m, l} [d_i, e_i])$  in the form

$$\prod_{p|2W} \left(1 - \frac{1}{p^2}\right)^{-1} \frac{g_2(a)g_2(b)}{g_7(a)g_7(b)} \prod_{i \neq m, l} \frac{[d_i, e_i]}{g_1([d_i, e_i]\varphi([d_i, e_i])} \prod_{j=m, l} \frac{1}{[d_j, e_j]\varphi([d_j, e_j])},$$

leaving us with a main term

$$\prod_{p|2W} \left(1 - \frac{1}{p^2}\right)^{-1} \frac{g_1(W)^2 Y_{N, W} \pi^2 N}{W \log^2 v} \cdot \sum_{\substack{d_1, \dots, d_k \\ e_1, \dots, e_k \\ W, [d_1, e_1], \dots, [d_k, e_k] \text{ coprime}}} \lambda_{d_1, \dots, d_k} \lambda_{e_1, \dots, e_k} \prod_{i \neq m, l} \frac{g_1([d_i, e_i])}{\varphi([d_i, e_i])} \prod_{j=m, l} \frac{1}{[d_j, e_j]\varphi([d_j, e_j])}.$$

Here we have defined

$$Y_{N, W} = \sum_{\substack{a, b \leq v \\ (a, W) = (b, W) = 1 \\ (a, b) = 1 \\ p|a, b \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)\mu(b)}{g_7(a)g_7(b)} \log \frac{v}{a} \log \frac{v}{b}$$



as in (5.3). The main term is of the form  $S_J$  for  $|J| = 2$ . By Lemma 6.6 it can be evaluated as

$$S_3^{(l,m)} = (1 + o(1)) \frac{Y_{N,W} g_1(W)^2 \pi^2 B^{k+2} N}{W \log^2 v} L_{k;m,l}^{(2)}(F),$$

where we have written

$$\prod_{p \nmid 2W} \left(1 - \frac{1}{p^2}\right)^{-1} = 1 + O(D_0^{-1}).$$

Evaluating  $Y_{N,W}$  as in Lemma 5.6, this simplifies to

$$S_3^{(l,m)} = (1 + o(1)) \frac{64 \left(\frac{\log R}{\log v}\right) B^k N}{\pi^2 W} L_{k;m,l}^{(2)}(F).$$

(iv) Rewrite  $S_4^{(m)}$  in the form

$$\sum_{\substack{d_1, \dots, d_k \\ e_1, \dots, e_k}} \lambda_{d_1, \dots, d_k} \lambda_{e_1, \dots, e_k} \sum_{\substack{N \leq n < 2N \\ n \equiv v_0 \pmod{W} \\ n \equiv 1 \pmod{4} \\ n \equiv -h_i \pmod{[d_i, e_i]} \forall i}} \rho^2(n + h_m).$$

Expanding out the definition of  $\rho^2(n)$  we see this is equal to

$$\frac{1}{\log^2 v} \sum_{\substack{d_1, \dots, d_k \\ e_1, \dots, e_k}} \lambda_{d_1, \dots, d_k} \lambda_{e_1, \dots, e_k} \sum_{a, b \leq v} \frac{\mu(a)\mu(b)}{g_2(a)g_2(b)} \log \frac{v}{a} \log \frac{v}{b} \sum_{\substack{N \leq n < 2N \\ n \equiv v_0 \pmod{W} \\ n \equiv 1 \pmod{4} \\ n \equiv -h_i \pmod{[d_i, e_i]} \forall i \\ n \equiv -h_m \pmod{[a, b]}} r^2(n + h_m).$$

Again, we may restrict to the case  $W, [d_1, e_1], \dots, [d_k, e_k], [a, b]$  are pairwise coprime. In this case the inner sum can be evaluated according to Lemma 5.5, taking  $q = W \prod_{i \neq m} [d_i, e_i]$  and  $d = [a, b][d_m, e_m]$ . We note that  $q \ll_\epsilon N^{\theta_2 + \epsilon}$  and  $d \ll N^{2\theta_1 + \theta_2}$ , and so the inner sum becomes

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

Bounding the rest of the sum trivially, we see the error term contributes  $O_\epsilon(N^{\frac{3}{4}+2(\theta_1+\theta_2)+\epsilon})$  which is small. For the main term, let

$$Z_{N,W}^{(1)} = \sum_{a,b \leq v} \frac{\mu(a)\mu(b)g_4([a,b])}{g_2(a)g_2(b)[a,b]} \log \frac{v}{a} \log \frac{v}{b},$$

$$Z_{N,W}^{(2)} = \sum_{a,b \leq v} \frac{\mu(a)\mu(b)g_4([a,b])}{g_2(a)g_2(b)[a,b]} \log \frac{v}{a} \log \frac{v}{b} \sum_{p|[a,b]} g_6(p)$$

be as in (5.4) and (5.5). We can express  $S_4^{(m)} = \Lambda_1 + \Lambda_2 + \Lambda_3 + \Lambda_4$  where

$$\Lambda_1 = \frac{g_3(W)Z_{N,W}^{(1)}N}{W \log^2 v} \left( \log N + A_2 + 2 \sum_{p|W} g_5(p) \right) T,$$

$$\Lambda_2 = \frac{2g_3(W)Z_{N,W}^{(1)}N}{W \log^2 v} \sum_{i \neq m} \sum_{\substack{D_0 < p \leq v \\ p \equiv 3 \pmod{4}}} g_5(p) T^{(p,i)},$$

$$\Lambda_3 = -\frac{2g_3(W)Z_{N,W}^{(1)}N}{W \log^2 v} \sum_{\substack{D_0 < p \leq v \\ p \equiv 3 \pmod{4}}} g_6(p) T^{(p,m)},$$

$$\Lambda_4 = -\frac{2g_3(W)Z_{N,W}^{(2)}N}{W \log^2 v} T$$

and

$$T = \sum_{\substack{d_1, \dots, d_k \\ e_1, \dots, e_k \\ W, [d_1, e_1], \dots, [d_k, e_k] \text{ coprime}}} \lambda_{d_1, \dots, d_k} \lambda_{e_1, \dots, e_k} \prod_{i \neq m} \frac{g_1([d_i, e_i])}{[d_i, e_i]} \frac{1}{[d_m, e_m]^2},$$

$$T^{(p,i)} = \sum_{\substack{d_1, \dots, d_k \\ e_1, \dots, e_k \\ p|[d_i, e_i] \\ W, [d_1, e_1], \dots, [d_k, e_k] \text{ coprime}}} \lambda_{d_1, \dots, d_k} \lambda_{e_1, \dots, e_k} \prod_{i \neq m} \frac{g_1([d_i, e_i])}{[d_i, e_i]} \frac{1}{[d_m, e_m]^2}.$$

$T$  is of the form  $S_J$  for  $|J| = 1$ , and so by Lemma 6.6 it can be evaluated as

$$T = (1 + o(1))B^{k+1}L_{k;m}^{(1)}(F).$$

To evaluate  $T^{(p,i)}$ , note by inclusion-exclusion we can write it as

$$\left( \sum'_{\substack{d_1, \dots, d_k \\ e_1, \dots, e_k \\ p|d_i}} + \sum'_{\substack{d_1, \dots, d_k \\ e_1, \dots, e_k \\ p|e_i}} - \sum'_{\substack{d_1, \dots, d_k \\ e_1, \dots, e_k \\ p|d_i, e_i}} \right) \lambda_{d_1, \dots, d_k} \lambda_{e_1, \dots, e_k} \prod_{i \neq m} \frac{g_1([d_i, e_i])}{[d_i, e_i]} \frac{1}{[d_m, e_m]^2},$$

where  $\sum'$  denotes the condition  $W, [d_1, e_1], \dots, [d_k, e_k]$  are pairwise coprime. Thus we see it is of the form  $S_{J,p,1,i} + S_{J,1,p,i} - S_{J,p,p,i}$  for  $|J| = 1$ . By Lemma 6.6 we conclude

$$T^{(p,i)} \ll \begin{cases} \frac{F_{\max}^2 B^{k+1} (\log \log R)^2}{p}, & \text{if } i \neq m \\ \frac{F_{\max}^2 B^{k+1} (\log \log R)^2}{p^2}, & \text{if } i = m \end{cases}$$

Now we note that

$$\begin{aligned} \sum_{\substack{D_0 < p \leq v \\ p \equiv 3 \pmod{4}}} \frac{g_5(p)}{p} &\ll \sum_{p > D_0} \frac{\log p}{p^2} \ll \frac{\log D_0}{D_0}, \\ \sum_{\substack{D_0 < p \leq v \\ p \equiv 3 \pmod{4}}} \frac{g_6(p)}{p^2} &\ll \sum_{p > D_0} \frac{\log p}{p^2} \ll \frac{\log D_0}{D_0}, \end{aligned}$$

and so, with our choice  $D_0 = (\log \log N)^3$ , the contributions from  $\Lambda_2$  and  $\Lambda_3$  are negligible. Because

$$\sum_{p|W} g_5(p) \ll \sum_{p < D_0} \frac{\log p}{p} \ll \log D_0,$$

we see that the only contribution to the main term comes from the  $\Lambda_1$  term corresponding to  $\log N$ , and  $\Lambda_4$ , leaving us with

$$S_4^{(m)} = (1 + o(1)) \frac{g_3(W) B^{k+1} N}{W \log^2 v} \left[ Z_{N,W}^{(1)} \log N - 2Z_{N,W}^{(2)} \right] L_{k;m}^{(1)}(F).$$

Evaluating these according to Lemma 5.6, and using the fact

$$\frac{g_3(W_1)}{g_1(W_1)^3} = \prod_{\substack{p < D_0 \\ p \equiv 1 \pmod{4}}} \left( 1 - \frac{1}{p^2} \right)^{-1} = \frac{3\zeta(2)}{8A^2} + O(D_0^{-1}) = \frac{\pi^2}{16A^2} + O(D_0^{-1}),$$

we obtain

$$S_4^{(m)} = (1 + o(1)) \frac{2\sqrt{\frac{\log R}{\log v}} \left( \frac{\log N}{\log v} + 1 \right) B^k N}{\pi W} L_{k;m}^{(1)}(F).$$

This finishes the proof of Proposition 6.2. □

Thus it remains to establish Lemma 6.6. First we require a few technical sieve lemmas. We list these in the following section.

## 7 Technical sieve sums lemmas

In the various sieve calculations that appear in the proof of Lemma 6.6, we will frequently encounter sums of the form

$$\sum_{\substack{n \leq X \\ p|n \Rightarrow p \equiv 3 \pmod{4}}} \mu^2(n) f(n),$$

where  $f$  is a multiplicative function satisfying  $f(p) = O(1/p)$ . We can evaluate sums of this type with the following lemmas.

**Lemma 7.1** (Technical sieve sum lemma). *Let  $A_1, A_2, L > 0$ . Let  $\gamma$  be a multiplicative function satisfying the sieve axioms*

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

and

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

for any  $2 \leq w \leq z$ . Let  $g$  be the totally multiplicative function defined on primes by  $g(p) = \frac{\gamma(p)}{p - \gamma(p)}$ . Finally, let  $G : [0, 1] \rightarrow \mathbb{R}$  be a piecewise differentiable function, and let  $G_{\max} = \sup_{t \in [0, 1]} (|G(t)| + |G'(t)|)$ . Then

$$\sum_{d < z} \mu^2(d) g(d) G\left(\frac{\log d}{\log z}\right) = c_\gamma \frac{(\log z)^{\frac{1}{2}}}{\Gamma(1/2)} \int_0^1 G(x) \frac{dx}{\sqrt{x}} + O_{A_1, A_2}(c_\gamma L G_{\max} (\log z)^{-\frac{1}{2}}),$$

where

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

Here, the implied constant in the Landau notation is independent of  $G$  and  $L$ .

**Proof** This is [5, Lemma 4] with slight changes to notation.  $\square$

To use this lemma in practice, we need to be able to evaluate the singular series  $c_\gamma$  which appears. In the next lemma we do this for a function  $\gamma(p)$  which covers the cases of interest to us.

**Lemma 7.2** (Evaluation of singular series). *Let*

$$\gamma(p) = \begin{cases} 1 + O(1/p) & \text{if } p \nmid W, p \equiv 3 \pmod{4}, \\ 0 & \text{otherwise,} \end{cases}$$

With the notation of Lemma 7.1, we have

$$c_\gamma = \frac{A}{\sqrt{L(1)}} \cdot \frac{\varphi(W_3)}{W_3} (1 + O(D_0^{-1}))$$

where  $A$  is the Ramanujan–Landau constant defined in (4.1).

**Proof** Let  $\gamma(p) = 1 + \alpha(p)$  where  $\alpha(p) = O(1/p)$ . Define the auxiliary function

$$\delta(p) = \begin{cases} 1 & \text{if } p \equiv 3 \pmod{4}, \\ 0 & \text{otherwise.} \end{cases}$$

One can easily show  $c_\delta = A/\sqrt{L(1)}$ . The result follows because

$$c_\gamma = c_\delta \prod_{\substack{p|W \\ p \equiv 3 \pmod{4}}} \left(1 - \frac{1}{p}\right) \prod_{\substack{p \nmid W \\ p \equiv 3 \pmod{4}}} \left(1 - \frac{\alpha(p)}{p-1}\right)^{-1}.$$

The latter product is  $1 + O(D_0^{-1})$  by our assumption  $\alpha(p) = O(1/p)$ . □

The next lemma collects both of these results together. First we recall the definition of the normalising constant from (6.8):

$$B = \frac{2A\varphi(W_3)(\log R)^{\frac{1}{2}}}{\pi W_3}. \tag{7.1}$$

**Lemma 7.3** (Evaluation of sieve sums). *Let  $f$  be a multiplicative function such that*

$$f(p) = \frac{1}{p} + O\left(\frac{1}{p^2}\right).$$

*Then for any piece-wise smooth function  $G$  we have*

$$\sum_{\substack{d \leq R \\ (d, W) = 1 \\ p|d \Rightarrow p \equiv 3 \pmod{4}}} \mu^2(d) f(d) G\left(\frac{\log d}{\log x}\right) = B \int_0^1 G(x) \frac{dx}{\sqrt{x}} + O\left(\frac{G_{\max} B}{D_0}\right).$$

**Proof** Let  $f(p) = 1/p + g(p)$  where  $g(p) = O(1/p^2)$ , and consider the function  $\gamma$  defined on primes by

$$\gamma(p) = \begin{cases} 1 - \frac{1}{p+1+pg(p)} & \text{if } p \nmid W, p \equiv 3 \pmod{4}, \\ 0, & \text{otherwise.} \end{cases}$$

With this choice of  $\gamma(p)$  we have

$$\frac{\gamma(p)}{p - \gamma(p)} = f(p).$$

Note that

$$\begin{aligned} \sum_{w \leq p \leq z} \frac{\gamma(p) \log p}{p} &= \sum_{\substack{w \leq p \leq z \\ p \equiv 3 \pmod{4} \\ p \nmid W}} \frac{\log p}{p} + O\left( \sum_{\substack{w \leq p \leq z \\ p \equiv 3 \pmod{4} \\ p \nmid W}} \frac{\log p}{p^2} \right) \\ &= \sum_{\substack{w \leq p \leq z \\ p \equiv 3 \pmod{4}}} \frac{\log p}{p} + O\left( \sum_{\substack{w \leq p \leq z \\ p \equiv 3 \pmod{4} \\ p \nmid W}} \frac{\log p}{p} \right) + O(1) \\ &= \frac{1}{2} \log \frac{z}{w} + O(\log D_0) + O(1). \end{aligned}$$

Therefore we can apply Lemma 7.1 with  $\gamma(p)$ , taking  $L \ll 1 + \log D_0$  and  $A_2$  a suitable constant. We obtain

$$\begin{aligned} \sum_{\substack{d \leq R \\ (d, W) = 1 \\ p|d \Rightarrow p \equiv 3 \pmod{4}}} \mu^2(d) f(d) G\left(\frac{\log d}{\log x}\right) &= \frac{c_\gamma (\log R)^{\frac{1}{2}}}{\Gamma(1/2)} \int_0^1 G(x) \frac{dx}{\sqrt{x}} \\ &\quad + O\left(\frac{G_{\max} c_\gamma (1 + \log D_0)}{(\log R)^{\frac{1}{2}}}\right). \end{aligned}$$

We can evaluate  $c_\gamma$  by Lemma 7.2 to find

$$c_\gamma = \frac{A}{\sqrt{L(1)}} \cdot \frac{\varphi(W_3)}{W_3} (1 + O(D_0^{-1}))$$

When we substitute this back into our expression we see the error incurred here contributes  $O(G_{\max} B/D_0)$  and dominates. The result follows as  $\Gamma(1/2)\sqrt{L(1)} = \pi/2$ . □

We highlight the following two results, the first of which follows immediately from Lemma 7.3, and the second of which is trivial.

