Geom. Funct. Anal. Vol. 33 (2023) 1173–1242
https://doi.org/10.1007/s00039-023-00635-w
Published online June 19, 2023
© 2023 The Author(s)

GAFA Geometric And Functional Analysis

ADDITIVE PROBLEMS WITH ALMOST PRIME SQUARES

Valentin Blomer, Lasse Grimmelt, Junxian Li and
Simon L. Rydin Myerson

Abstract. We show that every sufficiently large integer is a sum of a prime and two almost prime squares, and also a sum of a smooth number and two almost prime squares. The number of such representations is of the expected order of magnitude. We likewise treat representations of shifted primes $p-1$ as sums of two almost prime squares. The methods involve a combination of analytic, automorphic and algebraic arguments to handle representations by restricted binary quadratic forms with a high degree of uniformity.

1 Introduction

1.1 Statement of results. In his first paper, Hooley [Ho1] gave a proof, conditional upon the Generalized Riemann Hypothesis, of a conjecture made by Hardy and Littlewood some 30 years before [HL]: the number of representations of an integer n in the form

$$n = p + x_1^2 + x_2^2 \tag{1.1}$$

in primes p and non-zero integers x_1, x_2 as $n \to \infty$ is asymptotic to

$$\pi\, \mathfrak{S}(n)\mathrm{Li}(n), \qquad \text{where } \mathfrak{S}(n) = \prod_p \Big(1 + \frac{\chi_{-4}(p)}{p(p-1)}\Big) \prod_{p|n} \frac{(p-1)(p-\chi_{-4}(p))}{p^2 - p + \chi_{-4}(p)}. \tag{1.2}$$

In particular every sufficiently large number is representable in this form. The Generalized Riemann Hypothesis (GRH) was removed a few years later by Linnik [Li]

The first and third author were supported in part by DFG Grants BL 915/2-2 and BL 915/5-1 and Germany's Excellence Strategy Grant EXC-2047/1-390685813. The second author received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme, grant agreement no. 851318. The third and fourth author acknowledge support of the Max Planck Institute for Mathematics. The fourth author was supported by DFG project number 255083470 and by a Leverhulme Early Career Fellowship, and received support from the European Reseach Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant ID 648329)

Keywords and phrases: Additive problems, Almost primes, Smooth numbers, Sieves, Binary quadratic forms

Mathematics Subject Classification: 11N36, 11N75, 11P32, 11E16

who used his dispersion method to obtain the first unconditional proof of the Hardy-Littlewood problem. The proof could be greatly streamlined and simplified once the Bombieri-Vinogradov theorem became available, cf. e.g. [Ho2]. The best error term $O_A(n(\log n)^{-A})$ in the asymptotic formula was recently established by Assing, Blomer and Li [ABL] by invoking in addition the theory of automorphic forms.

In his last paper, Hooley [Ho3] returned to this topic and considered a refined version. He showed that every sufficiently large number is representable as a sum of a prime and two squares of *square-free* numbers and gave a lower bound for the number of such representations of the expected order of magnitude. At first sight this appears to be easy—square-free numbers can be detected by zero-dimensional sieve, so only a little bit of inclusion–exclusion is required. On second thought, a major issue emerges: to sieve out even a fixed number of squares of primes, one needs an asymptotic formula for (1.1) with $d_1 \mid x_1$, $d_2 \mid x_2$. However, the quadratic form $d_1^2x_1^2 + d_2^2x_2^2$ fails to have class number one, in fact very soon it even fails to have one class per genus, and hence there is no convolution formula for the number of ways it represents a given integer. Without such a formula, all approaches to the Hardy-Littlewood problem and its variations become very problematic. Hooley finds an ingenious (but rather ad hoc) way to circumvent this issue altogether for the particular problem at hand.

In this paper, we continue on this avenue. We introduce a general sieve process into the problem with the goal of showing that every sufficiently large integer is a sum of a prime and two squares of *almost primes*. The density of such representations requires us to have a good error term (a saving of at least $(\log x)^{3+\epsilon}$) in solutions to (1.1), which ultimately requires the study of primes in arithmetic progressions to large moduli that go beyond the range that GRH could handle individually. More importantly, in stark contrast to Hooley's result, the application of a sieve to make x_1, x_2 almost prime requires the coefficients d_1 and d_2 to be as large as some fixed power of n. In addition to the fact that the quadratic forms $d_1^2x_1^2 + d_2^2x_2^2$ fail to have class number one, this large range of uniformity for individual d_1, d_2 seems to require inevitably a count of solutions of (1.1) with a power-saving error term, which is only obtainable under some form of GRH.

We will give more details on the combination of algebraic, automorphic and analytic methods involved in a moment, but emphasize at this point that we are nevertheless able to solve this problem unconditionally.

Theorem 1.1. *There exists a constant $C > 0$ such that every sufficiently large integer $n \equiv 1, 3 \bmod 6$ can be represented in the form (1.1), where p is a prime and x_1, x_2 are integers all of whose prime factors are greater than $n^{1/C}$. The number of such representations is $\gg \mathfrak{S}(n)n(\log n)^{-3}$ with $\mathfrak{S}(n)$ as in (1.2). In particular, every sufficiently large integer $n \equiv 1, 3 \mod 6$ can be written as the sum of a prime and two squares of integers with no more than C prime factors.*

Theorem 1.1 gives the expected order of magnitude given the restrictions on x_1, x_2. We have made no effort to compute and optimize the value of C, although an

upper bound of size $C \leq 10^6$, say, can easily be obtained by careful book-keeping. Note that the congruence condition on n is necessary to guarantee solutions to (1.1) with $(x_1x_2, 6) = 1$ and p a prime greater than 3. At the cost of slightly more work it is possible to remove the condition modulo 6 if x_1, x_2 are permitted to have the small primes 2 and 3 as factors. Going further, our method can be generalized to deal with general fixed quadratic forms $F(x_1, x_2)$ of fundamental discriminant, instead of $x_1^2 + x_2^2$, cf. the remark after Lemma 2.2.

The technique used in the proof of Theorem 1.1 also gives the following result.

Theorem 1.2. *There exists a constant $C > 0$ such that the number of solutions to the equation*

$$p = x_1^2 + x_2^2 + 1 \tag{1.3}$$

in primes $p \leq x$ and x_1, x_2 all of whose prime factors are greater than $p^{1/C}$ is $\gg x(\log x)^{-3}$. In particular, there are infinitely many primes shifted by one that can be written as two squares of almost primes.

This should be compared with a beautiful recent result of Friedlander-Iwaniec [FI2] who considered (1.3) for a prime x_1 and an almost prime x_2, but without the additive shift. The multiplicativity in Gaußian primes plays a crucial rule in their proof. Adding a shift parameter ruins the multiplicative structure and so it is no surprise that our proof takes a different route, see Sect. 1.2. Note that Theorem 1.2 is slightly easier than Theorem 1.1 as the shift parameter is fixed (although of course one can prove a similar result with a more general shift f uniformly for $|f| \ll x$).

From a multiplicative point of view, complementary to primes are smooth numbers ("entiers friables" in French) all of whose prime factors are small. The smallness is measured in terms of a parameter y, and following the usual notation we let $\Psi(x, y) = \#S(x, y)$ where $S(x, y)$ denotes the set of positive integers $m \leq x$ such that $p \mid m$ implies $p \leq y$ for all primes p. Smooth numbers can appear quite frequently (e.g $\Psi(x, y) \gg x$ if $\log x \asymp \log y$), yet additive problems with smooth numbers are often surprisingly difficult, see e.g. [Ba, BBD, KMS]. The present case is no exception. Here we consider the equation

$$n = m + x_1^2 + x_2^2 \tag{1.4}$$

where x_1, x_2 are almost primes and m is y-smooth with $y \geq \mathcal{L}(n) := (\log n)^D$ for some sufficiently large constant D. We remark that the smoothing parameter y could be much smaller than those that are admissible in [Ba, BBD, KMS], which makes the problem quite delicate due to the sparseness of smooth numbers when y becomes such small. We have that

$$\Psi(n, y) \asymp nu^{-(1+o(1))u}, \quad u = \frac{\log n}{\log y}.$$

In particular, we have

$$\Psi(n, (\log n)^D) = n^{1-1/D+o(1)}.$$

To capture smooth numbers with such small density, we need to define the quantity

$$\mathcal{F}(n,y)=\prod_{p\mid n}\Big(1-\frac{\chi_{-4}(p)}{p}\Big)\prod_{\substack{p\mid n\\ p\le y}}\Big(1+\frac{\chi_{-4}(p)}{p^{\alpha}}\Big) \tag{1.5}$$

where $\alpha=\alpha(n,y)$ is the unique positive solution to the equation

$$\sum_{p\le y}\frac{\log p}{p^{\alpha}-1}=\log n. \tag{1.6}$$

Theorem 1.3. *There exist constants $C,D>0$ such that for any function g with $(\log n)^D\le g(n)\le n$, every sufficiently large integer n can be represented in the form (1.4) with $m\in S(n,g(n))$ and integers x_1,x_2 all of whose prime factors are greater than $n^{1/C}$. The number of such representations is $\gg \mathcal{F}(n,g(n))\Psi(n,g(n))(\log n)^{-2}$.*

In fact, in the representations we construct, the smooth number m will in addition be a square-free number times a power of 2. For orientation we remark that

$$\alpha=1-\frac{\log(u\log u)+O(1)}{\log y}\gg 1-1/D+o(1)\asymp 1$$

for $y\ge\mathcal{L}(n)$. The arithmetic factor could vary quite a lot when y is small as

$$\exp(-(\log n)^{1/D+o(1)})\ll\mathcal{F}(n,(\log n)^D)\ll\exp((\log n)^{1/D+o(1)})$$

and whilst when $\log y\ge(\log_2 n)^2$, we have $\mathcal{F}(n,y)\asymp 1$.

We mention one possible path for further work, namely Lagrange's four square theorem with restricted variables. That is, one seeks to solve $N=x_1^2+x_2^2+x_3^2+x_4^2$ with x_i restricted to thin subsets of the integers such as primes, almost primes, or smooth numbers. See for instance Brüdern-Fouvry [BF], Heath-Brown-Tolev [HBT] and Blomer-Brüdern-Dietmann [BBD]. It is feasible that the techniques in this paper can be helpful to study related problems.

1.2 Methods and ideas. While the statements of Theorems 1.1 and 1.3 are purely arithmetic, the proofs combine a variety of different methods from different fields. We need the algebraic structure theory of orders in quadratic number fields, the full machinery of automorphic forms including some higher rank L-functions, and of course the toolbox of analytic number theory.

Starting with the latter, the obvious approach to almost primes is the application of a sieve. Thus, for the proof of both theorems we need to analyze quite accurately the quantity

$$\sum_{\ell}r_{d_1,d_2}(n-\ell),\quad r_{d_1,d_2}(n)=\#\{(x,y)\in\mathbb{Z}^2\mid d_1^2x_1^2+d_2^2x_2^2=n\} \tag{1.7}$$

for square-free integers d_1,d_2 where ℓ runs either over primes or over smooth numbers. As mentioned above, in general there is no closed convolution formula for r_{d_1,d_2},

since there is typically more than one class per genus. For fixed d_1, d_2 one can lift the necessary ingredients to analyze (1.7) from [Iw1] or [Vi], but for almost primes we need strong uniformity where d_1, d_2 can be as large as some small power of n.

A robust substitute that may also be useful in other situations is as follows. Let F be a primitive binary quadratic form of discriminant $D = D_0 f^2 < 0$ (with D_0 fundamental) and let $\mathcal{O}_D \subseteq \mathbb{Q}(\sqrt{D})$ be the (not necessarily maximal) order of discriminant D. The set of equivalence classes of primitive forms of discriminant D is in bijection with the proper (i.e. invertible) ideal classes $\mathcal{C}_D$ of $\mathcal{O}_D$ (see [Co, Section 7]), and we denote by $C_F \in \mathcal{C}_D$ the class corresponding to F. Then for $(n, f) = 1$ we have the exact formula

$$\#\{(x_1, x_2) \in \mathbb{Z}^2 \mid F(x_1, x_2) = n\} = \frac{w}{\#\widehat{\mathcal{C}}_D} \sum_{\chi \in \widehat{\mathcal{C}}_D} \bar{\chi}(C_F) \sum_{\substack{N\mathfrak{a}=n \\ \mathfrak{a} \subseteq \mathcal{O}_K}} \chi(\mathfrak{a}) \tag{1.8}$$

where w is the number of units in $\mathcal{O}_D$. In our application, the underlying quadratic field is fixed: for all d_1, d_2 in (1.7) we obtain the Gaußian number field $\mathbb{Q}(i)$, but the discriminant of the order increases with d_1, d_2, in particular the order is highly non-maximal. This will cause technical challenges and involved computations of the corresponding main terms.

The key point is now to split the outer sum into two parts: class group characters χ of order at most two are genus characters and can be computed explicitly by some kind of convolution identity. For a character of order three or more, the inner sum is the n-th Hecke eigenvalue of a cusp form of weight 1 and level $|D|$. In this way, the theory of automorphic forms enters, and (1.8) can be seen as a version of Siegel's mass formula.

The upshot is that—at least morally—we need to understand a sum as in (1.7) in two cases: when the arithmetic function in question is a Fourier coefficients of some cusp form and when the arithmetic function is a Dirichlet convolution of two characters, in other words the Fourier coefficient of an Eisenstein series. In particular, this is yet another example where the originally purely arithmetic problem of representing n in the form (1.1), (1.3) or (1.4) is intimately linked to the theory of automorphic forms. Indeed, here we need to invoke the full power of the spectral theory of automorphic forms.

As an aside we remark that there is a different way to build a bridge to automorphic forms. Instead of applying the algebraic identity (1.8), one could model r_{d_1,d_2} by a suitably truncated expression coming from a formal application of the circle method. More rigorously, one could apply a δ-symbol method à la Heath-Brown [HB] and Duke-Friedlander-Iwaniec [DFI], followed by Poisson summation. This leads to sum of Kloosterman sums (for congruence subgroups of levels roughly $d_1^2 d_2^2$) which after an application of the Kuznetsov formula returns a similar expression of cuspidal and non-cuspidal Fourier coefficients. This can be seen as a double Kloosterman's refinement of the circle method. An application of (1.8), however, seems more direct.

To understand (1.7) when r_{d_1,d_2} is replaced by a cuspidal Fourier coefficient, we refine methods and results established in [ABL, Theorem 1.3] to obtain an unconditional power saving for *individual* pair (d_1, d_2) with the required degree of uniformity. However, the Eisenstein contribution requires much more (or at least rather different) work. As mentioned above, if we were going to handle each pair (d_1, d_2) *separately* for the Eisenstein contribution, we would ultimately require a zero-free strip for the relevant Dirichlet L-functions. To avoid this use of GRH, we instead incorporate the *sum* over d_1, d_2 from the sieve weights into the analysis of primes (or smooth numbers) in long arithmetic progressions (see Proposition 5.1 below). The corresponding generalizations and extensions of [ABL, Theorem 2.1] are given in Propositions 6.1 and 7.1 below. This is of independent interest and should find applications elsewhere.

The proof of Theorem 1.3 features additional difficulties. The arithmetic of smooth numbers is rather subtle, and it turns out that leading term in a suitable asymptotic formula of (1.7) for $\ell \in S(n, g(n))$ is not multiplicative in d_1, d_2. Multiplicativity does not fail by a lot, but we obtain secondary terms involving derivatives of Euler products. This makes the application of a sieve perhaps not impossible, but at least seriously problematic. There are some ways to handle non-multiplicative main terms without developing the theory in full generality (see [FI1, p. 37]). But the fact that we have two variables to sieve makes the monotonicity principle, in particular [FI1, Corollary 5.4/5.5], not applicable. We therefore take a different route and incorporate the sieve weights directly into the analysis. A prototype of this idea can be found in the proof of [FT3, Theorem 4], but such a strategy in the context of additive problems seems to be new. There are a number of interesting but more technical details that we will discuss as they arise.

Roadmap for the rest of the paper In Sect. 2, we give the explicit formula for r_{d_1,d_2} (see Lemma 2.2), which consists of a cuspidal part and an Eisenstein part. To deal with the cuspidal part, in Sect. 3 we prove general bounds for linear forms in Hecke eigenvalues over prime numbers (Theorem 3.3) and smooth numbers (Theorem 3.4) assuming a bilinear estimate (Proposition 3.1) and a τ_3-type estimate (Proposition 3.2) with Hecke eigenvalues, whose proofs are given in Sect. 4. We then give a general averaged bilinear estimates in Sect. 5 which will be used to handle the error terms when applying sieve methods. The proofs of Theorems 1.1/1.2 and 1.3 are given in Sects. 6 and 7 respectively. The most involved part is the rather involved analysis of the main term in the sieving process, which covers Sects. 6.2–6.3 in the prime number case and Sects. 7.4–7.7 in the smooth number case.

2 Algebraic considerations

We recall the classical theory on binary quadratic forms. For $f \in \mathbb{N}$ let $\mathcal{O}_f = \mathbb{Z} + f\mathbb{Z}[i] \subseteq \mathbb{Z}[i]$ be the unique order of discriminant $-4f^2$ (the notation is slightly different than in Sect. 1.2). We have $\mathcal{O}_f^* = \{\pm 1\}$ unless $f = 1$ in which case $\mathcal{O}_f^* =$

$\{\pm 1, \pm i\}$. We write $w_f = \#\mathcal{O}_f^*$. Its class group $\mathcal{C}_f$ is the set of proper (= invertible) fractional $\mathcal{O}_f$-ideals I_f modulo principal ideals P_f. This is a finite group of cardinality [Co, Theorem 7.24]

$$h_f = f \prod_{p|f} \Big(1 - \frac{\chi_{-4}(p)}{p}\Big) \begin{cases} 1/2, & f > 1, \\ 1, & f = 1. \end{cases} \tag{2.1}$$

We have a sequence of isomorphisms. Let $\tilde{\mathcal{C}}_f$ be the set of primitive positive binary quadratic forms with discriminant $-4f^2$ modulo the action of $\mathrm{SL}_2(\mathbb{Z})$. Let I_f^* be the set of fractional $\mathcal{O}_f$-ideals coprime to f (these are automatically proper [Co, Lemma 7.18(ii)]) and P_f^* the subgroup of its principal ideals. Let I^* be the set of fractional $\mathcal{O}_1 = \mathbb{Z}[i]$-ideals coprime to f and P^* the set of principal $\mathbb{Z}[i]$-ideals with a generator $\alpha \in \mathcal{O}_f$ coprime to f. Then we have ([Co, Theorem 7.7(ii), Exercise 7.17(a), Propositions 7.19, 7.20, 7.22] with the convention that $\Im(\beta/\alpha) > 0$)

$$\begin{array}{ccccccc} \tilde{\mathcal{C}}_f & \cong & \mathcal{C}_f = I_f/P_f & \cong & I_f^*/P_f^* & \cong & I^*/P^* \\ {\scriptstyle \frac{N(\alpha x - \beta y)}{N\mathfrak{a}}} & \leftarrow & \mathfrak{a} = [\alpha, \beta] & \leftarrow & \mathfrak{a} & \to & \mathfrak{a}\mathbb{Z}[i] \\ ax^2 + bxy + cy^2 & \to & [a, \tfrac{1}{2}(-b + \sqrt{-4f^2})] & & \mathfrak{a} \cap \mathcal{O}_f & \leftarrow & \mathfrak{a} \end{array} \tag{2.2}$$

The last isomorphism is norm-preserving ([Co, Proposition 7.20]). For a positive integer m and a primitive binary form F with corresponding ideal class C_F, the first isomorphism induces a bijection ([Co, proof of Theorem 7.7(iii)] or [BS, Theorem 5 in Chapter 2, Section 7.6]) between

$$\{(x_1, x_2) \mid F(x_1, x_2) = m\}/\mathcal{O}_f^* \longleftrightarrow \{\mathfrak{a} \text{ integral } \mathcal{O}_f\text{-ideal in the class } C_F \mid N\mathfrak{a} = m\}. \tag{2.3}$$

If $(m, f) = 1$, then the last isomorphism in (2.2) shows

$$\sum_{\substack{N\mathfrak{a}=m \\ \mathfrak{a}\subseteq\mathcal{O}_f}} 1 = \sum_{\substack{N\mathfrak{a}=m \\ \mathfrak{a}\subseteq\mathbb{Z}[i]}} 1 = \sum_{d|m} \chi_{-4}(d) = r(m) \tag{2.4}$$

where $r(m) = \mathbf{1} * \chi_{-4}$ is the usual sums-of-two-squares function.

The elements of order 2 in $\mathcal{C}_f$ are well-understood by genus theory. In particular, for square-free f the real characters χ of $\widehat{\mathcal{C}_f}$ are parametrized by odd divisors $D \mid f$, and for an ideal $\mathfrak{a} \in I_f^*$ (i.e. coprime to f) and a real character associated with the divisor D we have

$$\chi(\mathfrak{a}) = \chi_{D^*}(N\mathfrak{a}), \quad D^* := \chi_{-4}(D)D \in \{\pm D\}. \tag{2.5}$$

We write

$$G_f = \tau(f/(f, 2)) \tag{2.6}$$

for the number of genera.

From now on we focus on the form $F_{a,b} = a^2x_1^2 + b^2x_2^2$. We write $C_{a,b} \in \mathcal{C}_{ab}$ to denote the class corresponding to the form $F_{a,b}$. We will give an explicit formula for r_{d_1,d_2} defined in (1.7). We first notice that we can compute the real characters of $\tilde{\mathcal{C}}_{ab}$ explicitly for the class $C_{a,b}$.

LEMMA 2.1. *Let $a, b \in \mathbb{N}$ be coprime and square-free. Let $F_{a,b}$ be as above and let χ be a real character of $\tilde{\mathcal{C}}_{ab}$. Then $\chi(C_{a,b}) = 1$.*

Proof. Then $F_{a,b}$ is equivalent to the form $(a^2+b^2)x_1^2+2b^2x_1x_2+b^2x_2^2$, which under the isomorphism (2.2) corresponds to an ideal of norm a^2+b^2 coprime to ab, so if χ belongs to the divisor $D \mid ab$, then by (2.5) we have

$$\chi(C_{a,b}) = \chi_{D^*}(a^2+b^2) = 1.$$

REMARK. Alternatively one can explicitly compute the class group structure from the exact sequence [Co, (7.25), (7.27), Exercise 7.30]

$$1 \longrightarrow \{\pm 1\} \longrightarrow (\mathbb{Z}/f\mathbb{Z})^* \times \{\pm 1, \pm i\} \longrightarrow (\mathbb{Z}[i]/f\mathbb{Z}[i])^* \longrightarrow \mathcal{C}_f \longrightarrow 1.$$

For instance, if f is even (and squarefree), one can conclude by a computation based on the Chinese remainder theorem and the fact that the multiplicative group of a finite field is cyclic that $\mathrm{rk}_2(\mathcal{C}_f) = \mathrm{rk}_4(\mathcal{C}_f)$, so that the 2-part of $\mathcal{C}_f$ is a direct product of copies of $\mathbb{Z}/2^{k_j}\mathbb{Z}$ with $k_j \geq 2$ and hence all real characters are trivial on elements of order 2 (which are squares) and so a fortiori on diagonal forms.

From Lemma 2.1 we can give the explicit formula for r_{d_1,d_2} in (1.7). We recall the notation (2.1) and (2.6).

LEMMA 2.2. *Let d_1, d_2 be two square-free numbers, and write $\delta = (d_1, d_2), d_1' = d_1/\delta$, and $d_2' = d_2/\delta$. Let $\mathcal{L} \subseteq \mathbb{N}$ be a finite set and define*

$$\mathcal{S} := \sum_{l \in \mathcal{L}} r_{d_1,d_2}(l).$$

Then $\mathcal{S} = \mathcal{S}_{\leq 2} + \mathcal{S}_{\geq 3}$ where we have

$$\mathcal{S}_{\leq 2} = \sum_{\delta_1\delta_1' = d_1'} \sum_{\delta_2\delta_2' = d_2'} \frac{w_{\delta_1'\delta_2'}}{h_{\delta_1'\delta_2'}} G_{\delta_1'\delta_2'} \sum_{w \in \mathcal{G}_{\delta_1'\delta_2'}} \sum_{\substack{l \equiv 0 \,(\mathrm{mod}\,(\delta\delta_1\delta_2)^2) \\ l \equiv w \,(\mathrm{mod}\,\delta_1'\delta_2') \\ l \in \mathcal{L}}} r\Big(\frac{l}{\delta^2\delta_1^2\delta_2^2}\Big), \tag{2.7}$$

with

$$\mathcal{G}_{\delta_1'\delta_2'} = \big\{w \in (\mathbb{Z}/\delta_1'\delta_2'\mathbb{Z})^* : \chi_{p^*}(w) = 1 \text{ for all } 2 \nmid p \mid \delta_1'\delta_2'\big\}, \tag{2.8}$$

and

$$\mathcal{S}_{\geq 3} = \sum_{\delta_1\delta_1' = d_1'} \sum_{\delta_2\delta_2' = d_2'} \frac{w_{\delta_1'\delta_2'}}{h_{\delta_1'\delta_2'}} \sum_{\substack{\chi \in \widehat{\mathcal{C}}_{\delta_1'\delta_2'} \\ \mathrm{ord}\chi \geq 3}} \bar{\chi}(C_{\delta_1',\delta_2'}) \sum_{\substack{l \equiv 0 \,(\mathrm{mod}\,(\delta\delta_1\delta_2)^2) \\ (l,\delta_1'\delta_2') = 1 \\ l \in \mathcal{L}}} \lambda_\chi\Big(\frac{l}{\delta^2\delta_1^2\delta_2^2}\Big) \tag{2.9}$$

with a Hecke eigenvalue λ_χ of some holomorphic cuspidal newform of weight 1, character χ_{-4} and level dividing $4d_1^2d_2^2$.

Proof. Since δ, d_1', d_2' are pairwise coprime and square-free, we have

$$\mathcal{S} := \sum_{l\in\mathcal{L}} r_{d_1,d_2}(l) = \sum_{\delta^2|l\in\mathcal{L}} r_{d_1',d_2'}\Big(\frac{l}{\delta^2}\Big) = \sum_{\delta_1\delta_1'=d_1'}\sum_{\delta_2\delta_2'=d_2'}\sum_{\substack{(\delta\delta_1\delta_2)^2|l\in\mathcal{L}\\(l,\delta_1'\delta_2')=1}} r_{\delta_1',\delta_2'}\Big(\frac{l}{\delta^2\delta_1^2\delta_2^2}\Big). \tag{2.10}$$

Note that $(\delta_1',\delta_2')=1$, so the form $(\delta_1')^2x_1^2+(\delta_2')^2x_2^2$ is primitive and $f=\delta_1'\delta_2'$ is square-free. Moreover, the argument of $r_{\delta_1',\delta_2'}$ is (nonzero and) coprime to $\delta_1'\delta_2'$. We have by (2.3) and orthogonality of characters

$$r_{a,b}(m) = \frac{w_{ab}}{h_{ab}}\sum_{\chi\in\widehat{\mathcal{C}}_{ab}}\bar\chi(C_{a,b})\sigma(\chi,m),\quad \sigma(\chi,m)=\sum_{N\mathfrak{a}=m}\chi(\mathfrak{a}).$$

If χ is real and belongs to the odd divisor $D\mid\delta_1'\delta_2'$, then by (2.5) and (2.4) for $(m,\delta_1'\delta_2')=1$ we have

$$\sigma(\chi,m)=\chi_{D^*}(m)\#\{N\mathfrak{a}=m\}=\chi_{D^*}(m)r(m).$$

Combining this with Lemma 2.1, we obtain

$$\mathcal{S}_{\le 2} = \sum_{\delta_1\delta_1'=d_1'}\sum_{\delta_2\delta_2'=d_2'}\frac{w_{\delta_1'\delta_2'}}{h_{\delta_1'\delta_2'}}\sum_{2\nmid D|\delta_1'\delta_2'}\sum_{\substack{l\equiv 0 \,(\mathrm{mod}\ (\delta\delta_1\delta_2)^2)\\(l,\delta_1'\delta_2')=1\\ l\in\mathcal{L}}}\chi_{D^*}\Big(\frac{l}{\delta^2\delta_1^2\delta_2^2}\Big)r\Big(\frac{l}{\delta^2\delta_1^2\delta_2^2}\Big),$$

which yields (2.7) after summing over $2\nmid D\mid\delta_1'\delta_2'$.

If $\mathrm{ord}(\chi)\ge 3$, it remains to show that $\sigma(\chi,m)$ is a normalized Hecke eigenvalue of a weight 1 newform of level dividing $4d_1^2d_2^2$ and character χ_{-4}. This is a special case of automorphic induction for imaginary quadratic fields [AC], but most of it can be seen elementarily. We observe first that by splitting into ideal classes the function

$$f_\chi(z)=\sum_{\mathfrak{a}\subseteq\mathcal{O}_{\delta_1'\delta_2'}}\chi(\mathfrak{a})e(N\mathfrak{a}\cdot z)$$

is a linear combination of theta series corresponding to forms $F\in\tilde{\mathcal{C}}_{\delta_1'\delta_2'}$ and hence by [Iw4, Theorem 10.9] a modular form of weight 1, character χ_{-4} and level $4(\delta_1'\delta_2')^2$. Using (2.2), one can check directly from the definition of the Hecke operators T_m that

$$T_mf_\chi(z)=\sigma(\chi,m)f_\chi(z)$$

for $(m,\delta_1'\delta_2')=1$. Hence f_χ belongs to the space generated by some newform of level dividing $4(\delta_1'\delta_1')^2$, and for $(m,\delta_1'\delta_2')=1$ the numbers $\sigma(\chi,m)$ coincide with the Hecke eigenvalues.

It remains to show that f_χ is a cusp form for $\mathrm{ord}(\chi) \geq 3$. This is again a classical fact (also contained in [AC]) that can be proved in several ways: by the parametrization of Eisenstein series, a modular form is non-cuspidal if and only if the associated Dirichlet series factorizes into two Dirichlet L-functions, and this happens if and only if its Rankin–Selberg square L-function has a double pole at $s = 1$. The Rankin–Selberg square of a class group character χ is, up to finitely many Euler factors, $\zeta(s)L(s, \chi^2)$ which has a double pole at $s = 1$ if and only if $\mathrm{ord}(\chi) \leq 2$.

This completes the proof of the lemma.

Remark. The formula (2.7) depends on Lemma 2.1. For general quadratic forms not covered by Lemma 2.1 the summation over D can be carried out in the same way. This yields an expression similar to (2.7) with a set $\mathcal{G}_{\delta_1'\delta_2'}$ which may be different from the one defined in (2.8) but which has the same cardinality, namely $\phi(\delta_1'\delta_2')/G_{\delta_1'\delta_2'}$. This is a consequence of the fact that numbers coprime to the order can only be represented in one genus. More precisely, the Chinese Remainder Theorem induces a bijection $\mathcal{G}_{\delta_1'\delta_2'} \cong \prod_{p|\delta_1'\delta_2'} \mathcal{G}_p$ where $\mathcal{G}_p \subseteq (\mathbb{Z}/p\mathbb{Z})^*$ has cardinality $\phi(p)/2$ for odd p and cardinality 1 for $p = 2$. This is all the information needed in the subsequent arguments. In this way, our main results generalize to other quadratic forms than sums of two squares, although some extra care may be necessary at the last step of (2.10).

If we consider solutions to $n = p + x_1^2 + x_2^2$ where x_1, x_2 have no small prime factors other than a small square-free factor of r, then we will encounter binary quadratic forms of the shape $x^2 + (r^2y)^2$, which can be studied with our method while paying special attention to the real characters of $\mathcal{C}_f$ with non square-free f and the step in (2.10).

3 Additive problems in cusp forms

In this section we deal with the innermost sum of (2.9) when the set $\mathcal{L}$ consists of numbers $n-\ell$ where ℓ is either prime or smooth. We generally consider a holomorphic or Maaß cuspidal Hecke newform ϕ of level $4 \mid N$, character χ_{-4} and Hecke eigenvalues $\lambda(n)$ (which by self-adjointness of Hecke operators are real) whose archimedean parameter (weight or spectral parameter) we denote by μ. We start with two auxiliary results.

