

ON THE VARIANCE OF SQUAREFREE INTEGERS IN SHORT INTERVALS AND ARITHMETIC PROGRESSIONS

OFIR GORODETSKY, KAISA MATOMÄKI, MAKSYM RADZIWIŁŁ AND
BRAD RODGERS



Abstract. We evaluate asymptotically the variance of the number of squarefree integers up to x in short intervals of length $H < x^{6/11-\varepsilon}$ and the variance of the number of squarefree integers up to x in arithmetic progressions modulo q with $q > x^{5/11+\varepsilon}$. On the assumption of respectively the Lindelöf Hypothesis and the Generalized Lindelöf Hypothesis we show that these ranges can be improved to respectively $H < x^{2/3-\varepsilon}$ and $q > x^{1/3+\varepsilon}$. Furthermore we show that obtaining a bound sharp up to factors of H^ε in the full range $H < x^{1-\varepsilon}$ is equivalent to the Riemann Hypothesis. These results improve on a result of Hall (Mathematika 29(1):7–17, 1982) for short intervals, and earlier results of Warlimont, Vaughan, Blomer, Nunes and Le Boudec in the case of arithmetic progressions.

1 Introduction

1.1 Main results. An integer $n \geq 1$ is squarefree if it is not divisible by the square of a prime. By analogy with questions about prime numbers, a basic problem in analytic number theory is to understand the distribution of squarefree numbers in arithmetic progressions and in short intervals. Squarefree numbers ought to be a simpler, more regular sequence than primes, and yet they present distinct challenges; for instance we can determine whether n is prime in polynomial time [AKS04], but there is no known polynomial time algorithm to determine whether n is squarefree.

It was conjectured by Montgomery (see [Cro75]) that for any given $\varepsilon \in (0, 1/100)$, and $(a, q) = 1$,

$$\sum_{\substack{n \leq x \\ n \equiv a \pmod{q}}} \mu^2(n) = \frac{6}{\pi^2} \cdot \frac{x}{q} \prod_{p|q} \left(1 - \frac{1}{p^2}\right) + O_\varepsilon\left((x/q)^{1/4+\varepsilon}\right). \quad (1)$$

uniformly in $1 \leq q \leq x^{1-\varepsilon}$. This conjecture is difficult for two reasons. In the regime of large q of size roughly $x^{1-\varepsilon}$ the left-hand side contains only x^ε terms and even establishing an asymptotic is open¹ (see the work of Nunes [Num17] for the best result in this direction). In the regime of small q of size about x^ε establishing an

¹ In fact establishing that the left-hand side is positive for $q = x^{1-\varepsilon}$ is open!

asymptotic is easy but obtaining an error term as good as $O_\varepsilon((x/q)^{1/4+\varepsilon})$ is an open problem, even conditionally on the Generalized Riemann Hypothesis.

Analogously we conjecture that for any given $\varepsilon \in (0, \frac{1}{100})$, uniformly in $x^\varepsilon \leq H \leq x$,

$$\sum_{x < n \leq x+H} \mu^2(n) = \frac{6H}{\pi^2} + O_\varepsilon(H^{1/4+\varepsilon}). \quad (2)$$

Similarly to the case of arithmetic progressions, when H is close to x^ε no asymptotic estimates are known (see the work of Tolev [Tol06] and Filaseta and Trifonov [FT92] for the best unconditional results in this direction and [CE19, Thm. A.1], [Gra98] for results conditional on the ABC conjecture). Meanwhile for large H , say $H = x$, estimating (2) asymptotically is straightforward, but obtaining an error term $O_\varepsilon(x^{1/4+\varepsilon})$ is an open problem, even conditionally on the Riemann Hypothesis (see [Liu16] for the best result in this direction).

An important feature of both conjectures (1) and (2) is that the error term is significantly smaller than the square-root of the number of terms being summed, in contrast to what a naive probabilistic model predicts.

The conjectures (1) and (2) imply the Riemann Hypothesis, and they are almost certainly deeper than the Riemann Hypothesis. Nonetheless one can still hope to investigate them on average over residue classes for (1) or on average over short intervals for (2). Importantly, establishing (1) on average is easier when q is large than when q is small, since a large q allows for more averaging over the residue classes $a \pmod{q}$. Similarly establishing (2) on average is easier when H is small, since there are more non-overlapping short intervals $[x, x+H]$ to average over compared to the case when H is large. In fact when there is little averaging (i.e q small or H large), the averaged versions of (1) and (2) are not significantly easier than the non-averaged version, see Theorem 3 for a concrete manifestation of this.

In our first result we compute the variance of (2) on average over short intervals. We estimate the variance asymptotically thus making (on average) the error term in (2) more precise.

Theorem 1. *Let $\varepsilon \in (0, \frac{1}{100})$ be given. Let $X \geq 1$ and $1 \leq H \leq X^{6/11-\varepsilon}$. Then*

$$\frac{1}{X} \int_X^{2X} \left| \sum_{x < m \leq x+H} \mu^2(m) - \frac{6H}{\pi^2} \right|^2 dx = C\sqrt{H} + O_\varepsilon(H^{1/2-\varepsilon/16}) \quad (3)$$

with

$$C := \frac{\zeta(3/2)}{\pi} \prod_p \left(1 - \frac{3}{p^2} + \frac{2}{p^3} \right). \quad (4)$$

Assuming the Lindelöf Hypothesis (3) holds in the wider range $H \leq X^{2/3-\varepsilon}$.

We recall that the Lindelöf Hypothesis follows from the Riemann Hypothesis and asserts that for any given $\varepsilon > 0$ we have $|\zeta(\frac{1}{2} + it)| \ll_\varepsilon 1 + |t|^\varepsilon$ for all $t \in \mathbb{R}$. In Theorem 3 we will show that if we had (3) in the full range $H \leq X^{1-\varepsilon}$ then the Riemann Hypothesis would ensue.

Theorem 1 extends a theorem of Hall [Hal82] who showed that the asymptotic formula (3) holds in the range $H \leq X^{2/9-\varepsilon}$. We will now explain why the range $H = X^{1/2}$ can be considered a threshold in this problem. It is reasonable to conjecture that given $\varepsilon > 0$, for any $1 \leq h \leq x^{1-\varepsilon}$,

$$\sum_{n \leq x} \mu^2(n)\mu^2(n+h) - C(h)x = O_\varepsilon(x^{1/4+\varepsilon}) \quad (5)$$

with $C(h)$ a constant depending only on h . Summing this conjectural estimate over h recovers Theorem 1 but only in the range $H < X^{1/2-\varepsilon}$. Thus Theorem 1 exploits (unconditionally!) additional cancellations between the error terms in (5).²

We now describe the analogue of Theorem 1 for the distribution of squarefree numbers in arithmetic progressions with a given modulus. In this case for a given modulus q the parameter x/q has the same role as the length H of the short interval in Theorem 1. While the results are analogous they are harder to prove, as is often the case with q -analogues.

Theorem 2. *Let $\varepsilon \in (0, \frac{1}{100})$ be given. Let $x \geq q \geq x^{5/11+\varepsilon}$ be a prime. Then*

$$\begin{aligned} & \frac{1}{\varphi(q)} \sum_{(a,q)=1} \left| \sum_{\substack{m \leq x \\ m \equiv a \pmod{q}}} \mu^2(m) - \frac{6}{\pi^2} \cdot \frac{x}{q} \prod_{p|q} \left(1 - \frac{1}{p^2}\right)^{-1} \right|^2 \\ &= C \prod_{p|q} \left(1 + \frac{2}{p}\right)^{-1} \cdot \sqrt{\frac{x}{q}} + O_\varepsilon((x/q)^{1/2-\varepsilon/16}), \end{aligned} \quad (6)$$

where C is the same constant as in Theorem 1. Assuming the Generalized Lindelöf Hypothesis the claim holds in the wider range $q > x^{1/3+30\varepsilon}$.

We recall that the Generalized Lindelöf Hypothesis follows from the Generalized Riemann Hypothesis and asserts that for any given $\varepsilon > 0$ we have $|L(\frac{1}{2} + it, \chi)| \ll_\varepsilon 1 + (|q| + |t|)^\varepsilon$ for all $t \in \mathbb{R}$ and all characters $\chi \pmod{q}$.

For simplicity we have assumed in Theorem 2 that q is prime, but our methods are amenable to handling the general case of composite q with a bit more effort.

Once extended to composite q our Theorem 2 improves on results by Warlimont [War80] and Vaughan [Vau05] who obtain an asymptotic formula with an additional averaging over $q \leq Q$ in the range $x^{2/3} \leq Q = o(x)$. Moreover, for prime values of q , Theorem 2 improves on a succession of results by Blomer [Blo08], Nunes [Nun15] (see also [Par19]) and Le Boudec [LB18] who considered individual averages over $(a, q) = 1$ as we do in Theorem 2. In particular Nunes showed that (6) holds in the

² Using estimates of Tsang [Tsa85] for (5) recovers Theorem 1 in the range $H \leq X^{8/33-\varepsilon}$.

range $x^{31/41+\varepsilon} \leq q = o(x)$ and Le Boudec showed that the left-hand side of (6) is $O_\varepsilon((x/q)^{1/2+\varepsilon})$ for all $\varepsilon > 0$ in the range $x^{1/2} \leq q \leq x$.

Keating and Rudnick [KR16] obtained Theorems 1 and 2 in the context of function fields in the limit of a large field size. Their results hold in the (analogues of) the ranges $X^\varepsilon \leq H \leq X^{1-\varepsilon}$ and $x^\varepsilon \leq q \leq x^{1-\varepsilon}$. Our proofs of Theorem 1 and Theorem 2 can be adapted in the setting of a fixed base field and large degree limit. In fact our proofs of Theorems 1 and 2 were originally motivated by analogies with the function field setting. Since we ended up obtaining equally strong results in the setting of number fields we do not include the proofs in the function field setting.

Finally the next result shows that obtaining nearly optimal upper bounds for (3) in a complete range is equivalent to the Riemann Hypothesis.

Theorem 3. *The Riemann Hypothesis holds if and only if for every $\varepsilon \in (0, \frac{1}{100})$ and every $1 \leq H \leq X^{1-\varepsilon}$,*

$$\frac{1}{X} \int_X^{2X} \left| \sum_{x < m \leq x+H} \mu^2(m) - \frac{6H}{\pi^2} \right|^2 dx \ll_{\varepsilon, \delta} H^{1/2+\delta} \quad (7)$$

for every $\delta > 0$.

Following the proof of Theorem 3 one can show that for any smooth compactly supported Φ , conditionally on the Generalized Riemann Hypothesis

$$\frac{1}{\varphi(q)} \sum_{(a,q)=1} \left| \sum_{m \equiv a \pmod{q}} \mu^2(n) \Phi\left(\frac{n}{x}\right) - \frac{6}{\pi^2 \varphi(q)} \sum_{(m,q)=1} \Phi\left(\frac{m}{x}\right) \right|^2 \ll_{\varepsilon, \delta} (x/q)^{1/2+\delta} \quad (8)$$

for all $\delta > 0$ and uniformly in $1 \leq q \leq x^{1-\varepsilon}$ for any given $\varepsilon \in (0, \frac{1}{100})$. However it is not clear whether (8) implies the Generalized Riemann Hypothesis. Moreover replacing the smoothing Φ by sharp cut-offs appears to be difficult. For these reasons we decided not to pursue this further in the present paper.

Finally, we note that we have made no effort to optimize the exponents of error terms $O_\varepsilon(H^{1/2-\varepsilon/16})$ and $O_\varepsilon((x/q)^{1/2-\varepsilon/16})$ in Theorems 1 and 2. Better power saving estimates, in more restricted ranges, can be found in the papers [Hal82, Nun15].

1.2 Fractional Brownian motion. One notable feature of Theorems 1 and 2 is that while the expected count of squarefrees in a short interval (or likewise arithmetic progression) is of order H , the variance of these counts is of order $H^{1/2}$. For many other natural arithmetic sequences (e.g. primes) one conjectures that the variance of counts is of the same order of magnitude as the expected value of counts.

That the variance is of order $H^{1/2}$ in Theorems 1 and 2 speaks to the idea that the squarefree numbers are “less random” than (for example) the primes (cf. [CS13]). One may conjecture that higher moments are Gaussian (see [ACS17] for numerical

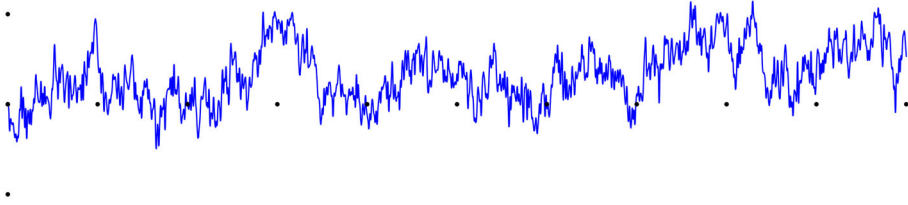


Figure 1: Partial sums of $\mu^2(n)$: depiction of (9) with $x = 2 \times 10^{15}$, $H = 44,721,359$ and $0 \leq t \leq 10$.

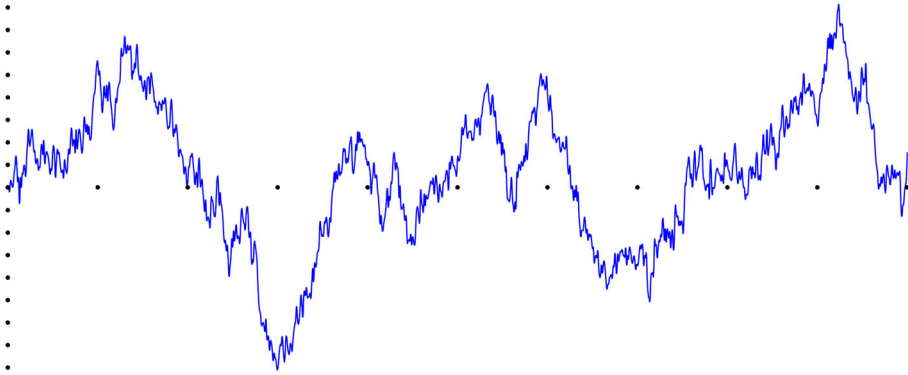


Figure 2: Partial sums of $\log p$: depiction of (10) with $x = 2 \times 10^{15}$, $H = 44,721,359$. and $0 \leq t \leq 10$.

evidence). For x drawn uniformly at random from $[X, 2X]$, one may even make the stronger conjecture that the process

$$t \mapsto \frac{1}{H^{1/4}} \sum_{x < n \leq x+tH} (\mu^2(n) - 1/\zeta(2)) \quad (9)$$

tends weakly, when suitably normalized by $H^{1/4}$, to a fractional Brownian motion with Hurst parameter $1/4$. See Fig. 1 for an illustration of the evolution of the partial sums (9). A formulation of this perspective seems to have been first made in [GH91]. This is in contrast to the analogous process generated by prime-counting, where one may conjecture the appearance of Hurst parameter $1/2$ —that is, usual Brownian motion. (See [She14] for a survey on fractional Brownian motion.) The evolution of the process

$$t \mapsto \frac{1}{H^{1/2}} \sum_{x < p \leq x+tH} (\log p - 1). \quad (10)$$

is depicted in Fig. 2. Both Figs. 1 and 2 depict the same range of parameters to make the comparison easier. The dots on Figs. 1 and 2 correspond to lattice points on the positive x -axis and on the (positive and negative) y -axis and indicate the difference in scales.

1.3 Conventions and notations. Throughout the rest of the paper we will allow the implicit constants in \ll and $O(\cdot)$ to depend on ε . Furthermore the notation $n \sim N$ in the subscript of a sum will mean that $N \leq n < 2N$.

2 Proofs of Theorems 1 and 2

We will show in this section how Theorems 1 and 2 follow from a number of technical propositions that are proven in Sects. 4–7.

The proof of Theorem 1 splits into two steps and depends on the identity

$$\mu^2(m) = \sum_{nd^2=m} \mu(d)$$

and the following two propositions.

PROPOSITION 1. *Let $\varepsilon \in (0, \frac{1}{100})$ be given. Let $X \geq 1$ and $X^\varepsilon \leq H \leq X^{2/3-\varepsilon}$. Let $H^{1+\varepsilon} \leq z \leq \min\{X/H^{1/2+\varepsilon}, H^{1/2-\varepsilon}X^{1/2}\}$. Then, as $X \rightarrow \infty$,*

$$\frac{1}{X} \int_X^{2X} \left| \sum_{\substack{x < nd^2 \leq x+H \\ d^2 \leq z}} \mu(d) - H \sum_{d^2 \leq z} \frac{\mu(d)}{d^2} \right|^2 dx = C\sqrt{H} + O(H^{1/2-\varepsilon/10}) \quad (11)$$

with C as in (4).