- (1) For multiplicative functions  $f$  satisfying  $f(p) = 1/p + O(1/p^2)$  we have the upper bound

$$\sum_{\substack{d \leq R \\ (d, W) = 1 \\ p|d \Rightarrow p \equiv 3 \pmod{4}}} \mu^2(d) f(d) \ll B. \tag{7.2}$$

(2) For multiplicative functions  $g$  satisfying  $g(p) = O(1/p^2)$  we have the upper bound

$$\sum_{\substack{d \leq R \\ (d, W)=1 \\ p|d \Rightarrow p \equiv 3 \pmod{4}}} \mu^2(d)g(d) \ll 1. \tag{7.3}$$

These sums will appear frequently in our calculations, and we will use these bounds without comment in the arguments which follow.

### 8 Establishing Lemma 6.6

Our attention now turns to establishing Lemma 6.6. We follow the combinatorial arguments used by Maynard—the steps which follow mirror those found in [9].

#### 8.1 Change of variables

Our first step to evaluating the sums appearing in Lemma 6.6 is to make a change of variables. We do so in the following proposition.

**Proposition 8.1** (Diagonalising the sieve sum). *With notation as in Lemma 6.6, denote by  $f^*, g^*$  the convolutions*

$$f^* = \mu * \frac{1}{f}, \quad g^* = \mu * \frac{1}{g}.$$

Define the diagonalising vectors  $y_{r_1, \dots, r_k}^{(J, p, m)} = y_{r_1, \dots, r_k}^{(J, p, m, f, g)}$  by

$$y_{r_1, \dots, r_k}^{(J, p, m)} = \left( \prod_{i \in I} \mu(r_i) f^*(r_i) \right) \left( \prod_{j \in J} \mu(r_j) g^*(r_j) \right) \sum_{\substack{d_1, \dots, d_k \\ r_i | d_i \ \forall i \\ p | d_m}} \lambda_{d_1, \dots, d_k} \prod_{i \in I} f(d_i) \prod_{j \in J} g(d_j).$$

Let  $y_{\max}^{(J, p, m)} = \sup_{r_1, \dots, r_k} |y_{r_1, \dots, r_k}^{(J, p, m)}|$  and  $\tilde{y}_{\max}^{(J, p, m)} = \sup_{\substack{r_1, \dots, r_k \\ (r_m, p)=1}} |y_{r_1, \dots, r_k}^{(J, p, m)}|$ . (Note that these coincide if  $p = 1$ .) Then we have

$$S_{J, p_1, p_2, m} = \sum_{\substack{u_1, \dots, u_k \\ (u_m, p_1 p_2)=1 \\ u_j=1 \ \forall j \in J}} \frac{(y_{u_1, \dots, u_k}^{(J, p_1, m)})(y_{u_1, \dots, u_k}^{(J, p_2, m)})}{\prod_{i \in I} f^*(u_i) \prod_{j \in J} g^*(u_j)} + E.$$

If  $m \in J$  then the (error) term  $E$  satisfies

$$E \ll B^{|J|} \left[ \frac{(\tilde{y}_{\max}^{(J,p_1,m)})(\tilde{y}_{\max}^{(J,p_2,m)})}{D_0} + \frac{(\tilde{y}_{\max}^{(J,p_1,m)})(\tilde{y}_{\max}^{(J,p_2,m)})}{(p_1 p_2 / (p_1, p_2))^2} + \frac{(\tilde{y}_{\max}^{(J,p_1,m)})(\tilde{y}_{\max}^{(J,p_2,m)})}{p_1^2} + \frac{(\tilde{y}_{\max}^{(J,p_1,m)})(\tilde{y}_{\max}^{(J,p_2,m)})}{p_2^2} \right].$$

If  $m \notin J$  then  $E$  satisfies a similar estimate, namely that which is obtained upon replacing all occurrences of  $p_i^2$  with  $p_i$  in the above expression, for  $i \in \{1, 2\}$ . Moreover, in both of these cases, we adopt the convention that if  $p_i = 1$  then any term in our expression for  $E$  involving  $p_i$  in the denominator may be omitted.

**Proof** Recall the definition of  $S_{J,p_1,p_2,m}$  given in Lemma 6.6:

$$S_{J,p_1,p_2,m} = \sum_{\substack{d_1, \dots, d_k \\ e_1, \dots, e_k \\ W, [d_1, e_1], \dots, [d_k, e_k] \text{ coprime} \\ p_1 | d_m, p_2 | e_m}} \lambda_{d_1, \dots, d_k} \lambda_{e_1, \dots, e_k} \prod_{i \in I} f([d_i, e_i]) \prod_{j \in J} g([d_j, e_j]). \tag{8.1}$$

We can write this in the form

$$S_{J,p_1,p_2,m} = \sum_{\substack{d_1, \dots, d_k \\ e_1, \dots, e_k \\ W, [d_1, e_1], \dots, [d_k, e_k] \text{ coprime} \\ p_1 | d_m, p_2 | e_m}} \lambda_{d_1, \dots, d_k} \lambda_{e_1, \dots, e_k} \prod_{i \in I} \frac{f(d_i) f(e_i)}{f((d_i, e_i))} \prod_{j \in J} \frac{g(d_j) g(e_j)}{g((d_j, e_j))}, \tag{8.2}$$

using multiplicativity of the functions  $f$  and  $g$ , together with the fact  $[d_i, e_i]$  is square-free for each  $i$ . We remark that because  $f, g$  are non-zero, the functions  $1/f, 1/g$  are well-defined. We note the convolution identities

$$\frac{1}{f((d_i, e_i))} = \sum_{u_i | d_i, e_i} f^*(u_i), \quad \frac{1}{g((d_j, e_j))} = \sum_{u_j | d_j, e_j} g^*(u_j)$$

for  $f^*$  and  $g^*$ . Substituting these into (8.2) and swapping the order of summation, we obtain

$$\sum_{u_1, \dots, u_k} \left( \prod_{i \in I} f^*(u_i) \right) \left( \prod_{j \in J} g^*(u_j) \right) \cdot \sum_{\substack{d_1, \dots, d_k \\ e_1, \dots, e_k \\ W, [d_1, e_1], \dots, [d_k, e_k] \text{ coprime} \\ u_i | d_i, e_i \ \forall i \in I \\ u_j | d_j, e_j \ \forall j \in J \\ p_1 | d_m, p_2 | e_m}} \lambda_{d_1, \dots, d_k} \lambda_{e_1, \dots, e_k} \prod_{i \in I} f(d_i) f(e_i) \prod_{j \in J} g(d_j) g(e_j).$$



From the support of the  $\lambda_{d_1, \dots, d_k}$ , we see the only restriction coming from the pairwise coprimality of  $W, [d_1, e_1], \dots, [d_k, e_k]$  is from the possibility  $(d_i, e_j) \neq 1$  for  $i \neq j$ . We can take care of this constraint by Möbius inversion: multiplying by  $\sum_{s_{i,j}|d_i, e_j} \mu(s_{i,j})$  for all  $i \neq j$ , we obtain

$$\sum_{u_1, \dots, u_k} \left( \prod_{i \in I} f^*(u_i) \right) \left( \prod_{j \in J} g^*(u_j) \right) \sum_{s_{1,2}, \dots, s_{k-1,k}} \left( \prod_{\substack{1 \leq i, j \leq k \\ i \neq j}} \mu(s_{i,j}) \right) \cdot \sum_{\substack{d_1, \dots, d_k \\ e_1, \dots, e_k \\ u_i | d_i, e_i \ \forall i \in I \\ u_j | d_j, e_j \ \forall j \in J \\ s_{i,j} | d_i, e_j \ \forall i \neq j \\ p_1 | d_m, p_2 | e_m}} \lambda_{d_1, \dots, d_k} \lambda_{e_1, \dots, e_k} \prod_{i \in I} f(d_i) f(e_i) \prod_{j \in J} g(d_j) g(e_j). \tag{8.3}$$

We may restrict to the case where  $s_{i,j}$  is coprime to  $s_{i,a}, s_{b,j}$  and  $u_i, u_j$ , for  $a \neq j$  and  $b \neq i$ , because the vectors  $\lambda_{d_1, \dots, d_k}$  are supported on square-free integers  $d = \prod_{i=1}^k d_i$ . Denote the sum over  $s_{i,j}$  with these conditions by  $\sum'$ . Define the diagonalising vectors

$$y_{r_1, \dots, r_k}^{(J,p,m)} = \left( \prod_{i \in I} \mu(r_i) f^*(r_i) \right) \left( \prod_{j \in J} \mu(r_j) g^*(r_j) \right) \sum_{\substack{d_1, \dots, d_k \\ r_i | d_i \ \forall i \\ p | d_m}} \lambda_{d_1, \dots, d_k} \prod_{i \in I} f(d_i) \prod_{j \in J} g(d_j). \tag{8.4}$$

From the support of  $\lambda_{d_1, \dots, d_k}$  we see that  $y_{r_1, \dots, r_k}^{(J,p,m)}$  is also supported on  $r_1, \dots, r_k$  with  $r = \prod_{i=1}^k r_i$  square-free,  $(r, W) = 1$  and  $p|r \Rightarrow p \equiv 3 \pmod{4}$ . We claim this change of variables is invertible. Indeed, from the definition (8.4), for  $d_1, \dots, d_k$  with  $\prod_{i=1}^k d_i$  square-free, we have

$$\begin{aligned} \sum_{\substack{r_1, \dots, r_k \\ d_i | r_i \ \forall i}} \frac{y_{r_1, \dots, r_k}^{(J,p,m)}}{\prod_{i \in I} f^*(r_i) \prod_{j \in J} g^*(r_j)} &= \sum_{\substack{r_1, \dots, r_k \\ d_i | r_i \ \forall i}} \prod_{i=1}^k \mu(r_i) \sum_{\substack{e_1, \dots, e_k \\ r_i | e_i \ \forall i \\ p | e_m}} \lambda_{e_1, \dots, e_k} \prod_{i \in I} f(e_i) \prod_{j \in J} g(e_j) \\ &= \sum_{e_1, \dots, e_k} 1_{p|e_m} \lambda_{e_1, \dots, e_k} \prod_{i \in I} f(e_i) \prod_{j \in J} g(e_j) \sum_{\substack{r_1, \dots, r_k \\ r_i | e_i \ \forall i \\ d_i | r_i \ \forall i}} \prod_{i=1}^k \mu(r_i) \\ &= 1_{p|d_m} \lambda_{d_1, \dots, d_k} \prod_{i \in I} \mu(d_i) f(d_i) \prod_{j \in J} \mu(d_j) g(d_j). \tag{8.5} \end{aligned}$$

With this transformation our sum (8.3) becomes

$$\sum_{u_1, \dots, u_k} \left( \prod_{i \in I} f^*(u_i) \right) \left( \prod_{j \in J} g^*(u_j) \right) \cdot \sum'_{s_{1,2}, \dots, s_{k-1,k}} \left( \prod_{\substack{1 \leq i, j \leq k \\ i \neq j}} \mu(s_{i,j}) \right) \left( \prod_{i \in I} \frac{\mu(a_i)\mu(b_i)}{f^*(a_i)f^*(b_i)} \right) \left( \prod_{j \in J} \frac{\mu(a_j)\mu(b_j)}{g^*(a_j)g^*(b_j)} \right) y_{a_1, \dots, a_k}^{(J, p_1, m)} y_{b_1, \dots, b_k}^{(J, p_2, m)},$$

where we have defined  $a_i = u_i \prod_{j \neq i} s_{i,j}$  and  $b_j = u_j \prod_{i \neq j} s_{i,j}$ . Because of our constraints on the  $s_{i,j}$  variables, we can use multiplicativity to write this as

$$\begin{aligned} & \sum_{u_1, \dots, u_k} \left( \prod_{i \in I} \frac{\mu^2(u_i)}{f^*(u_i)} \right) \left( \prod_{j \in J} \frac{\mu^2(u_j)}{g^*(u_j)} \right) \\ & \sum'_{s_{1,2}, \dots, s_{k-1,k}} \left( \prod_{\substack{i, j \in I \\ i \neq j}} \frac{\mu(s_{i,j})}{f^*(s_{i,j})^2} \right) \left( \prod_{\substack{i, j \in J \\ i \neq j}} \frac{\mu(s_{i,j})}{g^*(s_{i,j})^2} \right) \\ & \left( \prod_{i \in I, j \in J} \frac{\mu(s_{i,j})\mu(s_{j,i})}{f^*(s_{i,j})f^*(s_{j,i})g^*(s_{i,j})g^*(s_{j,i})} \right) y_{a_1, \dots, a_k}^{(J, p_1, m)} y_{b_1, \dots, b_k}^{(J, p_2, m)}. \end{aligned} \tag{8.6}$$

We now wish to reduce to the case when  $(a_m, p_1) = (b_m, p_2) = 1$ . Indeed, we will show the contribution from the alternative cases is negligible. Of course, depending on whether or not  $p_1 = 1$  and/or  $p_2 = 1$ , some (or all) of the analysis which follows is not necessary, and this accounts for the convention we assert in the statement of the proposition. First let us note the estimate

$$\frac{1}{f^*(p)} = \frac{f(p)}{1 - f(p)} = \frac{\frac{1}{p} + O\left(\frac{1}{p^2}\right)}{1 - \frac{1}{p} + O\left(\frac{1}{p^2}\right)} = \frac{1}{p} + O\left(\frac{1}{p^2}\right), \tag{8.7}$$

and similarly

$$\frac{1}{g^*(p)} = \frac{1}{p^2} + O\left(\frac{1}{p^3}\right). \tag{8.8}$$

Now, there are three cases to consider.