Proposition 3.1. *Let $M, X, Z \geq 1$, $\Delta \in \mathbb{N}$, and $f, \sigma \in \mathbb{Z}\backslash\{0\}$. Let α_m, β_n be sequences supported on $[M, 2M]$, $[X, 2X]$ respectively. Suppose that $|\alpha_m| \ll m^\varepsilon$ for any $\varepsilon > 0$ and $\alpha_m\beta_n$ vanishes unless $\sigma mn + f$ is in some dyadic interval $[\Xi, 2\Xi]$ with $|\sigma|MX/Z \ll \Xi \ll |\sigma|MX + |f|$. Then*

$$\sum_{\Delta|\sigma mn+f} \sum \alpha_m\beta_n\lambda\Big(\frac{\sigma mn+f}{\Delta}\Big) \ll P^A\|\beta\|\big(M^{11/14}X^{1/2} + M^{5/4}X^{3/8}\big)^{1+\varepsilon}$$

for some absolute constant $A > 0$ and any $\varepsilon > 0$, where $P = \Delta|\sigma|(1 + |\mu|)NZ(1 + |f|/MX)$.

We remark that the trivial bound is $\|\beta\|(MX^{1/2})^{1+\varepsilon}$, so Proposition 3.1 is non-trivial as soon as $M \ll X^{1/2-\delta}$ for some $\delta > 0$ as in [Pi, Theorem 1.2] when P is fixed.

PROPOSITION 3.2. *Let $\Delta, q, d \in \mathbb{N}, r, f \in \mathbb{Z} \setminus \{0\}$, $Z, X_1, X_2, X_3 \geq 1/2$. Suppose $(\Delta, q) = 1$ and that at least one of r, f is positive. Let G be a function with support on $\times[X_j, 2X_j]$ and $G^{(\nu_1,\nu_2,\nu_3)} \ll_{\boldsymbol{\nu}} Z^{\nu_1+\nu_2+\nu_3} X_1^{-\nu_1} X_2^{-\nu_2} X_3^{-\nu_3}$ for all $\boldsymbol{\nu} \in \mathbb{N}_0^3$. Write $X = |r|X_1X_2X_3$ and suppose that $G(k,l,m)$ vanishes unless $rklm + f$ is in some dyadic interval $[\Xi, 2\Xi]$ with $X/Z \ll \Xi \ll X + |f|$. Then*

$$\sum_{\substack{k,l,m \\ rklm \equiv -f \,(\mathrm{mod}\, \Delta) \\ rklm \equiv d \,(\mathrm{mod}\, q)}} G(k,l,m)\lambda\Big(\frac{rklm+f}{\Delta}\Big) \ll P^A X^{1-1/54+\varepsilon}$$

for some absolute constant A and any $\varepsilon > 0$, where $P = qr\Delta(1+|\mu|)NZ(1+|f|/X)$.

The trivial bound is $(X/|r|\Delta q)^{1+\varepsilon}$, and so Proposition 3.2 is non-trivial uniformly for P up to some small power of X. Better exponents could be obtained with more work, but these bounds suffice for our purpose. In principle this is a shifted convolution problem of τ_3 against λ, cf. e.g. [Mu, To].

We postpone the proofs of Propositions 3.1 and 3.2 to the next section and state the two main results of this section.

Theorem 3.3. *There exist absolute constants $A, \eta > 0$ with the following property. Let $\sigma, f \in \mathbb{Z}\setminus\{0\}$, $X, Z \geq 1$, $\Delta, q, d \in \mathbb{N}$ with $(\sigma q, \Delta) = 1$. Let G be a smooth function supported on $[X, 2X]$ with $G^{(\nu)} \ll_\nu (X/Z)^{-\nu}$ for $\nu \in \mathbb{N}_0$. Suppose that $G(n)$ vanishes unless $\sigma n + f$ is in some dyadic interval $[\Xi, 2\Xi]$ with $|\sigma|X/Z \ll \Xi \ll |\sigma|X + |f|$. Then*

$$\sum_{\substack{\sigma n \equiv -f \,(\mathrm{mod}\, \Delta) \\ n \equiv d \,(\mathrm{mod}\, q)}} \lambda\Big(\frac{\sigma n + f}{\Delta}\Big)\Lambda(n)G(n) \ll P^A X^{1-\eta}$$

where $P = q|\sigma|\Delta(1+|\mu|)NZ(1+|f|/|\sigma|X)$.

Proof. Based on Propositions 3.1 and 3.2, this could be proved similarly as in [ABL, Prop. 9.1] which in turn is the argument of [Pi, Section 9]. The proof in [Pi] uses Vaughan's identity to decompose Λ iteratively three times. This is slightly cumbersome, but has the advantage that we only need α_m to be supported on square-free integers in the bilinear estimate in Proposition 3.1, which could yield better error terms numerically. Here we give a shorter independent proof at the cost of slightly weaker error terms. Instead of using Vaughan's identity as in [Pi], we apply Heath-Brown's identity in the form

$$\Lambda(n) = \sum_{j=1}^{4} (-1)^{j-1} \binom{j}{4} \sum_{m_1,\dots,m_j \leq X^{1/4}} \prod_{i=1}^{j} \mu(m_i) \sum_{m_1\cdots m_j n_1 \cdots n_j = n} \log n_1 \tag{3.1}$$

for $1 \le n \le X$. Next we apply a smooth partition of unity to the variables m_i, n_i. To be precise, let $0 < \mathcal{Z} \le 1$ be a parameter to be chosen later and let $\psi(x)$ be a smooth function supported on $[-1-\mathcal{Z}, 1+\mathcal{Z}]$ that equals 1 on $[-1,1]$ and satisfies $\psi^{(j)}(t) \ll \mathcal{Z}^j$. Let $\mathcal{D} = \{(1+\mathcal{Z})^m, m \ge 0\}$. Then we have the smooth partition of unity

$$1 = \sum_{N \in \mathcal{D}} \psi_N(n) \tag{3.2}$$

for any natural number n, where $\psi_N(n) = \psi(n/N) - \psi((1+\mathcal{Z})n/N)$. Here we can choose $\mathcal{Z} = 1$ (later parts of the paper use the same argument with different choices of $\mathcal{Z}$) so that

$$\sum_{\substack{\sigma n \equiv -f \,(\mathrm{mod}\, \Delta) \\ n \equiv d \,(\mathrm{mod}\, q)}} \lambda\Big(\frac{\sigma n + f}{\Delta}\Big)\Lambda(n)G(n) \ll (\log X)^8 \max_{\substack{M_1,\dots,M_j \\ N_1,\dots,N_j}} |S|$$

where

$$S = \sum_{\substack{\sigma \prod_{i=1}^j m_i n_i \equiv -f \,(\mathrm{mod}\, \Delta) \\ \prod_{i=1}^j m_i n_i \equiv d \,(\mathrm{mod}\, q)}} \prod_{i=1}^j \mu(m_i)\psi_{M_i}(m_i) \log(n_1)\psi_{N_1}(n_1) \prod_{i=2}^j \psi_{N_i}(n_i) \times \lambda\Big(\frac{\sigma \prod_{i=1}^j m_i n_i + f}{\Delta}\Big) G\Big(\prod_{i=1}^j m_i n_i\Big).$$

Therefore it is enough to show that there exist some absolute constants $A, \eta > 0$ such that

$$S \ll P^A X^{1-\eta}. \tag{3.3}$$

We distinguish two cases. If there exists some i such that $X^{\eta_0} \le M_i \,(\text{or } N_i) \le X^{1/4+\eta_0}$ for some $\eta_0 > 0$, then we can split m_i or n_i into residue classes modulo q, use Mellin inversion to separate variables m_i, n_i (see e.g. [Pi, Section 7]) and then apply Proposition 3.1 to obtain

$$S \ll P^A X^{\varepsilon}\big(X^{1-3\eta_0/14} + X^{31/32+3\eta_0/8}\big).$$

Otherwise, we must have

$$R := \prod_{i : M_i \le X^{\eta_0}, N_i \le X^{\eta_0}} M_i N_i \le X^{8\eta_0}$$

and

$$\#\{1 \le i \le j : N_i \ge X^{1/4+\eta_0}\} \le 3, \quad 1 \le j \le 4,$$

in which case we can apply Proposition 3.2 to obtain

$$S \ll P^A X^{1-1/54+8A\eta_0+\varepsilon}.$$

Combining the bounds in these two cases and chossing η_0 suitably, we complete the proof of (3.3).

Our second main result treats the case of smooth numbers. As usual, $P^+(n)$ denotes the largest prime factor of n with the convention $P^+(1) = 1$.

Theorem 3.4. *There exist absolute constants $A, \eta > 0$ with the following property. Let $\sigma, f \in \mathbb{Z} \setminus \{0\}$, $X, Z, y \geq 1$, $\Delta, q, d \in \mathbb{N}$ with $(\sigma q, \Delta) = 1$. Let G be a smooth function supported on $[X, 2X]$ with $G^{(\nu)} \ll_\nu (X/Z)^{-\nu}$. Suppose that $G(n)$ vanishes unless $\sigma n + f$ is in some dyadic interval $[\Xi, 2\Xi]$ with $|\sigma|X/Z \ll \Xi \ll |\sigma|X + |f|$. Then*

$$\sum_{\substack{\sigma n \equiv -f \,(\mathrm{mod}\, \Delta) \\ n \equiv d \,(\mathrm{mod}\, q) \\ P^+(n) \leq y}} \lambda\Big(\frac{\sigma n + f}{\Delta}\Big) G(n) \ll P^A X^{1-\eta} d,$$

where $P = q|\sigma|\Delta(1 + |\mu|)NZ(1 + |f|/|\sigma|X)$.

Proof. The set-up follows along the lines in [FT1, Lemme 2.1] on averages of smooth numbers in arithmetic progressions to large moduli, which we also generalize (see Proposition 7.1). Throughout the proof, we use A to denote some absolute positive constant, not necessarily the same at each occurrence. For $t \geq 1$ let $g_t(n)$ be the characteristic function on numbers n with $P^+(n) \leq t$. The starting point is Buchstab's identity that we iterate. For $y \geq 1$ we have

$$g_y(n) = 1 - \sum_{\substack{p|n \\ p>y}} g_p\Big(\frac{n}{p}\Big) = 1 + \sum_{j=1}^{3} (-1)^j \sum_{\substack{p_1 \cdots p_j | n \\ y < p_1 < \cdots < p_j}} 1 + \sum_{\substack{p_1 \cdots p_4 | n \\ y < p_1 < \cdots < p_4}} g_{p_4}\Big(\frac{n}{p_1 \cdots p_4}\Big). \tag{3.4}$$

Thus

$$\sum_{\substack{\sigma n \equiv -f \,(\mathrm{mod}\, \Delta) \\ n \equiv d \,(\mathrm{mod}\, q) \\ P^+(n) \leq y}} \lambda\Big(\frac{\sigma n + f}{\Delta}\Big) G(n) = \sum_{\substack{\sigma n \equiv -f \,(\mathrm{mod}\, \Delta) \\ n \equiv d \,(\mathrm{mod}\, q)}} \lambda\Big(\frac{\sigma n + f}{\Delta}\Big) g_y(n) G(n)$$

$$= S_0 + \sum_{i=1}^{3} (-1)^i S_i + \overline{S},$$

say, where

$$S_0 = \sum_{\substack{\sigma n \equiv -f \,(\mathrm{mod}\, \Delta) \\ n \equiv d \,(\mathrm{mod}\, q)}} \lambda\Big(\frac{\sigma n + f}{\Delta}\Big) G(n),$$

$$S_i = \sum_{\substack{\sigma n \equiv -f \,(\mathrm{mod}\,\Delta) \\ y<p_1<\cdots<p_i \\ p_1\cdots p_i m \equiv d \,(\mathrm{mod}\, q)}} \lambda\Big(\frac{\sigma p_1\cdots p_i m + f}{\Delta}\Big) G(p_1\cdots p_i m),$$

$$\overline{S} = \sum_{\substack{\sigma n \equiv -f \,(\mathrm{mod}\,\Delta) \\ y<p_1<\cdots<p_4 \\ p_1\cdots p_4 m \equiv d \,(\mathrm{mod}\, q)}} \lambda\Big(\frac{\sigma p_1\cdots p_4 m + f}{\Delta}\Big) g_{p_4}(m) G(p_1\cdots p_4 m).$$

For S_0, we apply [ABL, Corollary 7.6] to obtain

$$S_0 = \sum_{k \equiv \bar{\Delta}(f+\sigma d) \,(\mathrm{mod}\, q)} \lambda(k) G\Big(\frac{\Delta k - f}{\sigma}\Big) \ll P^A. \tag{3.5}$$

For $\overline{S}$, we localize p_i in intervals $(P_j, P_{j+1}]$ where $P_j = y(1+\mathcal{Z})^j$ with some $\mathcal{Z} \le 1$ to be chosen later and $0 \le j \le J = 1 + \lfloor \frac{\log(X/y)}{\log(1+\mathcal{Z})} \rfloor$. Then

$$\overline{S} = \sum_{\substack{0 \le i_1 \le i_2 \le i_3 \le i_4 \le J \\ p_k \in (P_{i_k}, P_{i_k+1}] \\ \sigma p_1\cdots p_4 m \equiv -f \,(\mathrm{mod}\,\Delta) \\ p_1\cdots p_4 m \equiv d \,(\mathrm{mod}\, q)}} \lambda\Big(\frac{\sigma p_1\cdots p_4 m + f}{\Delta}\Big) g_{p_4}(m) G(p_1\cdots p_4 m) + O\Big(P^A X^{1+\varepsilon} \mathcal{Z}^{1/2}\Big). \tag{3.6}$$

To justify this, we see that $\overline{S}$ and the sum in (3.6) differ only by integers n that lie in $[X(1+\mathcal{Z})^{-4}, X(1+\mathcal{Z})^4]$ or have at least two prime factors p_i, p_j with $p_i < p_j < (1+\mathcal{Z})p_i$. The contribution from these integers can be bounded by

$$\begin{aligned}
&\Big(\sum_{X(1+\mathcal{Z})^{-4} < n \le X(1+\mathcal{Z})^4} + \sum_{\substack{p_i < p_j < (1+\mathcal{Z})p_i \\ p_i p_j \mid n \\ n \ll X}} \Big) \Big|\lambda\Big(\frac{\sigma n + f}{\Delta}\Big)\Big| \\
&\ll \Big(\Big(\sum_{X(1+\mathcal{Z})^{-4} < n \le X(1+\mathcal{Z})^4} 1 \Big)^{1/2} + \Big(\sum_{\substack{n \ll X \\ p_i p_j \mid n \\ p_i < p_j \le (1+\mathcal{Z})p_i}} 1 \Big)^{1/2} \Big) \\
&\qquad \times \Big(\sum_{\substack{n \ll X \\ \sigma n \equiv -f \,(\mathrm{mod}\,\Delta)}} \Big|\lambda\Big(\frac{\sigma n + f}{\Delta}\Big)\Big|^2 \Big)^{1/2} \\
&\ll P^A X^{1+\varepsilon} \mathcal{Z}^{1/2}.
\end{aligned}$$

Since $y < P_{i_1} \ll X^{1/4}$, we can apply Proposition 3.1 with $\mathbf{m} = p_1$ and $\mathbf{n} = |\sigma| p_2 \cdots p_4 m$ to the main term in (3.6) by splitting $\mathbf{m}, \mathbf{n}$ into residue classes modulo

q to obtain

$$\begin{aligned}\overline{S} &\ll \mathcal{Z}^{-4}X^{\varepsilon}P^{A}\Big(\frac{X}{y^{3/14}}+X^{31/32}\Big)+P^{A}X^{1+\varepsilon}\mathcal{Z}^{1/2}\\ &\ll P^{A}X^{1+\varepsilon}\max(y^{-1/42},X^{-1/288})\end{aligned} \tag{3.7}$$

upon choosing $\mathcal{Z}=\max(y^{-1/21},X^{-1/144})$.

For S_k, $1\le k\le 3$, we first replace the characteristic function on primes with the von Mangoldt function Λ. As above, we see that the contribution of higher prime powers is at most

$$\ll P^{A}X^{1+\varepsilon}y^{-1/2}$$

which can be absorbed in the existing bounds. Now similar to the proof of Theorem 3.3, we use Heath-Brown's identity (3.1) to decompose the prime variables. Again we split all variables into short intervals using the smooth partition (3.2) for some parameter $\mathcal{Z}$ to be chosen later. We first bound

$$\begin{aligned}\tilde{S}=\sum_{m}\sum_{m_i}\sum_{\substack{n_i\\ \sigma mm_1\cdots m_rn_1\cdots n_s\equiv -f\,(\mathrm{mod}\,\Delta)\\ mn_1\cdots m_rn_1\cdots n_s\equiv d\,(\mathrm{mod}\,q)}}&\psi_M(m)\prod_{i=1}^{r}\mu(m_i)\psi_{M_i}(m_i)\prod_{j=1}^{s}\psi_N(n_j)\\ &\times\lambda\Big(\frac{\sigma mm_1\cdots m_rn_1\cdots n_s+f}{\Delta}\Big)G(mm_1\cdots m_rn_1\cdots n_s)\end{aligned}$$

where $1\le r,s\le 12$, $M\prod_{i=1}^{r}M_i\prod_{j=1}^{s}N_j\asymp X$ and $M_i\ll X^{1/4}$. Again we distinguish two cases.

If there exists a subset $\mathcal{I}\subset\{M,M_i,N_j\}$ such that $K:=\prod_{I\in\mathcal{I}}I$ such that $X^{\eta_0}<K\le X^{1/4+\eta_0}$, then we can apply Proposition 3.1 to obtain

$$\tilde{S}\ll P^{A}X^{\varepsilon}\big(X^{1-3\eta_0/14}+X^{31/32+3\eta_0/8}\big).$$

If such $\mathcal{I}$ does not exist, then all the M_i's and possibly some of the M, N_j's can be combined into $R\ll X^{\eta_0}$ and the number of M,N_j's of size $\gg X^{1/4+\eta_0}$ is less than, or equal to three. So it is enough to bound

$$\sum_{r\ll R|\sigma|}\Big|\sum_{\substack{k,l,m\\ rklm\equiv -f\,(\mathrm{mod}\,\Delta)\\ rklm\equiv d\,(\mathrm{mod}\,q)}}\alpha_r\lambda\Big(\frac{rklm+f}{\Delta}\Big)\psi_K(k)\psi_L(l)\psi_M(m)G(rklm)\Big|$$

which by Proposition 3.2 is

$$\ll P^{A}X^{1-1/54+A\eta_0+\varepsilon}$$

for some absolute constant $A > 0$. Then by the same argument as above we see that

$$S_i \ll P^A X^\varepsilon (\mathcal{Z}^{-24} X^{1-1/54+A\eta_0+\varepsilon} + X\mathcal{Z}^{1/2}), \quad 1 \le i \le 3$$

and hence by choosing $\mathcal{Z} = X^{\eta_1}$ for some suitable $\eta_1 > 0$ we obtain

$$S_i \ll P^A X^{1-\eta_2}, \quad 1 \le i \le 3 \tag{3.8}$$

for some $\eta_2 > 0$. Combining (3.5), (3.7) and (3.8), we have shown that there exist some absolute constants $A, \eta > 0$ such that

$$\sum_{\substack{\sigma n \equiv -f \,(\mathrm{mod}\, \Delta) \\ n \equiv d \,(\mathrm{mod}\, q) \\ P^+(n) \le y}} \lambda\Big(\frac{\sigma n + f}{\Delta}\Big) G(n) \ll P^A X^\varepsilon \big(X^{1-\eta} + X y^{-\eta}\big). \tag{3.9}$$

On the other hand, when y is very small we can do better using the flexible factorization of smooth numbers to create some bilinear structure so that Proposition 3.1 can be applied directly. Recall [FT3, Lemme 3.1]: for any $M \ge 1$, every $n \ge M$ with $P^+(n) \le y$ has a unique representation in the form

$$n = lm, \quad P^+(l) \le P^-(m), \quad M \le m \le yM. \tag{3.10}$$

We can separate l and m in the condition $P^+(l) \le P^-(m)$ using [Dr1, eq. (3.37)] with an acceptable error. Thus by the same argument as before, we can localize m into short intervals and apply Proposition 3.1 to obtain

$$\begin{aligned} \sum_{\substack{\sigma n \equiv -f \,(\mathrm{mod}\, \Delta) \\ n \equiv d \,(\mathrm{mod}\, q) \\ P^+(n) \le y}} \lambda\Big(\frac{\sigma n + f}{\Delta}\Big) G(n) &\ll P^A X^\varepsilon \big(\mathcal{Z}^{-1}(XM^{-3/14} + (My)^{3/8} X^{7/8}) + X\mathcal{Z}^{1/2}\big) \\ &\ll P^A X^{1-1/66+\varepsilon} y^{1/22} \end{aligned} \tag{3.11}$$

by choosing $M = X^{7/33} y^{-7/11}, \mathcal{Z} = X^{1/33} y^{-1/11}$. Applying (3.9) when $y \ge X^\eta$ and (3.11) when $y \le X^\eta$ for some suitable $\eta > 0$ we obtain the theorem.

Remark. Arguing similarly as in the proofs of Theorems 3.3 and 3.4, we can replace $\Lambda(n)$ or $\mathbb{1}_{P^+(n) \le y}$ by $\tau_k(n)$ (for fixed k) and obtain a power saving bound *uniformly* in all parameters. Moreover, the proof works also for a wide class of arithmetic functions which possess a similar combinatorial decomposition (cf. e.g. [DT, FT4] as well as [MT, section 5.1]), such as generalized divisor functions $\tau_z(n)$ for any complex z and the indicator function of norm forms of abelian extensions of $\mathbb{Q}$.

4 Proofs of Propositions 3.1 and 3.2

We still owe the proofs of Propositions 3.1 and 3.2. Recall that P is the product of the "unimportant" parameters which has slightly different meanings in Proposition 3.1 and Proposition 3.2. For notational simplicity we write

$$R \preccurlyeq S \quad :\Longleftrightarrow \quad R \ll X^{\varepsilon} P^{A} S \tag{4.1}$$

for any $\varepsilon > 0$ and some $A > 0$, not necessarily the same at each occurrence.

4.1 Proof of Proposition 3.1. This is an extension of [ABL, Proposition 8.1] where we relax the condition α_m supported on square-free numbers at the cost of slightly weaker bounds. Let $\mathscr{B}(\alpha, \beta)$ denote the m, n-sum we want to bound. We use the notation (4.1) with $P = \Delta|\sigma|(1+|\mu|)NZ(1+|f|/MX)$. Note that we can assume without loss of generality that $\sigma = \{\pm 1\}$ as otherwise we can set $\tilde{\beta}_n = \beta_n \mathbb{1}_{\sigma|n}$. When α_m is supported on square-free numbers, then a straightforward extension of [ABL, Proposition 8.1] (replacing the F in the definition of set of moduli $\mathcal{Q}$ by $[N_1, D^2]\Delta_1$, with the notation $\Delta = \Delta_1\Delta_2$ and $\Delta_2 = (\Delta, (m_3 m_4)^{\infty})$ to accommodate the new parameter Δ) shows

$$\mathscr{B}(\alpha, \beta) \preccurlyeq \|\beta\| \big((XM)^{1/2} + X^{1/4} M^{3/2} + X^{3/8} M^{5/4} \big). \tag{4.2}$$

To deduce the result for general sequences, we write

$$\mathscr{B}(\alpha, \beta) = \sum_{\Delta | \sigma l^2 mn + f} \sum \mu^2(m) \alpha_{ml^2} \beta_n \lambda\Big(\frac{\sigma m l^2 n + f}{\Delta} \Big).$$

We choose a parameter $L > 1$. By Hölder's inequality, we see that the contribution from $l \geq L$ gives

$$\begin{aligned}
&\ll \Big(\sum_{L \leq l \ll M^{1/2}} \sum_{m \ll M/l^2} \sum_{n \asymp X} |\alpha_{ml^2} \beta_n|^{4/3} \Big)^{3/4} \\
&\qquad \times \Big(\sum_{L \leq l \ll M^{1/2}} \sum_{m \ll M/l^2} \sum_{n \asymp X} |\lambda\Big(\frac{\sigma m l^2 n + f}{\Delta} \Big)|^4 \Big)^{1/4} \\
&\preccurlyeq \Big(\sum_{L \leq l \ll M^{1/2}} \sum_{m \ll M/l^2} 1 \Big)^{3/4} \Big(\sum_n |\beta_n|^2 \Big)^{1/2} \Big(\sum_{n \asymp X} 1 \Big)^{1/4} (MX)^{1/4} \\
&\ll \|\beta\| \frac{MX^{1/2}}{L^{3/4}}
\end{aligned}$$

using the well-known bound for the fourth moment of Hecke eigenvalues

$$\sum_{n \leq Y} |\lambda(n)|^4 \leq Y^{1+\varepsilon} \sum_{n \leq Y} \frac{|\lambda(n)|^4}{n^{1+\varepsilon}} \preccurlyeq Y. \tag{4.3}$$

(The Dirichlet series $\sum_n \lambda_\phi(n)^4 n^{-s}$ equals $L(s, \mathrm{sym}^2\phi \times \mathrm{sym}^2\phi) L(s, \mathrm{sym}^2\phi)^2 \zeta(s)$ up to an Euler product that is absolutely convergent in $\Re s > 1/2 + 2\theta$ where $\theta \leq 7/64$

is an admissible exponent for the Ramanujan-Petersson conjecture. Hence if C_ϕ denotes the conductor of ϕ, then we conclude from standard bounds for L-functions at $s=1$ [XLi] the second bound in the above previous display.)

On the other hand, for $l \le L$ we combine n and l^2 to one variable and apply (4.2) with $\tilde\beta_n^{(l)} = \beta_{n/l^2}\mathbb{1}_{l^2|n}$ to obtain

$$\begin{aligned}&\sum_{l\le L}\sum_{m\asymp M/l^2}\sum_{n\asymp Xl^2}\mu^2(m)\alpha_{ml^2}\tilde\beta_n^{(l)}\lambda\Big(\frac{\sigma mn+f}{\Delta}\Big)\\ &\preccurlyeq \sum_{l\le L}\|\tilde\beta^{(l)}\|\Big((MX)^{1/2}+(Xl^2)^{3/4}(M/l^2)^{3/2}+(M/l^2)^{5/4}(Xl^2)^{3/8}\Big)\\ &\ll \|\beta\|\big(L(MX)^{1/2}+X^{3/4}M^{3/2}+M^{5/4}X^{3/8}\big).\end{aligned}$$

With $L=M^{2/7}$, we obtain

$$\mathscr{B}(\alpha,\beta)\preccurlyeq\|\beta\|(X^{1/2}M^{11/14}+X^{1/4}M^{3/2}+M^{5/4}X^{3/8}).$$

Here we can drop the middle term, because if it dominates the last term, then $M>X^{1/2}$ in which case the claim is trivially true. This completes the proof.

4.2 Proof of Proposition 3.2. This result has a precursor in [ABL, Theorem 2.5], but it turns out that here the extra congruence conditions modulo q and Δ are somewhat subtle and not completely straightforward to implement due to well-known difficulties in the Voronoi summation formula for general moduli. We therefore modify the proof, the biggest difference being that we replace the delta-symbol method ([ABL, Lemma 7.3]) of Duke-Friedlander-Iwaniec with Jutila's circle method.[1]

Let us denote by $\mathscr{S}$ the k,l,m-sum that we want to bound and recall that P is the product of the "unimportant" parameters (keeping in mind that the definition of P differs slightly from the previous proof). Assume without loss of generality $X_1\ge X_2\ge X_3$.

By Voronoi summation (cf. [ABL, Corollary 7.7]) it is easy to see that

$$\mathscr{S}\preccurlyeq (X/X_1)^{3/2}. \tag{4.4}$$

We can also write

$$\mathscr{S}=\sum_{\substack{k,l,m\\ rklm\equiv d\,(\mathrm{mod}\,q)}}\sum_n G(k,l,m)g(n)\lambda(n)\int_0^1 e\big(\alpha(rklm+f-\Delta n)\big)d\alpha.$$

where the redundant function $g(n)$ has support on $\frac12\Xi/\Delta\le n\le 3\Xi/\Delta$ and satisfies $g(n)=1$ for $\Xi/\Delta\le n\le 2\Xi/\Delta$ and $g^{(\nu)}(n)\ll(\Xi/\Delta)^{-\nu}$ for all $\nu\in\mathbb{N}_0$. We now employ Jutila's circle method as in the corollary to [Ju, Lemma 1], i.e. we insert the

[1] As an aside we remark that we could also follow Munshi's strategy [Mu], but note that his Lemma 6 needs to be corrected by a factor q_1 when $q_1=\tilde q_1$.

constant function into the α-integral and approximate it by a step function of small rational intervals. This device will simplify the argument especially with respect to the extra congruence condition on q. Let $C \gg X^{1/2}$ be a large parameter and $\mathcal{C} \subset [C, 2C]$ be the set of moduli which are coprime to $\Delta N q$. Define

$$L = \sum_{c \in \mathcal{C}} \phi(c) \gg C^2 (\Delta N q C)^{-\varepsilon}$$

and let δ be a real number such that $C^{-2} \ll \delta \ll C^{-1}$. Then we have (bounding the error term with Cauchy-Schwarz and Rankin–Selberg)

$$\mathscr{S} = \tilde{\mathscr{S}} + O\Big(\frac{(X\Xi/\Delta)^{1/2}}{\delta^{1/2} C} (XC)^{\varepsilon}\Big)$$

where

$$\tilde{\mathscr{S}} = \sum_{\substack{k,l,m \\ rklm \equiv d \,(\mathrm{mod}\ q)}} \sum_{n} G(k,l,m) g(n) \lambda(n) \times \frac{1}{L} \sum_{c \in \mathcal{C}} \sideset{}{^*}\sum_{a \,(\mathrm{mod}\ c)} \frac{1}{2\delta} \int_{-\delta}^{\delta} e\Big(\Big(\frac{a}{c} + \alpha\Big)(rklm + f - \Delta n)\Big) d\alpha.$$

We now apply Voronoi summation [KMV, Theorem A.4] to the n-sum and note that by construction $(\Delta N, c) = 1$ (this flexibility is the main advantage of the present set-up). We obtain

$$\sum_n g(n) e(\alpha \Delta n) \lambda(n) e\Big(\frac{a\Delta}{c} n\Big) = \frac{\xi}{c\sqrt{N}} \sum_{\pm} \chi_{-4}(\mp c) \sum_n \lambda(n) e\Big(\mp \frac{\overline{a \Delta N}}{c} n\Big) \mathcal{G}^{\pm}\Big(\frac{n}{c^2 N}\Big)$$

for $(a, c) = 1$, some constant $\xi \in S^1$ depending only on the cusp form ϕ and functions $\mathcal{G}^{\pm}$ which by [ABL, Lemma 7.5] satisfy uniformly in $|\alpha| \le \delta$ the bound

$$y^{\nu} \frac{d^{\nu} \mathcal{G}^{\pm}(y)}{dy^{\nu}} \ll_{A,\nu} \frac{\Xi}{\Delta} \Big(1 + \frac{1}{(y\Xi/\Delta)^{2\tau+\varepsilon}}\Big) \Big((1 + |\mu|)\Big(1 + \sqrt{\frac{y\Xi}{\Delta}}\Big)\Big)^{\nu} \times \Big(1 + \frac{y\Xi/\Delta}{(\delta\Xi/\Delta + 1)^2} + \frac{y\Xi/\Delta}{(1 + |\mu|)^2}\Big)^{-A}$$

where $\tau = 0$ if ϕ is holomorphic and $\tau = \Im t_\phi$ if ϕ is Maaß with spectral parameter t_ϕ. We choose

$$\delta = 1/X.$$

With this choice we can truncate the n-sum at $\preccurlyeq C^2/X$ and obtain

$$\mathscr{S} \preccurlyeq \frac{X^{3/2}}{C} + \frac{X}{L} \sup_{|\alpha| \le \delta} \sum_{\pm} \sum_{c \in \mathcal{C}} \sum_{n \preccurlyeq C^2/X} \frac{|\lambda(n)|}{c} \Big(1 + \frac{C^2}{nX}\Big)^{2\tau} \times \Big| \sum_{k,l,m} G(k,l,m) e\Big(\Big(\alpha + \frac{b}{q}\Big) rklm\Big) S(rklm + f, \pm\overline{\Delta} n, c)\Big|.$$

The factor $e(\alpha rklm)$ can be incorporated into G at essentially no cost, since $\alpha rklm \ll \delta X = 1$. When $X_1 \ll C$, we can bound the k,l,m-sum using [ABL, Lemma 7.2] to obtain

$$\preccurlyeq \frac{X(c,f,n)^{1/2}}{c^{1/2}} + \Big((c,f)^{1/2}c^{7/4} + X_1^{1/2}c^{3/2} + (X_2X_3)^{1/2}c^{5/4} + X_1(X_2X_3)^{1/2}c^{1/2}\Big)c_\square^{1/4}$$

where $c_\square$ is the squarefull part of c. We can now evaluate the sums over n (cf. [ABL, (7.1)]) and c (with Rankin's trick, cf. two displays after [ABL, (8.3)]) getting

$$\mathscr{S} \preccurlyeq \frac{X}{C^{1/2}} + C^{7/4} + X_1^{1/2}C^{3/2} + X^{1/3}C^{5/4} + X_1^{1/2}X^{1/2}C^{1/2} + \frac{X^{3/2}}{C}$$

(recall that $X_1 \geq X_2 \geq X_3$). At this point we make the admissible choice $C = X^{1/2+1/54}$ getting

$$\mathscr{S} \preccurlyeq X^{53/54} + X_1^{1/2}X^{7/9}. \tag{4.5}$$

Then the claim follows from (4.5) when $X^{1/3} \leq X_1 \leq X^{2/5}$ and from (4.4) when $X_1 \geq X^{2/5}$.