PROPOSITION 2. *Let $\varepsilon \in (0, \frac{1}{100})$ be given. Let $X \geq 1$ and $X^\varepsilon \leq H \leq X^{4/7-\varepsilon}$. Let $z \geq H^{4/3+\varepsilon}$. Then*

$$\frac{1}{X} \int_X^{2X} \left| \sum_{\substack{x < nd^2 \leq x+H \\ d^2 > z}} \mu(d) - H \sum_{2X \geq d^2 > z} \frac{\mu(d)}{d^2} \right|^2 dx \ll H^{1/2-\varepsilon/8}. \quad (12)$$

Assuming the Lindelöf Hypothesis, the claim holds in the wider range $X^\varepsilon \leq H \leq X^{2/3-\varepsilon}$ and $z \geq H^{1+\varepsilon}$.

Under the assumption of the Lindelöf Hypothesis, the above propositions cover all the possible values of d^2 for $X^\varepsilon \leq H \leq X^{2/3-\varepsilon}$. However, unconditionally they cover all the possible values of d^2 only for $X^\varepsilon \leq H \leq X^{6/(11+12\varepsilon)}$. It would be possible to improve on the exponent $4/7$ in Proposition 2, but this would not help. Similarly it should be possible to prove Proposition 1 only with the condition $H^{1+\varepsilon} \leq z \leq X/H^{1/2+\varepsilon}$ by adapting the proof of Proposition 3 below.

We note that only the terms d with $d^2 \in [H^{1-\varepsilon}, H^{1+\varepsilon}]$ contribute to the main term $C\sqrt{H}$ in Proposition 1.

Roughly speaking Proposition 1 depends only on “convex” inputs such as a Fourier expansion and a point-counting lemma, whereas Proposition 2 exploits large value estimates of Huxley and subconvexity and fourth moment estimates for the Riemann zeta-function.

Proof of Theorem 1 assuming Propositions 1 and 2. Let $\varepsilon \in (0, \frac{1}{100})$. If $H \leq X^\varepsilon$ then the result already follows from Hall's theorem. We can therefore assume that $H > X^\varepsilon$.

For $H \in [X^\varepsilon, X^{6/11-\varepsilon}]$, take $z = \min\{X/H^{1/2+\varepsilon}, H^{1/2-\varepsilon}X^{1/2}\}$. Note that $z \geq H^{4/3+\varepsilon}$. Denoting by \mathcal{I}_1 the left-hand side of (11) and by \mathcal{I}_2 the left-hand side of (12), we get, using Cauchy–Schwarz, that

$$\frac{1}{X} \int_X^{2X} \left| \sum_{x < nd^2 \leq x+H} \mu(d) - H \sum_{d^2 \leq 2X} \frac{\mu(d)}{d^2} \right|^2 dx = \mathcal{I}_1 + O(\sqrt{\mathcal{I}_1 \mathcal{I}_2} + \mathcal{I}_2).$$

Using the bounds in (11) and (12), we conclude that

$$\frac{1}{X} \int_X^{2X} \left| \sum_{x < nd^2 \leq x+H} \mu(d) - H \sum_{d^2 \leq 2X} \frac{\mu(d)}{d^2} \right|^2 dx = C\sqrt{H} + O(H^{1/2-\varepsilon/16}). \quad (13)$$

Notice that the tail $H \sum_{d^2 > 2X} \mu(d)/d^2$ is $\ll H/\sqrt{X}$. Hence the claim reduces to showing that

$$\frac{1}{X} \int_X^{2X} \left| \sum_{x < nd^2 \leq x+H} \mu(d) - H \sum_{d^2 \leq 2X} \frac{\mu(d)}{d^2} \right| \cdot \frac{H}{\sqrt{X}} + \left(\frac{H}{\sqrt{X}} \right)^2 dx \ll H^{1/2-\varepsilon/16}.$$

Applying Cauchy–Schwarz and (13) we see that the left hand side is $\ll H^{1/4}(H/\sqrt{X}) + H^2/X \ll H^{1/2-\varepsilon/16}$ since $H \leq X^{2/3-\varepsilon}$. \square

Likewise the proof of Theorem 2 splits into two steps and depends on the following propositions.

PROPOSITION 3. *Let $\varepsilon \in (0, \frac{1}{100})$. Let q be prime with $x^{1/3+30\varepsilon} \leq q \leq x^{1-\varepsilon}$ and let $(x/q)^{1+\varepsilon} \leq z \leq x^{-\varepsilon} \cdot \sqrt{qx}$. Then*

$$\frac{1}{\varphi(q)} \sum_{(a,q)=1} \left| \sum_{\substack{d^2 n \leq x, d^2 < z \\ d^2 n \equiv a \pmod{q}}} \mu(d) - \frac{1}{\varphi(q)} \sum_{\substack{d^2 n \leq x, d^2 < z \\ (d^2 n, q)=1}} \mu(d) \right|^2 = C\sqrt{x/q} + O\left((x/q)^{1/2-\varepsilon/16}\right) \quad (14)$$

with C as in (4).

PROPOSITION 4. *Let $\varepsilon \in (0, \frac{1}{100})$. Let $x \geq 1$ and $x^{3/7+\varepsilon} \leq q \leq x^{1-\varepsilon}$. Let $z \geq (x/q)^{4/3+\varepsilon}$. Then*

$$\frac{1}{\varphi(q)} \sum_{(a,q)=1} \left| \sum_{\substack{d^2 n \leq x, d^2 \geq z \\ d^2 n \equiv a \pmod{q}}} \mu(d) - \frac{1}{\varphi(q)} \sum_{\substack{d^2 n \leq x, d^2 \geq z \\ (d^2 n, q)=1}} \mu(d) \right|^2 \ll (x/q)^{1/2-\varepsilon/8}. \quad (15)$$

Assuming the Generalized Lindelöf Hypothesis, the claim holds in the wider range $x^{1/3+\varepsilon} \leq q \leq x^{1-\varepsilon}$ and $z \geq (x/q)^{1+\varepsilon}$.

The proof of Proposition 3 depends once again only on “convex” inputs: in this case Poisson summation and results on integer solutions to binary quadratic forms with positive discriminant. However the proof of Proposition 3 is more intricate than that of Proposition 1 due to a number of technical issues. The proof of Proposition 4 is similar to the proof of Proposition 2 and uses hybrid versions of Huxley’s large value estimates, subconvexity estimates for $L(s, \chi)$ and a hybrid fourth moment estimate.

The deduction of Theorem 2 from the above two proposition is identical to the deduction of Theorem 1 from Propositions 1 and 2. The only difference is that we use the result of Nunes to handle the case when $q > x^{1-\varepsilon}$ and we notice that for prime q ,

$$\begin{aligned} \frac{1}{\varphi(q)} \sum_{\substack{d^2 n \leq x \\ (d^2 n, q)=1}} \mu(d) &= \frac{1}{\varphi(q)} \sum_{\substack{d^2 \leq x \\ (d, q)=1}} \mu(d) \left(\left\lfloor \frac{x}{d^2} \right\rfloor - \left\lfloor \frac{x}{qd^2} \right\rfloor \right) \\ &= \frac{x}{q} \sum_{\substack{d^2 \leq x \\ (d, q)=1}} \frac{\mu(d)}{d^2} + O\left(\frac{\sqrt{x}}{q}\right) = \frac{6}{\pi^2} \frac{x}{q} \left(1 - \frac{1}{q^2}\right)^{-1} + O\left(\frac{\sqrt{x}}{q}\right) \end{aligned}$$

and the total error term incurred is x/q^2 which is $\leq x^{-\varepsilon} \sqrt{x/q}$ for $q > x^{1/3+\varepsilon}$.

Theorem 3 depends upon similar principles as Propositions 2 and 4. We prove Theorem 3 in Sect. 8.

Finally let us make a few remarks on the bottleneck that prevents us from pushing our result further. Taking $H = X^{6/11}$, we are unable to show the following estimate,

$$\frac{1}{X} \int_X^{2X} \left| \sum_{\substack{x \leq nd^2 \leq x+H \\ d^2 \sim X^{8/11}}} \mu(d) - H \sum_{d^2 \sim X^{8/11}} \frac{\mu(d)}{d^2} \right|^2 dx \ll_A \frac{\sqrt{H}}{\log^A X}.$$

Specifically opening $\mu(d)$ using Heath-Brown’s identity (see [HB82]) the only situation that we are not able to estimate is the one in which $\mu(d)$ is replaced by two smooth sums of equal length. Roughly speaking this corresponds to estimating,

$$\frac{1}{X} \int_X^{2X} \left| \sum_{\substack{x \leq na^2b^2 \leq x+H \\ a, b \sim X^{2/11}}} 1 - H \sum_{a, b \sim X^{2/11}} \frac{1}{a^2b^2} \right|^2 dx \ll \frac{\sqrt{H}}{\log^A X}$$

Opening the above expression into Dirichlet polynomials this is roughly equivalent to

$$\int_{|t| \leq X^{5/11}} \left| \sum_{n \sim X^{3/11}} \frac{1}{n^{1/2+it}} \sum_{a \sim X^{2/11}} \frac{1}{a^{1/2+2it}} \sum_{b \sim X^{2/11}} \frac{1}{b^{1/2+2it}} \right|^2 dt \ll \frac{X^{6/11}}{\log^A X}.$$

Applying the functional equation on the Dirichlet polynomial over n , and setting $Y = X^{12/11}$ we then see that obtaining the above estimate is equivalent to showing

that,

$$\int_{|t| \leq Y^{5/12}} \left| \sum_{n \sim Y^{1/6}} \frac{1}{n^{1/2+it}} \sum_{a \sim Y^{1/6}} \frac{1}{a^{1/2+2it}} \sum_{b \sim Y^{1/6}} \frac{1}{b^{1/2+2it}} \right|^2 dt \ll \frac{Y^{1/2}}{\log^A Y}.$$

If the $2it$ in the Dirichlet polynomial over a and b were replaced by it then we would be facing exactly the same bottleneck as in the case of improving Huxley's prime number theorem in short intervals (by a variant of the computations in [HB82], see also [Har07, Chapter 7]). In particular to make further progress we either need to find a way to improve Huxley's estimate or find a way to exploit the fact that the phases in two of the Dirichlet polynomials are $2it$ and not it . Unfortunately we do not see how to make progress on either of these questions.

3 Lemmas

3.1 Dirichlet polynomials and L -functions. Let us first collect some standard results on large values of Dirichlet polynomials and L -functions.

LEMMA 1 (Large-value theorem). *Let $N, T \geq 1$ and $V > 0$. Let $F(s) = \sum_{n \leq N} a_n n^{-s}$ be a Dirichlet polynomial and let $G = \sum_{n \leq N} |a_n|^2$. Let \mathcal{T} be a set of 1-spaced points $t_r \in [-T, T]$ such that $|F(it_r)| \geq V$. Then*

$$|\mathcal{T}| \ll (GNV^{-2} + T \min\{GV^{-2}, G^3NV^{-6}\})(\log 2NT)^6.$$

Proof. This follows from the mean-value theorem and Huxley's large value theorem, see e.g. [IK04, Theorem 9.7 and Corollary 9.9]. □

We will say that a set of tuples (t, χ) with χ a Dirichlet character and t a real number is *well-spaced* whenever it holds that if $(t, \chi) \neq (u, \chi')$ then either $\chi \neq \chi'$ or $|t - u| \geq 1$.

LEMMA 2 (Hybrid large-value theorem). *Let $q \in \mathbb{N}$, $N, T \geq 1$ and $V > 0$. Let $F(s, \chi) = \sum_{n \leq N} a_n \chi(n) n^{-s}$ be a Dirichlet polynomial, and let $G = \sum_{n \leq N} |a_n|^2$. Let \mathcal{T} be a set of well-spaced tuples (t_r, χ) with $t_r \in [-T, T]$ and with χ a primitive character of modulus q such that $|F(it_r, \chi)| \geq V$. Then*

$$|\mathcal{T}| \ll (GNV^{-2} + qT \min\{GV^{-2}, G^3NV^{-6}\}) \cdot (\log 2qNT)^{18}.$$

Proof. This follows e.g. from [IK04, Theorems 9.16 and 9.18 with $k = q$ and $Q = 1$]. □

LEMMA 3 (Fourth moment estimate). *Let $T \geq 2$. Then*

$$\int_{|t| \leq T} |\zeta(\frac{1}{2} + it)|^4 dt \ll T(\log T)^4.$$

Proof. See e.g. [Tit86, formula (7.6.1)]. \square

LEMMA 4 (Hybrid fourth moment estimate). *Let $T \geq 2$ and $q \geq 2$. Then*

$$\sum_{\chi} \int_{|t| \leq T} |L(\frac{1}{2} + it, \chi)|^4 dt \ll T \varphi(q) \log(Tq)^4,$$

where the sum is over all characters modulo q .

Proof. See e.g. [Mon71, Theorem 10.1]. \square

LEMMA 5 (Subconvexity estimate). *One has, for $|t| \geq 2$,*

$$\zeta(1/2 + it) \ll |t|^{1/6} (\log |t|)^2.$$

Proof. See e.g. [IK04, formula (8.22)]. \square

LEMMA 6 (Hybrid Weyl subconvexity). *For cube-free q , primitive characters $\chi \pmod{q}$ and $|t| \geq 2$,*

$$L(1/2 + it, \chi) \ll_{\varepsilon} (qt)^{1/6 + \varepsilon}$$

for any $\varepsilon > 0$.

Proof. See [PY19, Theorem 1.1]. \square

Of course the Lindelöf and Generalized Lindelöf Hypotheses would give us respectively that for any $\varepsilon > 0$, for $|t| \geq 2$ and any character $\chi \pmod{q}$,

$$\zeta(1/2 + it) \ll |t|^{\varepsilon}, \quad L(1/2 + it, \chi) \ll (q|t|)^{\varepsilon}.$$

LEMMA 7 (Hybrid mean-value theorem). *Let $a(n)$ be an arbitrary sequence of coefficients and $N, q \geq 1$ be integers and $T \geq 1$ real. Then, for any given $\varepsilon > 0$,*

$$\sum_{\chi \pmod{q}} \int_{|t| \leq T} \left| \sum_{n \leq N} a(n) \chi^2(n) n^{it} \right|^2 dt \ll q^{\varepsilon} (qT + N) \sum_{n \leq N} |a(n)|^2.$$

Proof. We notice that given a character $\psi \pmod{q}$ there are at most $\ll q^{\varepsilon}$ characters χ such that $\chi^2 = \psi$. Therefore the left-hand side of the claim is bounded by

$$\ll q^{\varepsilon} \sum_{\psi \pmod{q}} \int_{|t| \leq T} \left| \sum_{n \leq N} a(n) \psi(n) n^{it} \right|^2 dt$$

and the result follows from the standard hybrid mean-value theorem, see e.g. [Mon71, Theorem 6.4]. \square

3.2 Asymptotic estimates.

LEMMA 8. Fix $\varepsilon \in (0, \frac{1}{100})$. Let K_0 be a positive constant. Suppose that $W: \mathbb{R} \rightarrow \mathbb{C}$ is such that, for all $k, \ell \in \{0, 1, 2, 3, 4\}$, one has

$$|W^{(k)}(y)| \leq K_0 \frac{H^{\ell\varepsilon/4}}{(1+|y|)^\ell}, \quad \text{for all } y \in \mathbb{R}. \quad (16)$$

Let $z \geq H^{1+\varepsilon}$. Then

$$2H^2 \sum_{d_1^2, d_2^2 \leq z} \frac{\mu(d_1)\mu(d_2)}{d_1^2 d_2^2} \sum_{\lambda \geq 1} \left| W\left(\frac{H\lambda}{(d_1^2, d_2^2)}\right) \right|^2 = CH^{1/2} \pi \int_0^\infty |W(y)|^2 \sqrt{y} dy + O(H^{1/2-\varepsilon/5}), \quad (17)$$

where C is as in (4) and where the implied constant depends only on K_0 and ε .

Proof. The proof consists of two steps.