- (1) Suppose that  $p_1|a_m$  and  $p_2|b_m$ . This occurs if and only if  $p_1|u_m$  or  $p_1|s_{m,j}$  for some  $j \neq m$ , and  $p_2|u_m$  or  $p_2|s_{i,m}$  for some  $i \neq m$ . Suppose, for example, that  $p_1|u_m$  and  $p_2|u_m$ . Moreover let us assume that  $m \in J$ . Then one can bound the

contribution as follows:

$$\begin{aligned} &\ll (y_{\max}^{(J,p_1,m)})(y_{\max}^{(J,p_2,m)}) \left( \sum_{\substack{u \leq R \\ (u, \bar{W})=1 \\ p|u \Rightarrow p \equiv 3 \pmod{4}}} \frac{\mu^2(u)}{f^*(u)} \right)^{|I|} \\ &\cdot \left( \sum_{\substack{u \leq R \\ (u, \bar{W})=1 \\ p|u \Rightarrow p \equiv 3 \pmod{4}}} \frac{\mu^2(u)}{g^*(u)} \right)^{|J|-1} \left( \sum_{\substack{u_m \leq R \\ (u_m, \bar{W})=1 \\ p_1, p_2 | u_m \\ p|u_m \Rightarrow p \equiv 3 \pmod{4}}} \frac{\mu^2(u_m)}{g^*(u_m)} \right) \\ &\times \left( \sum_{s \geq 1} \frac{\mu^2(s)}{f^*(s)^2} \right)^{|I|^2 - |I|} \left( \sum_{s \geq 1} \frac{\mu^2(s)}{g^*(s)^2} \right)^{|J|^2 - |J|} \left( \sum_{s \geq 1} \frac{\mu^2(s)}{f^*(s)g^*(s)} \right)^{2|I||J|} \\ &\ll \frac{(y_{\max}^{(J,p_1,m)})(y_{\max}^{(J,p_2,m)})B^{|I|}}{(p_1 p_2 / (p_1, p_2))^2}. \end{aligned}$$

It is easy to see that this bound also holds in any of the other possible cases in which  $p_1|a_m$  and  $p_2|b_m$  and  $m \in J$ . If instead  $m \notin J$ , then again it is easy to see the contribution is

$$\ll \frac{(y_{\max}^{(J,p_1,m)})(y_{\max}^{(J,p_2,m)})B^{|I|}}{p_1 p_2 / (p_1, p_2)},$$

in all possible cases.

- (2) Suppose that  $p_1|a_m$  and  $p_2 \nmid b_m$ . If  $m \in J$ , then similarly to the above, one can bound the contribution by

$$\ll \frac{(y_{\max}^{(J,p_1,m)})(\tilde{y}_{\max}^{(J,p_2,m)})B^{|I|}}{p_1^2}.$$

If  $m \notin J$  then likewise one obtains a contribution

$$\ll \frac{(y_{\max}^{(J,p_1,m)})(\tilde{y}_{\max}^{(J,p_2,m)})B^{|I|}}{p_1}.$$

- (3) Finally, the case  $p_1 \nmid a_m$  and  $p_2|b_m$  proceeds as above, interchanging the roles of  $p_1$  and  $p_2$ .

Thus, we may now suppose that  $(a_m, p_1) = (b_m, p_2) = 1$ . From the support of the  $y_{a_1, \dots, a_k}^{(J, p_1, m)}$  we see there is no contribution from  $(s_{i, j}, W) \neq 1$  and so either  $s_{i, j} = 1$  or  $s_{i, j} > D_0$ . The contribution from  $s_{i, j} > D_0$  with  $i, j \in I$  is

$$\begin{aligned} &\ll (\tilde{y}_{\max}^{(J, p_1, m)})(\tilde{y}_{\max}^{(J, p_2, m)}) \left( \sum_{\substack{u \leq R \\ (u, \bar{W})=1 \\ p|u \Rightarrow p \equiv 3 \pmod{4}}} \frac{\mu^2(u)}{f^*(u)} \right)^{|I|} \left( \sum_{\substack{u \leq R \\ (u, \bar{W})=1 \\ p|u \Rightarrow p \equiv 3 \pmod{4}}} \frac{\mu^2(u)}{g^*(u)} \right)^{|J|} \\ &\quad \times \left( \sum_{s_{i, j} \geq D_0} \frac{\mu^2(s_{i, j})}{f^*(s_{i, j})^2} \right) \left( \sum_{s \geq 1} \frac{\mu^2(s)}{f^*(s)^2} \right)^{|I|^2 - |I| - 1} \left( \sum_{s \geq 1} \frac{\mu^2(s)}{g^*(s)^2} \right)^{|J|^2 - |J|} \left( \sum_{s \geq 1} \frac{\mu^2(s)}{f^*(s)g^*(s)} \right)^{2|I||J|} \\ &\ll \frac{(\tilde{y}_{\max}^{(J, p_1, m)})(\tilde{y}_{\max}^{(J, p_2, m)})B^{|I|}}{D_0}, \end{aligned}$$

This contribution will be negligible. The cases  $i, j \in J$  and  $i \in I, j \in J$  can be treated the same way. This leaves us with a main term

$$\sum_{\substack{u_1, \dots, u_k \\ (u_m, p_1 p_2) = 1}} \frac{(y_{u_1, \dots, u_k}^{(J, p_1, m)})(y_{u_1, \dots, u_k}^{(J, p_2, m)})}{\prod_{i \in I} f^*(u_i) \prod_{j \in J} g^*(u_j)}$$

To finish, we claim the contribution from  $u_j > 1$  is small whenever  $j \in J$ . Indeed if  $u_j > 1$  then it must be divisible by a prime  $p > D_0$  (with  $p \equiv 3 \pmod{4}$ ). So suppose  $|J| \geq 1$  and let  $j \in J$ . If  $u_j > 1$  we get a contribution

$$\begin{aligned} &\ll (\tilde{y}_{\max}^{(J, p_1, m)})(\tilde{y}_{\max}^{(J, p_2, m)}) \left( \sum_{\substack{u \leq R \\ (u, \bar{W})=1}} \frac{\mu^2(u)}{f^*(u)} \right)^{|I|} \left( \sum_{u \leq R} \frac{\mu^2(u)}{g^*(u)} \right)^{|J| - 1} \sum_{\substack{p > D_0 \\ p \equiv 3 \pmod{4}}} \left( \sum_{\substack{u_j < R \\ p|u_j}} \frac{\mu^2(u_j)}{g^*(u_j)} \right) \\ &\ll (\tilde{y}_{\max}^{(J, p_1, m)})(\tilde{y}_{\max}^{(J, p_2, m)})B^{|I|} \sum_{\substack{p > D_0 \\ p \equiv 3 \pmod{4}}} \frac{1}{g^*(p)} \left( \sum_{u_j} \frac{\mu^2(u_j)}{g^*(u_j)} \right) \\ &\ll \frac{(\tilde{y}_{\max}^{(J, p_1, m)})(\tilde{y}_{\max}^{(J, p_2, m)})B^{|I|}}{D_0}, \end{aligned}$$

which is small. Putting all of these facts together establishes Proposition 8.1. □

### 8.2 Transformation for $y_{r_1, \dots, r_k}^{(J, p, m)}$ and proof of Lemma 6.6 parts (i) and (ii)

Define

$$y_{r_1, \dots, r_k} = \left( \prod_{i=1}^k \mu(r_i) \varphi(r_i) \right) \sum_{\substack{d_1, \dots, d_k \\ r_i | d_i \forall i}} \frac{\lambda_{d_1, \dots, d_k}}{\prod_{i=1}^k d_i}.$$

and let  $y_{\max} = \sup_{r_1, \dots, r_k} |y_{r_1, \dots, r_k}|$ . By the inversion formula (8.5), our definition of  $\lambda_{d_1, \dots, d_k}$  in Proposition 6.2 is equivalent to taking

$$y_{r_1, \dots, r_k} = F\left(\frac{\log r_1}{\log R}, \dots, \frac{\log r_k}{\log R}\right). \tag{8.9}$$

We now wish to relate the more complicated diagonalisation vectors  $y_{r_1, \dots, r_k}^{(J, p, m)}$  to these simpler vectors. We first deal with the case when  $J = \emptyset$ , which is straightforward. By inspecting the proof of Proposition 6.2, it is clear that we only need to understand this case when  $f(p) = 1/p$ .

**Lemma 8.2** (Relating  $y_{r_1, \dots, r_k}^{(J, p_1, m)}$  to  $y_{r_1, \dots, r_k}$  when  $J = \emptyset$ ). *Suppose  $y_{r_1, \dots, r_k}^{(J, p_1, m)} \neq 0$ ,  $J = \emptyset$  and  $m \in \{1, \dots, k\}$ . Moreover suppose  $f(p) = 1/p$  for all primes  $p$ . Then the following hold.*

(1) *If  $p_1 | r_m$  then*

$$y_{r_1, \dots, r_k}^{(J, p_1, m)} = y_{r_1, \dots, r_k}.$$

(2) *If  $p_1 \nmid r_m$  then*

$$y_{r_1, \dots, r_k}^{(J, p_1, m)} = \frac{y_{r_1, \dots, p_1 r_m, \dots, r_k}}{\mu(p_1)\varphi(p_1)}.$$

**Proof** If  $f(p) = 1/p$  then  $f^*(p) = p - 1 = \varphi(p)$ . Hence we are assuming

$$y_{r_1, \dots, r_k}^{(J, p, m)} = \left(\prod_{i=1}^k \mu(r_i)\varphi(r_i)\right) \sum_{\substack{d_1, \dots, d_k \\ r_i | d_i \forall i \\ p | d_m}} \frac{\lambda_{d_1, \dots, d_k}}{\prod_{i=1}^k d_i}.$$

The result then follows by comparing with the definition of  $y_{r_1, \dots, r_k}$  given above.  $\square$

Thus, proceeding, we may suppose that  $|J| \geq 1$ . We may further suppose that  $f(p) \neq 1/p$  as this case is of no interest to us (again, this is clear by inspecting the proof of Proposition 6.2). The following proposition gives the result in full.

**Proposition 8.3** (Relating  $y_{r_1, \dots, r_k}^{(J, p_1, m)}$  to  $y_{r_1, \dots, r_k}$  when  $|J| \geq 1$ ). *Suppose  $y_{r_1, \dots, r_k}^{(J, p_1, m)} \neq 0$ ,  $|J| \geq 1$  and  $f(p) \neq 1/p$ . Then the following hold:*

(i) If  $m \in J$  and  $(r_m, p) = 1$  we have

$$\begin{aligned}
 y_{r_1, \dots, r_k}^{(J, p_1, m)} &= \frac{\mu(p_1) p_1 g(p_1)}{\varphi(p_1)} \left( \prod_{j \in J} \frac{r_j g(r_j) g^*(r_j)}{g^{**}(r_j)} \right) \\
 &\cdot \sum_{\substack{e_1, \dots, e'_m, \dots, e_k \\ r_i | e_i \ \forall i \neq m \\ r_m | e'_m \\ e_i = r_i \ \forall i \in I}} y_{e_1, \dots, p_1 e'_m, \dots, e_k} \left( \prod_{\substack{j \in J \\ j \neq m}} \frac{g^{**}(e_j)}{\varphi(e_j)} \right) \left( \frac{g^{**}(e'_m)}{\varphi(e'_m)} \right) \\
 &+ O\left( \frac{y_{\max} B^{|J|} \log \log R}{D_0 p_1^2} \right).
 \end{aligned}$$

(ii) If  $m \notin J$  and  $(r_m, p) = 1$  we have

$$\begin{aligned}
 y_{r_1, \dots, r_k}^{(J, p_1, m)} &= \frac{\mu(p_1) p_1 f(p_1)}{\varphi(p_1)} \left( \prod_{j \in J} \frac{r_j g(r_j) g^*(r_j)}{g^{**}(r_j)} \right) \\
 &\cdot \sum_{\substack{e_1, \dots, e'_m, \dots, e_k \\ r_i | e_i \ \forall i \neq m \\ e_i = r_i \ \forall i \in I \setminus \{m\} \\ e'_m = r_m}} y_{e_1, \dots, p_1 e'_m, \dots, e_k} \left( \prod_{j \in J} \frac{g^{**}(e_j)}{\varphi(e_j)} \right) + O\left( \frac{y_{\max} B^{|J|} \log \log R}{D_0 p_1} \right).
 \end{aligned}$$

(iii) If  $p_1 | r_m$  we have

$$\begin{aligned}
 y_{r_1, \dots, r_k}^{(J, p_1, m)} &= \left( \prod_{j \in J} \frac{r_j g(r_j) g^*(r_j)}{g^{**}(r_j)} \right) \\
 &\cdot \sum_{\substack{e_1, \dots, e_k \\ r_i | e_i \ \forall i \\ e_i = r_i \ \forall i \in I}} y_{e_1, \dots, e_k} \left( \prod_{j \in J} \frac{g^{**}(e_j)}{\varphi(e_j)} \right) + O\left( \frac{y_{\max} B^{|J|} \log \log R}{D_0} \right).
 \end{aligned}$$

Here  $f^{**}$  and  $g^{**}$  are defined by the convolutions

$$f^{**} = \iota \mu f * 1, \quad g^{**} = \iota \mu g * 1,$$

where  $\iota$  is the identity function,  $\iota(p) = p$ .