5 Averaged bilinear estimates

The main result of this section is an averaged bilinear estimate which will be used later to handle the error terms when applying sieve methods. Define

$$\mathfrak{u}_R(n;q,a,d,w) = \mathbb{1}_{\substack{n\equiv a \,(\mathrm{mod}\, q)\\ n\equiv w \,(\mathrm{mod}\, d)}} - \frac{1}{\phi(q)\phi(d)} \sum_{\substack{\chi \,(\mathrm{mod}\, q),\psi \,(\mathrm{mod}\, d)\\ \mathrm{cond}(\chi\psi)\leq R}} \chi(n\bar{a})\psi(n\overline{w}). \tag{5.1}$$

Note that $\mathfrak{u}_R(n;q,a,1,*)$ equals $\mathfrak{u}_R(n\bar{a},q)$ as defined in [ABL, Section 4]. Here we only use $\psi \,(\mathrm{mod}\, d)$ with $\mathrm{cond}(\psi) \leq R$ so that GRH can be avoided when estimating contributions from the second term in (5.1). However, we could not treat the contribution from each d *individually* as in [ABL, Proposition 4.1] and thus we make use of the *sum* over d as the following result indicates.

Proposition 5.1. *Let $M,N,Q,R \geq 1$, $a_1 \in \mathbb{Z}\backslash\{0\}$, write $x = MN$. Let $c \in \mathbb{N}$, $c_0 \in \mathbb{Z}$ with $(c_0,c) = 1$. Let α_m and β_n be two sequences supported in $m \in (M, 2M]$ and $n \in (N, 2N]$ such that for some $A \geq 1$ we have $\alpha_m \leq \tau(m)^A, \beta_n \leq \tau(n)^A$, and suppose $\lambda_d \ll \tau(d)^A$.*

Let $\eta > 0$ be any sufficiently small number. Then there exist $\delta = \delta(\eta) > 0$ and $D = D(\eta, A)$ with the following property. If

$$x^\eta \leq N \leq x^{1/4+\eta}, \quad Q \leq x^{1/2+\delta}, \quad c, R \leq x^\delta, \quad |a_1| \leq x^{1+\delta}, \quad |a_2| \leq x^\delta$$

then

$$\sum_{d\leq x^{\delta}}\lambda_d \sup_{\substack{w \,(\mathrm{mod}\, d)^{\times}\\ a_3|d}}\Big|\sum_{\substack{Q\leq q\leq 2Q\\ (q,a_1a_2d)=1\\ q\equiv c_0\,(\mathrm{mod}\, c)}}\sum_{\substack{m,n\\ (n,a_2)=1}}\alpha_m\beta_n u_R(mn;q,a_1\overline{a_2a_3},d,w)\Big|$$
$$\ll_{\eta,A} cx(\log x)^D R^{-1/3}.$$

We remark that the condition $(n,a_2)=1$ is important in this proposition for a few reasons: it is used by [Dr2, around eq. (5.11)] to reduce to the case of n being square-free as well as the modulo 1 reduction of the exponential phase in [Dr2, (5.24)]. More importantly, it is required in the application of the Kuznetsov formula when estimating sum of Kloosterman sums (e.g. [Dr2, the condition $(q,s)=1$ in Proposition 4.12]).

Proof. Following the proof of [ABL, Proposition 4.1], it is enough to prove for any smooth function $\gamma:\mathbb{R}_+\to[0,1]$ with

$$1_{Q\leq q\leq 2Q}\leq\gamma(q)\leq 1_{Q/2\leq q\leq 3Q/2},$$
$$\|\gamma^{(j)}\|_{\infty}\ll_j Q^{-j+B\delta j}\text{ for some } B\geq 0 \text{ and for all } j\geq 0$$

that under the assumptions in Proposition 5.1 as well the additional condition that β_n is supported on square-free integers we have

$$\mathcal{S}:=\sum_{d\asymp D}\lambda_d \sup_{\substack{w \,(\mathrm{mod}\, d)^{\times}\\ a_3|d}}\Big|\sum_{\substack{q\leq Q\\ (q,a_1a_2d)=1\\ q\equiv c_0\,(\mathrm{mod}\, c)}}\gamma(q)\sum_{\substack{m,n\\ (n,a_2)=1}}\sum\alpha_m\beta_n\mathfrak{u}_R(mn;q,a_1\overline{a_2a_3},d,w)\Big|$$
$$\ll\frac{cx(\log x)^{O(1)}}{R^{1/3}} \tag{5.2}$$

for $D\ll x^{\delta}$. The proof of (5.2) is similar to that in [ABL, Proposition 4.2] with some modifications. Here we have a sum over d which will only influence the application of the large sieve inequalities, as we obtain power saving error terms in all other parts of the argument for each d. Note that we also have an additional condition $(q,d)=1$ compared to that in [ABL, Proposition 4.2], which can be incorporated following the proof. The sup over $w\,(\mathrm{mod}\, d)$ and $a_3\mid d$ cause no issue as the main expressions will be independent of w,a_3.

To start with, we also assume β_n is supported on $n\equiv b_0\,(\mathrm{mod}\, c)$ as in the second display of the proof of Proposition 4.2 of [ABL], which makes certain coprime conditions easier to verify when applying [ABL, Theorem 2.1] (and costs a factor c in the final bound).

We now apply the Cauchy-Schwarz inequality in d,m with the bounds for λ_d and α_m. After majorizing the summation over m with the help of a suitable smooth

weight $\alpha(m)$ we get

$$\mathcal{S} \ll (\log x)^{O(1)}(DM)^{1/2}\Big(\sum_{d\asymp D}\sup_{w \,(\mathrm{mod}\, d)^\times \atop a_3|d}\sum_m \alpha(m) \times \Big|\sum_{q\atop (q,a_1a_2d)=1\atop q\equiv c_0 \,(\mathrm{mod}\, c)} \gamma(q)\sum_{n\atop (n,a_2)=1}\beta_n \mathfrak{u}_R(mn;q,a_1\overline{a_2a_3},d,w)\Big|^2\Big)^{1/2}$$

$$=: (\log x)^{O(1)}(DM)^{1/2}\Big(\sum_{d\asymp D}\sup_{w \,(\mathrm{mod}\, d)^\times \atop a_3|d}\mathcal{S}_1(w,a_3)-2\Re(\mathcal{S}_2(w,a_3))+\mathcal{S}_3(w,a_3)\Big)^{1/2},$$

where (if $(n,qd)\neq 1$ the above sum is empty)

$$\mathcal{S}_1(w,a_3)=\sum_{(q_i,a_1a_2d)=1\atop q_i\equiv c_0 \,(\mathrm{mod}\, c)}\gamma(q_1)\gamma(q_2)\sum_{n_i\asymp N\atop (n_i,q_ia_2d)=1}\beta_{n_1}\overline{\beta_{n_2}}\sum_{mn_i\equiv w \,(\mathrm{mod}\, d)\atop mn_i\equiv a_1\overline{a_2a_3} \,(\mathrm{mod}\, q_i)}\alpha(m),$$

$$\mathcal{S}_2(w,a_3)=\sum_{(q_i,a_1a_2d)=1\atop q_i\equiv c_0 \,(\mathrm{mod}\, c)}\frac{\gamma(q_1)\gamma(q_2)}{\phi(q_2)\phi(d)}\sum_{n_i\asymp N\atop (n_i,q_ia_2d)=1}\beta_{n_1}\overline{\beta_{n_2}} \times\sum_{\chi_2 \,(\mathrm{mod}\, q_2)\atop \psi_2 \,(\mathrm{mod}\, d)\atop \mathrm{cond}(\chi_2\psi_2)\le R}\chi_2(n_2\overline{a_1}a_2a_3)\psi_2(\bar w)\sum_{mn_1\equiv a_1\overline{a_2a_3} \,(\mathrm{mod}\, q_1)\atop mn_1\equiv w \,(\mathrm{mod}\, d)}\chi_2\psi_2(m)\alpha(m),$$

and

$$\mathcal{S}_3(w,a_3)=\sum_{(q_i,a_1a_2d)=1\atop q_i\equiv c_0 \,(\mathrm{mod}\, c)}\frac{\gamma(q_1)\gamma(q_2)}{\phi(q_1)\phi(q_2)\phi(d)^2}\sum_{n_i\asymp N\atop (n_i,q_ia_2d)=1}\beta_{n_1}\overline{\beta_{n_2}} \times\sum_{\chi_i \,(\mathrm{mod}\, q_i)\atop \psi_i \,(\mathrm{mod}\, d)\atop \mathrm{cond}(\chi_i\psi_i)\le R}\chi_1(n_1\overline{a_1}a_2a_3)\overline{\chi_2(n_2\overline{a_1}a_2a_3)}\psi_1\overline{\psi_2}(\bar w)\sum_m\chi_1\psi_1\overline{\chi_2\psi_2}(m)\alpha(m).$$

It is then enough to show (recall $MN=x$)

$$\sum_{d\asymp D}\sup_{w \,(\mathrm{mod}\, d)^\times \atop a_3|d}\Big(\mathcal{S}_1(w,a_3)-2\Re\mathcal{S}_2(w,a_3)+\mathcal{S}_3(w,a_3)\Big)\ll MN^2D^{-1}R^{-2/3}(\log x)^{O(1)}. \tag{5.3}$$

We can evaluate $\mathcal{S}_3$ as in [Dr2, Section 5.3.1] so that the m-sum becomes

$$\frac{\hat\alpha(0)}{[q_1,q_2]d}\sum_{h \,(\mathrm{mod}\, W)^\times}\chi_1\psi_1\overline{\chi_2\psi_2}(h)+O(W^\varepsilon R^{1/2})$$

where $W = [q_1, q_2]d$, and thus for $D, R \le x^\delta$ and uniformly for $|a_2| \le x^\delta$ we have

$$\mathcal{S}_3(w, a_3) = \hat{\alpha}(0)X_3 + O(x^\varepsilon R^{5/2}N^2/D^2)$$

where

$$X_3 = \sum_{\substack{(q_i, a_1a_2d)=1\\ q_i \equiv c_0 \,(\mathrm{mod}\, c)}} \frac{\gamma(q_1)\gamma(q_2)}{[q_1,q_2]d\phi((q_1,q_2))\phi(d)} \sum_{\substack{\chi \,(\mathrm{mod}\,(q_1,q_2)d)\\ \mathrm{cond}(\chi)\le R}} \sum_{\substack{n_i \asymp N\\ (n_i, q_ia_2d)=1}} \beta_{n_1}\overline{\beta_{n_2}}\chi(n_1\overline{n_2}).$$

Note that X_3 is independent of w, a_3, and the error term is acceptable for (5.3) (in fact much better than required).

Similarly we can evaluate $\mathcal{S}_2$ as in [Dr2, Section 5.3.2] so that for $D, R \le x^\delta$ and uniformly for $|a_2| \le x^\delta$ we have

$$\mathcal{S}_2(w, a_3) = \hat{\alpha}(0)X_3 + O(MN^2D^{-2}x^{-2\delta}).$$

We next evaluate $\mathcal{S}_1(w, a_3)$. After Poisson summation in m, we get the expected main term as $\hat{\alpha}(0)X_1$, where

$$X_1 = \sum_{\substack{(q_i, a_1a_2d)=1\\ q_i \equiv c_0 \,(\mathrm{mod}\, c)}} \frac{\gamma(q_1)\gamma(q_2)}{[q_1,q_2]d} \sum_{\substack{n_i \asymp N\\ (n_i, q_id)=1\\ n_1 \equiv n_2 \,(\mathrm{mod}\, d(q_1,q_2))}} \beta_{n_1}\overline{\beta_{n_2}},$$

which is also independent of w, a_3. The rest of the evaluation of $\mathcal{S}_1$ is similar to that in [ABL, Proposition 4.2] (with a_2 replaced by a_2a_3). The condition $(q_i, a_2d) = 1$ is equivalent to $(q_i, a_2a_3) = 1$ and $(q_i, \tilde{d}) = 1$ where $\tilde{d} = d/(d, (a_2a_3c)^\infty)$. Thus the additional condition $(q_i, \tilde{d}) = 1$ can be incorporated by following the proof of [ABL, Proposition 4.2] using Möbius inversion. To be precise, we replace the condition $\delta_i \mid a_1$ by the condition that $\delta_i \mid a_1\tilde{d}$ so that the condition $(\delta_2, n_0a_2a_3c) = 1$ is still satisfied when applying [ABL, Theorem 2.3] since $(n_i, d) = 1$. In conclusion, we have for $\delta = \delta(\eta), \kappa = \kappa(\eta)$ small enough, $D, R \le x^\delta$ and uniformly for $|a_2| \le x^\delta$ that

$$\mathcal{S}_1(w, a_3) = \hat{\alpha}(0)X_1 + O(MN^2D^{-2}x^{-2\delta}).$$

It remains to show that

$$\mathcal{X} := \sum_{d \asymp D} \sup_{\substack{w \,(\mathrm{mod}\, d)^\times\\ a_3 \mid d}} |X_1 - X_3| \ll N^2D^{-1}R^{-2/3}(\log x)^{O(1)}.$$

(Note that X_1, X_3 are independent of w, a_3.) The proof follows along the lines of [Dr2, Section 5.6]. We have

$$\mathcal{X} = \sum_{d \asymp D} \Big| \sum_{\substack{(q_i, a_1a_2d)=1\\ q_i \equiv c_0 \,(\mathrm{mod}\, c)}} \frac{\gamma(q_1)\gamma(q_2)}{[q_1,q_2]d\phi((q_1,q_2)d)\phi(d)} \sum_{\substack{\chi \,(\mathrm{mod}\,(q_1,q_2)d)\\ \mathrm{cond}(\chi) > R}} \sum_{\substack{n_i \asymp N\\ (n_i, q_ia_2d)=1}} \beta_{n_1}\overline{\beta_{n_2}}\chi(n_1\overline{n_2}) \Big|$$

$$\ll \log_2 x \sum_{d\asymp D} \sum_{\substack{q_i\asymp Q\\(q_i,a_1a_2d)=1}} \frac{1}{q_1q_2d^2} \sum_{\substack{\chi \text{ primitive}\\ \mathrm{cond}(\chi)>R\\ \mathrm{cond}(\chi)|(q_1,q_2)d}} \Big| \sum_{\substack{n_i\asymp N\\(n_i,q_ia_2d)=1}} \beta_{n_1}\overline{\beta_{n_2}}\chi(n_1\overline{n_2})\Big|.$$

Before we can estimate this with the large-sieve, we need to uncouple the variables. To do so, we write $\mathrm{cond}(\chi)=q\ell$ and detect the conditions $(n_i,q_id)=1$ by Möbius inversion. We thus estimate the above

$$\ll \log_2 x \sum_{\substack{\ell\ll D\\ q\ll Q\\ \ell q>R\\ (\ell,q)=1}} \sum_{\substack{r_1,r_2,s\\ r_i\ll Q/q\\ s\ll D/\ell}} \sum_{\substack{q_i\asymp Q, d\asymp D\\ r_iq|q_i\\ sl|d}} \frac{1}{q_1q_2d^2} \sum_{\substack{\chi \text{ primitive}\\ \chi \,(\mathrm{mod}\, q\ell)}} \Big| \sum_{\substack{n_i\\(n_i,a_2)=1}} \beta_{r_1sn_1}\overline{\beta_{r_2sn_2}}\chi(n_1\overline{n_2})\Big|$$

$$\ll \frac{(\log x)^3}{D} \sum_{\substack{r_1,r_2,s\\ r_i\ll Q\\ s\ll D}} \frac{1}{r_1r_2s} \sum_{\substack{\ell\ll D\\ q\ll Q\\ \ell q>R}} \frac{1}{q^2\ell} \sum_{\substack{\chi \text{ primitive}\\ \chi \,(\mathrm{mod}\, q\ell)}} \Big| \sum_{\substack{n_i\\(n_i,a_2)=1}} \beta_{r_1sn_1}\overline{\beta_{r_2sn_2}}\chi(n_1\overline{n_2})\Big|.$$

By Cauchy-Schwarz and the symmetry between n_1 and n_2, we arrive at

$$\mathcal{X} \ll \frac{(\log x)^{O(1)}}{D} \sum_{r\ll x} \frac{\tau(r)}{r} \sum_{\substack{\ell\ll D\\ q\ll Q\\ \ell q>R\\ (\ell,q)=1}} \frac{1}{q^2\ell} \sum_{\substack{\chi \text{ primitive}\\ \chi \,(\mathrm{mod}\, q\ell)}} \Big| \sum_{\substack{n\asymp N\\(n,a_2)=1}} \beta_{rn}\chi(n)\Big|^2.$$

By the large sieve inequality for Dirichlet characters ([Dr2, Lemma 3.3]) and partial summation, the contribution from $q>R^{1/3}$ can be bounded by

$$\ll (\log x)^{O(1)} \frac{1}{D}\Big(\frac{(D^2Q^2+N)N}{Q^2} + \int_{R^{1/3}}^{2Q} \frac{(D^2t^2+N)N}{t^3}dt\Big)$$
$$\ll (\log x)^{O(1)} \Big(ND + \frac{N^2}{DR^{2/3}}\Big).$$

Similarly the contribution from $\ell\geq R^{2/3}, q\leq R^{1/3}$ can be bounded by

$$\ll (\log x)^{O(1)} \frac{1}{D}\Big(\frac{(R^{2/3}t^2+N)N}{t}\Big|_{R^{2/3}}^{D} + \int_{R^{2/3}}^{2D} \frac{(R^{2/3}t^2+N)N}{t^2}dt\Big)$$
$$\ll (\log x)^{O(1)} \Big(R^{2/3}N + \frac{N^2}{DR^{2/3}}\Big).$$

Therefore, we have

$$\sum_{d\asymp D} \sup_{\substack{w \,(\mathrm{mod}\, d)^\times\\ a_3|d}} |X_1 - X_3| \ll (\log x)^{O(1)} \frac{N^2}{DR^{2/3}}$$

if $D\ll N/R^{1/3}$. Since $N\gg x^\eta$, we conclude that for δ small enough (in terms of η) equation (5.3) holds and so (5.2) follows. This completes the proof of Proposition 5.1.

6 Proofs of Theorems 1.1 and 1.2

6.1 Primes in arithmetic progressions. Before we start with the analysis of the main term, we give a result on averaged equidistribution of primes in long arithmetic progressions, which will be used to handle the error terms in the sieving process.

PROPOSITION 6.1. *There exists some absolute constant $\varpi > 0$ such that the following holds. Let $x \geq 2$, $c_0, c, d, C \in \mathbb{N}, (c_0, c) = 1, a_1, a_2 \in \mathbb{Z}\backslash\{0\}$ such that for*

$$Q \leq x^{1/2+\varpi}, \quad |a_1| \leq x^{1+\varpi}, \quad |a_2| \leq x^{\varpi}, \quad |\lambda_d| \ll \tau(d)^C$$

we have

$$\sum_{d \leq x^{\varpi}} \lambda_d \sup_{w \,(\mathrm{mod}\, d)^{\times}} \Big| \sum_{\substack{q \leq Q \\ (q, a_1 a_2 d) = 1 \\ q \equiv c_0 \,(\mathrm{mod}\, c)}} \Big(\sum_{\substack{n \leq x \\ n \equiv a_1 \overline{a_2} \,(\mathrm{mod}\, q) \\ n \equiv w \,(\mathrm{mod}\, d)}} \Lambda(n) - \frac{1}{\phi(qd)} \sum_{\substack{n \leq x \\ (n, qd) = 1}} \Lambda(n) \Big) \Big|$$

$$\ll_{C,A} cx(\log x)^{-A}.$$

Proof. This is the analogue of [ABL, Theorem 2.1] and can be proved in the exact same way as at the end of [ABL, Section 4]. We only need to replace [ABL, Proposition 4.1] with Proposition 5.1 with the choice $R = (\log x)^B$ for some large B depending on C, A. The contribution from τ_3-type sums is negligible by choosing ϖ small enough since [BFI, Lemma 2] saves a small power of x. Using $\sum_{d \leq x^{\kappa}} \lambda_d / \phi(d) \ll (\log x)^{O(1)}$ and the Siegel-Walfisz Theorem we also have that the contribution from $\chi \,(\mathrm{mod}\, qd)$ with $\mathrm{cond}(\chi) \leq R$ can be bounded by $xR(\log x)^{O(1)} \exp(-b\sqrt{\log x})$, which is negligible by the choice of R.

For future reference we remark that we can replace the von Mangoldt function by the characteristic function on primes.

6.2 Preparing for the sieve. In this section, we use κ to denote a sufficiently small (depending on ϖ in Proposition 6.1 and A, η as in Theorem 3.3) positive constant.

We use ε to denote an arbitrarily small positive constant.

Let $\mathbf{d} = (d_1, d_2)$ denote a pair of two square-free numbers with $d_1, d_2 \leq n^{\kappa}$. In preparation for a sieve we define

$$\mathcal{A}_{\mathbf{d}}(n) = \{(p, x_1, x_2) \mid p + d_1^2 x_1^2 + d_2^2 x_2^2 = n\}$$

(where of course p denotes a prime). As usual, we denote by $c_p(n)$ the Ramanujan sum. The key input for the sieve is the following:

PROPOSITION 6.2. *With the above notation, there exists some absolute constant $\kappa > 0$ such that uniformly for $d_1, d_2 \le n^\kappa$ and $\lambda_{\boldsymbol{d}} \ll (\tau(d_1 d_2))^C$ we have*

$$\sum_{d_1,d_2\le n^\kappa}\sum \lambda_{\boldsymbol{d}}\Big|\#\mathcal{A}_{\boldsymbol{d}}(n) - \mathrm{Li}(n)\mathfrak{S}_{\boldsymbol{d}}(n)\Big| \ll_{C,A} n(\log n)^{-A}, \tag{6.1}$$

where the singular series $\mathfrak{S}_{\boldsymbol{d}}(n)$ is given by

$$\begin{aligned}&\frac{\pi}{d_1 d_2}\prod_{p\mid(d_1,d_2)}\Big(1-\frac{c_p(n)}{p-1}\Big)\prod_{\substack{p\mid d_1d_2\\ p\nmid(d_1,d_2)\\ p\mid 2n}}\Big(1-\frac{\left(\frac{n}{p}\right)}{p-1}\Big)\\ &\quad\times\prod_{p\nmid d_1d_2n}\Big(1+\frac{\chi_{-4}(p)}{p(p-1)}\Big)\prod_{\substack{p\mid n\\ p\nmid 2d_1d_2}}\Big(1-\frac{\chi_{-4}(p)}{p}\Big).\end{aligned} \tag{6.2}$$

REMARK. One can check that this matches the main term coming from a formal application of the circle method. In particular, we see that $\mathfrak{S}_{\mathbf{d}}(n) = 0$ if $(d_1, d_2, n) > 1$ as it should.

Proof. Let $\delta = (d_1, d_2), d_1' = d_1/\delta, d_2' = d_2/\delta$ as in Lemma 2.2. By Lemma 2.2 and Theorem 3.3 (with κ sufficiently small depending on the constants A, η) we have

$$\#\mathcal{A}_{\mathbf{d}}(n) = \sum_{\delta_1\delta_1'=d_1'}\sum_{\delta_2\delta_2'=d_2'}\frac{w_{\delta_1'\delta_2'}}{h_{\delta_1'\delta_2'}}G_{\delta_1'\delta_2'}\sum_{w\in\mathcal{G}_{\delta_1'\delta_2'}}\sum_{\substack{p\equiv n\,(\mathrm{mod}\,(\delta\delta_1'\delta_2')^2)\\ n-p\equiv w\,(\mathrm{mod}\,\delta_1'\delta_2')}} r\left(\frac{n-p}{\delta^2\delta_1^2\delta_2^2}\right) + O(n^{1-\kappa}), \tag{6.3}$$

where $\mathcal{G}_{\delta_1'\delta_2'} \subseteq (\mathbb{Z}/\delta_1'\delta_2'\mathbb{Z})^*$ is as in (2.8) and $r = \mathbf{1} * \chi_{-4}$ as before. From (2.1) we obtain

$$\frac{w_{\delta_1'\delta_2'}}{h_{\delta_1'\delta_2'}} = \frac{4}{\delta_1'\delta_2'}\prod_{p\mid\delta_1'\delta_2'}\Big(1-\frac{\chi_{-4}(p)}{p}\Big)^{-1}$$

in all cases, including $\delta_1'\delta_2' = 1$. Note that we have

$$\frac{w_{\delta_1'\delta_2'}}{h_{\delta_1'\delta_2'}}G_{\delta_1'\delta_2'}\#\mathcal{G}_{\delta_1'\delta_2'} \ll (\tau(\delta_1'\delta_2'))^{O(1)}.$$

We apply the Dirichlet's hyperbola method to the r-function so that the innermost p-sum in (6.3) can for some parameter $Y > 0$ be written as

$$\sum_{Y\le p\le n}\sum_{\substack{ab=(n-p)/\delta^2\delta_1^2\delta_2^2\\ n-p\equiv w\,(\mathrm{mod}\,\delta_1'\delta_2'),}}\chi_{-4}(b) + O(Y^{1+\varepsilon}) = S_1 + S_2 + O(Y^{1+\varepsilon}),$$

say, where

$$S_1 = \sum_{\substack{a\le\sqrt{n}\\(a,\delta_1'\delta_2')=1}} \sum_{\substack{Y\le p\le n\\ p\equiv n \,(\mathrm{mod}\, a\delta^2\delta_1^2\delta_2^2)\\ n-p\equiv w \,(\mathrm{mod}\, \delta_1'\delta_2')}} \chi_{-4}\Big(\frac{n-p}{a\delta^2\delta_1^2\delta_2^2}\Big),$$

$$S_2 = \sum_{\substack{b\le \frac{n-Y}{\sqrt{n}\delta^2\delta_1^2\delta_2^2}\\(b,\delta_1'\delta_2')=1}} \sum_{\substack{Y\le p\le n-b\sqrt{n}\delta^2\delta_1^2\delta_2^2\\ p\equiv n \,(\mathrm{mod}\, b\delta^2\delta_1^2\delta_2^2)\\ n-p\equiv w \,(\mathrm{mod}\, \delta_1'\delta_2')}} \chi_{-4}(b).$$

We choose $Y = n^{1-3\kappa}$ so that the error term is acceptable after summing over $d_1, d_2 \le n^\kappa$. We will evaluate S_1 and S_2 on average over d_1, d_2 by Proposition 6.1. This requires some preparation. It turns out that the contribution from S_1 will not appear in the main term $\mathrm{Li}(n)\mathfrak{S}_{\mathbf{d}}(n)$, but as there are some complications from the powers of two, we start with a detailed treatment for S_1.

We use the notation

$$r_2 = (\delta\delta_1\delta_2, 2^\infty), \quad d = \frac{\delta\delta_1\delta_2}{r_2}, \quad d' = \frac{\delta_1'\delta_2'}{(\delta_1'\delta_2', 2)}, \quad a_2 = (a, 2^\infty), \quad a_d = (a, d^\infty).$$

We see that S_1 equals

$$\sum_{\substack{a_2\mid 2^\infty\\(a_2,\delta_1'\delta_2')=1}} \sum_{\substack{a\le\sqrt{n}/a_2\\(a,2\delta_1'\delta_2')=1}} \sum_{\substack{Y\le p\le n\\ p\equiv n \,(\mathrm{mod}\, aa_2r_2^2d^2)\\ n-p\equiv w \,(\mathrm{mod}\, \delta_1'\delta_2')}} \chi_{-4}(a)\chi_{-4}\Big(\frac{n-p}{a_2r_2^2}\Big)$$

$$= \sum_{\substack{a_2\mid 2^\infty\\(a_2,\delta_1'\delta_2')=1}} \sum_{a_d\mid d^\infty} \chi_{-4}(a_d) \sideset{}{^*}\sum_{u \,(\mathrm{mod}\, 4)} \chi_{-4}(u) \sum_{v \,(\mathrm{mod}\, 4)} \chi_{-4}(v) \times \sum_{\substack{a\le\sqrt{n}/a_2a_d\\(a,2d_1d_2)=1\\ a\equiv u \,(\mathrm{mod}\, 4)}} \sum_{\substack{Y\le p\le n\\ p\equiv n \,(\mathrm{mod}\, aa_2a_dr_2^2d^2)\\ n-p\equiv w \,(\mathrm{mod}\, d')\\ n-p\equiv a_2r_2^2v \,(\mathrm{mod}\, 4a_2r_2^2)}} 1.$$

Here we have used the fact that if $2 \mid \delta_1'\delta_2'$ then we must have $a_2r_2 = 1$ and $(w, 2) = 1$. The contribution when one of a_2, a_d is at least $n^{4\kappa}$ can be bounded by

$$\sum_{\substack{a\mid(2d)^\infty\\ a\ge n^{4\kappa}}} \sum_{q\le\sqrt{n}/a} \frac{n}{\phi(a)\phi(q)} \ll n^{1-3\kappa+\varepsilon},$$

which is negligible after summing over $d_1, d_2 \le n^\kappa$. So from now on we restrict to $a_2, a_d \le n^{4\kappa}$ (note that r_2, d', d are automatically at most $n^{2\kappa}$). We see that the

a,p-sum equals

$$\sum_{\substack{a\le\sqrt{n}/a_2a_d\\(a,ndd')=1\\a\equiv u\,(\mathrm{mod}\,4)}}\ \sum_{\substack{Y\le p\le n\\p\equiv n\,(\mathrm{mod}\,a)\\p\equiv n-w\,(\mathrm{mod}\,d')\\p\equiv n-va_2r_2^2\,(\mathrm{mod}\,4a_2r_2^2)\\p\equiv n\,(\mathrm{mod}\,a_dd^2)}} 1+O(n^{1/2+\varepsilon})$$

where the error term comes from the artificially added condition $(a,n)=1$. For the same reason we can restrict to $(n,d)=(n-w,d')=1$. After combining the last three congruence conditions modulo $4a_2r_2^2d'a_dd^2$ to a single condition and noting $d_1d_2\mid 4a_2r_2^2d'a_dd^2$ and $\lambda_{\mathbf{d}}\ll\tau(d_1d_2)^C$, we are in a position to apply Proposition 6.1 (since κ is sufficiently small in terms of ϖ) together with the prime number theorem to the two innermost sums and recast the previous display as

$$\mathbb{1}_{(n-w,d')=(n,d)=1}\sum_{\substack{a\le\sqrt{n}/a_2a_d\\a\equiv u\,(\mathrm{mod}\,4)\\(a,2d_1d_2n)=1}}\frac{\mathrm{Li}(n)}{\phi(a)\phi(d')\phi(4a_2a_dr_2^2d^2)}+r_{\mathbf{d}}$$

where the error terms $r_{\mathbf{d}}$ satisfy

$$\sum_{d_1,d_2\le n^{\kappa}}\lambda_{\mathbf{d}}\sum_{\delta_1\delta_1'=d_1'}\sum_{\delta_1\delta_2'=d_2'}\frac{w_{\delta_1'\delta_2'}}{h_{\delta_1'\delta_2'}}G_{\delta_1'\delta_2'}\sum_{w\in\mathcal{G}_{\delta_1'\delta_2'}}|r_{\mathbf{d}}|\ll x(\log x)^{-A}.$$

Plugging back, we obtain

$$\begin{aligned}S_1=\mathbb{1}_{(n-w,d')=(n,d)=1}&\sum_{\substack{a_2\mid 2^\infty\\a_2\le n^{4\kappa}}}\sum_{\substack{a_d\mid d^\infty\\a_d\le n^{4\kappa}}}\sum_{\substack{v\,(\mathrm{mod}\,4)\\(n-va_2r_2^2,4)=1}}\chi_{-4}(v)\sideset{}{^*}\sum_{u\,(\mathrm{mod}\,4)}\chi_{-4}(u)\\&\times\sum_{\substack{a\le\sqrt{n}/a_2a_d\\a\equiv u\,(\mathrm{mod}\,4)\\(a,2d_1d_2n)=1}}\frac{\mathrm{Li}(n)}{\phi(a)\phi(d')\phi(4a_2a_dr_2^2d^2)}+r_{\mathbf{d}}.\end{aligned}$$

But now the v-sum vanishes: with $h=a_2r_2^2\mid 2^\infty$ we have

$$\sum_{\substack{v\,(\mathrm{mod}\,4)\\(n-vh,4)=1}}\chi_{-4}(v)=\sum_{\ell\mid 4}\mu(\ell)\sum_{\substack{v\,(\mathrm{mod}\,4)\\vh\equiv n\,(\mathrm{mod}\,\ell)}}\chi_{-4}(v).$$

Since χ_{-4} is primitive, by [IK, (3.9)] the inner sum vanishes, unless $4\mid\ell/(\ell,h)$ (which can only happen if $h=1$), but then the Möbius function vanishes. We conclude the contribution from S_1 in (6.1) is acceptable.