The first step is to show that the sum in (17) can be completed into a sum over all d_1, d_2 without affecting the claimed asymptotic. We use the information (16) for $k=0$ and $\ell=0, 1$ to see that, for any $\nu > 0$,

$$\sum_{\lambda \geq 1} |W(\lambda/\nu)|^2 \ll \sum_{1 \leq \lambda \leq \nu H^{\varepsilon/4}} 1 + \sum_{\lambda > \nu H^{\varepsilon/4}} \frac{H^{\varepsilon/2} \nu^2}{\lambda^2} \ll \nu H^{\varepsilon/4}. \quad (18)$$

Hence

$$2H^2 \sum_{\substack{d_1 > z^{1/2} \\ \text{or} \\ d_2 > z^{1/2}}} \frac{\mu(d_1)\mu(d_2)}{d_1^2 d_2^2} \sum_{\lambda \geq 1} \left| W\left(\frac{H\lambda}{(d_1^2, d_2^2)}\right) \right|^2 \ll H^{1+\varepsilon/4} \sum_{\substack{d_1 > z^{1/2} \\ \text{or} \\ d_2 > z^{1/2}}} \frac{(d_1, d_2)^2}{d_1^2 d_2^2}.$$

Writing $(d_1, d_2) = d_0$ and $d_i = \delta_i d_0$ and utilizing symmetry and the lower bound for z , we see that this is

$$\ll H^{1+\varepsilon/4} \sum_{d_0 \geq 1} \frac{1}{d_0^2} \sum_{\substack{\delta_1 \geq H^{(1+\varepsilon)/2}/d_0 \\ \delta_2 \geq 1}} \frac{1}{\delta_1^2 \delta_2^2} \ll H^{1+\varepsilon/4} \sum_{d_0 \geq 1} \frac{1}{d_0^2} \min\left\{1, \frac{d_0}{H^{(1+\varepsilon)/2}}\right\} \ll H^{1/2-\varepsilon/5}.$$

Thus the left-hand side of (17) is

$$2H^2 \sum_{d_1, d_2 \geq 1} \frac{\mu(d_1)\mu(d_2)}{d_1^2 d_2^2} \sum_{\lambda \geq 1} \left| W\left(\frac{H\lambda}{(d_1^2, d_2^2)}\right) \right|^2 + O(H^{1/2-\varepsilon/5}). \quad (19)$$

The second step is to use contour integration to simplify (19). Define $g(x) = |W(e^x)|^2 e^x$, which is smooth and decays exponentially as $|x| \rightarrow \infty$. Now

$$\hat{g}(\xi) = \int_{-\infty}^{\infty} |W(e^x)|^2 e^x e(-x\xi) dx = \int_0^\infty |W(y)|^2 y^{-2\pi i \xi} dy,$$

and standard partial integration arguments show that (i) $\hat{g}(\xi)$ is entire and (ii) $\hat{g}(\xi) = O(H^{2\varepsilon}/(|\xi| + 1)^3)$ uniformly for $|\Im(\xi)| < 1/(2\pi)$. Fourier inversion implies, for $r > 0$,

$$|W(r)|^2 = r^{-1} \frac{1}{2\pi i} \int_{(c)} r^s \hat{g}\left(\frac{s}{2\pi i}\right) ds,$$

where the integral is over $\Re(s) = c$, and $-1 < c < 1$.

Hence, taking $c = -1/4$,

$$\begin{aligned} & 2H^2 \sum_{d_1, d_2 \geq 1} \frac{\mu(d_1)\mu(d_2)}{d_1^2 d_2^2} \sum_{\lambda \geq 1} \left| W\left(\frac{H\lambda}{(d_1^2, d_2^2)}\right) \right|^2 \\ &= \frac{H}{i\pi} \sum_{d_1, d_2} \frac{\mu(d_1)\mu(d_2)}{d_1^2 d_2^2} \sum_{\lambda \geq 1} \frac{(d_1, d_2)^2}{\lambda} \int_{(-1/4)} H^s \lambda^s (d_1, d_2)^{-2s} \hat{g}\left(\frac{s}{2\pi i}\right) ds. \end{aligned}$$

The range of s is such that the sums over both λ and d_1, d_2 can be taken inside the integral, and the above simplifies to

$$\begin{aligned} & \frac{H}{i\pi} \int_{(-1/4)} H^s \zeta(1-s) \prod_p \left(1 - \frac{2}{p^2} + \frac{1}{p^{2+2s}}\right) \hat{g}\left(\frac{s}{2\pi i}\right) ds \\ &= \frac{H}{i\pi} \int_{(-1/4)} H^s \zeta(1-s) \zeta(2+2s) \prod_p \left(1 - \frac{2}{p^2} + \frac{2}{p^{4+2s}} - \frac{1}{p^{4+4s}}\right) \hat{g}\left(\frac{s}{2\pi i}\right) ds. \end{aligned}$$

The Euler product in the last line converges absolutely for $\Re s > -3/4$. Therefore using Lemma 5 (noting that the same bound holds also for $\zeta(c + it)$ with $c \geq 1/2$) and bounds on \hat{g} we can push the contour integral above to the left to an integral over $\Re(s) = -3/4 + \varepsilon$. Because of the singularity from $\zeta(2 + 2s)$ at $s = -1/2$ the above then simplifies to

$$\begin{aligned} & H^{1/2} \zeta(3/2) \prod_p \left(1 - \frac{3}{p^2} + \frac{2}{p^3}\right) \hat{g}\left(-\frac{1}{4\pi i}\right) + O(H^{1/4+3\varepsilon}) \\ &= H^{1/2} \int_0^\infty |W(y)|^2 \sqrt{y} dy \zeta(3/2) \prod_p \left(1 - \frac{3}{p^2} + \frac{2}{p^3}\right) + O(H^{1/2-\varepsilon/5}). \end{aligned}$$

This verifies the lemma. □

We also have a minor variant:

LEMMA 9. Fix $\varepsilon \in (0, \frac{1}{100})$. Let $S(x) = \frac{\sin \pi x}{\pi x}$, defined by continuity at $x = 0$, and let $z \geq H^{1+\varepsilon}$. Then

$$2H^2 \sum_{d_1^2, d_2^2 \leq z} \frac{\mu(d_1)\mu(d_2)}{d_1^2 \cdot d_2^2} \sum_{\lambda \geq 1} S\left(\frac{H\lambda}{(d_1^2, d_2^2)}\right)^2 = CH^{1/2} + O(H^{1/2-\varepsilon/8}). \quad (20)$$

Proof. We first note that

$$\int_0^\infty S(y)^2 \sqrt{y} dy = \frac{1}{\pi}. \quad (21)$$

This identity follows from [GR14, formula 3.823].

Thus (20) is a variant of (17). Lemma 8 does not apply directly because S does not decay quickly enough. To overcome this issue, we let h be a smooth bump function such that $h(x) = 1$ for $|x| \leq 1$ and $h(x) = 0$ for $|x| \geq 2$. We introduce the function

$$W(y) = S(y)h(y/H^{\varepsilon/4})$$

which satisfies the hypothesis of Lemma 8 for our ε . On the other hand for such W

$$\begin{aligned} & 2H^2 \sum_{d_1^2, d_2^2 \leq z} \frac{\mu(d_1)\mu(d_2)}{d_1^2 \cdot d_2^2} \sum_{\lambda \geq 1} \left(S\left(\frac{H\lambda}{(d_1^2, d_2^2)}\right)^2 - W\left(\frac{H\lambda}{(d_1^2, d_2^2)}\right)^2 \right) \\ & \ll H^2 \sum_{d_1^2, d_2^2} \frac{1}{d_1^2 d_2^2} \sum_{\lambda \geq 1} \frac{1}{(H\lambda/(d_1^2, d_2^2))^2} \mathbf{1}\left(\frac{H\lambda}{(d_1^2, d_2^2)} \geq H^{\varepsilon/4}\right). \end{aligned}$$

We split the sum over d_1 and d_2 into the complementary ranges $(d_1^2, d_2^2) \leq H^{1-\varepsilon/4}$ and $(d_1^2, d_2^2) > H^{1-\varepsilon/4}$. In the second case we utilize that $\lambda > H^{\varepsilon/4-1}(d_1^2, d_2^2)$, and we see that the above is

$$\ll \sum_{(d_1, d_2)^2 \leq H^{1-\varepsilon/4}} \frac{(d_1, d_2)^4}{d_1^2 d_2^2} + H^{1-\varepsilon/4} \sum_{(d_1, d_2)^2 > H^{1-\varepsilon/4}} \frac{(d_1, d_2)^2}{d_1^2 d_2^2}.$$

Writing $(d_1, d_2) = d_0$ and $d_i = \delta_i d_0$, the above is

$$\ll \sum_{d_0 \leq H^{1/2-\varepsilon/8}} \sum_{\delta_1, \delta_2} \frac{1}{\delta_1^2 \delta_2^2} + H^{1-\varepsilon/4} \sum_{d_0 > H^{1/2-\varepsilon/8}} \frac{1}{d_0^2} \sum_{\delta_1, \delta_2} \frac{1}{\delta_1^2 \delta_2^2} \ll H^{1/2-\varepsilon/8}. \quad (22)$$

On the other hand,

$$\int_0^\infty S(y)^2 \sqrt{y} dy - \int_0^\infty W(y)^2 \sqrt{y} dy \ll \int_{H^{\varepsilon/4}}^\infty y^{-3/2} dy \ll H^{-\varepsilon/8}.$$

Combining this with the bound (22) and the identity (21) verifies (20) with error term of order $H^{1/2-\varepsilon/5} + H^{1/2-\varepsilon/8} \ll H^{1/2-\varepsilon/8}$. \square

3.3 Initial reductions on second moments. The following lemma will be used in the proof of Proposition 2.

LEMMA 10. *If $F: \mathbb{R} \rightarrow \mathbb{C}$ is square-integrable and $H \leq X$, then*

$$\int_X^{2X} |F(x+H) - F(x)|^2 dx \ll \sup_{\theta \in [\frac{H}{3X}, \frac{3H}{X}]} \int_X^{3X} |F(u+\theta u) - F(u)|^2 du$$

Proof. The proof can be found in a paper by Saffari and Vaughan [SV77, Page 25] but for the convenience of the reader we include the proof here.

First note that by the triangle inequality we have, for any $v \geq H$,

$$|F(x+H) - F(x)|^2 \ll |F(x+v) - F(x)|^2 + |F(x+v) - F(x+H)|^2.$$

Integrating this over $x \in [X, 2X]$ and $v \in [2H, 3H]$,

$$\begin{aligned} & H \int_X^{2X} |F(x+H) - F(x)|^2 dx \\ & \ll \int_{2H}^{3H} \int_X^{2X} |F(x+v) - F(x)|^2 dx dv \\ & \quad + \int_{2H}^{3H} \int_X^{2X} |F(x+v) - F(x+H)|^2 dx dv. \end{aligned}$$

By a change of variables the right-hand side is equal to

$$\begin{aligned} & \int_{2H}^{3H} \int_X^{2X} |F(x+v) - F(x)|^2 dx dv + \int_H^{2H} \int_{X+H}^{2X+H} |F(y+w) - F(y)|^2 dy dw \\ & \leq \int_H^{3H} \int_X^{3X} |F(x+v) - F(x)|^2 dx dv = \int_X^{3X} \int_H^{3H} |F(x+v) - F(x)|^2 dv dx. \end{aligned}$$

Changing the order of integration was justified by Fubini's theorem. Letting $v = \theta x$ in the inner integral of the last expression above, we see the right-hand side is equal to

$$\int_X^{3X} \int_{H/x}^{3H/x} |F(x+\theta x) - F(x)|^2 x d\theta dx \ll X \int_X^{3X} \int_{H/3X}^{3H/X} |F(x+\theta x) - F(x)|^2 d\theta dx.$$

Collecting everything and swapping the order of integration again, we obtain

$$H \int_X^{2X} |F(x+H) - F(x)|^2 dx \ll X \int_{H/3X}^{3H/X} \int_X^{3X} |F(u+\theta u) - F(u)|^2 dud\theta,$$

which immediately implies the claim. \square

We will frequently use the following immediate consequences of the orthogonality of characters: for any sequence b_n of complex numbers,

$$\begin{aligned} \frac{1}{\varphi(q)} \sum_{\substack{\chi \pmod{q} \\ \chi \neq \chi_0}} \left| \sum_n b_n \chi(n) \right|^2 &= \sum_{\substack{n_1 \equiv n_2 \pmod{q} \\ (n_1 n_2, q) = 1}} b_{n_1} \overline{b_{n_2}} - \frac{1}{\varphi(q)} \sum_{\substack{n_1, n_2 \\ (n_1 n_2, q) = 1}} b_{n_1} \overline{b_{n_2}} \\ &= \sum_{\substack{a \pmod{q} \\ (a, q) = 1}} \left| \sum_{n \equiv a \pmod{q}} b_n - \frac{1}{\varphi(q)} \sum_{(n, q) = 1} b_n \right|^2 \end{aligned} \quad (23)$$

and

$$\frac{1}{\varphi(q)} \sum_{\chi \pmod{q}} \left| \sum_n b_n \chi(n) \right|^2 = \sum_{\substack{n_1 \equiv n_2 \pmod{q} \\ (n_1 n_2, q) = 1}} b_{n_1} \overline{b_{n_2}} = \sum_{\substack{a \pmod{q} \\ (a, q) = 1}} \left| \sum_{n \equiv a \pmod{q}} b_n \right|^2. \quad (24)$$

3.4 Point-counting lemmas.

LEMMA 11. *Let $a, b \in \mathbb{N}$ be such that $\sqrt{b/a}$ is irrational. Let $\eta \in (0, 1]$ and $M \geq 1$. The number of $m \sim M$ such that*

$$\left\| m \sqrt{\frac{b}{a}} \right\| \leq \eta \quad (25)$$

is bounded by

$$\ll \eta M + \sqrt{\eta M} (ab)^{1/4} + 1.$$

Proof. We can clearly assume that $\eta^{1/2} (ab)^{1/4} \leq M^{1/2}$ since otherwise the claim is trivial. Assume we have a (reduced) rational approximation r/q with $r \in \mathbb{Z}$ and $q \in \mathbb{N}$ such that

$$\left| \sqrt{\frac{b}{a}} - \frac{r}{q} \right| \leq \frac{1}{q^2}. \quad (26)$$

Now, writing each $m \in (M, 2M]$ as $m = kq + \ell$ with $0 \leq \ell \leq q - 1$, we see that the number of solutions to (25) with $m \sim M$ is at most

$$\begin{aligned} & \sum_{\lfloor M/q \rfloor \leq k \leq 2M/q} \left| \left\{ 0 \leq \ell \leq q - 1 : \left\| (kq + \ell) \sqrt{\frac{b}{a}} \right\| \leq \eta \right\} \right| \\ & \ll \left(\frac{M}{q} + 1 \right) \max_{\xi \in [0, 1]} \left| \left\{ 0 \leq \ell \leq q - 1 : \left\| \ell \sqrt{\frac{b}{a}} + \xi \right\| \leq \eta \right\} \right| \\ & \ll \left(\frac{M}{q} + 1 \right) \max_{\xi \in [0, 1]} \left| \left\{ 0 \leq \ell \leq q - 1 : \left\| \frac{\ell r}{q} + \xi \right\| \leq \eta + 1/q \right\} \right| \\ & \ll \left(\frac{M}{q} + 1 \right) (q \cdot \eta + 1) \ll M\eta + \frac{M}{q} + q\eta + 1. \end{aligned}$$

Now since $\sqrt{b/a} = \sqrt{ab}/a$ is a quadratic irrational, the partial denominators in its continued fraction expansion have size at most $2\sqrt{ab}$ (see for instance [RS92, p. 44]). In particular this means that for any given $R \geq 1$, we can find $q \in [R, 3\sqrt{ab}R]$ such that (26) holds for some r coprime to q . Taking $R = M^{1/2}/(\eta^{1/2}(ab)^{1/4}) \geq 1$, we see that the number of solutions is indeed

$$\ll M\eta + M^{1/2}\eta^{1/2}(ab)^{1/4} + 1. \quad \square$$

LEMMA 12. *Let $a, b \in \mathbb{N}$ be such that $\sqrt{b/a}$ is irrational, and let $M_1, M_2, T \geq 1$. The number of solutions to*

$$|am_1^2 - bm_2^2| \leq \frac{bM_2^2}{T} \quad \text{with } m_1 \sim M_1 \text{ and } m_2 \sim M_2$$

is

$$\ll \frac{M_1M_2}{T} + \left(\frac{(M_1M_2)^{1/2}(ab)^{1/4}}{T^{1/2}} + 1 \right) \cdot \mathbf{1}_{M_2 < T}.$$

Proof. Dividing by b and factoring, we see that we need to count the number of solutions to

$$\left| \left(m_1 \sqrt{\frac{a}{b}} - m_2 \right) \left(m_1 \sqrt{\frac{a}{b}} + m_2 \right) \right| \leq \frac{M_2^2}{T}$$

Dividing by the second factor, we see that it suffices to count the number of solutions to

$$\left| m_1 \sqrt{\frac{a}{b}} - m_2 \right| \leq \frac{M_2}{T}.$$

If $M_2 \geq T$, we have M_1 choices for m_1 and after that $O(M_2/T)$ choices for m_2 , so in total M_1M_2/T solutions which is fine.