**Proof** We prove (i), with the rest proved in exactly the same way. Directly from the definition (8.4) we have

$$\frac{y_{r_1, \dots, r_k}^{(J, p_1, m)}}{\prod_{i \in I} \mu(r_i) f^*(r_i) \prod_{j \in J} \mu(r_j) g^*(r_j)} = \sum_{\substack{d_1, \dots, d_k \\ r_i | d_i \ \forall i \\ p_1 | d_m}} \lambda_{d_1, \dots, d_k} \prod_{i \in I} f(d_i) \prod_{j \in J} g(d_j). \tag{8.10}$$

From the inversion formula (8.5) and the definition of  $y_{r_1, \dots, r_k}$ , we see that the right hand side of (8.10) equals

$$\sum_{\substack{d_1, \dots, d_k \\ r_i | d_i \forall i \\ p_1 | d_m}} \left( \prod_{i \in I} \mu(d_i) d_i f(d_i) \right) \left( \prod_{j \in J} \mu(d_j) d_j g(d_j) \right) \sum_{\substack{e_1, \dots, e_k \\ d_i | e_i \forall i}} \frac{y_{e_1, \dots, e_k}}{\prod_{i=1}^k \varphi(e_i)}.$$

Swapping sums we obtain

$$\sum_{\substack{e_1, \dots, e_k \\ r_i | e_i \forall i \\ p_1 | e_m}} \frac{y_{e_1, \dots, e_k}}{\prod_{i=1}^k \varphi(e_i)} \sum_{\substack{d_1, \dots, d_k \\ r_i | d_i \forall i \\ d_i | e_i \forall i \\ p_1 | d_m}} \left( \prod_{i \in I} \mu(d_i) d_i f(d_i) \right) \left( \prod_{j \in J} \mu(d_j) d_j g(d_j) \right). \tag{8.11}$$

We can evaluate the inner sums using the convolution identities

$$f^{**}(n) = \sum_{d|n} \mu(d) d f(d), \quad g^{**}(n) = \sum_{d|n} \mu(d) d g(d).$$

We note that

$$f^{**}(p) = 1 - pf(p) = O\left(\frac{1}{p}\right), \tag{8.12}$$

and similarly

$$g^{**}(p) = 1 - pg(p) = 1 + O\left(\frac{1}{p}\right). \tag{8.13}$$

With our assumption  $f(p) \neq 1/p$ , we may suppose both of these functions are non-zero. Now, recall that we are assuming  $m \in J$ . Using these identities transforms (8.11) into

$$\frac{\mu(p_1) p_1 g(p_1)}{g^{**}(p_1)} \left( \prod_{i \in I} \frac{\mu(r_i) r_i f(r_i)}{f^{**}(r_i)} \right) \left( \prod_{j \in J} \frac{\mu(r_j) r_j g(r_j)}{g^{**}(r_j)} \right) \cdot \sum_{\substack{e_1, \dots, e_k \\ r_i | e_i \forall i \\ p_1 | e_m}} y_{e_1, \dots, e_k} \left( \prod_{i \in I} \frac{f^{**}(e_i)}{\varphi(e_i)} \right) \left( \prod_{j \in J} \frac{g^{**}(e_j)}{\varphi(e_j)} \right).$$

Here we are using the fact  $y_{e_1, \dots, e_k}$  is supported on square-free integers  $e = \prod_{i=1}^k e_i$ . Hence, from (8.10), it follows that

$$y_{r_1, \dots, r_k}^{(J, p_1, m)} = \frac{\mu(p_1)p_1g(p_1)}{\varphi(p_1)} \left( \prod_{i \in I} \frac{r_i f(r_i) f^*(r_i)}{f^{**}(r_i)} \right) \left( \prod_{j \in J} \frac{r_j g(r_j) g^*(r_j)}{g^{**}(r_j)} \right) \cdot \sum_{\substack{e_1, \dots, e_m', \dots, e_k \\ r_i | e_i \ \forall i \neq m \\ r_m | e_m'}} y_{e_1, \dots, p_1 e_m', \dots, e_k} \left( \prod_{i \in I} \frac{f^{**}(e_i)}{\varphi(e_i)} \right) \left( \prod_{\substack{j \in J \\ j \neq m}} \frac{g^{**}(e_j)}{\varphi(e_j)} \right) \frac{g^{**}(e_m')}{\varphi(e_m')}.$$

Here we have substituted  $e_m = p_1 e_m'$  and used the fact  $(r_m, p) = 1$ . Now, since

$$\frac{pf(p)f^*(p)}{\varphi(p)} = \frac{p}{p-1}(1-f(p)) = \frac{p}{p-1} \left[ 1 - \frac{1}{p} + O\left(\frac{1}{p^2}\right) \right] = 1 + O\left(\frac{1}{p^2}\right),$$

it follows that

$$\sup_{r_1, \dots, r_k} \prod_{j \in J} \frac{r_j f(r_j) f^*(r_j)}{\varphi(r_j)} \ll 1. \tag{8.14}$$

Similarly, since

$$\frac{pg(p)g^*(p)}{\varphi(p)} = \frac{p}{p-1}(1-g(p)) = \frac{p}{p-1} \left[ 1 + O\left(\frac{1}{p^2}\right) \right],$$

we have the bound

$$\sup_{r_1, \dots, r_k} \prod_{j \in J} \frac{r_j g(r_j) g^*(r_j)}{\varphi(r_j)} \ll \sup_{r_1, \dots, r_k} \prod_{j \in J} \frac{r_j}{\varphi(r_j)} \ll \log \log R. \tag{8.15}$$

Here we have used the fact  $r = \prod_{i=1}^k r_i \leq R$  and the standard estimate

$$\frac{n}{\varphi(n)} \ll \log \log n.$$

Now, if  $i \neq m$  then either  $e_i = r_i$  or  $e_i > D_0 r_i$ . Suppose  $e_{i_0} > D_0 r_{i_0}$  for some  $i_0 \in I$ . By first using multiplicativity of the sum over the  $e_i$  variables, and then using



estimates (8.14) and (8.15), we see that that this gives a contribution

$$\begin{aligned} &\ll \frac{y_{\max} \log \log R}{p_1^2} \left( \sum_{\substack{e_{i_0} > D_0 \\ (e_{i_0}, W)=1}} \frac{\mu^2(e_{i_0}) f^{**}(e_{i_0})}{\varphi(e_{i_0})} \right) \\ &\quad \cdot \left( \sum_{\substack{u \leq R \\ (u, W)=1 \\ p|u \Rightarrow p \equiv 3 \pmod{4}}} \frac{\mu^2(u) f^{**}(u)}{\varphi(u)} \right)^{|J|-1} \left( \sum_{\substack{u \leq R \\ (u, W)=1 \\ p|u \Rightarrow p \equiv 3 \pmod{4}}} \frac{\mu^2(u) g^{**}(u)}{\varphi(u)} \right)^{|J|} \\ &\ll \frac{y_{\max} B^{|J|} \log \log R}{D_0 p_1^2}, \end{aligned}$$

where we have used the fact  $f^{**}(p)/\varphi(p) = O(1/p^2)$  and  $g^{**}(p)/\varphi(p) = 1/p + O(1/p^2)$ , which follows from (8.12) and (8.13) respectively. This is small. Thus

$$\begin{aligned} y_{r_1, \dots, r_k}^{(J, p_1, m)} &= \frac{\mu(p_1) p_1 g(p_1)}{\varphi(p_1)} \left( \prod_{i \in I} \frac{r_i f(r_i) f^*(r_i)}{\varphi(r_i)} \right) \left( \prod_{j \in J} \frac{r_j g(r_j) g^*(r_j)}{g^{**}(r_j)} \right) \\ &\quad \cdot \sum_{\substack{e_1, \dots, e_m, \dots, e_k \\ r_i | e_i \ \forall i \neq m \\ r_m | e_m \\ e_i = r_i \ \forall i \in I}} y_{e_1, \dots, p_1 e'_m, \dots, e_k} \left( \prod_{\substack{j \in J \\ j \neq m}} \frac{g^{**}(e_j)}{\varphi(e_j)} \right) \left( \frac{g^{**}(e'_m)}{\varphi(e'_m)} \right) + O\left( \frac{y_{\max} B^{|J|} \log \log R}{D_0 p_1^2} \right). \end{aligned}$$

From the support restrictions on  $y_{r_1, \dots, r_k}^{(J, p_1, m)}$  we necessarily have  $(r_i, W) = 1$ . From the above, we see the first product may be replaced by  $1 + O(D_0^{-1})$ . This incurs an acceptable error

$$\ll \frac{y_{\max} \log \log R}{D_0 p_1^2} \left( \sum_{\substack{u \leq R \\ (u, W)=1 \\ p|u \Rightarrow p \equiv 3 \pmod{4}}} \frac{\mu^2(u) g^{**}(u)}{\varphi(u)} \right)^{|J|} \ll \frac{y_{\max} B^{|J|} \log \log R}{D_0 p_1^2},$$

which gives the result stated. □

We record the following useful corollary.

**Corollary 8.4** *With notation as in the statement of Proposition 8.1, the following estimates hold.*

(1) *If  $m \in J$  then*

$$\tilde{y}_{\max}^{(J, p_1, m)} \ll \frac{y_{\max} B^{|J|} \log \log R}{p_1^2}.$$

(2) *If  $m \notin J$  then*

$$\tilde{y}_{\max}^{(J, p_1, m)} \ll \frac{y_{\max} B^{|J|} \log \log R}{p_1}.$$

(3) We have

$$y_{\max}^{(J,p,m)} \ll y_{\max} B^{|J|} \log \log R.$$

**Proof** This follows easily from Lemma 8.2 and Proposition 8.1. (We note that in the special case  $J = \emptyset$ , in items (2) and (3) we could drop the extra  $\log \log R$  factor if required.)  $\square$

We are now in a position to prove the first two parts of Lemma 6.6.

**Proof of Lemma 6.6 parts (i) and (ii)** We prove part (i), in the case  $m \in J$ . The rest of the argument proceeds along the same lines. From Proposition 8.3 we have

$$S_{J,p_1,p_2,m} = \sum_{\substack{u_1, \dots, u_k \\ (u_m, p_1 p_2) = 1 \\ u_j = 1 \ \forall j \in J}} \frac{(y_{u_1, \dots, u_k}^{(J,p_1,m)})(y_{u_1, \dots, u_k}^{(J,p_2,m)})}{\prod_{i \in I} f^*(u_i) \prod_{j \in J} g^*(u_j)} + O\left(\frac{y_{\max}^2 B^{k+|J|} (\log \log R)^2}{(p_1 p_2 / (p_1, p_2))^2}\right).$$

Here we have used Corollary 8.4 to control the various error terms in the statement of the proposition. We have also used the fact  $|I| + |J| = k$ . It follows that

$$\begin{aligned} S_{J,p_1,p_2,m} &\ll (\tilde{y}_{\max}^{(J,p_1,m)})(\tilde{y}_{\max}^{(J,p_2,m)}) \left( \sum_{\substack{u \leq R \\ (u,W)=1 \\ p|u \Rightarrow p \equiv 3 \pmod{4}}} \frac{\mu^2(u)}{f^*(u)} \right)^{|I|} \\ &\quad + \frac{y_{\max}^2 B^{k+|J|} (\log \log R)^2}{(p_1 p_2 / (p_1, p_2))^2} \\ &\ll \frac{y_{\max}^2 B^{k+|J|} (\log \log R)^2}{(p_1 p_2 / (p_1, p_2))^2}, \end{aligned}$$

again using Corollary 8.4. This gives the result, as required.  $\square$

From now on we are only interested in the sums  $S_J = S_{J,1,1,m}$ , and in particular the cases  $|J| \in \{0, 1, 2\}$ . For ease of notation we let  $y_{r_1, \dots, r_k}^{(J)} = y_{r_1, \dots, r_k}^{(J,1,m)}$  and  $y_{\max}^{(J)} = \sup_{r_1, \dots, r_k} |y_{r_1, \dots, r_k}^{(J)}|$ . For future reference we note the bound

$$y_{\max}^{(J)} \ll y_{\max} B^{|J|} \log \log R, \tag{8.16}$$

proved above.

### 8.3 Relating vectors to functionals and proof of Lemma 6.6 part (iii)

We first prove the following corollary that follows from Proposition 8.3.

**Corollary 8.5** (Relating  $y_{r_1, \dots, r_k}^{(J)}$  to integral operators). *Let  $y_{r_1, \dots, r_k}$  be defined in terms of a fixed, smooth function  $F$ , supported on  $R_k = \{\vec{x} \in [0, 1]^k : \sum_{i=1}^k x_i \leq 1\}$ , by*

$$y_{r_1, \dots, r_k} = F\left(\frac{\log r_1}{\log R}, \dots, \frac{\log r_k}{\log R}\right).$$

Let

$$F_{max} = \sup_{(t_1, \dots, t_k) \in [0, 1]^k} |F(t_1, \dots, t_k)| + \sum_{i=1}^k \left| \frac{\partial F}{\partial t_i}(t_1, \dots, t_k) \right|.$$

Define the integral operators

$$I_{r_1, \dots, r_k; m}(F) = \int_0^1 F\left(\frac{\log r_1}{\log R}, \dots, x_m, \dots, \frac{\log r_k}{\log R}\right) \frac{dx_m}{\sqrt{x_m}},$$

$$I_{r_1, \dots, r_k; m, l}(F) = \int_0^1 \left( \int_0^1 F\left(\frac{\log r_1}{\log R}, \dots, x_l, \dots, x_m, \dots, \frac{\log r_k}{\log R}\right) \frac{dx_m}{\sqrt{x_m}} \right) \frac{dx_l}{\sqrt{x_l}}.$$

Then

(i) if  $J = \{m\}$  we have

$$y_{r_1, \dots, r_k}^{(J)} \Big|_{r_j=1 \forall j \in J} = B \left( \prod_{i \in I} \frac{\varphi(r_i)}{r_i} \right) I_{r_1, \dots, r_k; m}(F) + O\left(\frac{F_{max} B}{D_0}\right).$$

(ii) if  $J = \{m, l\}$  we have

$$y_{r_1, \dots, r_k}^{(J)} \Big|_{r_j=1 \forall j \in J} = B^2 \left( \prod_{i \in I} \frac{\varphi(r_i)}{r_i} \right)^2 I_{r_1, \dots, r_k; m, l}(F) + O\left(\frac{F_{max} B^2}{D_0}\right).$$

**Proof** First suppose  $J = \{m\}$ . From Proposition 8.3 we have

$$y_{r_1, \dots, r_k}^{(J)} \Big|_{r_j=1 \forall j \in J} = \sum_{e_m} \frac{g^{**}(e_m)}{\varphi(e_m)} F\left(\frac{\log r_1}{\log R}, \dots, \frac{\log e_m}{\log R}, \dots, \frac{\log r_k}{\log R}\right) + O\left(\frac{F_{max} B}{D_0}\right) \tag{8.17}$$

Consider the function

$$\gamma_1(p) = \begin{cases} \frac{p}{1 + \frac{\varphi(p)}{g^{**}(p)}} & \text{if } p \nmid W \prod_{i \in I} r_i \text{ and } p \equiv 3 \pmod{4}, \\ 0 & \text{otherwise.} \end{cases}$$

With this choice of  $\gamma_1(p)$  we have

$$\frac{\gamma_1(p)}{p - \gamma_1(p)} = \frac{g^{**}(p)}{\varphi(p)}$$

and one can easily check  $\gamma_1(p) = 1 + O(1/p)$ . By an argument identical to the proof of Lemma 7.3, we can evaluate the sum in (8.17) as

$$y_{r_1, \dots, r_k}^{(J)} \Big|_{r_j=1 \forall j \in J} = B \left( \prod_{i \in I} \frac{\varphi(r_i)}{r_i} \right) \int_0^1 F \left( \frac{\log r_1}{\log R}, \dots, t_m, \dots, \frac{\log r_k}{\log R} \right) \frac{dt_m}{\sqrt{t_m}} + O \left( \frac{F_{\max} B}{D_0} \right),$$

which proves (i). If  $J = \{m, l\}$  then, again using Proposition 8.3, we find

$$y_{r_1, \dots, r_k}^{(J)} \Big|_{r_j=1 \forall j \in J} = \sum_{e_m, e_l} \left( \prod_{j=m, l} \frac{g^{**}(e_j)}{\varphi(e_j)} \right) F \left( \frac{\log r_1}{\log R}, \dots, \frac{\log e_m}{\log R}, \dots, \frac{\log e_l}{\log R}, \dots, \frac{\log r_k}{\log R} \right) + O \left( \frac{F_{\max} B^2}{D_0} \right). \tag{8.18}$$

Take the sum over  $e_l$  first. Consider the function

$$\gamma_2(p) = \begin{cases} \frac{p}{1 + \frac{\varphi(p)}{g^{**}(p)}} & \text{if } p \nmid W \prod_{i \in I} r_i e_m \text{ and } p \equiv 3 \pmod{4}, \\ 0 & \text{otherwise.} \end{cases}$$

Reasoning as above, the sum on the right hand side of (8.18) becomes

$$B \left( \prod_{i \in I} \frac{\varphi(r_i)}{r_i} \right) \sum_{e_m} \frac{g^{**}(e_m)}{e_m} \int_0^1 F \left( \frac{\log r_1}{\log R}, \dots, \frac{\log e_m}{\log R}, \dots, t_l, \dots, \frac{\log r_k}{\log R} \right) \frac{dt_l}{\sqrt{t_l}} + O \left( \frac{F_{\max} B^2}{D_0} \right).$$

We can evaluate this sum in much the same way, this time using the function

$$\gamma_3(p) = \begin{cases} \frac{p}{1 + \frac{p}{g^{**}(p)}} & \text{if } p \nmid W \prod_{i \in I} r_i \text{ and } p \equiv 3 \pmod{4}, \\ 0 & \text{otherwise,} \end{cases}$$

to get the stated result. □

If we define the (identity) operator

$$I_{r_1, \dots, r_k}(F) = F \left( \frac{\log r_1}{\log R}, \dots, \frac{\log r_k}{\log R} \right),$$

then the results of Corollary 8.5 can be concisely written as

$$y_{r_1, \dots, r_k}^{(J)} \Big|_{r_j=1 \forall j \in J} = B^{|J|} \left( \prod_{i \in I} \frac{\varphi(r_i)}{r_i} \right)^{|J|} I_{r_1, \dots, r_k; J}(F) + O \left( \frac{F_{\max} B^{|J|}}{D_0} \right)$$

for  $|J| \in \{0, 1, 2\}$ . Here we have used  $I_{r_1, \dots, r_k; J}(F)$  to denote  $I_{r_1, \dots, r_k; j: j \in J}(F)$ . We are now in a position to prove the remaining part of Lemma 6.6.