We now turn to S_2 where the calculation is similar, and we obtain

$$S_2=\sum_{b_d\mid(\delta\delta_1\delta_2)^\infty}\chi_{-4}(b_d)\sum_{\substack{b\le\frac{n-Y}{\sqrt{n}(\delta\delta_1\delta_2)^2b_d}\\(b,d_1d_2)=1}}\chi_{-4}(b)\sum_{\substack{Y\le p\le n-bb_d\sqrt{n}\delta^2\delta_1^2\delta_2^2\\p\equiv n\,(\mathrm{mod}\,b_d\delta^2\delta_1^2\delta_2^2)\\p\equiv n\,(\mathrm{mod}\,b)\\n-p\equiv w\,(\mathrm{mod}\,\delta_1'\delta_2')}}1$$

$$= \sideset{}{^*}\sum_{u \,(\mathrm{mod}\ 4)} \chi_{-4}(u) \sum_{b_d \mid (\delta\delta_1\delta_2)^\infty} \chi_{-4}(b_d) \sum_{\substack{b \le \frac{n-Y}{\sqrt{n}(\delta\delta_1\delta_2)^2 b_d} \\ b \equiv u \,(\mathrm{mod}\ 4) \\ (b, d_1 d_2) = 1}} \sum_{\substack{Y \le p \le n - b b_d \sqrt{n} \delta^2 \delta_1^2 \delta_2^2 \\ p \equiv n \,(\mathrm{mod}\ b_d \delta^2 \delta_1^2 \delta_2^2) \\ p \equiv n \,(\mathrm{mod}\ b) \\ p \equiv n - w \,(\mathrm{mod}\ \delta_1' \delta_2')}} 1.$$

As before we restrict to $b_d \le n^{4\kappa}$ at the cost of an error of $O(x^{1-\kappa})$ (after summing over $d_1, d_2 \le n^\kappa$). The contribution from $(\delta\delta_1\delta_2, n) > 1$ is also negligible, so let us from now on assume $(\delta\delta_1\delta_2, n) = 1$. We cannot apply Proposition 6.1 to evaluate the b, p-sum on average over d_1, d_2 directly due to the dependence of the length of the intervals. To remedy this, we split the p-sum into intervals of the form $((1-\Delta)P, P]$ with $\Delta = (\log n)^{-K}$ for some sufficiently large $K = K(C)$. The condition $p \le n - b b_d \sqrt{n} \delta^2 \delta_1^2 \delta_2^2$ interferes in at most one of such intervals, whose contribution we estimate by the Brun-Titchmarsh inequality at the cost of an admissible error $O(n\Delta(\log x)^{O(1)})$ which is negligible by choosing K large enough in terms of C. For all other intervals we can split the b-sum in to intervals of the shape $((1-\Delta)P, P]$. The condition $b \le (n-Y)/\sqrt{n}(\delta\delta_1\delta_2)^2 b_d$ interferes in one of such intervals, whose contribution we can be bound by $O(nR(\log x)^{O(1)}\Delta)$ using the trivial bound for $\mathfrak{u}_R(n, q, a, d, w) \ll \frac{R\tau(q)\tau(d)}{\phi(q)\phi(d)}$. Here R is chosen to be some large power of $\log x$ in the proof of Proposition 6.1 and so the error term from separating the b-variable is negligible when K is sufficiently large. For the rest of the intervals, we can apply Proposition 6.1. Assembling the main term, we recast S_2 as

$$\begin{aligned} &\sideset{}{^*}\sum_{u \,(\mathrm{mod}\ 4)} \chi_{-4}(u) \sum_{\substack{b_d \mid (\delta\delta_1\delta_2)^\infty \\ b_d \le n^{4\kappa}}} \chi_{-4}(b_d) \\ &\qquad \times \sum_{\substack{b \le \frac{n-Y}{\sqrt{n}(\delta\delta_1\delta_2)^2 b_d} \\ b \equiv u \,(\mathrm{mod}\ 4) \\ (b, d_1 d_2 n) = 1}} \sum_{\substack{Y \le p \le n - b b_d \sqrt{n} \delta^2 \delta_1^2 \delta_2^2 \\ (p, b_d \delta^2 \delta_1^2 \delta_2^2) = 1 \\ (p, b) = 1 \\ (p, \delta_1' \delta_2') = 1}} \frac{\mathbb{1}_{(n-w, \delta_1'\delta_2') = (n, \delta\delta_1\delta_2) = 1}}{\phi(b)\phi(b_d\delta^2\delta_1^2\delta_2^2)\phi(\delta_1'\delta_2')} + r'_{\mathbf{d}} \\ &= \mathrm{Li}(n) \sum_{\substack{b_d \mid (\delta\delta_1\delta_2)^\infty \\ b_d \le n^{4\kappa}}} \frac{\chi_{-4}(b_d)}{\phi(b_d\delta^2\delta_1^2\delta_2^2)} \sum_{\substack{b \le \frac{n-Y}{\sqrt{n}(\delta\delta_1\delta_2)^2 b_d} \\ (b, d_1 d_2 n) = 1}} \frac{\chi_{-4}(b)}{\phi(b)} \frac{\mathbb{1}_{(n-w, \delta_1'\delta_2') = (n, \delta\delta_1\delta_2) = 1}}{\phi(\delta_1'\delta_2')} + r'_{\mathbf{d}}, \end{aligned}$$

where the error terms $r'_{\mathbf{d}}$ satisfy

$$\sum_{d_1, d_2 \le n^\kappa} \lambda_{\mathbf{d}} \sum_{\delta_1\delta_1' = d_1'} \sum_{\delta_1\delta_2' = d_2'} \frac{w_{\delta_1'\delta_2'}}{h_{\delta_1'\delta_2'}} G_{\delta_1'\delta_2'} \sum_{w \in \mathcal{G}_{\delta_1'\delta_2'}} |r'_{\mathbf{d}}| \ll x(\log x)^{-A}.$$

As a last step, we compute the b-sum in the main term in S_2 using [ABL, Lemma 5.2] and complete the sum over b_d, and recast the previous display up to an admissible

error as

$$\mathrm{Li}(n)L(1,\chi_{-4})\prod_p\Big(1+\frac{\chi_{-4}(p)}{p(p-1)}\Big)\mathfrak{c}(d_1d_2n)\sum_{b_d|(\delta\delta_1\delta_2)^\infty}\frac{\chi_{-4}(b_d)}{\phi(b_d\delta^2\delta_1^2\delta_2^2)}\frac{\mathbb{1}_{(n-w,d')=(n,\delta\delta_1\delta_2)=1}}{\phi(\delta_1'\delta_2')},$$

where

$$\mathfrak{c}(f)=\prod_{p|f}\Big(1+\frac{\chi_{-4}(p)}{p(p-1)}\Big)^{-1}\Big(1-\frac{\chi_{-4}(p)}{p}\Big). \tag{6.4}$$

This finishes the analysis of S_2.

We now return to (6.3) and summarize that there exists some absolute constant $\kappa>0$ such that

$$\sum_{d_1,d_2\le n^\kappa}\sum\lambda_{\mathbf{d}}\Big|\#\mathcal{A}_{\mathbf{d}}(n)-\mathrm{Li}(n)\mathfrak{S}_{\mathbf{d}}(n)\Big|\ll_{C,A} n(\log n)^{-A}$$

where the singular series $\mathfrak{S}_{\mathbf{d}}(n)$ is given by

$$\mathfrak{S}_{\mathbf{d}}(n)=\sum_{\substack{\delta_1\delta_1'=d_1'\\(\delta\delta_1\delta_2,n)=1}}\sum_{\delta_2\delta_2'=d_2'}\frac{4}{\delta_1'\delta_2'}\prod_{p|\delta_1'\delta_2'}\Big(1-\frac{\chi_{-4}(p)}{p}\Big)^{-1}\frac{G_{\delta_1'\delta_2'}}{\phi(\delta_1'\delta_2')}\sum_{\substack{w\,(\mathrm{mod}\,\delta_1'\delta_2')\\(n-w,\delta_1'\delta_2')=1\\w\in\mathcal{G}_{\delta_1'\delta_2'}}}1$$
$$\times\sum_{b_d|(\delta\delta_1\delta_2)^\infty}\frac{\chi_{-4}(b_d)}{\phi(b_d\delta^2\delta_1^2\delta_2^2)}L(1,\chi_{-4})\prod_p\Big(1+\frac{\chi_{-4}(p)}{p(p-1)}\Big)\mathfrak{c}(d_1d_2n).$$

The sum over $w\,(\mathrm{mod}\,\delta_1'\delta_2')$ can be computed using (2.8) as

$$\prod_{\substack{p|\delta_1'\delta_2'\\p\nmid 2n}}\Big(\frac{p-1}{2}-\frac{\chi_{p^*}(n)+1}{2}\Big)\prod_{\substack{p|\delta_1'\delta_2'\\p|n}}\Big(\frac{p-1}{2}\Big)\mathbb{1}_{(n-1,\delta_1'\delta_2',2)=1}.$$

It is now a straightforward exercise with Euler products, noting that $\chi_{p^*}(n)=(n/p)$ for odd p and $L(1,\chi_{-4})=\pi/4$, to obtain the expression (6.2). This completes the proof.

6.3 Completion of the proof of Theorem 1.1. With Proposition 6.2 available, Theorem 1.1 follows now easily from application of a sieve. Let $\mathcal{A}=(a_k)$ be a finite sequence of non-negative real numbers. We introduce the notation

$$S(\mathcal{A},z)=\sum_{(k,P_1(z))=1}a_k,\quad P_q(z)=\prod_{p\le z,p\nmid q}p,\quad A_c=\sum_{k\equiv 0\,(\mathrm{mod}\,c)}a_k,\quad A=A_1. \tag{6.5}$$

With these notations, we recall the following lemma from sieve theory (see e.g. [Iw2, Theorem 1]).

LEMMA 6.3. *Let $\gamma, L > 0$ be some fixed constants. Let $\Omega(c)$ be a multiplicative function satisfying $0 \le \Omega(p) < p$ and*

$$\prod_{w\le p\le w'}\Big(1-\frac{\Omega(p)}{p}\Big)^{-1} \le \Big(\frac{\log w'}{\log w}\Big)^{\gamma}\Big(1+\frac{L}{\log w}\Big) \tag{6.6}$$

for all $2 \le w \le w'$. Then we have

$$S(\mathcal{A},z) \ge A\prod_{p\le z}\Big(1-\frac{\Omega(p)}{p}\Big)\Big(f_\gamma(s)-\frac{e^{\sqrt{L}}Q(s)}{(\log\mathcal{D})^{1/3}}\Big) - \sum_{\substack{c\le\mathcal{D}\\ c\mid\widetilde{P}_1(z)}}\Big|A_c-\frac{\Omega(c)}{c}A\Big|,$$

where $s = \log\mathcal{D}/\log z$, $Q(s) < \exp(-s\log s + s\log\log 3s + O(s))$, and $f_\gamma(s)$ is some continuous function such that $0 < f_\gamma(s) < 1$ and $f(s) = 1 + O(e^{-s})$ as $s \to \infty$.

We are going to sieve the sequence $\mathcal{A} = \mathcal{A}(n) = (a_k)$ where $a_k = \#\{k = x_1x_2 : p + x_1^2 + x_2^2 = n\}$. Then from Proposition 6.2, there exists some $\kappa > 0$ such that for $c \le n^\kappa$ with $\mu^2(c) = 1$ we have

$$A_c = \mu(c)\sum_{\substack{\mathbf{d}\\ p\mid d_1d_2\Leftrightarrow p\mid c}}\mu(d_1)\mu(d_2)\frac{\omega(\mathbf{d};n)}{d_1d_2}\mathfrak{S}(n)\operatorname{Li}(n) + r_c, \tag{6.7}$$

where

$$\begin{aligned}
\mathfrak{S}(n) &= \pi\prod_{p\nmid n}\Big(1+\frac{\chi_{-4}(p)}{p(p-1)}\Big)\prod_{p\mid n}\Big(1-\frac{\chi_{-4}(p)}{p}\Big),\\
\omega(\mathbf{d};n) &= \prod_{p\mid(d_1,d_2)}\Big(1-\frac{c_p(n)}{p-1}\Big)\prod_{\substack{p\nmid d_1d_2\\ p\nmid(d_1,d_2)\\ p\mid 2n}}\Big(1-\frac{\left(\frac{n}{p}\right)}{p-1}\Big)\\
&\quad\times\prod_{\substack{p\nmid n\\ p\mid d_1d_2}}\Big(1+\frac{\chi_{-4}(p)}{p(p-1)}\Big)^{-1}\prod_{\substack{p\mid n\\ p\mid 2d_1d_2}}\Big(1-\frac{\chi_{-4}(p)}{p}\Big)^{-1}\\
&:= \prod_{p^\nu\parallel d_1d_2}\omega_\nu(p;n),
\end{aligned}$$

and

$$\sum_{c\le n^\kappa}\Big|A_c-\mu(c)\sum_{\substack{\mathbf{d}\\ p\mid d_1d_2\Leftrightarrow p\mid c}}\mu(d_1)\mu(d_2)\frac{\omega(\mathbf{d};n)}{d_1d_2}\mathfrak{S}(n)\operatorname{Li}(n)\Big| \ll_K n(\log n)^{-K} \tag{6.8}$$

for any $K > 0$. We next examine the multiplicative function $\omega_\nu(\bullet; n)$. For odd p we compute explicitly

$$\omega_1(p;n) = \begin{cases} \Big(1 - \dfrac{(n/p)}{p-1}\Big)\Big(1 + \dfrac{\chi_{-4}(p)}{p(p-1)}\Big)^{-1}, & p \nmid n, \\ \Big(1 - \dfrac{\chi_{-4}(p)}{p}\Big)^{-1}, & p \mid n, \end{cases}$$

and

$$\omega_2(p;n) = \begin{cases} \Big(1 - \dfrac{c_p(n)}{p-1}\Big)\Big(1 + \dfrac{\chi_{-4}(p)}{p(p-1)}\Big)^{-1}, & p \nmid n, \\ \Big(1 - \dfrac{\chi_{-4}(p)}{p}\Big)^{-1}, & p \mid n, \end{cases}$$

and for $p = 2$ we have

$$\omega_1(2;n) = 1, \quad \omega_2(2;n) = 1 - \frac{c_2(n)}{p-1}.$$

Let $\Omega(\bullet; n)$ be the multiplicative function defined by

$$\Omega(p;n) = 2\omega_1(p;n) - \omega_2(p;n)p^{-1}$$

so that we have from (6.7) and (6.8) that for any $\mathcal{D} \le n^\kappa$

$$\sum_{c \le \mathcal{D}} \mu^2(c) \Big| A_c - \frac{\Omega(c;n)}{c} A \Big| \ll n(\log n)^{-K}$$

where $A = \mathfrak{S}(n)\operatorname{Li}(n)$. Note that we have

$$0 \le \Omega(p;n) < 3, \quad \Omega(p;n) = 2 + O(1/p).$$

Thus there exists some absolute constant L such that (6.6) holds for $\gamma = 2$. We choose $\mathcal{D} = n^{\kappa_0}$ for some absolute small constant κ_0 (e.g. $\kappa_0 = \kappa/3$) and choose $C = C(\kappa_0)$ sufficiently large such that $f_2(C\kappa_0) > 0$ holds. Since $\Omega(2;n) = 2$ if and only if $n \equiv 0 \pmod 2$ and $\Omega(3;n) = 3$ if and only if $n \equiv 2 \pmod 3$, it follows that $0 \le \Omega(p) < p$ if $n \equiv 1, 3 \pmod 6$. Then, on taking $s = C\kappa_0$, it follows from Lemma 6.3 that for $n \equiv 1, 3 \pmod 6$ we have

$$\begin{aligned} &S(\mathcal{A}, n^{1/C}) \\ &\ge \mathfrak{S}(n)\operatorname{Li}(n) \prod_{p \le n^{1/C}} \Big(1 - \frac{\Omega(p;n)}{p}\Big)\Big(f_2(C\kappa_0) - \frac{e^{\sqrt{L}} Q(C\kappa_0)}{(\log \mathcal{D})^{1/3}}\Big) - O(n(\log n)^{-A}) \\ &\gg \mathfrak{S}(n)\operatorname{Li}(n)(\log n)^{-2}. \end{aligned}$$

This completes the proof of Theorem 1.1.

6.4 Proof of Theorem 1.2. Denote

$$\mathcal{B} = \mathcal{B}(x) = (b_k), \quad b_k = \#\{k = x_1x_2 : p = x_1^2 + x_2^2 + 1, p \le x\}.$$

Then we have

$$\#\{(p, x_1, x_2) : p = x_1^2 + x_2^2 + 1, P^-(x_1x_2) \ge p^{1/C}, p \le x\} \ge S(\mathcal{B}, x^{1/C}),$$

and thus it is enough to prove

$$S(\mathcal{B}, x^{1/C}) \gg x(\log x)^{-3} \tag{6.9}$$

for some sufficiently large C. We can follow the proof of Theorem 1.1 to obtain (6.9). To be more precise, let $B_c = B_c(x) = \sum_{k \equiv 0 \,(\mathrm{mod}\, c)} b_k$. Then for $\mu^2(c) = 1$ we have

$$B_c = \mu(c) \sum_{\substack{\mathbf{d}=(d_1,d_2) \\ p \mid d_1d_2 \Leftarrow p \mid c}} \mu(d_1)\mu(d_2)\#\mathcal{B}_{\mathbf{d}}(x),$$

where

$$\mathcal{B}_{\mathbf{d}}(x) = \{(p, x_1, x_2) : p = d_1^2x_1^2 + d_2^2x_2^2 + 1, p \le x\}.$$

After an application of Lemma 2.1 and Theorem 3.3, we see that uniformly for $d_1, d_2 \le x^\kappa$ with κ sufficiently small we have

$$\#\mathcal{B}_{\mathbf{d}}(x) = \sum_{\delta_1\delta_1' = d_1'} \sum_{\delta_2\delta_2' = d_2'} \frac{w_{\delta_1'\delta_2'}}{h_{\delta_1'\delta_2'}} G_{\delta_1'\delta_2'} \sum_{w \in \mathcal{G}_{\delta_1'\delta_2'}} \sum_{\substack{x^{1-\kappa} \le p \le x \\ p \equiv 1 \,(\mathrm{mod}\, (\delta\delta_1'\delta_2')^2) \\ p-1 \equiv w \,(\mathrm{mod}\, \delta_1'\delta_2')}} r\left(\frac{p-1}{\delta^2\delta_1^2\delta_2^2}\right) + O(x^{1-\kappa+\epsilon}).$$

Then we can follow the proof of Proposition 6.2 to evaluate the main term and obtain that there exists some absolute constant $\kappa > 0$ such that uniformly for $d_1, d_2 \le n^\kappa$ and $\lambda_{\mathbf{d}} \ll \tau(d_1d_2)^{C'}$ we have

$$\sum_{d_1,d_2 \le x^\kappa} \sum \lambda_{\mathbf{d}} \left|\#\mathcal{B}_{\mathbf{d}}(x) - \mathrm{Li}(x)\mathfrak{S}_{\mathbf{d}}(1)\right| \ll_{C',A} x(\log x)^{-A}, \tag{6.10}$$

where $\mathfrak{S}_{\mathbf{d}}(n)$ is as in (6.2). With (6.10), we can follow the proof and notation in Subsection 6.3 with n replaced by 1 in the singular series calculations to see that

$$\sum_{c \le n^\kappa} \left|B_c - \frac{\Omega(c;1)}{c}\mathfrak{S}(1)\,\mathrm{Li}(x)\right| \ll_A x(\log x)^{-A}$$

for any $A > 0$. Note that have $\Omega(p;1) < p$ for all p and $\mathfrak{S}(1) > 0$. Lemma 6.3 together with (6.4) then yields (6.9) for C sufficiently large. For the second claim in Theorem 1.2, we write

$$R_C(n) = \#\{(x_1, x_2) : n = x_1^2 + x_2^2 + 1, P^-(x_1x_2) \ge n^{1/C}\}.$$

Since $R_C(n) \le r(n-1) \ll n^\epsilon$, we see that

$$\sum_{p\le x} \mathbb{1}_{R_C(p)>0} \gg x^{-\epsilon}\#\{(p,x_1,x_2) : p = x_1^2+x_2^2+1, P^-(x_1x_2) \ge p^{1/C}, p\le x\}$$
$$\gg x^{-\epsilon} S(\mathcal{B}, x^{1/C}) \gg x^{1-\epsilon'},$$

which shows that there are infinitely many primes that can be written as one plus the sum of two squares of integers having no more than $C/2$ prime factors.

7 Proof of Theorem 1.3

The rest of the paper is devoted to the proof of Theorem 1.3. The basic idea is similar to the proof of Theorem 1.1, but it is technically and structurally substantially more involved. Again we start with the distribution of smooth numbers in arithmetic progressions to large moduli.

7.1 Smooth numbers in arithmetic progressions. Denote

$$\Psi(x,y) = \sum_{m\in S(x,y)} 1, \quad \Psi_q(x,y) = \sum_{\substack{m\in S(x,y)\\(m,q)=1}} 1.$$

We also use the usual notation

$$u = \frac{\log x}{\log y}, \quad H(u) = \exp\Big(\frac{u}{\log^2(u+1)}\Big), \quad u \ge 1.$$

For an integer n we set $n_y = \prod_{\substack{p^\nu \| n\\ p\le y}} p^\nu$ and $n_o = \frac{n}{(n,2^\infty)}$. We have the following analogue of Proposition 6.1 for smooth numbers.

PROPOSITION 7.1. *Suppose $|\lambda_d| \ll \tau(d)^C$ for some fixed $C > 0$. Then there exist some constants $\varpi, \delta, D > 0$ such that the following holds. Let $x \ge 2$, $c_0, c, d \in \mathbb{N}, (c_0,c) = 1, a_1, \ell \in \mathbb{Z}\backslash\{0\}$ such that for*

$$Q \le x^{1/2+\varpi}, \quad |a_1| \le x^{1+\varpi}, \quad |a_2| \le x^\varpi,$$
$$(\log x)^D \le y \le x, \quad a_2 \mid \ell, \quad \omega(\ell_y) \ll \log x$$

we have

$$\sum_{d\le x^\varpi} \lambda_d \sup_{w \,(\mathrm{mod}\, d)^\times} \Big| \sum_{\substack{q\le Q\\(q,a_1a_2d)=1\\ q\equiv c_0 \,(\mathrm{mod}\, c)}} \Big(\sum_{\substack{m\in S(x,y)\\ \mu^2(m_o)=(m,\ell)=1\\ m\equiv a_1\overline{a_2a_3} \,(\mathrm{mod}\, q)\\ m\equiv w \,(\mathrm{mod}\, d)}} 1 - \frac{1}{\phi(qd)} \sum_{\substack{m\in S(x,y)\\ \mu^2(m_o)=(m,qd\ell)=1}} 1\Big)\Big|$$
$$\ll_{C,A} c\Psi_\ell(x,y)(\log x)^{-A}$$

for any $A > 0$.

REMARK. One of the important features of Proposition 7.1 is that we make the dependence on ℓ explicit. This is necessary for our application since $\Psi_\ell(x,y) \asymp \prod_{p|\ell_y}(1-p^{-\alpha})\Psi(x,y)$ with $\alpha = \alpha(x,y)$ as in (1.6), which could be much smaller than $\Psi(x,y)(\log x)^{-A}$ when y is small. Furthermore, the condition $a_2 \mid \ell$ ensures that $(n,a_2)=1$ in an application of Proposition 5.1.

As a preparation for the proof we recall the following result that is a variation of the result in [Ha, Section 3.3] or [Dr1, Lemme 5].

LEMMA 7.2. *Suppose $|\lambda_q| \ll \tau(q)^C$ for some fixed $C>0$ and $\ell \in \mathbb{Z}\backslash\{0\}$. There exist some constants $D,\eta,\delta>0$ such that the following is true. If $(\log x)^D \le y \le x$, $q \le Q \le x$ and $\omega(\ell_y) \ll \log x$, we have*

$$\sum_{q\le Q}\frac{\lambda_q}{\phi(q)}\sum_{\substack{\chi \,(\mathrm{mod}\, q)\\ 1<\mathrm{cond}(\chi)\le x^\eta}}\Big|\sum_{\substack{m\in S(x,y)\\ (m,\ell)=1}}\chi(m)\Big| \ll_{C,A} \Psi_\ell(x,y)(H(u)^{-\delta}(\log x)^{-A}+y^{-\delta})$$

for any $A>0$.

Proof. When $\lambda_q=1$, $\ell=1$, this can be found in [Ha, Section 3.3]. For general coefficients λ_q we can follow the same argument using that

$$\sum_{q\le Q}\frac{|\lambda_q|}{\phi(q)} \ll (\log x)^{O(1)} \quad \text{and} \quad \lambda_q \ll q^\varepsilon \ \text{ for any } \varepsilon>0.$$

We now deal with general values for ℓ. We recall [IBT, Theorem 2.4, Lemme 3.1, eq (2.8)]: we have uniformly in $\mathcal{L}(x) \le y \le x$, $1 \le d \le x/y$ $P^+(\ell) \le y$ and $\omega(\ell) \le \sqrt{y}$ the estimate

$$\Psi_\ell(x/d,y) \ll \prod_{p|\ell}(1-p^{-\alpha})\frac{\Psi(x,y)}{d^\alpha} \tag{7.1}$$

where α is as in (1.6) satisfying

$$\alpha = \alpha(x,y) = 1 - \frac{\xi(u)+O(1)}{\log y}, \quad u = \frac{\log x}{\log y} \tag{7.2}$$

and $\xi(t)$ is defined implicitly by

$$e^{\xi(t)} = 1 + t\xi(t).$$

It is easy to see that $\xi(u) = \log(u\log u)+O(1)$. When $\exp((\log x)^{2/5}) < y \le x$, we have

$$\prod_{p|\ell_y}\Big(1-\frac{1}{p^\alpha}\Big) \gg \exp\Big(-\sum_{p\le y}p^{-\alpha}\Big) \gg \exp\Big(-\frac{u\log u}{\log y}\Big) \gg 1,$$

which implies that $\Psi_\ell(x,y) \asymp \Psi(x,y)$. Thus, assuming without loss of generality that $p \mid \ell \Rightarrow p \le y$, and after using Möbius inversion to detect the condition $(m,\ell)=1$, it is enough to prove that

$$\sum_{q\le Q} \frac{\lambda_q}{\phi(q)} \sum_{\substack{\chi \,(\mathrm{mod}\, q)\\ 1<\mathrm{cond}(\chi)\le x^\eta}} \sum_{d\mid \ell} \Big| \sum_{m\le x/d} \chi(m)\Big| \ll \Psi(x,y)\big(H(u)^{-\delta}(\log x)^{-A} + y^{\delta}\big). \quad (7.3)$$

The contribution from $d \ge x^{c_0}$ can be bounded trivially by $x^{1-c_0+\eta+\epsilon}$. The contribution from $d \le x^{c_0}$ can be bounded individually using the result for $\ell = 1$ (and adjusting the constant η) so that together with (7.1) we obtain the left hand side of (7.3) can be bounded by

$$\begin{aligned}
&\sum_{d\mid \ell} \sum_{q\le Q} \frac{\lambda_q}{\phi(q)} \sum_{\substack{\chi \,(\mathrm{mod}\, q)\\ 1<\mathrm{cond}(\chi)\le x^\eta}} \Big| \sum_{m\le x/d} \chi(m)\Big| \\
&\ll \sum_{\substack{d\mid \ell\\ d\le x^{c_0}}} \Psi\Big(\frac{x}{d}, y\Big)\Big(H\Big(u - \frac{\log d}{\log y}\Big)^{-\delta}(\log x)^{-A} + y^{-\delta}\Big) + O(x^{1-c_0+\eta+\epsilon}) \\
&\ll \sum_{d\mid \ell} \frac{1}{d^\alpha} d^{\frac{2\delta}{\log y(\log u)^2}} \Psi(x,y)\big(H(u)^{-\delta}(\log x)^{-A} + y^{-\delta}\big) + O(x^{1-c_0+\eta+\epsilon}) \\
&\ll \Psi(x,y)\big(H(u)^{-\delta}(\log x)^{-A} + y^{-\delta}\big),
\end{aligned}$$

where we used the the assumption of ℓ and the range of y in the last step.

It remains to incorporate the condition $(m,\ell)=1$ for $(\log x)^D \le y \le \exp((\log x)^{2/5})$. To this end, we first generalize [Ha, Proposition 1, Theorem 3] with $\Psi(x,y)$ replaced by $\Psi_\ell(x,y)$. This can be done by following the proof of [Ha, Proposition 1, Theorem 3] and replacing $L(s,\chi;y) = \prod_{p\le y}(1-\chi(p)p^{-s})^{-1}$ by

$$L_\ell(s,\chi;y) := \prod_{p\nmid \ell, p\le y} (1-\chi(p)p^{-s})^{-1}.$$

Using $\omega(\ell_y) \ll \log x$ we see that the contribution from $\sum_{(n,\ell)>1} \Lambda(n)\chi(n)n^{-\sigma}$ can be bounded by

$$\sum_{n \ll \log x \log_2 x} \Lambda(n) n^{-\sigma} \ll \frac{(\log x \log_2 x)^{1-\sigma}}{1-\sigma},$$

which is $o(\log x)$ when $\sigma \in [\alpha - 1/300, \alpha]$ using (7.2). Thus we still have the bound

$$|\log L_\ell(\sigma+it,\chi;y) - \log L_\ell(\alpha+it,\chi;y)| \le \frac{(\alpha-\sigma)\log x}{2},$$

which is enough to prove the analogue of [Ha, Theorem 3] with $\Psi(x,y)$ replaced by $\Psi_\ell(x,y)$ when the summand is restricted to $(m,\ell)=1$. We can also replace [Ha,

Smooth Numbers Result 1, 3] by [lBT, Theorem 2.1, 2.4]. Then we can follow the proof in [Ha, Section 3.3] in the case $\ell = 1$ and replace $\Psi(x, y)$ by $\Psi_\ell(x, y)$ so that condition $(m, \ell) = 1$ is preserved. On the way we need the estimate

$$\sum_{\substack{p\le y\\ p\nmid q\ell}} \frac{1-\Re(\chi(p)p^{-it})}{p^\alpha} \gg \frac{u}{\log^2(u+1)} \tag{7.4}$$

which was proved on [Ha, pp. 16-17] when $\ell = 1$. For general ℓ, we simply note that the contribution from $p \mid \ell$ can be trivially bounded by

$$\sum_{p\mid\ell} \frac{1}{p^\alpha} \ll \frac{(\sqrt{y}\log\sqrt{y})^{1-\alpha}}{(1-\alpha)\log y} \ll \sqrt{u}$$

so that (7.4) still holds since $u \gg (\log_2 x)^2$ by our current assumption on y.