If $M_2 < T$, then once m_1 is chosen there are at most two choices for m_2 . Therefore it suffices to count the number of integers $m_1 \sim M_1$ such that

$$\left\| m_1 \sqrt{\frac{b}{a}} \right\| \leq \frac{M_2}{T}.$$

The result now follows from Lemma 11. □

4 The range $H^{1+\varepsilon} \leq z \leq \min\{X/H^{1/2+\varepsilon}, H^{1/2-\varepsilon}X^{1/2}\}$ in the t -aspect: Proof of Proposition 1

In what follows we let S be the sinc function as defined in Lemma 9. Proposition 1 follows immediately combining the following proposition with Lemma 9.

PROPOSITION 5. Let $X^\varepsilon \leq H \leq X^{2/3-\varepsilon}$ and $H^{1+\varepsilon} \leq z \leq \min\{X^{1-\varepsilon}/H^{1/2}, H^{1/2-\varepsilon} X^{1/2}\}$. Then, as $X \rightarrow \infty$,

$$\begin{aligned} & \frac{1}{X} \int_X^{2X} \left| \sum_{d^2 \leq z} \mu(d) \sum_{x/d^2 < n \leq (x+H)/d^2} 1 - H \sum_{d^2 \leq z} \frac{\mu(d)}{d^2} \right|^2 dx \\ &= (1 + O(H^{-\varepsilon/2})) 2H^2 \sum_{k_1^2, k_2^2 \leq z} \frac{\mu(k_1)\mu(k_2)}{k_1^2 k_2^2} \sum_{\lambda \geq 1} S\left(\frac{H\lambda}{(k_1^2, k_2^2)}\right)^2 + O(H^{1/2-\varepsilon/3}). \end{aligned}$$

Proof. We prove a smoothed version of the claim first. Let $\sigma: \mathbb{R} \rightarrow \mathbb{R}$ be an absolutely integrable function such that $\hat{\sigma}$ is supported in the interval $[-BH^{\varepsilon/2}, BH^{\varepsilon/2}]$ for some constant B to be specified later. We first show that as $X \rightarrow \infty$,

$$\begin{aligned} & \frac{1}{X} \int_{-\infty}^{\infty} \sigma\left(\frac{x}{X}\right) \left| \sum_{d^2 \leq z} \mu(d) \sum_{x/d^2 < n \leq (x+H)/d^2} 1 - H \sum_{d^2 \leq z} \frac{\mu(d)}{d^2} \right|^2 dx \\ &= 2\hat{\sigma}(0)H^2 \sum_{k_1^2, k_2^2 \leq z} \frac{\mu(k_1)\mu(k_2)}{k_1^2 k_2^2} \sum_{\lambda \geq 1} S\left(\frac{H\lambda}{(k_1^2, k_2^2)}\right)^2 + O(H^{1/2-\varepsilon/3}). \end{aligned} \quad (27)$$

Here

$$\sum_{x/d^2 < n \leq (x+H)/d^2} 1 = H/d^2 + \psi(x/d^2) - \psi((x+H)/d^2), \quad (28)$$

where $\psi(y) = y - [y] - 1/2$ with $[y]$ the integral part of y . For ψ we have the Fourier expansion (see e.g. [IK04, (4.18)])

$$\psi(y) = -\frac{1}{2\pi i} \sum_{0 < |n| \leq N} \frac{1}{n} e(y n) + O(\min\{1, 1/(N\|y\|)\}). \quad (29)$$

We take $N = X^{10}$ and plug (29) into (28). The arising error term is $O(1/X^5)$ unless $\|x/d^2\| < X^{-5}$ or $\|(x+H)/d^2\| < X^{-5}$. Given this, it is easy to see that the error term leads to acceptable contribution to the left hand side (27).

Hence, the left hand side of (27) can be replaced by

$$\frac{1}{4\pi^2 X} \int_{-\infty}^{\infty} \sigma\left(\frac{x}{X}\right) \left| \sum_{d^2 \leq z} \mu(d) \sum_{0 < |n| \leq N} \frac{1}{n} e\left(\frac{nx}{d^2}\right) \left(1 - e\left(\frac{nH}{d^2}\right)\right) \right|^2 dx.$$

Expanding, this equals

$$\begin{aligned} & \frac{1}{4\pi^2} \sum_{d_1^2, d_2^2 \leq z} \sum_{0 < |n_1|, |n_2| \leq N} \mu(d_1)\mu(d_2) \frac{1}{n_1 n_2} \\ & \times \left(1 - e\left(\frac{n_1 H}{d_1^2}\right)\right) \overline{\left(1 - e\left(\frac{n_2 H}{d_2^2}\right)\right)} \hat{\sigma}\left(-X\left(\frac{n_1}{d_1^2} - \frac{n_2}{d_2^2}\right)\right). \end{aligned} \quad (30)$$

Owing to the support of $\hat{\sigma}$ this implies that we may restrict the sum in (30) to those integers for which

$$\left| \frac{n_1}{d_1^2} - \frac{n_2}{d_2^2} \right| \leq \frac{BH^{\varepsilon/2}}{X}. \quad (31)$$

We consider separately those (n_1, n_2, d_1, d_2) for which $n_1 d_2^2 = n_2 d_1^2$ and those for which this does not hold. In the first case parameterizing solutions in n_1 and n_2 by $n_1 = \lambda d_1^2 / (d_1^2, d_2^2)$ and $n_2 = \lambda d_2^2 / (d_1^2, d_2^2)$ for $\lambda \in \mathbb{Z} \setminus \{0\}$, we obtain

$$\frac{\hat{\sigma}(0)}{4\pi^2} \sum_{d_1^2, d_2^2 \leq z} \mu(d_1) \mu(d_2) \sum_{\lambda \neq 0} \frac{(d_1, d_2)^4}{d_1^2 d_2^2 \lambda^2} \left| 1 - e \left(\frac{\lambda H}{(d_1, d_2)^2} \right) \right|^2 + O \left(\frac{1}{X^5} \right),$$

where the error term comes from adding $|n_i| > N$ (for which surely $|\lambda| > X^8$). Here

$$\left| 1 - e \left(\frac{\lambda H}{(d_1, d_2)^2} \right) \right| = 2 \left| \sin \left(\frac{\lambda \pi H}{(d_1^2, d_2^2)} \right) \right|,$$

so we get the desired main term involving $S(\lambda H / (d_1^2, d_2^2))$.

Therefore it remains to show that the contribution of terms with $n_1 d_2^2 \neq n_2 d_1^2$ is negligible. Splitting n_j and d_j dyadically, we need to bound, for any $D_1, D_2 \leq z^{1/2}$ and any $N_1, N_2 \leq N$,

$$\min \left\{ \frac{1}{N_1}, \frac{H}{D_1^2} \right\} \min \left\{ \frac{1}{N_2}, \frac{H}{D_2^2} \right\} \sum_{\substack{n_1 \sim N_1 \\ n_2 \sim N_2}} \# \left\{ (d_1, d_2) : d_j \sim D_j, 0 < \left| \frac{n_1}{d_1^2} - \frac{n_2}{d_2^2} \right| \leq \frac{BH^{\varepsilon/2}}{X} \right\} \quad (32)$$

and we need a bound that is $O(H^{1/2-\varepsilon/2})$. Now

$$\begin{aligned} & \# \{ (d_1, d_2) : d_j \sim D_j, 0 < \left| \frac{n_1}{d_1^2} - \frac{n_2}{d_2^2} \right| \leq \frac{BH^{\varepsilon/2}}{X} \} \\ & \ll \# \left\{ (d_1, d_2) : d_j \sim D_j, 0 < \left| n_1 d_2^2 - n_2 d_1^2 \right| \leq 16 \frac{BD_2^2 H^{\varepsilon/2}}{X N_2} \cdot D_1^2 N_2 \right\}. \end{aligned} \quad (33)$$

Notice that there are no solutions unless

$$N_1 D_2^2 \asymp N_2 D_1^2. \quad (34)$$

We split into two cases according to whether $\sqrt{n_2/n_1}$ is quadratic irrational or instead rational. In the first case we can apply Lemma 12, which shows that the number of solutions (33) is

$$\ll \frac{H^{\varepsilon/2} D_1 D_2^3}{X N_2} + 1 + \frac{D_1^{1/2} D_2^{3/2} N_1^{1/4} H^{\varepsilon/4}}{X^{1/2} N_2^{1/4}}.$$

By (34) we can multiply the first term by $(D_1/D_2)(N_2/N_1)^{1/2}$ and the third term by $(D_1/D_2)^{1/2}(N_2/N_1)^{1/4}$ to obtain

$$\ll \frac{H^{\varepsilon/2} D_1^2 D_2^2}{X(N_1 N_2)^{1/2}} + 1 + \frac{D_1 D_2 H^{\varepsilon/4}}{X^{1/2}}.$$

Using this bound in (32), and summing over n_1 and n_2 , we note that the maximum is attained for $N_j = D_j^2/H$ and thus the contribution to (32) from $\sqrt{n_2/n_1}$ quadratic irrational is bounded by

$$\ll H^{\varepsilon/2} \left(\frac{D_1 D_2 H}{X} + 1 + \frac{D_1 D_2}{X^{1/2}} \right) = O(H^{1/2-\varepsilon/2})$$

since $D_1 \cdot D_2 \leq z \leq \min\{X/H^{1/2+\varepsilon}, H^{1/2-\varepsilon} X^{1/2}\}$.

In case $\sqrt{n_2/n_1}$ is rational, there exist $m, \ell_1, \ell_2 \in \mathbb{Z}$ such that $n_1 = m\ell_1^2$ and $n_2 = m\ell_2^2$. Hence, writing $r_1^2 = \ell_1^2 d_1^2$ and $r_2^2 = \ell_2^2 d_2^2$, we see that the contribution to (32) for $\sqrt{n_2/n_1}$ rational is bounded by

$$\begin{aligned} &\ll H^{\varepsilon/1000} \min\left\{\frac{1}{N_1}, \frac{H}{D_1^2}\right\} \min\left\{\frac{1}{N_2}, \frac{H}{D_2^2}\right\} \\ &\quad \times \sum_m \#\left\{(r_1, r_2): r_j \leq D_j \sqrt{N_j/m}, 0 < |r_1^2 - r_2^2| \leq \frac{BH^{\varepsilon/2} D_1^2 D_2^2}{mX}\right\}. \end{aligned} \quad (35)$$

Factoring $r_1^2 - r_2^2 = (r_1 - r_2)(r_1 + r_2)$ and dividing by the second factor, we see that the number of solutions (r_1, r_2) is

$$\ll \frac{BH^{\varepsilon/2} D_1^2 D_2^2}{mX} \log X$$

Summing over $m \ll \min\{N_1, N_2\}$ and using this bound in (35), the maximum in the resulting bound for (35) is attained for $N_j = D_j^2/H$. Hence we obtain that (35) is at most $H^{2+\varepsilon/2+\varepsilon/500}/X \leq H^{1/2-\varepsilon/2}$ since $H \leq X^{2/3-\varepsilon}$.

Let us now dispose of the smoothing σ : Take B to be a sufficiently large absolute constant that there exist integrable functions σ_- and σ_+ such that $\hat{\sigma}_-$ and $\hat{\sigma}_+$ have support $[-BH^{\varepsilon/2}, BH^{\varepsilon/2}]$, and

$$\sigma_- \leq \mathbf{1}_{[1,2]} \leq \sigma_+, \quad \text{and} \quad \left| \int \sigma_{\pm}(x) dx - 1 \right| \leq H^{-\varepsilon/2}.$$

(We allow σ_- and σ_+ to take negative values.) An explicit construction of such functions is given by the Beurling–Selberg majorant and minorant [Mon01, p. 273]. Applying (27) and these bounds,

$$\begin{aligned} &\frac{1}{X} \int_{-\infty}^{\infty} \mathbf{1}_{[1,2]} \left(\frac{x}{X} \right) \left| \sum_{d^2 \leq z} \mu(d) \sum_{x/d^2 \leq n \leq (x+H)/d^2} 1 - H \sum_{k^2 \leq z} \frac{\mu(k)}{k^2} \right|^2 dx \\ &= (1 + O(H^{-\varepsilon/2})) 2H^2 \sum_{k_1^2, k_2^2 \leq z} \frac{\mu(k_1)\mu(k_2)}{k_1^2 k_2^2} \sum_{\lambda \geq 1} S\left(\frac{H\lambda}{(k_1^2, k_2^2)}\right)^2 + O(H^{1/2-\varepsilon/3}). \quad \square \end{aligned}$$

5 The range $z \geq H^{4/3+\varepsilon}$ in the t -aspect: Proof of Proposition 2

We would like to establish that

$$\frac{1}{X} \int_X^{2X} \left| \sum_{\substack{x < nd^2 \leq x+H \\ d^2 > z}} \mu(d) - H \sum_{z < d^2 \leq 2X} \frac{\mu(d)}{d^2} \right|^2 dx \ll H^{1/2-\varepsilon/8}.$$

Splitting into dyadic ranges according to the size of d , it essentially suffices to show that, for each $D \in [z^{1/2}, (2X)^{1/2}]$, we have

$$\frac{1}{X} \int_X^{2X} \left| \sum_{\substack{x < nd^2 \leq x+H \\ d \sim D}} \mu(d) - H \sum_{d \sim D} \frac{\mu(d)}{d^2} \right|^2 dx \ll H^{1/2-\varepsilon/4}. \quad (36)$$

Let

$$A(x) := \sum_{\substack{nd^2 \leq x \\ d \sim D}} \mu(d) - x \sum_{d \sim D} \frac{\mu(d)}{d^2}.$$

Using this definition and Lemma 10, we see that the left-hand side of (36) is

$$\frac{1}{X} \int_X^{2X} |A(x+H) - A(x)|^2 dx \ll \frac{1}{X} \int_X^{3X} |A(u(1+\theta)) - A(u)|^2 du \quad (37)$$

for some $\theta \in [\frac{H}{3X}, \frac{3H}{X}]$. Choose w such that $e^w = 1 + \theta$, so that $w \asymp \frac{H}{X}$. By contour integration

$$A(e^y) = \frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} \frac{e^{ys}}{s} \zeta(s) M(2s) ds - e^y \sum_{d \sim D} \frac{\mu(d)}{d^2}, \quad (38)$$

where

$$M(s) := \sum_{d \sim D} \frac{\mu(d)}{d^s}.$$

Moving the contour to the line $\Re s = 1/2$ we notice that the residue from $s = 1$ cancels with the second term on the right-hand side of (38), and we obtain

$$\frac{A(e^{w+x}) - A(e^x)}{e^{x/2}} = \frac{1}{2\pi} \int_{\mathbb{R}} \frac{e^{w(\frac{1}{2}+it)} - 1}{\frac{1}{2} + it} e^{itx} \zeta(\frac{1}{2} + it) M(1 + 2it) dt.$$

Therefore, by Plancherel,

$$\int_0^\infty |A(e^{u+w}) - A(e^u)|^2 \cdot \frac{du}{e^u} \ll \int_{\mathbb{R}} \left| \frac{e^{w(\frac{1}{2}+it)} - 1}{\frac{1}{2} + it} \right|^2 \cdot |\zeta(\frac{1}{2} + it) M(1 + 2it)|^2 dt. \quad (39)$$