**Proof of Lemma 6.6 part (iii)** We can write the operators in the statement of Lemma 6.6 as follows:

$$\begin{aligned}
 L(F) &= \int_0^1 \dots \int_0^1 \left[ I(F) \right]^2 \prod_{i=1}^k \frac{dx_i}{\sqrt{x_i}}, \\
 L_m(F) &= \int_0^1 \dots \int_0^1 \left[ I_m(F) \right]^2 \prod_{\substack{i=1 \\ i \neq m}}^k \frac{dx_i}{\sqrt{x_i}}, \\
 L_{m,l}(F) &= \int_0^1 \dots \int_0^1 \left[ I_{m,l}(F) \right]^2 \prod_{\substack{i=1 \\ i \neq m, l}}^k \frac{dx_i}{\sqrt{x_i}},
 \end{aligned}$$

where we have defined

$$\begin{aligned}
 I(F) &= F(x_1, \dots, x_k), \\
 I_m(F) &= \int_0^1 F(x_1, \dots, x_k) \frac{dx_m}{\sqrt{x_m}}, \\
 I_{m,l}(F) &= \int_0^1 \left( \int_0^1 F(x_1, \dots, x_k) \frac{dx_m}{\sqrt{x_m}} \right) \frac{dx_l}{\sqrt{x_l}}.
 \end{aligned}$$

From Proposition 8.1 we see that

$$S_J = \sum_{\substack{u_1, \dots, u_k \\ u_j=1 \forall j \in J}} \frac{(y_{u_1, \dots, u_k}^{(J)})^2}{\prod_{i \in I} f^*(u_i)} + O\left(\frac{(y_{\max}^{(J)})^2 B^{|I|}}{D_0}\right). \tag{8.19}$$

From Corollary 8.5, for  $|J| \in \{0, 1, 2\}$ , we have

$$(y_{r_1, \dots, r_k}^{(J)})^2 \Big|_{r_j=1 \forall j \in J} = B^{2|J|} \left( \prod_{i \in I} \frac{\varphi(r_i)}{r_i} \right)^{2|J|} \left[ I_{r_1, \dots, r_k; J}(F) \right]^2 + O\left(\frac{F_{\max}^2 B^{2|J|}}{D_0}\right).$$

Substituting this into (8.19), and using (8.16), yields

$$S_J = B^{2|J|} \sum_{\substack{u_1, \dots, u_k \\ u_j=1 \forall j \in J \\ (u_i, u_j)=1 \forall i \neq j \\ (u_i, W)=1 \forall i \\ p|u_i \Rightarrow p \equiv 3 \pmod{4}}} \prod_{i \in I} \frac{\mu^2(u_i) \varphi(u_i)^{2|J|}}{f^*(u_i) u_i^{2|J|}} \left[ I_{u_1, \dots, u_k; J}(F) \right]^2 \\ + O\left( \frac{F_{\max}^2 B^{2|J|}}{D_0} \sum_{\substack{u_1, \dots, u_k \\ u_j=1 \forall j \in J \\ (u_i, W)=1 \forall i \\ p|u_i \Rightarrow p \equiv 3 \pmod{4}}} \prod_{i \in I} \frac{\mu^2(u_i)}{f^*(u_i)} \right) + O\left( \frac{F_{\max}^2 B^{k+|J|} (\log \log R)^2}{D_0} \right).$$

The first error contributes

$$\ll \frac{F_{\max}^2 B^{2|J|}}{D_0} \left( \sum_{\substack{u \leq R \\ (u, W)=1 \\ p|u \Rightarrow p \equiv 3 \pmod{4}}} \frac{\mu^2(u_i)}{f^*(u_i)} \right)^{|J|} \ll \frac{F_{\max}^2 B^{k+|J|}}{D_0}.$$

For the main term, if  $(u_i, u_j) \neq 1$  then they must be divisible by a prime  $q > D_0$  with  $q \equiv 3 \pmod{4}$ . In this case we get a contribution

$$\ll F_{\max}^2 B^{2|J|} \sum_{\substack{q > D_0 \\ q \equiv 3 \pmod{4}}} \sum_{\substack{u_1, \dots, u_k \\ u_j=1 \forall j \in J \\ (u_i, W)=1 \forall i \\ p|u_i \Rightarrow p \equiv 3 \pmod{4} \\ q|u_i, u_j}} \prod_{i \in I} \frac{\mu^2(u_i) \varphi(u_i)^{2|J|}}{f^*(u_i) u_i^{2|J|}} \\ \ll F_{\max}^2 B^{2|J|} \left( \sum_{\substack{u \leq R \\ (u, W)=1 \\ p|u \Rightarrow p \equiv 3 \pmod{4}}} \frac{\mu^2(u) \varphi(u)^{2|J|}}{f^*(u) u^{2|J|}} \right)^{|J|} \sum_{q > D_0} \frac{\varphi(q)^{4|J|}}{f^*(q)^2 q^{4|J|}} \ll \frac{F_{\max}^2 B^{k+|J|}}{D_0}.$$

Thus this constraint can be removed at the cost of a negligible error and we are left with

$$S_J = B^{2|J|} \sum_{\substack{u_1, \dots, u_k \\ u_j=1 \forall j \in J \\ (u_i, W)=1 \forall i \\ p|u_i \Rightarrow p \equiv 3 \pmod{4}}} \prod_{i \in I} \frac{\mu^2(u_i) \varphi(u_i)^{2|J|}}{f^*(u_i) u_i^{2|J|}} \left[ I_{u_1, \dots, u_k; J}(F) \right]^2 \\ + O\left( \frac{F_{\max}^2 B^{k+|J|} (\log \log R)^2}{D_0} \right).$$

Now, since

$$\frac{\varphi(p)^{2|J|}}{f^*(p)p^{2|J|}} = \frac{1}{p} + O\left(\frac{1}{p^2}\right),$$

we can evaluate this multidimensional sum by applying Lemma 7.3  $|I| = k - |J|$  times. We obtain

$$S_J = B^{k+|J|}L_J(F) + O\left(\frac{F_{\max}^2 B^{k+|J|}(\log \log R)^2}{D_0}\right).$$

This completes the proof of Lemma 6.6. □

**Acknowledgements** The author would like to thank James Maynard for introducing this problem to them in the first instance, and for being available for many helpful conversations thereafter. We would also like to thank the anonymous referee for a careful and thorough reading of the paper. The author is funded by an Engineering and Physical Sciences Research Council (EPSRC) Studentship.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

### Appendix A: Estimates for $r(n)$ and $r(n)r(n + h)$

We sketch proofs for Lemmas 5.3, 5.4 and 5.5. The first follows immediately from the following two results, due to Tolev [13, Theorem] and Plaksin (cf. discussion just before [10, Lemma 4]) respectively.

**Lemma A.1** *We have*

$$\sum_{\substack{n \leq N \\ n \equiv \Delta \pmod{Q}}} r(n) = \frac{\eta(Q, \Delta)}{Q^2} \pi N + O((Q^{\frac{1}{2}} + N^{\frac{1}{3}})(\Delta, Q)^{\frac{1}{2}} \tau^4(Q) \log^4 N), \tag{A.1}$$

where  $\eta(Q, \Delta) = \#\{1 \leq \alpha, \beta \leq Q : \alpha^2 + \beta^2 \equiv \Delta \pmod{Q}\}$ .

**Lemma A.2** *We have*

$$\sum_{\substack{n \leq N \\ n \equiv \Delta \pmod{Q}}} r(n) = \frac{A(Q, \Delta)}{Q} \pi N + O(P(N; Q, \Delta)), \tag{A.2}$$

where

$$A(Q, \Delta) = \left[ 1 + \sum_{\substack{k \\ 4|2^k|(Q,4\Delta)}} \chi\left(\frac{4\Delta}{2^k}\right) \right] \sum_{r|Q} \frac{\chi(r)}{r} c_r(\Delta), \tag{A.3}$$

and  $P(N; Q, \Delta)$  satisfies

$$\int_1^N P^2(y; Q, \Delta)dy \ll_{\epsilon} (QN)^{\epsilon} (N^{\frac{3}{2}} + QN).$$

Here  $c_r(\Delta) = \sum_{(a,r)=1} e^{\frac{2\pi ia\Delta}{r}}$  denotes the Ramanujan sum.

**Proof of Lemma 5.3** Equating (A.1) and (A.2), dividing through by  $\pi N$ , and letting  $N \rightarrow \infty$  we see that

$$\frac{\eta(Q, \Delta)}{Q^2} = \frac{A(Q, \Delta)}{Q}$$

holds for any fixed  $Q, \Delta$ . Thus we may write

$$\sum_{\substack{n \leq N \\ n \equiv \Delta \pmod{Q}}} r(n) = \frac{A(Q, \Delta)}{Q} + O((Q^{\frac{1}{2}} + N^{\frac{1}{3}})(\Delta, Q)^{\frac{1}{2}} \tau^4(Q) \log^4 N). \tag{A.4}$$

Let  $Q = 4qd$  and  $\Delta \pmod{Q}$  be the solution to the congruence system  $\Delta \equiv a \pmod{q}$ ,  $\Delta \equiv 1 \pmod{4}$ , and  $\Delta \equiv 0 \pmod{d}$ , where  $(a, q) = (d, q) = 1$  and  $q, d$  are square-free and odd. Then by multiplicativity of the Ramanujan sum and our assumptions about  $\Delta, q$  and  $d$  from (A.3) we obtain

$$A(Q, \Delta) = 2 \sum_{r|q} \frac{\chi(r)\mu(r)}{r} \sum_{r|d} \frac{\chi(r)\varphi(r)}{r} = 2g_1(q)g_2(d), \tag{A.5}$$

and so we obtain the result of Lemma 5.3 on dividing through by  $Q = 4qd$ . □

For Lemma 5.4 we have the following result.

**Lemma A.3** *Suppose that  $(a, q) = (a+h, q) = (d_1, q) = (d_2, q) = (d_1, d_2) = 1$ ,  $4|h$  where  $d_1, d_2, q$  are square-free and odd, of size  $\ll N^{O(1)}$ . Then for  $0 < h < N^{\frac{3}{4}}$  we have*

$$\sum_{\substack{n \leq N \\ n \equiv a \pmod{q} \\ n \equiv 1 \pmod{4} \\ d_1|n, d_2|n+h}} r(n)r(n+h) = \frac{g_1(q)^2}{q} \Gamma(h, d_1, d_2, q) \pi^2 N + R_2(N, d_1, d_2, q), \tag{A.6}$$

where

$$R_2(N, d_1, d_2, q) \ll_{\epsilon} q^{\frac{1}{2}} d_1 d_2 N^{\frac{3}{4}+\epsilon} + d_1^{\frac{1}{2}} d_2^{\frac{1}{2}} N^{\frac{5}{6}+\epsilon},$$

where

$$\Gamma(h, d_1, d_2, q) = \frac{g_2(d_1)g_2(d_2)}{d_1 d_2} \sum_{(r, 2q)=1} \frac{c_r(h)}{r^2} \frac{(d_1, r)(d_2, r)}{\Psi(d_1, r)\Psi(d_2, r)} \chi[(d_1^2, r)]\chi[(d_2^2, r)]$$



and  $g_1$  is the multiplicative function defined on primes by  $g_1(p) = 1 - \frac{\chi(p)}{p}$  and  $\Psi(d_1, r) = g_2((d_1, r/(r, d_1)))$ .