Proof of Proposition 7.1:. The proof follows by combining the proofs of [FT1, Lemme 2.1] and [ABL, Theorem 2.1]. We use ε as an arbitrary positive constant, not necessarily the same at each occurrence. Let $R = x^\eta$ for some sufficiently small η to be determined later. We use the notation $\mathfrak{u}_R(n; q, a, d, w)$ as defined in (5.1) to write

$$\sum_{d\le x^\varpi} \lambda_d \sup_{\substack{w \,(\mathrm{mod}\, d)^\times\\ a_3\mid d}} \Big| \sum_{\substack{q\le Q\\ (q,a_1a_2d)=1\\ q\equiv c_0 \,(\mathrm{mod}\, c)}} \Big(\sum_{\substack{m\in S(x,y)\\ \mu^2(m_o)=(m,\ell)=1\\ m\equiv a_1\overline{a_2a_3} \,(\mathrm{mod}\, q)\\ m\equiv w \,(\mathrm{mod}\, d)}} 1 - \frac{1}{\phi(qd)} \sum_{\substack{m\in S(x,y)\\ \mu^2(m_o)=(m,qd\ell)=1}} 1\Big)\Big|$$

$$\le \mathcal{S}_1 + \mathcal{S}_2,$$

where

$$\mathcal{S}_1 = \sum_{d\le x^\varpi} |\lambda_d| \sup_{\substack{w \,(\mathrm{mod}\, d)^\times\\ a_3\mid d}} \Big| \sum_{\substack{q\le Q\\ (q,a_1a_2d)=1\\ q\equiv c_0 \,(\mathrm{mod}\, c)}} \sum_{\substack{m\in S(x,y)\\ \mu^2(m_o)=(m,\ell)=1}} \frac{1}{\phi(q)\phi(d)} \times \sum_{\substack{\chi \,(\mathrm{mod}\, q)\\ \psi \,(\mathrm{mod}\, d)\\ 1<\mathrm{cond}(\chi\psi)\le R}} \chi(m\overline{a_1}a_2a_3)\psi(m\overline{w})\Big|,$$

$$\mathcal{S}_2 = \sum_{d\le x^\varpi} |\lambda_d| \sup_{\substack{w \,(\mathrm{mod}\, d)^\times\\ a_3\mid d}} \Big| \sum_{\substack{q\le Q\\ (q,a_1a_2d)=1\\ q\equiv c_0 \,(\mathrm{mod}\, c)}} \frac{1}{\phi(d)} \sum_{\substack{m\in S(x,y)\\ \mu^2(m_o)=(m,\ell)=1}} u_R(m; q, a_1\overline{a_2a_3}, d, q)\Big|.$$

For $\mathcal{S}_1$, we use Lemma 7.2 and $(q,d)=1$ to conclude that for η small enough,

$$\begin{aligned}\mathcal{S}_1 &\ll \sum_{d\le x^{\varpi}} |\lambda_d| \sup_{\substack{w \,(\mathrm{mod}\, d)^\times \\ a_3 | d}} \Big| \sum_{\substack{q\le Q \\ (q,a_2a_2d)=1 \\ q\equiv c_0 \,(\mathrm{mod}\, c)}} \frac{1}{\phi(qd)} \\ &\qquad \times \sum_{\substack{\chi \,(\mathrm{mod}\, q) \\ \psi \,(\mathrm{mod}\, d) \\ 1<\mathrm{cond}(\chi\psi)\le R}} \sum_{\substack{m\in S(x,y) \\ \mu^2(m_o)=(m,\ell)=1}} \chi(m\overline{a_1}a_2a_3)\psi(m\overline{w})\Big| \\ &\ll \sum_{k\le x^{\eta}} \sum_{q\le Qx^{\varpi}} \frac{\tilde{\lambda}_q}{\phi(q)} \sum_{\substack{\chi \,(\mathrm{mod}\, q) \\ 1<\mathrm{cond}(\chi)\le R}} \Big| \sum_{\substack{m\in S(x/k^2,y) \\ (m,\ell)=1}} \chi(m)\Big| + O(x^{1-\eta+\epsilon}) \\ &\ll \sum_{k\le x^{\eta}} \Psi_\ell(x/k^2,y)\Big(H\Big(u-\frac{2\log k}{\log y}\Big)^{-\delta}(\log x)^{-A} + y^{-\delta}\Big) \\ &\ll \Psi_\ell(x,y)(\log x)^{-A'}\end{aligned}$$

for some large constant A' depending on y,δ. Here we used (7.1), $y\ge(\log x)^D$ and $\alpha\ge 1-1/D+o(1)$ in the last step. Using ideas as in the proof of Theorem 3.4, we now complete the argument by showing that there exists some $\varpi,\delta_0>0$ such that

$$\mathcal{S}_2 \ll cx(\log x)^{O(1)}R^{-\delta_0}. \tag{7.5}$$

To prove (7.5), we consider separately the cases $y\ge x^\eta$ and $y<x^\eta$ for some sufficiently small $\eta>0$.

When $y\ge x^\eta$, we use Buchstab's identity (3.4) to write

$$\mathcal{S}_2 \le S_0 + \sum_{i=1}^{3} S_i + \overline{S},$$

where

$$\begin{aligned}S_0 = \sum_{d\le x^{\varpi}} |\lambda_d| \sup_{\substack{w \,(\mathrm{mod}\, d)^\times \\ a_3 | d}} \Big| \sum_{\substack{q\le Q \\ (q,a_1a_2d)=1 \\ q\equiv c_0 \,(\mathrm{mod}\, c)}} \Big(\sum_{\substack{m\equiv \overline{a_1}a_2a_3 \\ \mu^2(m_o)=(m,\ell)=1 \\ m\le x}} 1 \\ - \frac{1}{\phi(qd)} \sum_{\substack{\chi \,(\mathrm{mod}\, q) \\ \psi \,(\mathrm{mod}\, d) \\ \mathrm{cond}(\chi\psi)\le R}} \sum_{\substack{\mu^2(m_o)=1 \\ (m,\ell)=1 \\ m\le x}} \chi(m\overline{a_1}a_2a_3)\psi(m\overline{w})\Big)\Big|,\end{aligned}$$

$$S_j = \sum_{d\le x^{\varpi}} |\lambda_d| \sup_{\substack{w \,(\mathrm{mod}\, d)^\times \\ a_3|d}} \Big| \sum_{\substack{q\le Q \\ (q,a_1a_2d)=1 \\ q\equiv c_0 \,(\mathrm{mod}\, c)}} \sum_{\substack{y<p_1<\cdots<p_j \\ \mu^2(m_o p_1\cdots p_j)=1 \\ (mp_1\cdots p_j,\ell)=1 \\ mp_1\cdots p_j\le x}} u_R(mp_1\cdots p_j; q, a_1\overline{a_2a_3}, d, w)\Big|,$$

$$\overline{S} = \sum_{d\le x^{\varpi}} |\lambda_d| \sup_{\substack{w \,(\mathrm{mod}\, d)^\times \\ a_3|d}} \Big| \sum_{\substack{q\le Q \\ (q,a_1a_2d)=1 \\ q\equiv c_0 \,(\mathrm{mod}\, c)}} \sum_{\substack{y<p_1<\cdots<p_4 \\ \mu^2(m_o p_1\cdots p_4)=1 \\ (mp_1\cdots p_4,\ell)=1 \\ mp_1\cdots p_4\le x}} u_R(mp_1\cdots p_4; q, a_1\overline{a_2a_3}, d, w) g_{p_4}(m)\Big|.$$

Consider $\overline{S}$ first. Note that we can assume $y < p_1 \ll x^{1/4}$ since otherwise the sum is empty. The condition $\mu^2(m_o p_1\cdots p_4) = 1$ is equivalent to $\mu^2(m_o) = (m, p_1\cdots p_4) = 1$. We use Möbius inversion to detect $(m, p_1\cdots p_4) = 1$. Since we have that the primes $p_i > y \ge x^{\eta}$, the contribution from $k \mid p_1\cdots p_4$ with $k < x^{\eta/2}$ comes only from the term $k = 1$. Similar to the proof of Theorem 3.4, we localize m and p_i into intervals $m \in (M, M(1+\mathcal{Z})]$ and $p_i \in (P_j, P_{j+1}]$ with $P_j = y(1+\mathcal{Z})^j$ and with $\mathcal{Z} = R^{-\eta_0}$ for some $\eta_0 > 0$ so that

$$\overline{S} \ll \mathcal{Z}^{-4}(\log x)^{O(1)} \sup_{M, P_{j_i}} \tilde{\mathcal{S}} + x(\log x)^{O(1)}\mathcal{Z} + O(x^{1-\eta/2+\epsilon})$$

where

$$\tilde{\mathcal{S}} = \sum_{d\le x^{\varpi}} |\lambda_d| \sup_{\substack{w \,(\mathrm{mod}\, d)^\times \\ a_3|d}} \Big| \sum_{\substack{q\le Q \\ (q,a_1a_2d)=1 \\ q\equiv c_0 \,(\mathrm{mod}\, c)}} \sum_{\substack{p_i\in(P_{j_i},(1+\mathcal{Z})P_{j_i}] \\ m\in(M,M(1+\mathcal{Z})] \\ (mp_1\cdots p_4,\ell)=1 \\ 1\le i\le 4}} g_{p_4}(m)\mu^2(m_o) \times u_R(mp_1\cdots p_4; q, a_1\overline{a_2}, d, w)\Big|.$$

If $M\prod_{i=1}^4 P_{j_i} \le x^{1-\eta/2}$, we can bound $\tilde{\mathcal{S}} \ll x^{1-\eta/2+\varepsilon}$ trivially. Otherwise, we can apply the bilinear estimate Proposition 5.1 with $\boldsymbol{m} = p_1$ to $\tilde{\mathcal{S}}$ since $x^{\eta} \le y \le P_{j_1} \ll x^{1/4} \ll (M\prod_{i=1}^4 P_{j_i})^{1/4+\eta}$. We obtain

$$\overline{S} \ll \mathcal{Z}^{-4} cx(\log x)^{O(1)} R^{-1/3} + x(\log x)^{O(1)}\mathcal{Z} + O(x^{1-\eta/2+\epsilon})$$

as long as

$$Q \le x^{1/2+\varpi_0}, \quad c, R, |a_2| \le x^{\varpi_0}, \quad a_1 \le x^{1+\varpi_0} \tag{7.6}$$

for some constant $\varpi_0 > 0$. This suffices for the bound in (7.5) by choosing a suitable η_0.

For S_k, $1 \le k \le 3$, we can use Möbius to detect $(m, p_1\cdots p_j) = 1$ as before, replace the characteristic function of the primes by the von Mangoldt function,

apply Heath-Brown's identity (3.1) to each prime variable p_i and then localize the new variables so that we have

$$S_k \ll \mathcal{Z}^{-24}(\log x)^{O(1)} \sup_{\substack{M,M_i,N_j \\ 1\le r\le 12 \\ 1\le s\le 12}} \dot{\mathcal{S}} + x(\log x)^{(1)}\mathcal{Z} + O(x^{1-\eta/2+\epsilon}),$$

where

$$\dot{\mathcal{S}} = \sum_{d\le x^{\varpi}} \sup_{\substack{w \,(\mathrm{mod}\, d)^\times \\ a_3|d}} \Big| \sum_{\substack{q\le Q \\ (q,a_1a_2d)=1 \\ q\equiv c_0 \,(\mathrm{mod}\, c)}} \sum\sum\sum_{\substack{m\in(M,(1+\mathcal{Z})M] \\ m_i\in(M_i,(1+\mathcal{Z})M_i] \\ n_j\in(N_j,(1+\mathcal{Z})N_j] \\ (mm_in_j,\ell)=1\forall i,j}} \mu^2(m_o) \times \prod_{i=1}^{r}\mu(m_i)u_R\Big(m\prod_{i=1}^{r}m_i\prod_{j=1}^{s}n_j; q, a_1\overline{a_2a_3}, d, w\Big)\Big|$$

with

$$M\prod_{i=1}^{r}M_i\prod_{j=1}^{s}N_j \ll x, \quad M_i \le x^{1/4}.$$

As before, we can assume $M\prod_{i=1}^{r}M_i\prod_{j=1}^{s}N_j \ge x^{1-\eta/2}$ since otherwise we can bound $\dot{\mathcal{S}} \ll x^{1-\eta/2+\varepsilon}$ trivially. Now we argue similarly as in the proof of Theorem 3.4. If there is a subset $\mathcal{I}\subseteq\{M, M_i, N_j\}$ such that $K := \prod_{I\in\mathcal{I}} I$ such that $x^\eta \le K \le x^{1/4+\eta}$, then we can apply Proposition 5.1 to obtain

$$\dot{\mathcal{S}} \ll cx(\log x)^{O(1)}R^{-1/3}$$

as long as (7.6) is satisfied. Otherwise, we can use Möbius inversion to detect $\mu^2(m_o) = 1$ if necessary, and it is enough to show there exists some constant $\varpi > 0$ such that

$$\sup_{\substack{w \,(\mathrm{mod}\, d)^\times \\ a_3|d}} \Big| \sum_{\substack{q\le Q \\ (q,a_1a_2d)=1 \\ q\equiv c_0 \,(\mathrm{mod}\, c)}} \sum\sum\sum_{\substack{m\asymp M,k\asymp K,l\asymp L \\ (mkl,\ell)=1}} \alpha_r u_R(rmkl; q, a_1\overline{a_2a_3}, d, w)\Big| \ll x^{1-\varpi} \tag{7.7}$$

for $r \le x^\eta, x^{1-\eta/2} \ll rMKL \le x$ and $\alpha_r \ll r^\varepsilon$. After an application of Möbius inversion to detect the condition $(mkl, \ell) = 1$, we are left with

$$\sup_{\substack{w \,(\mathrm{mod}\, d)^\times \\ a_3|d}} \Big| \sum_{\lambda|\ell}\mu(d) \sum_{\substack{q\le Q \\ (q,a_1a_2d)=1 \\ q\equiv c_0 \,(\mathrm{mod}\, c)}} \sum\sum\sum_{m\asymp M,k\asymp K,l\asymp L} \alpha_r u_R(r\lambda mkl; q, a_1\overline{a_2a_3}, d, w)\Big| \ll x^{1-\varpi}$$

for $r \le x^\eta, x^{1-\eta/2} \ll r\lambda MKL \le x$ and $\alpha_r \ll r^\varepsilon$. The contribution from $\lambda > x^{\eta'}$ can be bounded trivially by $x^{1-\eta'/2}$. The contribution for any fixed $\lambda < x^{\eta'}$ can be rewritten and then estimated in the same way as in [ABL, eq. (4.5)], where ultimately

Deligne's estimates for exponential sums over algebraic varieties over finite fields are used. By choosing $\mathcal{Z}, \eta'$ suitably, we see that there exists some $\varpi, \delta_0 > 0$ such that $S_k \ll cx(\log x)^{O(1)} R^{-\delta_0}$ for $1 \le k \le 3$.

We can also treat S_0 by (7.7), which completes the proof of (7.5) when $y \ge x^\eta$.

When $y \le x^\eta$, we can apply [FT3, Lemme 3.1] in the form of (3.10) so that it is enough to show that there exists some absolute constant $\varpi, \delta_0 > 0$ such that

$$\begin{aligned}&\sum_{d\le x^{\varpi}} \lambda_d \sup_{\substack{w \,(\mathrm{mod}\, d)^\times \\ a_3|d}} \Big| \sum_{\substack{q\le Q \\ (q,a_1a_2d) \\ q\equiv c_0 \,(\mathrm{mod}\, c)}} \sum_{\substack{lm\le x,(lm,\ell)=1 \\ \mu^2((lm)_o)=1 \\ P^+(l)\le P^-(m) \\ M\le m\le yM}} g_y(l)g_y(m)u_R(lm;q,a_1\overline{a_2a_3},d,w)\Big| \\ &\ll cx(\log x)^{O(1)}R^{-\delta_0}.\end{aligned} \tag{7.8}$$

Since $P^+(l) \le P^-(m)$ we see that

$$\begin{aligned}&\sum_{d\le x^{\varpi}} \lambda_d \sup_{\substack{w \,(\mathrm{mod}\, d)^\times \\ a_3|d}} \Big| \sum_{\substack{q\le Q \\ (q,a_1a_2d) \\ q\equiv c_0 \,(\mathrm{mod}\, c)}} \sum_{\substack{lm\le x,(lm,\ell)=1 \\ \mu^2((lm)_o)=1 \\ P^+(l)\le P^-(m) \\ M\le m\le yM}} g_y(l)g_y(m) \\ &\qquad\qquad \times u_R(lm;q,a_1\overline{a_2a_3},d,w)\Big| \\ &\ll \sum_{d\le x^{\varpi}} |\lambda_d| \sup_{\substack{w \,(\mathrm{mod}\, d)^\times \\ a_3|d}} \Big| \sum_{\substack{q\le Q \\ (q,a_1a_2d) \\ q\equiv c_0 \,(\mathrm{mod}\, c)}} \sum_{\substack{lm\le x,(lm,\ell)=1 \\ P^+(l)<P^-(m) \\ M\le m\le yM}} \mu^2(l_o)g_y(l)\mu^2(m_o) \\ &\qquad\qquad \times g_y(m)u_R(lm;q,a_1\overline{a_2a_3},d,w)\Big| \\ &\quad + \sum_{d\le x^{\varpi}} |\lambda_d| \sup_{\substack{w \,(\mathrm{mod}\, d)^\times \\ a_3|d}} \Big| \sum_{\substack{q\le Q \\ (q,a_1a_2d) \\ q\equiv c_0 \,(\mathrm{mod}\, c)}} \sum_{p\le y} \sum_{\substack{lm\le x/p^2,(lm,\ell)=1 \\ P^+(l)<p<P^-(m) \\ M\le mp\le yM}} \mu^2(l_o)g_y(l) \\ &\qquad\qquad \times \mu^2(m_o)g_y(m)u_R(lm;q,a_1\overline{a_2a_3p^2},d,w)\Big|.\end{aligned}$$

After separating variables l and m as in [Dr1, eq. (3.37)] and localizing l, m into short intervals $l \asymp L$ and $m \asymp M$, we can apply Proposition 5.1 for the first sum above to obtain (7.8), as long as $x^{\eta_1} \le M$ and $My \le (LM)^{1/4+\eta_1}$, which can be satisfied by choosing $M = x^{\eta_2}$ for some suitable $\eta_2 > 0$ when $LM \ge x^{1-\eta_3}$. If $LM \le x^{1-\eta_3}$, we can obtain (7.8) trivially. For the second sum, we move the summation of $p \le y$ outside so that contribution from $p \le x^{\eta_4}$ can be estimated using Proposition 5.1

similarly as above and the contribution from $y \geq x^{\eta_4}$ can be bounded trivially. This complete the proof of (7.5) when $y \leq x^{\eta}$.

7.2 Implementing sieve weights. The most straightforward approach to Theorem 1.3 would be to use Proposition 7.1 to evaluate

$$\sum_{m \in S(n,y)} r_{d_1,d_2}(n-m)$$

for square-free numbers d_1, d_2 and $y = g(n)$, and then apply a sieve. Unfortunately the main term in the asymptotic formula is not a multiplicative function of $d_1 d_2$. As mentioned in the introduction, this makes it problematic to implement a sieve, so instead we will work with the sieve weights directly. We recall [Iw3, Lemma 3].

LEMMA 7.3. *Let $\gamma, L > 0$ be some fixed constants. Then there exists a sequence λ_c^- supported on square-free integers less than $\mathcal{D}$ such that*

$$\lambda_1^- = 1, \quad |\lambda_c^-| \leq 1, \quad \mu * 1 \geq \lambda^- * 1$$

and that for all multiplicative functions w satisfying $0 \leq w(p) < 1$ and

$$\prod_{W \leq p \leq W'} (1 - w(p))^{-1} \leq \Big(\frac{\log W'}{\log W}\Big)^{\gamma} \Big(1 + \frac{L}{\log w}\Big)$$

for all $2 \leq W \leq W'$ we have

$$\sum_{c | P(z)} \lambda_c^- w(c) \geq \prod_{p \leq z} (1 - w(p)) \Big(f_\gamma(s) + O\Big(\frac{e^{\sqrt{L} - s}}{(\log \mathcal{D})^{1/3}} \Big) \Big), \quad s = \frac{\log \mathcal{D}}{\log z} \geq 2,$$

where $f_\gamma(s)$ is some continuous function such that $0 < f_\gamma(s) < 1$ and $f_\gamma(s) = 1 + O(e^{-s})$ as $s \to \infty$.

The idea is to estimate the sieve weights λ_c^- with multiplicative factors $w(c)$ first to get a lower bound using Lemma 7.3, and then sum over the smooth numbers m. In order to avoid major technical difficulties, we need to prepare the set-up very carefully.

- Many estimates are sensitive to the prime factorization of m, and it simplifies our life to restrict m to be square-free. On the other hand, if $n \equiv 2 \pmod 4$, we will need $4 \mid m$ in order to represent n by m and two odd integer squares. Thus we only restrict m to have an odd square-free part. This restriction will not change the order of the number of solutions, but the fact that square-free numbers are not equidistributed among *all* residue classes will leads to different constants for different n.

- It turns out that a factor $\chi_{-4}(h)$ with $h \mid (n,m)$ appears in the main term. It is convenient to force $\chi_{-4}(h) > 0$ (so that a lower bound suffices), and thus we also restrict m to be coprime to all primes $p \equiv 3 \pmod 4$ dividing n. Since $n-m$ is a sum of two squares, this is a harmless maneuver and affects only square factors, but it makes an important difference for the computation if we implement this. Note that we could simplify the computations by restricting m to be coprime to n, but this could lose some order of magnitude for certain cases of n, e.g. when n is the product of the first few primes that are $1 \pmod 4$.
- Finally, the sieve weights behave a bit erratically for small primes, so we sieve them out directly by Möbius inversion.

With these general remarks in mind, we fix the following notation. We write

$$\mathfrak{N} = \prod_{\substack{p \mid n \\ p \equiv 3 \,(\mathrm{mod}\, 4)}} p, \quad m_o = \prod_{\substack{p^\nu \| m \\ p > 2}} p^\nu = \frac{m}{(m, 2^\infty)}.$$

Recall the notation (6.5). Let

$$\mathcal{A} = \mathcal{A}(n) = (a_k), \quad a_k = \#\Big\{k = x_1 x_2 : \begin{matrix} m + x_1^2 + x_2^2 = n,\ \mu^2(m_o) = 1, \\ P^+(m) \le y, \qquad (m, \mathfrak{N}) = 1 \end{matrix}\Big\}.$$

Then for c square-free, we have

$$\mathcal{A}_c = \mu(c) \sum_{\substack{\mathbf{d} \\ p \mid d_1 d_2 \Leftrightarrow p \mid c}} \mu(d_1)\mu(d_2) \sum_{\substack{m \le n \\ P^+(m) \le y \\ \mu^2(m_o) = (m, \mathfrak{N}) = 1}} r_{d_1, d_2}(n-m).$$

Let $\mathcal{Q}$ be some fixed absolute constant (we will later choose $\mathcal{Q} = 30$). Let λ_c^- be some lower bound sieve weights as in Lemma 7.3 of dimension γ, level $\mathcal{D} = n^{\kappa_0}$ for some sufficiently small κ_0 and $z = n^{1/C}$ for some sufficiently large C so that $f_\gamma(C\kappa_0) > 0$. Then we have the following lower bound for $S(\mathcal{A}, z)$:

$$\begin{aligned} S(\mathcal{A}, z) &= \sum_{\substack{m \le n \\ P^+(m) \le y \\ \mu^2(m_o) = (m, \mathfrak{N}) = 1}} \sum_{\substack{x_1^2 + x_2^2 = n - m \\ (x_1 x_2, \mathcal{Q}) = 1}} \sum_{c \mid (x_1 x_2, P_{\mathcal{Q}}(z))} \mu(c) \\ &\ge \sum_{\substack{m \le n \\ P^+(m) \le y \\ \mu^2(m_o) = (m, \mathfrak{N}) = 1}} \sum_{x_1^2 + x_2^2 = n - m} \sum_{\substack{u_1, u_2 \mid \mathcal{Q} \\ u_i \mid x_i}} \mu(u_1)\mu(u_2) \sum_{c \mid (x_1 x_2, P_{\mathcal{Q}}(z))} \lambda_c^- \\ &= \sum_{u \mid \mathcal{Q}} \sum_{c \mid P_{\mathcal{Q}}(z)} \lambda_c^- \mu(c) \sum_{\substack{d_1, d_2 \\ p \mid d_1 d_2 \Leftrightarrow p \mid uc}} \mu(d_1)\mu(d_2) \sum_{\substack{m \le n \\ P^+(m) \le y \\ \mu^2(m_o) = (m, \mathfrak{N}) = 1}} r_{d_1, d_2}(n-m). \end{aligned} \tag{7.9}$$

Our next goal is to evaluate the innermost sum using Lemma 2.2.

7.3 The cuspidal contribution. We first treat the contribution to the m-sum in (7.9) that corresponds to (2.9). To this end we need to show that there exists some absolute constant $\eta > 0$ (which eventually will depend on the constants in Theorem 3.4) such that

$$\sum_{\substack{m\le n\\ m\equiv n \,(\mathrm{mod}\,\Delta)\\ m\equiv w \,(\mathrm{mod}\, q)\\ \mu^2(m_o)=(m,\mathfrak{N})=1\\ P^+(m)\le y}} \lambda\Big(\frac{n-m}{\Delta}\Big) \ll n^{1-\eta} \tag{7.10}$$

uniformly for $\Delta, q \le n^\eta$ with $(\Delta w, q) = 1$ (specifically, in the notation of (2.9) we have $\Delta = (\delta\delta_1\delta_2)^2$, $q = \delta_1'\delta_2'$) and Hecke eigenvalues of a cusp form ϕ whose conductor is less than n^η. By Möbius inversion, the previous display equals

$$\sum_{\substack{d\le\sqrt{n}\\ (d,2\mathfrak{N})=1\\ P^+(d)\le y}} \mu(d) \sum_{\substack{m\le n/d^2\\ d^2m\equiv n \,(\mathrm{mod}\,\Delta)\\ d^2m\equiv w \,(\mathrm{mod}\, q)\\ (m,\mathfrak{N})=1\\ P^+(m)\le y}} \lambda\Big(\frac{n-d^2m}{\Delta}\Big).$$

The contribution from $d \ge n^{4\eta}$ can be bounded by

$$\sum_{d\ge n^{4\eta}} \Big(\sum_{m\le n/d^2} 1\Big)^{3/4} \Big(\sum_{m\le n/d^2} \Big|\lambda\Big(\frac{n-d^2m}{\Delta}\Big)\Big|^4\Big)^{1/4} \ll n^\varepsilon \sum_{d\ge n^{4\eta}} \frac{n}{d^{3/2}} \ll n^{1+\varepsilon-2\eta},$$

using the bound (4.3) for the fourth moment of Hecke eigenvalues and recalling that presently the conductor $C_\phi \le n^\eta$.

For $d \le n^{4\eta}$, we detect the condition $(m, \mathfrak{N}) = 1$ using Möbius inversion to write

$$\sum_{\substack{m\le n/d^2\\ d^2m\equiv n \,(\mathrm{mod}\,\Delta)\\ d^2m\equiv w \,(\mathrm{mod}\, q)\\ (m,\mathfrak{N})=1\\ P^+(m)\le y}} \lambda\Big(\frac{n-d^2m}{\Delta}\Big) = \sum_{\substack{\tau\mid\mathfrak{N}\\ P^+(\tau)\le y}} \mu(\tau) \sum_{\substack{m\le n/d^2\tau\\ d^2\tau m\equiv n \,(\mathrm{mod}\,\Delta)\\ d^2\tau m\equiv w \,(\mathrm{mod}\, q)\\ P^+(m)\le y}} \lambda\Big(\frac{n-d^2\tau m}{\Delta}\Big).$$

The contribution from $\tau \ge n^{4\eta}$ can be bounded by

$$n^{1+\varepsilon} \sum_{\substack{\tau\mid n\\ \tau\ge n^{4\eta}}} \frac{1}{\tau} \ll \frac{n^{1+\varepsilon}\tau(n)}{n^{4\eta}} \ll n^{1+2\varepsilon-4\eta}.$$

Both error terms are acceptable. Thus, for $d, \tau \le n^{4\eta}$ and automatically $(d\tau, q) = 1$ and $d_0 := (d^2\tau, \Delta) \mid n$, we are left with bounding

$$\sum_{\substack{m \le n/d^2\tau \\ m \equiv \frac{n}{d_0}\overline{\frac{d^2\tau}{d_0}} \,(\mathrm{mod}\, \frac{\Delta}{d_0}) \\ m \equiv w\overline{d^2\tau} \,(\mathrm{mod}\, q) \\ P^+(m) \le y}} \lambda\Big(\frac{n - d^2\tau m}{\Delta}\Big),$$

and the desired power saving as in (7.10) follows (after obvious smoothing) from Theorem 3.4.

7.4 Computing the main term I: summing over arithmetic progressions. We now return to (7.9) and insert the contribution corresponding to (2.7). In view of the bounds in the previous subsection, we have for $\mathcal{D} = n^{\kappa_0}$ for sufficiently small κ_0 (depending on the constants ϖ in Proposition 7.1 and η is as in (7.10))

$$S(\mathcal{A}, z) \ge \sum_{u \mid \mathcal{Q}} \sum_{c \mid P_{\mathcal{Q}}(z)} \lambda_c^- \mu(c) \sum_{\substack{d_1, d_2 \\ p \mid d_1 d_2 \Leftrightarrow p \mid uc}} \mu(d_1)\mu(d_2) S_{d_1,d_2} + O(n^{1-\kappa_0}) \tag{7.11}$$

where

$$S_{d_1,d_2} = \sum_{\substack{\delta = (d_1,d_2) \\ d_i' = d_i/\delta \\ \delta_1\delta_1' = d_1' \\ \delta_2\delta_2' = d_2'}} \sum \frac{4}{\delta_1'\delta_2'} \prod_{p \mid \delta_1'\delta_2'} (1 - \frac{\chi_{-4}(p)}{p})^{-1} G_{\delta_1'\delta_2'} \sum_{w \in \mathcal{G}_{\delta_1'\delta_2'}} \sum_{\substack{m \equiv n \,(\mathrm{mod}\, \delta^2\delta_1^2\delta_2^2) \\ n - m \equiv w \,(\mathrm{mod}\, \delta_1'\delta_2') \\ P^+(m) \le y \\ \mu^2(m_o) = (m, \mathfrak{N}) = 1}} r\Big(\frac{n-m}{\delta^2\delta_1^2\delta_2^2}\Big).$$

As in the proof of Proposition 6.2, we apply Dirichlet's hyperbola method to the r-function. Let η_0 be some small positive constant that will be chosen appropriately later. Let

$$x = n(1 - 1/Z), \quad Z = n^{\eta_0}$$

for some $\eta_0 > 0$ to be chosen later. In particular, in terms of orders of magnitude, n and x can (and will) be used interchangeably. Note that the contribution from $m \ge x$ or $m \le n - x$ can be bounded by $Z^{-1}n^{1+\varepsilon} \ll n^{1-\eta_0/2}$. Thus we can write

$$S_{d_1,d_2} =: S_1 + S_2 + O(n^{1-\eta_0/2}),$$

where

$$S_1 = \sum_{\substack{\delta = (d_1,d_2) \\ d_i' = d_i/\delta \\ \delta_1\delta_1' = d_1' \\ \delta_2\delta_2' = d_2'}} \sum \frac{4}{\delta_1'\delta_2'} \prod_{p \mid \delta_1'\delta_2'} \Big(1 - \frac{\chi_{-4}(p)}{p}\Big)^{-1} G_{\delta_1'\delta_2'}$$

$$\times \sum_{w\in\mathcal{G}_{\delta_1'\delta_2'}} \sum_{\substack{a\le\sqrt{n}\\(a,\delta_1'\delta_2')=1}} \sum_{\substack{n-x\le m\le x\\P^+(m)\le y\\ \mu^2(m_o)=(m,\mathfrak{N})=1\\ m\equiv n \,(\mathrm{mod}\, a\delta^2\delta_1^2\delta_2^2)\\ n-m\equiv w\,(\mathrm{mod}\,\delta_1'\delta_2')}} \chi_{-4}\Big(\frac{n-m}{a\delta^2\delta_1^2\delta_2^2}\Big),$$

$$S_2 = \sum_{\substack{\delta=(d_1,d_2)\\ d_i'=d_i/\delta\\ \delta_1\delta_1'=d_1'\\ \delta_2\delta_2'=d_2'}}\sum \frac{4}{\delta_1'\delta_2'} \prod_{p\mid\delta_1'\delta_2'} \Big(1-\frac{\chi_{-4}(p)}{p}\Big)^{-1} G_{\delta_1'\delta_2'}$$

$$\times \sum_{w\in\mathcal{G}_{\delta_1'\delta_2'}} \sum_{\substack{b\le \frac{x}{(\delta\delta_1\delta_2)^2\sqrt{n}}\\(b,\delta_1'\delta_2')=1}} \sum_{\substack{n-x\le m\le x\\ m\le n-b(\delta\delta_1\delta_2)^2\sqrt{n}\\P^+(m)\le y\\ \mu^2(m_o)=(m,\mathfrak{N})=1\\ m\equiv n \,(\mathrm{mod}\, b\delta^2\delta_1^2\delta_2^2)\\ n-m\equiv w\,(\mathrm{mod}\,\delta_1'\delta_2')}} \chi_{-4}(b).$$

Again we expect that S_1 gives a negligible contribution, as in the proof of Proposition 6.2, due to a particular behaviour at the prime 2 which makes the main term disappear, but this is not so easy to see in the present situation. We will carry out the computation for both S_1 and S_2 and combine them to a joint main term for which we obtain the desired lower bound. As before, the computation of S_1 is slightly harder, so we focus on this part in detail, while the computation of S_2 is similar, but technically slightly easier. Our first step is to establish the asymptotic evaluations (7.17) and (7.21) below.