Combining (37) and (39) we get after a change of variable,

$$\begin{aligned}
\frac{1}{X} \int_X^{2X} |A(x+H) - A(x)|^2 dx &\ll X \int_0^\infty |A(u(1+\theta)) - A(u)|^2 \frac{du}{u^2} \\
&\ll X \int_{\mathbb{R}} \left| \frac{e^{w(\frac{1}{2}+it)} - 1}{\frac{1}{2} + it} \right|^2 \cdot |\zeta(\frac{1}{2} + it)M(1+2it)|^2 dt \\
&\ll X \int_{\mathbb{R}} \min\left\{ \left(\frac{H}{X}\right)^2, \frac{1}{|t|^2} \right\} \cdot |\zeta(\frac{1}{2} + it)M(1+2it)|^2 dt.
\end{aligned} \tag{40}$$

By Lemma 5 the part with $|t| \geq X^2$ contributes

$$\ll X \int_{X^2}^\infty |t|^{-5/3+\varepsilon} dt = O(1).$$

On the other hand, the contribution of $|t| \leq X^2$ to the right-hand side of (40) is at most

$$\begin{aligned}
&\ll \frac{H^2}{X} \int_{|t| \leq 2X/H} |\zeta(\frac{1}{2} + it)M(1+2it)|^2 dt \\
&\quad + X \int_{X/H}^{X^2} \frac{1}{T^2} \cdot \frac{1}{T} \int_{T \leq |t| \leq 2T} |\zeta(\frac{1}{2} + it)M(1+2it)|^2 dt dT \\
&\ll H \left(\sup_{X/H \leq T \leq X^2} \frac{1}{T} \int_{|t| \leq T} |\zeta(\frac{1}{2} + it)M(1+2it)|^2 dt \right) + O(1).
\end{aligned} \tag{41}$$

Let us now prove the claim on the assumption of the Lindelöf Hypothesis. Applying Lindelöf and then the mean-value theorem (Lemma 7 with $q = 1$), we have for any choice of $\delta > 0$,

$$\begin{aligned}
\frac{H}{T} \int_{|t| \leq T} |\zeta(1/2 + it)M(1 + i2t)|^2 dt &\ll \frac{HT^\delta}{T} \int_{|t| \leq T} |M(1 + i2t)|^2 dt \\
&\ll \frac{HT^\delta}{T} (T + D) \cdot \frac{1}{D} \ll \frac{HT^\delta}{D} + \frac{HT^\delta}{T}.
\end{aligned}$$

Recall we have $D \geq z^{1/2} \geq H^{(1+\varepsilon)/2}$, $H \leq X^{2/3-\varepsilon}$ and $X/H \leq T \leq X^2$. Hence the above is

$$\ll H^{1/2-\varepsilon/2} T^\delta + \frac{H^{2-\delta}}{X^{1-\delta}} \ll H^{1/2-\varepsilon/4} + \frac{H^{2-\delta}}{H^{(1-\delta)/(2/3-\varepsilon)}} \ll H^{1/2-\varepsilon/4},$$

for δ sufficiently small. Applying this bound to (41) yields the claim.

Let us now prove the unconditional part of the proposition. First notice that the values of t for which $|M(1+2it)| \leq D^{-1/2+\varepsilon/16}$ contribute to (41) by Cauchy–Schwarz

and the fourth moment bound (Lemma 3) $O(H^{1+\varepsilon/16}D^{-1+\varepsilon/8}) = O(H^{1/2-\varepsilon/4})$, and therefore their contribution is always acceptable. Writing

$$S(V) = \{t \in [-T, T]: V \leq |M(1 + 2it)| < 2V\},$$

by dyadic splitting, it suffices to show that, for each $V \in [D^{-1/2}, 1]$ and $T \in [X/H, X^2]$, we have

$$\frac{H}{T}V^2 \int_{S(V)} |\zeta(1/2 + it)|^2 dt \ll H^{1/2-\varepsilon/3}.$$

Now by Lemma 1 we have

$$|S(V)| \ll (V^{-2} + T \min\{D^{-1}V^{-2}, D^{-2}V^{-6}\})(\log 2X)^6. \quad (42)$$

Consider first the case when the first term dominates here. Then by Lemma 5 we have

$$\frac{H}{T}V^2 \int_{S(V)} |\zeta(1/2 + it)|^2 dt \ll \frac{H}{T}T^{1/3+\varepsilon/2} \ll \frac{H}{T^{2/3-\varepsilon/2}} \ll \frac{H^{5/3}}{X^{2/3-\varepsilon/2}} \leq H^{1/2-\varepsilon/3}$$

since $H \leq X^{4/7-\varepsilon}$.

Consider now the case that the second term dominates in (42). Then, by Cauchy–Schwarz and the fourth moment estimate (Lemma 3),

$$\begin{aligned} \frac{H}{T}V^2 \int_{S(V)} |\zeta(1/2 + it)|^2 dt &\ll \frac{H}{T}V^2 |S(V)|^{1/2} \left(\int_{|t| \leq T} |\zeta(\frac{1}{2} + it)|^4 dt \right)^{1/2} \\ &\ll HV^2 \min\{D^{-1}V^{-2}, D^{-2}V^{-6}\}^{1/2} (\log 2X)^5 \ll H \min\{D^{-1/2}V, D^{-1}V^{-1}\} (\log 2X)^5 \\ &\ll H(D^{-1/2}V)^{1/2} (D^{-1}V^{-1})^{1/2} (\log 2X)^5 \ll HD^{-3/4} (\log 2X)^5 \ll Hz^{-3/8} (\log 2X)^5 \ll H^{1/2-\varepsilon/3} \end{aligned}$$

since $z \geq H^{4/3+\varepsilon}$. This finishes the proof of Proposition 2. \square

6 The range $(x/q)^{1+\varepsilon} \leq z < x^{-\varepsilon} \sqrt{qx}$ in the q -aspect: Proof of Proposition 3

By (23) Proposition 3 follows immediately from the following proposition.

PROPOSITION 6. *Let $\varepsilon \in (0, 1/100)$. Let q be prime with $x^{1/3+30\varepsilon} \leq q \leq x^{1-\varepsilon}$ and let $(x/q)^{1+\varepsilon} \leq z \leq x^{-\varepsilon} \sqrt{qx}$. Then*

$$\frac{1}{\varphi(q)} \sum_{\substack{\chi \pmod{q} \\ \chi \neq \chi_0}} \left| \sum_{\substack{d^2 \leq z \\ nd^2 \leq x}} \mu(d) \chi(d^2) \chi(n) \right|^2 = C \sqrt{qx} + O((x/q)^{-\varepsilon/16} \sqrt{qx}) \quad (43)$$

with C as in (4).

The proof of Proposition 6 is based on two Propositions that we now describe. Proposition 7 below will be used to introduce a smoothing into (43). Note that it gives an upper bound that is $o(\sqrt{qx})$ whenever $z = o(\sqrt{qx}/(\log x)^6)$ and the interval I has length $o(x/(\log x)^{12})$.

PROPOSITION 7. *Let q be prime with $q \leq x$, let $z \leq x$. Let $I \subset [1, 2x]$ be an interval. Then*

$$\frac{1}{\varphi(q)} \sum_{\substack{\chi \pmod{q} \\ \chi \neq \chi_0}} \left| \sum_{\substack{d^2 \leq z \\ nd^2 \in I}} \mu(d)\chi(d^2)\chi(n) \right|^2 \ll (\log x)^6 \cdot \left(z + \sqrt{|I|q} \right). \quad (44)$$

We will use the following proposition to evaluate the smoothed analogue of (43).

PROPOSITION 8. *Let $\varepsilon > 0$ be given. Let f be a smooth function such that f is compactly supported on $[0, 1]$ and $f(u) = 1$ for $(x/q)^{-\varepsilon/4} \leq u \leq 1 - (x/q)^{-\varepsilon/4}$ and for each integer $k \geq 0$, we have $f^{(k)}(u) \ll (x/q)^{\varepsilon k/4}$. Let $(x/q)^{1+\varepsilon} \leq z \leq x^{-\varepsilon} \sqrt{qx}$. Then for $x^{1/3+30\varepsilon} \leq q \leq x^{1-\varepsilon}$,*

$$\frac{1}{\varphi(q)} \sum_{\substack{\chi \pmod{q} \\ \chi \neq \chi_0}} \left| \sum_{\substack{n \geq 1 \\ d^2 \leq z}} f\left(\frac{nd^2}{x}\right) \mu(d)\chi(d^2)\chi(n) \right|^2 = C\sqrt{qx} + O((x/q)^{-\varepsilon/10} \sqrt{qx}),$$

where C is as in (4).

One way to construct f satisfying the assumptions of the proposition is to take $\phi(t)$ to be a smooth function which vanishes for negative t and has $\phi(t) = 1$ for t greater than 1, and then set $f(u) = \phi((x/q)^{\varepsilon/4}u)\phi((x/q)^{\varepsilon/4}(1-u))$.

With these two propositions at hand we are ready to prove Proposition 6.

6.1 Proof of Proposition 6. For $m \in \mathbb{N}$, set

$$A_m := \sum_{\substack{d^2 | m \\ d^2 \leq z}} \mu(d),$$

and let f be as described below Proposition 8. Then

$$\frac{1}{\varphi(q)} \sum_{\substack{\chi \pmod{q} \\ \chi \neq \chi_0}} \left| \sum_{\substack{n \leq x}} A_n \chi(n) \right|^2 = S_1 + O(\sqrt{S_1 S_2} + S_2),$$

where

$$S_1 := \frac{1}{\varphi(q)} \sum_{\substack{\chi \pmod{q} \\ \chi \neq \chi_0}} \left| \sum_n A_n \chi(n) f\left(\frac{n}{x}\right) \right|^2$$

and

$$S_2 := \frac{1}{\varphi(q)} \sum_{\substack{\chi \pmod{q} \\ \chi \neq \chi_0}} \left| \sum_{n \in I} A_n \chi(n) \left(1 - f\left(\frac{n}{x}\right)\right) \right|^2$$

where $I = I_1 \cup I_2$ with $I_1 := [1, x \cdot (x/q)^{-\varepsilon/4}]$ and $I_2 := [x - x \cdot (x/q)^{-\varepsilon/4}, x]$.

For $i = 1, 2$, define

$$B_i(\chi; t) = \sum_{\substack{n \in I_i \\ n < t}} A_n \chi(n).$$

By partial summation,

$$\begin{aligned} \sum_{n \in I_2} A_n \chi(n) \left(1 - f\left(\frac{n}{x}\right)\right) &= \int_{I_2} \left(1 - f\left(\frac{t}{x}\right)\right) dB_2(\chi; t) \\ &= \frac{1}{x} \int_{I_2} f'\left(\frac{t}{x}\right) B_2(\chi; t) dt + B_2(\chi; x). \end{aligned}$$

Hence,

$$\begin{aligned} &\frac{1}{\varphi(q)} \sum_{\chi \neq \chi_0} \left| \sum_{n \in I_2} A_n \chi(n) \left(1 - f\left(\frac{n}{x}\right)\right) \right|^2 \\ &\ll \frac{1}{\varphi(q)} \sum_{\chi \neq \chi_0} \left| \frac{1}{x} \int_{I_2} f'\left(\frac{t}{x}\right) B_2(\chi; t) dt \right|^2 + \frac{1}{\varphi(q)} \sum_{\chi \neq \chi_0} |B_2(\chi; x)|^2 \\ &\leq (x/q)^{\varepsilon/4} \cdot \frac{1}{\varphi(q)} \sum_{\chi \neq \chi_0} \frac{1}{x} \int_{I_2} |B_2(\chi; t)|^2 dt + \frac{1}{\varphi(q)} \sum_{\chi \neq \chi_0} |B_2(\chi; x)|^2. \end{aligned} \quad (45)$$

Now by Proposition 7 we have, for $t \in I_2$,

$$\frac{1}{\varphi(q)} \sum_{\chi \neq \chi_0} |B_2(\chi; t)|^2 \ll (\log x)^6 \cdot \left(x^{-\varepsilon} \sqrt{qx} + \sqrt{(t - (x - x \cdot (x/q)^{-\varepsilon/4})) \cdot q} \right).$$

Therefore (45) is

$$\begin{aligned} &\ll (\log x)^6 \cdot (x/q)^{\varepsilon/4} \cdot \frac{1}{x} \cdot \left(x \cdot (x/q)^{-\varepsilon/4} \cdot x^{-\varepsilon} \sqrt{qx} + x^{3/2} (x/q)^{-3\varepsilon/8} \sqrt{q} \right) \\ &\quad + (\log x)^6 (x/q)^{-\varepsilon/8} \sqrt{qx} \ll (\log x)^6 (x/q)^{-\varepsilon/8} \cdot \sqrt{qx}. \end{aligned}$$

A similar argument shows that

$$\frac{1}{\varphi(q)} \sum_{\chi \neq \chi_0} \left| \sum_{n \in I_1} A_n \chi(n) \left(1 - f\left(\frac{n}{x}\right)\right) \right|^2 \ll (\log x)^6 (x/q)^{-\varepsilon/8} \cdot \sqrt{qx}.$$

as well. By $(a + b)^2 \ll |a|^2 + |b|^2$ we conclude that

$$S_2 \ll (\log x)^6 (x/q)^{-\varepsilon/8} \sqrt{qx}$$

as needed. On the other hand we can compute S_1 by using Proposition 8 and this yields the claimed estimate. \square

6.2 Proof of Proposition 7. By Pólya's formula (see [MV77, Lemma 1]) for $I = [a, b]$ and any $\chi \neq \chi_0$ of modulus q ,

$$\sum_{n \in I/d^2} \chi(n) = \frac{\tau(\chi)}{2\pi i} \sum_{1 \leq |n| \leq q} \bar{\chi}(n) f_{I/d^2}(n) + O(\log q),$$

where

$$f_{I/d^2}(n) = \frac{1}{n} \cdot \left(e\left(\frac{na}{d^2q}\right) - e\left(\frac{nb}{d^2q}\right) \right) \ll g_{I/d^2}(n) := \begin{cases} \frac{|I|}{d^2q} & \text{if } |n| \leq \frac{d^2q}{|I|}, \\ \frac{1}{n} & \text{otherwise.} \end{cases}$$

We split d and n into dyadic intervals and bound the left-hand side of (44) by

$$(\log x)^2 \sup_{\substack{D \leq z^{1/2} \\ 1 \leq N \leq q}} \sum_{\chi \pmod{q}} \left| \sum_{d \sim D} \mu(d) \chi(d^2) \sum_{n \sim N} \bar{\chi}(n) f_{I/d^2}(n) \right|^2 + O(z(\log q)^2). \quad (46)$$

The error term is clearly acceptable. We bound the main term of (46) using a majorant principle—by going through the first equality in (24) we can replace coefficients $\mu(d)$ and $f_{I/d^2}(n)$ by their majorants. Hence we get the bound

$$\ll (\log x)^2 \sup_{\substack{D \leq z^{1/2} \\ 1 \leq N \leq q}} \sum_{\chi} \left| \sum_d \chi^2(d) V\left(\frac{d}{D}\right) \cdot \sum_n \chi(n) V(n/N) g_{I/D^2}(N) \right|^2$$

with V a smooth function supported on $[1/2, 4]$.

The contribution of the principal character and quadratic character is $\ll z(\log x)^4$ which is acceptable. On the remaining non-principal and non-quadratic characters we apply Cauchy–Schwarz giving the upper bound

$$\ll (\log x)^2 \sup_{\substack{D \leq z^{1/2} \\ 1 \leq N \leq q}} g_{I/D^2}(N)^2 \left(\sum_{\chi^2 \neq \chi_0} \left| \sum_n \chi^2(n) V\left(\frac{n}{D}\right) \right|^4 \right)^{1/2} \left(\sum_{\chi \neq \chi_0} \left| \sum_n \chi(n) V\left(\frac{n}{N}\right) \right|^4 \right)^{1/2}. \quad (47)$$

We claim that

$$\begin{aligned} \sum_{\chi^2 \neq \chi_0} \left| \sum_n \chi^2(n) V\left(\frac{n}{D}\right) \right|^4 &\ll qD^2 \cdot (\log x)^4 \\ \text{and } \sum_{\chi \neq \chi_0} \left| \sum_n \chi(n) V\left(\frac{n}{N}\right) \right|^4 &\ll qN^2 \cdot (\log x)^4. \end{aligned} \quad (48)$$

We explain the second bound in (48); the first bound is similar. Let \tilde{V} be the Mellin transform of V . Using contour integration, the decay of \tilde{V} , and Hölder, we get, for every $A \geq 1$,

$$\sum_{\chi \neq \chi_0} \left| \sum_n \chi(n) V\left(\frac{n}{N}\right) \right|^4 = \sum_{\chi \neq \chi_0} \left| \int_{\mathbb{R}} L(1/2 + it, \chi) \tilde{V}(1/2 + it) N^{1/2+it} dt \right|^4$$

$$\begin{aligned} &\ll_A N^2 \sum_{\chi \neq \chi_0} \left(\int_{\mathbb{R}} \left| L(1/2 + it, \chi) \right| (1 + |t|)^{-A} dt \right)^4 \\ &\ll_A N^2 \sum_{\chi \neq \chi_0} \int_{\mathbb{R}} \left| L(1/2 + it, \chi) \right|^4 (1 + |t|)^{-A} dt. \end{aligned}$$

A dyadic decomposition of the integration range and the fourth moment bound for Dirichlet L -functions (Lemma 4) yield the second part of (48).