**Proof of Lemma 5.4** Under the additional assumption that  $p|h \Rightarrow p|2q$ , then we see for  $r$  considered in the sum  $c_r(h) = \mu(r)$ . But now restricting to square-free  $r$ , since  $d_1, d_2$  are square-free  $\Psi(d_1, r) = \Psi(d_2, r) = 1$ . Thus in this case

$$\Gamma(h, d_1, d_2, q) = \frac{g_2(d_1)g_2(d_2)}{d_1d_2} \sum_{(r, 2q)=1} \frac{\mu(r)(d_1, r)(d_2, r)}{r^2} \chi[(d_1^2, r)]\chi[(d_2^2, r)]. \tag{A.7}$$

This is the form stated in Lemma 5.4. □

Now we briefly outline the proof of Lemma A.3. In [10, Lemma 4] a similar sum to (A.6) is considered, this time under the hypotheses  $p|d_1, d_2 \Rightarrow p \equiv 3 \pmod{4}$ ,  $q = 1$  and the congruence  $n \equiv 1 \pmod{4}$  is omitted. The proof of Lemma A.3 is similar to the proof found there, with few minor changes. The key point is to note that, using the convolution identity  $r = 4(\chi * 1)$  and complete multiplicativity of  $\chi$ , for  $n \equiv 1 \pmod{4}$  we can write

$$\begin{aligned} \frac{r(n)}{4} &= \sum_{m|n} \chi(m) = \sum_{\substack{m|n \\ m \leq \sqrt{N}}} \chi(m) + \sum_{\substack{m|n \\ m > \sqrt{N}}} \chi(m) = \sum_{\substack{m|n \\ m \leq \sqrt{N}}} \chi(m) + \sum_{\substack{l|n \\ l < \frac{n}{\sqrt{N}}}} \chi(l) \\ &= 2 \sum_{\substack{m|n \\ m \leq \sqrt{N}}} \chi(m) - \sum_{\substack{m|n \\ \frac{n}{\sqrt{N}} \leq m \leq \sqrt{N}}} \chi(m). \end{aligned} \tag{A.8}$$

Using this we may expand out  $r(n)$  in the sum (A.6). We obtain (after swapping the order of summation)

$$\sum_{\substack{n \leq N \\ n \equiv a \pmod{q} \\ n \equiv 1 \pmod{4} \\ d_1|n, d_2|n+h}} r(n)r(n+h) = 4 \sum_{m \leq \sqrt{N}} \chi(m) \left[ 2 \sum_{\substack{n \leq N \\ n \equiv a \pmod{q} \\ n \equiv 1 \pmod{4} \\ d_1|n, d_2|n+h \\ m|n}} r(n+h) - \sum_{\substack{n \leq m\sqrt{N} \\ n \equiv a \pmod{q} \\ n \equiv 1 \pmod{4} \\ d_1|n, d_2|n+h \\ m|n}} r(n+h) \right].$$

These congruences have a solution if and only if  $(m, q) = 1$  and  $(m, d_2)|h$ . In this case we can use the Chinese Remainder Theorem and write

$$\begin{aligned} \sum_{\substack{n \leq N \\ n \equiv a \pmod{q} \\ n \equiv 1 \pmod{4} \\ d_1|n, d_2|n+h}} r(n)r(n+h) &= 4 \sum_{\substack{m \leq \sqrt{N} \\ (m, q)=1 \\ (m, d_2)|h}} \chi(m) \left[ 2 \sum_{\substack{n \leq N \\ n \equiv \Delta \pmod{Q}}} r(n) - \sum_{\substack{n \leq m\sqrt{N} \\ n \equiv \Delta \pmod{Q}}} r(n) \right] \\ &\quad + O_\epsilon((N^{\frac{1}{2}} + h)N^\epsilon), \end{aligned} \tag{A.9}$$

where  $Q = 4q[m, d_1, d_2]$  and  $\Delta \pmod{Q}$  satisfies the congruence system  $\Delta \equiv a + h \pmod{q}$ ,  $\Delta \equiv 1 \pmod{4}$ ,  $\Delta \equiv h \pmod{m}$ ,  $\Delta \equiv h \pmod{d_1}$  and  $\Delta \equiv$

$0 \pmod{d_2}$ . The error term arises by estimating the intervals of length  $h$  left over: taking absolute values, using the divisor bound  $r(n) \ll n^\epsilon$ , and noting that in this regime  $h \ll N^{3/4}$ , we have to estimate sums of type

$$N^\epsilon \sum_{m \leq \sqrt{N}} \sum_{\substack{X < n \leq X+h \\ n \equiv h \pmod{m}}} 1.$$

Applying the standard estimate  $h/m + O(1)$  to the inner sum and then carrying out the summation over  $m$  yields the desired estimate (after redefining our choice of  $\epsilon$ ).

Now the inner sums in (A.9) consist of estimating  $r(n)$  in arithmetic progressions. It proves convenient to proceed using formula (A.4). We obtain

$$4\pi \sum_{\substack{m \leq \sqrt{N} \\ (m,q)=1 \\ (m,d_2)|h}} \chi(m) \left[ \frac{2N - m\sqrt{N}}{Q} \right] A(Q, \Delta) + R_2(N; d_1, d_2, q) + O_\epsilon((N^{1/2} + h)N^\epsilon), \tag{A.10}$$

where  $R_2$  is an error term. Using the fact  $(\Delta, Q) \leq d_1 d_2(m, h)$  and  $Q \ll mq d_1 d_2$  one can show  $R_2$  contributes

$$R_2(N; d_1, d_2, q) \ll q^{1/2} d_1 d_2 N^{3/4+\epsilon} + d_1^{1/2} d_2^{1/2} N^{5/6+\epsilon},$$

and so this error term dominates. Now one can estimate the main term in (A.10) using the same techniques found in [10, Lemma 4].

For Lemma 5.5 we have the following result.

**Lemma A.4** *Let  $(a, q) = (d, q) = 1$  where  $d, q$  are square-free and odd, of size  $\ll N^{O(1)}$ . Then we have*

$$\sum_{\substack{n \leq x \\ n \equiv a \pmod{q} \\ n \equiv 1 \pmod{4} \\ d|n}} r^2(n) = \frac{16H(1, d, 4q)}{\varphi(4q)} \left[ \log x + 2\gamma - 1 + \frac{H'(1, d, 4q)}{H(1, d, 4q)} \right] x + O_\epsilon(qx^{3/4+\epsilon}), \tag{A.11}$$

where

$$H(s, d, 4q) = \frac{L^2(s, \chi_4)G(s, 4q)A(s, d)}{\zeta(2s)},$$

$$G(s, 4q) = \prod_{p|4q} \left(1 - \frac{1}{p^s}\right) \left(1 + \frac{1}{p^s}\right)^{-1} \left(1 - \frac{\chi_4(p)}{p^s}\right)^2,$$

$$A(s, d) = \prod_{\substack{p|d \\ p \equiv 3 \pmod{4}}} \frac{1}{p^{2s}} \prod_{\substack{p|d \\ p \equiv 1 \pmod{4}}} \frac{4p^{2s} - 3p^s + 1}{p^{2s}(p^s + 1)}.$$

**Proof of Lemma 5.5** One obtains the main term as stated in Lemma 5.5 by taking the logarithmic derivative of  $H(s, d, 4q)$  (taking an appropriate branch-cut) and evaluating at  $s = 1$ . □

We now outline the proof of Lemma A.4. The proof uses a standard application of Perron’s formula. The following lemma, which combines [12, Theorem II.8.20] and [12, Theorem II.8.22], will prove necessary.

**Lemma A.5** *Let  $\mathcal{L} = \log(|t| + Q + 1)$  and let  $\chi_Q \pmod{Q}$  be a Dirichlet character. For  $\sigma \geq 1$  we have the following*

- (1) *If  $\chi_Q^2$  is complex then  $L(s, \chi_Q^2)^{-1} \ll \mathcal{L}^7$ .*
- (2) *If  $\chi_Q^2$  is real, non-trivial, then there exists an absolute constant  $c_0 > 0$  such that*

$$L(s, \chi_Q^2)^{-1} \ll \begin{cases} \mathcal{L}^6(\mathcal{L} + 1/|t|) & \text{if } |t| > c_0 Q^{-\frac{1}{2}}(\log 2Q)^{-2}, \\ Q^{\frac{1}{2}} & \text{if } |t| \leq c_0 Q^{-\frac{1}{2}}(\log 2Q)^{-2}. \end{cases}$$

Write  $Q = 4q$  and let  $\Delta \pmod{Q}$  be the unique solution to the congruences  $\Delta \equiv a \pmod{q}$  and  $\Delta \equiv 1 \pmod{4}$ . Since  $(\Delta, Q) = 1$  by character orthogonality we can write

$$\sum_{\substack{n \leq x \\ n \equiv \Delta \pmod{Q} \\ d|n}} r^2(n) = \frac{1}{\varphi(Q)} \sum_{\chi_Q \pmod{Q}} \overline{\chi(\Delta)} \sum_{\substack{n \leq x \\ d|n}} r^2(n) \chi_Q(n).$$

We study these sums by considering their generating series

$$F_d(s, \chi_Q) = \sum_{n=1}^{\infty} \frac{r^2(dn) \chi_Q(dn)}{(dn)^s}$$

Write  $F(s, \chi_Q) = F_1(s, \chi_Q)$ . Then

$$F(s, \chi_Q) = 16L^2(s, \chi_Q)L^2(s, \chi_4\chi_Q)L(2s, \chi_Q^2)^{-1}.$$

Here  $\chi_4$  denotes the unique non-trivial character  $\pmod{4}$ , and  $L(s, \chi_D)$  denotes the L series corresponding to the Dirichlet character  $\chi_D \pmod{D}$ . It follows that

$$\begin{aligned} \frac{F_d(s, \chi_Q)}{F(s, \chi_Q)} &= \prod_{p|d} \left( \sum_{k=1}^{\infty} \frac{r(p^k)^2 \chi_Q(p^k)}{p^{ks}} \right) \left( \sum_{k=0}^{\infty} \frac{r(p^k)^2 \chi_Q(p^k)}{p^{ks}} \right)^{-1} \\ &= \prod_{\substack{p|d \\ p \equiv 3 \pmod{4}}} \frac{\chi_Q(p)^2}{p^{2s}} \prod_{\substack{p|d \\ p \equiv 1 \pmod{4}}} \frac{\frac{4p^{2s}}{\chi_Q(p)^2} - \frac{3p^s}{\chi_Q(p)} + 1}{\frac{p^{2s}}{\chi_Q(p)^2} \left( \frac{p^s}{\chi_Q(p)} + 1 \right)} = A(s, d, Q) \end{aligned}$$

say, using the fact  $d$  is square-free and  $(d, Q) = 1$ . We note the divisor bound  $|r^2(n)\chi_Q(n)| \leq \tau^2(n) \leq e^C \frac{\log n}{\log \log n}$  for some explicit  $C > 0$ . We can apply an effective form of Perron’s formula (for example [12, Corollary II.2.4]), averaging over the height  $T$ , to obtain

$$\sum_{\substack{n \leq x \\ n \equiv \Delta \pmod{Q} \\ d|n}} r^2(n) = I(T; Q, \Delta, d) + O\left(\frac{x(\log x)^2}{T} + e^{\frac{2C \log x}{\log \log x}} \left(1 + \frac{x \log T}{T}\right)\right),$$

where

$$I = \frac{1}{\varphi(Q)} \sum_{\chi_Q \pmod{Q}} \overline{\chi_Q(\Delta)} \left[ \frac{1}{T} \int_T^{2T} \left( \frac{1}{2\pi i} \int_{c-it_0}^{c+it_0} \frac{16L^2(s, \chi_Q)L^2(s, \chi_4\chi_Q)A(s, d, Q)x^s}{L(2s, \chi_Q^2)s} ds \right) dt_0 \right].$$

Here  $c = 1 + 1/\log x$ . We move the contour to the region defined by  $[c - it_0, c + it_0]$ ,  $[1/2 + it_0, c + it_0]$ ,  $[1/2 - it_0, 1/2 + it_0]$  and  $[1/2 - it_0, c + it_0]$ . The integrand has a pole of order 2 at  $s = 1$ , coming from the trivial character  $\chi_0 \pmod{Q}$ . The residue of this pole is

$$H(1, d, Q) \left( \log x + 2\gamma - 1 + \frac{H'(1, d, Q)}{H(1, d, Q)} \right) x,$$

where

$$\begin{aligned} H(s, d, Q) &= \frac{16L^2(s, \chi_4)G(s, Q)A(s, d)}{\zeta(2s)}, \\ G(s, Q) &= \prod_{p|Q} \left(1 - \frac{1}{p^s}\right) \left(1 + \frac{1}{p^s}\right)^{-1} \left(1 - \frac{\chi_4(p)}{p^s}\right)^2, \\ A(s, d) &= \prod_{\substack{p|d \\ p \equiv 3 \pmod{4}}} \frac{1}{p^{2s}} \prod_{\substack{p|d \\ p \equiv 1 \pmod{4}}} \frac{4p^{2s} - 3p^s + 1}{p^{2s}(p^s + 1)}. \end{aligned}$$

We need bounds for the integrand in the region  $\sigma \geq 1/2$  and  $|t| \leq 2T$ . One can easily check the trivial bound  $|A(s, d, Q)| \ll 1$  (uniformly in  $d$  and  $q$ ). We require a lower bound for  $|L(2s, \chi_Q^2)|$  in this region. If  $\chi_Q$  is real then  $\chi_Q^2$  is the trivial character, and so we have

$$L(2s, \chi_Q^2) = \zeta(2s) \prod_{p|Q} \left(1 - \frac{1}{p^{2s}}\right).$$

In the region  $\sigma \geq \frac{1}{2}$  we can bound the product from below by  $\varphi(Q)/Q$ . Standard bounds for  $\zeta(s)$  on the 1-line (see for example [12, Theorem II.3.9]) then tell us that

$$L(2s, \chi_Q^2)^{-1} \ll \frac{Q}{\varphi(Q)} \log(|t| + 2) \ll_\epsilon (QT)^\epsilon.$$

If  $\chi_Q$  is complex then we have to be more careful. We use Lemma A.5, and see there is a minor technical complication where we must bound the contribution from  $s = 1/2 + it$  with  $|t| \leq 2c_0$  (say) separately (with this choice of cut-off we can use the bound  $L(2s, \chi_Q^2) \ll (QT)^\epsilon$  uniformly in  $\chi_Q$  for all the other contours). One can easily show the contribution from these values of  $s$  is  $\ll Q^{\frac{1}{2}}$ , which is small.

To bound the contribution from the other integrals, we can use the fourth-moment estimate for Dirichlet L-functions on the critical line in the form

$$\frac{1}{\varphi(D)T} \sum_{\chi \pmod{D}} \int_T^{2T} |L(1/2 + it, \chi)|^4 dt \ll (DT)^\epsilon,$$

(cf. [2, Theorem 1]) together with the Cauchy-Schwarz inequality. One can show that, with the choice  $T = x^{\frac{1}{4}}$ , the contours contribute  $\ll_\epsilon Qx^{\frac{3}{4} + \epsilon}$ .

### Appendix B: Auxiliary estimates for $\rho(n)$

To evaluate the sums appearing in Lemma 5.6 we use the Selberg–Delange method.

**Lemma B.1** (Selberg–Delange method). *Let  $F(s) = \sum_{n=1}^\infty a_n n^{-s}$  be a Dirichlet series such that the function  $G(s; z) = F(s)\zeta(s)^{-z}$  can be analytically continued to the region*

$$\sigma > 1 - \frac{c_0}{1 + \log(2 + |t|)} \tag{B.1}$$

for some positive  $c_0 > 0$  and  $z \in \mathbb{C}$  with  $|z| \leq A$ , and moreover satisfies the bounds  $|G(s; z)| \ll M(1 + |t|)^{1-\delta}$  for some  $\delta > 0$  in this region. Let  $Z_1(s; z) = [\zeta(s)(s - 1)]^z$ ,  $Z_2(s; z) = \frac{Z_1(s; z)}{s}$  (both holomorphic in the disc  $|s - 1| < 1$ ) and let

$$G(s; z)Z_1(s; z) = \sum_{k=0}^\infty \mu_k(z)(s - 1)^k,$$

$$G(s; z)Z_2(s; z) = \sum_{k=0}^\infty \eta_k(z)(s - 1)^k$$

be the Taylor series in this region. Then for any  $N \geq 0$  and  $|z| \leq A$  we have

$$\sum_{n \leq x} \frac{a_n}{n} \log \frac{x}{n} = (\log x)^{z+1} \left[ \sum_{k=0}^N \frac{\mu_k(z)}{\Gamma(z + 2 - k)(\log x)^k} + O(MR_N(x)) \right], \tag{B.2}$$

and moreover if  $a_n > 0$  then we also have

$$\sum_{n \leq x} a_n = x(\log x)^{z-1} \left[ \sum_{k=0}^N \frac{\eta_k(z)}{\Gamma(z - k)(\log x)^k} + O(MR_N(x)) \right], \tag{B.3}$$

where

$$R_N(x) = R_N(x; c_1, c_2) = e^{-c_1\sqrt{\log x}} + O\left(\frac{c_2 N + 1}{\log(x + 1)}\right)^{N+1} \tag{B.4}$$

for some suitable constants  $c_1, c_2 > 0$ . These positive constants, and the implicit constants in the Landau symbol, depend at most on  $c_0, \delta$  and  $A$ .