We would like to choose η_0 sufficiently small so that we can evaluate the sums by Proposition 7.1 on average of d_1, d_2, but this requires some preparation. We consider S_1 first. As before we write

$$r_2 = (\delta\delta_1\delta_2, 2^\infty), \quad d = \delta\delta_1\delta_2/r_2, \quad a_2 = (a, 2^\infty),$$
$$a_d = (a, d^\infty), \quad d' = \delta_1'\delta_2'/(\delta_1'\delta_2', 2),$$

so that a, m-sum in S_1 becomes

$$\sum_{\substack{a_2\mid 2^\infty\\(a_2,\delta_1'\delta_2')=1}} \sum_{a_d\mid d^\infty} \chi_{-4}(a_d) \sum_{u\,(\mathrm{mod}\,4)}^{*} \chi_{-4}(u) \sum_{v\,(\mathrm{mod}\,4)} \chi_{-4}(v)$$
$$\times \sum_{\substack{a\le\sqrt{n}/a_2a_d\\(a,2d_1d_2)=1\\ a\equiv u\,(\mathrm{mod}\,4)}} \sum_{\substack{n-x\le m\le x, P^+(m)\le y\\ \mu^2(m_o)=(m,\mathfrak{N})=1\\ m\equiv n\,(\mathrm{mod}\ aa_2a_dr_2^2d^2)\\ n-m\equiv w\,(\mathrm{mod}\,d')\\ \frac{n-m}{a_2r_2^2}\equiv v\,(\mathrm{mod}\,4)}} 1. \tag{7.12}$$

Here we have used the fact that if $2 \mid \delta_1'\delta_2'$, then we must have $r_2a_2 = 1$ and $(w, 2) = 1$ so that the condition $n - m \equiv w \,(\mathrm{mod}\, 2)$ is incorporated in the v-sum. As before, we

can restrict $a_2, a_d \le n^{\eta_0}$ at the cost of an error of size $O(n^{1+\varepsilon-\eta_0/2})$. The a,m-sum in (7.12) can be written as

$$\sum_{\substack{a\le\sqrt{n}/a_2a_d\\(a,2d_1d_2)=1\\a\equiv u\,(\mathrm{mod}\,4)}}\ \sum_{\substack{n-x\le m\le x, P^+(m)\le y\\ \mu^2(m_o)=(m,\mathfrak{N})=1\\ m\equiv n\,(\mathrm{mod}\,a)\\ m\equiv n\,(\mathrm{mod}\,a_dd^2)\\ m\equiv n-va_2r_2^2\,(\mathrm{mod}\,4a_2r_2^2)\\ m\equiv n-w\,(\mathrm{mod}\,d')}} 1. \tag{7.13}$$

By the Chinese remainder theorem, the congruence conditions modulo $a_dd^2, 4a_2r_2^2, d'$ can be written as a single congruence condition $\frac{m}{g_1}\equiv c_1\,(\mathrm{mod}\,F_1)$ with $(c_1,F_1)=1$, where

$$F_1=\frac{4a_2r_2^2a_dd^2d'}{g_1},\quad g_1=(n-va_2r_2^2,4a_2r_2^2)(n-w,d')(n,a_dd^2). \tag{7.14}$$

In order for the m-sum to be nonempty, we must have

$$P^+(g_1)\le y,\quad \mu((g_1)_o)^2=(g_1,\mathfrak{N})=1, \tag{7.15}$$

as well as

$$P^+((a,n))\le y,\quad \mu((a,n))^2=((a,n),\mathfrak{N})=1,$$

in which case we can re-write the a,m-sum in (7.13) as

$$\sum_{\substack{a\le\sqrt{n}/a_2a_d\\(a,2d_1d_2)=1\\a\equiv u\,(\mathrm{mod}\,4)}}\ \sum_{\substack{(n-x)/g_1\le m\le x/g_1\\P^+(m)\le y\\ \mu^2(m_o)=(m,(g_1)_o\mathfrak{N})=1\\ m\equiv n\bar{g}_1\,(\mathrm{mod}\,a)\\ m\equiv c_1\,(\mathrm{mod}\,F_1)}} 1=\sum_{\substack{h\mid n\\(h,2d_1d_2)=1\\ \mu^2(h)=(h,\mathfrak{N})=1\\ P^+(h)\le y}}\ \sum_{\substack{a\le\sqrt{n}/a_2a_dh\\(a,2d_1d_2)=1\\(a,n/h)=1\\ a\equiv u\bar{h}\,(\mathrm{mod}\,4)}}\ \sum_{\substack{n-x\le mg_1h\le x\\P^+(m)\le y\\ \mu^2(m_o)=(m_o,g_1h\mathfrak{N})=1\\ m\equiv\frac{n}{h}\bar{g}_1\,(\mathrm{mod}\,a)\\ m\equiv c_1\bar{h}\,(\mathrm{mod}\,F_1)}} 1.$$

Note that from the construction of the sieve weights we have $d_1,d_2\le\mathcal{QD}\ll n^{\kappa_0}$ and thus we automatically have $g_1\ll n^{2\eta_0+4\kappa_0}$, a small power of n. Also note that $g_1\mid(2d_1d_2)^\infty$, so that automatically $(g_1,h)=1$.

We can again restrict $h\le n^{\eta_0}$ with a total error of size $O(n^{1+\varepsilon-\eta_0/2})$. Finally we detect the conditions $\mu^2(m_o)=1$ by Möbius inversion. We summarize that the a,m-sum in (7.12) can, up to an admissible error of size $O(n^{1+\varepsilon-\eta_0/2})$, be replaced with

$$\sum_{\substack{h\mid n\\(h,2d_1d_2)=1\\ \mu^2(h)=(h,\mathfrak{N})=1\\ P^+(h)\le y\\ h\le n^{\eta_0}}}\ \sum_{\substack{a\le\sqrt{n}/a_2a_dh\\(a,2d_1d_2)=1\\(a,n/h)=1\\ a\equiv u\bar{h}\,(\mathrm{mod}\,4)}}\ \sum_{\substack{n-x\le mg_1h\le x\\ \mu^2(m_o)=1\\(m_o,g_1h\mathfrak{N})=1\\ P^+(m)\le y\\ m\equiv\frac{n}{h}\overline{g_1}\,(\mathrm{mod}\,a)\\ m\equiv c_1\bar{h}\,(\mathrm{mod}\,F_1)}} 1,$$

where $a_2, a_d \ll n^{\eta_0}, g_1 \ll n^{2\eta_0+4\kappa_0}$ are sufficiently small powers of n. We evaluate this as

$$\sum_{\substack{h|n\\(h,2d_1d_2)=1\\ \mu^2(h)=(h,\mathfrak{N})=1\\ P^+(h)\le y\\ h\le n^{\eta_0}}} \Big(\sum_{\substack{a\le \sqrt{n}/a_2a_dh\\(a,2d_1d_2)=1\\(a,n/h)=1\\ a\equiv u\bar{h} \,(\mathrm{mod}\, 4)}} \sum_{\substack{n-x\le mg_1h\le x\\ P^+(m)\le y\\ (m,aF_1)=1\\ \mu^2(m_o)=1\\ (m_o,g_1h\mathfrak{N})=1}} \frac{1}{\phi(a)\phi(F_1)} + \tilde{r}_1\Big), \tag{7.16}$$

where

$$\tilde{r}_1 = \sum_{\substack{a\le \sqrt{n}/a_2a_dh\\(a,2d_1d_2)=1\\(a,n/h)=1\\ a\equiv u\bar{h} \,(\mathrm{mod}\, 4)}} \sum_{\substack{n-x\le mg_1h\le x\\ \mu^2(m_o)=1\\ (m_o,g_1h\mathfrak{N})=1\\ P^+(m)\le y\\ m\equiv \frac{n}{h}\overline{g_1} \,(\mathrm{mod}\, a)\\ m\equiv c_1\bar{h} \,(\mathrm{mod}\, F_1)}} 1 - \sum_{\substack{a\le \sqrt{n}/a_2a_dh\\(a,2d_1d_2)=1\\(a,n/h)=1\\ a\equiv u\bar{h} \,(\mathrm{mod}\, 4)}} \sum_{\substack{n-x\le mg_1h\le x\\ P^+(m)\le y\\ (m,aF_1)=1\\ \mu^2(m_o)=1\\ (m_o,g_1h\mathfrak{N})=1}} \frac{1}{\phi(a)\phi(F_1)}.$$

We want to show that $\tilde{r}_1$ that is small on average using Proposition 7.1 with $\mathbf{q} = a, \mathbf{a}_1 = \frac{n}{h}, \mathbf{d} = F_1$. It is not enough to apply Proposition 7.1 with $\mathbf{a}_2 = g_1$ and sum over g_1 directly due to sparseness of smooth numbers when y is small and thus we need to be more precise about g_1. Note that $a_2r_2^2a_dd^2d' \mid g_1F_1$ and so $(h, d_1d_2) = 1$ and $(a, 2d_1d_2)$ is equivalent to $(a, g_1F_1) = 1$. The coefficients of h can be bounded by 1. The coefficients of g_1, F_1 is bounded by $\tau(dd')^C \ll \tau(g_1F_1)^C$. From (7.14), we see that if $p \mid g_1, p \nmid F_1$, then we must have either $p \mid (n-w, d')$ or $p \mid 2n$ and $p^2 \mid g_1$. If $p^2 \mid g_1$, we also have $p \mid 2n$. Therefore we can write $g_1 = g_0g_1'g_3$ where $g_1' = (n-w, d')$ is square-free, $g_0 \mid (2n)^\infty$ is square-full and $g_3 \mid F_1$. We can then change the order of summation to move h, g_0, g_1' outside the summation over u, c or equivalently d_1, d_2. After dividing the m-sum into short intervals to separate the variables m from g_1 similarly as in the proof of Proposition 6.2, we can apply Proposition 7.1 for the c, a_2, a_d, a, m-sum with $\mathbf{d} = F_1, \mathbf{a}_2 = g_0g_1', \mathbf{a}_3 = g_3$ for fixed h, g_0g_1' (with $h \le n^{\eta_0}$ and $g_0g_1' \ll n^{2\eta_0+4\kappa_0}$ for η_0, κ_0 sufficiently small) so that the total contribution of the error $\tilde{r}_1 = \tilde{r}_1(g_3)$ in (7.11) can be bounded by

$$\begin{aligned}\mathcal{E}_1 &\ll \sum_{\substack{h|n,h\nmid\mathfrak{N}\\ h\le n^{\eta_0}\\ P^+(h)\le y}} \sum_{\substack{g_1\le n^{2\eta_0+4\kappa_0}\\ P^+(g_1)\le y\\ g_1=g_0g_1'g_3\\ g_0|2^\infty n\\ p|g_0\Rightarrow p^2|g_0}} \sum_{\substack{F_1\ll n^{4\kappa_0}\\ g_3|F_1}} \tau(g_1F_1)^C \sum_{\substack{d'|F_1g_1'\\ \mu^2(d')=1}} \frac{\tau(d')}{d'} \sum_{\substack{w\in\mathcal{G}_{d'}\\ (n-w,d')=g_1'}} |\tilde{r}_1(g_3)|\\ &\ll \sum_{\substack{h|n,h\nmid\mathfrak{N}\\ h\le n^{\eta_0}\\ P^+(h)\le y}} \sum_{\substack{g_0g_1'\le n^{2\eta_0+4\kappa_0}\\ P^+(g_0g_1')\le y\\ g_0|2^\infty n\\ p|g_0\Rightarrow p^2|g_0}} \tau(g_0g_1')^C \sum_{F_1\ll n^{4\kappa_0}} \tau(F_1)^{2C}\end{aligned}$$

$$\times \max_{g_3|F_1} \sum_{\substack{d'|F_1 g_1' \\ \mu^2(d')=1}} \frac{\tau(d')}{d'} \sum_{\substack{w\in\mathcal{G}_{d'} \\ (n-w,d')=g_1'}} |\tilde{r}_1(g_3)|.$$

Note that

$$\sum_{\substack{d'|F_1 g_1' \\ \mu^2(d')=1}} \frac{\tau(d')}{d'} \sum_{\substack{w\in\mathcal{G}_{d'} \\ (n-w,d')=g_1'}} 1 \ll \frac{\tau(F_1 g_1')^2}{g_1'}$$

and thus for $y \geq \mathcal{L}(x) = (\log x)^D$ for some sufficiently large D, Proposition 7.1 gives

$$\mathcal{E}_1 \ll_C \sum_{\substack{h|n,h\nmid\mathfrak{N} \\ h\leq n^{\eta_0} \\ P^+(h)\leq y}} \sum_{\substack{g_0 g_1'\leq n^{2\eta_0+4\kappa_0} \\ P^+(g_0 g_1')\leq y \\ g_0|2^\infty n \\ p|g_0\Rightarrow p^2|g_0}} \frac{\tau(g_0 g_1')^C}{g_1'} \Psi_{hg_0g_1'\mathfrak{N}}\Big(\frac{x}{g_0 g_1' h}, y\Big)(\log x)^{-A}.$$

We recall (7.1), (7.2) and note that for $y \geq \mathcal{L}(x)$ we have $\alpha \geq 1 - \frac{1}{D} + o(1)$. Thus when $y \geq \mathcal{L}(x)$ we have

$$\begin{aligned}
\mathcal{E}_1 &\ll_C \sum_{\substack{h|n,h\nmid\mathfrak{N} \\ h\leq n^{\eta_0} \\ P^+(h)\leq y}} \sum_{\substack{g_0 g_1'\leq n^{2\eta_0+4\kappa_0} \\ P^+(g_0 g_1')\leq y \\ g_0|2^\infty n \\ p|g_0\Rightarrow p^2|g_0}} \frac{\tau(g_0 g_1')^C}{g_1'} \Psi_{hg_0g_1'\mathfrak{N}}\Big(\frac{x}{g_0 g_1' h}, y\Big)(\log x)^{-A} \\
&\ll \sum_{\substack{h|n,h\nmid\mathfrak{N} \\ P^+(h)\leq y}} \frac{1}{h^\alpha} \prod_{\substack{p|\mathfrak{N} \\ p\leq y}} (1-p^{-\alpha}) \sum_{\substack{g_0|2^\infty n \\ p|g_0\Rightarrow p^2|g_0}} \frac{\tau(g_0)^C}{g_0^\alpha} \sum_{g_1'\leq n^{2\eta_0+4\kappa_0}} \frac{\tau(g_1')^C}{(g_1')^{1+\alpha}} \Psi(x,y)(\log x)^{-A} \\
&\ll \prod_{\substack{p|n \\ p\leq y}} \Big(1 + \frac{\chi_{-4}(p)}{p^\alpha}\Big) \Psi(x,y)(\log x)^{-A}
\end{aligned}$$

by choosing A large enough in terms of C. Since $\prod_{p|n}(1+\frac{1}{p}) \ll \log_2 n$, we have

$$\mathcal{E}_1 \ll \mathcal{F}(n,y)\Psi(x,y)(\log x)^{-A}$$

for any $A > 0$ with $\mathcal{F}(n,y)$ defined as (1.5).

For the main term, we can substitute the main term in (7.16) back into (7.12) and evaluating the u-sum, we obtain (writing the condition (7.15) on g_1 in terms of the remaining variables and noting that $\mu^2(d') = 1, a_d \mid d^\infty$) that we can replace S_1 by $\tilde{S}_1$ in (7.11) up to a negligible error $\mathcal{E}_1$ where

$$\tilde{S}_1 = \sum_{\delta_1\delta_1'=d_1'} \sum_{\delta_2\delta_2'=d_2'} \frac{4}{\delta_1'\delta_2'} \prod_{p|\delta_1'\delta_2'} \Big(1 - \frac{\chi_{-4}(p)}{p}\Big)^{-1} G_{\delta_1'\delta_2'} \mathbb{1}_{(d,\mathfrak{N})=1} \mathcal{M}_1, \tag{7.17}$$

with

$$\mathcal{M}_1 = \sum_{\substack{a_2|2^\infty\\(a_2,\delta_1'\delta_2')=1\\a_2\le n^{\eta_0}}} \sum_{\substack{a_d|d^\infty, a_d\le n^{\eta_0}\\P^+((n,a_dd^2))\le y\\\mu^2((n,a_dd^2))=1}} \chi_{-4}(a_d) \sum_{v \,(\mathrm{mod}\, 4)} \chi_{-4}(v)$$
$$\times \sum_{\substack{w \,(\mathrm{mod}\, \delta_1'\delta_2')\\(w,\delta_1'\delta_2')=1\\w\in\mathcal{G}_{\delta_1'\delta_2'}\\P^+((n-w,d'))\le y\\(n-w,d',\mathfrak{N})=1}} \sum_{\substack{h|n\\(h,2d_1d_2\mathfrak{N})=1\\P^+(h)\le y\\\mu^2(h)=1\\h\le n^{\eta_0}}} \chi_{-4}(h) \sum_{\substack{a\le\sqrt{n}/a_2a_dh\\(a,2d_1d_2)=1\\(a,n/h)=1}} \sum_{\substack{n-x\le mg_1h\le x\\P^+(m)\le y\\(m,aF_1(g_1)_oh\mathfrak{N})=1\\\mu^2(m_o)=1}} \frac{\chi_{-4}(a)}{\phi(a)\phi(F_1)}.$$

Next we change the summation order of m and a, evaluate the a-sum by [ABL, Lemma 5.2] and then complete the a_2, a_d, h-sums with an admissible error to finally obtain (with e.g. $\eta_0 = 10\kappa_0$), uniformly for $d_1, d_2 \ll n^{\kappa_0}$ (with κ_0 sufficiently small), that

$$\mathcal{M}_1 = \frac{\pi}{4}\prod_p\Big(1+\frac{\chi_{-4}(p)}{p(p-1)}\Big) \sum_{\substack{a_2|2^\infty\\(a_2,\delta_1'\delta_2')=1}} \sum_{\substack{a_d|d^\infty\\P^+((n,a_dd^2))\le y\\\mu^2((n,a_dd^2))=1}} \chi_{-4}(a_d) \sum_{v \,(\mathrm{mod}\, 4)} \chi_{-4}(v)$$
$$\times \sum_{\substack{w \,(\mathrm{mod}\, \delta_1'\delta_2')\\w\in\mathcal{G}_{\delta_1'\delta_2'}\\P^+((n-w,d'))\le y\\(n-w,d',\mathfrak{N})=1}} \sum_{\substack{h|n\\(h,2d_1d_2\mathfrak{N})=1\\P^+(h)\le y\\\mu^2(h)=1}} \chi_{-4}(h) \sum_{\substack{m\le x/g_1h\\P^+(m)\le y\\(m,F_1(g_1)_oh\mathfrak{N})=1\\\mu^2(m_o)=1}} \frac{\mathfrak{c}(2d_1d_2mn/h)}{\phi(F_1)} + O(n^{1-3\kappa_0})$$

with $\mathfrak{c}(f)$ as in (6.4).

The same strategy can be applied to S_2. We give a sketch of the manipulations highlighting the differences from those for S_1. To begin with, we do not need to separate r_2 from $\delta\delta_1\delta_2$, so that the b, m-sum in S_2 becomes

$$\sum_{b_d|(\delta\delta_1\delta_2)^\infty} \chi_{-4}(b_d) \sideset{}{^*}\sum_{u \,(\mathrm{mod}\, 4)} \chi_{-4}(u) \sum_{\substack{b\le \frac{x}{b_d(\delta\delta_1\delta_2)^2\sqrt{n}}\\(b,d_1d_2)=1\\b\equiv u \,(\mathrm{mod}\, 4)}} \sum_{\substack{n-x\le m\le x, P^+(m)\le y\\m\le n-bb_d(\delta\delta_1\delta_2)^2\sqrt{n}\\\mu^2(m_o)=(m,\mathfrak{N})=1\\m\equiv n \,(\mathrm{mod}\, bb_d(\delta\delta_1\delta_2)^2)\\n-m\equiv w \,(\mathrm{mod}\, \delta_1'\delta_2')}} 1. \tag{7.18}$$

We can restrict $b_d \le n^{\eta_0}$ with an error of size $O(n^{1+\varepsilon-\eta_0/2})$. After splitting m into residue classes modulo d_1d_2, using the Chinese remainder theorem to rewrite the congruence condition modulo $b_d(\delta\delta_1\delta_2)^2\delta_1'\delta_2'$ as $m \equiv c_2 \,(\mathrm{mod}\, F_2)$ with

$$F_2 = \frac{b_d(\delta\delta_1\delta_2)^2\delta_1'\delta_2'}{g_2},\ g_2 = (n-w, \delta_1'\delta_2')(n, b_d\delta^2\delta_1^2\delta_2^2) \tag{7.19}$$

and then extracting (b,n), we can write the b,m-sum in (7.18) as

$$\mathbb{1}\sum_{\substack{P^+(g_2)\le y\\(g_2,\mathfrak{N})=1\\ \mu^2((g_2)_o)=1}}\sum_{\substack{h\mid n\\(h,d_1d_2\mathfrak{N})=1\\P^+(h)\le y\\ \mu^2(h)=1}}\chi_{-4}(h)\sum_{b\le \frac{x}{b_d(\delta\delta_1\delta_2)^2\sqrt{nh}}\\(b,d_1d_2)=1\\(b,n/h)=1\\ b\equiv u\,(\mathrm{mod}\,4)}\ \sum_{\substack{n-x\le mg_2h\le x, P^+(m)\le y\\ mg_2h\le n-bb_d(\delta\delta_1\delta_2)^2\sqrt{n}\\ \mu^2(m_o)=(m_o,g_2h\mathfrak{N})=1\\ m\equiv \frac{n}{h}\overline{g}_2\,(\mathrm{mod}\,b)\\ m\equiv c_2\overline{h}\,(\mathrm{mod}\,F_2)}}1.$$

We can again restrict $h\le n^{\eta_0}$ with a total error of size $O(n^{1+\varepsilon-\eta_0/2})$. Additionally, we need to separate b,m in the summation condition as well as their dependency on g_2,F_2 in order to apply Proposition 7.1. As usual, this can be achieved by splitting m into intervals of the form $[M,(1+\Delta)M]$ with $\Delta=R^{-\eta_0}$ so that at most one interval interferes with the condition $mg_2h\le n-bb_d\sqrt{n}(\delta\delta_1\delta_2)^2$, in which case the contribution can be bounded by $O(n^{1+\varepsilon}\Delta)$ trivially. Here $R=\exp(\sqrt{\log x})$. Similarly we can split the b-sum into intervals of the form $[M,(1+\Delta)M]$ so that the condition $b\le \frac{x}{b_d(\delta\delta_1\delta_2)^2\sqrt{nh}}$ only interferes with one interval in which case the contribution is also negligible. After Möbius inversion to detect $\mu^2(m_o)=1$, we can finally apply Proposition 7.1 so that the b,m-sum in the last display can be evaluated as

$$\sum_{b\le \frac{x}{b_d(\delta\delta_1\delta_2)^2\sqrt{nh}}\\(b,d_1d_2)=1\\(b,n/h)=1\\ b\equiv u\,(\mathrm{mod}\,4)}\ \sum_{\substack{n-x\le mg_2h\le x, P^+(m)\le y\\ mg_2h\le n-bb_d(\delta\delta_1\delta_2)^2\sqrt{n}\\ \mu^2(m_o)=(m,\mathfrak{N})=1\\ (m,bF_2(g_2)_oh\mathfrak{N})=1}}\frac{1}{\phi(b)\phi(F_2)}+\tilde{r}_2'. \tag{7.20}$$

Here the total contribution from the error terms $\tilde{r}_2'$ in (7.11) can be bounded using Proposition 7.1 the same way as in $\mathcal{E}_1$ so that

$$\begin{aligned}\mathcal{E}_2 &:= \sum_{u\mid\mathcal{Q}}\sum_{c\mid P_{\mathcal{Q}}(z)}\lambda_c^-|\mu(c)|\sum_{\substack{d_1,d_2\\ p\mid d_1d_2\Leftrightarrow p\mid uc}}|\mu(d_1)\mu(d_2)|\sum_{\substack{\delta=(d_1,d_2)\\ d_i'=d_i/\delta\\ \delta_1\delta_1'=d_1'\\ \delta_2\delta_2'=d_2'}}\sum\frac{4}{\delta_1'\delta_2'}\\ &\quad\times\prod_{p\mid\delta_1'\delta_2'}\Big(1-\frac{\chi_{-4}(p)}{p}\Big)^{-1}G_{\delta_1'\delta_2'}\sum_{w\in\mathcal{G}_{\delta_1'\delta_2'}}|\tilde{r}_2'|\\ &\ll_C \prod_{\substack{p\mid n\\ p\le y}}\Big(1+\frac{\chi_{-4}(p)}{p^\alpha}\Big)\Psi(x,y)(\log x)^{-A}\ll_C \mathcal{F}(n,y)\Psi(x,y)(\log x)^{-A+1}\end{aligned}$$

for any $A>0$. After changing the order of summation, evaluating the b-sum in (7.20) by [ABL, Lemma 5.2] (note that the upper limit for b_d is always $\gg n^{1/2-\eta_0}$ unless $|n-mg_2h|\le n^{1-\eta_0}$ whose contribution can be bounded by $O(n^{1-\eta_0/2})$) and completing the b_d,h-sums, we obtain, uniformly for $d_1,d_2\le n^{\kappa_0}$, that we can replace S_2 by $\tilde{S}_2$ in (7.11) up to a negligible error $\mathcal{E}_2$ where

$$\tilde{S}_2=\sum_{\delta_1\delta_1'=d_1'}\sum_{\delta_2\delta_2'=d_2'}\frac{4}{\delta_1'\delta_2'}\prod_{p\mid\delta_1'\delta_2'}(1-\frac{\chi_{-4}(p)}{p})^{-1}G_{\delta_1'\delta_2'}\mathbb{1}_{(\delta\delta_1\delta_2,\mathfrak{N})=1}\mathcal{M}_2 \tag{7.21}$$

and

$$\mathcal{M}_2 = \frac{\pi}{4}\prod_p\Big(1+\frac{\chi_{-4}(p)}{p(p-1)}\Big)\sum_{\substack{b_d\mid(\delta\delta_1\delta_2)^\infty\\ P^+((n,b_d(\delta\delta_1\delta_2)^2))\le y\\ \mu^2((n,b_d(\delta\delta_1\delta_2)^2)_o)=1}}\chi_{-4}(b_d)$$
$$\times\sum_{\substack{w\,(\mathrm{mod}\,\delta_1'\delta_2')\\ w\in\mathcal{G}_{\delta_1'\delta_2'}\\ P^+((n-w,\delta_1'\delta_2'))\le y\\ (n-w,\delta_1'\delta_2',\mathfrak{N})=1}}\sum_{\substack{h\mid n\\ (h,d_1d_2)=1\\ P^+(h)\le y\\ \mu^2(h)=1\\ (h,\mathfrak{N})=1}}\chi_{-4}(h)\sum_{\substack{m\le x/g_2h\\ P^+(m)\le y\\ (m,F_2(g_2)_oh\mathfrak{N})=1\\ \mu^2(m_o)=1}}\frac{\mathfrak{c}(d_1d_2mn/h)}{\phi(F_2)}+O(n^{1-3\kappa_0}).$$

7.5 Computing the main term II: Euler products. We substitute (7.17) and (7.21) back into (7.11). By choosing $\mathcal{D}=n^{\kappa_0}$ sufficiently small, it follows from (7.9) together with (7.10) that

$$S(\mathcal{A},z)\ge\pi\prod_p\Big(1+\frac{\chi_{-4}(p)}{p(p-1)}\Big)(S_1^-+S_2^-)+O(\mathcal{E}_1+\mathcal{E}_2)+O(n^{1-\kappa_0}),\quad(7.22)$$

where

$$S_1^-=\sum_{u\mid\mathcal{Q}}\sum_{c\mid P_{\mathcal{Q}}(z)}\lambda_c^-\mu(c)\sum_{\substack{d_1,d_2\\ p\mid d_1d_2\Leftrightarrow p\mid uc}}\mu(d_1)\mu(d_2)\sum\sum_{\substack{\delta=(d_1,d_2)\\ \delta_i\delta_i'=d_i'=d_i/\delta\\ r_2d=\delta\delta_1\delta_2,r_2\mid 2,(d,2\mathfrak{N})=1\\ d'=\delta_1'\delta_2'/(\delta_1'\delta_2',2)}}\frac{1}{\delta_1'\delta_2'}$$
$$\times\prod_{p\mid\delta_1'\delta_2'}\Big(1-\frac{\chi_{-4}(p)}{p}\Big)^{-1}\sum_{\substack{a_2\mid 2^\infty\\ (a_2,\delta_1'\delta_2')=1}}\sum_{\substack{a_d\mid d^\infty\\ P^+((n,a_dd^2))\le y\\ \mu^2((n,a_dd^2))=1}}\chi_{-4}(a_d)\sum_{v\,(\mathrm{mod}\,4)}\chi_{-4}(v)$$
$$\times G_{d'}\sum_{\substack{w\,(\mathrm{mod}\,d')\\ w\in\mathcal{G}_{d'}\\ P^+((n-w,d'))\le y\\ (n-w,d',\mathfrak{N})=1}}\sum_{\substack{h\mid n\\ (h,2d_1d_2\mathfrak{N})=1\\ P^+(h)\le y\\ \mu^2(h)=1}}\chi_{-4}(h)\sum_{\substack{m\le x/g_1h\\ P^+(m)\le y\\ (m,F_1(g_1)_oh\mathfrak{N})=1\\ \mu^2(m_o)=1}}\frac{\mathfrak{c}(2d_1d_2mn/h)}{\phi(F_1)}$$

with F_1,g_1 as in (7.14), and

$$S_2^-=\sum_{u\mid\mathcal{Q}}\sum_{c\mid P_{\mathcal{Q}}(z)}\lambda_c^-\mu(c)\sum_{\substack{d_1,d_2\\ p\mid d_1d_2\Leftrightarrow p\mid uc}}\mu(d_1)\mu(d_2)\sum\sum_{\substack{\delta=(d_1,d_2)\\ \delta_i\delta_i'=d_i'=d_i/\delta\\ d=\delta\delta_1\delta_2,(d,\mathfrak{N})=1\\ d'=\delta_1'\delta_2'}}\frac{1}{\delta_1'\delta_2'}$$
$$\times\prod_{p\mid\delta_1'\delta_2'}\Big(1-\frac{\chi_{-4}(p)}{p}\Big)^{-1}G_{d'}\sum_{\substack{b_d\mid d^\infty P^+((n,b_dd^2))\le y\\ \mu^2((n,b_dd^2)_o)=1}}\chi_{-4}(b_d)$$

$$\times \sum_{\substack{w \,(\mathrm{mod}\, d') \\ w \in \mathcal{G}_{d'} \\ P^+((n-w,d'))\le y \\ (n-w,d',\mathfrak{N})=1}} \sum_{\substack{h \mid n \\ (h,d_1d_2\mathfrak{N})=1 \\ P^+(h)\le y \\ \mu^2(h)=1}} \chi_{-4}(h) \sum_{\substack{m \le x/g_2h \\ P^+(m)\le y \\ (m,F_2(g_2)_o h\mathfrak{N})=1 \\ \mu^2(m_o)=1}} \frac{\mathfrak{c}(d_1d_2mn/h)}{\phi(F_2)}$$

with F_2, g_2 as in (7.19). Now we rename mg_i as m, change the order of summation and evaluate S_1^-, S_2^- in roughly the form

$$\sum_{\substack{h \mid n \\ (\cdots)}} \sum_{\substack{m \le x/h \\ (\cdots)}} \sum_{\substack{u \mid \mathcal{Q} \\ (u,h)=1}} \sum_{\substack{c \mid P(z) \\ (c,\mathcal{Q}h)=1}} \lambda_c^- w_{m,n,h}(c)$$

for some multiplicative function $w_{m,n,h}(c)$. In particular, we postpone the summation over m to the last moment. Precisely, we write

$$\begin{aligned} S_1^- = \sum_{\substack{h \mid n \\ P^+(h)\le y \\ \mu^2(h)=1 \\ (h,\mathfrak{N})=1}} \chi_{-4}(h) \sum_{\substack{m \le x/h \\ (m,\mathfrak{N})=1 \\ P^+(m)\le y \\ \mu^2(m_o h)=1}} \sum_{\substack{u \mid \mathcal{Q} \\ (u,h)=1}} \sum_{\substack{c \mid P(z) \\ (c,\mathcal{Q}h)=1}} \lambda_c^- \mu(c) \sum_{\substack{d_1,d_2 \\ p \mid d_1d_2 \Leftrightarrow p \mid uc}} \\ \mu(d_1)\mu(d_2)\mathfrak{c}\Big(2d_1d_2\frac{n}{h}m\Big) \sum\sum_{\substack{\delta=(d_1,d_2),\delta_i\delta_i'=d_i'=d_i/\delta \\ r_2 d=\delta\delta_1\delta_2, r_2 \mid 2,(d,2)=1 \\ d'=\delta_1'\delta_2'/(\delta_1'\delta_2',2)}} \frac{G_{d'}}{\delta_1'\delta_2'} \prod_{p \mid \delta_1'\delta_2'} \Big(1-\frac{\chi_{-4}(p)}{p}\Big)^{-1} \\ \times \sum\sum_{\substack{a_2,a_d,v,w \\ (a_2,a_d,v,w)\in\mathcal{R}_1}}\sum\sum \frac{\chi_{-4}(a_d)\chi_{-4}(v)}{\phi(F_1)} \end{aligned} \tag{7.23}$$

where $\mathfrak{c}(f)$ is as in (6.4), F_1 is as in (7.14), and

$$\mathcal{R}_1 = \left\{ \begin{array}{c} a_2 \mid 2^\infty, a_d \mid d^\infty \\ v \,(\mathrm{mod}\, 4), w \,(\mathrm{mod}\, d') \end{array} : \begin{array}{cc} (a_2, \delta_1'\delta_2') = 1 & w \in \mathcal{G}_{d'} \\ (m, 4a_2r_2^2) = (n - va_2r_2^2, 4a_2r_2^2) & (w,d') = 1 \\ (m, a_dd^2) = (n, a_dd^2) & (m,d') = (n-w,d') \end{array} \right\}.$$

To see this, we recall the conditions (7.15) and observe that after replacing mg_1 with m, the new conditions $P^+(m) \le y$, $\mu^2(m_o) = 1$, $(m, \mathfrak{N}) = 1$ take care of the old conditions $P^+((n, a_dd^2)) \le y$, $P^+((n-w,d')) \le y$, $\mu^2((n,a_dd^2)) = (n, a_dd^2, \mathfrak{N}) = (n-w,d',\mathfrak{N}) = 1$. Moreover, the old condition $(m, F_1) = 1, mg_1 \le x/h$ is equivalent to the new condition $(m, 4a_2r_2^2a_dd^2d') = g_1, m \le x/h$ after replacing mg_1 with m.