Using (48) in (47), we obtain an upper bound

$$\begin{aligned} &\ll (\log x)^6 \sup_{\substack{D < z^{1/2} \\ 1 \leq N \leq q}} g_{I/D^2}(N)^2 qDN \\ &\ll (\log x)^6 \sup_{\substack{D \leq \sqrt{|I|/q} \\ 1 \leq N \leq q}} g_{I/D^2}(N)^2 qDN + (\log x)^6 \sup_{\substack{\sqrt{|I|/q} < D \leq z^{1/2} \\ 1 \leq N \leq q}} g_{I/D^2}(N)^2 qDN. \end{aligned}$$

Recalling the definition of $g_{I/D^2}(N)$, we see that on the last line, the first N -supremum is attained for $N = 1$ and the second N -supremum is attained for $N = D^2q/|I|$, and we get the bound

$$\ll (\log x)^6 \sup_{D \leq \sqrt{|I|/q}} qD + (\log x)^6 \sup_{\sqrt{|I|/q} < D \leq z^{1/2}} |I|/D \ll (\log x)^6 \sqrt{|I|q}$$

and the claim follows. □

6.3 Proof of Proposition 8. We apply Poisson summation (see e.g. [IK04, formula (4.26)]) in the sum over n , getting

$$\sum_n \chi(n) f\left(\frac{nd^2}{x}\right) = \tau(\chi) \cdot \frac{x}{qd^2} \sum_{\ell} \bar{\chi}(\ell) \hat{f}\left(\frac{x\ell}{d^2q}\right).$$

Therefore we have to asymptotically estimate

$$\frac{q}{\varphi(q)} \cdot \frac{x^2}{q^2} \sum_{\chi \neq \chi_0} \left| \sum_{\substack{d^2 \leq z \\ \ell \in \mathbb{Z}}} \frac{\mu(d)}{d^2} \chi(d^2) \bar{\chi}(\ell) \hat{f}\left(\frac{x\ell}{d^2q}\right) \right|^2 \tag{49}$$

$$= \frac{x^2}{q} \sum_{\substack{n_1, n_2 \in \mathbb{Z} \\ d_1^2, d_2^2 \leq z \\ d_1^2 n_1 \equiv d_2^2 n_2 \pmod{q} \\ (d_1 d_2 n_1 n_2, q) = 1}} \frac{\mu(d_1)}{d_1^2} \frac{\mu(d_2)}{d_2^2} \cdot \hat{f}\left(\frac{xn_2}{d_1^2q}\right) \overline{\hat{f}\left(\frac{xn_1}{d_2^2q}\right)} + O(zx^{2\varepsilon/3}), \tag{50}$$

and where $O(zx^{2\varepsilon/3})$ comes from the principal character and from replacing $\varphi(q)$ by q . We note that since $z \leq x^{-\varepsilon} \sqrt{qx}$ this contribution is acceptable. Notice that we can add and remove the restrictions $d_1, d_2 > x^{1/2-\varepsilon/6}/\sqrt{q}$ and $|n_1|, |n_2| \leq x^{\varepsilon/3} \cdot zq/x$ at will because they cost us a negligible error term that is $\ll_A x^{-A}$ for any given $A > 0$. Moreover note that n_1 and n_2 now traverse all of \mathbb{Z} .

We now separate the set of tuples (n_1, n_2) into

$$\mathcal{M} := \{(k_1^2 m, k_2^2 m) : m \in \mathbb{Z} \text{ squarefree}, k_1, k_2 \in \mathbb{N}\}$$

and the complement. The $(n_1, n_2) \in \mathcal{M}$ contribute to a main term that is relatively easy to compute. On the other hand we will bound the contribution of $(n_1, n_2) \notin \mathcal{M}$.

6.3.1 The main term $(n_1, n_2) \in \mathcal{M}$. The conditions $d_1^2 n_1 \equiv d_2^2 n_2 \pmod{q}$ and $(n_1 n_2, q) = 1$ in the sum in (50) imply that if $(n_1, n_2) \in \mathcal{M}$ then $d_1^2 k_1^2 \equiv d_2^2 k_2^2 \pmod{q}$ and therefore $d_1 k_1 \equiv \pm d_2 k_2 \pmod{q}$. This implies that $d_1 k_1 = d_2 k_2$ since $d_j k_j \leq \sqrt{z} \cdot \sqrt{x^{\varepsilon/3} z q / x} = x^{\varepsilon/6} z \cdot \sqrt{q/x}$ and this is $\leq q/3$ because $z \leq x^{-\varepsilon} \sqrt{q x}$. We conclude that the contribution of $(n_1, n_2) \in \mathcal{M}$ is given by

$$\frac{x^2}{q} \sum_{k_1, k_2} \sum_{\substack{d_1 k_1 = d_2 k_2 \\ d_1^2, d_2^2 \leq z \\ (d_1 d_2 k_1 k_2, q) = 1}} \frac{\mu(d_1) \mu(d_2)}{d_1^2 \cdot d_2^2} \sum_{(m, q) = 1} \mu^2(m) \hat{f}\left(\frac{x k_2^2 m}{d_1^2 q}\right) \overline{\hat{f}\left(\frac{x k_1^2 m}{d_2^2 q}\right)}. \quad (51)$$

We now parametrize the equation $d_1 k_1 = d_2 k_2$ by dividing by (d_1, d_2) on both sides so that

$$k_1 = \frac{d_2 \ell}{(d_1, d_2)} \text{ and } k_2 = \frac{d_1 \ell}{(d_1, d_2)} \text{ with } \ell \in \mathbb{N}.$$

Plugging this and noticing that each non-negative integer can be written uniquely as $\ell^2 m$ with m squarefree, we can re-write (51) as

$$\frac{2x^2}{q} \sum_{\substack{d_1^2, d_2^2 \leq z \\ (d_1 d_2, q) = 1}} \frac{\mu(d_1) \mu(d_2)}{d_1^2 \cdot d_2^2} \sum_{\substack{\ell \geq 1 \\ (\ell, q) = 1}} \left| \hat{f}\left(\frac{x \ell}{q(d_1^2, d_2^2)}\right) \right|^2.$$

Note that we can drop the condition $(d_1 d_2, q) = 1$ as q is prime and $d_1, d_2 < q$. Likewise since $\ell \geq q$ contribute $O_A(x^{-A})$, we can drop the condition $(\ell, q) = 1$ and apply Lemma 8 with $W = \hat{f}$ and $H = x/q$ to see that the above is

$$C \sqrt{q x} \cdot \pi \int_0^\infty |\hat{f}(y)|^2 \sqrt{y} dy + O((x/q)^{-\varepsilon/8} \sqrt{q x}).$$

Let $F(u) = \mathbf{1}_{[0,1]}(u)$. We have,

$$\hat{f}(y) - \hat{F}(y) \ll \min\{(x/q)^{-\varepsilon/4}, |y|^{-1}\},$$

with the bound $(x/q)^{-\varepsilon/4}$ for the difference between these two Fourier transforms following from the fact that $\|f - F\|_{L^1} \ll (x/q)^{-\varepsilon/4}$, and the bound $1/|y|$ following from the fact that the total variation of the function $f - F$ is bounded by an absolute constant. Likewise

$$\hat{f}(y) \ll (1 + |y|)^{-1} \text{ and } \hat{F}(y) \ll (1 + |y|)^{-1}.$$

Hence

$$\begin{aligned} & \int_0^\infty |\hat{f}(y)|^2 \sqrt{y} dy - \int_0^\infty |\hat{F}(y)|^2 \sqrt{y} dy \\ & \ll \int_0^\infty \min\{(x/q)^{-\varepsilon/4}, y^{-1}\} (1+y)^{-1} \sqrt{y} dy \ll (x/q)^{-\varepsilon/8}. \end{aligned}$$

Putting these estimates together, and using the relation $|\hat{F}(\xi)| = |S(\xi)|$ and the integral identity (21), we see that (51) is

$$C\sqrt{qx} + O((x/q)^{-\varepsilon/8}\sqrt{xq}).$$

6.3.2 The off-diagonal $(n_1, n_2) \notin \mathcal{M}$. Let us focus on bounding the contribution of $(n_1, n_2) \notin \mathcal{M}$. We recall that the contribution of $d_1 \leq x^{1/2-\varepsilon/6}/q^{1/2}$ to (49) is negligible and likewise the contribution of $d_2 \leq x^{1/2-\varepsilon/6}/q^{1/2}$ is negligible. We now partition d_1, d_2 into intervals $[D_1, 2D_1]$ and $[D_2, 2D_2]$ with $x^{1/2-\varepsilon/6}/q^{1/2} \leq D_1, D_2 \leq \sqrt{x}$. The total contribution of $(n_1, n_2) \notin \mathcal{M}$ with $d_1 \in [D_1, 2D_1]$ and $d_2 \in [D_2, 2D_2]$ to (49) is bounded by

$$\frac{x^2}{q} \cdot \frac{1}{D_1^2 D_2^2} \sum_{(n_1, n_2) \notin \mathcal{M}} V\left(\frac{n_1}{N_1}\right) V\left(\frac{n_2}{N_2}\right) \sum_{\substack{d_1^2 n_1 \equiv d_2^2 n_2 \pmod{q} \\ (d_1, d_2, n_1, n_2, q) = 1}} V\left(\frac{d_1}{D_1}\right) V\left(\frac{d_2}{D_2}\right) \quad (52)$$

with V a smooth non-negative compactly supported function such that $V(x) \geq 1$ for $x \in [-2, 2]$ and $D_1, D_2 > x^{1/2-\varepsilon/6}/q^{1/2}$ and $N_1 \leq x^{\varepsilon/3} D_2^2 q/x$ and $N_2 \leq x^{\varepsilon/3} D_1^2 q/x$.

We now split into two cases according to the size of $D_1 D_2$:

6.3.3 Case $D_1 D_2 \geq x^{1+2\varepsilon}/q$. In this case we do not use the condition $(n_1, n_2) \notin \mathcal{M}$. Dropping this condition and using Dirichlet characters we can re-write (52) as

$$\begin{aligned} & \frac{x^2}{q\varphi(q)} \frac{1}{D_1^2 D_2^2} \sum_{\chi^2 \neq \chi_0} \left(\sum_{n_1} \chi(n_1) V\left(\frac{n_1}{N_1}\right) \right) \left(\sum_{n_2} \bar{\chi}(n_2) V\left(\frac{n_2}{N_2}\right) \right) \\ & \times \left(\sum_{d_1} \chi^2(d_1) V\left(\frac{d_1}{D_1}\right) \right) \left(\sum_{d_2} \bar{\chi}^2(d_2) V\left(\frac{d_2}{D_2}\right) \right) + O\left(\frac{x^2}{q^2} \cdot \frac{N_1 N_2}{D_1 D_2}\right) \quad (53) \end{aligned}$$

and where the $O(\cdot)$ term corresponds to the contribution of the characters with $\chi^2 = \chi_0$. Note that this contribution is acceptable since

$$\frac{x^2}{q^2} \cdot \frac{N_1 N_2}{D_1 D_2} \ll x^{2\varepsilon/3} D_1 D_2 \ll x^{2\varepsilon/3} z \ll x^{-\varepsilon/3} \sqrt{qx}.$$

Now we express each of the sums in (53) using a contour integral, and using Hölder's inequality this allows us to bound (53) by

$$\frac{x^2}{q^2} \cdot \frac{\sqrt{N_1 N_2 D_1 D_2}}{D_1^2 D_2^2} \sum_{\chi} \int_{|u| \leq x^{\varepsilon/3}} |L(\frac{1}{2} + iu, \chi)|^4 du + x^{-\varepsilon/3} \sqrt{qx}.$$

By the fourth moment bound (Lemma 4) the first term is

$$\ll \frac{x^2}{q^2} \cdot \frac{\sqrt{N_1 N_2 D_1 D_2}}{D_1^2 D_2^2} q x^{\varepsilon/2} \ll x^{5\varepsilon/6} \frac{x}{\sqrt{D_1 D_2}} \ll x^{-\varepsilon/6} \sqrt{q x}$$

since $D_1 D_2 \geq x^{1+2\varepsilon}/q$.

6.3.4 Case $D_1 D_2 < x^{1+2\varepsilon}/q$. In this case we notice that since $D_1, D_2 > x^{1/2-\varepsilon/6}/\sqrt{q}$ we have $D_1, D_2 \leq (x/q)^{1/2} x^{3\varepsilon}$ and in particular $N_1, N_2 \ll x^{7\varepsilon}$. We notice that if $(n_1, n_2) \notin \mathcal{M}$ and $n_1 d_1^2 \equiv n_2 d_2^2 \pmod{q}$ then $n_1 d_1^2 = n_2 d_2^2 + q\ell$ with $0 < |\ell| \ll x^{1+13\varepsilon}/q^2$. We now fix n_1, n_2, ℓ —there are $\ll x^{1+27\varepsilon}/q^2$ possible choices. We shall show that the number of solutions in $|d_1|, |d_2| \ll (x/q)^{1/2} x^{3\varepsilon}$ to $n_1 d_1^2 - n_2 d_2^2 = q\ell$ is bounded by $\ll x^{9\varepsilon}$ which will be sufficient.

First of all note that we can assume that $(n_1, n_2, q\ell) = 1$. Indeed, q cannot divide $n_1 n_2$ as $n_1 n_2 = o(q)$, and so letting $g = (n_1, n_2, q\ell)$ we have $g \mid \ell$ and the problem reduces to one where (n_1, n_2, ℓ) is replaced with $(n'_1, n'_2, \ell') = (n_1, n_2, \ell)/g$ and now $(n'_1, n'_2, q\ell') = 1$.

Notice that $f(x_1, y_1) = n_1 x_1^2 - n_2 y_1^2$ is a primitive binary quadratic form with discriminant $d = 4n_1 n_2 > 0$. Denote by $\varepsilon_{n_1 n_2}$ the real number $x_0/2 + y_0 \sqrt{n_1 n_2}$ where (x_0, y_0) is the solution in positive integers to the equation $x_0^2 - 4n_1 n_2 y_0^2 = 4$ for which $x_0 + y_0 \sqrt{d}$ is least. Note that $\varepsilon_{n_1 n_2} \geq 3/2$.

Let (x_1, y_1) be a solution to $f(x_1, y_1) = q\ell$ with $x_1, y_1 \ll (x/q)^{1/2} x^{3\varepsilon}$. We notice that in this situation

$$(x_1, y_1) \in \bigcup_{1 \leq m \leq \log x} T_m^+ \cup T_m^-$$

where

$$T_m^+ = \left\{ (x, y) \in \mathbb{Z}^2 : f(x, y) = q\ell \text{ and } \sqrt{n_1} x > \sqrt{n_2} y \text{ and } \varepsilon_{n_1 n_2}^{2m-2} \leq \left| \frac{\sqrt{n_1} x + \sqrt{n_2} y}{\sqrt{n_1} x - \sqrt{n_2} y} \right| < \varepsilon_{n_1 n_2}^{2m} \right\}$$

and

$$\begin{aligned} T_m^- &= \left\{ (x, y) \in \mathbb{Z}^2 : f(x, y) = q\ell \text{ and } \sqrt{n_1} x < \sqrt{n_2} y \text{ and } \varepsilon_{n_1 n_2}^{2m-2} \leq \left| \frac{\sqrt{n_1} x + \sqrt{n_2} y}{\sqrt{n_1} x - \sqrt{n_2} y} \right| < \varepsilon_{n_1 n_2}^{2m} \right\} \\ &= \{(x, y) \in \mathbb{Z}^2 : (-x, -y) \in T_m^+\}. \end{aligned}$$

The reason for this is that $\sqrt{n_1} x_1 + \sqrt{n_2} y_1 \ll x^{7\varepsilon} (x/q)^{1/2}$ and

$$|\sqrt{n_1} x_1 - \sqrt{n_2} y_1| = \frac{q\ell}{\sqrt{n_1} x_1 + \sqrt{n_2} y_1} \gg \frac{q^{3/2}}{x^{1/2+7\varepsilon}} \gg 1.$$

Moreover by Lemma 13 of [MW02] we have $\#T_m^+ = \#T_1^+$ for all $m \geq 1$, and trivially $\#T_m^- = \#T_m^+$ for all $m \geq 1$.