The second statement is exactly [12, Theorem II.5.2]. The first statement follows by the same proof with a few minor changes.

Lemma 5.6 follows from suitable applications of Lemma B.1. We sketch the details in each case.

**Proof of Lemma 5.6** (i) For  $X_{N,W}$  recall from (5.2) the definition

$$X_{N,W} = \sum_{\substack{a \leq v \\ (a,W)=1 \\ p|a \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)}{a} \log \frac{v}{a}. \tag{B.5}$$

Consider the generating series, for  $\sigma > 1$

$$h_1(W, s) = \sum_{\substack{n=1 \\ (n,W)=1 \\ p|n \Rightarrow p \equiv 1 \pmod{4}}}^{\infty} \frac{\mu(n)}{n^s} = \prod_{\substack{p \nmid W \\ p \equiv 1 \pmod{4}}} \left(1 - \frac{1}{p^s}\right) = \frac{K_1(s)G_1(W, s)}{(\zeta(s)L(s, \chi_4))^{\frac{1}{2}}},$$

where

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

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

(taking the positive determination of the square-root in the first instance). Both of these functions are analytic for  $\sigma > 3/4$  (say).  $K_1(s)$  is bounded in this region, and  $G_1(W, s)$  satisfies

$$\begin{aligned} |G_1(W, s)| &\leq \prod_{\substack{p|W \\ p \equiv 1 \pmod{4}}} \left(1 - \frac{1}{p^{\frac{3}{4}}}\right)^{-1} = \exp\left[-\sum_{\substack{p \leq D_0 \\ p \equiv 1 \pmod{4}}} \log\left(1 - \frac{1}{p^{3/4}}\right)\right] \\ &= \exp \sum_{\substack{p \leq D_0 \\ p \equiv 1 \pmod{4}}} \left[\frac{1}{p^{3/4}} + O\left(\frac{1}{p^{3/2}}\right)\right] \ll \exp \sum_{\substack{p \leq D_0 \\ p \equiv 1 \pmod{4}}} \frac{D_0^{1/4}}{p} \\ &\ll \exp(D_0^{1/4} \log \log D_0). \end{aligned}$$

We remark that with the choice  $D_0 = (\log \log N)^3$ , we certainly have the bound

$$\exp(D_0^{1/4} \log \log D_0) \ll_\epsilon (\log N)^\epsilon$$

for any fixed  $\epsilon > 0$ , and using this estimate, it is a simple task to verify that all the error terms which follow are indeed controlled. (Recall  $v = N^\theta$  for some fixed  $\theta > 0$ .)

From the classical zero-free region for Dirichlet L-functions,  $L(s)^{-1}$  can be analytically continued to a region of the form (B.1) for some  $c_0$  and moreover satisfies a bound  $L(s)^{-1} \ll \log(2 + |t|)$  in this region (cf. [3, Chapter 14]). Thus we can apply Lemma B.1 with  $M = \exp(D_0^{1/4} \log \log D_0)$ ,  $z = -\frac{1}{2}$ ,  $N = 0$  and some suitable choices of  $c_0, \delta$ , to obtain

$$X_{N,W} = \frac{K_1(1)G_1(W)}{\Gamma(3/2)\sqrt{L(1, \chi_4)}} (\log v)^{\frac{1}{2}} + O\left(\frac{\exp(D_0^{1/4} \log \log D_0)}{(\log v)^{\frac{1}{2}}}\right).$$

Here we have written  $G_1(W) = G_1(W, 1)$  (and this convention will be continued below the fold). This simplifies to the stated result. Note that

$$G_1(W) = \prod_{\substack{p \leq D_0 \\ p \equiv 1 \pmod{4}}} \left(1 - \frac{1}{p}\right)^{-1} \asymp (\log D_0)^{\frac{1}{2}}$$

by Mertens' theorem.

(ii) For  $Y_{N,W}$  recall from (5.3) the definition

$$Y_{N,W} = \sum_{\substack{a,b \leq v \\ (a,W)=(b,W)=1 \\ (a,b)=1 \\ p|a,b \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)\mu(b)}{g_7(a)g_7(b)} \log \frac{v}{a} \log \frac{v}{b}. \tag{B.6}$$

Take the sum over  $b$  on the inside. We can evaluate

$$\sum_{\substack{b \leq v \\ (b,aW)=1 \\ p|b \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(b)}{g_7(b)} \log \frac{v}{b}$$

using the generating function, for  $\sigma > 1$

$$h_2(aW, s) = \sum_{\substack{n=1 \\ (n,aW)=1 \\ p|n \Rightarrow p \equiv 1 \pmod{4}}}^\infty \frac{\mu(n)n}{n^s g_7(n)} = h_1(aW, s) K_2(s) G_2(aW, s),$$

where

$$K_2(s) = \prod_{p \equiv 1 \pmod{4}} \left( 1 + \frac{1}{(p^s - 1)(p + 1)} \right),$$

$$G_2(aW, s) = \prod_{\substack{p|aW \\ p \equiv 1 \pmod{4}}} \left( 1 + \frac{1}{(p^s - 1)(p + 1)} \right)^{-1}.$$

Arguing as above, Lemma B.1 yields

$$Y_{N,W} = T_1 + T_2 + O(T_3),$$

this time taking  $M = \exp(D_0^{1/4} \log \log D_0)$ ,  $z = -1/2$  and  $N = 1$ . Here

$$T_3 = \frac{M}{(\log v)^{\frac{3}{2}}} \sum_{\substack{a \leq v \\ (a, W) = 1 \\ p|a \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu^2(a)\tau(a)}{g_7(a)} \log \frac{v}{a} \ll \frac{M}{(\log v)^{\frac{1}{2}}} \prod_{\substack{p \leq v \\ p \equiv 1 \pmod{4}}} \left( 1 + \frac{2}{p + 1} \right)$$

$$\ll \frac{M}{(\log v)^{\frac{1}{2}}} \prod_{\substack{p \leq v \\ p \equiv 1 \pmod{4}}} \left( 1 + \frac{1}{p} \right)^2 \ll M(\log v)^{\frac{1}{2}} = \exp(D_0^{1/4} \log \log D_0)(\log v)^{\frac{1}{2}}.$$

To evaluate  $T_2$ , we need a handle on the second Taylor series coefficient. With notation as in the statement of Lemma B.1, this is labelled  $\mu_1(-1/2)$  and is precisely equal to

$$G(1)Z'_1(1; -1/2) + G'(1)Z_1(1; -1/2),$$

where

$$G(s) = \frac{K_1(s)G_1(aW, s)K_2(s)G_2(aW, s)}{L(s, \chi_4)^{1/2}}$$

(taking an appropriate branch-cut). Here all derivatives are taken with respect to  $s$ . By a series expansion, one can directly show that  $Z_1(1; -1/2) = 1$  and  $Z'_1(1; -1/2) = -\gamma/2$ , where  $\gamma$  is the Euler-Mascheroni constant. We now note



the following:

$$G_1(p)G_2(p) = \frac{g_7(p)}{p},$$

$$G'_1(d) = G_1(d) \sum_{\substack{p|d \\ p \equiv 1 \pmod{4}}} \frac{\log p}{p-1},$$

$$G'_2(d) = G_2(d) \sum_{\substack{p|d \\ p \equiv 1 \pmod{4}}} \frac{\log p}{p(p-1)}.$$

Here we have defined  $G'_i(d) = \frac{dG_i(d,s)}{ds} \Big|_{s=1}$  for  $i \in \{1, 2\}$ . These results are valid for primes  $p \equiv 1 \pmod{4}$ , and for any square-free  $d$ . Thus one can write  $T_2$  as a finite linear combination

$$T_2 = \frac{1}{(\log v)^{\frac{1}{2}}} \sum_i c_i \alpha_i(W_1) R_i,$$

where  $c_i \in \mathbb{R}$  are bounded,

$$\alpha_i(W_1) \in \left\{ \frac{g_7(W_1)}{W_1}, \frac{g_7(W_1)}{W_1} \sum_{p|W_1} \frac{\log p}{p-1}, \frac{g_7(W_1)}{W_1} \sum_{p|W_1} \frac{\log p}{p(p-1)} \right\},$$

and  $R_i$  is one of

$$\sum_{\substack{a \leq v \\ (a, W) = 1 \\ p|a \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)}{a} \log \frac{v}{a} \ll (\log D_0)^{1/2} (\log v)^{\frac{1}{2}}$$

or

$$\sum_{\substack{D_0 < q < v \\ q \equiv 1 \pmod{4}}} \frac{\log q}{q-1} \sum_{\substack{a \leq v \\ (a, W) = 1 \\ q|a \\ p|a \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)}{a} \log \frac{v}{a} \ll \exp(D_0^{1/4} \log \log D_0) (\log v)^{1/2},$$

or finally

$$\sum_{\substack{D_0 < q < v \\ q \equiv 1 \pmod{4}}} \frac{\log q}{q(q-1)} \sum_{\substack{a \leq v \\ (a, W) = 1 \\ q|a \\ p|a \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)}{a} \log \frac{v}{a} \ll \frac{(\log D_0)^{3/2} (\log v)^{1/2}}{D_0^2}.$$

Here  $q$  denotes a prime variable (a convention that will also be used below the fold). The first estimate follows directly from our expression for  $X_{N,W}$  found above. For the last two estimates, we note that the inner sum appearing in both is precisely  $X_{N,W} - X_{N,qW}$ , and from our work above it follows that

$$\begin{aligned} X_{N,W} - X_{N,qW} &= \frac{K_1(1)}{\Gamma(3/2)\sqrt{L(1, \chi_4)}}(\log v)^{\frac{1}{2}}(G_1(W) - G_1(qW)) \\ &\quad + O\left(\frac{\exp(D_0^{\frac{1}{4}} \log \log D_0)}{(\log v)^{\frac{1}{2}}}\right) \\ &\ll \frac{(\log D_0)^{\frac{1}{2}}(\log v)^{\frac{1}{2}}}{q} \\ &\quad + \frac{\exp(D_0^{\frac{1}{4}} \log \log D_0)}{(\log v)^{\frac{1}{2}}}. \end{aligned}$$

Now, standard estimates for sums over primes such as

$$\sum_{\substack{q > D_0 \\ q \equiv 1 \pmod{4}}} \frac{\log q}{q(q-1)} \ll \frac{\log D_0}{D_0},$$

yields the results stated.

Now, using the fact

$$\sum_{p|W_1} \frac{\log p}{p-1} \ll \log D_0,$$

and

$$\frac{g_7(W_1)}{W_1} = \prod_{\substack{p \leq D_0 \\ p \equiv 1 \pmod{4}}} \left(1 + \frac{1}{p}\right) \ll (\log D_0)^{\frac{1}{2}},$$

it follows that the total contribution from  $T_2$  is

$$T_2 \ll (\log D_0)^{3/2} \exp(D_0^{1/4} \log \log D_0)(\log v)^{1/2}$$

which is negligible.

Finally, we can write

$$T_1 = \frac{K_1(1)K_2(1)G_1(W_1)G_2(W_1)(\log v)^{\frac{1}{2}}}{\Gamma(3/2)\sqrt{L(1, \chi_4)}} \sum_{\substack{a \leq v \\ (a,W)=1 \\ p|a \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)}{a} \log \frac{v}{a}$$

This last sum is exactly  $X_{N,W}$ . Using our bound from part (i), and also the fact

$$K_2(1)G_2(W_1) = \prod_{\substack{p > D_0 \\ p \equiv 1 \pmod{4}}} \left(1 + \frac{1}{p^2 - 1}\right) = 1 + O(D_0^{-1}),$$

we arrive at a final estimate of

$$Y_{N,W} = (1 + O(D_0^{-1})) \left[ \frac{K_1(1)G_1(W)}{\Gamma(3/2)\sqrt{L(1, \chi_4)}} \right]^2 \log v + O((\log D_0)^{3/2} \exp(D_0^{1/4} \log \log D_0) (\log v)^{1/2}).$$

(iii) For  $Z_{N,W}^{(1)}$  recall from (5.4) the definition

$$Z_{N,W}^{(1)} = \sum_{\substack{a, b \leq v \\ (a, W) = (b, W) = 1 \\ p|a, b \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)\mu(b)g_4([a, b])}{g_2(a)g_2(b)[a, b]} \log \frac{v}{a} \log \frac{v}{b}. \tag{B.7}$$

We write this as

$$\sum_{\substack{a \leq v \\ (a, W) = 1 \\ p|a \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)g_4(a)}{g_2(a)a} \log \frac{v}{a} \sum_{d|a} \frac{d}{g_4(d)} \sum_{\substack{b \leq v \\ (b, W) = 1 \\ (b, a) = d \\ p|b \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(b)g_4(b)}{g_2(b)b} \log \frac{v}{b}.$$

Now substitute  $b = md$ . The inner sums can be rewritten as

$$\sum_{d|a} \frac{\mu(d)}{g_2(d)} \sum_{\substack{m \leq v/d \\ (m, aW) = 1 \\ p|m \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(m)g_4(m)}{g_2(m)m} \log \frac{v}{md},$$

where we have used the fact  $a$  is square-free. To handle the inner sum we consider the generating series

$$h_3(aW, s) = \sum_{\substack{n=1 \\ (n, aW) = 1 \\ p|n \Rightarrow p \equiv 1 \pmod{4}}}^{\infty} \frac{\mu(n)g_4(n)}{n^s g_2(n)} = \frac{K_3(s)K_4(s)G_3(aW, s)G_4(aW, s)}{\zeta(s)L(s, \chi_4)},$$

where

$$\begin{aligned}
 K_3(s) &= \left(1 - \frac{1}{2^s}\right)^{-1} \prod_{p \equiv 3 \pmod{4}} \left(1 - \frac{1}{p^{2s}}\right)^{-1} \prod_{p \equiv 1 \pmod{4}} \left(1 - \frac{1}{(p^s - 1)^2}\right), \\
 K_4(s) &= \prod_{p \equiv 1 \pmod{4}} \left(1 + \frac{5p - 3}{(p + 1)(2p - 1)(p^s - 2)}\right), \\
 G_3(aW, s) &= \prod_{\substack{p|aW \\ p \equiv 1 \pmod{4}}} \left(1 - \frac{2}{p^s}\right)^{-1}, \\
 G_4(aW, s) &= \prod_{\substack{p|aW \\ p \equiv 1 \pmod{4}}} \left(1 + \frac{5p - 3}{(p + 1)(2p - 1)(p^s - 2)}\right)^{-1}.
 \end{aligned}$$

These four functions are analytic in the region  $\sigma > 3/4$  (say), where they all satisfy the bound  $O(1)$  except for  $G_3$ . Since  $(a, W) = 1$ , by multiplicativity we can write  $G_3(aW, s) = G_3(a, s)G_3(W, s)$ . As above, we can bound  $|G_3(W, s)| \ll \exp(D_0^{1/4} \log \log D_0)$  in this region. On the other hand, note that

$$1 \leq \left(1 - \frac{2}{p^{3/4}}\right)^{-1} \leq 2$$

whenever  $p \geq 7$ . Since we are assuming  $a$  is square-free, it follows that we can bound  $|G_3(a, s)| \ll \tau(a)$ .