Note also that $\mathcal{G}_{\delta_1'\delta_2'} \simeq \mathcal{G}_{d'}$. Similarly, we write

$$\begin{aligned} S_2^- = & \sum_{\substack{h|n \\ P^+(h)\le y \\ \mu^2(h)=1 \\ (h,\mathfrak{N})=1}} \chi_{-4}(h) \sum_{\substack{m\le x/h \\ (m,\mathfrak{N})=1 \\ P^+(m)\le y \\ \mu^2(m_o h)=1}} \sum_{\substack{u|\mathcal{Q} \\ (u,h)=1}} \sum_{\substack{c|P(z) \\ (c,\mathcal{Q}h)=1}} \lambda_c^- \mu(c) \sum_{\substack{d_1,d_2 \\ p|d_1d_2 \Leftrightarrow p|uc}} \mu(d_1)\mu(d_2)\mathfrak{c}\Big(d_1d_2\frac{n}{h}m\Big) \\ & \times \sum_{\substack{\delta=(d_1,d_2) \\ \delta_i\delta_i'=d_i'=d_i/\delta \\ d=\delta\delta_1\delta_2 \\ d'=\delta_1'\delta_2'}} \sum\sum \frac{G_{d'}}{\delta_1'\delta_2'} \prod_{p|\delta_1'\delta_2'} \Big(1-\frac{\chi_{-4}(p)}{p}\Big)^{-1} \sum_{\substack{b_d,w \\ (b_d,w)\in\mathcal{R}_2}} \sum\sum \frac{\chi_{-4}(b_d)}{\phi(F_2)} \end{aligned} \tag{7.24}$$

where F_2 is as in (7.19) and

$$\mathcal{R}_2 = \left\{ \begin{matrix} b_d \mid d^\infty \\ w \pmod{d'} \end{matrix} : \begin{matrix} (m, b_d d^2) = (n, a_d d^2) & (w,d')=1 \\ w \in \mathcal{G}_{d'} & (m,d')=(n-w,d') \end{matrix} \right\}.$$

Our next aim is to write the second line in (7.23) as an Euler product which will be given in (7.29) below. The v-sum modulo 4 in (7.23) contributes

$$\begin{aligned} & \sum_{\substack{v \pmod 4 \\ (n-va_2r_2^2, 4a_2r_2^2)=(m,4a_2r_2^2)}} \frac{\chi_{-4}(v)}{\phi(\frac{4a_2r_2^2}{(n-va_2r_2^2,4a_2r_2^2)})} \\ & = \mathbb{1}_{2a_2r_2^2|m} \mathbb{1}_{a_2r_2^2\|n} \chi_{-4}\Big(\frac{n}{a_2r_2^2}\Big) \times \begin{cases} -1 & 4a_2r_2^2 \nmid m, \\ 1 & 4a_2r_2^2 \mid m. \end{cases} \end{aligned} \tag{7.25}$$

In particular, we see that for each r_2, only one a_2 with $(a_2, \delta_1'\delta_2') = 1$ gives a non-zero contribution. The w-sum together with $G_{d'}$ can be evaluated as (treat each prime $p \mid d'$ at a time)

$$\begin{aligned} & G_{d'} \sum_{\substack{w \pmod{d'} \\ w\in\mathcal{G}_{d'} \\ (n-w,d')=(m,d')}} \frac{1}{\phi\big(d'/(n-w,d')\big)} \\ & = \mathbb{1}_{(d',mn)=1} \prod_{\substack{p|d' \\ p|mn}} (1+\chi_{p^*}(n)) \prod_{\substack{p|d' \\ p\nmid mn}} \Big(1-\frac{1+\chi_{p^*}(n)}{p-1}\Big). \end{aligned} \tag{7.26}$$

Finally, the sum over $a_d \mid d^\infty$ gives

$$\sum_{\substack{a_d|d^\infty \\ (n,a_dd^2)=(m,a_dd^2)}} \frac{\chi_{-4}(a_d)}{\phi(a_dd^2/(n,a_dd^2))} = \prod_{p|d} \mathfrak{d}(p;m,n) \tag{7.27}$$

where

$$\mathfrak{d}(p;m,n)=\begin{cases}\frac{1}{(p-1)(p-\chi_{-4}(p))}, & p\nmid mn,\\ \frac{p}{(p-1)(p-\chi_{-4}(p))}, & \nu_p(m)=\nu_p(n)=1,\\ 0, & \text{else.}\end{cases}\tag{7.28}$$

Here we use crucially that (the odd part of) m is square-free. For more general m, the formula would look much more complicated, in particular the contribution could be much larger which would create considerable technical problems in applying Lemma 7.3.

We summarize: the second line in (7.23) can be written as an Euler product

$$\sum_{\substack{\delta_1'\delta_1=d_1'\\ \delta_2'\delta_2=d_2'}}\frac{G_{\delta_1'\delta_2'}}{\delta_1'\delta_2'}\prod_{p\mid\delta_1'\delta_2'}\Big(1-\frac{\chi_{-4}(p)}{p}\Big)^{-1}\sum_{a_2,a_d,v,w}(\dots)=:\prod_{p^\nu\|d_1d_2}\mathfrak{f}_1(p,\nu;m,n)\tag{7.29}$$

where the Euler factor is composed of (7.25), (7.26) and (7.27).

We compute $\mathfrak{f}_1(2,\nu;m,n)$ explicitly using (7.25). If $2\nmid d_1d_2$, then $r_2=1$ and only $a_2=2^{\nu_2(n)}$ contributes. If $2\|d_1d_2$, then we have two cases: when $2\|\delta_1'\delta_2'$, we have $r_2=1$ and only $a_2=1$ contributes (since we require $(a_2,\delta_1'\delta_2')=1$); and when $2\|\delta_1\delta_2$, we have $r_2=2$ and only $a_2=2^{\nu_2(n)-2}$ contributes. If $2\mid(d_1,d_2)$, then $r_2=2$ and only $a_2=2^{\nu_2(n)-2}$ contributes. Precisely, we have

$$\begin{aligned}&\mathfrak{f}_1(2,\nu;m,n)\\ &=\begin{cases}\mathbb{1}_{2^{\nu_2(n)+1}\mid m}\chi_{-4}(n_o)(1-2\cdot\mathbb{1}_{2^{\nu_2(n)+2}\nmid m}), & \nu=0,\\ \frac{\mathbb{1}_{2\mid m}}{2}\chi_{-4}(n)(-1)^{\mathbb{1}_{4\nmid m}}+\mathbb{1}_{2^{\nu_2(n)+1}\mid m}\mathbb{1}_{4\mid n}\chi_{-4}(n_o)(1-2\cdot\mathbb{1}_{2^{\nu_2(n)+2}\nmid m}), & \nu=1,\\ \mathbb{1}_{2^{\nu_2(n)+1}\mid m}\mathbb{1}_{4\mid n}\chi_{-4}(n_o)(1-2\cdot\mathbb{1}_{2^{\nu_2(n)+2}\nmid m}), & \nu=2.\end{cases}\end{aligned}\tag{7.30}$$

We can also compute $\mathfrak{f}_1(p,\nu;m,n)$ for $2\nmid p$ using (7.26) and (7.27) as

$$\begin{aligned}&\mathfrak{f}_1(p,\nu;m,n)\\ &=\begin{cases}1, & \nu=0,\\ \frac{1}{(p-1)(p-\chi_{-4}(p))}+\frac{1}{p}(1-\frac{\chi_{-4}(p)}{p})^{-1}(1-\frac{1+\chi_{p^*}(n)}{p-1}), & p\nmid mn,\nu=1,\\ \frac{1}{p}(1-\frac{\chi_{-4}(p)}{p})^{-1}\big(1+\chi_{p^*}(n)\big), & p\mid mn,p\nmid(m,n),\nu=1,\\ \frac{p}{(p-1)(p-\chi_{-4}(p))}, & p\mid(m,n),\nu=1,\\ \mathfrak{d}(p;m,n), & \nu=2.\end{cases}\end{aligned}\tag{7.31}$$

For S_2^-, we have analogous results with $\mathfrak{f}_1$ replaced by $\mathfrak{f}_2$ where

$$\mathfrak{f}_2(2,\nu;m,n)=\begin{cases}1, & \nu=0,\\ \frac{1}{2}\mathbb{1}_{2\mid mn,2\nmid(m,n)}+\mathbb{1}_{(m,4)=(n,4)}\frac{1}{\phi(4/(n,4))}, & \nu=1,\\ \mathbb{1}_{(m,4)=(n,4)}\frac{1}{\phi(4/(n,4))}, & \nu=2,\end{cases}\tag{7.32}$$

(noting that if $2 \mid \delta_1'\delta_2'$ then the w-sum is zero unless $2|mn, 2 \nmid (m,n)$; and if $2 \mid \delta\delta_1\delta_2$ then the 2-part in the b_d-sum becomes $\frac{1}{\phi(4/(n,4))}\mathbb{1}_{(m,4)=(n,4)}$), and for $2 \nmid p$ we have

$$\mathfrak{f}_2(p,\nu;m,n) = \mathfrak{f}_1(p,\nu;m,n). \tag{7.33}$$

7.6 Computation of the main term III: the sieve weights. We are now prepared to further evaluate S_1^- and S_2^-. We will bound $S_1^- + S_2^-$ from below by (7.6) using properties of the lower bound sieve weights in Lemma 7.3.

To begin with, we substitute (7.29) into (7.23). We still keep h and m fixed and consider the u, c, d_1, d_2-sum in (7.23) which now equals

$$\sum_{\substack{u|\mathcal{Q}\\(u,h)=1}}\sum_{\substack{c|P(z)\\(c,\mathcal{Q}h)=1}}\lambda_c^-\mu(c)\sum_{\substack{d_1,d_2\\p|d_1d_2\Leftrightarrow p|uc}}\mu(d_1)\mu(d_2)\mathfrak{c}\Big(2d_1d_2\frac{n}{h}m\Big)\prod_{p^\nu\|d_1d_2}\mathfrak{f}_1(p,\nu;m,n).$$

We recast this as (noting $\mathfrak{c}(2) = 1$)

$$\begin{aligned}\mathfrak{c}\Big(\frac{n}{h}m\Big)\sum_{\substack{u|\mathcal{Q}\\(u,h)=1}}\sum_{\substack{c|P(z)\\(c,\mathcal{Q}h)=1}}\lambda_c^-\mu(c)\sum_{\substack{d_1,d_2\\p|d_1d_2\Leftrightarrow p|uc}}\mu(d_1)\mu(d_2)\\ \times\prod_{\substack{p^\nu\|d_1d_2\\p\nmid 2}}\mathfrak{w}(p,\nu;m,n,h)\prod_{2^\nu\|d_1d_2}\mathfrak{f}_1(2,\nu;m,n)\end{aligned}$$

(of course each of the products over primes has only one factor, since ν is unique) where

$$\mathfrak{w}(p,\nu;m,n,h) = \begin{cases}\mathfrak{c}(p)\mathfrak{f}_1(p,\nu;m,n), & p\nmid 2\frac{n}{h}m,\\ \mathfrak{f}_1(p,\nu;m,n), & 2\nmid p\mid \frac{n}{h}m.\end{cases} \tag{7.34}$$

We observe that the sum over d_1, d_2 gives

$$\tilde{\mathfrak{f}}_1(2;m,n;u)\prod_{\substack{p|uc\\p\nmid 2}}(-2\mathfrak{w}(p,1;m,n,h)+\mathfrak{w}(p,2;m,n,h))$$

where

$$\tilde{\mathfrak{f}}_1(2;m,n;u) = \begin{cases}-2\mathfrak{f}_1(2,1;m,n)+\mathfrak{f}_1(2,2;m,n), & 2\mid u,\\ \mathfrak{f}_1(2,0;m,n), & 2\nmid u.\end{cases} \tag{7.35}$$

Using the formulas (7.32) and (7.33) instead of (7.30) and (7.31), the same computation evaluates the corresponding u, c, d_1, d_2-sums of S_2^- in (7.24) as

$$\sum_{\substack{u|\mathcal{Q}\\(u,h)=1}}\sum_{\substack{c|P(z)\\(c,\mathcal{Q}h)=1}}\lambda_c^-\mu(c)\sum_{\substack{d_1,d_2\\p|d_1d_2\Leftrightarrow p|uc}}\mu(d_1)\mu(d_2)\mathfrak{c}\Big(d_1d_2\frac{n}{h}m\Big)\prod_{p^\nu\|d_1d_2}\mathfrak{f}_2(p,\nu;m,n)$$

$$= \mathfrak{c}\Big(\frac{n}{h}m\Big) \sum_{\substack{u|\mathcal{Q}\\(u,h)=1}} \sum_{\substack{c|P(z)\\(c,\mathcal{Q}h)=1}} \lambda_c^- \mu(c) \tilde{\mathfrak{f}}_2(2;m,n;u) \times \prod_{2\nmid p\nmid uc} (-2\mathfrak{w}(p,1;m,n,h) + \mathfrak{w}(p,2;m,n,h))$$

where $\tilde{\mathfrak{f}}_2(2;m,n;u)$ is defined in the same way as (7.35), replacing the index 1 with the index 2. Combining the above calculations, we have shown

$$S_1^- + S_2^- = \sum_{j=1}^{2} \sum_{\substack{h|n\\P^+(h)\le y\\ \mu^2(h)=(h,\mathfrak{N})=1}} \chi_{-4}(h) \sum_{\substack{hm\le x\\P^+(m)\le y\\ \mu^2(mh)=1}} \mathfrak{c}\Big(\frac{n}{h}m\Big) \sum_{\substack{u|\mathcal{Q}\\(u,h)=1}} \sum_{\substack{c|P(z)\\(c,\mathcal{Q}h)=1}} \lambda_c^- \mu(c) \times \tilde{\mathfrak{f}}_j(2;m,n;u) \prod_{2\nmid p\mid cu} (-2\mathfrak{w}(p,1;m,n,h) + \mathfrak{w}(p,2;m,n,h)).$$

We are now finally in a position to apply Lemma 7.3 to bound from below the sum over the sieve weights. We make the following important observations. From (7.34), (7.31), (7.28), (6.4) we have

$$\begin{aligned} 0 &\le 2\mathfrak{w}(p,1;m,n,h) - \mathfrak{w}(p,2;m,n,h) \\ &= \left\{ \begin{array}{ll} \frac{1}{p}(2 - \frac{2\chi_{p^*}(n)+1}{p-1})(1+\frac{\chi_{-4}(p)}{p(p-1)})^{-1}, & p\nmid nm \\ \frac{2}{p}(1-\frac{\chi_{-4}(p)}{p})^{-1}(1+\chi_{p^*}(n)), & p\mid \frac{n}{h}m, p\nmid (m,n) \\ \frac{1}{p-1}(1-\frac{\chi_{-4}(p)}{p})^{-1}\mathbb{1}_{\nu_p(n)=1}, & p\mid (m,n) \end{array} \right\} \le \frac{5}{p}, \end{aligned} \tag{7.36}$$

for $p\nmid 2h$. So we can choose $\mathcal{Q}=30$ and use a sieve of dimension $\gamma=5$. Secondly, since $h\mid n$, $(h,2\mathfrak{N})=1$, we have $\chi_{-4}(h)=1$, so that a lower bound for the sum of sieve weights with multiplicative coefficients suffices. This is an important device in the argument. Now applying Lemma 7.3 and (7.22), we obtain

$$S(\mathcal{A},n^{1/C}) \ge \mathcal{S}(n)\Big(f_5(C\kappa_0) + O\Big(\frac{e^{-C\kappa_0}}{(\log\mathcal{D})^{1/3}}\Big)\Big) + O_A(\mathcal{F}(n,y)\Psi(x,y)(\log x)^{-A}), \tag{7.37}$$

where f_5 has the same meaning as in Lemma 7.3, $\mathcal{D}=n^{\kappa_0}$ for some small $\kappa_0>0$ and C a large constant such that $f_5(C\kappa_0)>0$ and

$$\mathcal{S}(n) = \sum_{\substack{h|n\\P^+(h)\le y\\ \mu^2(h)=1\\(h,2\mathfrak{N})=1}} \sum_{\substack{hm\le x\\P^+(m)\le y\\(m,\mathfrak{N})=1\\ \mu^2(m_o h)=1}} \mathfrak{c}\Big(\frac{n}{h}m\Big)\mathfrak{F}_{\mathcal{Q}}(m,n,h) \times \prod_{\substack{p\le z\\p\nmid \mathcal{Q}h}} (1-2\mathfrak{w}(p,1;m,n,h)+\mathfrak{w}(p,2;m,n,h)),$$

where $\mathfrak{F}_{\mathcal{Q}}(m,n,h) = \prod_{\substack{p\mid \mathcal{Q}\\ p\nmid h}} \mathfrak{F}_p(m,n,h)$ with

$$\mathfrak{F}_2(m,n,h) = \sum_{j=1}^{2} (\mathfrak{f}_j(2,0;m,n) - 2\mathfrak{f}_j(2,1;m,n) + \mathfrak{f}_j(2,2;m,n)) \tag{7.38}$$

(the right hand side is independent of h) and

$$\mathfrak{F}_p(m,n,h) = \begin{cases} (1 - 2\mathfrak{w}(p,1;m,n,h) + \mathfrak{w}(p,2;m,n,h)), & p \nmid 2h, \\ 1, & p \mid h. \end{cases} \tag{7.39}$$

Recalling (7.36), we define multiplicative functions $\mathfrak{g}_i$ supported on square-free integers with $\mathfrak{g}_i(p) = 1$ except for $p \mid P_{\mathcal{Q}}(z)$ where

$$\begin{aligned} \mathfrak{g}_0(p) &= 1 - \Big(\frac{2}{p} - \frac{2\chi_{p^*}(n)+1}{p(p-1)}\Big)\Big(1 + \frac{\chi_{-4}(p)}{p(p-1)}\Big)^{-1}, \\ \mathfrak{g}_1(p) &= 1 - \frac{2}{p}\Big(1 + \chi_{p^*}(n)\Big)\Big(1 - \frac{\chi_{-4}(p)}{p}\Big)^{-1}, \\ \mathfrak{g}_2(p) &= 1 - \frac{1}{p-1}\Big(1 - \frac{\chi_{-4}(p)}{p}\Big)^{-1}\mathbb{1}_{\nu_p(n)=1}, \end{aligned} \tag{7.40}$$

so that

$$\begin{aligned} \mathcal{S}(n) = \sum_{\substack{h\mid n\\ P^+(h)\le y\\ \mu^2(h)=(h,2\mathfrak{N})=1}} \sum_{\substack{hm\le x\\ P^+(m)\le y\\ (m,\mathfrak{N})=1\\ \mu^2(m_o h)=1}} \mathfrak{c}\Big(\frac{n}{h}m\Big)\mathfrak{F}_{\mathcal{Q}}(m,n,h) \prod_{\substack{p\nmid nm\\ p\nmid h}} \mathfrak{g}_0(p) \\ \times \prod_{\substack{p\mid \frac{n}{h}m\\ p\nmid h(m,n)}} \mathfrak{g}_1(p) \prod_{\substack{p\mid (m,n)\\ p\nmid h}} \mathfrak{g}_2(p). \end{aligned}$$

We rename mh by m, h by hg where $h \mid (m,n), h \nmid \mathcal{Q}$, and $2 \nmid g \mid (m,n,\mathcal{Q})$. In this way we can re-write $\mathcal{S}(n)$ as

$$\begin{aligned} \prod_{p\mid P_{\mathcal{Q}}(z)} \mathfrak{g}_0(p) \sum_{\substack{m\le x\\ P^+(m)\le y\\ (m,\mathfrak{N})=1\\ \mu^2(m_o)=1}} \sum_{\substack{g\mid (n,m,\mathcal{Q})\\ (g,2\mathfrak{N})=1}} \mathfrak{F}_{\mathcal{Q}}\Big(\frac{m}{g},n,g\Big) \prod_{\substack{p\mid \mathcal{Q}\\ p\mid \frac{n}{g}\frac{m}{g}}} \mathfrak{c}(p) \sum_{\substack{h\mid (m,n)\\ (h,\mathcal{Q}\mathfrak{N})=1}} \frac{1}{\mathfrak{g}_0(h)} \\ \times \prod_{\substack{p\mid (m,n)\\ p\nmid h}} \frac{\mathfrak{g}_2(p)}{\mathfrak{g}_0(p)} \prod_{\substack{p\nmid mn\\ p\nmid (m,n)}} \frac{\mathfrak{g}_1(p)}{\mathfrak{g}_0(p)} \prod_{\substack{p\nmid \mathcal{Q}\\ p\mid \frac{n}{h}\frac{m}{h}}} \mathfrak{c}(p) \end{aligned}$$

$$= \prod_{p\mid P_{\mathcal{Q}}(z)} \mathfrak{g}_0(p) \sum_{\substack{r\mid 2^\infty\mathcal{Q}\\(r,\mathfrak{N})=1}} \Big(\sum_{\substack{g\mid(n,r,\mathcal{Q})\\ \mu^2(g)=1\\(g,2\mathfrak{N})=1}} \mathfrak{F}_{\mathcal{Q}}\Big(\frac{r}{g},n,g\Big) \prod_{\substack{p\mid\mathcal{Q}\\ p\mid \frac{n}{g}\frac{r}{g}}} \mathfrak{c}(p)\Big)$$

$$\times \sum_{\substack{m\le x/r\\ P^+(m)\le y\\ (m,\mathcal{Q}\mathfrak{N})=1\\ \mu^2(m)=1}} \Big(\sum_{h\mid(m,n)} \frac{1}{\mathfrak{g}_0(h)} \prod_{\substack{p\mid(m,n)\\ p\nmid h}} \frac{\mathfrak{g}_2(p)}{\mathfrak{g}_0(p)} \prod_{\substack{p\mid mn\\ p\nmid(m,n)}} \frac{\mathfrak{g}_1(p)}{\mathfrak{g}_0(p)} \prod_{\substack{p\nmid\mathcal{Q}\\ p\nmid \frac{n}{h}\frac{m}{h}}} \mathfrak{c}(p)\Big).$$

We first see that the sum over g can be written as (note that $\mathfrak{F}_p(\bullet,\bullet,p)=1$)

$$\prod_{p\mid\mathcal{Q}} \Big(\mathfrak{F}_p(r,n,1) \prod_{p\mid rn} \mathfrak{c}(p) + \mathbb{1}_{\substack{p\mid(n,r)\\ p\nmid 2\mathfrak{N}}} \cdot \prod_{p\mid \frac{n}{p}\frac{r}{p}} \mathfrak{c}(p)\Big) =: \mathfrak{F}_{\mathcal{Q}}(r,n) = \prod_{p\mid\mathcal{Q}} \mathfrak{F}_p(r,n).$$

With $\mathcal{Q}=30$, $\mu^2(r_o)=1$, this becomes

$$\mathfrak{F}_2(r,n,1)\mathfrak{F}_3(r,n,1) \prod_{3\mid nr} \mathfrak{c}(3) \Big(\mathfrak{F}_5(r,n,1) \prod_{5\mid nr} \mathfrak{c}(5) + \mathbb{1}_{5\mid(n,r)} \cdot \prod_{p\mid\frac{n}{5}} \mathfrak{c}(5)\Big).$$

Using (7.38), (7.30), (7.32) and (7.39), (7.36), we compute explicitly

$$\mathfrak{F}_2(r,n) = \begin{cases} \frac{1}{2}\mathbb{1}_{2\nmid r}, & n\equiv 1 \,(\mathrm{mod}\, 2),\\ 2\mathbb{1}_{8\mid r}, & n\equiv 2 \,(\mathrm{mod}\, 8),\\ 2\mathbb{1}_{4\parallel r}, & n\equiv 6 \,(\mathrm{mod}\, 8),\\ \mathbb{1}_{2\parallel r}, & n\equiv 0 \,(\mathrm{mod}\, 4). \end{cases}, \quad \mathfrak{F}_3(r,n) = \begin{cases} \frac{4}{5}\mathbb{1}_{3\nmid r}, & n\equiv 1 \,(\mathrm{mod}\, 3),\\ \frac{8}{5}\mathbb{1}_{3\mid r}, & n\equiv 2 \,(\mathrm{mod}\, 3),\\ \frac{4}{5}\mathbb{1}_{3\nmid r}, & n\equiv 0 \,(\mathrm{mod}\, 3), \end{cases} \tag{7.41}$$

(note that if $3\mid n$ then $(r,3)=1$ since $(r,\mathfrak{N})=1$) and (with $\mathfrak{c}(5)=\frac{16}{21}$)

$$\mathfrak{F}_5(r,n) = \begin{cases} \mathfrak{c}(5)\mathbb{1}_{5\nmid r}, & n\equiv 1,4 \,(\mathrm{mod}\, 5),\\ \mathfrak{c}(5)\mathbb{1}_{5\mid r} + \frac{4}{7}\mathbb{1}_{5\nmid r}, & n\equiv 2,3 \,(\mathrm{mod}\, 5),\\ 2\mathfrak{c}(5)\mathbb{1}_{\substack{5\mid n\\ 5\mid r}} + \frac{1}{2}\mathfrak{c}(5)\mathbb{1}_{5\nmid r}, & n\equiv 0 \,(\mathrm{mod}\, 5). \end{cases} \tag{7.42}$$

In particular, we can always find some residue class $r_0 \,(\mathrm{mod}\, 120)$ such that $\mathfrak{F}_{\mathcal{Q}}(r,n) > 0$ for $r\equiv r_0 \,(\mathrm{mod}\, 120)$.

Noting that $(m,\mathcal{Q}) = \mu^2(m) = 1$ and $\mathfrak{g}_i(p)=1$ if $p\mid\mathcal{Q}$, we next evaluate the inner sum over h as

$$
\begin{aligned}
&\sum_{h|(m,n)} \frac{1}{\mathfrak{g}_0(h)} \prod_{\substack{p|(m,n)\\ p\nmid h}} \frac{\mathfrak{g}_2(p)}{\mathfrak{g}_0(p)} \prod_{\substack{p|mn\\ p\nmid(m,n)}} \frac{\mathfrak{g}_1(p)}{\mathfrak{g}_0(p)} \prod_{\substack{p\nmid\mathcal{Q}\\ p|\frac{n}{h}\frac{m}{h}}} \mathfrak{c}(p)\\
&\quad= \sum_{h|(m,n)} \frac{1}{\mathfrak{g}_0(h)} \prod_{p|h} \frac{\mathfrak{g}_0(p)}{\mathfrak{g}_2(p)} \prod_{\substack{p|h\\ \nu_p(n)=1}} \frac{1}{\mathfrak{c}(p)} \prod_{p|(m,n)} \frac{\mathfrak{g}_2(p)}{\mathfrak{g}_0(p)} \prod_{\substack{p|m\\ p\nmid n}} \frac{\mathfrak{g}_1(p)}{\mathfrak{g}_0(p)} \mathfrak{c}(p)\\
&\qquad\times \prod_{\substack{p|n\\ p\nmid m}} \frac{\mathfrak{g}_1(p)}{\mathfrak{g}_0(p)} \prod_{\substack{p|n\\ p\nmid m\mathcal{Q}}} \mathfrak{c}(p) \prod_{p|(m,n)} \mathfrak{c}(p)\\
&\quad= \sum_{h|(m,n)} \Big(\frac{1}{\mathfrak{g}_0(h)} \prod_{p|h} \frac{\mathfrak{g}_0(p)}{\mathfrak{g}_2(p)} \prod_{\substack{p|h\\ \nu_p(n)=1}} \frac{1}{\mathfrak{c}(p)}\Big)\\
&\qquad\times \prod_{p|(m,n)} \frac{\mathfrak{g}_2(p)}{\mathfrak{g}_1(p)} \prod_{\substack{p|m\\ p\nmid n}} \frac{\mathfrak{g}_1(p)}{\mathfrak{g}_0(p)} \mathfrak{c}(p) \prod_{\substack{p\nmid\mathcal{Q}\\ p|n}} \frac{\mathfrak{g}_1(p)}{\mathfrak{g}_0(p)} \mathfrak{c}(p)\\
&\quad= \prod_{\substack{p|(m,n)\\ \nu_p(n)=1}} \Big(1+\frac{1}{\mathfrak{c}(p)\mathfrak{g}_2(p)}\Big) \prod_{\substack{p|(m,n)\\ \nu_p(n)>1}} \Big(1+\frac{1}{\mathfrak{g}_2(p)}\Big)\\
&\qquad\times \prod_{p|(m,n)} \frac{\mathfrak{g}_2(p)}{\mathfrak{g}_1(p)} \prod_{\substack{p|m\\ p\nmid n}} \frac{\mathfrak{g}_1(p)}{\mathfrak{g}_0(p)} \mathfrak{c}(p) \prod_{\substack{p\nmid\mathcal{Q}\\ p|n}} \frac{\mathfrak{g}_1(p)}{\mathfrak{g}_0(p)} \mathfrak{c}(p)\\
&\quad=: \mathfrak{G}(m;n) \prod_{\substack{p\nmid\mathcal{Q}\\ p|n}} \frac{\mathfrak{g}_1(p)}{\mathfrak{g}_0(p)} \mathfrak{c}(p),
\end{aligned}
\tag{7.43}
$$

where $\mathfrak{G}(m;n)$ is a multiplicative function of m with $\mathfrak{G}(p^\nu;n)=0$ for $\nu\geq 2$ and $\mathfrak{G}(p;n)=0$ for $p\mid\mathcal{Q}\mathfrak{N}$. We summarize

$$
\mathcal{S}(n) = \prod_{p|P_{\mathcal{Q}}(z)} \mathfrak{g}_0(p) \sum_{\substack{r|2^\infty\mathcal{Q}\\ (r,\mathfrak{N})=1}} \mathfrak{F}_{\mathcal{Q}}(r,n) \prod_{\substack{p\nmid\mathcal{Q}\\ p|n}} \frac{\mathfrak{g}_1(p)}{\mathfrak{g}_0(p)} \mathfrak{c}(p) \sum_{\substack{m\leq x/r\\ P^+(m)\leq y}} \mathfrak{G}(m;n). \tag{7.44}
$$

7.7 Computation of the main term IV: summing over smooth square-free numbers. Our last goal is a lower bound for $\mathcal{S}(n)$ based on the formula (7.44).