The solutions belonging to T_1^+ are *primary* for the quadratic form $n_1 x_1^2 - n_2 x_2^2$ of discriminant $4n_1 n_2$ (see p. 101 of [SW06] for the definition of primary). By Theorem

4.1 of [SW06] the number of (x_1, y_1) for which there exists a quadratic form g of discriminant $4n_1n_2$ such that $g(x_1, y_1) = q\ell$ and such that (x_1, y_1) is primary for g , is either 0 or given by

$$m \prod_{p|m} \left(1 - \frac{1}{p} \left(\frac{4n_1n_2/m^2}{p} \right) \right) \cdot \sum_{k|\frac{q\ell}{m^2}} \left(\frac{d_0}{k} \right),$$

for particular integers m and d_0 with $m^2 \mid (q\ell, 4n_1n_2)$. Using the divisor bound $\#\{k: k \mid n\} \ll_\varepsilon n^{\varepsilon/100}$, we find that this is

$$\ll (n_1n_2)^{1/2+\varepsilon/100} (q\ell)^{\varepsilon/100} \ll x^{8\varepsilon}.$$

We conclude therefore that $\#T_1^+ \ll x^{8\varepsilon}$ and therefore the number of solutions (x_1, y_1) with $|x_1|, |y_1| \ll (x/q)^{1/2} x^{3\varepsilon}$ to the equation $f(x_1, y_1) = q\ell$ is bounded by $\ll \log x \cdot \#T_1^+ \ll x^{9\varepsilon}$ as claimed. It follows therefore that the total number of solutions to $n_1d_1^2 - n_2d_2^2 = q\ell$ with $n_i \sim N_i$, $d_i \sim D_i$ for $i = 1, 2$ is $\ll x^{1+36\varepsilon}/q^2$.

We conclude therefore that (52) is

$$\ll \frac{x^2}{q} \cdot \frac{1}{D_1^2 D_2^2} \cdot \frac{x^{1+36\varepsilon}}{q^2} \ll \frac{x^{1+40\varepsilon}}{q} \ll x^{-\varepsilon} \sqrt{qx}$$

since $q > x^{1/3+30\varepsilon}$. □

7 The range $z \geq (x/q)^{4/3+\varepsilon}$ in the q -aspect: Proof of Proposition 4

Splitting into dyadic segments and recalling (23), we can bound the left-hand side of (15) by a constant times

$$\log x \sup_{\sqrt{z} \leq D \leq \sqrt{x}} \frac{1}{\varphi(q)^2} \sum_{\chi \neq \chi_0} \left| \sum_{\substack{nd^2 \leq x \\ d \sim D}} \mu(d) \chi(n) \chi(d^2) \right|^2.$$

Expressing the condition $nd^2 \leq x$ using a contour integral (see [MV07, Cor. 5.3]) the above is bounded by

$$\ll \log x \sup_{\sqrt{z} \leq D \leq \sqrt{x}} \frac{1}{\varphi(q)^2} \sum_{\chi \neq \chi_0} \left| \int_{|t| \leq x} L\left(\frac{1}{2} + it, \chi\right) M(1 + 2it, \chi^2) \cdot \frac{x^{1/2+it}}{1/2+it} dt \right|^2 + O((x/q)^{1/2-\varepsilon/8}),$$

(in fact a better error term can be obtained but we do not need to keep track of it) where

$$M(1 + 2it, \chi^2) = \sum_{d \sim D} \frac{\mu(d) \chi^2(d)}{d^{1+2it}}.$$

Applying Cauchy–Schwarz and splitting according to the values of t we can bound the main term above as

$$\ll x(\log x)^3 \sup_{\substack{\sqrt{z} \leq D \leq \sqrt{x} \\ 1 \leq T \leq x}} \frac{1}{\varphi(q)^2} \sum_{\chi \neq \chi_0} \frac{1}{T} \int_{-T}^T |L(\tfrac{1}{2} + it, \chi)|^2 \cdot |M(1 + 2it, \chi^2)|^2 dt. \quad (54)$$

Let us now prove the claim on the assumption of the Generalized Lindelöf Hypothesis. Applying Generalized Lindelöf and then the hybrid mean-value theorem (Lemma 7) we have for any choice of $\delta > 0$,

$$\begin{aligned} & \frac{x}{\varphi(q)^2} \sum_{\chi \neq \chi_0} \frac{1}{T} \int_{-T}^T |L(\tfrac{1}{2} + it, \chi)|^2 |M(1 + 2it, \chi^2)|^2 dt \\ & \ll \frac{x(qT)^\delta}{q^2 T} \sum_{\chi \neq \chi_0} \int_{-T}^T |M(1 + 2it, \chi^2)|^2 dt \\ & \ll \frac{xT^\delta q^{2\delta}}{q^2 T} (qT + D) \cdot \frac{1}{D} \ll T^\delta q^{2\delta} \left(\frac{x}{qD} + \frac{x}{q^2 T} \right). \end{aligned} \quad (55)$$

Since $q, T \leq x \leq (x/q)^{O(1)}$, for sufficiently small δ we have $T^\delta q^{2\delta} \leq (x/q)^{\varepsilon/100}$. Recalling also that $D \geq z^{1/2} \geq (x/q)^{(1+\varepsilon)/2}$ and $q \geq x^{1/3+\varepsilon}$, we see that (55) is

$$\ll x^{\varepsilon/100} \left(\left(\frac{x}{q} \right)^{1/2-\varepsilon/2} + \frac{x}{q^2} \right) \ll (x/q)^{1/2-\varepsilon/3}.$$

Applying this estimate to (54) yields the claim.

Let us now consider the unconditional part of the claim. Let

$$S_{T,q}(V) := \{(t, \chi) : V \leq |M(1 + 2it, \chi^2)| \leq 2V \quad |t| \leq T \quad \chi \pmod{q}\}.$$

Note that for $D \geq \sqrt{z} \geq (x/q)^{1/2+\varepsilon}$, the values of $t \in [-T, T]$ for which $|M(1 + 2it, \chi^2)| \leq D^{-1/2+\varepsilon/4}$ contribute to (54) by Cauchy–Schwarz and the fourth moment bound (Lemma 4) $O((\log x)^5 x D^{-1+\varepsilon/2}/q) = O((x/q)^{1/2-\varepsilon/2})$. Additionally, $|M(1 + 2it)| \leq \sum_{d \sim D} 1/d \leq 2$. Therefore it suffices to show that for each $\sqrt{x} \geq D \geq \sqrt{z}$, $V \in [D^{-1/2}, 1]$, and $T \in [1, x]$, we have

$$\frac{xV^2}{\varphi(q)} \sum_{\chi \neq \chi_0} \frac{1}{T} \int_{t: (t, \chi) \in S_{T,q}(V)} |L(\tfrac{1}{2} + it, \chi)|^2 dt \ll (x/q)^{-\varepsilon/8} (\log x)^{-4} \cdot \sqrt{qx}. \quad (56)$$

By Lemma 2 we have,

$$|S_{T,q}(V)| \ll (V^{-2} + qT \min\{D^{-1}V^{-2}, D^{-2}V^{-6}\}) \cdot (\log x)^{18}. \quad (57)$$

Here $|S_{T,q}(V)|$ is the measure of $S_{T,q}(V)$, where the set of $\chi \pmod{q}$ is endowed with the counting measure. Consider first the case when the first term dominates. Then, by Lemma 6, we see that the left-hand side of (56) is

$$\ll \frac{xV^2}{\varphi(q)} \cdot \frac{|S_{T,q}(V)|}{T} \cdot (qT)^{1/3+\varepsilon/4} \ll \frac{x}{\varphi(q)} \cdot \frac{1}{T} \cdot (qT)^{1/3+\varepsilon/3}$$

$$\ll \frac{x}{q^{2/3-\varepsilon/2}} \ll (x/q)^{-\varepsilon/8} (\log x)^{-4} \cdot \sqrt{qx}$$

since $q > x^{3/7+\varepsilon}$. Note that the factor $(\log x)^{18}$ in (57) was absorbed in the exponent of qT .

Consider now the case that the second term dominates in (57). Then by Cauchy–Schwarz and the hybrid fourth moment estimate (Lemma 4),

$$\begin{aligned} & \frac{xV^2}{\varphi(q)} \sum_{\chi \neq \chi_0} \frac{1}{T} \int_{t: (t,\chi) \in S_{T,q}(V)} |L(\tfrac{1}{2} + it, \chi)|^2 dt \\ & \ll \frac{xV^2}{T\varphi(q)} \cdot |S_{T,q}(V)|^{1/2} \cdot \left(\sum_{\chi \neq \chi_0} \int_{-T}^T |L(\tfrac{1}{2} + it, \chi)|^4 dt \right)^{1/2} \\ & \ll (\log x)^{11} \cdot xV^2 \cdot \min\{D^{-1}V^{-2}, D^{-2}V^{-6}\}^{1/2} \\ & \ll (\log x)^{11} \cdot x \min\{D^{-1/2}V, D^{-1}V^{-1}\} \\ & \ll (\log x)^{11} \cdot x \cdot (D^{-1/2}V)^{1/2} \cdot (D^{-1}V^{-1})^{1/2} \\ & \ll x(\log x)^{11} \cdot D^{-3/4} \ll (x/q)^{-\varepsilon/8} (\log x)^{-3} \sqrt{qx} \end{aligned}$$

since $D \geq \sqrt{z} \geq (x/q)^{2/3+\varepsilon/2}$.

8 Conditional estimates: Proof of Theorem 3

The proof of Theorem 3 splits into two parts since two assertions are made.

8.1 Proof that the Riemann Hypothesis implies (7). The proof follows the same ideas as the proof of Proposition 2. The claim (7) is already proved for $H \leq X^{2/3-\varepsilon}$, so we may assume $H > X^{2/3-\varepsilon}$. We return to (41) and consider first the case $D \geq H^{(1-\delta)/2}$. Note that the Riemann Hypothesis implies

$$M(1+2it) \ll_{\delta} D^{-1/2+\delta/2}, \quad (58)$$

for $|t| \leq X^2$. Now (41), Cauchy–Schwarz and the fourth moment bound for the Riemann zeta function (Lemma 3) imply

$$\frac{1}{X} \int_X^{2X} \left| \sum_{\substack{x < nd^2 \leq x+H \\ d \sim D}} \mu(d) - H \sum_{d \sim D} \frac{\mu(d)}{d^2} \right|^2 dx \ll_{\delta} (\log X)^2 H / D^{1-\delta}.$$

For $D \geq H^{(1-\delta)/2}$, the right-hand side is $\ll_{\delta} (\log X) H^{1/2+\delta-\delta^2/2}$. Splitting dyadically for $D \in [H^{(1-\delta)/2}, X^{1/2}]$ and using the tail bound

$$\sum_{d^2 > 2X} \frac{\mu(d)}{d^2} \ll_{\delta} \frac{1}{X^{3/4-\delta/10}},$$

valid under Riemann Hypothesis, we see that the Riemann Hypothesis implies

$$\begin{aligned} & \frac{1}{X} \int_X^{2X} \left| \sum_{\substack{x < nd^2 \leq x+H \\ d^2 \geq H^{1-\delta}}} \mu(d) - H \sum_{d^2 \geq H^{1-\delta}} \frac{\mu(d)}{d^2} \right|^2 dx \\ & \ll (\log X)^2 \sup_{H^{(1-\delta)/2} \leq D \leq X^{1/2}} \frac{1}{X} \int_X^{2X} \left| \sum_{\substack{x < nd^2 \leq x+H \\ d \sim D}} \mu(d) - H \sum_{d \sim D} \frac{\mu(d)}{d^2} \right|^2 dx \quad (59) \\ & + \frac{1}{X} \int_X^{2X} \left| H \sum_{d^2 > 2X} \frac{\mu(d)}{d^2} \right|^2 dx \ll_{\delta} H^{1/2+\delta}. \end{aligned}$$

On the other hand, estimating the n -sum on the left-hand side by $H/d^2 + O(1)$, we see that

$$\frac{1}{X} \int_X^{2X} \left| \sum_{\substack{x < nd^2 \leq x+H \\ d^2 \leq H^{1/2}}} \mu(d) - H \sum_{d^2 \leq H^{1/2}} \frac{\mu(d)}{d^2} \right|^2 dx \ll H^{1/2}.$$

Hence the claim follows once we have shown that, for any $D \in [H^{1/4}, H^{(1-\delta)/2}]$, we have

$$\frac{1}{X} \int_X^{2X} \left| \sum_{\substack{x < nd^2 \leq x+H \\ d \sim D}} \mu(d) - H \sum_{d \sim D} \frac{\mu(d)}{d^2} \right|^2 dx \ll_{\delta} H^{1/2}.$$

Notice that we can attach to the n variable a dummy function $f(nD^2/X)$ with f a smooth function supported in $[1/20, 20]$ and such that $f(y) = 1$ for $y \in [1/10, 10]$.

Similarly to the proof of Proposition 2, write

$$A(x) := \sum_{\substack{nd^2 \leq x \\ d \sim D}} f\left(\frac{n}{X/D^2}\right) \mu(d) - x \sum_{d \sim D} \frac{\mu(d)}{d^2}.$$

By contour integration we have, for $e^y \in [X, 2X]$ and $w \leq 1/100$,

$$A(e^{y+w}) - A(e^y) = \frac{1}{2\pi i} \int_{1/2-i\infty}^{1/2+i\infty} e^{ys} \cdot \frac{e^{ws} - 1}{s} N_1(s) M(2s) ds - e^y (e^w - 1) \sum_{d \sim D} \frac{\mu(d)}{d^2}, \quad (60)$$

where

$$M(s) := \sum_{d \sim D} \frac{\mu(d)}{d^s} \quad \text{and} \quad N_1(s) := \sum_m \frac{1}{m^s} \cdot f\left(\frac{m}{X/D^2}\right).$$

Write also

$$N_2(s) := \int_{\mathbb{R}} \frac{1}{u^s} \cdot f\left(\frac{u}{X/D^2}\right) du$$

and note that, for $e^y \in [X, 2X]$, we have by contour integration

$$\begin{aligned} \frac{1}{2\pi i} \int_{1/2-i\infty}^{1/2+i\infty} e^{ys} \cdot \frac{e^{ws} - 1}{s} N_2(s) M(2s) ds &= \sum_{d \sim D} \mu(d) \int_{\mathbb{R}} f\left(\frac{u}{X/D^2}\right) 1_{e^y \leq ud^2 \leq e^{y+w}} du \\ &= e^y (e^w - 1) \sum_{d \sim D} \frac{\mu(d)}{d^2}. \end{aligned}$$

Plugging this into (60) and arguing as in the proof of Proposition 2, we see that, for some $w \asymp H/X$, we have

$$\begin{aligned} &\frac{1}{X} \int_X^{2X} |A(x+H) - A(x)|^2 dx \\ &\ll X \int_{\mathbb{R}} \left| \frac{e^{w(\frac{1}{2}+it)} - 1}{\frac{1}{2}+it} \right|^2 \cdot |(N_1(\frac{1}{2}+it) - N_2(\frac{1}{2}+it)) M(1+2it)|^2 dt. \end{aligned} \tag{61}$$

By Poisson summation

$$\begin{aligned} N_1(\frac{1}{2}+it) &= \sum_m \frac{1}{m^{1/2+it}} \cdot f\left(\frac{m}{X/D^2}\right) = \sum_{\ell} \int_{-\infty}^{\infty} \frac{1}{u^{1/2+it}} \cdot f\left(\frac{u}{X/D^2}\right) e(\ell u) du \\ &= N_2(\frac{1}{2}+it) + \frac{X}{D^2} \sum_{\ell \neq 0} \int_{1/20}^{20} \left(\frac{D^2}{yX}\right)^{1/2+it} f(y) e\left(\frac{\ell y X}{D^2}\right) dy. \end{aligned}$$

By partial integration (taking antiderivatives of $e(\ell y X/D^2)$), this implies that, for $|t| < X/D^{2+\delta/100}$,

$$|N_1(\frac{1}{2}+it) - N_2(\frac{1}{2}+it)| \ll_A X^{-A},$$

for any $A > 0$. Therefore the part of the integral (61) with $|t| < X/D^{2+\delta/100}$ is completely negligible.