By similar arguments to part (i) we can apply Lemma B.1 taking  $z = -1$ ,  $M = \tau(a) \exp(D_0^{1/4} \log \log D_0)$  and suitable  $\delta, c_0$ . Note that for  $N \geq 1$  the terms  $\mu_k(z)(\log x)^{z+1-k} / \Gamma(z + 2 - k)$  are 0 for  $1 \leq k \leq N$ . Hence we can choose  $N = \lfloor (\log x) / ec_1 \rfloor$  (for some suitable  $c_1 > 0$ ) to balance the error terms, yielding a stronger error term of the form  $O(Me^{-c_1 \sqrt{\log x}})$ . Using this choice of  $N$  we obtain

$$\sum_{\substack{m \leq v/d \\ (m, aW) = 1 \\ p|m \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(m)g_4(m)}{g_2(m)m} \log \frac{v}{md} = \frac{K_3(1)K_4(1)G_3(aW)G_4(aW)}{L(1, \chi_4)} + O\left(Me^{-c_1 \sqrt{\log v/d}}\right).$$

This error term contributes

$$\ll \exp(D_0^{1/4} \log \log D_0) \sum_{\substack{a \leq v \\ p|a \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu^2(a)g_4(a)\tau(a)}{g_2(a)a} \log \frac{v}{a} \sum_{d|a} \frac{\mu^2(d)}{g_2(d)} e^{-c_1 \sqrt{\log v/d}}.$$

Swapping sums yields

$$\sum_{\substack{d \leq v \\ p|d \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu^2(d)g_4(d)\tau(d)}{g_2(d)^2 d} e^{-c_1 \sqrt{\log v/d}} \sum_{\substack{m \leq v/d \\ (m, d) = 1 \\ p|m \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu^2(m)g_4(m)\tau(m)}{g_2(m)m} \log \frac{v}{md}$$

The inner sum can be bounded by

$$\begin{aligned} &\ll \log \frac{v}{d} \prod_{\substack{p \leq v/d \\ p \equiv 1 \pmod{4}}} \left( 1 + \frac{2(4p^2 - 3p + 1)}{p(p + 1)(2p - 1)} \right) \ll \log \frac{v}{d} \prod_{\substack{p \leq v/d \\ p \equiv 1 \pmod{4}}} \left( 1 + \frac{4}{p} \right), \\ &\ll \log \frac{v}{d} \prod_{\substack{p \leq v/d \\ p \equiv 1 \pmod{4}}} \left( 1 + \frac{1}{p} \right)^4 \ll \left( \log \frac{v}{d} \right)^3 \ll e^{c_2 \sqrt{\log v/d}} \end{aligned}$$

for some suitably small  $c_2 > 0$ . Note this final bound is valid for all  $d \leq v$ . Using the fact  $g_4(d)/g_2(d)^2 \leq 1$ , we see the total error is

$$\ll \exp(D_0^{1/4} \log \log D_0) \sum_{\substack{d \leq v \\ p|d \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu^2(d)\tau(d)}{d} e^{-c_3 \sqrt{\log v/d}} \tag{B.8}$$

for some  $0 < c_3 < c_1$ . For this sum, we use the generating series, for  $\sigma > 1$

$$\sum_{\substack{n=1 \\ p|n \Rightarrow p \equiv 1 \pmod{4}}}^{\infty} \frac{\mu^2(n)\tau(n)}{n^\sigma} = \prod_{p \equiv 1 \pmod{4}} \left( 1 + \frac{2}{p^\sigma} \right) = \frac{\zeta(\sigma)L(\sigma, \chi_4)K(\sigma)}{\zeta(2\sigma)L(2\sigma, \chi_4)},$$

where

$$K(\sigma) = \left( 1 + \frac{1}{2^\sigma} \right)^{-1} \prod_{p \equiv 3 \pmod{4}} \left( 1 + \frac{1}{p^{2\sigma}} \right)^{-1} \prod_{p \equiv 1 \pmod{4}} \left( 1 + \frac{1}{p^\sigma(p^\sigma + 2)} \right).$$

By the second part of Lemma B.1, taking  $z = 1$  and  $N = \lfloor (\log x)/ec_4 \rfloor$  for some suitable  $c_4 > 0$  we get

$$\sum_{\substack{d \leq v \\ p|d \Rightarrow p \equiv 1 \pmod{4}}} \mu^2(d)\tau(d) = \frac{L(1, \chi_4)K(1)}{\zeta(2)L(2, \chi_4)} v + O(v e^{-c_4 \sqrt{\log v}}).$$

Now by splitting the sum at  $v^{1-\epsilon}$  and using partial summation one can show that

$$\sum_{\substack{d \leq v \\ p|d \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu^2(d)\tau(d)}{d} e^{-c_3 \sqrt{\log v/d}} \ll 1 + (\epsilon + e^{-c_3 \sqrt{\epsilon \log v}}) \log v.$$

Hence, taking  $\epsilon = \frac{(\log \log v)^3}{\log v}$ , we see the error (B.8) is  $O(\exp(D_0^{1/4} \log \log D_0)(\log \log v)^3)$ , which is small. Let  $\gamma_1(a) = \sum_{d|a} \mu(d)/g_2(d)$ . We obtain a main term

$$\frac{K_3(1)K_4(1)G_3(W)G_4(W)}{L(1, \chi_4)} \sum_{\substack{a \leq v \\ (a, W)=1 \\ p|a \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)g_4(a)\gamma_1(a)G_3(a)G_4(a)}{g_2(a)a} \log \frac{v}{a}.$$

For this sum we consider the generating series, for  $\sigma > 1$

$$h_4(W, s) = \sum_{\substack{n=1 \\ (n, W)=1 \\ p|n \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(n)g_4(n)\gamma_1(n)G_3(n)G_4(n)}{n^s g_2(n)} = h_1(W, s)K_5(s)G_5(W, s).$$

where

$$K_5(s) = \prod_{p \equiv 1 \pmod{4}} \left( 1 - \frac{(p-1)^2}{(p^s-1)(2p-1)(2p^2-p+1)} \right),$$

$$G_5(W, s) = \prod_{\substack{p|W \\ p \equiv 1 \pmod{4}}} \left( 1 - \frac{(p-1)^2}{(p^s-1)(2p-1)(2p^2-p+1)} \right)^{-1}.$$

Applying Lemma B.1 with  $N = 0$  gives

$$Z_{N,W}^{(1)} = C_W (\log v)^{\frac{1}{2}} + O(\exp(D_0^{1/4} \log \log D_0)(\log \log v)^3),$$

where

$$C_W = \frac{K_1(1)K_3(1)K_4(1)K_5(1)G_1(W)G_3(W)G_4(W)G_5(W)}{\Gamma(3/2)L(1, \chi_4)^{\frac{3}{2}}} \tag{B.9}$$

which reduces to the stated result (note that

$$K_3(1)G_3(W) = \frac{A^2}{g_1(W_1)^2} \prod_{\substack{p > D_0 \\ p \equiv 1 \pmod{4}}} \left( 1 - \frac{1}{(p-1)^2} \right) = \frac{A^2}{g_1(W_1)^2} (1 + O(D_0^{-1}))$$

to get the form stated). We note for future reference that  $|C_W| \asymp (\log D_0)^{3/2}$ .

(iv) For  $Z_{N,W}^{(2)}$  recall from (5.5) the definition

$$Z_{N,W}^{(2)} = \sum_{\substack{a, b \leq v \\ (a, W)=(b, W)=1 \\ p|a, b \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)\mu(b)g_4([a, b])}{g_2(a)g_2(b)[a, b]} \log \frac{v}{a} \log \frac{v}{b} \sum_{p|[a, b]} g_6(p). \tag{B.10}$$

We can write

$$\begin{aligned}
 Z_{N,W}^{(2)} &= \sum_{\substack{D_0 < q \leq v \\ q \equiv 1 \pmod{4}}} g_6(q) \sum_{\substack{a,b \leq v \\ (a,W)=(b,W)=1 \\ q|[a,b] \\ p|a,b \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)\mu(b)g_4([a,b])}{g_2(a)g_2(b)[a,b]} \log \frac{v}{a} \log \frac{v}{b} \\
 &= \sum_{\substack{D_0 < q \leq v \\ q \equiv 1 \pmod{4}}} g_6(q) \left[ T_1 + T_2 - T_3 \right].
 \end{aligned}$$

where  $q$  is prime, and

$$\begin{aligned}
 T_1 &= \sum_{\substack{a,b \leq v \\ (a,W)=(b,W)=1 \\ q|a \\ p|a,b \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)\mu(b)g_4([a,b])}{g_2(a)g_2(b)[a,b]} \log \frac{v}{a} \log \frac{v}{b}, \\
 T_2 &= \sum_{\substack{a,b \leq v \\ (a,W)=(b,W)=1 \\ q|b \\ p|a,b \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)\mu(b)g_4([a,b])}{g_2(a)g_2(b)[a,b]} \log \frac{v}{a} \log \frac{v}{b}, \\
 T_3 &= \sum_{\substack{a,b \leq v \\ (a,W)=(b,W)=1 \\ q|a,b \\ p|a,b \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a)\mu(b)g_4([a,b])}{g_2(a)g_2(b)[a,b]} \log \frac{v}{a} \log \frac{v}{b}.
 \end{aligned}$$

$T_1$  can be evaluated similarly to part (iii) to give

$$T_1 = \frac{C_W \beta_1(q)}{q} (\log v/q)^{\frac{1}{2}} + O\left(\frac{\exp(D_0^{1/4} \log \log D_0) (\log \log v)^3}{q}\right),$$

where

$$\beta_1(q) = \frac{\mu(q)g_4(q)\gamma_1(q)G_1(q)G_3(q)G_4(q)G_5(q)}{g_2(q)} = -\frac{q(4q^2 - 3q + 1)}{2(q - 1)(2q^2 - 2q + 1)}.$$

$T_2$  can be evaluated similarly. For  $T_3$ , write  $a = a'q, b = b'q$  then  $[a, b] = [qa', qb'] = q[a', b']$  so that

$$T_3 = \frac{\mu^2(q)g_4(q)}{g_2(q)^2q} \sum_{\substack{a',b' \leq v/q \\ (a',qW)=(b',qW)=1 \\ p|a',b' \Rightarrow p \equiv 1 \pmod{4}}} \frac{\mu(a')\mu(b')g_4([a', b'])}{g_2(a')g_2(b')[a', b']} \log \frac{v}{a'} \log \frac{v}{b'}.$$

This is the same form as  $Z_{N,W}^{(1)}$ . By the same considerations as above we obtain

$$T_3 = \frac{C_W \beta_2(q)}{q} (\log v/q)^{\frac{1}{2}} + O\left(\frac{\exp(D_0^{1/4} \log \log D_0) (\log \log v)^3}{q}\right),$$

where

$$\beta_2(q) = \frac{\mu^2(q) g_4(q) G_1(q) G_3(q) G_4(q) G_5(q)}{g_2(q)^2} = \frac{q^2(4q^2 - 3q + 1)}{2(q - 1)^2(2q^2 - 2q + 1)}.$$

Thus we obtain

$$Z_{N,W}^{(2)} = C_W \sum_{\substack{D_0 < q \leq v \\ q \equiv 1 \pmod{4}}} \frac{2\beta_1(q) - \beta_2(q)}{q} g_6(q) (\log(v/q))^{\frac{1}{2}} + O\left(\exp(D_0^{1/4} \log \log D_0) (\log \log v)^3 \sum_{\substack{D_0 < q \leq v \\ q \equiv 1 \pmod{4}}} \frac{g_6(q)}{q}\right),$$

where  $C_W$  is defined as in (B.9). Recalling the definition of  $g_6(q)$ , we see the error term contributes

$$\ll \exp(D_0^{1/4} \log \log D_0) (\log \log v)^3 \log v.$$

Now note that

$$\frac{2\beta_1(q) - \beta_2(q)}{q} g_6(q) = -\frac{(3q - 2)(2q + 1) \log q}{2(q + 1)(2q^2 - 2q + 1)} = -\frac{3 \log q}{2q} + O\left(\frac{\log q}{q^2}\right)$$

This error contributes

$$\ll C_W \sum_{\substack{D_0 < q \leq v \\ q \equiv 1 \pmod{4}}} \frac{\log q}{q^2} (\log v)^{\frac{1}{2}} \ll C_W (\log v)^{\frac{1}{2}}$$

which is small. We are left with a main term

$$-\frac{3C_W}{2} \sum_{\substack{D_0 < q \leq v \\ q \equiv 1 \pmod{4}}} \frac{\log q}{q} \left(\log \frac{v}{q}\right)^{\frac{1}{2}}.$$

By partial summation one can show

$$\sum_{\substack{D_0 < q \leq v \\ q \equiv 1 \pmod{4}}} \frac{\log q}{q} \left(\log \frac{v}{q}\right)^{\frac{1}{2}} = \frac{1}{3} (\log v)^{\frac{3}{2}} + O((\log v)^{\frac{1}{2}} \log D_0),$$



so that

$$Z_{N,W}^{(2)} = -\frac{C_W}{2}(\log v)^{\frac{3}{2}} + O(\exp(D_0^{1/4} \log \log D_0)(\log \log v)^3 \log v).$$

This simplifies to the stated result.  $\square$

## References

1. Banks, W.D., Freiberg, T., Maynard, J.: On limit points of the sequence of normalised prime gaps. *Proc. Lond. Math. Soc.* **113**(4), 515–539 (2016)
2. Bui, H.M., Heath-Brown, D.R.: A note on the fourth moment of Dirichlet  $L$ -functions. *Acta Arith.* **141**, 335–344 (2010)
3. Davenport, H.: *Multiplicative Number Theory*, Springer GTM **74**
4. Goldston, D.A., Pintz, J., Yıldırım, C.Y.: Primes in tuples I. *Ann. Math.* **170**(2), 819–862 (2009)
5. Goldston, D.A., Graham, S.W., Pintz, J., Yıldırım, C.Y.: Small gaps between products of two primes. *Proc. Lond. Math. Soc.* **98**(3), 741–774 (2009)
6. Granville, A.: Primes in intervals of bounded length. *Bull. Am. Math. Soc.* **52**(2), 171–222 (2015)
7. Hooley, C.: On the intervals between numbers that are sums of two squares. *Acta Math.* **127**, 279–297 (1971)
8. Jakobson, D.: Quantum limits on flat tori. *Ann. Math.* **145**(2), 235–266 (1997)
9. Maynard, J.: Small gaps between primes. *Ann. Math.* **181**(1), 383–413 (2015)
10. Plaksin, V.A.: The distribution of numbers representable as the sum of two squares. *Math. USSR Izv.* **31**(1), 171–191 (1988)
11. Polymath, D.H.J.: Variants of the Selberg sieve, and bounded intervals containing many primes. *Res. Math. Sci.* **1**(12), 83 (2014)
12. Tenenbaum, G.: Introduction to analytic and probabilistic number theory. *Am. Math. Soc. GMT* **163**
13. Tolev, D.I.: On the remainder term in the circle problem in an arithmetic progression. *Proc. Steklov Inst. Math.* **276**, 261–274 (2012)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.