In order to carry out the sum over m in (7.44), it is useful to define $\mathfrak{h}(\bullet;n)=\mathfrak{G}(\bullet;n)*\mu$. Recall that the definition of $\mathfrak{G}(m;n)$ is implicit in (7.43), so that $\mathfrak{h}(p^\nu;n)=0$ if $\nu>2$, and for $\nu\leq 2$ we have

$$\mathfrak{h}(p;n)=\begin{cases}(1+\frac{1}{\mathfrak{c}(p)\mathfrak{g}_2(p)})\frac{\mathfrak{g}_2(p)}{\mathfrak{g}_1(p)}-1, & p\mid n, p\nmid \mathcal{Q}\mathfrak{N}, \nu_p(n)=1,\\ (1+\frac{1}{\mathfrak{g}_2(p)})\frac{\mathfrak{g}_2(p)}{\mathfrak{g}_1(p)}-1, & p\mid n, p\nmid \mathcal{Q}\mathfrak{N}, \nu_p(n)>1,\\ -1, & p\mid \mathcal{Q}\mathfrak{N},\\ \frac{\mathfrak{g}_1(p)}{\mathfrak{g}_0(p)}\mathfrak{c}(p)-1, & p\nmid n\mathcal{Q}\end{cases}$$

and

$$\mathfrak{h}(p^2;n)=-\mathfrak{h}(p,n).$$

In particular, from (7.40) and (6.4) we have

$$\mathfrak{h}(p;n)=\begin{cases}1+\frac{4}{p-3}, & p\mid n, p\nmid \mathcal{Q}\mathfrak{N},\\ -1, & p\mid \mathcal{Q}\mathfrak{N},\\ O(\frac{1}{p}), & p\nmid n\mathcal{Q},\end{cases}\qquad \mathfrak{h}(p^2;n)=O(1), \tag{7.45}$$

where we can take 5 as implied constants. Using (7.1) and $\alpha\geq 1-\frac{1}{D}+o(1)$ with D sufficiently large, the contribution from $r\geq K=n^{1/100}$ in (7.44) can be bounded by

$$\Psi(x,y)\sum_{t_1t_2^2\geq K}\frac{(t_1,n)}{t_1(t_1t_2^2)^{\alpha}}\ll \Psi(x,y)\frac{1}{K^{2\alpha/5}}\sum_{t_1,t_2}\frac{(t_1,n)}{t_1(t_1t_2^2)^{3\alpha/5}}\ll K^{-1/3}n^{\varepsilon}\Psi(x,y). \tag{7.46}$$

To evaluate the contribution from $r\leq K$, we consider two cases depending on the size of y.

7.7.1 Case I: $y\geq \exp((\log_2 x)^2)$. We write

$$\sum_{\substack{m\leq x/r\\ P^+(m)\leq y}}\mathfrak{G}(m;n)=\sum_{\substack{t\leq x/r\\ P^+(t)\leq y}}\mathfrak{h}(t;n)\Psi\left(\frac{x}{tr},y\right). \tag{7.47}$$

We can truncate the t-sum using (7.1) as in (7.46) so that the contribution from $t\geq K$ (7.47) can be bounded by $K^{-1/3}n^{\varepsilon}\Psi(x,y)$. For $t,r\leq K$ we evaluate the sum by [FT2, Theorem 2 with $q=1$] getting

$$\sum_{\substack{t\leq K\\ P^+(t)\leq y}}\mathfrak{h}(t;n)\frac{x}{tr}\int_{0^-}^{\infty}\rho\Big(\frac{\log x}{\log y}-\frac{\log(tr)}{\log y}-v\Big)dR(y^v)\Big(1+O\big(\exp(-(\log y)^{0.55})\big)\Big). \tag{7.48}$$

where $R(x)=[x]/x$. Strictly speaking, the application of [FT2, Theorem 2] requires $y\leq x/K^2\leq x/(tr)$, which we can assume without loss of generality since $\Psi(x,y)\asymp\Psi(x,x)$ for $y>x^{1/2}$, say.

To estimate the error term, we use (7.1) again, and obtain the bound

$$\begin{aligned}\sum_{t\le K}|\mathfrak{h}(t;n)|\Psi\Big(\frac{x}{tr},y\Big)\exp(-(\log y)^{0.55})\\ \ll\prod_{p|n}\Big(1+\frac{O(1)}{p^\alpha}\Big)\Psi(x,y)\exp(-(\log y)^{0.55})\\ \ll\Psi(x,y)\frac{(\log n)^{O(1)}}{\exp((\log y)^{0.55})}\ll\frac{\Psi(x,y)}{(\log x)^{100}}\end{aligned}\tag{7.49}$$

for $y\ge\exp((\log_2 x)^2)$.

We now focus on the main term in (7.48) which we substitute back as the m-sum in (7.44). By partial summation, the main term equals

$$\prod_{p|P_{\mathcal{Q}}(z)}\mathfrak{g}_0(p)\cdot x\int_{0^-}^{\infty}\rho(u-v)dW(y^v)\tag{7.50}$$

where

$$W(U)=\frac{1}{U}\prod_{\substack{p\nmid\mathcal{Q}\\p|n}}\frac{\mathfrak{g}_1(p)}{\mathfrak{g}_0(p)}\mathfrak{c}(p)\sum_{\substack{r\le K\\r|2^\infty\mathcal{Q}\\(r,\mathfrak{N})=1}}\mathfrak{F}_{\mathcal{Q}}(r,n)\sum_{\substack{t\le K\\P^+(t)\le y}}\mathfrak{h}(t;n)\sum_{m\le U/tr}1.$$

It is clear that $W(U)=0$ for $U<1$ and $W(U)=c+O(1/U)$ for some constant c (depending on n and z). In preparation for an application of [FT1, Lemme 4.1] we define

$$M(s)=\int e^{-sv}dW(e^v).$$

We compute $M(s)$ as

$$M(s)=\mathfrak{S}_K(n,s)\frac{s}{s+1}\zeta(s+1)$$

where

$$\mathfrak{S}_K(n,s)=\prod_{\substack{p\nmid\mathcal{Q}\\p|n}}\frac{\mathfrak{g}_1(p)}{\mathfrak{g}_0(p)}\mathfrak{c}(p)\sum_{\substack{r\le K\\r|2^\infty\mathcal{Q}\\(r,\mathfrak{N})=1}}\frac{\mathfrak{F}_{\mathcal{Q}}(r,n)}{r^{s+1}}\sum_{\substack{t\le K\\P^+(t)\le y}}\frac{\mathfrak{h}(t;n)}{t^{s+1}}.$$

Recalling (7.40) and (6.4) we have

$$\frac{60}{67}\le\frac{\mathfrak{g}_1(p)}{\mathfrak{g}_0(p)}\mathfrak{c}(p)=1-\frac{\chi_{-4}(p)}{p}+O(p^{-2}),\quad p\mid n,\ p\nmid\mathcal{Q}=30.\tag{7.51}$$

We can extend r, t to infinity getting

$$\mathfrak{S}(n,s) = \prod_{\substack{p\nmid \mathcal{Q}\\ p\mid n}} \frac{\mathfrak{g}_1(p)}{\mathfrak{g}_0(p)}\mathfrak{c}(p)\sum_{k=0}^{\infty}\frac{\mathfrak{F}_2(2^k,n)}{2^{k(s+1)}}$$
$$\times \prod_{\substack{p\mid \mathcal{Q}\\ (p,2\mathfrak{N})=1}} \Big(\mathfrak{F}_p(1,n)+\frac{\mathfrak{F}_p(p,n)}{p^{s+1}}\Big)\prod_{p\le y}\Big(1+\frac{\mathfrak{h}(p;n)}{p^{s+1}}+\frac{\mathfrak{h}(p^2;n)}{p^{s+2}}\Big)$$

where uniformly in $|s|\le 1/2$ we have (once again by Rankin's trick)

$$\mathfrak{S}(n,s)-\mathfrak{S}_K(n,s)\ll \frac{1}{K^{1/2}}\prod_{p\mid n}\Big(1+\frac{1}{p^{\Re s+1/2}}\Big)\ll \frac{1}{K^{1/4}}=n^{-1/400}, \quad (7.52)$$

and by Cauchy's formula this bound remains true for all fixed derivatives with respect to s. Note that by (7.45) we have

$$\frac{6}{7}\le 1+\frac{\mathfrak{h}(p;n)}{p}+\frac{\mathfrak{h}(p^2;n)}{p^2}=1+\frac{\chi_{-4}(p)}{p}+O(p^{-2}), \quad p\mid n. \quad (7.53)$$

Combing (7.41), (7.42), (7.45), (7.51) and (7.53), we obtain for all $j\ge 0$ that

$$\frac{d^j}{ds^j}\mathfrak{S}(n,s)|_{s=0}\asymp \prod_{\substack{p\mid n\\ p\le \min(y,z)}}\Big(1+O\Big(\frac{(\log p)^j}{p^2}\Big)\Big)\prod_{\substack{p\mid n\\ p\ge \min(y,z)}}\Big(1+O\Big(\frac{(\log p)^j}{p}\Big)\Big)\ll 1$$

for $y\ge \exp((\log_2 x)^2)$ and $\log z\asymp \log n$, and we also see $\mathfrak{S}(n,0)\gg 1$.

We are now prepared to apply [FT1, Lemme 4.1] to evaluate (7.50), which we complement with [FT2, Lemme 4.1] for bounds of $\rho^{(j)}(u)$. Using again $y\ge \mathcal{L}(n)$ and $\log z\asymp \log n$ and recalling the error terms (7.49), (7.46) and (7.52), we conclude

$$\mathcal{S}(n)\asymp \Big(\prod_{p\mid P_{\mathcal{Q}}(z)}\mathfrak{g}_0(p)\Big)\mathfrak{S}(n,0)x\rho(u)\Big(1+O\Big(\frac{\log\log n}{\log y}\Big)\Big)+O\Big(\frac{\Psi(x,y)}{(\log x)^{100}}\Big)$$
$$\gg \frac{x\rho(u)}{(\log z)^2}+O\Big(\frac{\Psi(x,y)}{(\log x)^{100}}\Big)\asymp \frac{\Psi(x,y)}{(\log z)^2}.$$

We plug this lower bound back into (7.37), noting that $\mathcal{F}(n,y)\asymp 1$ when $y\ge \exp((\log_2 x)^2)$, and choose $C=C(\kappa_0)$ a sufficiently large constant so that $f_5(C\kappa_0)>0$ getting

$$S(\mathcal{A},n^{1/C})\gg (\log z)^{-2}\Psi(x,y)$$

provided that $y\ge \exp((\log_2 x)^2)$ and $\mathcal{D}=n^{\kappa_0}$ for some sufficiently small $\kappa_0>0$.

This concludes the proof of Theorem 1.3 when $y\ge \exp((\log_2 x)^2)$.

REMARK. We close this computation by a short comment on the "singular series" $\mathfrak{S}(n,0)$ encountered in the final lower bound, which might look a bit unexpected. This is due to our restrictions on m, in particular the fact that we impose a square-free condition on m. As square-free numbers are not equidistributed among all residue classes, the singular series $\mathfrak{S}(n,0)$ will genuinely depend on the factorization of n in our set-up. For example, we can see from (7.41) that there are fewer solutions to $n = m + x^2 + y^2$ with $(xy,3) = 1$, $\mu^2(m_o) = 1$ when $n \equiv 2 \pmod 3$ than in the other cases, whilst the number of solutions to $n = m + x^2 + y^2$ with $(xy,3) = 1$ and no extra conditions on m would be approximately equal for all residue classes of $n \pmod 3$.

7.7.2 Case II: $(\log x)^D \le y \le \exp((\log_2 x)^2)$.
When y is small, the factor $\prod_{p|n}(1+\frac{O(1)}{p^\alpha})$ in (7.49) can be much bigger than a large power of $\log x$ and so we cannot bound $\mathfrak{h}(t;n)$ by absolute values. Instead, we apply Perron's formula to evaluate the m-sum in (7.44) directly as follows. Write

$$\mathcal{S}_0(n,x/r) := \sum_{\substack{m\le x/r\\ P^+(m)\le y}} \mathfrak{G}(m;n) = \frac{1}{2\pi i}\int_{(\alpha)} G(s;n)\Big(\frac{x}{r}\Big)^s \frac{ds}{s} \tag{7.54}$$

where $G(s;n) = \sum_{m=1}^\infty \mathfrak{G}(m;n)n^{-s}$ and as before α is defined in (1.6). (Recall that for D large enough we have $9/10 \le \alpha \le 1$.) Denote

$$\zeta_m(s,y) = \prod_{\substack{p\le y\\ p\nmid m}} (1-p^{-s})^{-1}, \quad \zeta(s,y) = \zeta_1(s,y).$$

Using (7.53) we can write

$$G(s;n) = \sum_{m=1}^\infty \frac{\mathfrak{h}(m;n)}{m^s}\zeta(s,y) = \prod_{p|n_y}\Big(1-\frac{1}{p^s}\Big)^{-\chi_{-4}(p)}\zeta(s,y)\tilde{G}(s;n) \tag{7.55}$$

for some function $\tilde{G}(s;n)$ which is holomorphic and uniformly bounded in $\Re(s) \ge 2/3$, say. We use the saddle point method to evaluate (7.54). The main term comes from a small neighbourhood of the point $s = \alpha$. Precisely, we truncate the integral in (7.54) at height $1/\log y$ to write

$$\mathcal{S}_0(n,x/r) = \frac{1}{2\pi i}\int_{\alpha-i/\log y}^{\alpha+i/\log y} G(s;n)\Big(\frac{x}{r}\Big)^s\frac{ds}{s} + E(x,y), \tag{7.56}$$

say. To handle the error term $E(x,y)$ we need some estimate of $G(s;n)$ when $s = \alpha + i\tau$ for $|\tau| \ge 1/\log y$. This can be done following the proof of the bound for $\zeta_m(s,y)/\zeta_m(\alpha;y)$ in [IBT, (4.42)]. In fact, the bound for $\zeta_m(s,y)/\zeta_m(\alpha;y)$ follows from the case of $m = 1$ together with bounds for the contribution from primes

satisfying $p \mid m$. In our case, we can bound the contribution from $p \mid n_y$ in the same way (since $\omega(n_y) \ll \sqrt{y}$ for $y \geq \mathcal{L}(x)$) and thus we can obtain for $s = \alpha + i\tau$ that

$$\frac{G(s;n)}{G(\alpha;n)} \ll \begin{cases} \exp\Big(-\frac{c_0 y}{\log y}\log\Big(1+\frac{\tau^2\sigma_2}{y/\log y}\Big)\Big), & |\tau| \leq 1/\log y, \\ \exp\Big(\frac{-c_0\tau^2 u}{(1-\alpha)^2+\tau^2}\Big), & 1/\log y < |\tau| \leq Y_\varepsilon \end{cases} \tag{7.57}$$

where $Y_\varepsilon = \exp\big((\log y)^{3/2-\varepsilon}\big)$ and

$$\sigma_k = \frac{d^k}{ds^k}\log\zeta(s,y)\Big|_{s=\alpha}. \tag{7.58}$$

Note that we have (see e.g. [lBT, eq. (3.9)])

$$\sigma_k \asymp u(\log y)^k \asymp \log x(\log y)^{k-1} \tag{7.59}$$

for $k \geq 1$. Similarly to the proof of [lBT, eq. (4.49)], we can show that for $1 \leq z \leq Y_\varepsilon$ we have

$$S_0(n, x+x/z) - S_0(n,x) \ll x^\alpha G(\alpha;n)(1/z + e^{-cu}) \tag{7.60}$$

for some small $c > 0$. Set

$$R := e^{-c'u/(\log 2u)^2} + 1/Y_\varepsilon \tag{7.61}$$

for some $c' > 0$. With (7.57) and (7.60), we can follow [HT, Lemma 10] to bound the contribution from $1/\log y \leq |\tau| \leq T$ and $|\tau| \geq T$ in (7.54) respectively with $T := 1/R^2$, so that (7.56) holds with

$$E(x,y) \ll (x/r)^\alpha G(\alpha;n)R,$$

which can be bounded using (7.61) and (7.59) by

$$E(x,y) \ll \frac{x^\alpha G(\alpha;n)}{r^\alpha \alpha\sqrt{\sigma_2 u}}. \tag{7.62}$$

It remains to evaluate the integral in (7.56). Let $T_0 = (u^{1/3}\log y)^{-1}$. Using (7.57), the contribution from $|\tau| \leq T_0$ can be bounded (see also [HT, p. 281])

$$\begin{aligned} \frac{x^\alpha}{r^\alpha}\frac{G(\alpha;n)}{\alpha}\int_{T_0}^{1/\log y}\Big(1+\frac{\tau^2\sigma_2}{y/\log y}\Big)^{-c_0 y/\log y}d\tau &\ll \frac{x^\alpha}{r^\alpha}\frac{G(\alpha;n)}{\alpha\sqrt{\sigma_2}}\int_{T_0\sqrt{\sigma_2}} e^{-c_0\tau^2/2}d\tau \\ &\ll \frac{x^\alpha}{r^\alpha}\frac{G(\alpha;n)e^{-c_1 u^{1/3}}}{\alpha\sqrt{\sigma_2}}. \end{aligned} \tag{7.63}$$

The contribution from $|\tau| \leq T_0$ in the integral (7.56) gives the main term. We can show that

$$\frac{1}{2\pi i}\int_{\alpha-T_0}^{\alpha+T_0} G(\alpha;n)\frac{(x/r)^s}{s}ds = \frac{x^\alpha G(\alpha;n)}{r^\alpha\alpha\sqrt{2\pi\sigma_2}}\Big(1+O\Big(\frac{1}{u^{1/3}}\Big)\Big). \tag{7.64}$$

To prove (7.63), we compare it with the estimation for $\Psi(x,y)$. Using [HT, last equation on p. 280 in the proof of Lemma 11] which states that uniformly for $x \geq y \geq 2$ we have

$$\frac{1}{2\pi i}\int_{\alpha-iT_0}^{\alpha+iT_0}\zeta(s,y)\frac{x^s}{s}ds = \frac{x^\alpha\zeta(\alpha,y)}{\alpha\sqrt{2\pi\sigma_2}}\Big(1+O\Big(\frac{1}{u}\Big)\Big),$$

we can write

$$\begin{aligned}&\frac{1}{2\pi i}\int_{\alpha-iT_0}^{\alpha+iT_0}G(\alpha;n)\frac{(x/r)^s}{s}ds\\ =&\frac{x^\alpha G(\alpha;n)}{r^\alpha\alpha\sqrt{2\pi\sigma_2}}\Big(1+O\Big(\frac{1}{u}\Big)\Big)+\frac{1}{2\pi i}\int_{\alpha-iT_0}^{\alpha+iT_0}\Big(\frac{G(s;n)}{\zeta(s,y)}-\frac{G(\alpha;n)}{\zeta(\alpha,y)}\Big)\zeta(s,y)\Big(\frac{x}{r}\Big)^s\frac{ds}{s}.\end{aligned} \tag{7.65}$$

Thus it remains to show that second term in (7.65) is of smaller order. Set

$$\gamma(s) := \log\frac{G(s;n)}{\zeta(s,y)}.$$

Then we have uniformly for $1/2 < \sigma < 1$ and $\Re(s) = \sigma$ that

$$|\gamma'(s)| \leq \gamma'(\sigma) \ll \sum_{p|n_y}\frac{\log p}{p^\sigma-1} \ll (\log n)^{1-\sigma}, \tag{7.66}$$

$$|\gamma''(s)| \leq \gamma'(\sigma) \ll \sum_{p|n_y}\frac{(\log p)^2p^\sigma}{(p^\sigma-1)^2} \ll (\log n)^{1-\sigma}\log_2 n. \tag{7.67}$$

We have the Taylor expansion around $\tau = 0$ for $|\tau| \leq T_0 \ll (\sigma_3)^{-1/3}$ (see e.g. [lBT, eq. (4.54)])

$$\frac{\zeta(s,y)x^s}{s} = \frac{\zeta(\alpha,y)x^\alpha}{s}e^{-\tau^2\sigma_2/2}(1+O(\tau^3\sigma_3+\tau/\alpha))$$

where σ_i is as in (7.58), and for $|\tau| \leq T_0$ we have

$$\begin{aligned}\frac{G(s;n)}{\zeta(s,y)}-\frac{G(\alpha;n)}{\zeta(\alpha,y)} &= \frac{G(\alpha;n)}{\zeta(\alpha;n)}\exp\Big(i\tau\gamma'(\alpha)+O(\tau^2\gamma''(\alpha))-1\Big)\\ &= \frac{G(\alpha;n)}{\zeta(\alpha;n)}\Big(i\tau\gamma'(\alpha)+O\big(\tau^2(|\gamma''(\alpha)|+\gamma'(\alpha)^2)\big)\Big).\end{aligned}$$

Using this together with the formulas

$$\int_{-T_0}^{T_0}\tau e^{-\sigma_2\tau^2/2}d\tau = 0, \quad \int_{-T_0}^{T_0}|\tau|^k e^{-\sigma_2\tau^2/2}d\tau \ll_k \sigma_2^{-(k+1)/2} \quad (k \geq 1),$$

we see that

$$\frac{1}{2\pi i}\int_{\alpha-iT_0}^{\alpha+iT_0}\Big(\frac{G(s;n)}{\zeta(s,y)}-\frac{G(\alpha;n)}{\zeta(\alpha,y)}\Big)\zeta(s,y)\Big(\frac{x}{r}\Big)^s\frac{ds}{s}$$
$$\ll\frac{x^\alpha G(\alpha;n)}{r^\alpha\alpha\sqrt{\sigma_2}}\left(|\gamma'(\alpha)|\Big(\frac{\sigma_3}{\sigma_2^2}+\frac{1}{\sigma_2\alpha}\Big)+\frac{|\gamma''(\alpha)|+\gamma'(\alpha)^2}{\sigma_2}\right).$$

Using (7.59), (7.66) and (7.67) this can be bounded by (for $\alpha\geq 9/10$)

$$\frac{x^\alpha G(\alpha;n)}{r^\alpha\alpha\sqrt{\sigma_2}}\Big(\frac{|\gamma'(\alpha)|}{u\log y}+\frac{|\gamma''(\alpha)|+(\gamma'(\alpha))^2}{u(\log y)^2}\Big)\ll\frac{x^\alpha}{r^\alpha}\frac{G(\alpha;n)}{\alpha\sqrt{\sigma_2}}u^{-1/3}.$$

This completes the proof of (7.63).

Combining (7.56), (7.62), (7.63) and (7.64), we see that

$$\mathcal{S}_0(n,x/r)=\frac{x^\alpha G(\alpha;n)}{r^\alpha\alpha\sqrt{2\pi\sigma_2}}\Big(1+O\Big(\frac{1}{u^{1/3}}\Big)\Big).$$

Recalling (7.51), (7.55) and $\Psi(x,y)\asymp\frac{x^\alpha\zeta(\alpha,y)}{\alpha\sqrt{2\pi\sigma_2}}$ (see e.g. [HT, Theorem 1]), we conclude that when $y\geq\mathcal{L}(x)$ the formula (7.44) becomes

$$\mathcal{S}(n)=\prod_{p|P_{\mathcal{Q}}(z)}\mathfrak{g}_0(p)\sum_{\substack{r|2^\infty\mathcal{Q}\\(r,\mathfrak{N})=1}}\mathfrak{F}_{\mathcal{Q}}(r,n)\prod_{\substack{p\nmid\mathcal{Q}\\p|n}}\frac{\mathfrak{g}_1(p)}{\mathfrak{g}_0(p)}\mathfrak{c}(p)\sum_{\substack{m\leq x/r\\P^+(m)\leq y}}\mathfrak{G}(m;n)$$
$$\asymp\prod_{p|n}\Big(1-\frac{\chi_{-4}(p)}{p}\Big)\prod_{\substack{p|n\\p\leq y}}\Big(1+\frac{\chi_{-4}(p)}{p^\alpha}\Big)\Psi(x,y)\Big(1+O\Big(\frac{1}{u^{1/3}}\Big)\Big).$$

We plug this lower bound back into (7.37) and choose $C=C(\kappa_0)$ a sufficiently large constant so that $f_5(C\kappa_0)>0$ getting

$$S(\mathcal{A},n^{1/C})\gg(\log z)^{-2}\mathcal{F}(n,y)\Psi(x,y)$$

provided that $y\geq\mathcal{L}(x)$ and $\mathcal{D}=n^{\kappa_0}$ for some sufficiently small $\kappa_0>0$.

This concludes the proof of Theorem 1.3 when $\mathcal{L}(x)\leq y\leq\exp((\log_2 x)^2)$.

Acknowledgments

We would like to thank the referee for a careful reading of the manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL.

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

[AC] J. Arthur, L. Clozel, *Simple Algebras, Base Change, and the Advanced Theory of the Trace Formula*, Annals of Math. Studies **120**, Princeton University Press 1990

[ABL] E. Assing, V. Blomer, J. Li, *Uniform Titchmarsh divisor problems*, Adv. Math. **393** (2021), 108076.

[Ba] A. Balog, *On additive representation of integers*, Acta Math. Hungar. **54** (1989), 297-301.

[BBD] V. Blomer, J. Brüdern, R. Dietmann, *Sums of smooth squares*, Compos. Math. **145** (2009), 1401-1441.

[BFI] E. Bombieri, J. B. Friedlander and H. Iwaniec, *Primes in arithmetic progressions to large moduli. II*, Math. Ann. **277** (1987) 361-393.

[BS] Z. I. Borevich, I. R. Shafarevich, *Number Theory*, Academic Press 1966

[BF] J. Brüdern, É. Fouvry, *Lagrange's four squares theorem with almost prime variables*, J. Reine Angew. Math. **454** (1994), 59–96.

[lBT] R. de la Bretèche and G. Tenenbaum, *Propriétés statistiques des entiers friables*, Ramanujan J. **9** (2005), 139–202.

[Co] D. A. Cox, *Primes of the form $x^2 + ny^2$*, John Wiley, 1989.

[Dr1] S. Drappeau, *Théorèmes de type Fouvry-Iwaniec pour les entiers friables*, Compositio Math. **151** (2015), 828–862.

[Dr2] S. Drappeau, *Sums of Kloosterman sums in arithmetic progressions, and the error term in the dispersion method*, Proc. Lond. Math. Soc. **114** (2017), 684-732.

[DT] S. Drappeau, B. Topacogullari, *Combinatorial identities and Titchmarsh's divisor problem for multiplicative functions*, Algebra Number Theory **13** (2019), 2383–2425.

[DFI] W. Duke, J. Friedlander, H. Iwaniec. *Bounds for automorphic L-functions*, Invent. Math. **112** (1993), 1–8.

[FT1] É. Fouvry, G. Tenenbaum, *Diviseurs de Titchmarsh des entiers sans grand facteur premier*, in: Analytic number theory (Tokyo, 1988), Lecture Notes in Mathematics **1434** (Springer, Berlin, 1990), 86–102.

[FT2] É. Fouvry, G. Tenenbaum, *Entiers sans grand facteur premier en progressions arithmétiques*, Proc. London Math. Soc. (3) **63** (1991), 449–494.

[FT3] E. Fouvry, G. Tenenbaum, *Répartition statistique des entiers sans grand facteur premier dans les progressions arithmétiques*, Proc. London Math. Soc. **72** (1996), 481-514.

[FT4] E. Fouvry, G. Tenenbaum, *Multiplicative functions in large arithmetic progressions and applications* Trans. Amer. Math. Soc. **375** (2022), no. 1, 245-299.

[FI1] J. B. Friedlander, H. Iwaniec, *Opera de cribro*, Coll. Publ. **57** (2010), AMS, Providence, RI.

[FI2] J. B. Friedlander and H. Iwaniec, *Coordinate distribution of Gaussian primes*, J. Eur. Math. Soc. **24** (2022), 737–772.

[Ha] A. Harper, *Bombieri-Vinogradov and Barban-Davenport-Halberstam type theorems for smooth numbers*, arXiv:1208.5992

[HL] G. H. Hardy, J. E. Littlewood, *Some problems of partitio numerorum; III: on the expression of a number as a sum of primes*, Acta Math. **44** (1922), 1-70.

[HB] D. R. Heath-Brown, *A new form of the circle method, and its application to quadratic forms*, J. Reine Angew. Math. **481** (1996), 149-206.

[HBT] D. R. Heath-Brown, D. I. Tolev, *Lagrange's four squares theorem with one prime and three almost-prime variables*, J. Reine Angew. Math. **558** (2003), 159–224.

[HT] A. Hildebrand, G. Tenenbaum, *On integers free of large prime factors*. Trans. Amer. Math. Soc. **296** (1986), 265-290.

[Ho1] C. Hooley, *On the representation of a number as the sum of two squares and a prime*, Acta Math. **97** (1957), 189-210.

[Ho2] C. Hooley, *Applications of sieve methods to the theory of numbers*, Cambridge Tracts in Mathematics **70**, Cambridge University Press 1976

[Ho3] C. Hooley, *On the representation of a number as the sum of a prime and two squares of square-free numbers*, Acta Arith. **182** (2018), 201-229.

[Iw1] H. Iwaniec, *Primes of the type $\phi(x,y)+A$ where ϕ is a quadratic form*, Acta Arith. **21** (1972), 203-234.

[Iw2] H. Iwaniec, *Rosser's sieve*, Acta Arith. **36** (1980), 171-202.

[Iw3] H. Iwaniec, *A new form of the error term in the linear sieve*, Acta Arith. **37** (1980), 307-320.

[Iw4] H. Iwaniec, *Topics in classical automorphic forms*, Grad. Stud. Math. **17** (1997), AMS.

[IK] H. Iwaniec, E. Kowalski, *Analytic Number Theory*, AMS Colloquium Publications **53**, Providence 2004

[Ju] M. Jutila, *A variant of the circle method*, in: Sieve methods, exponential sums, and their application in number theory (Cardiff 1995), 245-254.

[KMS] H. Ki, H. Maier, A. Sankaranarayanan, *Additive problems with smooth integers*, Acta Arith. **175** (2016), 301-319.

[KMV] E. Kowalski, P. Michel, J. VanderKam, *Rankin-Selberg L-functions in the level aspect*, Duke Math. J. **114** (2002), 123-191.

[XLi] X. Li, *Upper bounds on L-functions at the edge of the critical strip*, IMRN 2010, 727-755.

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

[MT] K. Matomäki, J. Teräväinen. *On the Möbius function in all short intervals*, J. Eur. Math. Soc. **25** (2023), 1207–1225.

[Mu] R. Munshi, *Shifted convolution of divisor function d_3 and Ramanujan τ function*, Ramanujan Math. Soc. Lect. Notes Ser. **20** (2013), 251-260.

[Pi] N. Pitt, *On an analogue of Titchmarsh's divisor problem for holomorphic cusp forms*, J. Amer. Math. Soc. **26** (2013), 735-776.

[To] B. Topacogullari, *The shifted convolution of divisor functions*, Q. J. Math. **67** (2016), 331-363.

[Vi] A. I. Vinogradov, *General Hardy-Littlewood equation*, Mat. Zametki **1** (1967), 189-197.

V. Blomer and J. Li
Mathematisches Institut, Endenicher Allee 60, 53115 Bonn, Germany.
blomer@math.uni-bonn.de
J. Li
jli135@math.uni-bonn.de

L. Grimmelt
Mathematical Institute, Andrew Wiles Building, University of Oxford, Oxford OX2 6GG, UK.
grimmelt@maths.ox.ac.uk

S. L. Rydin Myerson
Mathematics Institute, Zeeman Building, University of Warwick, Coventry CV4 7AL, UK.
simon.rydin-myerson@warwick.ac.uk

Received: June 29, 2022
Revised: February 17, 2023
Accepted: February 20, 2023