On the other hand the part with $|t| \geq X^{10}$ contributes only $O(1)$ to the left-hand side of (61) by estimating $|N_j(1/2+it)|$ and $|M(1+it)|$ trivially.

Furthermore, assuming the Riemann Hypothesis, we have by contour integration, for $|t| \in [X/D^{2+\delta/100}, X^{10}]$,

$$|N_1(\frac{1}{2}+it)| \ll \sup_{|t|/2 \leq |u| \leq 2X^{10}} |\zeta(\frac{1}{2}+iu)| \ll_{\delta} X^{\delta/100}$$

and

$$|M(1+2it)| \ll_{\delta} D^{-1/2+\delta/100}.$$

Furthermore, by partial integration we have, for $|t| \in [X/D^{2+\delta/100}, X^{10}]$,

$$|N_2(\frac{1}{2}+it)| \ll_{\delta} X^{\delta/100}$$

Hence the part with $|t| \in [X/D^{2+\delta/100}, X^{10}]$ contributes to (61)

$$\ll_{\delta} X \int_{X/D^{2+\delta/100} \leq |t| \leq X^{10}} \frac{1}{|t|^2} X^{\delta/50} D^{-1+\delta/50} dt \ll_{\delta} DX^{\delta/10} \ll_{\delta} H^{1/2}$$

since $D \leq H^{(1-\delta)/2}$ and $H \geq X^{1/2}$. \square

8.2 Proof that (7) implies the Riemann Hypothesis. Suppose that (7) holds for $H = X^{1-\delta}$. Then, by Cauchy–Schwarz,

$$\int_{\mathbb{R}} \Phi\left(\frac{x}{X}\right) \left(\frac{1}{H} \sum_{x < m \leq x+H} \mu^2(m)\right) dx = \frac{6X\widehat{\Phi}(0)}{\pi^2} + O_{\delta}(X^{1/4+3\delta}).$$

with Φ an arbitrary, but not identically zero smooth function compactly supported in $[1/2, 3]$ (one could even enforce that $\widehat{\Phi}(0) = 0$ to simplify the above expression but we didn't find any significant advantage in doing this). Therefore,

$$\frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} \frac{\zeta(s)}{\zeta(2s)} X^s \cdot \Psi_{H/X}(s) ds - \frac{6X\widehat{\Phi}(0)}{\pi^2} = O_{\delta}(X^{1/4+3\delta}), \quad (62)$$

where uniformly in $1/100 < \Re s < 100$, for any given $A > 1$,

$$\begin{aligned} \Psi_{H/X}(s) &:= \frac{1}{s} \cdot \frac{X}{H} \int_{\mathbb{R}} \left(\Phi\left(x - \frac{H}{X}\right) - \Phi(x)\right) x^s dx \\ &= \frac{1}{s} \sum_{1 \leq j \leq A} \frac{(-1)^j}{j!} \cdot \left(\frac{H}{X}\right)^{j-1} \int_{\mathbb{R}} \Phi^{(j)}(x) x^s dx + O_A(X^{-\delta A}) \\ &= -\frac{1}{s} \int_{\mathbb{R}} \Phi'(x) x^s dx + O_A\left(\frac{H}{X} \cdot (1 + |\Im s|)^{-A} + X^{-\delta A}\right). \end{aligned} \quad (63)$$

By integration by parts the main term is equal to $\widetilde{\Phi}(s)$ where $\widetilde{\Phi}(s)$ is the Mellin transform of Φ . The reader may also verify that we have the exact relation $\Psi_{H/X}(1) = \widetilde{\Phi}(1)$.

Suppose that the Riemann Hypothesis fails. Then $\zeta(s)/\zeta(2s)$ has a pole in the strip $\frac{1}{4} < \sigma < \frac{1}{2}$ (e.g. $s = \rho/2$ with $\rho = \beta + i\gamma$ the zeros of $\zeta(s)$ with smallest $\gamma > 0$ among all zeros of $\zeta(s)$ with $\beta \in (\frac{1}{2}, 1)$). Let $\Theta > \frac{1}{4}$ denote the supremum of the real part of poles of $\zeta(s)/\zeta(2s)$ lying in the strip $\frac{1}{4} < \sigma < \frac{1}{2}$. Choose $\delta > 0$ to be sufficiently small so that $\frac{1}{4} + 3\delta \leq \Theta - \delta/2$.

Pick now s_0 a pole of $\zeta(s)/\zeta(2s)$ with $\Re s_0 \in (\Theta - \delta/50, \Theta]$ and the smallest positive imaginary part. We can assume without loss of generality that Φ is chosen so that $\widetilde{\Phi}(s_0) \neq 0$. Indeed if it were the case that $\widetilde{\Phi}(s_0) = 0$ then pick a $c \in (0, 1)$ such that $\widetilde{\Phi}(c + s_0) \neq 0$ and consider $x^c \Phi(x)$ in place of $\Phi(x)$.

We shall shift the contour in (62) to the line $\sigma = \Theta + \delta/8$. Note that for any fixed values of X and H , from the definition (63) and integration by parts, we have $\Psi_{H/X}(s) \ll_A (1 + |\Im s|)^{-A}$ uniformly for $1/100 < \Re s < 100$ for all $A \geq 1$. Furthermore for $\Re s \in [\sigma, 2]$ and s bounded away from 1 we get the bound $\zeta(s)/\zeta(2s) \ll_{\delta}$

$(1 + |\Im s|)^C$, where C is a constant which depends only on δ . (This follows because we may bound $\zeta(s)$ using a convexity bound (see e.g. [Tit86, Sec. 5.1]) and we may bound $1/\zeta(2s)$ using the estimate $\log \zeta(2s) \ll_\delta \log(|\Im(2s)| + 2)$ in this region, which follows from a well-known estimate on the logarithmic derivative of the zeta function (e.g. [Tit86, Thm 9.6 (A)]) and the fact that in this region $|2s - \rho'| \geq \delta/8$ whenever $\zeta(\rho') = 0$.) Thus

$$\frac{1}{2\pi i} \int_{(2)} \frac{\zeta(s)}{\zeta(2s)} X^s \cdot \Psi_{H/X}(s) ds = \frac{1}{2\pi i} \int_{(\sigma)} \frac{\zeta(s)}{\zeta(2s)} X^s \cdot \Psi_{H/X}(s) ds + \frac{6}{\pi^2} X \Psi_{H/X}(1).$$

Applying (62) and (63) we find

$$\frac{1}{2\pi i} \int_{(\sigma)} \frac{\zeta(s)}{\zeta(2s)} \cdot X^s \tilde{\Phi}(s) ds = O_\delta(X^{\Theta-\delta/2}) + O_\delta(X^{1/4+3\delta}).$$

By choice of $\delta > 0$ the error term is bounded by $O_\delta(X^{\Theta-\delta/2})$. Therefore setting

$$A(X) := \sum_n \mu^2(n) \Phi\left(\frac{n}{X}\right) - \frac{6}{\pi^2} \hat{\Phi}(0) X \cdot \mathbf{1}_{[1,\infty)}(X),$$

we have for $X \geq 1$,

$$A(X) = \frac{1}{2\pi i} \int_{(\sigma)} \frac{\zeta(s)}{\zeta(2s)} \cdot X^s \tilde{\Phi}(s) ds = O_\delta(X^{\Theta-\delta/2}).$$

Thus there exists a constant $c = c(\Theta, \delta)$ such that,

$$|A(x)| \leq cx^{\Theta-\delta/50} \tag{64}$$

for all $x \geq 0$ (note that for $0 < x < 1/100$ we have that $A(x)$ vanishes). Let us start by observing that for $\Re s > 1$,

$$\begin{aligned} \int_0^\infty A(x)x^{-s-1} dx &= \sum_{n \geq 1} \mu^2(n) \int_0^\infty \Phi\left(\frac{n}{x}\right) x^{-s-1} dx - \frac{6\hat{\Phi}(0)}{\pi^2} \cdot \frac{1}{s-1} \\ &= \sum_{n \geq 1} \mu^2(n) \cdot n^{-s} \tilde{\Phi}(s) - \frac{6\tilde{\Phi}(1)}{\pi^2} \cdot \frac{1}{s-1} \\ &= \frac{\zeta(s)}{\zeta(2s)} \cdot \tilde{\Phi}(s) - \frac{6\tilde{\Phi}(1)}{\pi^2} \cdot \frac{1}{s-1}. \end{aligned} \tag{65}$$

The function $\int_0^\infty A(x)x^{-s-1} dx$ is analytic in $\Re s > \Theta - \delta/50$ by (64). Therefore, by (65) and analytic continuation,

$$\frac{\zeta(s)}{\zeta(2s)} \cdot \tilde{\Phi}(s) - \frac{6\tilde{\Phi}(1)}{\pi^2} \cdot \frac{1}{s-1} \tag{66}$$

is analytic in the region $\Re s > \Theta - \delta/50$. This however contradicts that (66) has a pole at s_0 and $\Re s_0 \in (\Theta - \delta/50, \Theta]$. \square

Acknowledgments

We would like to thank Bingrong Huang and Francesco Cellarosi for useful conversations, and the anonymous referees for their helpful comments. OG was supported by the European Research Council (ERC) under the European Union's 2020 research and innovation programme (ERC Grant Agreement No. 786758). KM was supported by Academy of Finland Grant No. 285894. MR acknowledges partial support of a Sloan fellowship and of NSF Grant DMS-1902063. BR received partial support from NSF Grant DMS-1854398 and an NSERC grant. Parts of this research were done during visits to Centre de Recherches Mathématiques and Oberwolfach and we thank these institutions for their hospitality.

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/>.

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

References

- [ACS17] M. AVDEEVA, F. CELLAROSI, and Y. G. SINAI. Ergodic and statistical properties of \mathfrak{B} -free numbers. *Theory of Probability & Its Applications*, 61(4):569–589, 2017.
- [AKS04] M. AGRAWAL, N. KAYAL, and N. SAXENA. PRIMES is in P. *Annals of Mathematics (2)*, 160(2):781–793, 2004.
- [Blo08] V. BLOMER. The average value of divisor sums in arithmetic progressions. *Q. J. Math.*, 59(3):275–286, 2008.
- [CE19] D. CARMON and A. ENTIN. On square-free values of large polynomials over the rational function field. *Mathematical Proceedings of the Cambridge Philosophical Society*, page 1–17, 2019.
- [Cro75] M. J. CROFT. Square-free numbers in arithmetic progressions. *Proc. London Math. Soc. (3)*, 30:143–159, 1975.
- [CS13] F. CELLAROSI and Y. G. SINAI. Ergodic properties of square-free numbers. *Journal of the European Mathematical Society*, 15(4):1343–1374, 2013.
- [FT92] M. FILASETA and O. TRIFONOV. On gaps between squarefree numbers. II. *Journal of the London Mathematical Society (2)*, 45(2):215–221, 1992.
- [GH91] G. R. GRIMMETT and R. R. HALL. The asymptotics of random sieves. *Mathematika*, 38(2):285–302, 1991.
- [GR14] I. S. GRADSHTEYN and I. M. RYZHIK. *Table of integrals, series, and products*. Academic Press, New York, 2014.

- [Gra98] A. GRANVILLE. *ABC* allows us to count squarefrees. *Internat. Math. Res. Notices*, (19):991–1009, 1998.
- [Hal82] R. R. HALL. Squarefree numbers on short intervals. *Mathematika*, 29(1):7–17, 1982.
- [Har07] G. HARMAN. *Prime-detecting sieves*, volume 33 of *London Mathematical Society Monographs Series*. Princeton University Press, Princeton, NJ, 2007.
- [HB82] D. R. HEATH-BROWN. Prime numbers in short intervals and a generalized Vaughan identity. *Canadian J. Math.*, 34(6):1365–1377, 1982.
- [IK04] H. IWANIEC and E. KOWALSKI. *Analytic number theory*, volume 53. American Mathematical Society, 2004.
- [KR16] J. KEATING and Z. RUDNICK. Squarefree polynomials and Möbius values in short intervals and arithmetic progressions. *Algebra Number Theory*, 10(2):375–420, 2016.
- [LB18] P. LE BOUDEC. On the distribution of squarefree integers in arithmetic progressions. *Math. Z.*, 290(1-2):421–429, 2018.
- [Liu16] H-Q. LIU. On the distribution of squarefree numbers. *Journal of Number Theory*, 159:202–222, 2016.
- [Mon71] H. L. MONTGOMERY. *Topics in multiplicative number theory*. Lecture Notes in Mathematics, Vol. 227. Springer, Berlin, 1971.
- [Mon01] H. L. MONTGOMERY. Harmonic analysis as found in analytic number theory. In *Twentieth century harmonic analysis—a celebration (Il Ciocco, 2000)*, volume 33 of *NATO Sci. Ser. II Math. Phys. Chem.*, pages 271–293. Kluwer, Dordrecht, 2001.
- [MV77] H. L. MONTGOMERY and R. C. VAUGHAN. Exponential sums with multiplicative coefficients. *Invent. Math.*, 43(1):69–82, 1977.
- [MV07] H. L. MONTGOMERY and R. C. VAUGHAN. *Multiplicative number theory. I. Classical theory*, volume 97 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2007.
- [MW02] H. MUZAFFAR and K. S. WILLIAMS. A restricted Epstein zeta function and the evaluation of some definite integrals. *Acta Arith.*, 104(1):23–66, 2002.
- [Nun15] R. M. NUNES. Squarefree numbers in arithmetic progressions. *J. Number Theory*, 153:1–36, 2015.
- [Nun17] R. M. NUNES. On the least squarefree number in an arithmetic progression. *Mathematika*, 63(2):483–498, 2017.
- [Par19] T. PARRY. A variance for k -free numbers in arithmetic progressions of given modulus. *arXiv preprint arXiv:1912.04683*, 2019.
- [PY19] I. PETROW and M. P. YOUNG. The fourth moment of Dirichlet L -functions along a coset and the Weyl bound. *arXiv preprint arXiv:1908.10346*, 2019.
- [RS92] A. M. ROCKETT and P. SZÜSZ. *Continued fractions*. World Scientific Publishing Co., Inc., River Edge, NJ, 1992.
- [She14] G. SHEVCHENKO. Fractional Brownian motion in a nutshell. *arXiv preprint arXiv:1406.1956*, 2014.
- [SV77] B. SAFFARI and R. C. VAUGHAN. On the fractional parts of x/n and related sequences. II. *Annales de l’institut Fourier (Grenoble)*, 27(2):v, 1–30, 1977.
- [SW06] Z-H. SUN and K. S. WILLIAMS. On the number of representations of n by $ax^2 + bxy + cy^2$. *Acta Arithmetica*, 122(2):101–171, 2006.
- [Tit86] E. C. TITCHMARSH. *The theory of the Riemann zeta-function*. The Clarendon Press, Oxford University Press, New York, second edition, 1986. Edited and with a preface by D. R. Heath-Brown.

- [Tol06] D. I. TOLEV. On the distribution of r -tuples of squarefree numbers in short intervals. *Int. J. Number Theory*, 2(2):225–234, 2006.
- [Tsa85] K. M. TSANG. The distribution of r -tuples of squarefree numbers. *Mathematika*, 32(2):265–275 (1986), 1985.
- [Vau05] R. C. VAUGHAN. A variance for k -free numbers in arithmetic progressions. *Proceedings of the London Mathematical Society*, 91(3):573–597, 2005.
- [War80] R. WARLIMONT. Squarefree numbers in arithmetic progressions. *Journal of the London Mathematical Society*, 2(1):21–24, 1980.

OFIR GORODETSKY

Mathematical Institute, University of Oxford, Oxford OX2 6GG, UK.

ofir.goro@gmail.com

KAISA MATOMÄKI

Department of Mathematics and Statistics, University of Turku, 20014 Turku, Finland.

ksmato@utu.fi

MAKSYM RADZIWIŁŁ

Department of Mathematics, Caltech, 1200 E California Blvd, Pasadena, CA 91125, USA.

maksym.radziwill@gmail.com

BRAD RODGERS

Department of Mathematics and Statistics, Queen's University, Kingston, ON K7L 3N6, Canada.

brad.rodgers@queensu.ca

Received: June 26, 2020

Revised: November 18, 2020

Accepted: December 9, 2020